Stochastic Interpolants in Hilbert Spaces

James Boran Yu

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Master of Philosophy in Machine Learning and Machine Intelligence

Pembroke College August 2025

To my loving family.

Declaration

I, James Boran Yu of Pembroke College, being a candidate for the MPhil in Machine Learning and Machine Intelligence, hereby declare that this report and the work described in it are my own work, unaided except as may be specified below, and that the report does not contain material that has already been used to any substantial extent for a comparable purpose.

TODO: Signed, Date

TODO: Software declaration

TODO: Word count

James Boran Yu
August 2025

Acknowledgements

TODO Acknowledgements

Abstract

TODO ABSTRACT!

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Motivation and Overview

TODO!

1.2 Contributions

This thesis develops a novel framework for generative modelling on function spaces. Our primary contributions are as follows.

1. We formulate stochastic interpolants directly in infinite-dimensional settings, which forms the core of our proposed framework.

2. We provide a rigourous theoretical analysis, establishing sufficient conditions under which the framework is well-posed and satisfies critical theoretical guarantees.

3. We translate these theoretical insights into practical design principles to improve the algorithm's performance.

4. We demonstrate our framework's effectiveness for solving partial differential equation (PDE)-based forward and inverse problems, achieving results competitive with state-of-the-art approaches but with reduced inference time.

5. Finally, we outline areas for further research, such as extending our theoretical guarantees under more relaxed assumptions and developing novel practical designs.

1.3 Outline

This thesis is structured as follows.

Chapter 1 provides the motivation and overview for this thesis.

Chapter 2 presents the necessary groundwork for this thesis: we provide an overview of stochastic interpolants in their original finite-dimensional setting, as proposed by Albergo et al. (2023a), and contrast this with diffusion models for generative modelling (Song et al., 2021). We then give an overview of the key mathematical concepts necessary to generalise stochastic interpolants to infinite-dimensional spaces, and provide a review of related works in generative modelling on function spaces.

Chapter 3 introduces our core framework: a formulation of stochastic interpolants directly in infinite dimensions. We present a Hilbert space-valued SDE and justify its suitability for generative modelling and prove sufficient conditions for the well-posedness of such an SDE. We provide a training objective and relate this to an error bound of the learned generative process. From this theoretical analysis, we describe how our framework is useful for solving both forward and inverse problems and identify key design principles informing the implemention of our method.

?? details an application of our framework for solving PDE-based forward and inverse problems. We describe the datasets and methods used, and compare our results with current state-of-the-art stochastic and deterministic solvers.

?? describes the merits of our work, as well as some limitations and potential areas for further work.

TODO: mention optimal transport in future work

TODO: make sure you frame the entire paper from the pov of bayesian inverse problems

TODO: add detail!

Chapter 2

Background and Preliminaries

In this chapter, we establish the conceptual and mathematical preliminaries to lay the necessary groundwork to formally generalise stochastic interpolants to function spaces. To achieve this, we structure our discussion as follows.

1. We begin by presenting diffusion models (DMs; Song et al., 2021) in finite dimensions.

2. Then, we describe key advantages of the stochastic interpolants framework over DMs, and present a form of stochastic interpolants in their original finite-dimensional context, as proposed by Albergo et al. (2023a).

3. We define Hilbert spaces as the underlying setting for our analysis, and present an overview of the key mathematical concepts necessary to describe random variables and stochastic differential equations (SDEs) in Hilbert spaces. Given these concepts, we outline key challenges in extending stochastic interpolants to infinite dimensions in Hilbert spaces.

4. Finally, we provide a review of related works which generalise DDPM and SBDM to function spaces, highlighting the relationship of these methods with their finite-dimensional counterparts.

2.1 Diffusion Models in Finite Dimensions

Diffusion models (DMs; Song et al., 2021) are a family of generative models achieving remarkable empirical success across a broad range of domains. To generate data x distributed according to a target measure μ_{target} on N-dimensional Euclidean space $\mathbb{R}^N$, we define two stochastic processes on a finite time interval $[0,T]$. For a drift coefficient $f : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$

and diffusion coefficient $g : [0,T] \to \mathbb{R}_{>0}$, the *diffusion process* $\mathbb{X} = \{X_t\}_{t\in[0,T]}$, is the solution to the following *forward SDE*:

$$\mathrm{d}X_t = f(t,X_t)\,\mathrm{d}t + g(t)\,\mathrm{d}W_t\,, \quad X_0 \sim \mu_{\text{target}},$$

where $\mathbb{W} := \{W_t\}_{t\in[0,T]}$ is a standard dimensional Wiener process.

