Variational Inference for Stochastic Control of
Infinite Dimensional Systems

George I. Boutselis*, Marcus Pereira* and Evangelos A. Theodorou
Autonomous Control and Decision Systems Laboratory
School of Aerospace Engineering
Georgia Institute of Technology

Abstract

This paper develops a variational inference framework for control of infinite dimensional stochastic systems. We employ a measure theoretic approach which relies on the generalization of Girsanov’s theorem, as well as the relation between relative entropy and free energy. The derived control scheme is applicable to a large class of stochastic, infinite dimensional systems, and can be used for trajectory optimization and model predictive control. Our work opens up new research avenues at the intersection of stochastic control, inference and information theory for dynamical systems described by stochastic partial differential equations.

1 Introduction

In many practical applications, one faces the problem of controlling dynamical systems represented by stochastic partial differential equations (SPDEs). Examples can be found, for instance, in fluid mechanics, open quantum systems, turbulence, plasma physics and partially observable stochastic control Chow [2007], Da Prato and Zabczyk [2014], Mikulevicius and Rozovskii [2004], G. Dumont and Longtine [2017], Pardoux [1980], Bang et al. [1994], Cont [2005], Knopf and Weber [2017].

Despite the importance of such applications, the majority of works on computational stochastic control has been dedicated to finite dimensional systems. These are systems represented by stochastic differential equations (SDEs), and can be found in a plethora of applications from robotics and autonomous systems, to computational neuroscience, biology and finance. In contrast, the literature is lacking works on scalable/implementable control schemes for stochastic, infinite dimensional systems. To this end, this paper tries to bridge the gap between theory and implementation of stochastic control in infinite dimensions. Our approach is based on the free energy-relative entropy duality, and utilizes elements from stochastic calculus in Hilbert spaces. The resulting methodology avoids restrictive assumptions about the problem formulation, and can be applied to a broad class of semilinear SPDEs.

Previous work in the area of control of SPDEs has focused on very specific systems, and typically consists of theoretical results on the existence and uniqueness of solutions. References Prato and Debussche [1999] and Feng [2006] share some common characteristics with our paper. In particular, the former work investigates explicit solutions of the Hamilton-Jacobi-Bellman (HJB) equation for the stochastic Burgers equation. The derivation is based on the exponential transformation of the value function, as well as the transformation of the backward HJB equation into a forward Kolmogorov equation. Then, the explicit solution is recovered through the forward Feynman-Kac lemma and a probabilistic representation of the value function. The work in Feng [2006] extends the large deviation theory to infinite dimensional systems, and creates connections to HJB theory. The analysis therein shows that a free energy-like function corresponds to the value function of a deterministic optimal control problem under a specific cost functional. This connection is established by proving

*Equal contribution

Preprint. Work in progress.
that the aforementioned free energy-like function satisfies the HJB equation of an infinite horizon
deterministic optimal control problem.

On the computational side, the work in [Lou et al. 2009] proposes a model predictive control
methodology for nonlinear dissipative SPDEs. The key idea lies in model reduction; that is, the
transformation of the original SPDE into a set of coupled stochastic differential equations. Once this
finite dimensional representation is obtained, a model predictive control methodology is developed
and is then applied on the Kuramoto-Sivanshinsky equation. Another work on computational control
of the aforementioned SPDE can be found in [Gomes et al. 2017]. This approach shares similarities
with the one in [Lou et al. 2009], in that a finite dimensional representation of the SPDE is utilized,
rendering thus the use of standard control theory feasible.

To the best of our knowledge, the framework developed in this paper is the first step towards explicitly
designing implementable, numerical stochastic control algorithms in infinite dimensions. In contrast
to prior work (see [Lou et al. 2009], [Gomes et al. 2017]), the proposed approach treats SPDEs as
time-indexed stochastic processes taking values in an infinite dimensional space. The core of our
methodology relies on sampling stochastic paths from the dynamics, and computing the associated
trajectory costs. Grounded on the theory of stochastic calculus in Hilbert spaces, we are not restricted
to any particular finite representation of the original system. Besides the theoretical implications, this
fact is also beneficial from a computational standpoint. Specifically, the obtained expressions for our
control updates are independent of the method used to actually simulate the SPDEs. This further
implies that the required sampled paths can be obtained by employing the scheme that is more suitable
to each particular problem setup (e.g., finite differences, Galerkin methods or finite elements). Finally,
we note that this work can be considered as a generalization of the Path Integral and information
theoretic control method [Todorov 2009], [Theodorou and Todorov 2012], [Theodorou 2015], [Kappen
2005]. As such, the proposed stochastic control algorithm can be efficiently applied in a Model
Predictive Control (MPC) fashion, and inherits the ability to deal with non-quadratic cost functions
and nonlinear dynamics.

