



硕士学位论文



论文题目 基于 Mix-IS 算法的 SRAM 设计及良率分析

研究生姓名 李二亮

指导教师姓名 张立军

专业名称 测试计量技术及仪器

研究方向 集成电路设计

论文提交日期 2015 年 4 月

苏州大学学位论文独创性声明

本人郑重声明：所提交的学位论文是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含其他个人或集体已经发表或撰写过的研究成果，也不含为获得苏州大学或其它教育机构的学位证书而使用过的材料。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人承担本声明的法律责任。

论文作者签名： 李二亮 日 期： 2015.6.19

苏州大学学位论文使用授权声明

本人完全了解苏州大学关于收集、保存和使用学位论文的规定，即：学位论文著作权归属苏州大学。本学位论文电子文档的内容和纸质论文的内容相一致。苏州大学有权向国家图书馆、中国社科院文献信息情报中心、中国科学技术信息研究所（含万方数据电子出版社）、中国学术期刊（光盘版）电子杂志社送交本学位论文的复印件和电子文档，允许论文被查阅和借阅，可以采用影印、缩印或其他复制手段保存和汇编学位论文，可以将学位论文的全部或部分内容编入有关数据库进行检索。

涉密论文 ☐

本学位论文属 在____年____月解密后适用本规定。

非涉密论文 ☒

论文作者签名： 李亮 日期： 2015.6.19
导师签名： 沈国芳 日期： 2015.6.19

基于 Mix-IS 算法的 SRAM 设计及良率分析

中文摘要

SRAM 是一种重要的存储器，具有速度快、功耗低、可靠性高等优点，被广泛应用于系统级芯片。随着 CMOS 工艺的不断进步，SRAM 的性能不断提高的同时，其生产成本也在不断下降。然而，当工艺尺寸降低到 100nm 以后，工艺参数的随机变化引起的 MOS 管阈值电压波动，对 SRAM 稳定性的影响越来越严重，成为限制 SRAM 良率提高的重要因素之一。正因为如此，近年来在 SRAM，特别是全定制 SRAM 设计过程中，充分考虑工艺变化的设计思想（variation-aware design）成为业界的一大热点。传统的 PVT 分析方法已经不能满足先进工艺下对 SRAM 高良率的要求，采用蒙特卡罗方法分析工艺参数变化对 SRAM 稳定性及性能的影响已被广泛接纳。

本文首先分析了 CMOS 工艺以及工艺参数变化对 SRAM 稳定性的影响。随之，鉴于高密度的 SRAM 对存储单元失效率的严格要求，探讨了蒙特卡罗分析方法存在的收敛速度慢、仿真时间长等问题。为解决这些题，本文基于混合重要性采样算法，实现了一种快速蒙特卡罗方法，该方法在保证仿真精度的基础上，极大地高了仿真速度。将本文所实现的快速蒙特卡罗方法应用到 40G 工艺 32M SRAM 测试芯片设计中，通过对 SRAM 存储单元稳定性的分析，选择稳定性最好的存储单元，并应用到测试片设计中。最后对所设计测试片进行的一系列仿真数据表明，对于工艺参数变化的影响，测试片仍然具有良好的稳定性，并且存储阵列可以保持非常高的良率。

关键词：静态随机存储器；工艺参数变化；良率；混合重要性采样方法

作 者：李二亮

指导老师：张立军

SRAM Design and Yield Analysis Based on Mixture Importance Sampling

Abstract

Static random access memory plays a critical role in memory due to its high speed, robust reliability and low power consumption. It is widely used in SOCs. With the development of CMOS process, the performance of SRAM is continuously improved, and costs continue to decrease. As the technology scales down to 100nm and below, the random fluctuations of MOS threshold voltage begin to limit the exaltation of SRAM yield seriously. Because of those, variation-aware design has become a hot spot in the design of SRAM in recent years, especially in the full-custom design of SRAM. The traditional analytical method of PVT already can not satisfy the need of high yield of SRAM, and using the Monte Carlo method analyze process variation's effects on the stability and performance of an SRAM has been widely accepted.

This dissertation analyses the process impact on SRAM stability. Based on the stringent requirement of SRAM cell failure rate in high-density SRAM, the thesis discusses the disadvantages of traditional Monte Carlo analytical method, such as slow convergence, long simulation time and so on. To solve this problem, the thesis realizes a Fast Monte Carlo analytical method based on Mixture Importance Sampling, which greatly improved simulation speed. The proposed Fast Monte Carlo analytical method is applied to a 32M SRAM test chip design based on a 40nm CMOS process. After the stability analysis of SRAM cells, the most stable one is used to the design of the test chip. Simulation results indicate that the SRAM test chip has a robust stability and keeps high yield under the influence of process variations.

Keywords: SRAM; Process Variations; Yield; Mixture Importance Sampling

Written by: Li Er-Liang

Supervised by: Zhang Li-Jun

目 录

第一章 绪 论.....	1
1.1 半导体存储器概述	1
1.2 研究背景及意义	2
1.3 SRAM 结构	3
1.4 SRAM 存储单元	4
1.4.1 数据保持.....	6
1.4.2 数据读出.....	6
1.4.3 数据写入.....	6
1.5 本文主要工作及章节安排	7
第二章 亚 40 纳米 SRAM 稳定性分析	9
2.1 工艺参数波动变化介绍	9
2.2 SRAM 存储单元失效机制	11
2.2.1 读操作失效.....	11
2.2.2 写操作失效.....	12
2.2.3 数据保持失效.....	13
2.2.4 访问时间失效.....	14
2.3 SRAM 存储单元性能参数	14
2.3.1 静态噪声容限.....	14
2.3.2 读噪声容限.....	15
2.3.3 写裕度.....	16
2.4 本章小结	16
第三章 快速蒙特卡罗算法的实现.....	18
3.1 蒙特卡罗方法及其应用	18
3.1.1 蒙特卡罗方法基本思想.....	18
3.1.2 蒙特卡罗方法在 CMOS 电路中的应用	19

3.1.3 传统蒙特卡罗方法的局限性.....	21
3.2 重要性采样算法	22
3.2.1 重要性采样算法原理.....	22
3.2.2 偏置函数典型构造方法.....	23
3.3 混合重要性采样算法的实现	25
3.3.1 混合重要性采样算法原理.....	26
3.3.2 平移量 μ_s 的确定	27
3.3.3 抽样方法.....	28
3.3.4 仿真效率比较.....	29
3.4 本章小结	31
第四章 SRAM 测试片设计	32
4.1 SRAM 测试片总体结构.....	32
4.2 SRAM 存储单元设计	34
4.2.1 存储单元设计方法.....	34
4.2.2 存储阵列设计.....	37
4.3 灵敏放大器设计	38
4.3.1 灵敏放大器架构.....	39
4.3.2 灵敏放大器参数确定策略.....	42
4.4 外围电路设计	43
4.4.1 行译码器.....	43
4.4.2 列复用电路.....	44
4.4.3 输入缓冲电路.....	45
4.4.4 输出缓冲电路.....	45
4.4.5 升压电路.....	46
4.5 本章小结	47
第五章 版图设计及仿真数据分析.....	48
5.1 版图设计	48
5.2 系统仿真与分析	49

5.2.1 功能仿真.....	50
5.2.2 读写模式最小工作电压仿真.....	51
5.2.3 漏电流.....	52
5.2.4 存储阵列良率分析.....	52
5.3 本章小结	53
第六章 总结与展望.....	54
6.1 总结	54
6.2 展望	55
参考文献.....	56
攻读硕士期间发表的论文.....	60
致 谢.....	61

ETD 论文管理系统

第一章 绪 论

20 世纪 80 年代，CMOS 工艺开始成为集成电路设计工艺的主流，极大地推动了微电子产业的发展，集成电路的集成规模和性能也得到了空前的提高。随着 CMOS 工艺的不断进步，集成电路的成本也在不断下降，促使了个人计算机、手机等消费电子逐渐普及到一般民众的日常生活中。如今，CMOS 工艺已经进入到深亚微米时代，40nm，28nm 也已经开始被广泛应用到数字电路设计中，电路设计以及可靠性研究也进入了新时代。

1.1 半导体存储器概述

近几十年，半导体存储器迅速发展，是半导体产业的重要组成部分。目前，在全世界范围内，半导体存储器占半导体产业近22%的比例^[1]，是几乎所有电子系统的基本组成之一。半导体存储器是片上系统（System on Chip, SoC）中非常重要的组成部分之一^[2]，随着半导体工艺制程的不断进步，以及系统对性能要求的越来越高，半导体存储器在SoC上所占的面积比例也越来越高，如图1.1所示。如今，在SoC中，超过90%的面积被各种类型的半导体存储器占据。

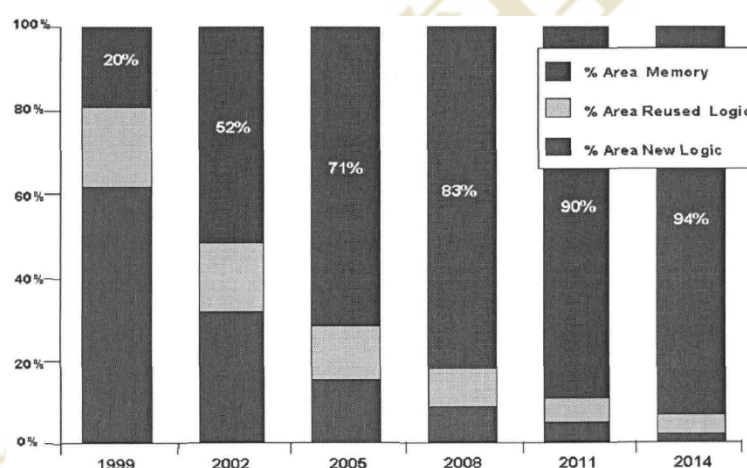


图 1.1 半导体存储器在 SoC 芯片中的比重^[3]

半导体存储器的种类繁多，通常情况下，按照其断电后存储的数据是否丢失，可以将半导体存储器分为两类：非易失性存储器（Nonvolatile Memory, NVM）和易失

性存储器（Volatile Memory, VM）。

非易失性存储器的特点是电源中断后，存储器所存储的数据不丢失，而且重新获得供电后，所存储的数据仍然能够被成功读取。常见的非易失性存储器包括 ROM、EPROM、EEPROM、FLASH 等。随着技术的发展，一些新类型的非易失性存储器陆续出现，如 FeRAM、SONOS 等。

易失性存储器的特点是电源中断后，存储器所存储的数据丢失，常见的易失性存储器主要包括动态随机存储器（Dynamic Random Access Memory, DRAM）和静态随机存储器（Static Random Access Memory, SRAM）。DRAM 利用电容上的电荷来存储数据，但由于电容上电荷丢失的原因，需要定时刷新。DRAM 具有集成密度高和成本低的特点，主要用于计算机的主存。与 DRAM 不同，SRAM 利用交叉耦合的锁存器来存储数据，因此存储数据稳定，不需要反复刷新。因此，SRAM 具有功耗低、速度快的特点，常用于 SoC 上的高速缓冲存储器（Cache），作为处理器与主存之间的一级存储器。

1.2 研究背景及意义

1965 年，英特尔（intel）创始人之一戈登·摩尔提出了著名的摩尔定律，即一颗集成电路芯片上所集成的晶体管数目，每隔 18 个月就翻一倍，性能也将提升一倍。这个预言在接下来的近半个世纪得到了很好的印证。随着半导体工艺的不断发展，晶体管尺寸变得越来越小，半导体存储器的速度变得越来越快，功耗变得越来越低，成本也越来越低^[4]。

在 CMOS 电路设计中，当工艺尺寸下降到 100nm 以下时，CMOS 工艺变化引起 MOS 管阈值电压的随机波动，已经对 SRAM 的稳定性产生了不可忽略的影响。在 40nm、28nm 以下，工艺参数变化已经严重限制了 SRAM 良率的提高。因此，在工艺参数变化的影响下，如何保证所设计的 SRAM 的稳定性，是每个设计工程师所面临和需要解决的问题之一。

为了帮助设计者分析工艺参数变化对 CMOS 电路的影响，一般在 90nm 以下，大型的工艺厂都会提供 MOS 管的蒙特卡罗（Monte Carlo, MC）模型库，其参数来自大量器件分析测试，将蒙特卡罗模型应用于实际电路中，就可以得到电路对器件参数的敏感程度，得到电路的强壮程度，对设计有直接的指导意义，特别是深亚微米条

件下，对于高精度电路设计，蒙特卡罗分析是必不可少的辅助工具^[5]。

然而，随着 SRAM 集成密度越来越大，对单个 SRAM 存储单元失效率仿真精度的要求越来越高，传统蒙特卡罗分析方法因为需要的仿真次数过于庞大，仿真时间过长，已经变得越来越不实用。例如对一个 SRAM 关键路径进行传统 Monte Carlo 仿真，要在 95% 的置信区间内得到 6-sigma 的精度，需要 1×10^{12} 次采样。而用 spice 进行仿真，完成 2×10^4 次采样大概需要 7 天时间，所以 1×10^{12} 次的传统 Monte Carlo 仿真在注重设计周期的今天是不切实际的。因此，为了提高仿真效率，基于传统蒙特卡罗算法，人们提出了快速蒙特卡罗算法（Fast Monte Carlo, FMC）。

1.3 SRAM 结构

常用的 SRAM 结构如图 1.2 所示，主要包括存储单元阵列（Memory Array），输入输出电路（I/O），地址译码电路（Address Decoder），控制电路（Control Circuit）四个部分。

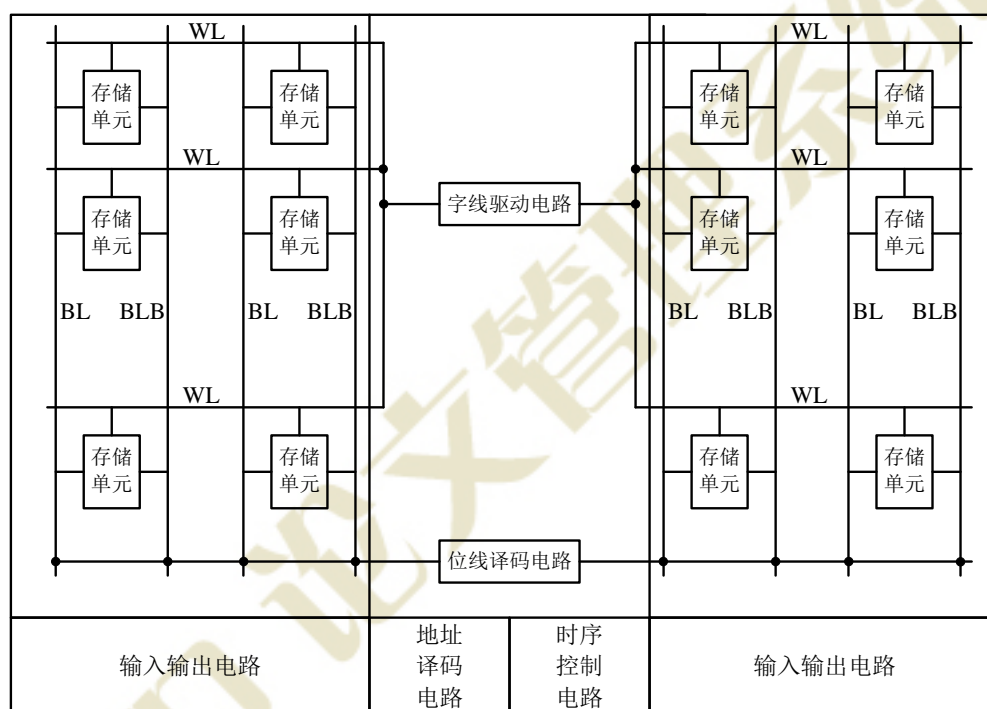


图 1.2 SRAM 结构图

存储单元阵列电路是 SRAM 的核心部分，对于较大容量的 SRAM 而言，存储单元阵列决定了整个 SRAM 芯片的面积、功耗以及读写性能等。存储单元是存储单元

阵列的基本组成单位, SRAM 存储单元一般有 6 管、8 管等不同形式, 通过字线 (Word Line, WL) 和位线 (Bit Line, BL) 相互连接。输入输出电路一般又可以分为三部分: 写驱动电路, 用于写操作时将数据写入 SRAM 存储单元; 灵敏放大器 (Sense Amplifier, SA), 用于读操作时放大微小的位线电压差; 位线预充电电路, 用于将处于空闲状态的位线电压拉到工作电压。地址译码器根据外部输入的地址信号, 选中特定的 WL 和 BL, 两者交叉的位置即是选中的存储单元, 可对其进行读或写操作。控制电路的主要作用是协调 SRAM 内部不同部分工作, 包括灵敏放大器使能信号, 读写使能信号以及开闭预充电信号等。

1.4 SRAM 存储单元

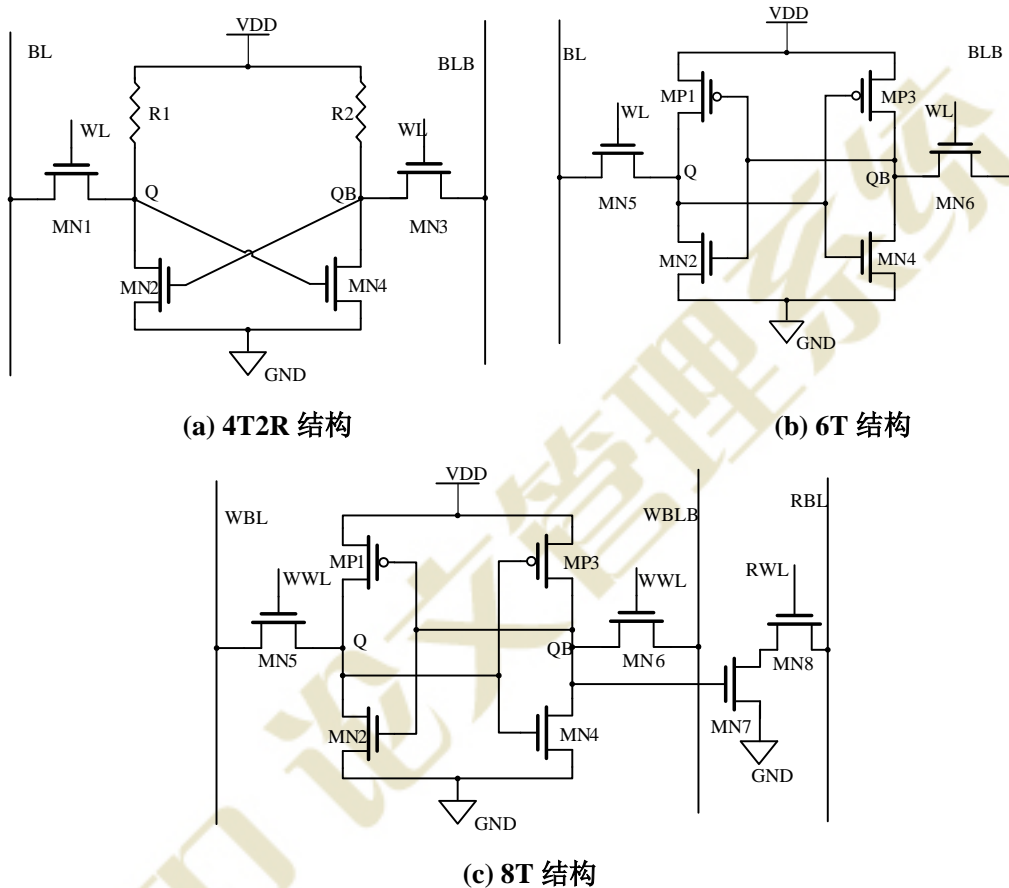


图 1.3 SRAM 存储单元

SRAM 存储单元有多种结构, 典型的有 4T2R、6T 和 8T 结构, 如图 1.3 所示。4T2R 结构是由 4 个 NMOS 管和 2 个电阻构成, 使用两个反相器以正反馈形式连接存

储数据, 每个反相器由一个 NMOS 管和一个电阻组成, Q 和 QB 点为存储节点。4T2R 存储单元具有集成密度高、存取速度快的特点。但由于 MN2 和 MN4 中必定有一个导通, 导通的 NMOS 管和电阻一起构成了静态回路, 在数据保持模式下, 静态功耗过大。在静态功耗是 SRAM 主要功耗来源的今天, 这种结构逐渐退出了历史舞台。

