arXiv:2403.01562v1 [math.PR] 3 Mar 2024

On Some Extensions of the Boué-Dupuis Variational Formula

Amarjit Budhiraja

March 5, 2024

Abstract

The Boué-Dupuis variational formula gives a representation for log Laplace transforms of bounded measurable functions of a finite dimensional Brownian motion on a compact time interval as an infimum of a suitable cost over a collection of non-anticipative control processes. This variational formula has proved to be very useful in studying a variety of large deviation problems. In this article we collect some extensions of this basic result that have appeared in disparate venues in studying a broad range of large deviation questions. Some of these results can be found in a unified way in the recent book [10], while others, to date, have been scattered at various places in the literature. The latter category includes, in particular, variational representations, when the stochastic dynamical system of interest has in addition to a driving Lévy noise, another source of randomness, e.g. due to a random initial condition; when the functionals of interest depends on infinite-length paths of a Lévy process; when the noise process is a Gaussian process with long-range dependence, e.g. a fractional Brownian motion, etc. The goal of this survey article is to present these diverse variational formulas in a systematic manner.

Keywords. Variational representations, fractional Brownian motion, Hilbert space valued Brownian motion, Lévy process, large deviations, Laplace asymptotics.

In memory of Professor K.R. Parthasarathy.

1 Introduction

The paper [2] established a variational formula for positive bounded functionals of paths of a finite dimensional Brownian motion over a compact time interval that has proved to be extremely useful in studying a wide range of large deviation problems. This result, now well nown as the Boué-Dupuis variational formula, says the following. Fix $T \in (0, \infty)$. Let $W = \{W(t) : 0 \leq t \leq T\}$ be a d-dimensioanl standard Wiener process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $\{\mathcal{G}_t\}_{0\leq t\leq T}$ the $\mathbb{P}$-augmentation of the filtration $\{\sigma\{W(s) : 0 \leq s \leq t\}\}_{0\leq t\leq T}$. Let $\mathcal{A}$ denote the class of $\mathcal{G}_t$-predictable processes $v : [0, T] \times \Omega \to \mathbb{R}^d$ that satisfy $\int_0^T \|v(t)\|^2 dt < \infty$ a.s. Let $F : \mathcal{C}_0^T(\mathbb{R}^d) \to \mathbb{R}$ be a bounded measurable map. Here and throughout, for a Polish space $\mathcal{E}$, $\mathcal{C}_0^T(\mathcal{E})$ will denote the space of continuous functions from $[0, T]$ to $\mathcal{E}$ equipped with the uniform topology. Then

$$-\log \mathbb{E} \exp\{-F(W)\} = \inf_{v\in\mathcal{A}} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|^2 ds + F(W + \int_0^{\cdot} v(s)ds)\right]. \tag{1.1}$$

In this we article we collect several important extensions of this result which have proved to be useful in a range of problems, including those from infinite dimensional stochastic dynamical systems, stochastic partial differential equations, jump-diffusions and other stochastic systems driven by Lévy processes, Gaussian processes with long-range dependence, mean-field interacting particle systems, stochastic particle systems on discrete lattices, random graph models, etc. Several of these extensions can be found in the recent book[10], but there are a few others, such as those for fractional Brownian motions, or those for functionals involving additional randomness than that

from a Lévy process, or representations on an infinite time horizon, which have found use in different problems, that to-date have not been recorded in a single place. Our goal in this survey article is to introduce these various representations in a systematic way and explain the different contexts where they have been used.

The following notation and terminology will be used. A filtration $\{\mathcal{F}_t\}$ on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to satisfy *the usual conditions* if the filtration is right continuous and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets. On a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ a d-dimensional Brownian motion B is said to be a $\{\mathcal{F}_t\}$- Brownian motion (or a $\{\mathcal{F}_t\}$- Wiener process) if $\{B(t) - B(s) : t \geq s\}$ is independent of $\mathcal{F}_s$ for all $s \geq 0$. For a random variable X with values in some Polish space $\mathcal{E}$, $\mathcal{L}(\mathcal{E})$ will denote the probability distribution of X which is an element of $\mathcal{P}(\mathcal{E})$, namely the space of probability measures on $\mathcal{E}$. Borel σ-field on $\mathcal{E}$ will be denoted as $\mathcal{B}(\mathcal{E})$. For $\gamma, \theta \in \mathcal{P}(\mathcal{E})$, the relative entropy of γ with respect to θ, denoted as $R(\gamma\|\theta)$ is defined to be $\int_{\mathcal{E}} \log \frac{d\gamma}{d\theta} d\gamma$ when $\gamma \ll \theta$ and ∞ otherwise. For a sequence $\{X_n\}$ of $\mathcal{E}$ valued random, the convergence of X_n to an $\mathcal{E}$ valued random variable X in distribution will be denoted as $X_n \Rightarrow X$. Occasionally, to emphasize the dependence on the probability measure, the expected value on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ will be written as $\mathbb{E}_{\mathbb{P}}$. For a Hilbert space $(H, \langle \cdot, \cdot \rangle)$, $L^2([0,T] : H)$ will denote the Hilbert space of functions $f : [0,T] \to H$ such that $\int_0^T \|f(s)\|^2 ds < \infty$, where for $x \in H$, $\|x\|^2 = \langle x, x \rangle$. The Hilbert space $L^2([0,\infty) : H)$ is defined similarly. We will frequently use same notation for related but different objects in different sections of the article. For example the class of controls will usually be denoted as $\mathcal{A}$ but their precise definition will change from section to section.

2 General Filtrations and Infinite Dimensional Brownian Motions.

Results in this section can be found in [3, 5]. They also appear in the recent book [10]. Although Theorem 2.2 is not explicitly given in these references, it can be deduced easily from the results therein. We provide proof details for reader's convenience.

The representation given in 1.1 requires the class of controls to be adapted with respect to the (augmented) Brownian filtration. In the study of large deviation problems for mean field interacting particle systems [4, 8, 9, 11] it is useful to have an extension of the representation which allows for a filtration that is larger than the Brownian filtration. Such a representation is important in the proof of the lower bound in the large deviation principle for the path occupation measure associated with such interacting particle systems. It turns out that an optimal point μ of the rate function, over a given set, corresponds to the probability law of the state process, in a controlled stochastic system, which apriori is not known to be adapted to the filtration generated by the driving noise. This subtle but crucial issue requires a variational representation that allows for a larger filtration.

An extension in a different direction is motivated by systems driven by infinite dimensional Brownian motions that are the basic models in the field of stochastic partial differential equations. In order to study large deviation properties of such systems, it becomes important to establish a representation that is applicable to the various models of infinite dimensional Brownian motions, such as a Hilbert space valued Brownian motion, a cylindrical Brownian motion, space-time Brownian sheet, etc.

A representation that allows for both of these features was obtained in [3, 5]. We only present the representation in the setting of a Hilbert space valued Brownian motion and refer the reader to [10] for the other related settings.

We begin with some basic definitions. Fix $T < \infty$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which is given a filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ satisfying the usual conditions. Let $(H, \langle \cdot, \cdot \rangle)$ be a real separable Hilbert space and let Q be a strictly positive, symmetric, trace class operator on H. We recall the definition of a Hilbert space valued Wiener process.

Definition 2.1. *A continuous H-valued stochastic process $\{W(t)\}_{0 \leq t \leq T}$ is called a Q-Wiener process with respect to the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ if the following hold: (i) For every nonzero $h \in H$, $\langle Qh, h \rangle^{-1/2} \langle W(t), h \rangle$ is a one dimensional standard Wiener process. (ii) For every $h \in H$, $\langle W(t), h \rangle$ is an $\mathcal{F}_t$-martingale.*

Define $H_0 \doteq Q^{1/2}H$. Then H_0 is a Hilbert space with inner product defined as

$$\langle h, k\rangle_0 \doteq \langle Q^{-1/2}h, Q^{-1/2}k\rangle, \text{ for } h, k \in H_0.$$

Denote by $\tilde{\mathcal{A}}$ the collection of all $\mathcal{F}_t$-predictable $u : [0,T] \times \Omega \to H_0$ that satisfy $\int_0^T \|u(t)\|_0^2 dt < \infty$. Also denote by $\mathcal{G}_t$ the $\mathbb{P}$-augmentation of the Brownian filtration $\sigma\{W(s) : 0 \le s \le t\}$ and let $\mathcal{A}$ be the subcollection of $\tilde{\mathcal{A}}$ consisting of $\mathcal{G}_t$-predictable H_0 valued processes. Then the following representation can be found in [3, 5, 10].

Theorem 2.1. *Let $F : \mathcal{C}_0^T(H) \to \mathbb{R}$ be a bounded measurable map. Then*

$$-\log \mathbb{E}\exp\{-F(W)\} = \inf_{v\in\mathcal{R}} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^{\cdot} v(s)ds)\right], \tag{2.1}$$

where $\mathcal{R} \in \{\mathcal{A}, \tilde{\mathcal{A}}\}$.

Thus the representation in particular says that the infimum on the right side taken over the larger collection $\tilde{\mathcal{A}}$ is same as that over the subcollection $\mathcal{A}$. In fact the infimum can be taken over a much more smaller subcollection, a fact which is very useful when one would like to use weak convergence arguments in proofs of large deviation principles. To describe this subcollection, let for $M \in (0,\infty)$, define

$$S_M \doteq \left\{u \in L^2([0,T] : H_0) : \int_0^T \|u(s)\|_0^2 ds \le M\right\}.$$

Denote by $\mathcal{A}_{b,M}$ (resp. $\tilde{\mathcal{A}}_{b,M}$) the subcollection of $\mathcal{A}$ (resp. $\tilde{\mathcal{A}}$) consisting of u that take values in S_M. Let $\mathcal{A}_b \doteq \cup_{M>0}\mathcal{A}_{b,M}$ and $\tilde{\mathcal{A}}_b \doteq \cup_{M>0}\tilde{\mathcal{A}}_{b,M}$. Then [10, Theorem 8.3] shows that (2.1) in fact holds for any $\mathcal{R} \in \{\mathcal{A}, \tilde{\mathcal{A}}, \mathcal{A}_b, \tilde{\mathcal{A}}_b\}$.

One can take the infimum over even a further smaller class, namely the collection of simple processes. This fact was crucially exploited in [1] in the proof of the large deviation principle for a Brownian interacting particle system with local interactions. This work, in the proof of the large deviation upper bound, required certain estimates on Dirichlet forms associated with the controlled state processes, which in turn relied on the smoothness properties of the density functions of the laws of these controlled processes. The proof of these regularity properties given in [1] made key use of the fact that the controls are piecewise constant. Thus the result that in the variational representation the infimum can be taken over such controls played a central role in the proof.