The rest of the paper is organized as follows: In section 2 we provide some important definitions
and theorems on infinite dimensional stochastic systems. In section 3 we discuss the free energy
and relative entropy relation. Based on this connection, section 4 derives our stochastic control
method by performing inference in Hilbert spaces. Furthermore, in subsection 4.1 we develop an
iterative version of our framework, which is subsequently tested in simulation in section 5. Section 6
concludes the paper.

2 Preliminaries - Stochastic Calculus

In this paper we consider infinite dimensional stochastic systems of the following form:
\[ dX = \alpha X(t)dt + F(X(t))dt + G(X(t))dW(t), \quad X(0) = \xi \]  
(1)
defined on the probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with filtration \(\mathcal{F}_t, t \geq 0,\) for the time interval \(t \in [0, T]\).

Let, \(H\) and \(U\) Hilbert spaces, then \(\alpha : D(\alpha) \subset H \to H\) is an infinitesimal generator, \(\xi\) is an \(\mathcal{F}_0\)-measurable \(H\)-valued random variable, while \(F : H \to H\) and \(G : U \to H\) are nonlinear mappings
that satisfy properly formulated Lipschitz and linear growth conditions (associated with the existence
and uniqueness of solutions for infinite dimensional stochastic systems - see [Da Prato and Zabczyk
2014] Theorem 7.2)). The term \(W(t) \in U\) corresponds to a Q-Wiener process that is defined based
on the following proposition (see [Da Prato and Zabczyk 2014] Chapter 4)). We use the notation
\(X(\cdot, \omega)\) to denote a state trajectory.

**Proposition 2.1.** Let \(\{e_i\}_{i=1}^\infty\) be a complete orthonormal system for the Hilbert Space \(U\) such that
\(Qe_i = \lambda_i e_i\). Here, \(\lambda_i\) is the eigenvalue of \(Q \in L(U)\) that corresponds to eigenvector \(e_i\), and \(L(U)\)
denotes the space of linear operators acting on \(U\). Then, a Q-Wiener process \(W(t) \in U\) satisfies the
following properties:

i) \(W\) is a Gaussian process on \(U\) with mean and variance:
\[ \mathbb{E}[W(t)] = 0, \quad \mathbb{E}[W(t)W(t)] = tQ, \quad t \geq 0. \]  
(2)

ii) For arbitrary \(t \geq 0\), \(W\) has the following expansion:
\[ W(t) = \sum_{j=1}^\infty \sqrt{\lambda_j} \beta_j(t)e_j, \]  
(3)
where \( \beta_j(t) \) are real valued brownian motions that are mutually independent on \((\Omega, \mathcal{F}, \mathbb{P})\).

In this paper we will make use of Girsanov’s theorem for systems evolving on Hilbert spaces. To this end, let us introduce the Hilbert space \( U_0 := Q^{1/2}(U) \subset U \) with inner product: \((u, v)_{U_0} := \langle Q^{-1/2}u, Q^{-1/2}v \rangle_U \), \( \forall u, v \in U_0 \). The following proposition is from Da Prato and Zabczyk [2014, Theorem 10.18]:

**Proposition 2.2** (Girsanov). Let \( \Omega \) be a sample space with a \( \sigma \)-algebra \( \mathcal{F} \). Consider the following \( H \)-valued stochastic processes:

\[
\begin{align*}
\text{d}X &= (\omega'X + F(X))dt + G(X)dW(t), \quad X(0) = x \\
\text{d}\tilde{X} &= (\omega\tilde{X} + F(\tilde{X}))dt + \tilde{B}(\tilde{X})dt + G(\tilde{X})dW(t), \quad \tilde{X}(0) = x,
\end{align*}
\]

where \( W \in U \) is a \( Q \)-Wiener process with respect to the measure \( \mathbb{P} \). Moreover, \( \forall t \in C([0, T]; H) \) let the law of \( X \) defined as \( \mathcal{L}(X(\cdot, \omega)) \in \Gamma := \mathbb{P}(\omega \in \Omega | X(\cdot, \omega) \in \Gamma) \). Similarly, the law of \( \tilde{X} \), is defined as \( \mathcal{L}(\tilde{X}(\cdot, \omega)) \in \Gamma := \mathbb{P}(\omega \in \Omega | \tilde{X}(\cdot, \omega) \in \Gamma) \). Then

\[
\mathcal{L}(\tilde{X}(\cdot, \omega)) = \mathbb{E}_\mathbb{P} \left[ \exp \left( \int_0^T (\psi(s), dW(s))_{U_0} - \frac{1}{2} \int_0^T ||\psi(s)||_{U_0}^2 ds \right) | X(\cdot) \in \Gamma \right],
\]

where \( \psi(t) := G^{-1}(X(t))B(X(t)) \in U_0 \). Here, we write for brevity \( \mathcal{L}(\omega) \equiv \mathcal{L}(X(\cdot, \omega) \in \Gamma) \).