6T 结构是最常见的 SRAM 存储单元, 由 6 个晶体管组成。MOS 管 MP1 和 MP3 称为负载管, NMOS 管 MN2 和 MN4 称为驱动管, NMOS 管 MN5 和 MN6 称为传输管, Q 和 QB 为两个存储节点。MP1 与 MN2, MP3 与 MN4 分别组成两个反相器, 这两个反相器首尾相连, 形成了一个交叉耦合的锁存器, 将数据保存在存储节点 Q 和 QB。与 4T2R 结构比较, 6T 结构是用两个 PMOS 管替换 4T2R 结构中的两个电阻, 隔断了静态回路, 降低了静态功耗。

8T 结构的存储单元比 6T 结构存储单元多两个 NMOS 管, 主要区别是拥有一条写数据线 WWL 和一条读数据线 RWL, 将读操作和写操作分开。随着工艺的进步, 为了保持同样的读写裕度, 8T 存储单元的驱动管可以按照比例缩小, 在最差情况下, 可以比 6T 结构节省 22% 的功耗^[6]。另外, 当工艺尺寸下降到一定节点时, 8T 存储单元面积将与 6T 存储单元相同^{[7][8]}, 甚至更小。如图 1.4 所示, 工作电压为 0.8V 时, 在 65nm 工艺节点处, 8T 存储单元与 6T 存储单元面积已基本相同。另外, 在读操作时 WWL 关闭, 所以不会影响到存储的数据, 具有高的读可靠性^[9]。

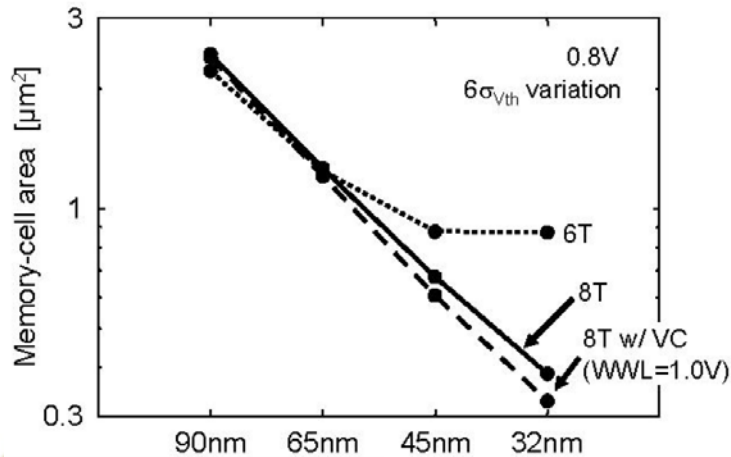


图 1.4 6T 与 8T 存储单元面积比较^[8]

本文所设计 SRAM 采用的是 6T 存储单元, 因此仅对 6T 存储单元进行分析。作为一种重要类型的存储器, SRAM 最重要的功能是存储数据, 为了实现这个功能, SRAM 需要支持三种基本操作: 数据保持、数据读出以及数据写入^[10]。

1.4.1 数据保持

存储单元处于数据保持状态时，假设 Q 点存储数据“0”，为低电平，QB 点存数据“1”，为高电平。（BL，BLB）被预充到高电平，WL 为低电平，MN5 和 MN6 处于截止状态，存储节点存储的数据保持不变。

1.4.2 数据读出

存储单元处于数据读出状态时，WL 电压由低电平拉至高电平，传输管 MN5 和 MN6 导通。由于 QB 点电平为高，MN2 导通。MN2 和 MN5 在 BL 与 GND 之间形成了一条通路，BL 向 GND 放电，而 BLB 则维持高电平不变，因此，在 BL 与 BLB 之间形成了电压差 ΔV 。通过灵敏放大器，将这个电压差放大至输出端口。SRAM 读操作简化图如图 1.5 所示。

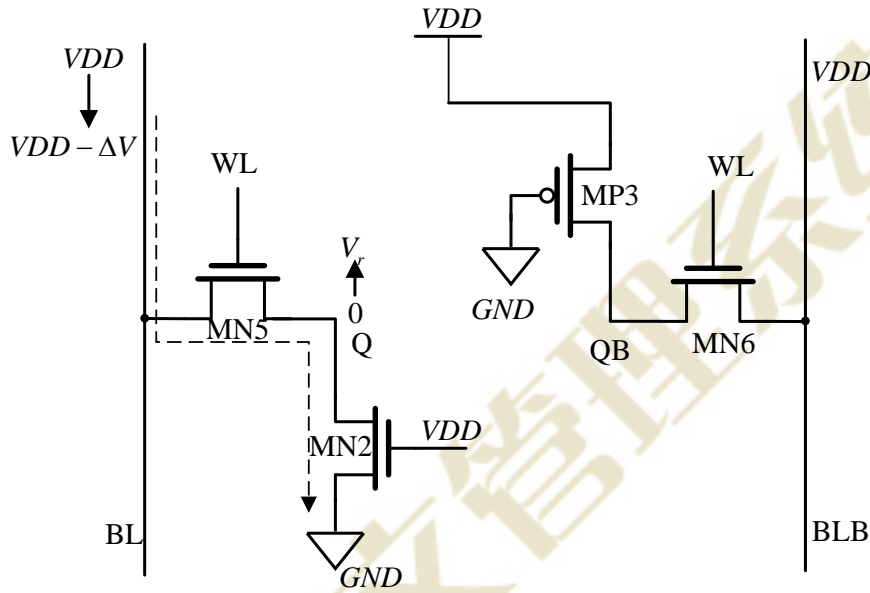


图 1.5 SRAM 读操作简化图

1.4.3 数据写入

Q 点存储初始数据“0”，QB 存储初始数据“1”，假设往存储单元内写入相反的数据，即 Q 点写入“1”，QB 点写入“0”。对存储单元进行写操作时，WL 电压被拉至高电平，MN5 和 MN6 导通。BLB 的电压由 VDD 变成 GND，则电源电压 VDD 通过 MP3 和 MN6 向 BLB 放电，存储节点 QB 的电压由 VDD 下降到 V_w ，当 V_w 低于 MP1 和 MN2 组成的反相器翻转电压时，就会导致 Q 和 QB 点数据的翻转，即将对应的数

据写入。SRAM 写操作简化图如图 1.6 所示。

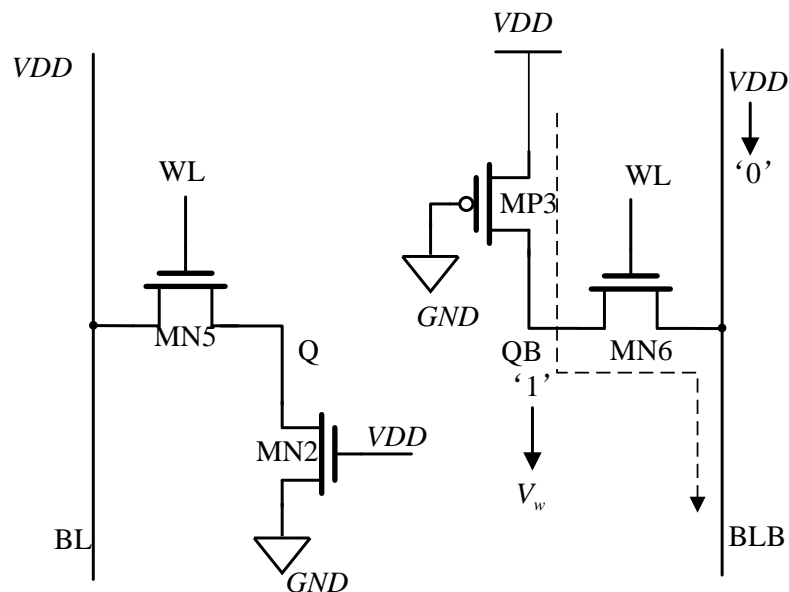


图 1.6 SRAM 写操作简化图

1.5 本文主要工作及章节安排

随着 CMOS 工艺尺寸的减小，工艺参数变化对 SRAM 稳定性产生了越来越严重的影响。另外，对于高集成密度的 SRAM，利用传统蒙特卡罗算法来分析工艺参数变化对 SRAM 良率的影响，所呈现的不足也越来越明显。针对这些问题，本文进行了改善和研究，并应用到实际设计中。

文章主要工作包括：

- 1) 分析了工艺参数变化对 SRAM 稳定性的影响。
- 2) 分析了传统蒙特卡罗方法在 CMOS 电路中的应用和存在仿真效率低的问题。
- 3) 实现了混合重要性采样算法，有效地解决上述存在的问题。
- 4) 基于所实现混合重要性采样算法的分析，设计了一颗 32M 的 SRAM 测试片。
- 5) 对所设计测试片进行了基础功能检查、读写模式下最小工作电压仿真、漏电流检查、以及存储阵列的良率分析，验证了测试片的稳定性。

本论文主要分为五章，具体内容安排如下：

第一章介绍了选题的意义和背景，SRAM 的基本结构以及操作原理，阐述了本论文的主要内容和安排。

第二章介绍并分析了 CMOS 工艺参数变化的来源，以及工艺参数变化对 SRAM

存储单元稳定性的影响。

第三章分析了传统蒙特卡罗方法在深亚微米 VLSI 设计中的应用和呈现的问题。并在此基础上，实现了一种基于混合重要性采样算法的快速蒙特卡罗分析方法。

第四章基于 40G 工艺，设计了一颗 32M 的 SRAM 测试芯片。以第三章实现的快速蒙特卡罗方法为指导，以不同参数标准，对几种存储单元进行稳定性上的比较，选取稳定性最好的存储单元应用到测试片的设计中。

第五章主要在 40G CMOS 工艺下，进行基础功能检查，读写模式下最小工作电压仿真，漏电流检查，以及存储阵列的良率分析，验证测试片的稳定性。

第六章对全文进行总结，并在此基础上提出展望。

第二章 亚 40 纳米 SRAM 稳定性分析

SRAM 的稳定性指的是其抗外部噪声和工艺参数变化干扰的能力。作为现代 SOC 和微处理器的关键组成部分，SRAM 的稳定性对于整个系统非常重要。对于 100nm 以下的 CMOS 工艺，工艺参数随机变化导致 MOS 管阈值电压的波动愈发严重，威胁到 SRAM 的稳定性，导致 SRAM 良率下降。当工艺制程进入 40nm，这种影响将更加严重。

2.1 工艺参数波动变化介绍

集成电路生产过程中，工艺参数随机变化无处不在，其大致可以分为两种：全局工艺参数变化和局部工艺参数变化。其中存在于不同 Lot 之间，不同 wafer 之间，不同 die 之间的称为全局工艺参数变化；存在于相同 die 上的不同器件之间的称为局部工艺参数变化。正是由于这些工艺参数差异的存在，才导致 MOS 管在性能上产生了一些差异^[11]。

工艺参数变化导致 MOS 管性能的波动，主要表现在导致 MOS 管阈值电压 V_{th} 产生了随机波动^{[12][13]}。这一影响主要来自局部工艺参数变化，并且，局部工艺参数变化又可以分为系统性变化和随机性变化。系统性变化一般是可以预测的，依赖于确定性的因素，例如周围拓扑环境、版图结构等^[14]。与系统性变化相比，随机性变化是不可预测的，主要由制造时的随机不确定性引起的，如随机掺杂波动(RDF)，因此随机性变化对于设计中关键模块的良率起十分重要的作用。

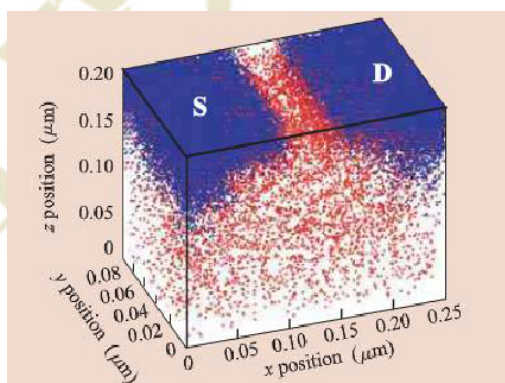


图 2.1 随机掺杂的蒙特卡罗 3D 仿真模型^[15]

在进行注入和热退火工艺时,沟道中注入离子的数量和位置具有随机波动性,这个原因导致的 MOS 管阈值电压的随机变化,称为随机掺杂波动(RDF)^[16]。如图 2.1 所示,为随机掺杂的蒙特卡罗 3D 仿真模型。除此之外,导电沟道的有效长度波动^[17]、栅氧厚度不均匀也是导致 MOS 管阈值电压波动的主要原因。

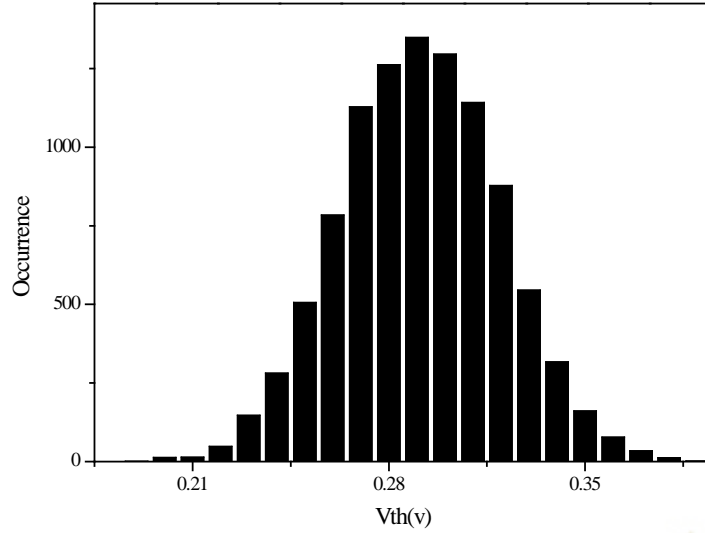


图 2.2 执行 10000 次 MC 仿真得到的 NMOS 管阈值电压的分布图

如图 2.2 所示,通过执行 10000 次蒙特卡罗仿真得到的 n 型 MOS 管阈值电压的分布图。从图中可以看出,在工艺参数变化的影响下,晶体管阈值电压服从高斯分布。

另外,工艺参数变化导致 MOS 管阈值电压的变化与晶体管栅极面积成反比关系。对于 SRAM 来说,为了减小面积的开销,通常采用工艺允许的最小尺寸来设计存储单元。因此随机掺杂波动效应对 SRAM 存储单元的影响更为明显,可能会导致 MOS 管阈值失配,使存储单元失效。

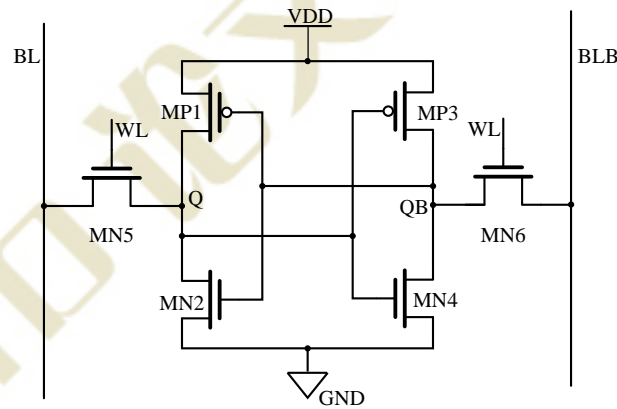


图 2.3 6T 存储单元

2.2 SRAM 存储单元失效机制

工艺参数的随机波动最主要的表现是使晶体管的阈值电压发生了偏移, 导致相邻晶体管间产生很大的失配, 影响存储单元的稳定性, 使存储单元失效。SRAM 存储单元的失效机制主要有四种情况, 包括: 1) 读操作失效。2) 写操作失效。3) 数据保持失效。4) 访问时间失效。下面将对这几种失效机制逐个分析。

2.2.1 读操作失效

结合图 1.5, Q 点存储数据“0”, 当 SRAM 存储单元进行读操作时, 由于 MN5 与 MN2 等效电阻的分压效果, Q 点电压由 0 升高到 V_r , 如果 V_r 达到 MN4 的阈值电压, 将会导致存储数据的翻转, 即破坏了保存的数据。因此, 为了保证读操作的稳定性, 必须保证 V_r 最大值小于 MN4 的阈值电压。考虑到工艺参数的随机变化对单元内晶体管驱动能力的影响, 可能导致 V_r 高于设计值, 或者 MN4 阈值电压减小, 都容易导致读操作失效。因此, 读失效的概率可以描述为:

$$P_{RF} = P(V_r > V_{Mn4}) \quad (2-1)$$

随着 BL 电压的降低, V_r 也会越来越小, 所以在读操作刚开始的时候, V_r 达到最大值。不考虑工艺参数变化, V_r 最大值可以通过求解电流方程获得^[18], 为了简单起见, 忽略 MN5 上的体效应, 有:

$$k_{n,MN5} \left((VDD - V_r - V_{Tn}) V_{DSATn} - \frac{V_{DSATn}^2}{2} \right) = k_{n,MN2} \left((VDD - V_{Tn}) V_r - \frac{V_r^2}{2} \right) \quad (2-2)$$

求解 V_r 的一元二次方程, 可以得到:

$$V_r = \frac{V_{DSATn} + CR(VDD - V_{Tn}) - \sqrt{V_{DSATn}^2(1 + CR) + CR^2(VDD - V_{Tn})^2}}{CR} \quad (2-3)$$

上式中, CR 称为单元比 (Cell Ratio), 定义为

$$CR = \frac{W_2 / L_2}{W_5 / L_5} \quad (2-4)$$

其中, W_2 / L_2 为 MN2 的宽长比, W_5 / L_5 为 MN5 的宽长比。采用 40G 工艺设计时, V_r 与 CR 的关系如图 2.4 所示。从图中可以看出, CR 越大, V_r 就越小。因为, 在 40G 工艺中, 下拉管 MN4 的阈值电压为 232.36mv, 如果要使 V_r 小于这个值, 必

须保证 CR 大于 0.82。

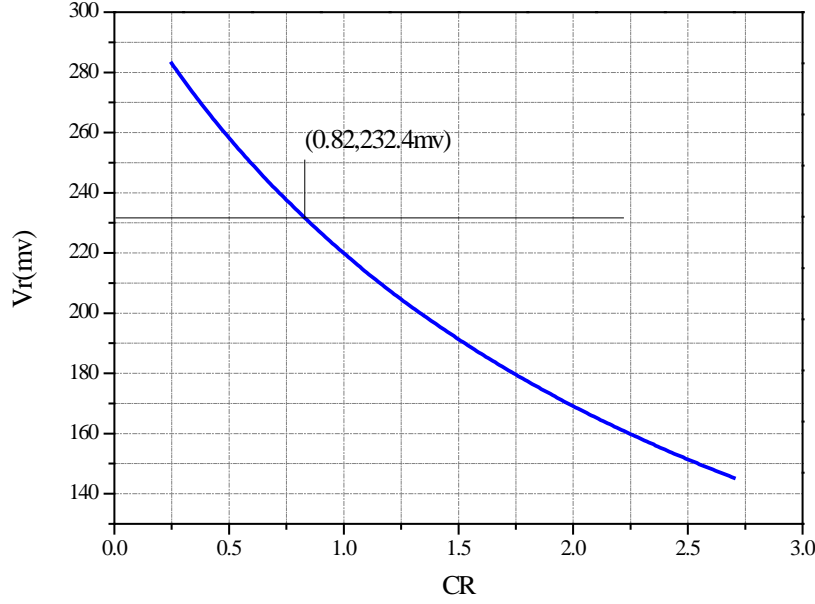


图 2.4 读操作时节点电压 V_r 与 CR 的关系

2.2.2 写操作失效

结合图 1.6, 当 SRAM 存储单元进行读操作时, 存储节点 QB 处的电压, 由 VDD 下降到 V_w , 如果在 WL 为高的时间 T_{WL}^w 内, V_w 的值不能低于 MN2 的阈值电压, 则数据将不能被写入, 即出现了写操作失效。从原理上讲, 把 Q 点电压下拉至低于 MP3 和 MN4 组成的反相器开关阈值就足以保证开始切换, 但考虑到噪声容限时, 更为保险的是要求把 Q 点电压下拉至低于 MN2 的阈值电压^[18]。因此, SRAM 写失效概率可表示为:

$$P_{WF} = P(T_w > T_{WL}^w) \quad (2-5)$$

其中, T_w 定义为 QB 点电压由 VDD 降到 MN2 阈值电压所需要的时间。由于工艺参数的变化, 使 V_w 与 MN2 阈值电压的波动, 导致设计中可以正确进行的写操作, 实际操作时也可能出现失效。

一般情况下, 不考虑工艺参数变化, V_w 的临界值可以通过所希望翻转点处的电流相等来求解^[18]:

$$k_{n,MN6} \left((VDD - V_{Tn}) V_w - \frac{V_w^2}{2} \right) = k_{p,MP3} \left((VDD - |V_{Tp}|) V_{DSATp} - \frac{V_{DSATp}^2}{2} \right) \quad (2-6)$$

