PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 20, Fasc. 1 (2000), pp. 39–61

# A VARIATIONAL REPRESENTATION FOR POSITIVE FUNCTIONALS OF INFINITE DIMENSIONAL BROWNIAN MOTION

BY

AMARJIT BUDHIRAJA* (NOTRE DAME, IN) AND PAUL DUPUIS** (PROVIDENCE, RI)

*Abstract.* A variational representation for positive functionals of a Hilbert space valued Wiener process $(W(\cdot))$ is proved. This representation is then used to prove a large deviations principle for the family $\{\mathcal{G}^\varepsilon(W(\cdot))\}_{\varepsilon>0}$, where $\mathcal{G}^\varepsilon$ is an appropriate family of measurable maps from the Wiener space to some Polish space.

**Key words and phrases:** Large deviations, Laplace principle, stochastic control, cylindrical Brownian motion, stochastic evolution equations, infinite dimensional stochastic calculus.

**1. Introduction.** The theory of large deviations is one of the classical areas in probability and statistics (see for example [23], [7], [6], [14], [11]). The book [10] develops an approach to this topic that is based on proving the convergence of solutions to variational problems. The starting point for this approach is the fact that the large deviation principle (LDP) is equivalent to what is called a Laplace principle (see Definition 4.2 below) if the underlying space is Polish. This is a consequence of Varadhan's lemma [24] and Bryc's converse to Varadhan's lemma [2]. We refer the reader to [10] for the elementary proof. A key step in the approach is the representation of the pre-limit normalized expectations in the statement of the Laplace principle by value functions (minimal cost functions) of certain stochastic optimal control problems. The large deviation problem then reduces to verifying the convergence of these value functions and identifying the limits. This latter problem is well suited to the application of weak convergence methods.

The prototype of the representation is the following ([10], Proposition 1.4.2). Let $(\mathcal{V}, \mathcal{A})$ be a measurable space, $k$ a bounded measurable function

* This research was supported in part by the National Science Foundation (NSF-DMI-9812857) and the University of Notre Dame Faculty Research Program.

** This research was supported in part by the National Science Foundation (NSF-DMS-9704426) and the Army Research Office (DAAD19-99-1-0223).

mapping $\mathscr{V}$ into $\boldsymbol{R}$, and $\theta$ a probability measure on $\mathscr{V}$. Then

$$-\log \int_{\mathscr{V}} e^{-k}\, d\theta = \inf_{\gamma \in \mathscr{P}(\mathscr{V})} \{R(\gamma \,\|\, \theta) + \int_{\mathscr{V}} k d\gamma\}, \tag{1.1}$$

where $\mathscr{P}(\mathscr{V})$ is the space of all probability measures on $(\mathscr{V}, \mathscr{A})$, and $R(\cdot\,\|\,\cdot)$ denotes the relative entropy function (see Section 3 for the definition of relative entropy). For many interesting examples the right-hand side of the expression above can be written as the value function of an appropriate stochastic control problem (cf. [10] and [1]). For example, if $\mathscr{V}$ is $\mathscr{C}([0, T]: \boldsymbol{R}^n)$ and $\theta$ is the Wiener measure, then it is proved in [1] that

$$-\log \int_{\mathscr{V}} e^{-k}\, d\theta = \inf_{v \in \mathscr{A}} \int_{\mathscr{V}} \left( \frac{1}{2} \int_0^T \|v(s)\|^2 + f\big(W + \int_0^{\cdot} v(s)\, ds\big) \right) d\theta, \tag{1.2}$$

where $\mathscr{A}$ is the space of square integrable predictable (with respect to the Wiener filtration) processes.