**Proof.** Define the process:

\[
\hat{W}(t) := W(t) - \int_0^t \psi(s)ds.
\]

Based on Da Prato and Zabczyk [2014, Theorem 10.18], \( \hat{W} \) is a \( Q \)-Wiener process with respect to a measure \( \mathbb{Q} \) determined by:

\[
d\mathbb{Q}(\omega) = \exp \left( \int_0^T (\psi(s), dW(s))_{U_0} - \frac{1}{2} \int_0^T ||\psi(s)||_{U_0}^2 ds \right) \mathbb{P}(\omega) d\mathbb{P} \approx \exp \left( \int_0^T (\psi(s), d\hat{W}(s))_{U_0} + \frac{1}{2} \int_0^T ||\psi(s)||_{U_0}^2 ds \right) \mathbb{P}(\omega) d\mathbb{P}.
\]

Now, using eq. 7, eq. 4 is rewritten as:

\[
\text{d}X = (AX + F(X))dt + G(X)d\hat{W}(t) = (AX + F(X))dt + B(X)dt + G(X)d\hat{W}(t) \tag{9}
\]

Notice that the above SPDE has the same form as \( \text{(5)} \). Therefore, under the introduced measure \( \mathbb{Q}, X \) becomes equivalent to \( \mathbb{P} \). However, under the measure \( \mathbb{P} \), the SPDE in \( \text{(9)} \) behaves as the original system in \( \text{(4)} \). In other words, eqs. \( \text{(4)} \) and \( \text{(9)} \) describe the same system on \((\Omega, \mathcal{F}, \mathbb{P})\). From the uniqueness of solutions and the aforementioned reasoning, one has

\[\mathbb{P}(\{X \in \Gamma\}) = \mathbb{Q}(\{X \in \Gamma\}).\]

The result follows from \( \text{(8)} \). \( \square \)

To conclude this section, we note that when \( \lambda_j = 1, \forall j \), \( W(t) \) corresponds to a cylindrical Wiener process (space-time white noise). In that case, the series in \( \text{(3)} \) converges in another Hilbert space \( U_1 \supset U \), when the inclusion \( i : U \rightarrow U_1 \) is Hilbert-Schmidt. For more details see Da Prato and Zabczyk [2014].

### 3 Relative Entropy and Free Energy Dualities in Hilbert Spaces

In this section we provide the relation between free energy and relative entropy. The relation is valid for general probability measures including measures defined on path spaces induced by infinite dimensional stochastic systems. Here we will consider the general measures \( \mathcal{L} \) and \( \tilde{\mathcal{L}} \).

**Definition 3.1.** (Free Energy) Let \( \mathcal{L} \in \mathcal{P} \) a probability measure and let the function \( J \equiv J(X(\cdot, \omega)) : L^p \rightarrow \mathbb{R}_+ \) be a measurable function. Then the following term:

\[
V = \frac{1}{\mu} \log_e \int_{\Omega} \exp(\mu J) d\mathcal{L}(\omega), \tag{10}
\]

is called the free energy\(^2\) of \( J \) with respect to \( \mathcal{L} \) and \( \mu \in \mathbb{R} \).

\(^2\)The function \( \log_e \) denotes the natural logarithm.
with respect to $Q$ which has statistical mechanics interpretation. The equilibrium probability measure has the classical

**Theorem 3.1.** Let $L \in \mathcal{P}$ and $\hat{L} \in \mathcal{P}$ then the relative entropy of $\hat{L}$ with respect to $L$ is defined as:

$$S(\hat{L} \mid | L) = \left\{ \begin{array}{ll} -\int_{\Omega} \frac{d\hat{P}(\omega)}{dP(\omega)} \log \frac{d\hat{P}(\omega)}{dP(\omega)} d\mathcal{L}(\omega), & \text{if } \hat{L} \ll L, \\
+\infty, & \text{otherwise}, \end{array} \right.$$

where “$\ll$” denotes absolute continuity of $\hat{L}$ with respect to $L$ and $L_1$ denotes the space of Lebesgue measurable functions on $[0, \infty)$. We say that $\hat{L}$ is absolutely continuous with respect to $L$ and we write $\hat{L} \ll L$ if $\hat{L}(B) = 0 \Rightarrow L(B) = 0$, $\forall B \in \mathcal{F}$.