求解 V_w 得到:

$$V_w = VDD - V_{Tn} - \sqrt{(VDD - V_{Tn})^2 - 2 \frac{\mu_p}{\mu_n} PR \left((VDD - |V_{Tp}|) V_{DSATp} - \frac{V_{DSATp}^2}{2} \right)} \quad (2-7)$$

上式中的 PR 称作上拉比, 定义为

$$PR = \frac{W_3 / L_3}{W_6 / L_6} \quad (2-8)$$

其中, W_3 / L_3 为 MP3 的宽长比, W_6 / L_6 为 MN6 的宽长比, 且 PR 越小, 写入数据越容易。采用 40G 工艺时, V_w 与 PR 的关系如图 2.5 所示。从图中可以看出, PR 越大, V_w 也越大。MN2 的阈值电压为 232.36mv, 因此为了保证数据能够被写入存储单元中, 必须保证 PR 的值小于 1.51。

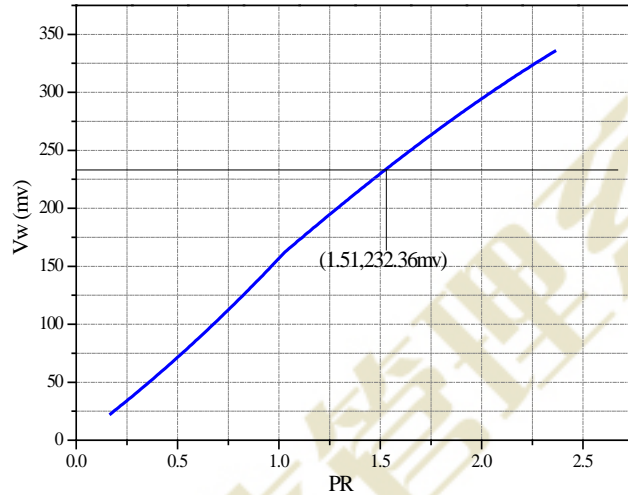


图 2.5 写操作时节点电压 V_w 与 PR 的关系

2.2.3 数据保持失效

在稳定模式下, 一般情况下, 为了减少漏电流, 系统会降低 SRAM 存储单元的电压。如果电压降低导致存储数据翻转, 就会出现数据保持失效^[19]。因此, 数据保持失效概率可以表示为:

$$P_{HF} = P(V_{DDH \min} > V_{HOLD}) \quad (2-9)$$

其中, $V_{DDH \min}$ 为存储单元数据保持电压, V_{HOLD} 为 standby 状态下的存储单元工作电压。SRAM 电压降低时, 存储'1'的节点电压会随着降低, 如果负载管导通电流

不够大，驱动管有漏电流，节点电压降低加剧。如果电压低于 MP3-MN4 反相器的翻转电压，数据就翻转。然而，假如工艺参数变化导致 MN2 阈值电压减小，MP1 阈值电压增大，同时 MN4 阈值电压增大，MP3 阈值电压减小（这会导致 MP3-MN4 反相器翻转电压升高），原本正常的稳定电压下，数据也会丢失。

2.2.4 访问时间失效

存储单元访问时间 T_{access} 定义为位线 BL 和 BLB 发展电压差 ($\Delta V=0.1VDD$) 所需时间。如果发展 $0.1VDD$ 所需时间大于最大可容忍时间 T_{MAX} ，灵敏放大器就不能正确放大信号，从而造成读出数据错误。因此，访问时间失效概率可以表示为：

$$P_{AF} = P(T_{ACCESS} > T_{MAX}) \quad (2-10)$$

访问时间会随着传输管或驱动管阈值电压的增大而升高，因此，传输管或驱动管的阈值电压增大，会导致访问时间失效。

2.3 SRAM 存储单元性能参数

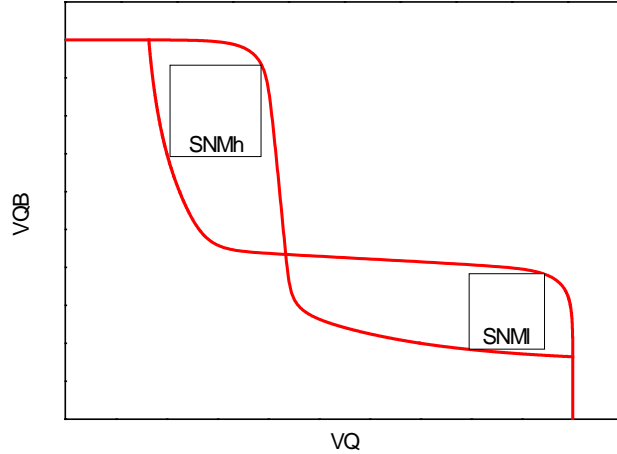
芯片量产时存在的工艺参数变化，以及工作环境中电压、温度波动等因素，都会对 SRAM 存储单元的正常工作的产生一定的影响，甚至会导致 SRAM 出现错误。因此在 SRAM 存储单元设计时，会针对读写等性能波动设计一定的裕度，使影响因素在一定范围内波动时，SRAM 存储单元仍然能够正常工作。常用的衡量 SRAM 存储单元性能参数有：静态噪声容限 (Static Noise Margin, SNM)，读噪声容限 (Read Noise Margin, RNM)，写裕度 (Write Margin, WM)。

2.3.1 静态噪声容限

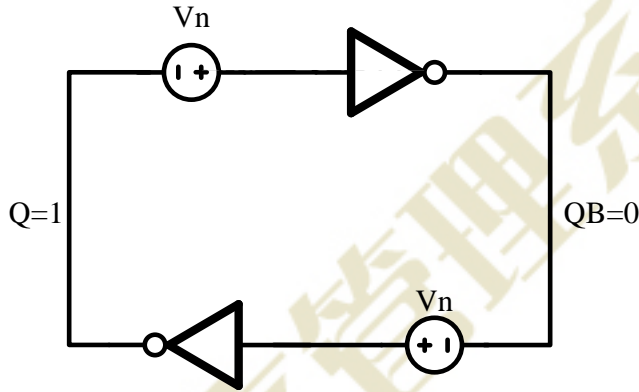
SRAM 存储单元的 SNM 定义为，在存储数据不变的前提下，存储节点所能承受最大噪声的电压值。因此，SNM 可以衡量数据保持状态下，存储单元保持数据的能力，且 SNM 越大，则存储单元数据保持能力越强。测量 SNM 的常用方法是通过存储单元中反相器的电压传输特性曲线 (Voltage Transfer Curve, VTC)，如图 2-6(a) 所示，两个反相器的 VTC 曲线交叉构成的图形称为蝶形曲线 (Butterfly Curve) ^{[20][21]}。SNMh 和 SNMI 分别为蝶形曲线上半部分和下半部分所能容纳最大正方形的边长，SRAM 存储单元 SNM 定义为 SNMh 和 SNMI 中最小值^[22]。若存储单元中两个反相器完全对称，则 SNMh 与 SNMI 相等。但由于工艺参数变化等因素的影响，往往会造成

两个反相器不对称，两条 VTC 会往不同方向偏移，SNM 便会减小。因此，当影响因素足够大时，会出现 $SNM=0$ 的情况，这样的存储单元将不能稳定保存数据。

测量 SNM 的常用电路如图 2.6(b)所示，噪声源 V_n 幅值从 0 开始缓慢增加，直到双稳态电路翻转，则对应的噪声值即为此存储单元的 SNM。



(a) 蝶形曲线



(b) SNM 测试电路图

图 2.6 SRAM 存储单元 SNM 定义图

2.3.2 读噪声容限

RNM 是衡量 SRAM 存储单元读操作稳定性的一个重要参数。和 SNM 定义类似，RNM 被定义为，读操作时，SRAM 存储单元的存储节点上所能承受的最大噪声^[23]。读失效往往是一种单边干扰，所以测试电路如图 2.7 所示^[14]。与 SNM 测试电路不同的，RNM 测试电路仅有一边加有噪声源，并且值得注意的是，RNM 测量是在读操作状态下进行，而 SNM 测量则是在数据保持状态下进行的。RNM 越大，SRAM 存储

单元的读操作就越稳定。

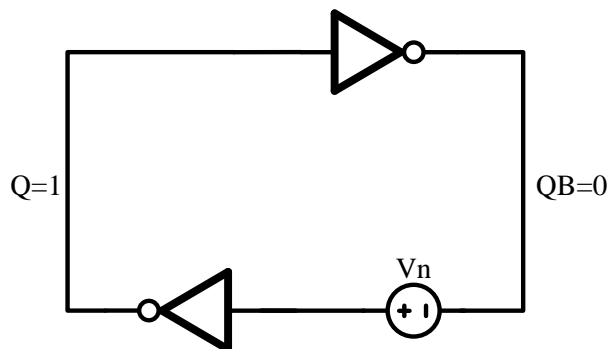


图 2.7 RNM 测试电路图

2.3.3 写裕度

对于 SRAM 存储单元来说，不仅要求能够稳定地读取数据，还要能够稳定地将数据写入。根据小节 1.4.3 中的内容，存储单元写操作时，BLB 的电压被拉至 GND。一般情况下，BLB 电压只要低到能使 QB 点电压 V_w 达到反相器的翻转电压即可完成写操作。因此，WM 可以定义为，能够保证数据被成功写入前提下，BLB 上的最高电压^[24]，如图 2.8 所示。WM 可以用来衡量存储单元写入数据能力，且 WM 越大，SRAM 存储单元写操作越稳定。

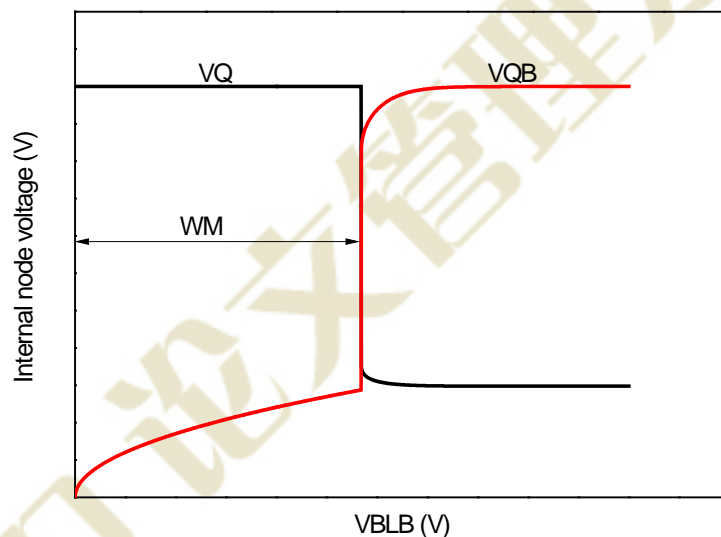


图 2.8 SRAM 存储单元 WM 定义图

2.4 本章小结

本章首先讨论了 CMOS 工艺参数变化的来源，导致 MOS 管阈值电压随机波动的

最主要原因是导电沟道中离子的随机掺杂波动，另外，导电沟道的有效长度和栅氧厚度不均匀也是主要原因。分析了 SRAM 存储单元操作的四种失效机制：读失效、写失效、访问时间失效、数据保持失效，给出了统计上分析这四种失效的概率公式。并且，在不考虑工艺参数变化的条件下，为保证读写功能稳定，SRAM 存储单元的单元比 CR 和上拉比 PR 应该满足条件： $CR > 0.82$ ， $PR < 1.51$ 。最后，介绍了衡量 SRAM 存储单元稳定性的几个重要参数：静态噪声容限，读噪声容限，写裕度。

ETD 论文管理系统

第三章 快速蒙特卡罗算法的实现

在工程设计中, 计算机仿真扮演着十分重要的角色。对于芯片设计来说, 设计的周期长, 成本高。因此, 在芯片投产之前的各个阶段, 都离不开计算机仿真这个可靠的手段。利用计算机仿真有效地缩短了工程设计周期, 降低了设计成本, 提高了设计可靠性。在进行计算机仿真时, 仿真算法的选取非常重要。选取仿真算法时, 要综合算法本身的复杂度和算法的估计器方差收敛情况考量, 复杂度低、方差小的仿真算法, 可以减少仿真样本数, 提高仿真效率。

作为 20 世纪十大算法之一, 蒙特卡罗算法被广泛应用于解决各领域的复杂问题。在 SRAM 设计中, 蒙特卡罗方法则被有效地用来分析工艺参数变化对 SRAM 性能和良率的影响。然而, 随着集成电路工艺的进步, SRAM 集成规模的不断提高, 传统蒙特卡罗方法采样多, 收敛慢, 仿真时间极长等缺点逐渐突显。为了弥补这一不足, 本章在传统蒙特卡罗的基础上, 基于业界已有的研究理论, 实现了混合重要性采样算法 (Mixture Importance Sampling, Mix-IS)。

3.1 蒙特卡罗方法及其应用

蒙特卡罗 (Monte Carlo, MC) 方法, 也称作计算机随机模拟方法。通常被用来验证随机模型是否正确或者用来分析难以实现的随机过程, 是一种通过产生大量的随机样本, 然后对所产生的样本进行统计分析的方法。蒙特卡罗方法具有适用解决随机性问题的优点, 在 SRAM 设计中, 被广泛用于分析存储单元等模块的稳定性^[25-27]。

3.1.1 蒙特卡罗方法基本思想

蒙特卡罗方法是一种将频率近似概率的方法, 其基本思想是: 在分析物理、数学或工程技术等方面的问题时, 首先构造一个随机过程或者概率模型, 使其中的某个参数 (如某个随机变量的数学期望或某个事件发生的概率) 等于所要解决问题的解; 然后对所构造的过程或模型进行随机抽样试验, 以计算所求参数的统计特征; 最后给出所求解问题的近似解。使用蒙特卡罗方法解决问题, 主要分为三个步骤^[28]:

第一, 构造或描述概率过程。如果所求解问题本身就具有随机性质, 那么只需要

正确的描述和模拟这个概率过程；如果所求解的本身是个确定性问题，那么就必须先人为构造一个随机过程，并且使该随机过程的某个参数等于所求解问题的解。

第二，实现从已知分布抽样。概率模型都可以看作是由各种各样的概率分布构成的，因此，可以从之前构造或描述的概率模型中产生已知概率分布的随机变量（或随机向量），对这些已知的概率分布进行抽样，并用抽样得到的数据来计算所求参数的统计特征。

第三，建立各种估计量。根据所求参数的统计特征，给出所求问题解的近似值，是一种以频率近似为概率的方法。

3.1.2 蒙特卡罗方法在 CMOS 电路中的应用

在 CMOS 工艺中，考虑到工艺参数变化对晶体管电学特性的影响，一般会将晶体管分成 3 种情况，即快速情况（FAST），典型情况（Typical），慢速情况（Slow）。因此，对应到由 NMOS 和 PMOS 管组成的电路中便会有 TT、SS、FF、FS、SF 五种工艺拐点，如图 3.1 所示。

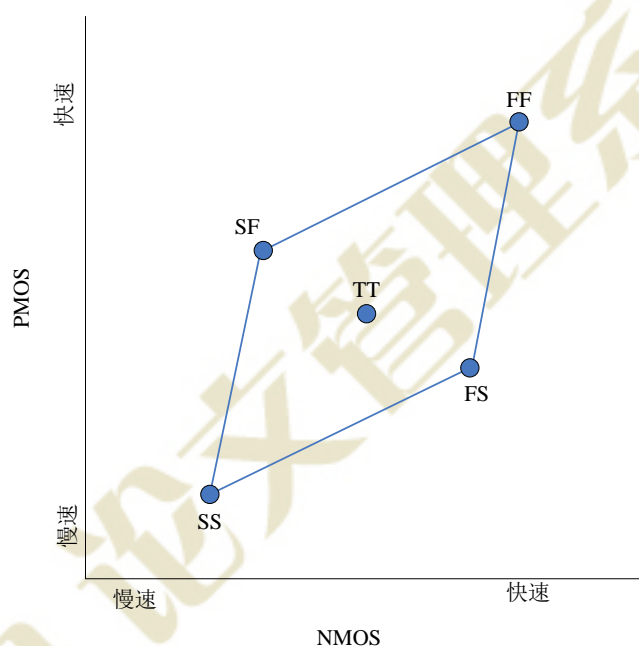


图 3.1 工艺拐点

对于电路设计者而言，在较早的 CMOS 工艺中，考虑到环境偏差（温度和电压）情况下，只要确保电路在图 3-1 中的 5 个工艺拐点处仿真通过，就基本可以保证电路的稳定性和良率。然而，对于 100nm 以后的 CMOS 工艺而言，由于工作电压的降低

以及工艺参数变化对 MOS 管阈值电压影响的加剧，这种做法已经不能保证芯片的稳定性和良率。

在深亚微米条件下，为了尽可能地模拟实际流片中工艺参数变化对晶体管电学特性的影响，工艺厂商一般都会提供晶体管的蒙特卡罗模型，电路设计者可以通过蒙特卡罗方法分析工艺参数变化对电路的影响，设计出可靠的电路，保证芯片的良率。

在 SRAM 电路中，存储单元作为芯片中最小的器件，受工艺参数变化的影响尤为明显。蒙特卡罗方法能够对 SRAM 在不同条件下的稳定性进行模拟仿真，提高良率。在器件方面，使用蒙特卡罗方法分析存储单元晶体管阈值电压的随机变化情况对 Write Margin 产生的影响^[29]，可以确定阈值电压的最佳情况。在电路单元优化方面，可以使用蒙特卡罗分析得到灵敏放大器新架构电路的器件尺寸及各个电压变化对整体性能和良率的影响^[30]。在 SRAM 低功耗优化方面，可以用蒙特卡罗方法来估计存储单元的数据保持电压（Data Retention Voltage, DRV）^[31]。并且使用蒙特卡罗方法对电压测试电路进行验证能够确定 SRAM 保证功能正确的最低电压工作点^{[32][33]}。在 SRAM 抗辐射稳定性方面，基于不同工艺的 SRAM 对于辐射的敏感程度可以使用蒙特卡罗方法进行模拟仿真对比^{[34][35]}。具体的应用方法如下：

假设 $X = [x_1, x_2, \dots, x_N]^T$ 是一个 N 维随机向量，其中， $x_i (i=1,2,\dots,N)$ 是表示不同的工艺参数偏差的随机变量， X 的概率密度函数定义为 $p(X)$ 。首先，不妨假设 $x_i (i=1,2,\dots,N)$ 相互独立，且均服从标准正态分布，即 $x_i \sim N(0,1)$ 。 x_i 的概率密度函数为

$$f(x_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right) \quad (3-1)$$

则 X 的概率密度函数 $p(X)$ 可以表示为

$$p(X) = \prod_{i=1}^N \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right) \right) \quad (3-2)$$

在联合高斯分布中，可以通过主成分分析方法（Principal Component Analysis, PCA）将相关的随机变量转换为公式(3-2)中互相独立的随机变量，因此，上述假设并不失一般性^[24]。

因此，电路的失效率 P_f 可以表示为

$$P_f = \int_{\Omega} p(X) \cdot dX \quad (3-3)$$

其中， Ω 表示 X 分布范围内使电路失效的区域，那么电路的良率就是 $1 - P_f$ 。

通常情况下，造成工艺参数变化和 CMOS 电路失效的因素很多，也很复杂，因此通过求解公式(3-3)来计算 CMOS 电路良率是一个不切实际的做法。采用蒙特卡罗方法则会将问题变得简单很多。

在蒙特卡罗方法中，假设 $f(X)$ 为可以表征 CMOS 电路性能的某个函数， f_0 是与 $f(X)$ 相关的判断 CMOS 电路失效与否的一个临界值。 $I(X)$ 为判断一个电路失效与否的指示函数，定义为

$$I(X) = \begin{cases} 1, fail & f(X) \leq f_0 \\ 0, pass & f(X) > f_0 \end{cases} \quad (3-4)$$

对应到公式(3-3)中，当 $X \notin \Omega$ ，则 $I(X)=0$ ，CMOS 电路正常工作；当 $X \in \Omega$ ，则 $I(X)=1$ ，CMOS 电路失效。

用蒙特卡罗方法来估计失效率，关键是从 X 概率密度函数 $p(X)$ 中抽取 M 个样本 $\{X_1, X_2, \dots, X_M\}$ ，通过这些采样数据，可以用 SPICE 仿真得到对应的电路性能值 $\{f_1, f_2, \dots, f_M\}$ ，最后，通过与临界值 f_0 的比较，得到对应的指示函数值 $\{I_1, I_2, \dots, I_M\}$ ，则用蒙特卡罗算法估计的电路失效率可以表示为

$$\tilde{P}_f^{MC} = \frac{1}{M} \sum_{i=1}^M I(X_i) \quad (3-5)$$

\tilde{P}_f^{MC} 的方差为

$$\sigma_{\tilde{P}_f^{MC}}^2 = \frac{1}{M} \left(\tilde{P}_f^{MC} - P_f^{MC} \right)^2 \quad (3-6)$$