We now present this result.

Definition 2.2. *A process $v \in \tilde{\mathcal{A}}$ (resp. $\mathcal{A}$) is called simple, if there exists a $k \in \mathbb{N}$, $N \in \mathbb{N}$ and $0 = t_1 \le t_2 \le \cdots \le t_{k+1} = T$, such that*

$$v(s,\omega) \doteq \sum_{j=1}^{k} X_j(\omega)\mathbf{1}_{(t_j,t_{j+1}]}(s), \ (s,\omega) \in [0,\infty) \times \Omega,$$

where, for each $j = 1, \ldots, k$, X_j is a real $\mathcal{F}_{t_j}$ (resp. $\mathcal{G}_{t_j}$) measurable random variable satisfying $|X_j| \le N$. We denote the collection of all such simple processes as $\tilde{\mathcal{A}}_s$ (resp. $\mathcal{A}_s$).

The following result says that the infimum in (2.1) in fact holds for any $\mathcal{R} \in \{\mathcal{A}, \tilde{\mathcal{A}}, \mathcal{A}_b, \tilde{\mathcal{A}}_b, \mathcal{A}_s, \tilde{\mathcal{A}}_s\}$. Although the result follows readily from the results in [3], below we give a brief argument for reader's convenience.

Theorem 2.2. *Let $F : \mathcal{C}_0^T(H) \to \mathbb{R}$ be a bounded measurable map. Then*

$$-\log \mathbb{E}\exp\{-F(W)\} = \inf_{v\in\mathcal{R}} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^{\cdot} v(s)ds)\right],$$

where $\mathcal{R} \in \{\mathcal{A}_s, \tilde{\mathcal{A}}_s\}$.

Proof. First suppose that $F : \mathcal{C}_0^T(H) \to \mathbb{R}$ is a continuous and bounded map. In that case, from [3, Lemma 3.5], we have that

$$\inf_{v\in\mathcal{A}_s} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right] = \inf_{v\in\tilde{\mathcal{A}}} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right]$$

and so the statement in the theorem follows from Theorem 2.1. Now consider a general real bounded measurable map F on $\mathcal{C}_0^T(H)$. Then there is a sequence of continuous maps $F_n : \mathcal{C}_0^T(H) \to \mathbb{R}$ such that $\|F_n\|_\infty \le \|F\|_\infty < \infty$ and $F_n \to F$ $\mathbb{P}$-a.s. Fix $\varepsilon > 0$ and choose $v^* \in \tilde{\mathcal{A}}_b$ such that

$$\mathbb{E}\left[\frac{1}{2}\int_0^T \|v^*(s)\|_0^2 ds + F(W + \int_0^\cdot v^*(s)ds)\right] \le \inf_{v\in\tilde{\mathcal{A}}_b} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right] + \varepsilon.$$

Since $v^* \in \tilde{\mathcal{A}}_b$, the law of $W + \int_0^\cdot v^*(s)ds$ is mutually absolutely continuous with respect to that of W and so $F_n(W + \int_0^\cdot v^*(s)ds)$ converges a.s. to $F(W + \int_0^\cdot v^*(s)ds)$. Thus

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v^*(s)\|_0^2 ds + F_n(W + \int_0^\cdot v^*(s)ds)\right] \\
&\le \inf_{v\in\tilde{\mathcal{A}}_b} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right] + \varepsilon.
\end{aligned}$$

Using the first part of the proof, choose $v_n \in \mathcal{A}_s$ such that

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2}\int_0^T \|v_n(s)\|_0^2 ds + F_n(W + \int_0^\cdot v_n(s)ds)\right] \\
&\le \mathbb{E}\left[\frac{1}{2}\int_0^T \|v^*(s)\|_0^2 ds + F_n(W + \int_0^\cdot v^*(s)ds)\right] + n^{-1}.
\end{aligned}$$

From [10, Lemma 2.5] it follows that

$$\lim_{n\to\infty} \mathbb{E}\left|F_n(W + \int_0^\cdot v_n(s)ds) - F(W + \int_0^\cdot v_n(s)ds)\right| = 0.$$

Combining the last three displays we have

$$\begin{aligned}
&\inf_{v\in\mathcal{A}_s} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right] \\
&\le \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v_n(s)\|_0^2 ds + F(W + \int_0^\cdot v_n(s)ds)\right] \\
&\le \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v^*(s)\|_0^2 ds + F_n(W + \int_0^\cdot v^*(s)ds)\right] \\
&\le \inf_{v\in\tilde{\mathcal{A}}_b} \mathbb{E}\left[\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + F(W + \int_0^\cdot v(s)ds)\right] + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the result follows. □

3 Functionals of Brownian Motions with Additional Randomness

In some situations one is interested in studying the large deviations behavior of a stochastic dynamical system for which, in addition to a random driving noise, there is another source of randomness, coming e.g. from the initial state of the stochastic system. Two such settings were studied in [1, 11]. In order to motivate the extension considered in this section, we describe the setting in [11] in some detail.

3.1 Brownian particle systems with killing.

Let $\{X_i\}_{i\geq 1}$ be a sequence of i.i.d. exponential random variables with rate 1 and let $\{B_i(t), t \geq 0\}_{i\geq 1}$ be independent d-dimensional standard Brownian motions independent of $\{X_i\}_{i\geq 1}$. Define for $t \geq 0$ the random sub-probability measure $\mu^n(t)$ as the solution to the following equation

$$\mu^n(t) = \frac{1}{n}\sum_{i=1}^{n} \delta_{B_i(t)} \mathbf{1}_{\left\{X_i > \int_0^t \langle \zeta,\, \mu^n(s)\rangle\, ds\right\}}. \tag{3.1}$$

Here $\zeta : \mathbb{R}^d \to \mathbb{R}$ is a continuous function with sub-quadratic growth. Since a.s., we can enumerate $\{X_i\}_{i=1}^n$ in a strictly increasing order, the unique solution of (3.1) can be written explicitly in a recursive manner. It can be checked (see [11, Theorem 2.1]) that $\mu^n \doteq \{\mu^n(t)\}_{t\in[0,T]}$ converges, in the Skorokhod path space $\mathcal{D} \doteq \mathbb{D}([0,T] : \mathcal{M}(\mathbb{R}^d))$, where $\mathcal{M}(\mathbb{R}^d)$ is the space of sub-probability measures on $\mathbb{R}^d$ equipped with the weak convergence topology, in probability to μ where for $t > 0$, $\mu(t)$ has density $u(t,\cdot)$ given as the solution of

$$\frac{\partial u(t,x)}{\partial t} = \frac{1}{2}\Delta u(t,x) - \langle \zeta,\, u\rangle\, u(t,x),\ (t,x) \in (0,\infty)\times\mathbb{R}^d,\ \lim_{t\downarrow 0} u(t,\cdot) = \delta_0(\cdot),$$

Such particle systems are motivated by problems in biology, ecology, chemical kinetics, etc. For example, the simplest case where $\zeta \equiv 1$ corresponds to the case where the killing rate is proportional to the total number of particles alive and models a setting in which particles compete for a common resource. More general functions ζ are of interest as well and one interpretation of $\zeta(x)$ is the amount of resource consumed by a particle in state x. Similar particle systems arise in problems of mathematical finance as models for self exciting correlated defaults.

In studying large deviation properties of the sequence $\{\mu^n\}$ of $\mathcal{D}$ valued random variables one needs to understand the asymptotics of Laplace functionals of the form

$$-\frac{1}{n}\log \mathbb{E}[\exp(-nf(\mu^n))]$$

for real continuous and bounded functions f on $\mathcal{D}$. In view of the weak convergence approach to the study of large deviation problems[10, 13], it is then natural to derive a suitable controlled representation for the above Laplace functional and understand tightness and weak convergence properties of the various terms in this representation. Note that μ^n can be viewed as a measurable functional of the nd-dimensional Brownian motion $\mathbf{B}^n = (B_1, \ldots B_n)$ and an independent $\mathbb{R}^n_+$ dimensional random variable $\mathbf{X}^n = (X_1, \ldots, X_n)$ and so one can write

$$-\frac{1}{n}\log \mathbb{E}[\exp(-nF_n(\mathbf{B}^n, \mathbf{X}^n))]$$

where F_n is a suitable real bounded and measurable map on $\mathcal{C}_0^T(\mathbb{R}^{nd}) \times \mathbb{R}^n_+$. In the case where F_n only depends on $\mathbf{B}^n$, we saw that useful variational representations are given by (1.1) and Theorem 2.1. However the setting where the function F_n also depends on $\mathbf{X}^n$ is not covered by the above results.

For such settings one needs a variational representation that allows for functionals that, in addition to depending on a Brownian motion, depend also on an additional source of randomness. Such a representation was given in [11, Proposition 4.1]. We present this result below.

3.2 A Variational Representation with Additional Randomness.

Let $T < \infty$, and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space as in the previous section on which are given a d-dimensional standard $\mathcal{F}_t$-Brownian motion W and an $\mathcal{F}_0$-measurable random variable X, which takes values in a Polish space $\mathbb{S}$ and has probability law ρ. We present below a variational representation for $-\log \mathbb{E}\left[\exp\left(-F(W,X)\right)\right]$, where F is a real bounded measurable map on $\mathcal{C}_0^T(\mathbb{R}^d) \times \mathbb{S}$. Note that since X is $\mathcal{F}_0$-measurable, W and X are independent.

Consider the probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ on which we are given a d-dimensional standard Brownian motion $\bar{W}$ and an $\mathbb{S}$-valued random variable $\bar{X}$, which is independent of $\bar{W}$, with law Π. Let

$\{\hat{\mathcal{F}}_t\}$ be any filtration on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ satisfying the usual conditions, such that $\bar{W}$ is still a standard d-dimensional Brownian motion with respect to $\{\hat{\mathcal{F}}_t\}$ and $\bar{X}$ is $\hat{\mathcal{F}}_0$-measurable. One example of such a filtration is $\{\bar{\mathcal{F}}_t \doteq \sigma\{\bar{\mathcal{F}}_t^{\bar{W},\bar{X}} \cup \bar{\mathcal{N}}\}\}$, where $\bar{\mathcal{N}}$ is the collection of all $\bar{\mathbb{P}}$-null sets and $\bar{\mathcal{F}}_t^{\bar{W},\bar{X}} \doteq \sigma\{\bar{X}, \bar{W}(s) : s \leq t\}$. Let $\Upsilon_\Pi \doteq (\bar{\Omega}, \bar{\mathcal{F}}, \{\hat{\mathcal{F}}_t\}, \bar{\mathbb{P}})$ and consider the following collection of processes

$$\mathcal{A}(\Upsilon_\Pi) \doteq \left\{ u : \text{the process } u \text{ is } \hat{\mathcal{F}}_t\text{-predictable and } \bar{\mathbb{E}} \int_0^T \|u(s)\|^2 \, ds < \infty \right\}.$$

The following is [11, Proposition 4.1].