The free energy and relative entropy relationship is expressed by the theorem that follows:

**Theorem 3.1.** Let $(\Omega, \mathcal{F})$ be a measurable space. Consider $L, \hat{L} \in \mathcal{P}$ and the definitions of free energy and relative entropy as expressed in definitions 3.1 and 3.2. Under the assumption that $Q_X \ll \hat{P}_X$, the following inequality holds:

$$-\frac{1}{\rho} \log \mathbb{E}_L \left[ \exp(-\rho J) \right] \leq \left[ \mathbb{E}_L(\rho) - \frac{1}{\rho} S(\hat{L} \mid | L) \right],$$

where $\mathbb{E}_L, \mathbb{E}_{\hat{L}}$ is the expectation under the probability measure $L, \hat{L}$ respectively and $\rho \in \mathbb{R}_+$ and $J : L^p \rightarrow \mathbb{R}_+$. The inequality in (11) is the so called Legendre Transform.

By defining the free energy as temperature $T = \frac{1}{\rho}$ the Legendre transformation has the form:

$$V \leq E - TS,$$

which has statistical mechanics interpretation. The equilibrium probability measure has the classical form:

$$d\mathcal{L}^\omega = \exp(-\rho J) d\mathcal{L}(\omega) \\frac{1}{\Omega} \exp(-\rho J) d\mathcal{L}(\omega),$$

(13)

To verify that the measure in (13) is the optimal measure it suffices to substitute (13) in (11) and show that the inequality collapses to an equality [Theodorou 2015]. The statistical physics interpretation of inequality (12) is that, maximization of entropy results in reduction of the available energy. At the thermodynamic equilibrium the entropy reaches its maximum and the inequality collapses to equality. It can be shown that when the measures $\hat{L}$ and $L$ are associated to paths generated by control and uncontrolled semi-linear SPDEs, then the free energy is value function that satisfies the HJB equation of an infinite dimensional stochastic optimal control problem. This observation motivates the use of (13) for the development of stochastic control algorithms.

**4 Variational Inference and Control in Hilbert Spaces**

In this section we will derive our numerical algorithm for controlling stochastic infinite dimensional systems. To simplify our expressions, we will consider without loss of generality SPDEs with additive noise. Let the uncontrolled and controlled version of an $H$-valued process be given respectively by:

$$dX(t) = (\mathcal{A}X + F(X(t)))dt + \frac{1}{\sqrt{\rho}} dW(t), \quad \text{and} \quad \hat{X}(t) = (\mathcal{A}\hat{X} + F(X(t)) + \mathcal{W}(t))dt + \frac{1}{\sqrt{\rho}} dW(t)$$

(14)

both with initial condition: $X(0) = \hat{X}(0) = \xi$. Here, $W \in U = H$ is a $Q$-Wiener process on $(\Omega, \mathcal{F}, \mathcal{P})$ with covariance operator $Q \in L(U)$. As in the previous section, the uncontrolled dynamics are equivalent to:

$$dX(t) = (\mathcal{A}X + F(X(t)) + \mathcal{W}(t))dt + \frac{1}{\sqrt{\rho}} d\hat{W}(t),$$

(15)

with respect to $\mathcal{P}$. Here, $\hat{W}$ is a $Q$-Wiener process with respect to another measure $Q$. The law of the uncontrolled states, $\mathcal{L}(\cdot)$, defines a measure on the path space via (13) as $\mathcal{L}(\omega) := \mathcal{F}(\omega) X(\cdot, \omega) \in \Gamma$. Similarly, the law of controlled trajectories is $\hat{\mathcal{L}}(\omega) := \mathcal{P}(\omega) \hat{X}(\cdot, \omega) \in \Gamma$. Finally, we suppose that there exists an optimal controller $\mathcal{W}^*$ which corresponds to the law of optimal trajectories, $\hat{\mathcal{L}}^*(\cdot)$.
In this section we derive controllers by formulating a new optimization problem in which we make use of the measure theoretic approach. We are looking for a control input \( U \) that minimizes the distance to the optimal path law. That is:

\[
U^* = \arg\max_{U} S(\mathcal{L}^*|\tilde{\mathcal{P}})
\]  