Our main interest in the present paper is the study of large deviations for infinite dimensional stochastic differential equations (SDEs). Such equations arise in a wide range of applications (see [5], [25], [15], [17]). The problem of proving Wentzell–Friedlin type large deviation estimates for such SDEs has been studied by a number of authors, including [13], [5], [4], [21], [22], [17]. When the diffusion coefficient is constant, the proofs in these papers basically follow from the contraction principle. In the general case where the diffusion coefficient is not constant and the contraction principle cannot be applied, discretization arguments as in the original work of Wentzell and Friedlin are used. A feature that is common to all the different models considered is their representation as a dynamical system driven by some type of infinite dimensional Brownian motion. In this paper we will use the stochastic control and weak convergence approach to obtain the LDP for the family $\{\mathscr{G}^\varepsilon(W(\cdot))\}_{\varepsilon>0}$, where $\mathscr{G}^\varepsilon$ is an appropriate family of measurable maps from the Wiener space to some Polish space and $W(\cdot)$ is a Hilbert space valued Wiener process. This is done in Theorem 4.4. The key assumption on the family $\{\mathscr{G}^\varepsilon\}$ is Assumption 4.3. Assumption 4.3 (ii) essentially says that the level sets of the rate function are compact. Assumption 4.3 (i) is the crucial condition that needs to be verified in various applications of this result and is a statement on the weak convergence of a certain family of random variables. This condition is at the core of the weak convergence approach to the study of large deviations. Using the above result we are able to obtain Wentzell–Freidlin type large deviation results for a wide class of stochastic dynamical systems driven by a small noise, infinite dimensional Wiener process. We refer the reader to [3], where Hilbert space valued small noise diffusions with quite general coefficients and stochastic evolution equations with a multiplicative noise are studied in detail. In fact, the conditions imposed on the coefficients are precisely the ones that are required for the existence of a unique strong (resp. mild) solution. The proofs of these large deviation results, which are essentially based on the verification of Assumption 4.3, are quite different from the proofs in [13], [5], [4], [21], [22], [17]. Furthermore, the approach taken in this paper gives a unified method for studying large deviations for a wide range of stochastic dynamical systems driven by an infinite dimensional Brownian motion.

The crucial step in the proof of the LDP mentioned above is a variational representation for positive functionals of an infinite dimensional Brownian motion, proved in Theorem 3.6 (see equation (3.14)). It may be worth observing that in our representation we allow the class $\mathscr{A}$ to consist of processes predictable with respect to a larger filtration than that generated by the Wiener process. This relaxation is of importance in some control applications. The starting point of the proof of the representation is (1.1). One of the main issues can be described as follows. Suppose that $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, \theta)$ is a probability space with a filtration satisfying the usual hypothesis, $H$ is a separable Hilbert space, and $(W_t, \mathscr{F}_t)$ is an $H$-valued Wiener process (to be described precisely in Section 2) on $\Omega$. Let $\gamma \in \mathscr{P}(\mathscr{V})$ be such that

$$\frac{d\gamma}{d\theta} = \exp\left\{\int_0^T \psi(s)\, dW(s) - \frac{1}{2}\int_0^T \|\psi(s)\|_0^2\, ds\right\}$$

for an appropriate predictable process $\psi(\cdot)$. Then the expression on the right-hand side of (1.1), i.e.,

$$R(\gamma \,\|\, \theta) + \int_\Omega k\, d\gamma,$$

equals

$$E^\gamma\left[\frac{1}{2}\int_0^T \|\psi(s)\|_0^2\, ds + k\Big(W(\cdot) + \int_0^{\cdot} \psi(s)\, ds\Big)\right],$$

where $E^\gamma$ denotes the expectation on the space $\mathscr{V}$ with respect to the original probability measure $\gamma$. Thus, roughly speaking, in order to obtain the desired representation we need to replace the expectation with respect to the original probability measure $\gamma$ with the expectation with respect to the probability measure $\theta$. This key step is undertaken in Lemma 3.5.

The paper is organized as follows. In Section 2 we recall some facts about Hilbert space valued Brownian motions and weak convergence criteria for probability measures on Hilbert spaces. Section 3 is devoted to the proof of our main representation theorem. In Section 4 we formulate and prove the general large deviation result for the family $\{\mathscr{G}^\varepsilon(W(\cdot))\}$.

**2. Preliminaries.** Let $(\Omega, \mathscr{F}, \theta)$ be a probability space with an increasing family of right continuous $\theta$-complete sigma fields $\{\mathscr{F}_t\}_{0 \leqslant t \leqslant T}$. We begin with the definition of a Hilbert space valued Wiener process. Let $(H, \langle \cdot, \cdot \rangle)$ be a real separable Hilbert space. Let $Q$ be a strictly positive, symmetric, trace class operator (cf. [9]) on $H$.

DEFINITION 2.1. An $H$-valued stochastic process $\{W(t), 0 \leqslant t \leqslant T\}$ is called a *Q-Wiener process* with respect to $\{\mathscr{F}_t\}$ if the following conditions hold:

1. For every non-zero $h \in H$, $\langle Qh, h\rangle^{-1/2} \langle W(t), h\rangle$ is a one-dimensional standard Wiener process.