Let μ_t be the law (marginal distribution) of X_t and let $p_t : \mathbb{R}^N \to \mathbb{R}$ be the density of μ_t with respect to the Lebesgue measure. Under some mild regularity conditions (Anderson, 1982) we may define a *time-reversed process* $\overline{\mathbb{X}} = \left\{\overline{X}_t\right\}_{t\in[0,T]}$, which when solved backwards in time from $\overline{X}_T \sim \mu_T$ yields a sample $\overline{X}_0 \sim \mu_{\text{target}}$:

$$\mathrm{d}\overline{X}_t = \left(f(t,\overline{X}_t) - g^2(t)\nabla \log p_t(\overline{X}_t)\right)\mathrm{d}t + g(t)\,\mathrm{d}\overline{W}_t\,, \quad \overline{X}_T \sim \mu_T, \tag{2.1}$$

where $\overline{\mathbb{W}} := \left\{\overline{W}_t\right\}_{t\in[0,T]}$ is a standard Wiener process when time flows backwards from $t = T$ to 0, and $\nabla \log p_t(x)$ is the *score* of the marginal distribution at time t, namely, the spatial derivative of the log-density of X_t.

By learning a time-dependent score network $s_\theta(t,x)$ and plugging this in place of $\nabla \log p_t(x)$ in Equation (2.1), we may generate approximate samples from μ_{target}, provided we have samples from μ_T.

To ensure that μ_T is a simple and tractable distribution, f and g are typically chosen such that the forward process systematically transforms data $X_0 \sim \mu_{\text{target}}$ into a Gaussian $\mathrm{N}(0,\sigma_T^2 I_N)$. However, this transformation is only guaranteed to be perfect asymptotically as $T \to \infty$. In a practical implementation, we must terminate time at a finite time step T. This introduces a bias during sampling, since the final condition for the time-reversed SDE is not a Gaussian at time T.

For example, *score-based diffusion models* (SBDMs) are a special case of DMs in which the forward SDE is an Ornstein-Uhlenbeck process. In this case, the law of X_t converges to a standard Gaussian $\mathrm{N}(0,I_N)$ in the limit $t \to \infty$.

$$\mathrm{d}X_t = -X_t d_t + \sqrt{2}\,\mathrm{d}W_t\,, \quad X_0 \sim \mu_{\text{target}}.$$

While a larger T bridges the data closer to a Gaussian, a smaller T helps improve the learned approximation $s_\theta(t,x)$ of the score and leads to more tractable sampling when solving the reverse process. Hence, a tradeoff must be found when choosing T (see, for example, Franzese et al., 2023).

2.2 Stochastic Interpolants in Finite Dimensions

Stochastic interpolants (SIs) are a class of generative models which provide the following improvements in flexibility over DMs:

1. SIs can bridge between any two arbitrary distributions determined *a priori*, as opposed to between a single target distribution and a fixed noise prior. Moreover, the source and target distributions can be coupled, allowing SIs to model a joint probability law between source and target data. This provides a powerful and flexible framework, where a single trained model can perform unconditional generation in addition to solving both forward and inverse tasks within a Bayesian setting.

2. The interpolation is constructed on a finite time horizon, in contrast to DMs which rely on an asymptotic convergence to the simple noise prior. By design, this has two advantages: it removes approximation bias from the terminal distribution and eliminates the need to tune the time horizon as a hyperparameter.

3. The interpolation path is an explicit design choice, allowing us to construct simple bridges (e.g., linear trajectories) between the two distributions. This contrasts with DMs, where the trajectory is an emergent property determined by the specific SDE. Simple, low-curvature paths are easier for numerical solvers to approximate accurately, which can lead to greater sampling efficiency with fewer function evaluations.