(16)

Under the parameterization \( \mathcal{U} = \mathcal{U}(X(t); \theta) \) the problem above will take the form:

\[
\theta^* = \arg\max \left[ -\int_{\Omega} \frac{d\mathcal{L}^*(\omega)}{d\tilde{\mathcal{P}}(\omega)} \log \frac{d\mathcal{L}^*(\omega)}{d\tilde{\mathcal{P}}(\omega)} d\tilde{\mathcal{P}}(\omega) \right] = \arg\min \left[ \int_{\Omega} \log \frac{d\mathcal{L}^*(\omega)}{d\tilde{\mathcal{P}}(\omega)} d\tilde{\mathcal{P}}(\omega) \right].
\]

(17)

To perform the optimization we will consider the chain rule property for the Radon-Nikodym derivative. For instance, this results in the following expression:

\[
\frac{d\tilde{\mathcal{L}}}{d\mathcal{L}} = \frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} = \exp \left( \int_{0}^{T} \langle \psi(s), dW(s) \rangle_{t_0} - \frac{1}{2} \int_{0}^{T} \|\psi(s)\|_{t_0}^2 ds \right),
\]

(18)

where \( \psi(t) := \sqrt{\mathcal{P}} \mathcal{U}(t) \in U \). In this paper we will parameterize our infinite dimensional control as follows:

\[
\mathcal{U}(t) = \sum_{i=1}^{N} m_i u_i(t) \in U \equiv H,
\]

(19)

so that

\[
\mathcal{U}(t)(x) = \sum_{i=1}^{N} m_i(x) u_i(t) = m(x)^T u(t) \in \mathbb{R}.
\]

(20)

Here, \( m_i \in U \) are design functions that specify how the actuation is incorporated into the infinite dimensional dynamical system. Under this parameterization, the change of measure between the two SPDEs takes the form:

\[
\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} = \exp \left( \sqrt{\mathcal{P}} \int_{0}^{T} u(t)^T \tilde{m}(t) - \rho \frac{1}{2} \int_{0}^{T} u(t)^T M u(t) dt \right),
\]

(21)

where

\[
\tilde{m}(t) := \left[ \langle m_1, dW(t) \rangle_{t_0}, \ldots, \langle m_N, dW(t) \rangle_{t_0} \right]^T \in \mathbb{R}^N,
\]

\[
M \in \mathbb{R}^{N \times N}, \quad (M)_{ij} := \langle m_i, m_j \rangle_{t_0}.
\]

(22)

The following theorem provides the optimal control \( u^* \) for the case of the controlled SPDEs of the form in \([14]\).

**Lemma 4.1.** (Variational Stochastic Control) Consider the controlled SPDE in \([14]\) and let the following objective function:

\[
u^* = \arg\max_{\nu} S(\mathcal{L}^*|\tilde{\mathcal{P}})
\]

(24)

The probability measure \( \mathcal{L}^* \) is induced by the optimally controlled SPDE in \([14]\) and has the form:

\[
d\mathcal{L}^*(\omega) = \frac{\exp(-\rho J(Xq))d\mathcal{L}(\omega)}{\int_{\Omega} \exp(-\rho J(X)) d\mathcal{L}(\omega)},
\]

(25)

The probability measure \( \tilde{\mathcal{L}} \) is induced by controlled trajectories of the SPDEs when infinite dimensional control \( \mathcal{U}(t) \) is determined by \([26]\) and \( u(t) \) in \((20)\) is parameterized as follows:

\[
u(t) = u_i = u(t_i) \quad \text{if} \quad i\Delta t \leq t < (i+1)\Delta t, \quad \forall t \in [0, T]
\]

(26)

with \( i = \{0, 1, \ldots, L\} \). Under the aforementioned representation, the optimal control is provided by the following expression:

\[
u^*_i = \frac{1}{\sqrt{\rho \Delta t}} M^{-1} \mathcal{L}^* \left[ \frac{\exp(-\rho J)}{\mathcal{L}^*[\exp(-\rho J)]} \delta u \right], \quad \text{and} \quad \delta u := \int_{t_i}^{t_{i+1}} \tilde{m}(t).
\]

(27)
Proof. Under the parameterization $\mathcal{U}(x, t) = m(x)^T u(t)$ the problem above will take the form:

$$u^* = \arg\min \left[ \int_\Omega \log \frac{d\mathcal{L}^u}{d\mathcal{L}^0}(\omega) d\mathcal{L}(\omega) \right] = \arg\min \left[ \int_\Omega \log \frac{d\mathcal{L}^u}{d\mathcal{L}^0}(\omega) d\mathcal{L}(\omega) \right].$$