Theorem 3.1. *Let F be a real bounded measurable map on $\mathcal{C}_0^T(\mathbb{R}^d) \times \mathbb{S}$. Then*

$$\begin{aligned} &-\log \mathbb{E}\left[\exp\left(-F(W,X)\right)\right] \\ &= \inf_{\Pi, \Upsilon_\Pi} \inf_{u \in \mathcal{A}(\Upsilon_\Pi)} \left\{ R(\Pi\|\rho) + \bar{\mathbb{E}}\left[\frac{1}{2}\int_0^T \|u(s)\|^2 \, ds + F\left(\bar{W} + \int_0^\cdot u(s)\, ds, \bar{X}\right)\right]\right\}, \end{aligned}$$

where the outer infimum is over all $\Pi \in \mathcal{P}(\mathbb{S})$ and all systems Υ_Π.

Remark 3.2. *Denote for $M \in (0,\infty)$, by $\mathcal{A}_{b,M}(\Upsilon_\Pi)$ the subcollection of $\mathcal{A}(\Upsilon_\Pi)$ consisting of controls u that take value in S_M. Let $\mathcal{A}_b(\Upsilon_\Pi) \doteq \cup_{M>0}\mathcal{A}_{b,M}(\Upsilon_\Pi)$. Also, let $\mathcal{A}_s(\Upsilon_\Pi)$ be the subcollection of $\mathcal{A}(\Upsilon_\Pi)$ consisting of $\hat{\mathcal{F}}_t$-adapted simple processes. Then the collection $\mathcal{A}(\Upsilon_\Pi)$ in the statement of Theorem 3.1 can be replaced by either $\mathcal{A}_b(\Upsilon_\Pi)$ or $\mathcal{A}_s(\Upsilon_\Pi)$.*

Furthermore, a similar representation can be written for an infinite dimensional Brownian motion, e.g. for a H-valued Wiener process as in Section 2.

For details of the proof of Theorem 3.1 we refer the reader to [11] but we make some comments on the proof idea.

Consider the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, where $\tilde{\Omega} = \mathcal{C}_0^T(\mathbb{R}^d)$, $\tilde{\mathcal{F}} = \mathcal{B}(\mathcal{C}_0^T(\mathbb{R}^d))$ and $\tilde{\mathbb{P}}$ is the d-dimensional Wiener measure. Namely, under $\tilde{\mathbb{P}}$ the canonical coordinate process $\tilde{W} \doteq \{\tilde{W}(t,\tilde{\omega}) \doteq \tilde{\omega}(t), 0 \leq t \leq T\}$ is a standard d-dimensional Brownian motion with respect to the filtration $\{\tilde{\mathcal{F}}_t^{\tilde{W}} \doteq \{\sigma\{\tilde{W}(s) : s \leq t\}\}$. Let $\{\tilde{\mathcal{F}}_t\}$ be the augmented filtration, namely $\tilde{\mathcal{F}}_t \doteq \sigma\{\tilde{\mathcal{F}}_t^{\tilde{W}} \cup \tilde{\mathcal{N}}\}$ and $\tilde{\mathcal{N}}$ is the collection of all $\tilde{\mathbb{P}}$-null sets. Define

$$\tilde{F}(x) \doteq -\log \tilde{\mathbb{E}}\left[\exp\left(-F(\tilde{W}, x)\right)\right], \quad x \in \mathbb{S}.$$

From the independence between W and X we see that

$$-\log \mathbb{E}\left[\exp\left(-F(B,X)\right)\right] = -\log \mathbb{E}\left[\exp\left(-\tilde{F}(X)\right)\right].$$

Applying classical results of Donsker–Varadhan (cf. [10, Proposition 2.2]) to the right side, we have the following representation formula from the above equality

$$-\log \mathbb{E}\left[\exp\left(-F(B,X)\right)\right] = \inf_{\Pi \in \mathcal{P}(\mathbb{S})} \left[R(\Pi\|\rho) + \int_{\mathbb{S}} \tilde{F}(x)\Pi(dx)\right]. \tag{3.2}$$

Consider the collection of processes

$$\mathcal{A} \doteq \left\{ u : \text{the process } u \text{ is } \tilde{\mathcal{F}}_t\text{-predictable and } \tilde{\mathbb{E}} \int_0^T \|u(s)\|^2 \, ds < \infty \right\}.$$

From (1.1) we now have the following variational formula

$$\tilde{F}(x) = \inf_{u \in \mathcal{A}} \tilde{\mathbb{E}}\left[\frac{1}{2}\int_0^T \|u(s)\|^2 \, ds + F\left(\tilde{W} + \int_0^\cdot u(s)\, ds, x\right)\right],$$

which together with (3.2) gives

$$\begin{aligned} &-\log \mathbb{E}\left[\exp\left(-F(B,X)\right)\right] \\ &= \inf_{\Pi \in \mathcal{P}(\mathbb{S})} \left\{ R(\Pi\|\rho) + \int_{\mathbb{S}} \inf_{u \in \mathcal{A}} \tilde{\mathbb{E}}\left[\frac{1}{2}\int_0^T \|u(s)\|^2 \, ds + F\left(\tilde{W} + \int_0^\cdot u(s)\, ds, x\right)\right] \Pi(dx)\right\}. \end{aligned}$$

The above representation is inconvenient to use directly in large deviation proofs and the main challenge in using it, in comparison to the representation in Theorem 3.1 is the presence of the infimum under the integral on the right side. Nevertheless, the above representation is the starting point of the proof of Theorem 3.1 which uses discretization and measurable selection arguments to show not only that the integral and the infimum can be interchanged in the representation but also that the minimal filtration $\tilde{\mathcal{F}}_t$ can be replaced by a larger filtration with respect to which which B is a martingale. The fact that one can allow for a general filtration turns out to be crucial in the proof of the lower bound of the large deviation proof for the collection $\{\mu^n\}$ described earlier in the section.

4 Infinite Time Horizon

In some problems one needs a variational formula for functionals of the paths of a Brownian motion over an infinite time horizon. We describe one such situation below.

4.1 Empirical Measures of State Processes Driven by a Fractional Brownian Motion.

One situation where the need to consider infinite-length paths of Brownian motion arises naturally is in the study of path empirical measures of ergodic solutions of stochastic differential equations driven by a fractional Brownian motion (fBM). The term ergodicity here needs to be interpreted suitably as the solution processes are not Markov, however there is a natural way to augment the state of the system with an infinite length path of a Brownian motion to get a Markovian state descriptor (see [15]). The starting point of obtaining such a Markovian description is the Mandelbrot - Van Ness decomposition of a fBM in terms of a two sided Brownian motion. This decomposition is as follows.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is given a real two-sided Wiener process $\{W(t) : -\infty < t < \infty\}$. Such a process is characterized by the property that for every $t \in \mathbb{R}$, $\{W^t(s) \doteq W(t-s) - W(t)\}_{s\geq 0}$ is a standard Brownian motion independent of $\{W(t+u) : u \geq 0\}$. Fix $H \in (0,1)$. Then a real fractional Brownian motion with Hurst parameter H, denoted as $B_H(\cdot)$, has the following representation: For $t \geq 0$,

$$B_H(t) = \int_{-\infty}^{0} \mathcal{G}(r)(dW(r+t) - dW(r)) = \int_{-\infty}^{0} (\mathcal{G}(r+t) - \mathcal{G}(r))dW(r) + \int_0^t \mathcal{G}(r-t)dW(r),$$

where $\mathcal{G}(r) = \alpha_H r^{H-1/2}$, for $r \geq 0$, and α_H is a suitable constant.

One is interested in stochastic differential equations (SDE) driven by such a driving noise. For simplicity of presentation, considered the simplest such equation which describes a fractional Ornstein-Uhlenbeck process:

$$dX(t) = -\alpha X(t)dt + dB_H(t).$$

The ergodicity behavior of $\{X(t), t \geq 0\}$ and in fact of much more general processes given as solutions of SDE driven by a fBM have been studied in [15] and in several subsequent works. These results, in particular, characterize the law of the limit of the occupation measures $\{L^T, T > 0\}$ as $T \to \infty$, where

$$L^T(\cdot) \doteq \frac{1}{T}\int_0^T \delta_{X(s)}\, ds,\ T > 0.$$

In order to study the large deviation behavior of the above collection of $\mathcal{P}(\mathbb{R})$ valued random variables, one needs to characterize the asymptotic behavior of Laplace functionals of the following form:

$$-\frac{1}{T}\log \mathbb{E}\exp\{-Tf(L^T)\}$$

where $f : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ is a bounded continuous map. Note that L^T is a measurable map of the infinite path of a standard Brownian motion, namely the collection $\{W^T(s) = W(T-s) - W(T) : s \geq 0\}$. Thus, for a suitable bounded measurable map $F_T : \mathcal{C}_0^\infty(\mathbb{R}) \to \mathbb{R}$,

$$-\frac{1}{T}\log \mathbb{E}\exp\{-Tf(L^T)\} = -\frac{1}{T}\log \mathbb{E}\exp\{-TF_T(W^T)\}.$$

Here, for a Polish space $\mathcal{E}$, $\mathcal{C}_0^\infty(\mathcal{E})$ denotes the space of continuous functions from $[0,\infty)$ to $\mathcal{E}$, equipped with the local uniform topology. Thus, in order to study asymptotic behavior of such a quantity, it is of interest to develop a variational representation for functionals of infinite-length paths of a standard Brownian motion.

4.2 Representation for Infinte Length Brownian Paths.

We will only provide detailed arguments for a one dimensional Brownian motion. Analogous results can be established for a Brownian motion with values in a separable Hilbert space by combining the arguments here with those in [3]. We also remark that, in order to study ergodicity properties of $X(\cdot)$ and other fractional diffusions, [15] and subsequent papers view W^T as an element of a suitable path space $\mathcal{H}_H$ of Hölder continuous functions. Since the Borel σ-fields on $\mathcal{H}_H$ and $\mathcal{C}_0^\infty(\mathbb{R})$ are the same, one immediately obtains from Theorem 4.1 below also a representation for real bounded measurable functionals on $\mathcal{H}_H$.