由于 $E\left(\tilde{P}_f^{MC}\right) = E\left(\frac{1}{M} \sum_{i=1}^M I(X_i)\right) = \frac{1}{M} \cdot M \cdot E(I(X_i)) = P_f$ ，因此， \tilde{P}_f^{MC} 是 P_f 的无偏估计量。

3.1.3 传统蒙特卡罗方法的局限性

用蒙特卡罗方法来估计电路失效率，实现简单，只需要建立 X 概率模型，并抽取 X 的若干样本，就可以得到失效率的估计值。另外，仿真结果误差可控，根据仿

真结果期待的最大相对误差，需要的样本数就可以定量或半定量的算出。根据估计理论，

$\left[\tilde{P}_f^{MC} - z_{\alpha/2} \cdot \sigma_{\tilde{P}_f^{MC}}, \tilde{P}_f^{MC} + z_{\alpha/2} \cdot \sigma_{\tilde{P}_f^{MC}} \right]$ 是电路失效率 P_f 置信水平为 $(1-\alpha)\%$ 的置信空间，令其等于区间 $\left[\tilde{P}_f^{MC} - \delta \tilde{P}_f^{MC}, \tilde{P}_f^{MC} + \delta \tilde{P}_f^{MC} \right]$ ，其中 δ 为误差百分比^[36]。代入公式(3-6)可得，

$$M = \left(\frac{z_{\alpha/2}}{\delta} \right)^2 \cdot \frac{1 - \tilde{P}_f^{MC}}{\tilde{P}_f^{MC}} \quad (3-7)$$

从公式(3-7)可以看出，当 \tilde{P}_f^{MC} 非常低时， M 与 \tilde{P}_f^{MC} 成反比。因此，用传统蒙特卡罗算法来估计高精度电路的失效率时，需要非常多的采样点，使得仿真的时间过长，仿真效率大大降低。例如，对于一个 10M 的 SRAM，在良率要满足一定标准的前提下，存储单元的失效率要小于 10^{-10} ，用传统蒙特卡罗算法来估计其失效率是不现实的。

3.2 重要性采样算法

为了弥补传统蒙特卡罗方法越来越明显的不足，在传统蒙特卡罗方法的基础上，人们提出了快速蒙特卡罗算法，也称作重要性采样算法（Importance Sampling, IS）。重要性采样算法的基本思想是，通过引入一个偏置函数，增大稀有事件在有限的样本中发生的比例，与传统蒙特卡罗算法相比，大大减少了抽取样本的数目，提高了仿真效率。

3.2.1 重要性采样算法原理

在用重要性采样算法来估计 CMOS 电路失效率时，要构造一个函数 $g(X)$ 来代替公式(3-3)中的 $p(X)$ 进行抽样。因此，公式(3-3)可以改写为

$$P_f = \int_{\Omega} p(X) \cdot dX = \int_{\Omega} \frac{p(X)}{g(X)} \cdot g(X) \cdot dX = \int_{\Omega} w(X) \cdot g(X) \cdot dX \quad (3-8)$$

其中， $w(X) = p(X)/g(X)$ 称为权重函数，所构造的 $g(X)$ 称为偏置函数或重要性密度函数。

通过从偏置函数 $g(X)$ 中抽取 M 个样本 $\{X_1, X_2, \dots, X_M\}$ ，得到样本权重

$\{w(X_1), w(X_2), \dots, w(X_M)\}$ ，与传统蒙特卡罗算法类似，利用重要性采样算法估计的电路失效率可以表示为

$$\tilde{P}_f^{IS} = \frac{1}{M} \cdot \sum_{i=1}^M I(X_i) w(X_i) \quad (3-9)$$

权重函数 $w(X)$ 用来调整采样点输出，以确保对失效率重要性采样估计是无偏估计。

表 3.1 对重要性采样与蒙特卡罗的方差收敛速度做了对比。

表 3.1 IS MC 与传统 MC 方差收敛速度对比^[28]

偏置函数	权重函数	方差收敛速度 V
$p(X) = g(X)$	$w(X) = 1$	$V_{IS} = V_{MC}$
$p(X) < g(X)$	$w(X) < 1$	$V_{IS} > V_{MC}$
$p(X) > g(X)$	$w(X) > 1$	$V_{IS} < V_{MC}$

因此，重要性采样算法的关键在于偏置函数 $g(X)$ 的选择，合理的偏置函数可以在保证仿真精度不受影响的前提下，减少仿真次数，缩短仿真时间，提高仿真效率；相反，如果偏置函数选择不合理，会造成重要性采样算法的仿真效率和精度比传统蒙特卡罗方法更差。

3.2.2 偏置函数典型构造方法

最简单的偏置函数 $g(X)$ ，可以构造为一个均匀分布，但是随着随机向量维数的增加，仿真效率大大降低。常见的偏置函数构造方法可分为 3 类^[28]：

1. 均值平移变换 (Mean Translation)

均值平移变换，又称为改进型重要性采样 (Improved Importance Sampling, IIS)，具有简单、有效、不受维数影响的优点。将随机变量 X 的概率密度函数 $p(X)$ 沿着某一向量 S 向失效区域平移，所得到的函数作为偏执函数 $g(X)$ 。假设 $S = (\mu_1, \mu_2, \dots, \mu_N)$ ，则 $g(X)$ 可以表示为

$$g(X) = p(X - S) = \prod_{i=1}^N \left(\frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} (x_i - \mu_i)^2 \right) \right) \quad (3-10)$$

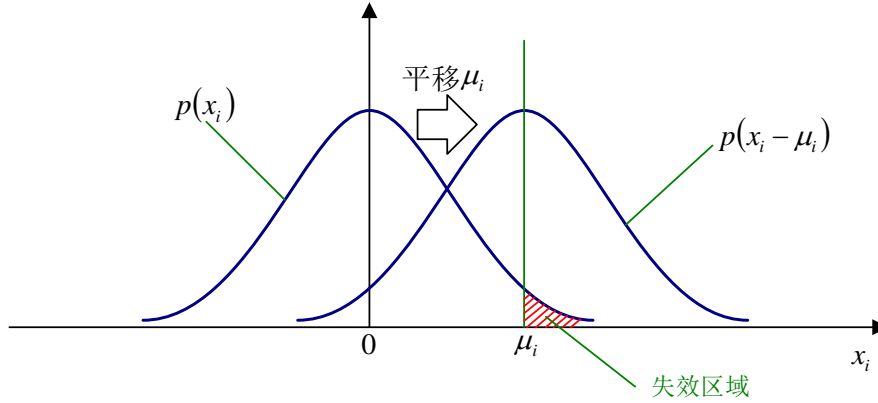


图 3.2 均值平移变换

如图 3.2 所示为随机向量 X 中各个随机变量 x_i 的概率密度函数 $p(x_i)$ 以及平移后得到的函数 $p(x_i - \mu_i)$ ，即 $g(x_i)$ 。可以看出，通过对 $g(x_i)$ 进行抽样，增大了抽取到失效区域中样本的概率。

均值平移构造 $g(X)$ 的关键是找到最理想的平移向量 S ，这个往往是一个比较复杂的过程。另外，偏置函数是由概率密度函数单向平移得到，当存在多个不连续的失效区域时，结果将会导致抽取的样本覆盖不充分。

2. 方差缩放变换 (Variance Scaling)

方差缩放变换，又称为传统重要性采样 (Conventional Importance Sampling, CIS)。对于单变量情况，如图 3.3 所示，该方法通过对变量 x 概率密度函数进行缩放来构造的偏置函数，缩放后得到的偏置函数形式为：

$$g(x) = \frac{1}{a} \cdot p\left(\frac{x}{a}\right) \quad (3-11)$$

通常取 $a > 1$ ，权重函数 $w(x) = a \cdot \frac{p(x)}{p(x/a)}$ ， $\frac{1}{a^2}$ 为随机变量 x 的方差缩放因子。

通过增加工艺参数偏差的方差来增大失效次数的比例。

同样，对于多维随机向量 $X = (x_1, x_2, \dots, x_n)$ 的情况，假设缩放向量为 $A = (a_1, a_2, \dots, a_n)$ ，缩放后得到的典型偏置函数形式为：

$$g(X) = A \cdot p(A \odot X) \quad (3-12)$$

式中， $A \odot X = (a_1 x_1, a_2 x_2, \dots, a_n x_n)$ ，通常取 $a_i < 1$ ， a_i^2 是随机变量 x_i 的方差缩放因子。

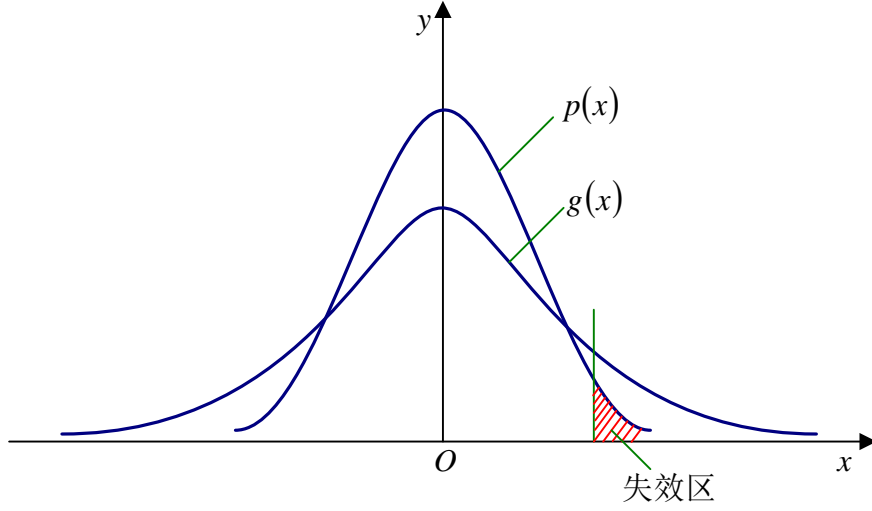


图 3.3 方差缩放变换

利用方差缩放变换构造偏置函数，对于 n 维随机向量 X ，缩放效果会在 n 维空间展开。将会导致重要性采样算法的效率随着 n 的增加而下降。

3. 指数扭转变换 (Exponential Twisting)

指数扭转是一种应用比较广泛的构造偏置函数方法，通过指数扭转，可以得到比较理想的偏置函数。基本思想是将所有可能的偏置函数归为一个参数类，然后在这一限制类内选择最优的指数变换 (Optional Exponential Change of Measure, OECM)。其一般形式如下：

$$g_a(x) = \frac{e^{ax} p(x)}{M(a)} \quad (3-13)$$

式中， $M(a) = E_f(e^{ax}) = \int e^{ax} p(x) dx$ ， $a \in R^d$ ， a 是扭转参数

在仿真过程中，为了方便选取，常将(3-13)所示的指数扭转偏置函数简化为下式表示：

$$g_a(x) = ce^{\theta x} f(x) \quad (3-14)$$

其中， c 是标准化常数。

3.3 混合重要性采样算法的实现

重要性采样算法的关键在于偏置函数的构造，理论上存在最优偏置函数，可实现零方差，即每个仿真实验能够计算出正确期望值。但是在实际应用中，由于涉及到真

实电路失效率，无法求得，只能取近似最优或次优的偏置函数。

在选择仿真算法时，不仅要考虑该算法方差的收敛情况，还要考虑算法本身的复杂度。前者决定了仿真所需的抽取样本个数，后者决定了单个样本需要的仿真时间，实际中要折中考虑。3.2.1 小节介绍了三种偏置函数的常用构造方法，构造理想的偏置函数的过程往往比较复杂，导致单个样本仿真时间过长。本小节基于常用的均值平移法的基础上，实现了文献[36]中的混合重要性采样算法（Mixture Importance Sampling, Mix-IS），所实现混合重要性采样算法可根据不同工艺的蒙特卡罗 MOS 管模型，手动选择模型中所定义的随机变量进行抽样。具有一定的可移植性，且有偏置函数构造过程简单，采样点分布全面，不受随机变量维数影响等优点。

3.3.1 混合重要性采样算法原理

均值平移法是构造偏置函数的常用方法之一，如图 3.2 所示，将工艺参数变量 x 的概率密度函数 $p(x)$ 向失效区域平移距离 μ_s ，即 $p(x - \mu_s)$ 。混合重要性采样算法更进一步，采用不同的分布混合构造偏置函数，对于单变量 x ，形式如下

$$g(x) = \alpha \cdot p(x) + \beta \cdot U(x) + (1 - \alpha - \beta) \cdot p(x - \mu_s) \quad (3-15)$$

式中， $U(x)$ 为均匀分布， α, β 为系数，并且满足 $0 < \alpha + \beta < 1$ 。 α 和 β 的大小取决于 μ_s ，一般情况下，偏移量 μ_s 取值越大， α, β 越小。对于多维随机变量 X ，混合重要性采样算法偏置函数与 3-15 类似。

如图 3.4 所示，当存在多个不连续的失效区域时，混合重要性采样算法的偏置函数可以覆盖到所有失效区域，抽取的样本更全面。

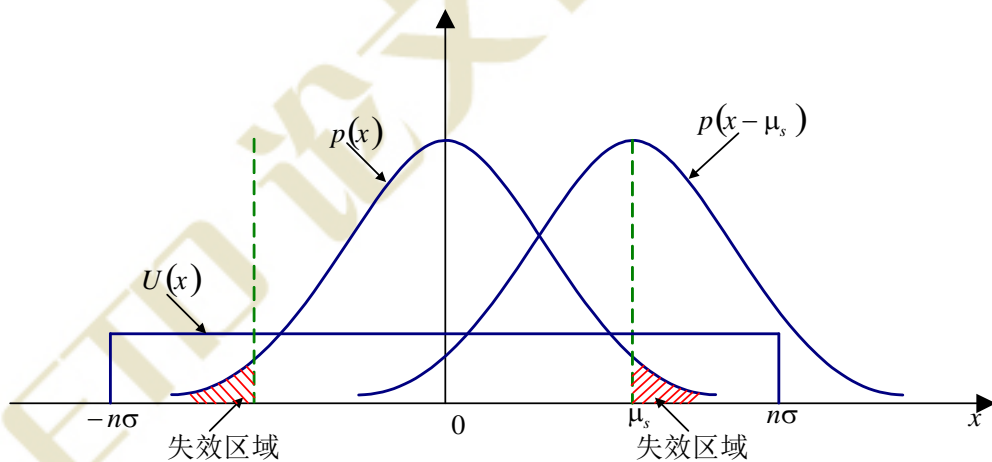


图 3.4 混合重要性采样算法原理

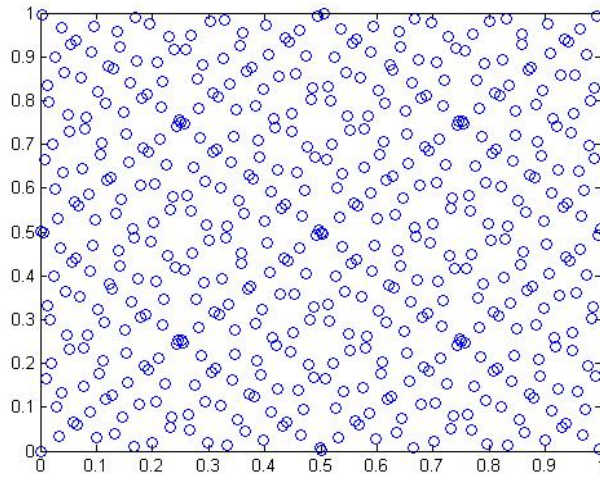
3.3.2 平移量 μ_s 的确定

对于一维随机变量 x ，确定平移量 μ_s 的方法如下，在随机变量 x 的分布空间均匀抽取 l 个样本 $(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_l)$ ，为了方便表述而又不失一般性，假设其中前 k 个样本落在失效区域，则 μ_s 的取值为这个 k 个样本的均值，即

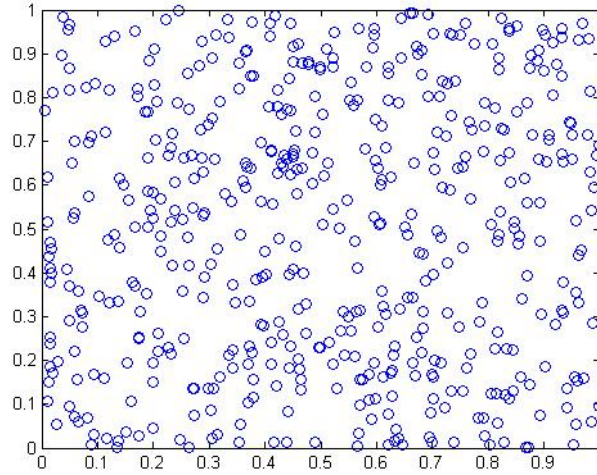
$$\mu_s = \frac{\sum_{i=1}^k \hat{x}_i}{k} \quad (3-16)$$

类似的，对于 n 维随机向量 $X = [x_1, x_2, \dots, x_n]$ ，确定平移向量 $S = [\mu_1, \mu_2, \dots, \mu_n]$ ，即确定 x_i 对应的平移量 μ_i 。在 X 分布空间均匀抽取 L 个样本 $(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_L)$ ，假设前 K 个为失效样本，则 μ_i 是这 K 个样本中对应 x_i 样本值的平均值。

为了保证抽取样本在参数空间分布的均匀性，在抽样时，采用Sobol序列法来取代均匀分布的随机序列法。Sobol序列是一种准随机低差异序列，其分布性均匀性好，耗时少。如图3.5所示，在边长为1的正方形内，分别用Sobol序列法和均匀分布随机序列法抽取512个采样点，与均匀分布随机序列法相比，Sobol序列法采样更均匀^[37]。



(a) Sobol 序列法



(b) 均匀分布随机序列法

图 3.5 Sobol 序列法与均匀分布随机序列法的采样

3.3.3 抽样方法

n 维随机向量 X 中各元素之间相互独立，因此， X 的偏置函数等于各元素偏置函数的乘积，即

$$g(X) = \prod_{i=1}^n g(x_i) \quad (3-17)$$

在 $g(X)$ 上对 X 抽样，也就是在各 $g(x_i)$ 上对 x_i 抽样。由于 $g(x_i)$ 是由均匀分布和高斯分布按一定比例混合构造的分布，按照常规产生高斯随机数的抽样方法难以实现。本章实现的混合重要性采样算法，通过在区间 $(0,1)$ 中产生一个服从均匀分布的随机数 r ，然后根据 r 与 α 和 $\alpha + \beta$ 在区间 $(0,1)$ 中的位置关系选择不同的分布函数进行样本抽样，具体如下：

Case 1: 如果 $r < \alpha$ ，按照标准正态分布 $p(x_i)$ 抽取 x_i 样本；

Case 2: 如果 $\alpha < r < \alpha + \beta$ ，按照均匀分布 $U(x_i)$ 抽取 x_i 样本；

Case 3: 如果 $\alpha + \beta < r < 1$ ，按照高斯分布 $p(x_i - \mu_i)$ 抽取 x_i 样本；

选取到不同的抽样概率密度函数的分布律，如表 3.2 所示。从抽样概率密度函数数学期望的角度来看，抽样函数就是公式 3-15。

表 3.2 抽样概率密度函数的分布律

抽样概率密度函数	$p(x_i)$	$U(x_i)$	$p(x_i - \mu_i)$
概率	α	β	$1 - \alpha - \beta$

3.3.4 仿真效率比较

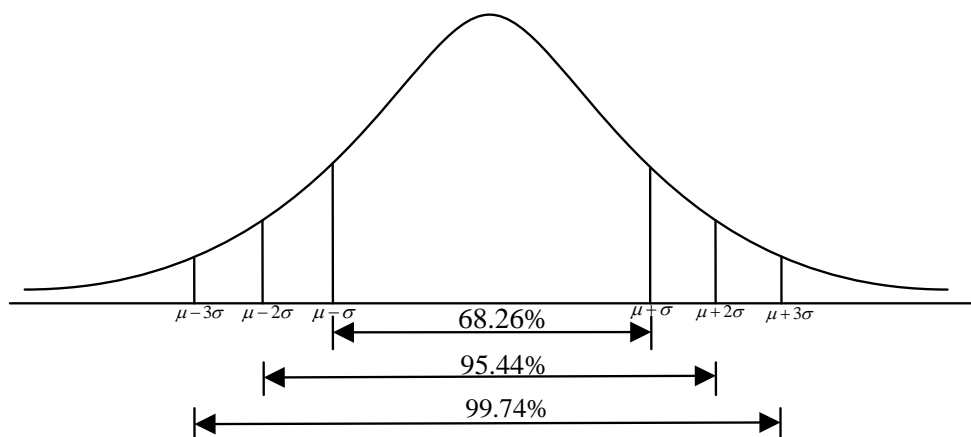


图 3.6 高斯分布

假设随机变量 x 服从高斯分布，即 $x \sim N(\mu, \sigma^2)$ 。尽管 x 的取值范围是 $(-\infty, \infty)$ ，但是它的值落在区间 $(\mu - n\sigma, \mu + n\sigma)$ ， $(n=1,2,3...)$ 内的累积概率是一个定值，如图 3.6 所示。因此，在估计 SRAM 良率时，常用 $n\sigma$ 来表示良率为对应区间内的累积概率。