By using (21) minimization of the last expression is equivalent to the minimization of the expression:

$$\mathbb{E}_{\mathcal{L}^u} \left[ \log \frac{d\mathcal{L}(\omega)}{d\mathcal{L}^0}(\omega) \right] = -\sqrt{\rho} \mathbb{E}_{\mathcal{L}^u} \left[ \int_0^T u(t)^T \tilde{m}(t) \right] + \frac{1}{2} \rho \int_0^T u(t)^T \mathbf{M}u(t) dr.$$

The goal is to find the function $u^*(\cdot)$ which minimizes. However, since we inevitably apply the control in discrete time it suffices to consider the class of step functions:

$$\mathbb{E}_{\mathcal{L}^u} \left[ \log \frac{d\mathcal{L}(\omega)}{d\mathcal{L}^0}(\omega) \right] = -\sqrt{\rho} \sum_{i=0}^{L-1} u_i^T \mathbb{E}_{\mathcal{L}^u} \left[ \int_{t_i}^{t_{i+1}} \tilde{m}(t) \right] + \frac{1}{2} \rho \sum_{i=0}^{L-1} u_i^T \mathbf{M}u_i \Delta t,$$

where we have used the fact that $\mathbf{M}$ is symmetric and constant with respect to time. Minimization of the expression above with respect to $u_i$ results in:

$$u_i^* = \frac{1}{\sqrt{\rho} \Delta t} \mathbf{M}^{-1} \mathbb{E}_{\mathcal{L}^u} \left[ \int_{t_i}^{t_{i+1}} \tilde{m}(t) \right]. \quad (28)$$

Since we cannot sample from the $\mathcal{L}^u$, we need to change the expectation to be an expectation with respect to the uncontrolled dynamics, $\mathcal{L}^0$. We can then directly sample trajectories from $\mathcal{L}^0$ to approximate the controls. The change in expectation is achieved by applying the Radon-Nikodym derivative. The result is equation (27). \hfill \Box

### 4.1 Iterative Control of SPDEs

We derive an iterative scheme that can be used for stochastic optimization and be implemented in a receding horizon fashion. In particular, let us consider the controlled dynamics at iteration $i^{th}$ given by:

$$dX^{(i)}(t) = (dX^{(i)} + F(X^{(i)}) + \mathcal{U}^{(i)}(t))dt + \frac{1}{\sqrt{\rho}} dW(t), \quad (29)$$

where $\mathcal{U}^{(i)}(t)$ is the control at the $i^{th}$ iteration. As we have already shown, the uncontrolled dynamics can be equivalently written as:

$$dX(t) = (dX + F(X(t)))dt + \frac{1}{\sqrt{\rho}} dW(t) = (dX + F(X(t)) + \mathcal{U}^{(i)}(t))dt + \frac{1}{\sqrt{\rho}} dW^{(i)}(t), \quad (30)$$

where $W^{(i)}$ is a $Q$-Wiener process with respect to some measure $\mathbb{Q}^{(i)}$ with:

$$W^{(i)}(t) := W(t) - \int_0^t \rho \mathcal{U}^{(i)}(s) ds. \quad (31)$$

Again here we define the path measure $\mathcal{L}^{(i)} := \mathbb{P}(\omega | X^{(i)}(\cdot), \omega) \in \Gamma$ induced by (29) and the path measure $\mathcal{L}^{(0)} := \mathbb{P}(\omega | X^{(0)}(\cdot), \omega) \in \Gamma$ induced by (30). Then according to (22) we have:

$$\frac{d\mathcal{L}^{(i)}}{d\mathcal{L}^{(0)}} = \exp \left( \sqrt{\rho} \sum_{k=0}^{L-1} u^{(i)\top}_k \tilde{m}^{(i)}(t) + \rho \frac{L-1}{2} \sum_{k=0}^{L-1} u_k^{(i)\top} \mathbf{M}u_k^{(i)} \Delta t \right), \quad (32)$$

where

$$\mathbb{R}^N \ni \tilde{m}^{(i)}(t) := \begin{bmatrix} \langle m_1, dW^{(i)}(t) \rangle_{U_0}, \ldots, \langle m_N, dW^{(i)}(t) \rangle_{U_0} \end{bmatrix}^\top, \quad (33)$$

**Lemma 4.2.** (Iterative Stochastic Control) Consider the controlled SPDE in (14) and the parameterization of the control as specified by (20) and (26). The iterative control scheme is given by the following expression:

$$u_j^{(i+1)} = u_j^{(i)} + \frac{1}{\sqrt{\rho} \Delta t} \mathbf{M}^{-1} \mathbb{E}_{\mathcal{L}^{(i)}} \left[ \exp(-\rho \tilde{J}) \delta u_j^{(i)} \right], \quad \text{and} \quad \delta u_j^{(i)} = \int_{t_j}^{t_{j+1}} \tilde{m}^{(i)}(t). \quad (34)$$
The expectation in \((34)\) is taken with respect to the probability path measure \(\mathcal{L}^{(i)}\) induced by sampled trajectories generated using \((29)\).