Let $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered probability space where the filtration satisfies the usual conditions. Let $\{W(t)\}_{t\geq 0}$ be a standard $\mathcal{F}_t$-Wiener process on this space. The $\mathbb{P}$-augmentation of the filtration $\{\sigma\{W(s) : 0 \leq s \leq t\}\}_{t\geq 0}$ will be denoted as $\{\mathcal{G}_t\}_{t\geq 0}$. Note that $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \geq 0$. We denote by $\mathcal{A}$ (resp. $\tilde{\mathcal{A}}$) the collection of all $\mathbb{R}$-valued $\mathcal{G}_t$-predictable (resp. $\mathcal{F}_t$-predictable) processes $v : [0,\infty) \times \Omega \to \mathbb{R}$ that satisfy

$$\mathbb{P}\left\{\int_0^\infty |v(t)|^2 dt < \infty\right\} = 1.$$

It will be convenient to consider the following subcollections of $\mathcal{A}$ and $\tilde{\mathcal{A}}$ as well. Let, for $M \in (0,\infty)$,

$$S_M \doteq \left\{u \in L^2([0,\infty) : \mathbb{R}) : \int_0^\infty |u(s)|^2 ds \leq M\right\},$$

and $\mathcal{A}_{b,M}$ (resp. $\tilde{\mathcal{A}}_{b,M}$) be elements of $\mathcal{A}$ (resp. $\tilde{\mathcal{A}}$) that take values in S_M a.s. Define

$$\mathcal{A}_b \doteq \cup_{M\in\mathbb{N}} \mathcal{A}_{b,M},\ \tilde{\mathcal{A}}_b \doteq \cup_{M\in\mathbb{N}} \tilde{\mathcal{A}}_{b,M}.$$

The following is the main variational representation of the section.

Theorem 4.1. *Let $F : \mathcal{C}_0^\infty(\mathbb{R}) \to \mathbb{R}$ be a bounded and measurable map. Then*

$$-\log \mathbb{E}\exp\{-F(W)\} = \inf_{v\in\mathcal{R}} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right],$$

where $\mathcal{R} \in \{\mathcal{A}, \tilde{\mathcal{A}}, \mathcal{A}_b, \tilde{\mathcal{A}}_b\}$.

In preparation for the proof of the theorem we give sime preliminary results. For $v \in \tilde{\mathcal{A}}_b$, and $T \in [0,\infty)$, let

$$L_\infty(v) \doteq \exp\left\{\int_0^\infty v(s)dW(s) - \frac{1}{2}\int_0^\infty |v(s)|^2 ds\right\},\ L_T(v) \doteq \exp\left\{\int_0^T v(s)dW(s) - \frac{1}{2}\int_0^T |v(s)|^2 ds\right\}.$$

The following result is a simple consequence of Girsanov's theorem.

Lemma 4.2. *Let $v \in \tilde{\mathcal{A}}_b$. Then $\mathbb{E}L_\infty(v) = 1$. Define the probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ as*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \doteq L_\infty(v) = \exp\left\{\int_0^\infty v(s)dW(s) - \frac{1}{2}\int_0^\infty |v(s)|^2 ds\right\}.$$

Then

$$\tilde{W}(t) \doteq W(t) - \int_0^t v(s)ds,\ t \geq 0$$

is a $\mathcal{F}_t$-Wiener process on $(\Omega, \mathcal{F}, \mathbb{Q}, \{\mathcal{F}_t\}_{t\geq 0})$.

Proof. Fix $v \in \tilde{\mathcal{A}}_b$. Then there is a $M < \infty$ such that $v \in \tilde{\mathcal{A}}_{b,M}$. Since, for every $T \in [0,\infty)$, $\int_0^T |v(s)|^2 ds \le \int_0^\infty |v(s)|^2 ds \le M$, we have from Novikov's criterion (cf. [19, Corollary VIII.1.16]) $\mathbb{E}(L_T(v)) = 1$ for all $T \in [0,\infty)$. Also, using Cauchy-Schwarz inequality, it is easy to verify that $\sup_{T\in[0,\infty)} \mathbb{E}(L_T(v))^2 \le e^M < \infty$ which shows that $\{L_T(v), T \ge 0\}$ is a uniformly integrable $\mathcal{F}_t$-martingale on $(\Omega, \mathcal{F}, \mathbb{P})$. Noting that $L_T(v) \to L_\infty(v)$ in probability, we now see that in fact the convergence holds in L^1 as well and consequently $\mathbb{E}L_\infty(v) = 1$ proving the first statement in the lemma. Also, it follows that, for $T \in [0,\infty)$, $E(L_\infty(v) \mid \mathcal{F}_T) = L_T(v)$ a.s. By Girsanov's theorem (cf. [17, Theorem 3.5.1]) it now follows that $\tilde{W}$ is a $\mathcal{F}_t$-Wiener process on $(\Omega, \mathcal{F}, \mathbb{Q}, \{\mathcal{F}_t\}_{t\ge 0})$. This completes the proof of the second statement. □

The following collection of simple processes will be useful.

Definition 4.1. *A process $v \in \tilde{\mathcal{A}}$ (resp. $\mathcal{A}$) is called simple, if there exists a $T \in (0,\infty)$, $k \in \mathbb{N}$, $N \in \mathbb{N}$ and $0 = t_1 \le t_2 \le \cdots \le t_{k+1} = T$, such that*

$$v(s,\omega) \doteq \sum_{j=1}^{k} X_j(\omega)\mathbf{1}_{(t_j,t_{j+1}]}(s),\ (s,\omega) \in [0,\infty)\times\Omega,$$

where, for each $j = 1,\ldots,k$, X_j is a real $\mathcal{F}_{t_j}$ (resp. $\mathcal{G}_{t_j}$) measurable random variable satisfying $|X_j| \le N$. We denote the collection of all such simple processes as $\tilde{\mathcal{A}}_s$ (resp. $\mathcal{A}_s$).

Note that $\mathcal{A}_s \subset \mathcal{A}_b$, $\tilde{\mathcal{A}}_s \subset \tilde{\mathcal{A}}_b$.

Since a simple process is zero after some finite time horizon, the known results for finite time horizon filtrations (see e.g. [10, Lemma 8.7 and Equation (8.14)]) give the following result. For $v \in \mathcal{A}_b$, denote $W^v(\cdot) \doteq W(\cdot) - \int_0^\cdot v(s)ds$ and define the probability measure $\mathbb{Q}^v$ on $(\Omega, \mathcal{F})$ as $d\mathbb{Q}^v \doteq L_\infty(v)d\mathbb{P}$.

Lemma 4.3. *For every $v_0 \in \tilde{\mathcal{A}}_s$, a bounded measurable map $F : \mathcal{C}_0^\infty(\mathbb{R}) \to \mathbb{R}$, and $\varepsilon > 0$ there is a $v \in \mathcal{A}_s$ such that*

$$\mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right] \le \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v_0(s)|^2 ds + F(W + \int_0^\cdot v_0(s)ds)\right] + \varepsilon.$$

Furthermore, for every $v \in \mathcal{A}_s$ there is a $\tilde{v} \in \mathcal{A}_s$ such that $\mathbb{Q}^{\tilde{v}} \circ (W^{\tilde{v}}, \tilde{v})^{-1} = \mathbb{P} \circ (W, v)^{-1}$.

In the above lemma when talking about the distribiution of $v, \tilde{v}$, these are regarded as random variables in S_M, for a suitable $M \in (0,\infty)$, where this space is equipped with the weak toplogy inherited from that on the Hilbert space $L^2([0,\infty) : \mathbb{R})$.

We now complete the proof of Theorem 4.1.

Proof of Theorem 4.1. We begin by proving the upper bound

$$-\log \mathbb{E}\exp\{-F(W)\} \le \inf_{v\in\mathcal{R}} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right] \tag{4.1}$$

Since $\tilde{\mathcal{A}}$ is the largest collection, it suffices to show the result with $\mathcal{R} = \tilde{\mathcal{A}}$. We will first show the inequality in (4.1) (without the infimum) for an arbitrary $v \in \tilde{\mathcal{A}}_s$. Note that for this, in view of the first part of Lemma 4.3, it suffices to establish the inequality for an arbitrary $v \in \mathcal{A}_s$. Consider now such a v. From the second part of Lemma 4.3 it follows that there is a $\tilde{v} \in \mathcal{A}_s$ such that $\mathbb{Q}^{\tilde{v}} \circ (W^{\tilde{v}}, \tilde{v})^{-1} = \mathbb{P} \circ (W, v)^{-1}$. From the Donsker-Varadhan variational formula (cf. [10, Proposition 2.2(a)]) we have that

$$-\log \mathbb{E}\exp\{-F(W)\} \le R(\mathbb{Q}^{\tilde{v}}\|\mathbb{P}) + \int_\Omega F(W)d\mathbb{Q}^{\tilde{v}}.$$

Note that

$$\begin{aligned} R(\mathbb{Q}^{\tilde{v}}\|\mathbb{P}) &= \int_\Omega \log\left(\frac{d\mathbb{Q}^{\tilde{v}}}{d\mathbb{P}}\right) d\mathbb{Q}^{\tilde{v}} = \mathbb{E}_{\mathbb{Q}^{\tilde{v}}}\left(\int_0^\infty \tilde{v}(s)dW(s) - \frac{1}{2}\int_0^\infty |\tilde{v}(s)|^2 ds\right) \\ &= \mathbb{E}_{\mathbb{Q}^{\tilde{v}}}\left(\int_0^\infty \tilde{v}(s)dW^{\tilde{v}}(s) + \frac{1}{2}\int_0^\infty |\tilde{v}(s)|^2 ds\right) = \mathbb{E}\left(\int_0^\infty v(s)dW(s) + \frac{1}{2}\int_0^\infty |v(s)|^2 ds\right) \end{aligned}$$

$$= \frac{1}{2}\mathbb{E}\int_0^\infty |v(s)|^2 ds,$$

where the fourth equality uses the second part of Lemma 4.3. Also, by another application of this result, we have

$$\int_\Omega F(W)d\mathbb{Q}^{\tilde{v}} = \int_\Omega F(W^{\tilde{v}} + \int_0^\cdot \tilde{v}(s)ds)d\mathbb{Q}^{\tilde{v}} = \mathbb{E}F(W + \int_0^\cdot v(s)ds).$$