Each of these merits is demonstrated in a function generation setting in **??**: we show that our framework is highly effective for solving PDE-based forward and inverse problems. Notably, this is achieved on a strict finite time interval, and with fewer function evaluations and reduced inference time.

Having stated the key merits of SIs over DMs, we now introduce SIs in their finite-dimensional setting, as proposed by Albergo et al. (2023a,b). To establish the necessary context for our subsequent development in infinite dimensions, the following discussion captures the conceptual essence of SIs in finite dimensions. A formal and detailed presentation of the specific regularity conditions in our infinite-dimensional setting will be provided in Chapter 3.

Let μ be a joint measure on $\mathbb{R}^N \times \mathbb{R}^N$ with marginals μ_0 and μ_1. We draw a (possibly coupled) pair of random variables $\xi = (\xi_0, \xi_1) \sim \mu$, where we refer to μ_0 as the *source* and μ_1 as the *target distribution*. Furthermore, we denote by $\mu_{1|0}(\xi_0)$ the law of ξ_1, conditional on ξ_0 and similarly for $\mu_{0|1}(\xi_1)$. We will refer to these, respectively, as the *conditional target* and *conditional source* distributions. The case where the source and target distributions are uncoupled, $\xi_0 \perp \xi_1$, is a special case of this general setting.

Let z be a standard N-dimensional Gaussian, distributed independently of ξ. A *stochastic interpolant* is a family of random variables $\{x_t\}_{t\in[0,1]}$ indexed by time $t \in [0,1]$:

$$x_t = I(t,\xi) + \gamma(t)z, \quad t \in [0,1],$$

where

1. $I : [0,1] \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}^N$ is an *interpolating function* satisfying $I(0,\xi) = \xi_0$ and $I(1,\xi) = \xi_1$ for all $\xi \in \mathbb{R}^N \times \mathbb{R}^N$
2. $\gamma : [0,1] \to \mathbb{R}_{\geq 0}$ satisfies $\gamma(0) = \gamma(1) = 0$ and $\gamma(t) > 0$ for all $t \in (0,1)$
3. Both I and γ are continuously differentiable with respect to time t. We denote their time derivatives respectively by $\dot{I}$ and $\dot{\gamma}$.

Intuitively, the boundary conditions on I and γ ensure that the law of the stochastic interpolant matches the source and target distributions at the endpoints, $x_0 \sim \mu_0$ and $x_1 \sim \mu_1$. For intermmediate times $t \in (0,1)$, the law of x_t is equal to that of a deterministic path between ξ_0 and ξ_1, corrupted by scaled Gaussian noise.

To bridge from μ_0 to μ_1, we choose a positive constant $\varepsilon > 0$ and define a *forward SDE* as follows:

$$\mathrm{d}X_t = \left(\mathbb{E}\left[\dot{I}(t,\xi) + \dot{\gamma}(t)z \mid x_t = X_t\right] + \varepsilon\nabla\log p_t(X_t)\right)\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t\,, \quad X_0 = \xi_0, t \in [0,1]$$

where p_t is the density of the law of the interpolant x_t at time t, with respect to the Lebesgue measure. Under suitable regularity conditions, it can be shown that conditional on $X_0 = \xi_0$, the law of X_t at any time $t \in [0,1]$ is equal to the law of x_t conditional on ξ_0. Hence, by solving the forward SDE, we generate a sample from the conditional target distribution $\mu_{1|0}(\xi_0)$.

Similarly, we define a *time-reversed SDE* which, when solved backwards in time starting from $\overline{X}_1 = \xi_1$, gives a sample from the conditional source distribution $\mu_{0|1}(\xi_1)$:

$$\mathrm{d}\overline{X}_t = \left(\mathbb{E}\left[\dot{I}(t,\xi) + \dot{\gamma}(t)z \mid x_t = X_t\right] - \varepsilon\nabla\log p_t(X_t)\right)\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}\overline{W}_t\,, \quad \overline{X}_1 = \xi_1, t \in [0,1].$$

In the special case where $\varepsilon = 0$, the forward and time-reversed SDEs collapse to a *probability flow ODE*, where the source of stochasticity only comes from the initial/final conditions, in contrast to $\varepsilon > 0$ where additional noise is injected by the Wiener process.

