Chapter 1.

Introduction

1.1 Motivation

Artificial neural networks represent a sea change in computing. They have successfully been used in a wide range of applications, from protein-folding in Tsaban et al. (2022), knot theory in Davies et al. (2022), and extracting data from gravitational waves in Zhao et al. (2023).

As neural networks become more ubiquitous, we see that the number of parameters required to train them increases, which poses two problems: accessibility on low-power devices and the amount of energy needed to train these models, see for instance Wu et al. (2022) and Strubell et al. (2019). Parameter estimates become increasingly crucial in an increasingly climate-challenged world. That we know strict and precise upper bounds on parameter estimates tells us when training becomes wasteful, in some sense, and when, perhaps, different approaches may be needed.

Our goal in this dissertation is threefold:

(i) Firstly, we will take something called Multi-Level Picard first introduced in E et al. (2019) and E et al. (2021), and in particular, the version of Multi-Level Picard that appears in Hutzenthaler et al. (2021). We show that dropping the drift term and substantially simplifying the process still results in convergence of the method and polynomial bounds for the number of computations required and rather nice properties for the approximations, such as integrability and measurability.

(ii) We will then go on to realize that the solution to a modified version of the heat equation has a solution represented as a stochastic differential equation by Feynman-Kac and further that a version of this can be realized by the modified multi-level Picard technique mentioned in Item (i), with certain simplifying assumptions since we dropped