Combining the last two displays, we now have the inequality in (4.1) (without the infimum) for any $v \in \mathcal{A}_s$ and thus, as discussed previously, also for any $v \in \tilde{\mathcal{A}}_s$. Now consider a $v \in \tilde{\mathcal{A}}$. Without loss of generality, $\mathbb{E}\int_0^\infty |v(s)|^2 ds < \infty$. Thus, for each $n \in \mathbb{N}$ there is a $T_n < \infty$ such that $\mathbb{E}\int_{T_n}^\infty |v(s)|^2 ds \leq n^{-1}$. Also, we can find (cf. [16, Lemma II.1.1]]), for each $n \in \mathbb{N}$, a $v_n \in \tilde{\mathcal{A}}_s$ that satisfying $v_n(t) = 0$ for all $t \geq T_n$ such that

$$\mathbb{E}\int_0^\infty |v_n(t) - v(t)|^2 dt = \mathbb{E}\int_0^{T_n} |v_n(t) - v(t)|^2 dt + \mathbb{E}\int_{T_n}^\infty |v(t)|^2 dt \leq 2n^{-1}. \tag{4.2}$$

Note that

$$\sup_n \int_0^\infty |v_n(t)|^2 dt \leq 2\int_0^\infty |v(t)|^2 dt + 4 < \infty.$$

From the fact that the inequality in (4.1) holds for any $v \in \mathcal{A}_s$, we have

$$-\log \mathbb{E}\exp\{-F(W)\} \leq \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v_n(s)|^2 ds + F(W + \int_0^\cdot v_n(s)ds)\right]. \tag{4.3}$$

Note that, from (4.2), for any $T < \infty$,

$$E\sup_{0\leq t\leq T} |\int_0^t v_n(s)ds - \int_0^t v(s)ds|^2 \leq TE\int_0^\infty |v_n(s) - v(s)|^2 ds \to 0, \quad \text{as } n \to \infty.$$

Thus we have that $\theta_n \doteq \mathcal{L}(W + \int_0^\cdot v_n(s)ds) \to \mathcal{L}(W + \int_0^\cdot v(s)ds) \doteq \theta$, as $n \to \infty$, as probability measures on $\mathcal{C}_0^\infty(\mathbb{R})$. Also, denoting the probability law of W by θ_0, note that $\theta_0 = \mathbb{Q} \circ (W + \int_0^\cdot v_n(s)ds)$, where $\mathbb{Q}$ is as in Lemma 4.2 with v there replaced by $-v_n$. Thus

$$\begin{aligned} R(\theta_n\|\theta_0) &= R(\mathbb{P} \circ (W + \int_0^\cdot v_n(s)ds)^{-1}\|\mathbb{Q} \circ (W + \int_0^\cdot v_n(s)ds)^{-1}) \\ &\leq R(\mathbb{P}\|\mathbb{Q}) = \mathbb{E}\left(\int_0^\infty v_n(s)dW(s) + \frac{1}{2}\int_0^\infty |v_n(s)|^2 ds\right) = \mathbb{E}\int_0^\infty |v_n(s)|^2 ds. \end{aligned}$$

Thus $\sup_n R(\theta_n\|\theta_0) < \infty$. This together with $\theta_n \to \theta$ now implies (see [10, Lemma 2.5]) that $F(W + \int_0^\cdot v_n(s)ds) \Rightarrow F(W + \int_0^\cdot v(s)ds)$. Together with the fact from (4.2) that $\mathbb{E}\int_0^\infty |v_n(t))|^2 dt$ converges to $\mathbb{E}\int_0^\infty |v(t))|^2 dt$, we now have from (4.3) that this inequality holds with v_n replaced by v. Since $v \in \mathcal{A}$ is arbitrary, this proves the upper bound in (4.1) for $\mathcal{R} = \tilde{\mathcal{A}}$ and therefore for any $\mathcal{R} \in \{\mathcal{A}, \mathcal{A}_b, \tilde{\mathcal{A}}, \tilde{\mathcal{A}}_b\}$.

We now establish the complementary lower bound:

$$-\log \mathbb{E}\exp\{-F(W)\} \geq \inf_{v\in\mathcal{R}} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right]$$

It suffices to show the inequality for $\mathcal{R} = \mathcal{A}_b$ since that is the smallest collection. Consider first the case where for some $T \in (0, \infty)$, and a bounded measurable map $F_T : \mathcal{C}_0^T(\mathbb{R}) \to \mathbb{R}$, $F(w(\cdot)) = F_T(w(\cdot \wedge T))$ for all $w \in \mathcal{C}_0^\infty(\mathbb{R})$. In this case, denoting the subcollection of $\mathcal{A}_b$ that consists of predictable $v : [0, \infty) \times \Omega \to \mathbb{R}$ that satisfy $v(t, \omega) = 0$ for $t > T$ by $\mathcal{A}_{b,T}$, we have (cf. [10, Section 8.1.4]) that

$$-\log \mathbb{E}\exp\{-F(W)\} \geq \inf_{v\in\mathcal{A}_{b,T}} \mathbb{E}\left[\frac{1}{2}\int_0^T |v(s)|^2 ds + F_T\left(W(\cdot \wedge T) + \int_0^{\cdot\wedge T} v(s)ds\right)\right]$$

$$\geq \inf_{v\in\mathcal{A}_b} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right].$$

Next, let $F : \mathcal{C}_0^\infty \to \mathbb{R}$ be a general bounded measurable map. Let $\{T_n\}_{n\geq 1}$ be an increasing sequence such that $T_n \uparrow \infty$. Define $\mathcal{G}_n \doteq \sigma\{W(s) : s \leq T_n\}$. Note that by the martingale convergence theorem $F_n \doteq \mathbb{E}(F \mid \mathcal{G}_n)$ converges a.s. to F as $n \to \infty$. Also $\|F_n\|_\infty \leq \|F\|_\infty$. Furthermore, since F_n is $\mathcal{G}_n$ measurable, there is a bounded measurable map $\tilde{F}_n : \mathcal{C}_0^{T_n}(\mathbb{R}) \to \mathbb{R}$ such that $F_n(W(\cdot)) = \tilde{F}_n(W(\cdot \wedge T_n))$ a.s. This also says that for any $v \in \mathcal{A}_b$, $F_n(W(\cdot) + \int_0^\cdot v(s)ds) = \tilde{F}_n(W(\cdot \wedge T_n) + \int_0^{\cdot\wedge T_n} v(s)ds)$ a.s. It thus follows that

$$-\log \mathbb{E}\exp\{-F_n(W)\} \geq \inf_{v\in\mathcal{A}_b} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F_n(W + \int_0^\cdot v(s)ds)\right]$$

a.s. Let $v_n \in \mathcal{A}_b$ be n^{-1}-optimal for the right side. Then we have

$$-\log \mathbb{E}\exp\{-F_n(W)\} \geq \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v_n(s)|^2 ds + F_n(W + \int_0^\cdot v_n(s)ds)\right] - n^{-1}.$$

The left side in the above display converges to $-\log \mathbb{E}\exp\{-F(W)\}$ as $n \to \infty$. Denoting by θ_n (resp. θ) the probability law of $W + \int_0^\cdot v_n(s)ds$ (resp. W) on $\mathcal{C}_0^\infty$, we have that, for every $n \in \mathbb{N}$,

$$R(\theta_n\|\theta) \leq \frac{1}{2}E\int_0^\infty |v_n(s)|^2 ds \leq 2\|F\|_\infty + 1.$$

This shows that (cf. [10, Lemma 2.5])

$$\lim_{n\to\infty}\left|F_n(W + \int_0^\cdot v_n(s)ds) - F(W + \int_0^\cdot v_n(s)ds)\right| = 0.$$

Combining the above

$$\begin{aligned}
-\log \mathbb{E}\exp\{-F(W)\} &= \lim_{n\to\infty} -\log \mathbb{E}\exp\{-F_n(W)\} \\
&\geq \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v_n(s)|^2 ds + F_n(W + \int_0^\cdot v_n(s)ds)\right] \\
&= \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v_n(s)|^2 ds + F(W + \int_0^\cdot v_n(s)ds)\right] \\
&\geq \inf_{v\mathcal{A}_b} \mathbb{E}\left[\frac{1}{2}\int_0^\infty |v(s)|^2 ds + F(W + \int_0^\cdot v(s)ds)\right].
\end{aligned}$$

This completes the proof of the lower bound and thus the result follows. □

Remark 4.4. *As in the proof of Theorem 2.2, the infimum in Theorem 4.1 can be taken over $\mathcal{A}_s$ or $\tilde{\mathcal{A}}_s$. Also, a similar representation can be written for an infinite dimensional Brownian motion, e.g. for a H-valued Wiener process as in Section 2.*

5 Functionals of Fractional Brownian Motion

A representation similar to (1.1) can be obtained for more general Gaussian processes than a Brownian motion. Fractional Brownian motions are an important class of Gaussian processes with long-range dependence that arise in many applications. In this section we present a variational formula for functionals of a fractional Brownian motion given in [7] that has been used for studying large deviation properties of small noise stochastic differental equations with a fractional Brownian motion. The variational formula relies on a general result from [22] on a variational representation for random functionals on abstract Wiener spaces. We begin by presenting this latter result and then describe how a representation for functionals of a fractional Brownian motion can be deduced from it.

5.1 Abstract Wiener Space Representation.

For simplicity we consider the time horizon $[0,1]$. The case of a general finite time horizon $[0,T]$ can be treated similarly. Let $(\mathbb{W}, \|\cdot\|_{\mathbb{W}})$ be a separable Banach space. and let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}}, \|\cdot\|_{\mathbb{H}})$ be a separable Hilbert space densely and continuously embedded in $\mathbb{W}$. Let μ be a centered, unit variance Gaussian measure over $\mathbb{W}$ (cf. [18]). Identifying in the usual manner the dual space $\mathbb{H}^*$ with itself, $\mathbb{W}^*$ may be viewed as a dense linear subspace of $\mathbb{H}$ and we have that for any $\ell \in \mathbb{W}^*$ and $h \in \mathbb{H}$, $\ell(h) = \langle \ell, h\rangle_{\mathbb{H}}$. Denoting the embedding map from $\mathbb{H}$ into $\mathbb{W}$ as $i_{\mathbb{H}}$, the collection $(i_{\mathbb{H}}, \mathbb{H}, \mathbb{W}, \mu)$ is referred to as an abstract Wiener space and $\mathbb{H}$ is called to Cameron-Martin space of this abstract Wiener space (cf. [14]).