2.3 Stochastic Processes in Hilbert Spaces

A *Hilbert space* H is a vector space equipped with a scalar-valued inner product $\langle f, g\rangle_H$, which is *complete* with respect to the norm $\|f\|_H := \sqrt{\langle f, f\rangle_H}$ induced by this inner product, that is, every H-valued Cauchy sequence converges in H-norm to an element in H. The choice of a Hilbert space, as opposed to a more general Banach space, is justified by the fact that the inner product provides essential geometric structure, giving rise to the concept of orthogonality.

Throughout, we let H be an infinite dimensional Hilbert space satisfying the following two properties:

1. H is *real*, meaning that all scalars, including inner products, are real-valued.
2. H is *separable*, which has the implication that there exists a *countable* orthonormal basis for H.

We develop our framework by viewing functions as vectors living in H. Hence, we use the terms *vector* and *function* interchangeably.

To fix ideas, an example of such a Hilbert space is the set of all square-integrable functions defined on the finite interval $[0,1]$, and equipped with the inner product

$$\langle f, g\rangle_H = \int_{[0,1]} f(x)g(x)\,\mathrm{d}x.$$

One countable orthonormal basis for this Hilbert space is the trigonometric Fourier series $1 \cup \left\{\sqrt{2}\cos 2\pi nx, \sqrt{2}\sin 2\pi nx\right\}_{n=1}^{\infty}$.

2.3.1 Gaussian Measures in Hilbert Spaces

A random variable x is distributed according to a *Gaussian measure* on a real, separable Hilbert space H if, for all $f \in H$, the inner product $\langle f, x\rangle_H \in \mathbb{R}$ is distributed according to a one-dimensional Gaussian. Such a Gaussian measure is completely determined by its mean $m \in H$ and a *covariance operator*, defined as a bounded, self-adjoint, positive-semidefinite, linear operator $C : H \to H$ which satisfies:

$$\langle Cf, g\rangle_H = \langle f, Cg\rangle_H = \mathrm{Cov}\left[\langle f, x\rangle_H \langle g, x\rangle_H\right] = \mathbb{E}\left[\langle f-m, x\rangle_H \langle g-m, x\rangle_H\right],$$

for all $f, g \in H$. Hence we denote the law of x by $\mathrm{N}(m, C)$.

Let $\{e_n\}_{n=1}^{\infty}$ be an orthonormal basis of eigenvectors of C with corresponding eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$. We call C *trace-class*, if

$$\mathrm{Tr}(C) := \sum_{n=1}^{\infty} \langle Ce_n, e_n \rangle_H = \sum_{n=1}^{\infty} \lambda_n < \infty.$$

This condition is critical in infinite dimensions: for a Gaussian to be supported on H, its expected squared norm must be finite, and this value is equal to $\|m\|_H^2 + \mathrm{Tr}(C)$. A Gaussian with non-trace-class noise will have samples which are almost-surely unbounded in norm and hence do not belong to the Hilbert space H. Crucially, this rules out the isotropic Gaussian $\mathrm{N}(0, I)$. Hence, we focus only on the case of Gaussians with trace-class covariance.

For a covariance operator C, the *Cameron-Martin space*, H_C, is an (infinite-dimensional) subspace of H defined as the image of H under $C^{\frac{1}{2}}$. Equipped with the inner product $\langle f, g \rangle_{H_C} := \left\langle C^{-\frac{1}{2}} f, C^{-\frac{1}{2}} g \right\rangle_H$, the Cameron-Martin space is also a Hilbert space. Criticially in infinite dimensions, H_C is a strict subspace of H when C is trace-class: there exist vectors in H which are not in H_C. To see why, since C's eigenvalues sum to zero the eigenvalues of $C^{-\frac{1}{2}}$ diverge to infinity and hence $C^{-\frac{1}{2}}$ is unbounded on H. H_C consists only of vectors who.....

Intuitively.

Chapter 3

My third chapter

TODO

References

Albergo, M. S., Boffi, N. M., and Vanden-Eijnden, E. (2023a). Stochastic interpolants: A unifying framework for flows and diffusions.

