

SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY —
INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Thesis title

Author

SCHOOL OF COMPUTATION,
INFORMATION AND TECHNOLOGY —
INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics

Thesis title

Titel der Abschlussarbeit

Author:	Author
Examiner:	Supervisor
Supervisor:	Advisor
Submission Date:	Submission date

I confirm that this master's thesis is my own work and I have documented all sources and material used.

Munich, Submission date

Author

Abstract

Contents

Abstract	iii
1 Introduction	2
1.1 Section	2
2 Bayesian Deep Learning	3
2.1 General Concepts	3
2.2 Priors	4
2.3 Inference	4
2.4 Properties of Bayesian Neural Network Posteriors	4
3 Flow Models	5
4 Geometry of Neural Networks	6
4.1 Symmetries of Neural Networks	6
4.2 Canonical Representations of Neural Networks	6
5 Graph Neural Networks & MetaNets	7
6 Generative Models in Weight-Space	8
7 Method & Design Choices	9
8 Results	10
9 Discussion	11
10 Conclusion & Future Work	12
Abbreviations	13
List of Figures	14
List of Tables	15

Bibliography

16

Notation	Explanation
θ	Parameters of a neural network
$(x, y) \in \mathcal{D}$	Dataset with inputs x and labels/targets y

Table 1: Summary of notation used throughout the thesis.

1 Introduction

1.1 Section

2 Bayesian Deep Learning

A primary use case for a generative model over neural network weights is in Bayesian deep learning, where it can allow efficient inference by transporting the prior distribution to the posterior. Thus, to motivate the rest of the discussion, we first give an overview of concepts from Bayesian deep learning. Section 2.1 is a general introduction. It is followed by a review of inference methods (Laplace approximations, variational inference, MCMC-based methods) typically used for Bayesian neural networks, and we conclude with a review of literature around Bayesian deep learning particularly relevant for our work in Section 2.4.

2.1 General Concepts

In typical Bayesian fashion, Bayesian deep learning (refer to (MacKay, 1992; Neal, 1996) for foundational work and (Goan and Fookes, 2020; Arbel et al., 2023) for more recent reviews) aims to quantify the uncertainty in neural networks through probability distributions over their parameters, rather than obtaining a single solution by an SGD-like optimization method. Then, given the *posterior* distribution $p(\theta|\mathcal{D})$ over weights θ conditioned on the dataset \mathcal{D} , predictions are obtained via *Bayesian model averaging*:

$$p(y|x, \mathcal{D}) = \mathbb{E}_{\theta \sim p(\theta|\mathcal{D})} [p(y|x, \theta)] = \int p(y|x, \theta) p(\theta|\mathcal{D}) d\theta. \quad (2.1)$$

Therefore, Bayesian inference over neural network weights consists of three steps:

1. Specify prior $p(\theta)$.
2. Compute/sample posterior $p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta)$.
3. Average predictions over the posterior.

The last step only requires forward passes through the model and thus is straightforward. The first two steps require deeper consideration.

2.2 Priors

Specifying a prior mainly consists of two choices: specifying an architecture, and specifying a probability distribution over the weights. The distribution is typically taken to be an isotropic Gaussian, which is an uninformative prior but straightforward to work with.

Different architectural decisions also result in different functions, even if the flattened weight vectors are identical, meaning that the choice of an architecture further specifies a prior in function space. As a simple example, keeping the depth and width of a neural network constant, even just changing the activation function from a ReLU to a sigmoid results in a different distribution of functions. The functional distribution can also be specified in a more deliberate way; e.g. a translation-invariant convolutional neural network, or a group-equivariant network (Cohen and Welling, 2016) puts probability mass only on functions satisfying certain equivariance constraints depending on the task at hand.

We keep this discussion short since the choice of a prior has tangential impact in the rest of the presentation, and we refer to recent reviews such as (Fortuin, 2022) for a more detailed treatment of Bayesian neural network priors.

2.3 Inference

Laplace

VI

DE

MCMC introduction

HMC

Highlight distinction between chain-based and transformation based sampling, using normalizing flows and connections to optimal transport etc

– Stochastic –

SG-MCMC

SG-HMC

2.4 Properties of Bayesian Neural Network Posteriors

How they are like Boltzmann distributions, not arbitrary.

3 Flow Models

4 Geometry of Neural Networks

4.1 Symmetries of Neural Networks

Also introduce mode connectivity here, before canonical reps

Introduce permutation and scaling symmetries, and maybe other symmetries if we end up using them.

4.2 Canonical Representations of Neural Networks

Mainly based on the deep toroids paper

5 Graph Neural Networks & MetaNets

6 Generative Models in Weight-Space

7 Method & Design Choices

8 Results

9 Discussion

10 Conclusion & Future Work

Abbreviations

List of Figures

List of Tables

1	Summary of notation used throughout the thesis.	1
---	---	---

Bibliography

- Arbel, Julyan et al. (2023). *A Primer on Bayesian Neural Networks: Review and Debates*. DOI: [10.48550/arXiv.2309.16314](https://doi.org/10.48550/arXiv.2309.16314). arXiv: 2309.16314 [cs, math, stat].
- Cohen, Taco and Max Welling (2016). "Group Equivariant Convolutional Networks." In: *Proceedings of The 33rd International Conference on Machine Learning*. PMLR, pp. 2990–2999.
- Fortuin, Vincent (2022). "Priors in Bayesian Deep Learning: A Review." In: *International Statistical Review* 90.3, pp. 563–591. ISSN: 0306-7734, 1751-5823. DOI: [10.1111/insr.12502](https://doi.org/10.1111/insr.12502).
- Goan, Ethan and Clinton Fookes (2020). "Bayesian Neural Networks: An Introduction and Survey." In: vol. 2259, pp. 45–87. DOI: [10.1007/978-3-030-42553-1_3](https://doi.org/10.1007/978-3-030-42553-1_3). arXiv: 2006.12024 [cs, stat].
- MacKay, David J C (1992). "Bayesian Methods for Adaptive Models." In.
- Neal, Radford M. (1996). *Bayesian Learning for Neural Networks*.