We recall the notion of a filtration on an abstract Wiener space from [20, 21]. A collection of projection operators $\{\pi_t, t \in [0,1]\}$ on $\mathbb{H}$ are referred to as a continuous and strictly monotonic resolution of the identity in $\mathbb{H}$ if the following hold (i) $\pi_0 = 0$, $\pi_1 = \mathrm{Id}$, where Id is the identity operator, (ii) for $0 \le s < t \le 1$, $\pi_s\mathbb{H}$ is a strict subset of $\pi_t\mathbb{H}$, (iii) for any $h \in \mathbb{H}$ and $t \in [0,1]$, $\lim_{s\to t} \pi_s h = \pi_t h$.

By denseness of $\mathbb{W}^*$ in $\mathbb{H}$ we can find for every $h \in \mathbb{H}$ a sequence $\{h_n\} \subset \mathbb{W}^*$ such that $h_n \to h$. This says that the sequence $\{h_n\}$ regarded as a sequence of random variables on the probability space $(\mathbb{W}, \mathcal{B}(\mathbb{W}), \mu)$ is a Cauchy sequence in $L^2(\mu)$ and thus must converge to a limit $\delta(h)$ in $L^2(\mu)$. This limit is referred to as the Skorohod integral of h and we occasionally write $\delta(h)(w) = \langle h, w\rangle$ for $h \in \mathbb{H}$ and $w \in \mathbb{W}$.

Define the filtration $\{\mathcal{F}_t\}_{0\le t\le 1}$ as $\mathcal{F}_t \doteq \sigma\{\delta(\pi_t h), h \in \mathbb{H}\} \vee \mathcal{N}$, where $\mathcal{N}$ is the collection of all μ-null sets in $\mathcal{B}(\mathbb{W})$. We will regard this as a filtration on the μ-complete probability space $(\mathbb{W}, \mathcal{F}, \mu)$, where $\mathcal{F} \doteq \mathcal{F}_1$. Expecteation on this probability space will be denoted as $\mathbb{E}$.

An $\mathbb{H}$ valued random variable v on the above probability space is said to be $\mathcal{F}_t$-adapted if for every $t \in [0,1]$ and $h \in \mathbb{H}$, $\langle \pi_t h, v\rangle_{\mathbb{H}}$ is $\mathcal{F}_t$ measurable. The collection of adapted $\mathbb{H}$ valued square integrable random variables (i.e $\mathbb{E}\|v\|^2_{\mathbb{H}} < \infty$) is denoted by $\mathcal{H}^a$. For $N \in \mathbb{N}$, denote by S_N the ball of radius N in $\mathbb{H}$ and let $\mathcal{H}^a_{b,N}$ the subcollection of $\mathcal{H}^a$ consisting of v that take values in S_N. Let $\mathcal{H}^a_b \doteq \cup_{N>0}\mathcal{H}^a_{b,N}$. For $t \in [0,1]$, let $\mathcal{C}_t$ be the collection of cylindrical functions of the form

$$F(w) = g(\langle \pi_t h_1, w\rangle, \cdots, \langle \pi_t h_n, w\rangle),\ g \in C^\infty_b(\mathbb{R}^n),\ h_1, \dots h_n \in \mathbb{W}^*.$$

Here $C^\infty_b(\mathbb{R}^n)$ is the space of real infinitely differentiable functions on $\mathbb{R}^n$, with the function and all its derivatives bounded. Note that such a F is $\mathcal{F}_t$ measurable. A $v \in \mathcal{H}^a$ is said to be simple if it takes values in S_N for some $N < \infty$ and for some $0 = t_0 < t_1 < \cdots < t_n = 1$, $\{h_i\}_{0\le i\le n-1} \subset \mathbb{H}$, and $\xi_i \in \mathcal{C}_{t_i}$, $i = 0, 1, \dots n-1$,

$$v(w) = \sum_{i=0}^{n-1} \xi_i(w)(\pi_{t_{i+1}} - \pi_{t_i})h_i.$$

We denote the collection of all such simple v as $\mathcal{S}^a$. The following is the main representation from [22] (see Theorem 3.2 therein).

Theorem 5.1. *Let F be a real bounded and measurable function on $\mathbb{W}$. Then*

$$-\log \mathbb{E}(e^{-F}) = \inf_{v\in\mathcal{R}} \mathbb{E}\left(\frac{1}{2}\|v\|^2_{\mathbb{H}} + F(\cdot + v)\right)$$

where $\mathcal{R} \in \{\mathcal{H}^a, \mathcal{S}^a\}$.

5.2 Fractional Brownian Motion Representation.

We will now use the representation in Theorem 5.1 to provide a more explicit representation for a fractional Brownian motion. We begin by recalling some basic definitions. For $H \in (0,1)$, a d-dimensional fractional Brownian motion (fBm) $B^H = \{B^H_t : t \in [0,1]\}$ with Hurst parameter H defined on some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a centered Gaussian process whose covariance matrix $R_H = (R^{i,j}_H)_{1\le i,\, j\le d}$ is given by

$$R^{i,j}_H(s,t) = \mathbb{E}(B^{H,i}_s B^{H,j}_t) = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H})\delta_{i,j},\ s,\, t \in [0,1],$$

where δ is the Kronecker delta function. When $H = \frac{1}{2}$, the above process is simply a d-dimensional standard Brownian motion. From the above covariance formula it is immediate that

$$\mathbb{E}(|B_t^H - B_s^H|^2) = d|t-s|^{2H},\ t, s \in [0,1].$$

From the above property along with Kolmogorov's continuity criterion it follows that the sample paths of B^H are a.s. Hölder continuous of order β for all $\beta < H$.

In what follows $(\Omega, \mathcal{F}, \mathbb{P})$ will denote the canonical probability space, where $\Omega = C_0([0,1] : \mathbb{R}^d)$ is the space of continuous functions null at time 0, equipped with topology of uniform convergence, $\mathcal{F} = \mathcal{B}(C_0([0,1] : \mathbb{R}^d))$ is the Borel σ-algebra and $\mathbb{P}$ is the unique d-dimensional probability measure such that the canonical process $B^H = \{B_t^H(\omega) = \omega(t) : t \in [0,1]\}$ is a d-dimensional fractional Brownian motion with Hurst parameter H. Consider the canonical filtration given by $\{\mathcal{F}_t^H : t \in [0,1]\}$, where $\mathcal{F}_t^H = \sigma\{B_s^H : 0 \le s \le t\} \vee \mathcal{N}$ and $\mathcal{N}$ is the set of the $\mathbb{P}$-negligible events.

We now recall a representation for a fBM in terms of a standard Brownian motion and a suitable kernel function (cf. [12]). Let $F(a,b,c;z)$ denote the Gauss hypergeometric function defined for any a, b, c, $z \in \mathbb{C}$ with $|z| < 1$ and $c \neq 0, -1, -2, \ldots$ by

$$F(a,b,c;z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} z^k,$$

where $(a)_0 = 1$ and $(a)_k = a(a+1)\ldots(a+k-1)$ is the Pochhammer symbol.

Let $B = \{B_t = (B_t^1, \ldots, B_t^d),\ t \in [0,1]\}$ be a standard d-dimensional Brownian motion. Then it is well known (cf. [12]) that the process

$$B_t^H = \int_0^1 K_H(t,s) dB_s,\ t \in [0,1]$$

defines a fBm with Hurst parameter H, where

$$K_H(t,s) = k_H(t,s) \mathbf{1}_{[0,t]}(s),$$

for $0 \le s \le t$

$$k_H(t,s) = \frac{c_H}{\Gamma\left(H + \frac{1}{2}\right)} (t-s)^{H-\frac{1}{2}} F\left(H - \frac{1}{2}, \frac{1}{2} - H, H + \frac{1}{2}; 1 - \frac{t}{s}\right),$$

$c_H = \left[\frac{2H\Gamma\left(\frac{3}{2}-H\right)\Gamma\left(H+\frac{1}{2}\right)}{\Gamma(2-2H)}\right]^{1/2}$, and $\Gamma(\cdot)$ is the gamma function.

Define $\mathcal{H}_H = \{(K_H \dot{h}^1, \ldots, K_H \dot{h}^d) : \dot{h} = (\dot{h}^1, \ldots, \dot{h}^d) \in L^2([0,1] : \mathbb{R}^d)\}$, where

$$(K_H f)(t) \doteq \int_0^1 K_H(t,s) f(s) ds \text{ for } f \in L^2([0,1] : \mathbb{R}^d),\ t \in [0,1].$$

That is, any $h \in \mathcal{H}_H$ can be represented as

$$h(t) = (K_H \dot{h})(t) = \int_0^1 K_H(t,s) \dot{h}(s) ds,$$

for some $\dot{h} \in L^2([0,1] : \mathbb{R}^d)$. Define a scalar inner product on $\mathcal{H}_H$ by

$$\langle h, g \rangle_{\mathcal{H}_H} = \langle K_H \dot{h}, K_H \dot{g} \rangle_{\mathcal{H}_H} = \langle \dot{h}, \dot{g} \rangle_{L^2}.$$

Then $\mathcal{H}_H$ is a separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_H}$. It can be checked (cf. [7, Remark 1]) that $\mathcal{H}_H$ is a subset of $\Omega = C_0([0,1] : \mathbb{R}^d)$ and the embedding map is continuous. We now present the main variational representation for functionals of fBm from [7]. For $0 < N < \infty$, let $S_N = \{v \in \mathcal{H}_H : \frac{1}{2}\|v\|_{\mathcal{H}_H}^2 \le N\}$.

Let $\mathcal{A}$ be the collection of all measurable $v : \Omega \to \mathcal{H}_H$ such that there is a $\mathcal{F}_t^H$-predictable $\dot{v} : [0,T] \times \Omega \to \mathbb{R}$ such that $\mathbb{E} \int_0^1 \|\dot{v}(t)\|^2 dt < \infty$ and $v\omega) = K_H \dot{v}(\omega)$ a.s. The subcollection of $\mathcal{A}$ consisting of v that take values in S_N will be denoted as $\mathcal{A}_{b,N}$ and, as before, $\mathcal{A}_b = \cup_{N>0} \mathcal{A}_{b,N}$. The following is [7, Proposition 2].

Theorem 5.2. *Let F be a real bounded measurable function on Ω. Then we have*

$$-\log \mathbb{E}(e^{-F(B^H)}) = \inf_{v\in\mathcal{R}} \mathbb{E}\left(f(B^H+v) + \frac{1}{2}\|v\|^2_{\mathcal{H}_H}\right),$$

where $\mathcal{R} \in \{\mathcal{A}, \mathcal{A}_b\}$.