Proof. In order to derive the iterative scheme, we perform one step of importance sampling. In particular, instead of sampling from the uncontrolled SPDE \((14)\) to evaluate the expectation in \((27)\) we sample using the controlled SPDE \((29)\). In addition, we modify \((27)\) so that to perform the appropriate change of measure between the uncontrolled version of infinite dimensional dynamics and the controlled version at iteration \(i\). Next we modify equations \((27)\) by considering \((32)\) and \((31)\).

\[
\delta u_{j}^{(i+1)} = \frac{1}{\sqrt{\rho \Delta t}} \mathbf{M}^{-1} \mathbb{E}_{\mathcal{L}^{(i)}} \left[ \frac{\exp(-\rho J)}{d\mathcal{L}^{(i)}} \frac{\exp(-\rho J)}{d\mathcal{L}^{(i)}} \delta u_{j} \right].
\]

Regarding \(\delta u_{j}\), one has:

\[
\left( \int_{t_{j}}^{t_{j+1}} \mathbf{m}(t) \right)_{i} = \int_{t_{j}}^{t_{j+1}} \langle m_{i}, dW(t) \rangle_{\mathcal{U}_{0}} = \int_{t_{j}}^{t_{j+1}} \langle m_{i}, dW(t) \rangle_{\mathcal{U}_{0}} + \sqrt{\rho} \int_{t_{j}}^{t_{j+1}} \langle m_{i}, dW(t) \rangle_{\mathcal{U}_{0}} + \sqrt{\rho} \left[ \langle m_{1}, m_{1} \rangle_{\mathcal{U}_{0}}, ..., \langle m_{1}, m_{N} \rangle_{\mathcal{U}_{0}} \right] u_{j}^{(i)} \Delta t.
\]

It follows that:

\[
\int_{t_{j}}^{t_{j+1}} \mathbf{m}(t) = \int_{t_{j}}^{t_{j+1}} \mathbf{m}(t) + \sqrt{\rho} \Delta t u_{j}^{(i)}.
\]

Substitution of the Radon-Nikodym derivative yields the final result in \((34)\). Note that under \(Q^{(i)}\) renders \(W^{(i)}\) a standard \(Q\)-Wiener process.

For the purposes of implementation we will approximate the optimal controls \((34)\) as:

\[
\langle \delta u_{j}^{(i)} \rangle_{i} := \sum_{x=1}^{R} \langle m_{i}, \sqrt{\lambda_{x}} e_{x} \rangle_{\mathcal{U}_{0}} \Delta \theta x^{(i)}(t_{j}),
\]

where \(\Delta \theta x^{(i)}(t_{j}) \sim \mathcal{N}(0, \Delta t)\) under \(Q^{(i)}\). Next we discuss the application of the iterative stochastic control on two examples of SPDEs.

5 Experiments

In this section, we present simulation results on two infinite dimensional stochastic systems. The first systems is the stochastic Heat equation and the second system is the Nagumo SPDE. The iterative stochastic optimal control is used for open loop trajectory optimization and for MPC.

**Heat SPDE:** The 1-D stochastic heat equation with homogeneous Dirichlet boundary conditions can be used to simulate the diffusion of heat along a rod insulated on the sides and exposed to freezing conditions at the end points. Our experiments consisted of achieving desired temperature levels at specific positions along a rod in the presence of space-time stochastic disturbing forces. As seen in Fig. [1] the MPC has robust performance compared to open-loop controller with the mean temperature profile closer to the desired temperature levels and tighter sigma bounds in the presence of space-time white noise.