Albergo, M. S., Goldstein, M., Boffi, N. M., Ranganath, R., and Vanden-Eijnden, E. (2023b). Stochastic interpolants with data-dependent couplings. *arXiv preprint arXiv:2310.03725*.

Anderson, B. D. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12(3):313–326.

Franzese, G., Rossi, S., Yang, L., Finamore, A., Rossi, D., Filippone, M., and Michiardi, P. (2023). How much is enough? a study on diffusion times in score-based generative models. *Entropy*, 25(4):633.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. (2021). Score-based generative modeling through stochastic differential equations.

Appendix A

How to install LaTeX

Windows OS

TeXLive package - full version

1. Download the TeXLive ISO (2.2GB) from https://www.tug.org/texlive/
2. Download WinCDEmu (if you don't have a virtual drive) from http://wincdemu.sysprogs.org/download/
3. To install Windows CD Emulator follow the instructions at http://wincdemu.sysprogs.org/tutorials/install/
4. Right click the iso and mount it using the WinCDEmu as shown in http://wincdemu.sysprogs.org/tutorials/mount/
5. Open your virtual drive and run setup.pl

or

Basic MikTeX - TeX distribution

1. Download Basic-MiKTeX(32bit or 64bit) from http://miktex.org/download
2. Run the installer
3. To add a new package go to Start » All Programs » MikTex » Maintenance (Admin) and choose Package Manager

4. Select or search for packages to install

TexStudio - TeX editor

1. Download TexStudio from
http://texstudio.sourceforge.net/#downloads

2. Run the installer

Mac OS X

MacTeX - TeX distribution

1. Download the file from
https://www.tug.org/mactex/

2. Extract and double click to run the installer. It does the entire configuration, sit back and relax.

TexStudio - TeX editor

1. Download TexStudio from
http://texstudio.sourceforge.net/#downloads

2. Extract and Start

Unix/Linux

TeXLive - TeX distribution

Getting the distribution:

1. TexLive can be downloaded from
http://www.tug.org/texlive/acquire-netinstall.html.

2. TexLive is provided by most operating system you can use (rpm,apt-get or yum) to get TexLive distributions

Installation

1. Mount the ISO file in the mnt directory

   ```
   mount -t iso9660 -o ro,loop,noauto /your/texlive####.iso /mnt
   ```

2. Install wget on your OS (use rpm, apt-get or yum install)

3. Run the installer script install-tl.

   ```
   cd /your/download/directory
   ./install-tl
   ```

4. Enter command ‘i’ for installation

5. Post-Installation configuration:
   http://www.tug.org/texlive/doc/texlive-en/texlive-en.html#x1-320003.4.1

6. Set the path for the directory of TexLive binaries in your .bashrc file

For 32bit OS

For Bourne-compatible shells such as bash, and using Intel x86 GNU/Linux and a default directory setup as an example, the file to edit might be

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/i386-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

For 64bit OS

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/x86_64-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
```

```
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

Fedora/RedHat/CentOS:

```
sudo yum install texlive
sudo yum install psutils
```

SUSE:

```
sudo zypper install texlive
```

Debian/Ubuntu:

```
sudo apt-get install texlive texlive-latex-extra
sudo apt-get install psutils
```

Appendix B

Installing the CUED class file

LaTeX.cls files can be accessed system-wide when they are placed in the <texmf>/tex/latex directory, where <texmf> is the root directory of the user's TeXinstallation. On systems that have a local texmf tree (<texmflocal>), which may be named "texmf-local" or "localtexmf", it may be advisable to install packages in <texmflocal>, rather than <texmf> as the contents of the former, unlike that of the latter, are preserved after the LaTeXsystem is reinstalled and/or upgraded.

It is recommended that the user create a subdirectory <texmf>/tex/latex/CUED for all CUED related LaTeXclass and package files. On some LaTeXsystems, the directory look-up tables will need to be refreshed after making additions or deletions to the system files. For TeXLive systems this is accomplished via executing "texhash" as root. MIKTeXusers can run "initexmf -u" to accomplish the same thing.

Users not willing or able to install the files system-wide can install them in their personal directories, but will then have to provide the path (full or relative) in addition to the filename when referring to them in LaTeX.