Remark 5.3. *Although not discussed here, one can also replace the above infimum by the infimum over the subclass of $\mathcal{A}_b$ consisting of all simple processes.*

We provide a brief proof sketch of this result. We refer the reader to [7] for details. The basic idea is to identify a suitable abstract Wiener space and then apply Theorem 5.1.

It is easy to verify that functions in $\mathcal{H}_H$ are H-Hölder continuous (cf. [7, Lemma 2]), in particular we have the following fact: Any $h \in \mathcal{H}_H$ is in $C^H([0,1]:\mathbb{R}^d)$, and $\|h\|_\infty \le \|h\|_{\mathcal{H}_H}$ and $\|h\|_H \le \|h\|_{\mathcal{H}_H}$.

Next, recall that $\Omega = C_0([0,1]:\mathbb{R}^d)$ is a Banach space equipped with the sup-norm $\|\cdot\|_\infty$. Let Ω^* be its topological dual. We now introduce the abstract Wiener space associated with a fractional Brownian motion. The following result is taken from [12, Theorem 3.3].

Lemma 5.4. *If we identify $L^2([0,1]:\mathbb{R}^d)$ and its dual, we have the following diagram*

$$\Omega^* \xrightarrow{i_H^*} \mathcal{H}_H^* \xrightarrow{K_H^*} L^2([0,1]:\mathbb{R}^d) \xrightarrow{K_H} \mathcal{H}_H \xrightarrow{i_H} \Omega$$

Where i_H is the injection from $\mathcal{H}_H$ into Ω, and K_H^* and i_H^* are the respective adjoints.

(a) The injection i_H embeds $\mathcal{H}_H$ densely into Ω, and $\mathcal{H}_H$ is the Cameron-Martin space of the abstract Wiener space $(i_H, \mathcal{H}_H, \Omega, \mathbb{P})$ in the sense of Gross [14].

(b) The restriction of K_H^* to Ω^* can be represented by

$$(K_H^*\eta)(s) = \int_0^1 K_H(t,s)\eta(dt) = \left(\int_0^1 K_H(t,s)\eta_1(dt), \ldots, \int_0^1 K_H(t,s)\eta_d(dt)\right),$$

for any $\eta = (\eta_1, \ldots, \eta_d) \in \Omega^*$.

As before, denote for $h \in \mathcal{H}_H$ its Skorohod integral by $\delta(h)$. Recall the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t^H\})$ introduced above and recall that $\{B_t^H\}$ is the canonical coordinate process on $(\Omega, \mathcal{F})$.

Define the family $\{\pi_t^H,\, t \in [0,1]\}$ of projection operators in $\mathcal{H}_H$ by

$$\pi_t^H h = \pi_t^H(K_H\dot{h}) = K_H(\dot{h}\mathbf{1}_{[0,t]}),\ h \in \mathcal{H}_H. \tag{5.1}$$

The following result is a consequence of [12, Proposition 4.4, Theorems 4.3 and 4.8].

Lemma 5.5. *(a) For any $t \in [0,1]$, $\mathcal{F}_t^H = \sigma\{\delta(\pi_t^H h),\, h \in \mathcal{H}_H\} \vee \mathcal{N}$.*

(b) For any $\mathcal{H}_H$-valued $\{\mathcal{F}_t^H,\, t \in [0,1]\}$-adapted stochastic process u there is a $\{\mathcal{F}_t^H,\, t \in [0,1]\}$-predictable process $\dot{u} : [0,T] \times \Omega \to \mathbb{R}^d$ such that $u = K_H\dot{u}$ a.s.

One can now complete the proof of Theorem 5.2 as follows. The collection $\{\pi_t^H,\, t \in [0,1]\}$ introduced in (5.1) defines a continuous and strictly monotonic resolution of the identity in $\mathcal{H}_H$ in the sense described previously in the section. Also, using Lemma 5.5, the classes $\mathcal{A}$ and $\mathcal{A}_b$ in this section are the same as the classes $\mathcal{H}^a$ and $\mathcal{H}_b^a$ introduced earlier in the section, when specialized to the abstract Wiener space for the setting of a fBM. The result is now immediate from Theorem 5.1. □

6 Lévy Noise.

In this section we present the analogues of the Boué-Dupuis variational formulas for Lévy processes. The results from this section are from [6]. They can also be found in [10, Section 8.2]. The statement in Theorems 6.1 and 6.2 that $\mathcal{R}$ can be $\mathcal{A}_s$ or $\tilde{\mathcal{A}}_s$, although does not appear in these references, it can be easily deduced from the arguments there along the lines of the proof of Theorem 2.2. We omit the details.

A Lévy noise is composed of a two random components, one consisting of a finite or an infinite dimensional Brownian motion and the other described through a suitable Poisson random measure(PRM). We begin by first presenting a representation of functionals of a PRM and then we will give the representation for the more general case where the functional depends on both a PRM and an infinite dimensional Brownian motion.

6.1 Representation for a Poisson Random Measure.

Let $\mathcal{X}$ be a locally compact Polish space and let $\Sigma(\mathcal{X})$ denote the space of all locally-finite measures ν on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. The space $\Sigma(\mathcal{X})$ is equipped with the standard vague convergence topology. Fix $T \in (0, \infty)$ and let $\mathcal{X}_T = [0, T] \times \mathcal{X}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{0 \le t \le T}$ satisfying the usual assumptions. Fix $\nu \in \Sigma(\mathcal{X})$ and let $\nu_T \doteq \lambda_T \times \nu$ where λ_T is the Lebesgue measure on $[0, T]$.

Let N be a $\mathcal{F}_t$- Poisson random measure with points in $\mathcal{X}_T$ and intensity measure ν_T. Recall this means that N is a $\Sigma(\mathcal{X}_T)$ valued random variable with the following properties (i) for every $t \in [0, T]$ and $A \in \mathcal{B}([0, t] \times \mathcal{X})$, $N(A)$ is $\mathcal{F}_t$-measurable, (ii) for every $t \in [0, T]$ and $A \in \mathcal{B}((t, T] \times \mathcal{X})$, $N(A)$ is independent of $\mathcal{F}_t$, (iii) for any $k \in \mathbb{N}$ and $A_1, \dots A_k \in \mathcal{B}(\mathcal{X}_T)$ such that $A_i \cap A_j = \emptyset$ when $i \neq j$, and $\nu_T(\cup_{i=1}^k A_i) < \infty$, we have $N(A_1), \dots N(A_k)$ are mutually independent Poisson random variables with parameters $\nu_T(A_1), \dots \nu_T(A_k)$, respectively.

Fix $\theta > 0$. Letting $\mathbb{M} \doteq \Sigma(\mathcal{X})$, we will denote by $\mathbb{P}_\theta$ the unique probability measure on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ under which the the canonical map $N : \mathbb{M} \to \mathbb{M}$ defined as $N(m) = m$ is PRM with intensity $\theta \nu_T$. In typical large deviation problems of interest θ plays the role of scaling parameter and one is interested in behavior of the system as θ approaches some limiting value (e.g. ∞ or 0). In such problems one is interested in characterizing the asymptotic behavior of Laplace functionals of the form $-\log \mathbb{E}_\theta \exp\{-F(N)\}$, where $F : \mathbb{M} \to \mathbb{R}$ is some bounded measurable map.

We will now present the variational representation from [6] that, together with weak convergence methods, allows one to study such asymptotic behavior in many problem settings (cf. references in [10]). In formulating this variational representation it will be convenient to work with a PRM that is defined on a larger point space.

Let $\mathcal{Y} \doteq \mathcal{X} \times [0, \infty)$ and $\mathcal{Y}_T \doteq [0, T] \times \mathcal{Y}$. Let $\bar{\mathbb{M}}$ denote the space of locally finite measures on $\mathcal{Y}_T$. and let $\bar{P}$ be the unique measure on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$ such that the canonical map $\bar{N} : \bar{\mathbb{M}} \to \bar{\mathbb{M}}$ defined as $\bar{N}(m) \doteq m$, $m \in \bar{\mathbb{M}}$, is a PRM with points in $\mathcal{Y}_T$ and intensity measure $\bar{\nu}_t \doteq \lambda_T \times \nu \times \lambda_\infty$, where λ_∞ is the Lebesgue measure on $[0, \infty)$.

Let $\mathcal{G}_t$ denote the augmentation of $\sigma\{\bar{N}((0, s] \times A) : 0 \le s \le t, A \in \mathcal{B}(\mathcal{Y})\}$ with all $\bar{\mathbb{P}}$ null sets in $\mathcal{B}(\bar{\mathbb{M}})$ and denote by $\mathcal{PF}$ the $\mathcal{G}_t$-predictable σ-field on $[0, T] \times \bar{\mathbb{M}}$.

Recall that the representation for functionals of Brownian motions, given e.g. in (1.1), involves L^2-controls v and controlled noise processes $W + \int_0^\cdot v(s)ds$. We now introduce the controls and controlled noises that will play the analogous role in the variational representations for PRM. First the space of controls, which we denote by $\mathcal{A}$ will be the collection of all maps $v : [0, T] \times \bar{\mathbb{M}} \times \mathcal{X}$ that are $(\mathcal{PF} \otimes \mathcal{B}(\mathcal{X})) \backslash \mathcal{B}([0, \infty))$ measurable. For notational ease, for $(t, \omega, x) \in [0, T] \times \bar{\mathbb{M}} \times \mathcal{X}$, we will occasionally suppress ω and write $v(t, \omega, x)$ as $v(t, x)$. For $v \in \mathcal{A}$, define a counting process N^v on $\mathcal{X}_T$ as

$$N^v((0, t] \times U) \doteq \int_{(0,t] \times U} \int_{(0,\infty)} \mathbf{1}_{[0, v(s,x)]}(r) \bar{N}(ds \times dx \times dr)$$

for all $t \in [0, T]$ and $U \in \mathcal{B}(\mathcal{X})$. When, for some $\theta > 0$, $v(s, \omega, x) = \theta$ for all $(s, \omega, x) \in \mathcal{X}_T \times \bar{\mathbb{M}}$, we will write N^v as simply N^θ. The process N^v can be thought of as the controlled noise process associated with the control v.

Next, recall that the Brownian motion representation involved a quadratic cost term given as a suitable L^2-norm. In the representation for a PRM the quadratic function is replaced by the

following superlinear (but sub-quadratic) function $\ell : [0,\infty) \to [0,\infty)$ defined as

$$\ell(r) \doteq r\log r - r + 1,\ r \in [0,\infty).$$

For $v \in \mathcal{A}$, define

$$L_T(v) \doteq \int_{\mathcal{X}_T} \ell(v(t,x))\nu_T(dt \times dx).$$

This is the cost term associated with a control v that will replace the quadratic cost term we see in Brownian motion representation formulas, for the case of a PRM.