本小节通过对 6 管 SRAM 存储单元的 SNM 分析，来比较混合重要性采样算法与蒙特卡罗算法的优势。如图 3.7 所示为，执行 60000 次蒙特卡罗仿真，得到 SNM 分布的概率密度拟合曲线。可以看出，SNM 服从高斯分布，且均值为 $\mu = 374.395\text{mv}$ ，标准差 $\sigma = 15.762\text{mv}$ 。

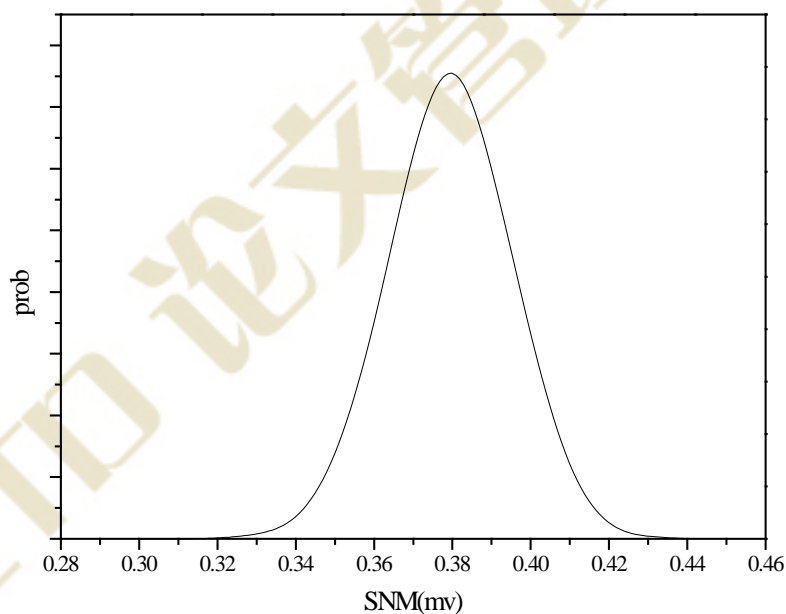


图 3.7 通过 60000 次蒙特卡罗仿真得到的 SNM 概率密度曲线

表 3.3 蒙特卡罗算法和混合重要性采样算法仿真速度比较

	N_{MC} #Monte Carlo Simulations	N_{IS} #Mix IS Simulations	Speedup $= N_{MC} / N_{IS}$
1σ	8.26e2	2.8e2	3e0
2σ	8.06e3	6.8e2	1.2e1
3σ	1.42e5	9.2e2	1.54e2
4σ	6.10e6	9.3e2	6.56e3
5σ	6.74e8	1.81e3	3.72e5
6σ	1.92e11	4.87e4	3.94e6

表 3.3 列出了 SNM 分布于 $1\sigma \sim 6\sigma$ 区间的情况，在置信度为 95%，精度为 90% 条件下的，蒙特卡罗算法和混合重要性采样算法所需要的仿真个数。第二列和第三列分别为用蒙特卡罗和混合重要性采样来估计 SNM 分布在 $1\sigma \sim 6\sigma$ 的概率，所需要的仿真个数，第四列为两者速度的比较。可以看出，在良率较低条件下，两者仿真速度差距不大，混合重要性采样算法的速度优势并不明显。随着良率的提高，混合重要性采样算法的仿真速度优势越来越明显，可以达到数个数量级的速度提高。

表 3.4 则列出了图 3.7 中 SNM 概率密度曲线和混合重要性采样算法仿真得到的 $1\sigma \sim 6\sigma$ 对应上限的值，其中，混合重要性采样算法执行了 3000 次。通过对第二列和第三列中数据的比较，可以看出，混合重要性采样算法分析结果与 SNM 的统计结果的误差保持在 1% 以内。因此，混合重要性采样性算法可以用来有效而精确地对小概率事件进行分析。

表 3.4 SNM 概率密度曲线和混合重要性采样算法仿真得到的 $1\sigma \sim 6\sigma$ 分布区间对应上限的值

	SNM 概率密度曲线	Mix-IS 仿真
1σ	358.6329mv	360.1mv
2σ	342.8710mv	344.1mv
3σ	327.1092mv	328.9mv
4σ	311.3474mv	311.9mv
5σ	295.5856mv	295.8mv
6σ	279.8237mv	280.1mv

3.4 本章小结

本章首先介绍了蒙特卡罗方法的基本思想，以及在 CMOS 电路中的应用。分析了随着 SRAM 容量的不断增加，传统蒙特卡罗对大容量 SRAM 阵列可靠性进行仿真存在采样点多，收敛慢，仿真时间极长等问题。因此，为了解决上述问题，人们提出了混合重要性采样算法，通过引入偏置函数，增大稀有事件在样本总数中发生的个数，从而减少采样个数，缩短仿真时间。本章介绍了 3 种常用的构造偏置函数，分析了重要性采样算法偏置函数构造难度大的问题。最后，为了解决混合重要性采样算法的仿真效率对偏置函数的依赖，本章基于混合重要性采样算法，实现了一种快速蒙特卡罗方法。并通过仿真验证算法的有效性和可靠性。

第四章 SRAM 测试片设计

随着 CMOS 新工艺的不断复杂化，给电路设计带来了巨大的隐患。因此，在电路设计阶段，要尽可能充分考虑到工艺变化对电路的影响。本章基于 40G 工艺，设计一颗 32 兆（Mega, M）的 SRAM 测试芯片，用来验证该工艺的稳定性。在设计过程中，针对受工艺参数变化影响严重的 SRAM 存储单元，利用第三章所实现的快速蒙特卡罗算法，选择最稳定的存储单元尺寸，应用到测试片的设计中，使测试片在工艺参数变化的影响下，依然能够保持足够高的良率。

4.1 SRAM 测试片总体结构

本文所设计的 SRAM 容量为 32M，由 16 个独立的 2M 单元组成，布局结构如图 4.1 所示。

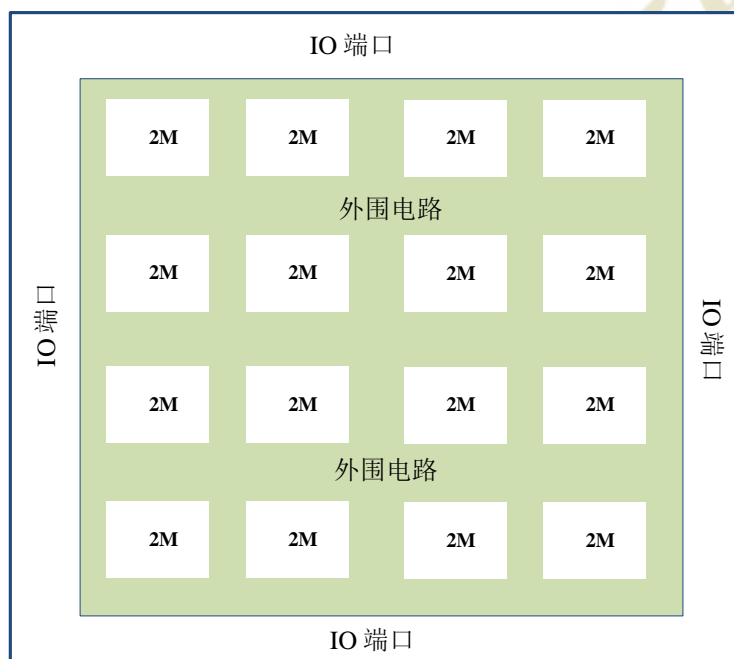


图 4.1 芯片布局结构

从图中可以得到，设计主要包含三个部分：

- 1) 2M 单元设计；
- 2) 外围电路

3) 输入输出端口布局设计;

将 32M 测试芯片分割为 2M 可以更好地进行电源管理, 时序控制, 并且也能很好地减少设计复杂度, 方便进行后期错误测试。每个 2M 存储单元为 2 个 1M 存储单元组成, 因此 1M 存储单元是存储块部分设计基础。

外围电路的设计主要是完成这些功能: 驱动端口与内部 2M 单元, 选择输出到外部的单元, 选择启动的单元, 在驱动输出端口时进行第一级升压, 如图 4.2 所示。

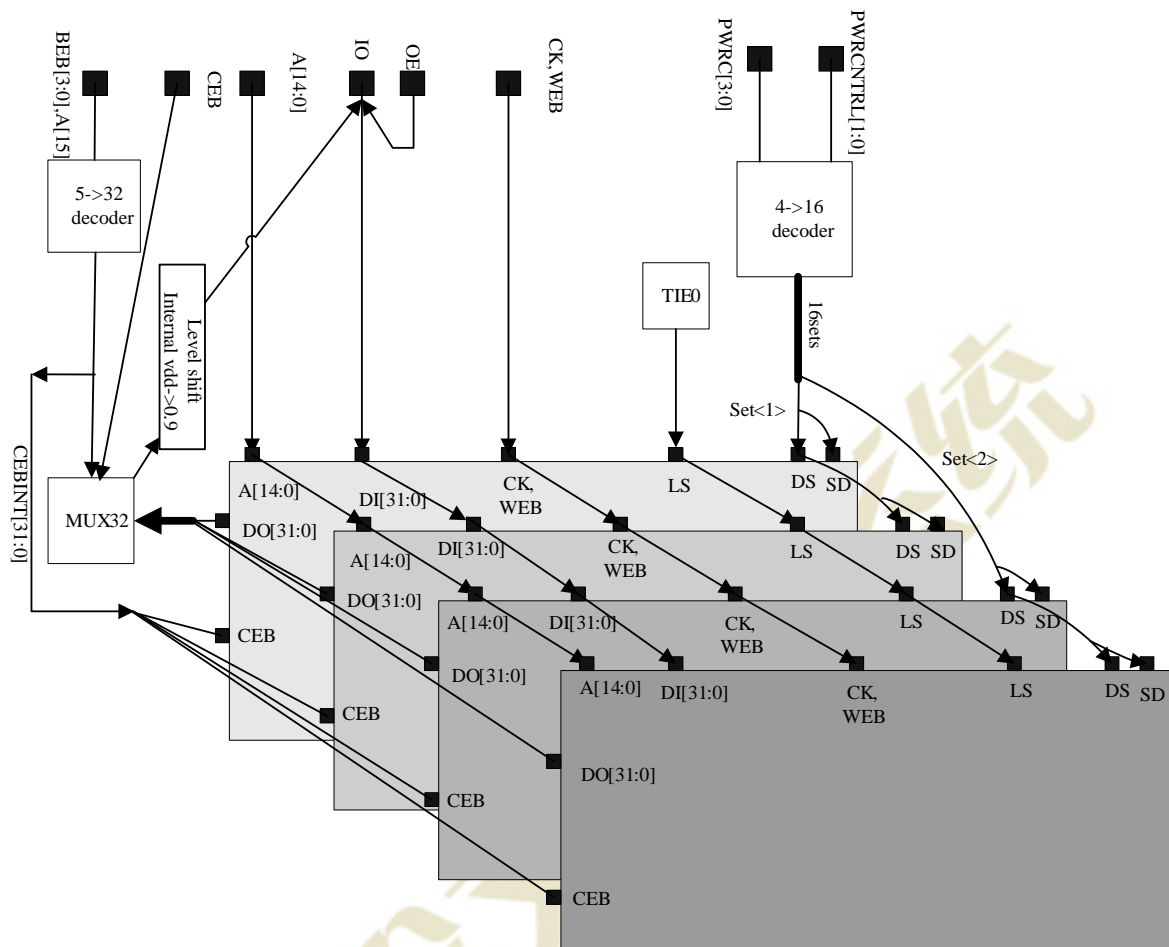


图 4.2 外围电路结构图

输入输出端口布局设计, 主要完成信号端口分布, 电源端口分布以及 ESD 保护单元的分布与配置。

为了方便测试, 并且更好地进行电源管理。本文所设计的 SRAM 测试片中存储阵列和外围电路的电源分开。如图 4.3 所示, (a)中 VSS1 为存储单元和外部逻辑电路的接地电压, VSS2 为输入输出端口接地电压。(b)中 VDDC 为 1M 存储单元中存储阵列的工作电压, VDDP1 为 1M 存储单元中存储阵列的外围电路的工作电压, VDDP2

为外围逻辑电路的工作电压，VDDP3 为升压电路工作电压，VDDP4 为输入输出端口工作电压。

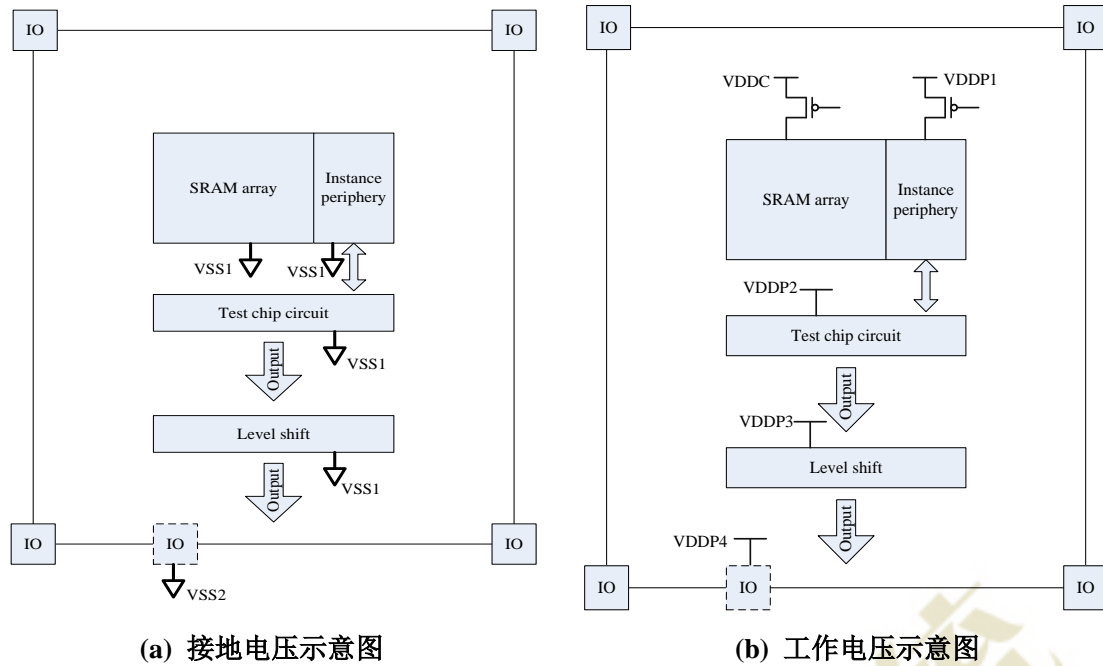


图 4.3 电源管理示意图

4.2 SRAM 存储单元设计

在一个 SRAM 芯片上，存储阵列占到了整个面积的 80% 以上。为了降低面积开销，对于每个工艺，存储单元都会采用可生产的最小尺寸的器件。由于随机掺杂波动等工艺偏差与器件的面积成反比关系，因此，相比于其他电路，存储单元的稳定性更容易受到工艺参数偏差影响^[38]。

4.2.1 存储单元设计方法

SRAM 存储单元是整个测试片设计中最为核心的模块，它的可靠与否是关系到测试片良率的关键因素。根据第二章的相关分析，稳定的存储单元应该有较强的 SNM、RNM 和 WM，但是以上关系之间往往是相互制约的，即在 SRAM 存储单元中，读操作的稳定性越高，数据的写入将越困难；相反，数据写入越容易，读操作的稳定性越差。因此，在存储单元设计中，为了保证读操作和写操作的稳定性，应该折中考虑，平衡上述几个性能。

本文所设计 SRAM 测试片存储单元结构采用的是最常用的 6 管架构。图 4.4、图 4.5、图 4.6 分别为在 40G 工艺下，SNM、RNM、WM 与 PR、CR 之间的三维关系图。

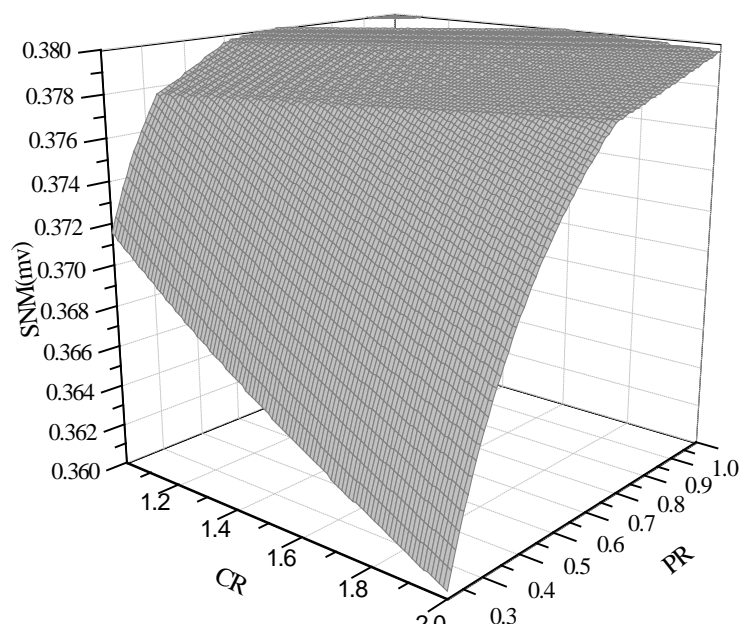


图 4.4 SNM 与 PR、CR 关系

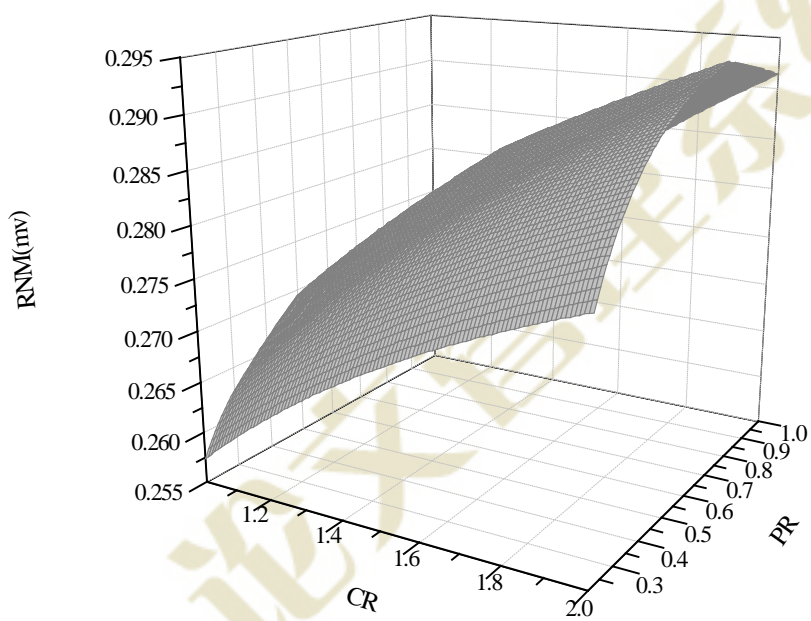


图 4.5 RNM 与 PR、CR 关系

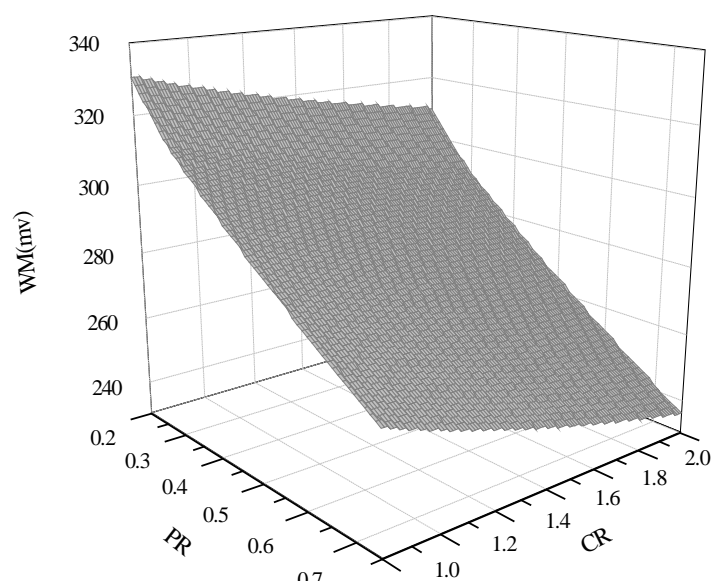


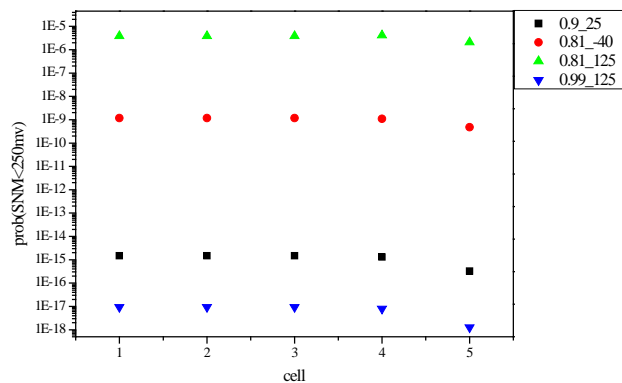
图 4.6 WM 与 PR、CR 关系

SRAM 存储单元设计要满足读操作 CR (>0.82) 和写操作 PR (<1.51) 的要求, 综合考虑来选择合适的 PR 和 CR。结合图 4.4~图 4.6, 存储单元晶体管尺寸选择如表 4.1 所示。对应的 CR 与 PR 满足要求, 并且有较大的 SNM、RSNM、WM。