**Nagumo SPDE:** The stochastic Nagumo equation with homogeneous Neumann boundary conditions is a reduced model for wave propagation of the voltage \(u\) in the axon of a neuron [Lord et al., 2014]. The Nagumo equation is expressed as follows:

\[
u_{t} = \varepsilon u_{xx} + u(1-u)(u-\alpha) + \sigma dW(t), \quad u(x,0) = u(x,a) = 0, u(0,x) = (1 + \exp(-(2 - x)/\sqrt{2}))^{-1}
\]
The parameter $\alpha$ determines the speed of a wave traveling down the length of the axon and $\varepsilon$ the rate of diffusion. From simulating the deterministic version of the above PDE for $a = 10$, $\varepsilon = 1$ and $\alpha = -0.5$, we observed that it requires about 10 seconds for the wave to propagate to the end of the axon. An open-loop infinite-dimensional controller was employed to accelerate the propagation of the voltage and to suppress the propagation of the voltage in about 2.5 seconds. The plots shown in the figure below demonstrate the achievement of desired behavior in the axon.

6 Conclusions

We present an information theoretic formulation for stochastic optimal control of infinite dimensional dynamical systems. The analysis relies on concepts drawn from the theory of stochastic calculus in Hilbert spaces, the relative entropy and free energy relation and its connections to stochastic dynamic programming. The resulting algorithm can be used for stochastic trajectory optimization and MPC for a large class of systems with dynamics governed by SPDEs. The work in this paper is a generalization of the path integral and information theoretic control to infinite dimensional spaces and is a significant step towards the development of scalable and real time control algorithms for infinite dimensional stochastic systems. Future directions involve the theoretical analysis of the convergence, application to higher order infinite dimensional systems, fully nonlinear SPDEs and application to real systems.

References

P.L. Chow. *Stochastic Partial Differential Equations*. Advances in Applied Mathematics, Taylor & Francis, 2007. ISBN 9781584884439. URL https://books.google.com/books?id=x901s88HJ0QC
G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2014. ISBN 9780521385299. URL https://books.google.com/books?id=Sid6pwAACAAJ

R. Mikulevicius and B. L. Rozovskii. Stochastic navier-stokes equations for turbulent flows. *SIAM Journal on Mathematical Analysis*, 35(5):1250–1310, 2004.

A. Payeur, G. Dumont and A. Longtin. A stochastic-field description of finite-size spiking neural networks. *PLOS Computational Biology*, 13, 08 2017.

E. Pardoux. Stochastic partial differential equations and filtering of diffusion processes. *Stochastics*, 3(1-4):127–167, 1980.

O. Bang, P. L. Christiansen, F. If, K. Ø. Rasmussen, and Y. B. Gaididei. Temperature effects in a nonlinear model of monolayer scheibe aggregates. *Phys. Rev. E*, 49:4627–4636, May 1994.

Rama Cont. Modeling term structure dynamics: An infinite dimensional approach. *International Journal of Theoretical and Applied Finance*, 08(03):357–380, 2005.

P. Knopf and J. Weber. Optimal control of a Vlasov-Poisson plasma by fixed magnetic field coils. *ArXiv e-prints*, 2017.

Giuseppe Da Prato and Arnaud Debussche. Control of the stochastic burgers model of turbulence. *SIAM Journal on Control and Optimization*, 37(4):1123–1149, 1999.

Jin Feng. Large deviation for diffusions and hamilton-jacobi equation in hilbert spaces. *Ann. Probab.*, pages 321–385, 01 2006. doi: 10.1214/009117905000000567.

Y. Lou, G. Hu, and P. D. Christofides. Model predictive control of nonlinear stochastic pdes: Application to a sputtering process. In *2009 American Control Conference*, June 2009.

S.N. Gomes, S. Kalliadasis, D.T. Papageorgiou, G.A. Pavliotis, and M. Pradas. Controlling roughening processes in the stochastic kuramoto-sivashinsky equation. *Physica D: Non-linear Phenomena*, 348:33 – 43, 2017. ISSN 0167-2789. doi: https://doi.org/10.1016/j.physd.2017.02.011. URL http://www.sciencedirect.com/science/article/pii/S0167278916301567

E. Todorov. Efficient computation of optimal actions. *Proceedings of the national academy of sciences*, 106(28):11478–11483, 2009.

E.A Theodorou and E. Todorov. Relative entropy and free energy dualities: Connections to path integral and kI control. In *the Proceedings of IEEE Conference on Decision and Control*, pages 1466–1473, Dec 2012.

E. A. Theodorou. Nonlinear stochastic control and information theoretic dualities: Connections, interdependencies and thermodynamic interpretations. *Entropy*, 17(5):3352, 2015.

H. J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, 11:P11011, 2005.

Gabriel J. Lord, Catherine E. Powell, and Tony Shardlow. *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2014. doi: 10.1017/CBO9781139017329.