In the case of the Brownian motion, it was useful to consider the smaller collection of controls than $\mathcal{A}$, namely $\mathcal{A}_b$, which consisted of controls for which one can invoke the Girsanov theorem. We will consider an analogous smaller collection of controls here as well. Let $\{K_n\}_{n\in\mathbb{N}}$ be an increasing sequence of compact sets of $\mathcal{X}$ such that $\cup_{n\geq 1} K_n = \mathcal{X}$. For $M \in (0,\infty)$, let

$$\begin{aligned}\mathcal{A}_{b,M} \doteq \{v \in \mathcal{A} : L_T(v) \leq M,\ &\text{for some } n \in \mathbb{N}, v(t,x) \in [1/n, n],\\ &\text{and } v(t,\omega,x) = 1 \text{ if } x \in K_n^c,\ \text{for all } (t,\omega) \in [0,T] \times \bar{\mathbb{M}}\}.\end{aligned}$$

A process in $\mathcal{A}_{b,M}$ is called simple if the following holds: There exist $n, l, n_1, \ldots n_l \in \mathbb{N}$; a partition $0 = t_0 < t_1 < \cdots < t_l = T$; for each $i = 1, \ldots, l$ a disjoint measurable partition$\{E_{ij}\}_{1\leq j\leq n_i}$ of K_n; $\mathcal{G}_{t_{i-1}}$- measurable random variables X_{ij}, $1 \leq i \leq l$, $1 \leq j \leq n_i$, such that $X_{ij} \in [1/n, n]$; and for $(t, \bar{m}, x) \in [0,T] \times \Omega \times \mathcal{X}$

$$v(t,\bar{m},x) = \mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{l}\sum_{j=1}^{n_i} \mathbf{1}_{(t_{i-1},t_i]}(t) X_{ij}(\bar{m}) \mathbf{1}_{E_{ij}}(x) + \mathbf{1}_{K_n^c}(x) 1_{(0,T)}(t).$$

We denote the collection all such simple processes as $\mathcal{A}_{s,M}$. Define $\mathcal{A}_b \doteq \cup_{M>0}\mathcal{A}_{b,M}$ and $\mathcal{A}_s \doteq \cup_{M>0}\mathcal{A}_{s,M}$. As in the Brownian motion case, when the canonical filtration $\{\mathcal{G}_t\}$ is replaced by a larger filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ under which $\bar{N}$ is a $\mathcal{F}_t$-PRM with the same intensity (in such a case of course $\bar{N}$ is defined on a larger probability space than the canonical space $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$), we will denote the classes analogous to $\mathcal{A}$, $\mathcal{A}_b$ and $\mathcal{A}_s$ as $\tilde{\mathcal{A}}$, $\tilde{\mathcal{A}}_b$, and $\tilde{\mathcal{A}}_s$, respectively. The following result is from [6] (see also [10, Theorem 8.12]).

Theorem 6.1. *Let $F : \mathbb{M} \to \mathbb{R}$ be a bounded measurable map. Then for any $\theta > 0$*

$$-\log E_\theta \exp\{-F(N)\} = \inf_{v\in\mathcal{R}} \bar{\mathbb{E}}[\theta L_T(v) + G(N^{\theta v})],$$

where $\mathcal{R} \in \{\mathcal{A}, \tilde{\mathcal{A}}, \mathcal{A}_b, \tilde{\mathcal{A}}_b, \mathcal{A}_s, \tilde{\mathcal{A}}_s\}$.

6.2 General Lévy Process Representation.

Finally, we present the representation for a general Lévy noise. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ satisfying the usual conditions and assume that this probability space supports the processes introduced below. With Q as in Section 2, let W be a a Q-Wiener process with respect to the filtration $\mathcal{F}_t$. Let $\nu, \mathcal{X}, \mathcal{X}_T, \mathcal{Y}_T, \bar{\nu}_T$ be as introduced above. Let $\bar{N}$ be a $\mathcal{F}_t$-PRM with points in $\mathcal{Y}_T$ with intensity measure ν_T. Assume that for all $0 \leq s \leq t < \infty$, $(\bar{N}((s,t] \times \cdot)), W(t) - W(s))$ is independent of $\mathcal{F}_s$. Let $\mathcal{PF}$ be the $\mathcal{F}_t$-predictable σ-field on $[0,T] \times \Omega$. Denote by $\tilde{\mathcal{A}}^W$ and $\tilde{\mathcal{A}}_b^W$ the collections $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{A}}_b$ introduced below Defintiion 2.1. Also, denote by $\tilde{\mathcal{A}}^N$ and $\tilde{\mathcal{A}}_b^N$ the collections $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{A}}_b$ introduced earlier in the current section. Let $\tilde{\mathcal{A}}_b \doteq \tilde{\mathcal{A}}_b^W \times \tilde{\mathcal{A}}_b^N$ and $\tilde{\mathcal{A}} \doteq \tilde{\mathcal{A}}^W \times \tilde{\mathcal{A}}^N$. For $v = (\psi, \varphi) \in \tilde{\mathcal{A}}$, let $L_T^W(\psi) \doteq \frac{1}{2}\int_0^T \|\psi(s)\|_0^2 ds$, where the norm $\|\cdot\|_0$ is as introduced in Section 2, and let $L_T^N(\varphi) \doteq \int_{\mathcal{X}_T} \ell(\varphi(t,x))\nu_T(dt \times dx)$. Also, set $L_T(v) \doteq L_T^W(\psi) + L_T^N(\varphi)$. Next, for $\psi \in \tilde{\mathcal{A}}^W$, let $W^\psi(t) \doteq W(t) + \int_0^t \psi(s)ds$, $t \in [0,T]$.

The following is [10, Theorem 8.19].

Theorem 6.2. *Let $F : \mathcal{C}_0^T(H) \times \mathbb{M} \to \mathbb{R}$ be a bounded measurable map. Then for $\theta \in (0,\infty)$*

$$-\log \mathbb{E}\exp\{-F(W, N^\theta)\} = \inf_{v=(\psi,\varphi)\in\mathcal{R}} E\left[\theta L^T(v) + F(W^{\sqrt{\theta}}, N^{\theta\psi})\right],$$

where $\mathcal{R} \in \{\tilde{\mathcal{A}}_b, \tilde{\mathcal{A}}\}$.

Remark 6.3. *In a similar manner as in Section 3 one can give a variational representation for functionals that, in addition to depending on a Brownian motion and a PRM, also depend on another independent random variable. Also, in a manner similar to Section 4 one can obtain a variational representation for functionals that depend on an infinite path of a Brownian motion and a PRM, namely a function of* $\{(W(t), N^\theta((0,,t] \times \cdot)), 0 \leq t < \infty\}$. *We omit the details.*

Acknowledgements Research supported in part by the NSF (DMS-2134107 and DMS-2152577).

References

[1] S. Banerjee, A. Budhiraja, and M. Perlmutter, *A new approach to large deviations for the Ginzburg-Landau model*, Electronic Journal of Probability **25** (2020).

[2] M. Boué and P. Dupuis, *A variational representation for certain functionals of Brownian motion*, The Annals of Prob. **26** (1998), 1641–1659.

[3] A. Budhiraja and P. Dupuis, *A variational representation for positive functionals of infinite dimensional Brownian motion*, Prob. Math. Statist. **20** (2000), 39–61.

[4] A. Budhiraja, P. Dupuis, and M. Fischer, *Large deviation properties of weakly interacting processes via weak convergence methods*, Ann. Probab. **40** (2012), 74–1–2.

[5] A. Budhiraja, P. Dupuis, and V. Maroulas, *Large deviations for infinite dimensional stochastic dynamical systems*, Ann. Probab. **36** (2008), 1390–1420.

[6] A. Budhiraja, P. Dupuis, and V. Maroulas, *Variational representations for continuous time processes*, Ann. de l'Inst. H. Poincaré **47** (2011), 725–747.

[7] A. Budhiraja and X. Song, *Large deviation principles for functionals of fractional Brownian motions*, Frontiers of Statistics and Data Science, eds. Subhashis Ghoshal and Anindya Roy, Springer, New York, NY, 2023.

[8] A. Budhiraja and M. Conroy, *Asymptotic behavior of stochastic currents under large deviation scaling with mean field interaction and vanishing noise*, Annali Scuola Normale Superiore-Classe Di Scienze (2022).

[9] A. Budhiraja and M. Conroy, *Empirical measure and small noise asymptotics under large deviation scaling for interacting diffusions*, Journal of Theoretical Probability (2022), 1–55.

[10] A. Budhiraja and P. Dupuis, *Analysis and Approximation of Rare Events*, Representations and Weak Convergence Methods. Series Prob. Theory and Stoch. Modelling **94** (2019).

[11] A. Budhiraja, W.-T. Fan, and R. Wu, *Large deviations for Brownian particle systems with killing*, Journal of Theoretical Probability **31** (2018), no. 3, 1779–1818.

[12] L. Decreusefond and A. S. Üstünel, *Stochastic analysis of the fractional Brownian motion*, Potential analysis **10** (1999), 177–214.

[13] P. Dupuis and R. S. Ellis., *A Weak Convergence Approach to the Theory of Large Deviations*, John Wiley & Sons, New York, 1997.

[14] L. Gross, *Abstract Wiener Spaces*, Proceedings of the fifth berkeley symposium on mathematical statistics and probability, volume 2: Contributions to probability theory, part 1, 1967, pp. 31–43.

[15] M. Hairer, *Ergodicity of stochastic differential equations driven by fractional Brownian motion*, Ann. Prob. (2005), 703–758.

[16] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam, 1981.

[17] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

[18] H.-H. Kuo, *Gaussian Measures in Banach Spaces*.

[19] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Vol. 293, Springer Science & Business Media, 2013.

[20] A. S. Üstünel and M. Zakai, *Transformation of Measure on Wiener Space*, Springer Science & Business Media, 2013.

[21] A. S. Üstünel and M. Zakai, *The construction of filtrations on abstract Wiener space*, Journal of Functional Analysis **143** (1997), no. 1, 10–32.

[22] X. Zhang, *A variational representation for random functionals on abstract Wiener spaces*, Journal of Mathematics of Kyoto University **49** (2009), no. 3, 475–490.

A. Budhiraja
Department of Statistics and Operations Research, and
School of Data Science and Society
University of North Carolina
Chapel Hill, NC 27599, USA
email: budhiraj@email.unc.edu