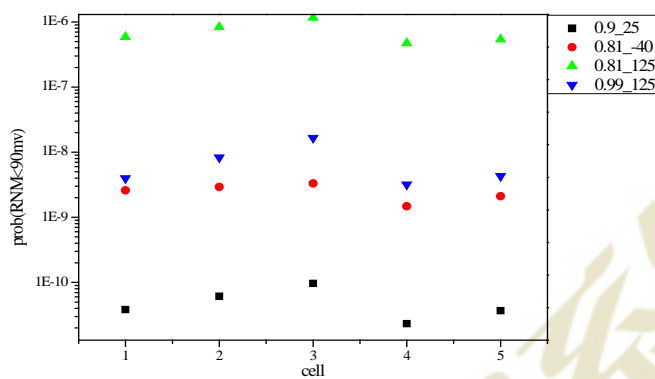
表 4.1 SRAM 存储单元晶体管尺寸 (nm)

晶体管	Cell1	Cell2	Cell3	Cell4	Cell5
MP1/MP3 (W _{ppu} /L _{ppu})	70/55	70/55	70/55	70/55	75/55
MN2/MN4 (W _{npd} /L _{npd})	235/55	235/55	235/55	245/55	235/55
MN5/MN6 (W _{npg} /L _{npg})	175/65	185/65	195/65	185/65	185/65

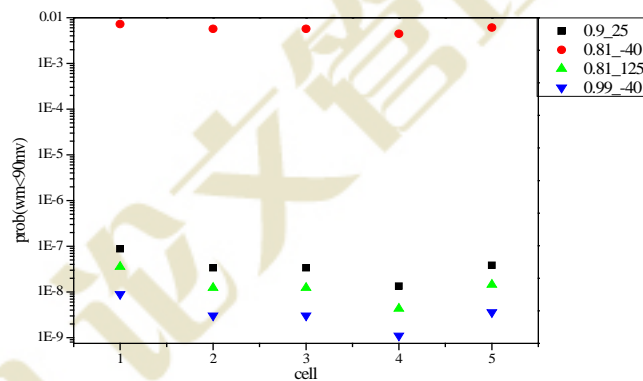
利用本文第三章所实现的快速蒙特卡罗算法, 从统计的角度对表 4.1 中所列 5 种存储单元的 SNM、RNM、WM 进行比较, 所得结果如图 4.7 所示, 其中 0.9_25 表示工作电压为 0.9V, 温度为 25°C, 其他同理。图(a)为不同温度和电压下, 存储单元 SNM 小于 250mV 的概率, 相比其他 cell, cell5 在数据保持模式下, 稳定性最好。图(b)为不同温度和电压下, 存储单元 RNM 小于 90mV 的概率, cell4 的读操作稳定性最好。图(c)为不同温度和电压下, 存储单元 WM 小于 90mV 的概率, cell3 的写操作稳定性最好。对于 SRAM 来说, 最主要的失效来源是读操作失效, 因此, 综合考虑, 选择 cell4 作为本文所示设计测试片的存储单元尺寸。



(a) 温度和电压对 SNM 的影响



(b) 温度和电压对 RNM 的影响



(c) 温度和电压对 WM 的影响

图 4.7 温度和电压对 SRAM 存储单元性能的影响

4.2.2 存储阵列设计

对于相同存储容量的存储器，一根字线和位线上所连接的存储单元个数越多，字

线和位线的负载电容也就越大，严重影响到存储器所消耗的功率和读取数据的速度。因此，为了减小字线和位线上连接的存储单元个数，在设计时，通常会对存储器的存储阵列进行分割，减小字线和位线的负载电容，以达到降低功耗和提高速度的目的。

本文所设计的 32M 存储器是以 1M 存储单元为基础，1M 存储单元的存储阵列被分割成 16 个 128×512 的存储阵列，即单根字线上最多连接 512 个存储单元，单根位线上最多连接 128 个存储单元，如图 4.8 所示。每个 128×512 对应着各自的局部控制逻辑电路、字线驱动电路、局部输入输出电路。全局控制逻辑电路用于选择所要访问的 128×512 存储阵列，并且，在读操作时，将局部输入输出电路的输出数据送到全局输入输出电路；在写操作时，将全局输入输出电路的输入数据送到局部输入输出电路。另外，为了保证长字线的驱动能力，在字线上设计有缓冲器。

128X512	字线驱动	128X512
局部输入输出电路	局部控制逻辑	局部输入输出电路
128X512	字线驱动	128X512
128X512	字线驱动	128X512
局部输入输出电路	局部控制逻辑	局部输入输出电路
128X512	字线驱动	128X512
128X512	字线驱动	128X512
局部输入输出电路	局部控制逻辑	局部输入输出电路
128X512	字线驱动	128X512
128X512	字线驱动	128X512
局部输入输出电路	局部控制逻辑	局部输入输出电路
128X512	字线驱动	128X512
全局输入输出电路	全局控制逻辑	全局输入输出电路

图 4.8 1M SRAM 分割图

4.3 灵敏放大器设计

在 SRAM 电路中，灵敏放大器是关键组成部分之一。灵敏放大器通过将 BL 与

BLB 之间微小的电压差 ΔV 放大，并且快速地送到输出端口。通常情况下， ΔV 远低于电源电压，这样做有效地加快了读取速度，并且减小了读操作功耗。同时又要保证灵敏放大器能够正确识别这个 ΔV ，因此，一个快速、可靠的灵敏放大器对于 SRAM 的读取速度以及良率有着重要的作用。

4.3.1 灵敏放大器架构

为了保证 SRAM 能够正确读取，灵敏放大器必须满足一定的要求。首先，输入端的最小分辨电压必须小于 ΔV ；其次，输入端为最小电压时，灵敏放大器必须在放大时间内提供稳定的输出^[39]。按照电路中是否包含锁存器结构，可以将灵敏放大器分为两类：线性灵敏放大器和锁存灵敏放大器。

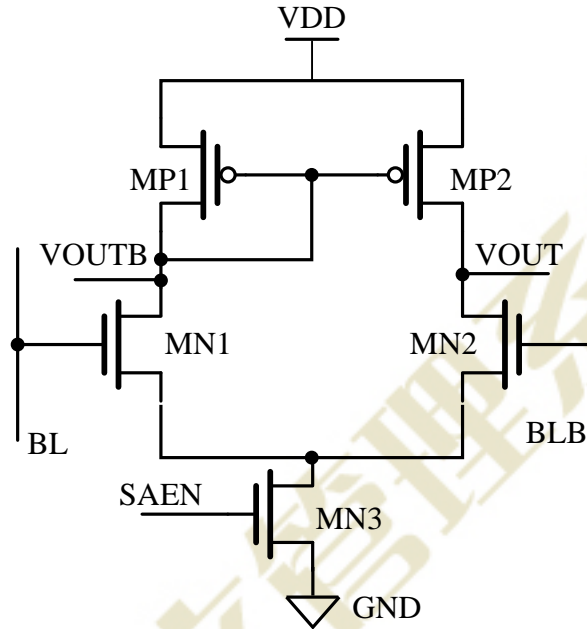


图 4.9 线性灵敏放大器电路

如图 4.9 所示，为线性灵敏放大器的基本结构。MN1 和 MN2 的栅极连接位线，作为放大器的输入端；MP1 和 MP2 构成一组电流镜，SAEN 为线性放大器的启动信号，VOUT 和 VOUTB 为放大器的输出端。当 SAEN 信号为低时，VOUT 和 VOUTB 均被预充至高电平，此时放大器不工作。当位线 BL 与 BLB 形成足够的电压差时，SAEN 被拉至高电平，MN1、MN2 和 MN3 均开启，VOUTB 和 VOUT 下拉，产生通过 MN1 和 MN2 的电流，在电流镜的作用下，位线上微小的电压差被放大到输出端。线性灵敏放大器的增益可以表示为

$$A = G_m * R_{ON} / R_{OP} \quad (4-1)$$

其中, G_m 为 MN1/MN2 的跨导, R_{ON} 为 MN1 和 MN2 的输出电阻, R_{OP} 为 MP1/MP2 的输出电阻。因此其增益主要由 MN1 和 MN2 决定。

线性灵敏放大器对输入噪声具有较好的抑制性, 而且结构简单^[10]。缺点是存在静态功耗, 速度较慢, 并且结构上并未完全匹配。

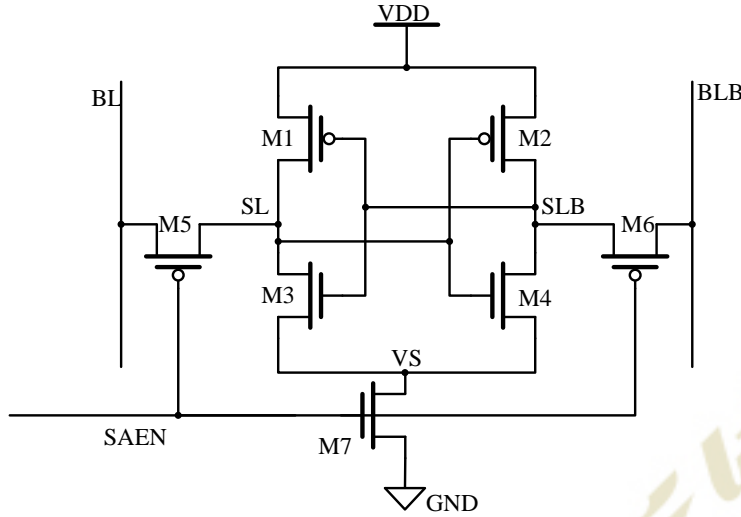


图 4.10 锁存结构电压灵敏放大器

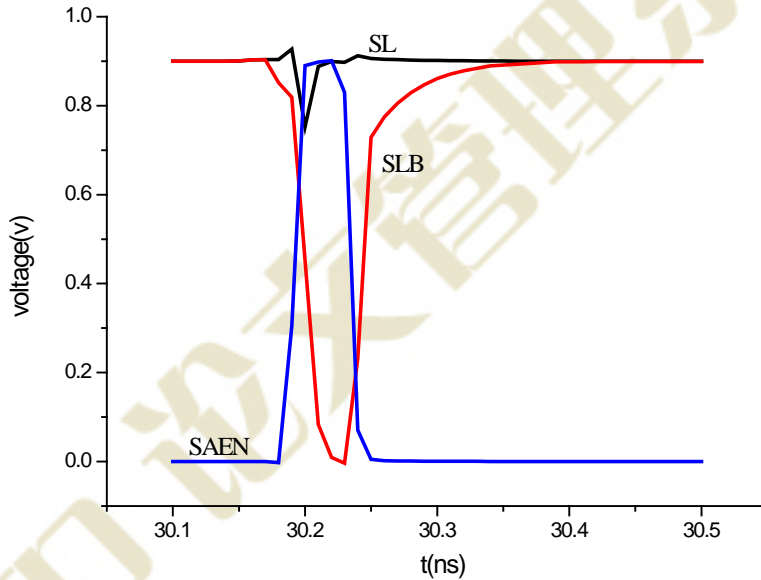


图 4.11 锁存灵敏放大器正常工作波形图

与线性灵敏放大器相比, 锁存灵敏放大器具有速度快, 功耗低等优点, 为业界广泛使用。本文所设计的 SRAM 中灵敏放大器采用就是锁存结构的电压灵敏放大器,

如图 4.10 所示。基本工作原理如下，初始状态，SAEN 为低电平，PMOS 管 M5 和 M6 打开，位线电压可以传输到节点 SL 与 SLB 处，NMOS 管 M7 关闭，此时放大器无效，不存在电流通路，内部处于高阻态。当位线形成足够的电压差后（假设 $BL=VDD$ ， $BLB=VDD-\Delta V$ ），SAEN 变为高电平，M5 和 M6 关闭，切断了放大器与位线负载电容的联系，M7 打开，M4 栅极电压比 M3 栅极电压高，因此，M4 的导通电流比 M3 的大，SLB 点放电速度比 SL 点更快，SLB 比 SL 早低于 M1 与 M2 构成的反相器的翻转电压，开启了交叉耦合反相器的正反馈，使 SL 点快速升高到 VDD ，SLB 点电压则被拉到低，在 SL 与 SLB 之间形成了全摆幅输出电压，仿真波形图如图 4.11 所示。

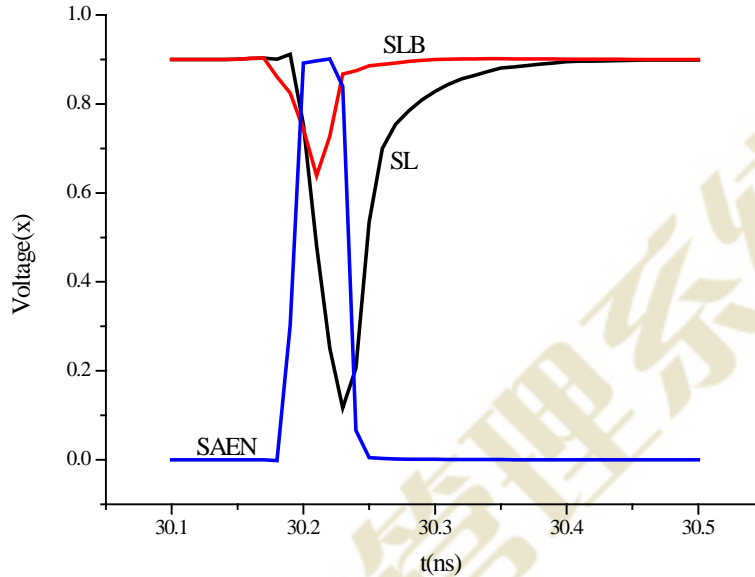


图 4.12 锁存灵敏放大器失效波形图

理论上，对于左右完全匹配的灵敏放大器，只要初始电压差 $\Delta V > 0$ 就能保证被正确放大。但是，由于工艺参数的随机波动导致晶体管阈值电压的随机波动，如果出现晶体管 M3 的阈值电压比 M4 的阈值电压低的情况，尽管 M4 栅极电压比 M3 栅极电压高，仍然有可能出现 M3 的导通电流比 M4 的大，此时，SL 放电比 SLB 更快，导致灵敏放大器往错误的方向放大数据，仿真波形图如图 4.12 所示。因此，为了保证灵敏放大器的稳定性，都会要求 ΔV 必须大于一个能使灵敏放大器正确工作的最小分辨值。在不考虑特殊效应的情况下，可以通过简单的计算来推出这个最小分辨值^[5]：

初始状态，M1 和 M2 关闭，M3 和 M4 处于饱和状态，则有

$$Id_3 = \beta(VDD - \Delta V - VS - V_{th3})^2 W_3 / L_3 \quad (4-2)$$

$$Id_4 = \beta(VDD - VS - V_{th4})^2 W_4 / L_4 \quad (4-3)$$

考虑沟道宽 W ，沟长 L 引起的阈值电压 V_{th} 波动。如果要求灵敏放大器正确工作，必须满足

$$Id_3 \leq Id_4 \quad (4-4)$$

将(4-2)和(4-3)代入(4-4)，化简得：

$$\Delta V \geq (VDD - VS) \left(1 - \sqrt{\frac{W_4 / L_4}{W_3 / L_3}} \right) + V_{th4} \sqrt{\frac{W_4 / L_4}{W_3 / L_3}} - V_{th3} \quad (4-5)$$

设 $V_{th3} = V_{th}$ ， $V_{th4} = V_{th} + \Delta V_{th}$ ， $W_4 / L_4 = m$ ， $W_3 / L_3 = n$

$$\text{则有 } \Delta V \geq (VDD - VS - V_{th}) \left(1 - \sqrt{m/n} \right) + \Delta V_{th} \sqrt{m/n} \quad (4-6)$$

从(4-6)式可以看出，在实际流片过程中，器件尺寸和阈值电压的波动都会影响到灵敏放大器最小分辨电压，为了保证灵敏放大器的良率，都会给最小分辨值留有一定的裕度，这种做法是以增大面积开销和降低性能为代价。因此，在灵敏放大器的设计过程中，要做好稳定性、面积以及性能的折中。

4.3.2 灵敏放大器参数确定策略

当灵敏放大器的架构确定之后，放大器的稳定性取决于晶体管的尺寸。图 4.10 中灵敏放大器的晶体管尺寸及错误率如表 4.2 所示，采用 45nm 模型^[40]。

表 4.2 灵敏放大器晶体管尺寸

单元	最小尺寸	优化后的尺寸
M1/M2	120nm/40nm	240nm/45nm
M3/M4	120nm/40nm	360nm/90nm
M5/M6	120nm/40nm	240nm/45nm
M7	120nm/40nm	120nm/45nm
错误率 (%)	6.29	0.2

M3 与 M4 阈值电压的不匹配是导致灵敏放大器失配最主要的原因，为了减小工艺参数偏差对 M3 和 M4 阈值电压的影响，在确定灵敏放大器各晶体管尺寸时，应该将这两个晶体管尺寸设计的大一些。

M7 的宽度对于灵敏放大器的速度影响最大，通过增大 M7 的宽度，减小放电延

时，加快灵敏放大器的速度。M5、M6 的宽度对于读操作速度也有一定的影响，增大 M5 和 M6 的宽度，减小位线到 SL 和 SLB 点电压的传输延时，对于提高 SRAM 读操作速度有一定的帮助。灵敏放大器开始工作时，M1 和 M2 是关闭的，因此对于整个灵敏放大器的失调的影响不大。

图 4.10 中锁存灵敏放大器的晶体管最终尺寸如表 4.3 所示。

表 4.3 灵敏放大器晶体管尺寸

单元	尺寸	与最小尺寸倍数
M1/M2	840nm/80nm	7/2
M3/M4	1.68um/160nm	14/4
M5/M6	1.6um/65nm	13/1.625
M7	1.8um/60nm	15/1.5

4.4 外围电路设计

在 SRAM 电路中，除了存储单元和灵敏放大器以外，还有很多重要的外围电路。比如，行译码电路，列复用电路，输入缓冲电路，输出缓冲电路等。这些电路的共同工作，才能保证 SRAM 的正常操作。

4.4.1 行译码器

SRAM 能根据外部随机输入的地址进行存取，因此必然要有地址译码电路。地址译码电路的速度和功耗对 SRAM 的整体性能有着重要的影响。译码电路的结构一般都较为简单，针对输入地址位数的不同，可以用传统的 3-8 译码器、2-4 译码器、1-2 译码器来进行组合。当 SRAM 规模比较大时，为了减小面积，减少金属走线，通常都会有多级译码。本文所设计的测试片采用三级译码，第一级，选中 32M 中所要操作的 1M 存储单元，需要 5 位地址信号。第二级，选中所要 1M 存储单元中 128×512 单元，由图 4.6 所示，需要 3 位行地址；第三级，选中所要驱动的字线，需要 7 位行地址，第三级行译电路示意图如图 4.13 所示。译码电路共有三组输出，每个字线驱动电路通过三输入与非门，与每组中的一条连接，则一共可以驱动 $8 \times 4 \times 4 = 128$ 条不同的字线，如图 4.14 所示。