2. For every $h \in H$, $W(t, h) \doteq \langle W(t), h\rangle$ is an $\mathscr{F}_t$-martingale.

Define $H_0 \doteq Q^{1/2} H$. Clearly, $H_0$ is a Hilbert space with the inner product

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

Denote the norms in $H$ and $H_0$ by $\|\cdot\|$ and $\|\cdot\|_0$, respectively. Since $Q$ is a trace class operator, the identity mapping from $H_0$ to $H$ is Hilbert–Schmidt. This Hilbert–Schmidt embedding of $H_0$ in $H$ will play a central role in many of the arguments to follow. One consequence of the embedding is that if $v^{(n)}$ is a sequence in $H_0$ such that $v^{(n)} \to 0$ weakly in $H_0$, then $\|v^{(n)}\| \to 0$. For an exposition of stochastic calculus with respect to an $H$-valued Wiener process we refer the reader to [5]. Other useful references are [19], [20], and [17].

The following two theorems are crucial ingredients to the proofs in this paper. Although the first theorem is standard, the second requires some elementary modifications of standard arguments. A sketch of the proof is provided for the sake of completeness.

Let $\{\mathscr{G}_t\}_{0 \leqslant t \leqslant T}$ be the $\theta$-completion of the filtration generated by $\{W(s)\colon 0 \leqslant s \leqslant t\}_{0 \leqslant t \leqslant T}$. We denote the space of square integrable random variables on $(\Omega, \mathscr{F}, \theta)$ by $L^2(\theta)$ and the subspace of random variables which are $\mathscr{G}_T$-measurable by $L^2_W(\theta)$. Also, define $\mathscr{A}$ to be the class of $H_0$-valued $\mathscr{F}_t$-predictable processes $\phi$ that satisfy

$$\theta\Big\{\int_0^T \|\phi(s)\|_0^2\, ds < \infty\Big\} = 1. \tag{2.1}$$

Finally, let

$$\mathscr{A}^W \doteq \{\phi \in \mathscr{A}\colon \phi \text{ is } \mathscr{G}_t\text{-predictable}\}.$$

We refer the reader to Chapter 4 of [5] for the definition of stochastic integrals of elements of $\mathscr{A}$ with respect to $W$.

THEOREM 2.2. *Let $\psi \in \mathscr{A}$ be such that*

$$E\left(\exp\left\{\int_0^T \psi(s)\, dW(s) - \frac{1}{2}\int_0^T \|\psi(s)\|_0^2\, ds\right\}\right) = 1.$$

*Then the process*

$$\tilde{W}(t) \doteq W(t) - \int_0^t \psi(s)\,ds, \qquad t \in [0, T],$$

*is a Q-Wiener process with respect to* $\{\mathscr{F}_t\}$ *on* $(\Omega, \mathscr{F}, \gamma)$, *where* $\gamma$ *is the probability measure defined by*

$$\frac{d\gamma}{d\theta} = \exp\left\{\int_0^T \psi(s)\,dW(s) - \frac{1}{2}\int_0^T \|\psi(s)\|_0^2\,ds\right\}.$$

For the proof see Theorem 10.14 of [5].

THEOREM 2.3. *Let* $(M(t), \mathscr{G}_t)$ *be a real-valued local martingale with right continuous paths having left limits. Then there exists* $\phi \in \mathscr{A}^W$ *such that for all* $0 \leqslant t \leqslant T$

$$M(t) = M(0) + \int_0^t \phi(s)\,dW(s) \ \ a.s.$$

Proof. The proof is adapted from [16]. We consider only the case where $M(t)$ is a mean zero square integrable martingale. The general statement in the theorem follows by the usual localization arguments (cf. Problem 3.4.16 of [18]). Let $L^2([0, T]: H_0)$ denote the class of all measurable maps $\eta$: $[0, T] \to H_0$ for which $\int_0^T \|\eta(s)\|_0^2\,ds$ is finite. For $0 \leqslant t \leqslant T$ and $\eta \in L^2([0, T]: H_0)$ define

$$\beta^{(\eta)}(t) \doteq \exp\left\{\int_0^t \eta(s)\,dW(s) - \frac{1}{2}\int_0^t \|\eta(s)\|_0^2\,ds\right\}.$$

Applying Itô's formula (cf. Theorem 4.17 of [5]) we have

$$\beta^{(\eta)}(t) = 1 + \int_0^t \beta^{(\eta)}(s)\,\eta(s)\,dW(s) \qquad \text{for all } 0 \leqslant t \leqslant T. \tag{2.2}$$

Since $(2\int_0^t \eta(s)\,dW(s), \mathscr{G}_t)$ is a real-valued martingale with quadratic variation process $\int_0^t \|2\eta(s)\|_0^2\,ds$, it follows that $\beta^{(2\eta)}$ is a non-negative local martingale, and hence a supermartingale. Observing that

$$(\beta^{(\eta)}(t))^2 = \beta^{(2\eta)}(t)\exp\left(\int_0^t \|\eta(s)\|_0^2\,ds\right),$$

we have

$$\sup_{0 \leqslant t \leqslant T} E(\beta^{(\eta)}(t))^2 \leqslant \exp\left(\int_0^T \|\eta(s)\|_0^2\,ds\right) < \infty.$$

This implies that $\int_0^t \eta(s)\,\beta^{(\eta)}(s)\,dW(s)$ is a square integrable $\mathscr{G}_t$-martingale, and hence, by (2.2), $\beta^{(\eta)}(T) - 1$ is a mean zero square integrable random variable.