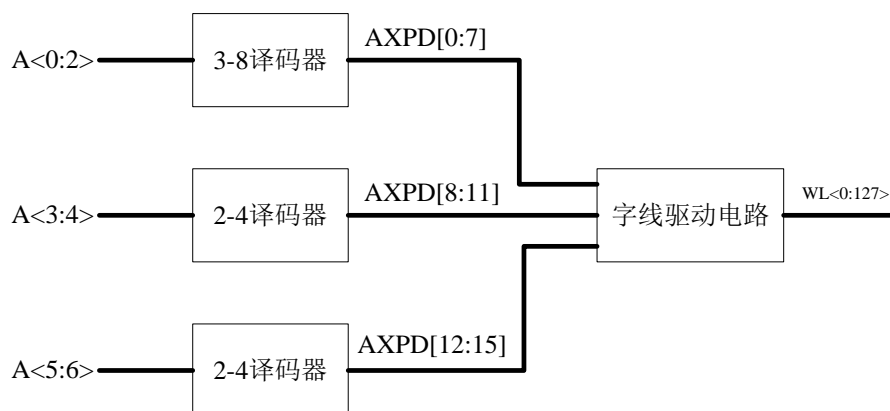


图 4.13 第三级行地址译码电路示意图

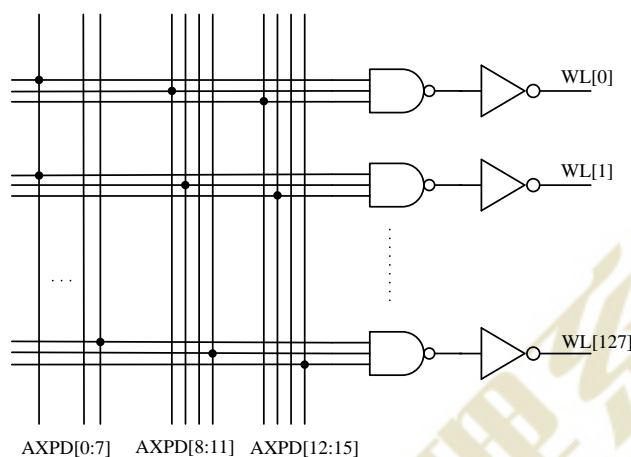


图 4.14 字线驱动电路

4.4.2 列复用电路

SRAM 存储器中，通过字线和位线来选中所要进行操作的存储单元，字线的选中主要是依靠行译码电路和字线驱动电路，而位线的选中要用到列复用电路。列复用电路是根据控制信号选通相应的开关，将存储单元的位线与读写数据线连接起来，从而实现列选中，如图 4.15 所示。对于本文中 1M 存储单元，共有 1024 列，32 个并行 IO 口，因此要实现 32 选 1 的列复用电路。本文中的列复用电路采用两层设计。第一层包含 4 个 8 选 1 选择电路；第二层包含 1 个 4 选 1 选择电路。

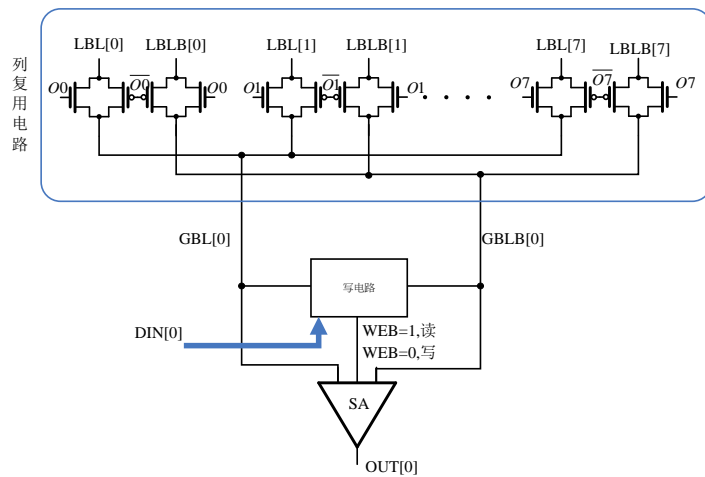


图 4.15 列复用电路

4.4.3 输入缓冲电路

输入缓冲电路采用常用的电路形式，主要由驱动电路和触发器组成，如图 4.16 所示。主要作用是 SRAM 写操作时，将输入的数据转为双端口，驱动被选中的位线，因此在设计时，需要设定合适的晶体管尺寸，既能保证有足够的驱动能力使打开的存储单元翻转，又不会造成面积的浪费。

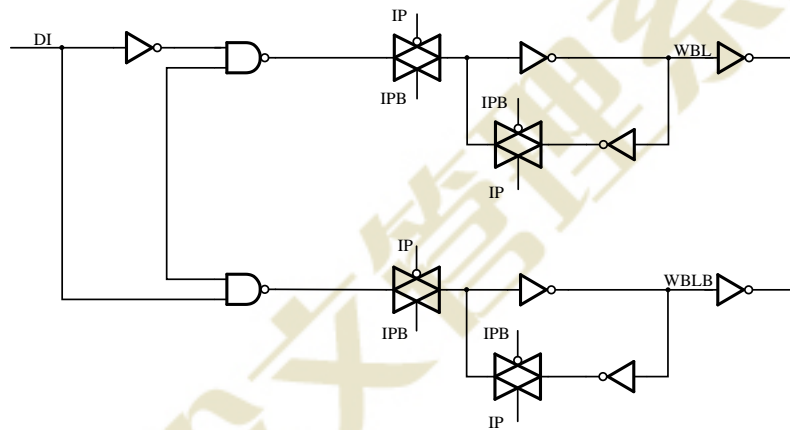


图 4.16 输入缓冲电路

4.4.4 输出缓冲电路

输出缓冲电路的主要作用是驱动输出，一般情况下，输出数据总线是被共用的，因此，需要有一个可供选择的高阻态，如图 4.17 所示。OE 为输出使能信号，当 OE=0 时，OEB=1，则与 DO 连接的上拉管和下拉管均关闭，DO 输出为高阻态。对于上述的上拉管和下拉管，在设计时要根据 DO 后的负载，调整晶体管尺寸。

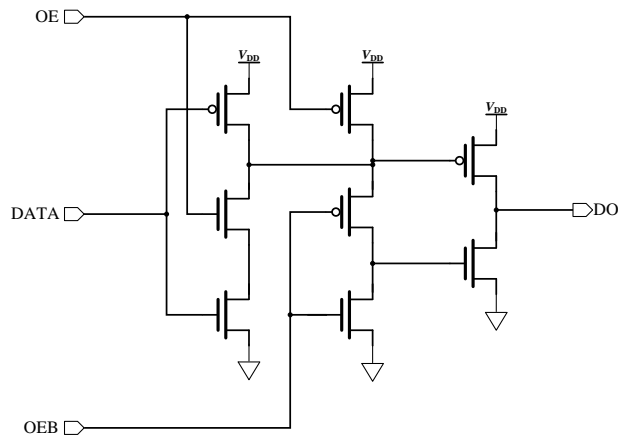


图 4.17 输出缓冲电路

4.4.5 升压电路

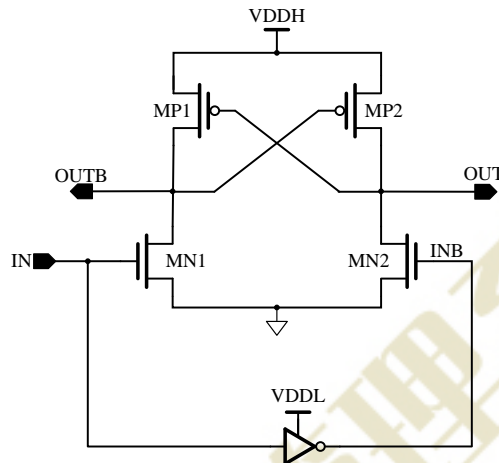


图 4.18 电压转换电路

考虑到测试芯片低压工作特性，为使芯片低压失效出现在内部 6 管存储单元上，而不是因为驱动输出端口电压过低造成的。当测试低压特性时，内部电压会不断下降，如没有升压电路，则驱动端口的电压会不断下降，而导致驱动失败。增加后则是会始终保持驱动端口的是一个较高的电压 $VDDH$ ，而不是 $VDDL$ ，如图 4.18 所示。当 IN 为 0 时， INB 为 $VDDL$ ，使 $MN2$ 导通，将 OUT 点下拉为 0，同时 OUT 点下拉为 0 会开启 $MP1$ ，对 $OUTB$ 点充电至高电位 $VDDH$ 。此电路不论输出 0，还是 1，均先用 $VDDL$ 驱动的 NMOS 下拉，反馈充电另一边来操作，避免了 $VDDL$ 与 $VDDH$ 之间的漏电。

4.5 本章小结

本章设计了一颗 32M 的 SRAM 测试芯片，每个最基础单元为 1M，而每 2M 单元实现独立可测，每 2M 的存储阵列电源可以单独关闭，降低整个芯片的漏电流对于电源的影响；存储阵列和外围电路的电源实行分开管理；32 个双向并行 IO 口。芯片的存储单元采用的最常用的 6 管结构，通过对 SNM、RNM、WM 与 CR 和 PR 三维关系的仿真，初步确定了 5 组晶体管尺寸不同的存储单元，利用第三章所实现的快速蒙特卡罗方法对这 5 组存储单元的 SNM、RNM、WM 进行比较，选择最稳定的存储单元应用到测试片设计中。灵敏放大器采用的是锁存结构，本章分析了灵敏放大器的失效机制，给出了灵敏放大器尺寸调整策略。最后，本章介绍了行译码电路，列复用电路，输入缓冲电路，输出缓冲电路，升压电路等外围电路的设计。

第五章 版图设计及仿真数据分析

5.1 版图设计

版图是集成电路设计的物理实现，是整个设计中非常关键的一步，版图实现的好坏很大程度上影响着整个系统的性能^[41]。版图中应该有良好的电源密度，使芯片内部电压得到保证，使其不论是静态情况下或者动态情况下，都不因为电源波动而使功能出现任何问题。另外，在版图设计中应尽量避免长的栅极走线。

图 5.1 为整个 32M SRAM 测试片版图，32 颗 1M 的存储单元按照 8 行 4 列布局排列；最外围为 IO 端口；其他则为测试片的外围电路。

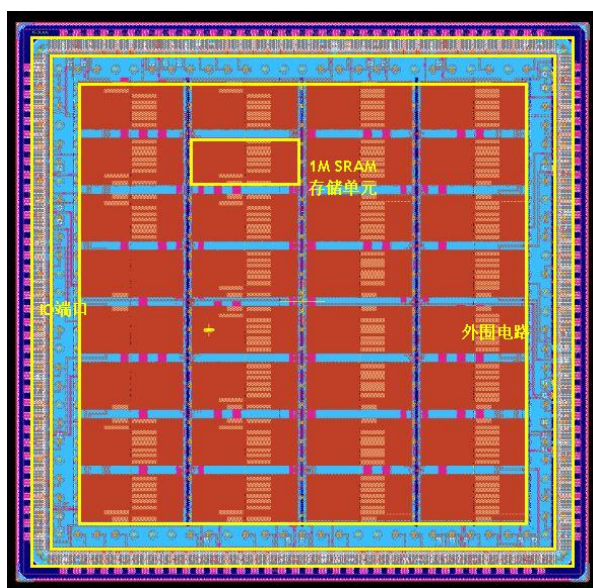


图 5.1 32M SRAM 测试片版图

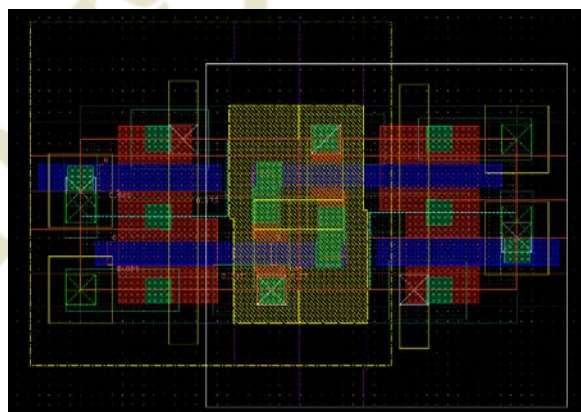


图 5.2 存储单元版图

图 5.2 为 SRAM 存储单元版图, 采用标准阈值电压的 MOS 管, 单元尺寸为 $1.04\mu\text{m}\times 0.36\mu\text{m}$, 面积为 $0.374\mu\text{m}^2$;

图 5.3 灵敏放大器版图, 由于灵敏放大器的输入信号是微小电压差, 受失配影响严重, 因此在进行版图设计时, 要严格按照两边对称原则来设计。

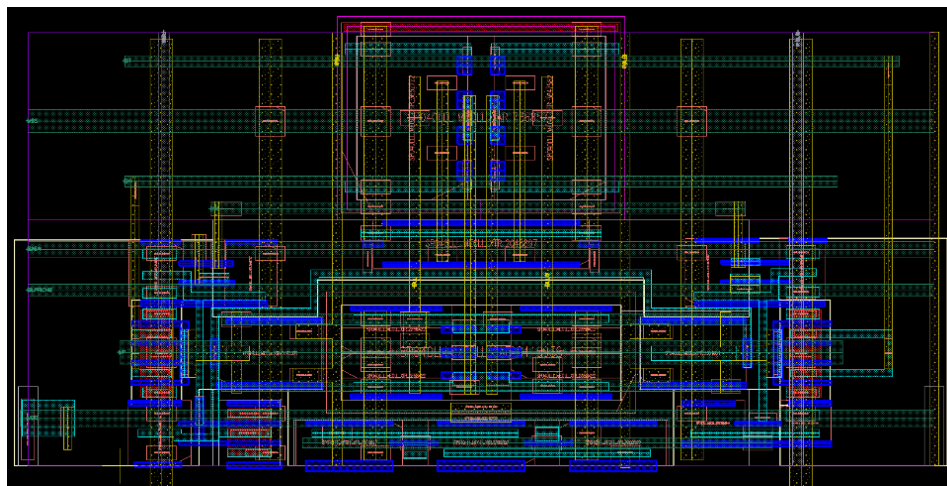


图 5.3 灵敏放大器版图

5.2 系统仿真与分析

设计验证是 VLSI 设计流程中重要的环节。电路的仿真可以帮助设计者确定产品中存在的问题, 以及是否达到设计要求。本文在对 SRAM 的仿真中, 我们采用 HSPICE 作为验证工具。HSPICE 具有收敛性好、模型参数精确等优点, 并且可以进行蒙特卡罗分析。

为了保证存储块设计的强壮性, 保证芯片的良率在可预期的工艺变化都能得到保证。我们在不同工艺拐点(PSNF, PFNS, PSNS, PFNF, PTNT), 温度(-40°C , 25°C , 125°C)及电源电压($0.9\text{V}\pm 10\%$)下进行了基础功能验证以良率分析。还通过对不同工艺拐点下, 基础功能仿真和读写模式下的最小工作电压仿真, 来验证存储块设计的强壮性。除此之外, 还对所设计 SRAM 测试片的各组电压进行了漏电流的仿真。

由于所设计 SRAM 测试片输入信号较多, 若要全面考虑到各种情况的输入激励组合很庞大, 在此仅选择几种比较典型的情况加以分析。为了方便接下来的表述, 假设 VDD 为电压 VDDC、VDDP1、VDDP2、VDDP3、VDDP4 构成的集合, 即 $\text{VDD}=\{\text{VDDC}, \text{VDDP1}, \text{VDDP2}, \text{VDDP3}, \text{VDDP4}\}$ 。

5.2.1 功能仿真

为了验证所设计 SRAM 测试片在不同工艺拐点条件下，能够正常地进行读写操作，仿真了 3 种工艺角的情况，即 25°C，1.0*VDD，TT 工艺角；125°C，0.9*VDD，SS 工艺角；-40°C，1.1*VDD，FF 工艺角。

在 25°C，1.0*VDD，TT 情况下的仿真结果如图 5.4 所示。CK 是时钟信号，OE 与 WEB 分别表示输出使能信号（高电平有效）和读写使能信号（高电平：读操作；低电平：写操作）。当 WEB=0，OE=0 时，SRAM 测试片进行写操作，依次往地址为 800e、800f、800d 中写入数据 5555_5555、aaaa_aaaa、ffff_ffff。当 WEB=1，OE=1 时，SRAM 测试片进行读操作，将地址为 800e、800f、800d 依次读出。数据读出与 CK 的延时为 2.512ns。

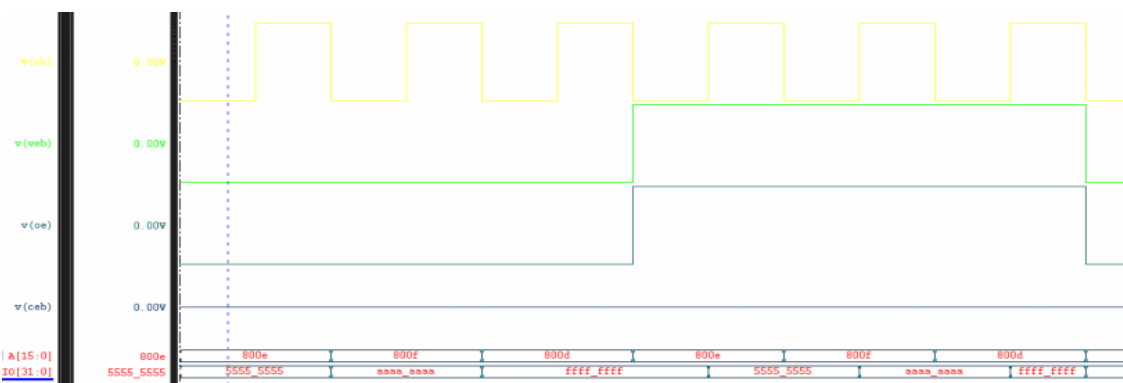


图 5.4 TT 下 SRAM 测试片读写时序

在 125°C，0.9*VDD，SS 情况下的仿真结果如图 5.5 所示。读写操作内容与 TT 情况下一致。数据读出与 CK 的延时为 4.947ns。

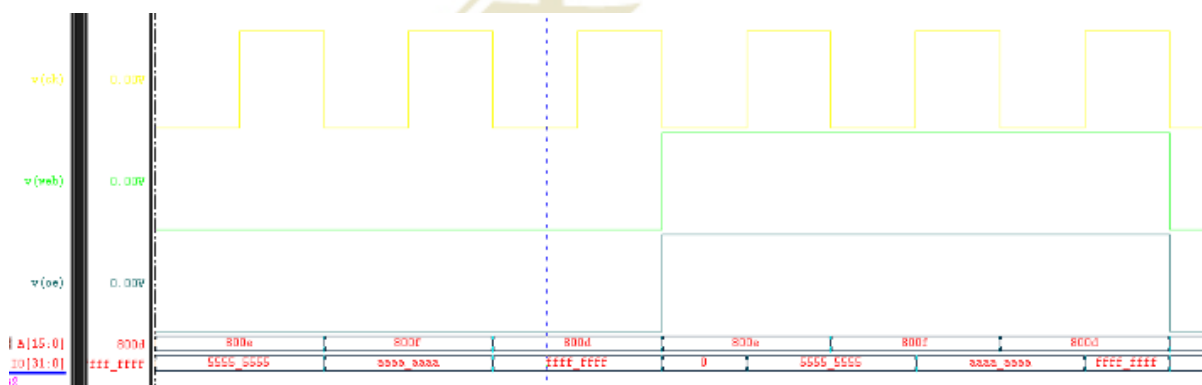


图 5.5 SS 下 SRAM 测试片读写时序

在 -40°C ， $1.1 \times \text{VDD}$ ，FF 情况下的仿真结果如图 5.6 所示。读写操作内容与 TT 情况下一致。数据读出与 CK 的延时为 2.045ns 。

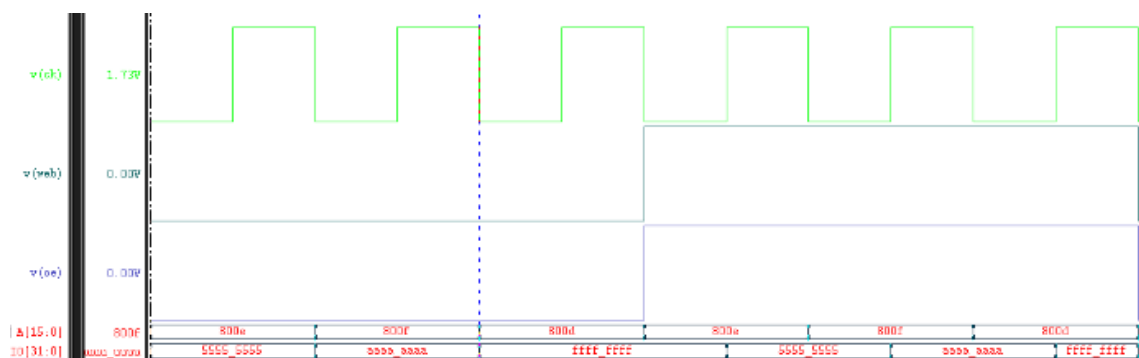


图 5.6 FF 下 SRAM 测试片读写时序

5.2.2 读写模式最小工作电压仿真

为了保证较低的工作电压下存储块的稳定性，本文设计要求，在存储块电压为 0.6V 条件下，存储阵块仍然能够正常进行读写操作。因此，本小节对读写模式下的测试片进行了最小电压仿真。如图 5.7 所示，为读写模式最小电压仿真示意图，值得注意的是，为了排除 IO 端口电压的变化对存储块读写功能的影响，VDDP4 和 VDDP3 分别保持 1.8V 和 0.9V 不变，VDDC、VDDP1 和 VDDP2 应连接一起，从 0.9V 递减，直到出现读写操作失效为止，此时存储块的工作电压为读写模式下的最小工作电压。

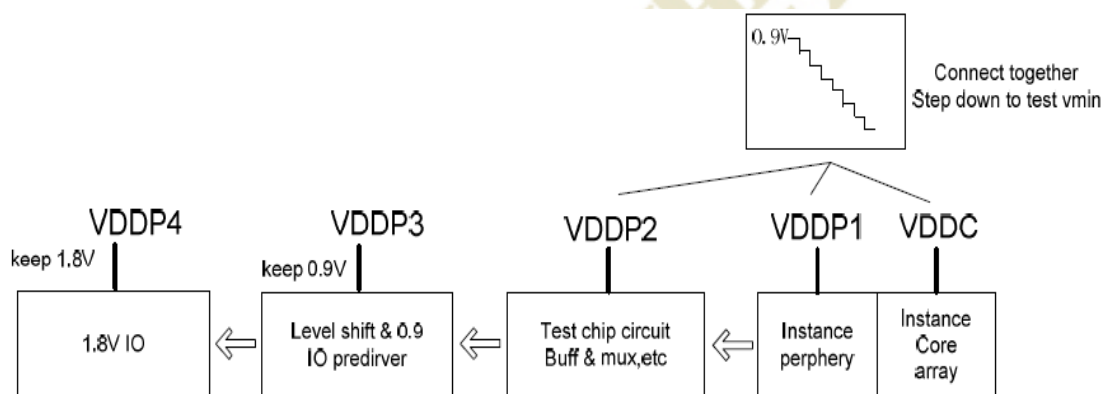


图 5.7 读和写模式最小工作电压仿真示意图

为了更全面的了解 SRAM 测试片读写模式下的最小工作电压，本文针对 3 个工艺角进行了仿真，即 25°C ，TT 工艺； 125°C ，SS 工艺； -40°C ，FF 工艺。如表 5.1 所示为读写模式下最小工作电压的仿真结果。第二行所列为不同工艺角对应的读写模式下的最小工作电压，全都小于 600mv ，因此满足设计要求。第三行中 T_{cq} 为数据输

出与时钟 CK 上升沿的延时。

表 5.1 读写模式下最小工作电压仿真结果

Corner	TT_25	SS_125	FF_-40
Vmin	572mv	589mv	594mv
Tcq	6.071ns	7.057ns	3.630ns

5.2.3 漏电流

表 5.2 所示为不同工艺、电压和温度(PVT)下, 测试片 2M 存储单元每组电压对应的漏电流大小。表 5.3 为 TT 0.9V 25℃条件下, 1 个芯片上 2M 存储单元部分电压对应的漏电流仿真数据与测试数据比较的结果, 电压 VDDPC~VDDP4 所对应的漏电流满足要求。

表 5.2 不同 PVT 下测试片的漏电流

	电源	FF 0.99V 125℃	FF 0.99V -40℃	TT 0.9V 25℃	SS 0.81V -40℃
Standby (mA)	VDDP1	687.7	42.1	27.6	51.3
	VDDP2	67.5	2.52	1.67	2.32
	VDDP3	0.938	0.46	0.275	0.196
	VDDP4	0.573	0.047	0.014	0.035
	VDDC	12.2	1.16	0.68	0.793
	VSS1	766.5	45.2	30.3	55.2
	VSS2	0.383	0.048	0.015	0.012

表5.3 TT 0.9V 25℃条件下, 部分电压对应的漏电流测试结果

Test Item	Simulation Result For One Instance	Silicon data	Result
ISB VDDC	0.68mA	1.3198mA	Normal
ISB VDDP1	27.6mA	19.5272mA	Normal
ISB VDDP2	1.67mA	6.9370mA	Normal
ISB VDDP3	0.275mA	1.2452mA	Normal
ISB VDDP4	0.014mA	0.035mA	Normal

5.2.4 存储阵列良率分析

考虑到工艺参数变化的影响, 利用第三章所实现的混合重要性采样算法, 分别对不同电压 (0.9*VDDC, 1.0*VDDC, 1.1*VDDC)、不同温度 (-40℃, 25℃, 125℃)

情况下，32M 的 SRAM 存储阵列的良率进行分析。

表 5.4 存储阵列良率分析结果

电压(V) 温度(°C)	0.81	0.9	0.99
-40	99.6%	100%	100%
25	99.8%	100%	100%
125	99.9%	100%	100%

仿真结果如表 5.4 所示，在电压为 0.9V 和 0.99V 的情况下，在不同温度的影响下，32M 的 SRAM 存储阵列的良率可以稳定在 100%。在电压较低的 0.81V 情况下，32M 的 SRAM 存储阵列良率可以达到 99.5% 以上。因此，在工艺参数变化的影响下，所设计 SRAM 测试片存储阵列在不同温度和电压下可以保持高的良率。

5.3 本章小结

本章在 40G CMOS 工艺下，对所设计的 SRAM 测试片进行了一系列模拟仿真。通过基础功能仿真波形可以看出，在不同 PVT 条件下，SRAM 读写操作稳定。通过读写模式下最小工作电压的仿真，对于比较典型的 3 个工艺角，即 TT 25°C、SS 125°C、FF -40°C，SRAM 可正常读写的最小工作电压均小于 0.6V，满足设计要求。在漏电流检查中，对 2M 单元中的各组电压下的漏电流进行仿真，并将 TT 0.9V 25°C 条件下的仿真结果与测试结果比较，漏电流大小满足要求。最后，本章利用所实现快速蒙特卡罗方法对 SRAM 存储阵列进行良率分析，结果表明，所设计 SRAM 测试片在工艺参数变化，以及温度电压波动的影响下，良率能够保持在很高的值。

第六章 总结与展望

6.1 总结

SRAM 具有速度快、功耗低等优点，是高性能微处理器中非常关键的一部分^[42]。随着 CMOS 工艺的不断进步，SRAM 的集成密度越来越大，性能也不断地得到提高。然而，对于 CMOS 工艺进入深亚微米的今天，工艺尺寸的减小，也给 SRAM 设计带来了一些挑战。由于工艺参数的变化，尤其是局部工艺参数的随机变化，导致 MOS 管阈值电压的随机波动，造成同一个 die 上相邻晶体管的失配，已经开始成为限制 SRAM 良率提高的一个瓶颈。

本文首先分析了 SRAM 工艺参数变化的来源，以及工艺参数变化对于 SRAM 的稳定性的影响。然后，针对 HSPICE 蒙特卡罗分析方法对于高密度、低失效率的 SRAM 分析中存在的收敛慢，采样点过多，仿真时间过长等致命问题，基于文献[36]提出的 Mixture Importance Sampling 算法，实现了一种 Fast Monte Carlo 方法。对于低失效率 SRAM，仿真速度可以提高数个数量级。另外，基于 40G 工艺，设计了一颗 32M 的 SRAM 测试芯片，主要用来测试工艺的稳定性。在设计过程中，利用所实现的 Fast Monte Carlo 算法分析工艺参数变化给 SRAM 存储单元良率带来的影响，在面积均为 $0.374 \mu\text{m}^2$ 前提下，选择良率最高的存储单元应用到测试片中。所设计的 32M SRAM 测试芯片由 16 个独立可测的 2M 存储单元组成，实现了电源分开管理，具有 32 个双向并行 IO 端口，并且，提供了丰富的测试模式和测试端口。

最后，对所设计 SRAM 测试片在不同电压和温度下进行了基础功能验证，以及读写模式下最小工作电压仿真，仿真结果符合设计要求。另外，对各组电压进行漏电流检查，并将部分电压对应的漏电流与测试数据比较。考虑到工艺参数变化，还利用了所实现的快速蒙特卡罗算法对存储阵列进行良率分析，结果表明在温度和电压在一定范围内波动的情况下，存储阵列可以保持很高的良率。因此，在快速蒙特卡罗算法的指导下，所设计的 SRAM 测试片在不同工艺角下可以保持比较高的稳定性，并且在工艺参数变化的影响下，存储阵列有很高的良率。

6.2 展望

本文实现 Mixture Importance Sampling 算法,并将该算法应用到 SRAM 设计过程中。但是比较遗憾的是,由于时间原因,本文所设计测试芯片的仿真数据未能与测试数据进行比较。希望等测试片后期的测试数据出来之后,结合测试数据再对快速 Monte Carlo 算法以及测试芯片的设计进行进一步的优化。此外,在应用所实现的算法来设计 SRAM 存储单元过程中,可以根据文献[25]所提供方法,设计出高性能或者低功耗的 SRAM 存储单元,但由于时间原因,本文并没有深入研究,希望在以后的工作中可以弥补这个缺憾。

一般说来,SRAM 的面积越大,受工艺参数变化的影响就越小,但是造成的面积开销就会越大。在未来的工作中,SRAM 设计必须同时考虑到稳定性、成本与性能的折中。

参考文献

- [1] G. Atwood, S. DeBoer, K. Prall, et al. A Semiconductor Memory Development and Manufacturing Perspective. European Solid State Circuit Conference (ESSCIRC), ESSCIRC 2014-40th, Sept.2014, pp. 1-6
- [2] D. J. Rennie, T. Shakir, M. Sachdev. Design Challenges in Nanometeric Embedded Memories. Signals, Circuits and Systems (SCS), 2009 3rd International Conference on, Nov. 2009, pp. 1-8
- [3] E. J.Marinissen, B. Prince, D. Keltel-Schulz, et al. Challenges in Embedded memory design and test. Design, Automation and Test in Europe, 2005. Proceeding, 2005:722-727
- [4] Neil H.E. Weste 等著, 汪东等译, “CMOS 超大规模集成电路设计”, 中国电力出版社, 2006.
- [5] 张一平, 深亚微米灵敏放大器设计, 苏州大学硕士学位论文, 2008 年 5 月。
- [6] J. F. Zhu, N. Bai, J. H. Wua. A low active and leakage power SRAM using a read and write divided and BIST programmable timing control circuit. Microelectronics Journal , Vol. 44, 2013, PP.283–291
- [7] L. A. Wen, Z. T. Li, Y. Li. Single-ended, robust 8T SRAM cell for low-voltage operation. Microelectronics Journal, Vol. 44, 2013, pp.718–728
- [8] Y. Morita, H. Fujiwara, H. Noguchi, et al. An Area-Conscious Low-Voltage-Oriented 8T-SRAM Design under DVS Environment. VLSI Circuit, 2007 IEEE Symposium on.IEEE, 2007:256-257
- [9] L. Chang, R. K. Montoye, Y. Nakamura. An 8T-SRAM for Variability Tolerance and Low-Voltage Operation in High-Performance Caches. IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 43, NO. 4, APRIL 2008
- [10] 张强, 65nm SRAM 的设计, 浙江大学硕士学位论文, 2010 年 1 月。
- [11] Th. Fischer, T. Nirschl, B. Lemaitre, and D. Schmitt-Landisiedl. “Modeling of the parametric yield in decananometer SRAM-Arrays”, Advances in Radio Science, 4, 281-285, 2006
- [12] H. Yamauchi. A Discussion on SRAM Circuit Design Trend on Deeper

- Nanometer-Scale Technologies. IEEE TRANSACTIONS ON VLSI SYSTEMS, VOL. 18, NO. 5, MAY 2010
- [13] B. Cheng, S. Roy, and A. Asenov, "Impact of intrinsic parameter fluctuations on SRAM cell design," in Proc. Int. Solid-State Integr. Circuit Technol. Conf., 2006, pp. 1290–1292.
- [14] 郑丹丹, 嵌入式CPU的纳米尺度SRAM设计研究, 浙江大学博士学位毕业论文, 2009年6月
- [15] K. Bernstein, D.J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Person, N. J. Rohrer, "High-Performance CMOS variability in the 65-nm regime and beyond", IBM J. RES. & DEV., 2006, VOL. 50 NO 4/5.
- [16] V. Gupta, M. Anis, Statistical Design of the 6T SRAM Bit Cell. IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, VOL. 57, NO. 1, JANUARY 2010
- [17] J. Hartmann. Towards a New Nanoelectronic Cosmology. IEEE International Solid-State Circuit Conference, 2007, 01. 3
- [18] Jan M. Rabaey 等著, 周润德等译, "数字集成电路—电路、系统与amp;设计", 电子工业出版社, 2013.
- [19] 张金峰, 亚 65 纳米 SRAM 的稳定性研究与设计, 苏州大学硕士学位论文, 2008 年 5 月。
- [20] E. Seevinck, F. J. List, and Lohstroh. Static-noise margin analysis of MOS SRAM cells, IEEE journal of Solid-State Circuits, vol.22, OCT. 1987, pp.748-754.
- [21] W. Wang, Z. Yu, and K. Choi. High SNM 6T CNFET SRAM Cell Design Considering Nanotube Diameter and Transistor Ratio. 2011 IEEE international Conference on EIT, Mankato, MN, 15-17 May 2011, pp. 1-4
- [22] 肖海林, 高可靠静态存储器研究, 哈尔滨工业大学硕士学位论文, 2007 年 7 月。
- [23] T. Data, S. Hagiwara, K. Masu, and T. Sato. Robust Importance Sampling for Efficient SRAM Yield Analysis, 2010 11th International Symposium on Quality Electronic Design, San Jose, CA, 2010, pp. 15-21
- [24] T.S. Doorn, E.J.W. ter Maten, and J.A. Croon. Importance Sampling Monte Carlo simulations for accurate estimation of SRAM yield. In proc. 34th ESSCIRC, Sep. 2008, pp. 230-233

- [25] S. Mukhopadhyay, H. Mahmoodi, and K. Roy. Statistical design and optimization of SRAM cell for yield enhancement. Computer Aided Design, 2004. ICCAD-2004.IEEE/ACM International Conference on, 2004:10-13.
- [26] K. Agarwal and S. Nassif. Statistical analysis of SRAM cell stability. Design Automation Conference, 2006 43rd ACM/IEEE, 2006: 57-62.
- [27] J. Wang, A. Singhee, R. A. Rutenbar, et al. Statistical modeling for minimum standby supply voltage of a full SRAM array. Solid State Circuit Conference, 2007. ESSCIRC 2007. 33rd European, 2007:400-403.
- [28] 郭瑛, 高速铁路移动通信系统仿真中的重要性采样技术研究, 北京交通大学硕士学位论文, 2012 年。
- [29] H. Makino, S. Nakata, H. Suzuki, et al. Reexamination of SRAM Cell Write Margin Definitions in View of Predicting the Distribution. IEEE Transactions on Circuits and Systems II: Express Briefs, April 2011, vol.58, no.4, pp:230-234
- [30] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel. Yield and Speed Optimization of a Latch-Type Voltage Sense Amplifier. IEEE Journal of solid-state circuits, July 2004, vol. 39, no. 7, pp:1148-1158
- [31] W. Jiajing, A. Singhee, R.A. Rutenbar, et al. Two Fast Methods for Estimating the Minimum Standby Supply Voltage for Large SRAMs. IEEE Transactions on Computer-Aided Design of Integrated Circuit and Systems, Dec. 2010, vol.29,no.12, pp:1908-1920
- [32] B. Y. Farah, M. Mohammad. Determining the minimum energy operating point for embedded SRAM memory. 2011 IEEE Workshop on Signal Processing Systems (SiPS), Oct. 2011, pp:112-116
- [33] C. Wu, L. Zhang, Z. Lu et al. Leakage reduction of sub-55nm SRAM based on a feedback monitor scheme for standby voltage scaling. 2010 International SoC DesignConference (ISOCC), Nov. 2010, pp: 315-318
- [34] S. Uznanski, G. Gasiot, P. Roche, et al. Heavy Ion Characterization and Monte Carlo Simulation on 32 nm CMOS Bulk Technology. IEEE Transactions on Nuclear Science, Dec. 2011, vol.58, no.6, pp: 2652-2657
- [35] H. Mostafa, M. Anis, M. Elmasry. A Bias-Dependent Model for the Impact of Process Variations on the SRAM Soft Error Immunity. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, Nov. 2011, vol.19, no.11, pp:2130-2134

- [36] R. Kanj, R. Joshi, and S. Nassif. Mixture Importance sampling and its application to the analysis of SRAM designs in the presence of SRAM designs in the presence of rare failure events. Design Automation Conference, 2006 43rd ACM/IEEE, 2006:69-72.
- [37] 许磊, 张兰. 基于 Sobol 序列的量子粒子群求解电力经济调度. 西南师范大学学报, 2013.4
- [38] K. Agarwal, S. Nassif. The Impact of Random Device Variation on SRAM Cell Stability in Sub-90-nm CMOS Technologies. IEEE Transactions on VLSI Systems, Jan. 2008, pp. 86-97
- [39] 方海涛, 高速低功耗嵌入式 SRAM 的设计, 华中科技大学硕士学位论文, 2012 年 5 月。
- [40] Oghenekarho Okobiah. EXPLORING PROCESS-VARIATION TOLERANT DESIGN OF NANOSCALE SENSE AMPLIFIER CIRCUITS. Thesis Prepared for the Degree of MASTER OF SCIENCE, UNIVERSITY OF NORTH TEXAS, DEC. 2010
- [41] 吴浩, Flash 存储器中的电荷泵系统研究和设计, 苏州大学硕士学位论文, 2007 年 5 月。
- [42] 吴晨, 基于数据保持电压的低功耗 SRAM 设计, 苏州大学硕士学位论文, 2011 年 5 月。

攻读硕士期间发表的论文

[1] 李二亮, 张立军, 李友忠, 张其笑, 姜伟, 胡玉青. 基于补偿电路的 SRAM 读操作跟踪电路设计. 电子设计工程, 2015 年 5 月第 9 期.

[2] Li Erliang, Zhang Lijun, Li Youzhong, Zhang Qixiao, Jiang Wei, Hu Yuqing. A Novel Read Tracking Circuit Design Based on Compensation Circuit. Solid-State and Integrated Circuit Technology (ICSICT), 2014 12th IEEE International Conference on, Guilin, Oct. 2014, pp.1-3

ETD 论文管理系统

致 谢

时间如白驹过隙，三年的研究生生涯一晃而过。在本论文即将完成之际，感谢我的导师张立军老师。本文是在张老师的悉心指导下完成的，从选题到完成，无不渗透着张老师的热情关怀和精心指导。在攻读硕士学位这三年里，张老师敏锐的洞察力、渊博的学识、严谨的治学之道使我受益匪浅。

感谢存储器资深设计师郑坚斌先生在设计过程中，给予我耐心的指导和宝贵的意见，同时他强烈的事业心和忘我的工作精神都深深的鼓励着我，是我步入工作岗位的学习榜样。

感谢教育过我的老师，在我成长的道理上基于的培养和关怀。

感谢我的同学，朋友，以及实验室的所有师兄们对我学习和生活上的帮助，为我在这边的学习营造了良好的氛围。感谢师兄李有忠，师弟胡玉青为论文提出了宝贵的修改意见和进行了繁琐的文字检查工作。

感谢我的家人，尤其是我的父母，一如既往的支持我，照顾我，鼓励我，使我能够安心完成学业。

李二亮

2015 年 4 月于苏州大学