Let $\mathcal{M}$ denote the class of all square integrable random variables of the form $X = \int_0^T \gamma(s)\,dW(s)$ for some $\gamma \in \mathcal{A}^W$. Clearly, $\beta^{(\eta)}(T)-1$ is in $\mathcal{M}$ for all $\eta$ as above.

We assert now that $\mathcal{M}$ is all of $L_W^2(\theta)$. To see this let $Y$ be an arbitrary mean zero square integrable $\mathcal{G}_T$-measurable random variable. Suppose that $Y$ is orthogonal to $\mathcal{M}$, i.e., $E(YX) = 0$ for all $X \in \mathcal{M}$. Since $\mathcal{M}$ is a closed subspace of $L_W^2(\theta)$, to prove $\mathcal{M} = L_W^2(\theta)$ we need only show that $Y \equiv 0$. Let $\{\lambda_k\}$ be the sequence of eigenvalues of $Q$ and let $\{e_k\}$ be a complete orthonormal system (CONS) of corresponding eigenvectors. Let $N_0$ denote the set of positive integers. Suppose that $l \in N_0$, $N \in N_0$, $0 = t_1 \leqslant \cdots \leqslant t_N \leqslant T$, and that $\{\alpha_k\}_{k=1}^N$ is a sequence of reals. By taking $\eta$ to be the appropriate step function and using $EY = 0$, we see that

$$E\Big(Y \exp\Big\{\sum_{k=1}^{N-1} \alpha_k\big(W(t_{k+1}, e_l) - W(t_k, e_l)\big)\Big\}\Big) = 0.$$

This proves that $E[Y \mid W(t, e_l);\ 0 \leqslant t \leqslant T] = 0$ for all $l \in N_0$. In a similar manner we see that, for all $m \in N_0$, $E[Y \mid W(t, e_l);\ 0 \leqslant t \leqslant T,\ 1 \leqslant l \leqslant m] = 0$. The assertion now follows on observing that $\mathcal{G}_T \equiv \sigma\{W(\cdot, e_l);\ l \in N_0\}$ (cf. Proposition 4.1 of [5]).

Finally, let $M(t)$ be a mean zero square integrable martingale. Then there exists $\gamma \in \mathcal{A}^W$ such that $M(T) - M(0) = \int_0^T \gamma(s)\,dW(s)$ a.s. The proof now follows by taking conditional expectations with respect to $\mathcal{G}_t$ and using the martingale properties of $M(t)$ and the stochastic integral. ∎

Finally, in this section we will record two results which will be used in Section 3 in proving tightness for a sequence of Hilbert space valued processes. The first of these results is due to Aldous (cf. [25]). Let $(\mathscr{E}, d)$ be a Polish space. We denote by $\mathscr{C}([0, T]: \mathscr{E})$ the Polish space of continuous maps from $[0, T]$ to $\mathscr{E}$ equipped with the uniform convergence topology.

THEOREM 2.4. *Let $\{X^{(n)}\}$ be a sequence of processes with paths in $\mathscr{C}([0, T]: \mathscr{E})$. Suppose that $\{X^{(n)}(t)\}$ is tight for each rational $t \in [0, T]$ and that, for any sequence of stopping times $\{\tau_n\}$ such that $\tau_n \leqslant T$ and any sequence of non-negative numbers $\{\delta_n\}$ converging to zero as $n \to \infty$,*

$$d\big(X^{(n)}(\tau_n + \delta_n), X^{(n)}(\tau_n)\big) \to 0 \text{ in probability} \quad \text{as } n \to \infty.$$

*Then $\{X^{(n)}\}$ is tight.*

The proof of the following theorem can be found in [17].

THEOREM 2.5. *Let $K$ be a separable Hilbert space and let $\{e_i\}$ be a CONS in $K$. Let $\{\mu^{(n)}\}$ be a sequence of probability measures on $(K, \mathscr{B}(K))$. Then $\{\mu^{(n)}\}$ is tight if and only if*

1. *for all $N > 0$*

$$\lim_{A \to \infty} \sup_n \mu^{(n)}\{x \in K:\ \max_{1 \leqslant i \leqslant N} |\langle x, e_i \rangle| > A\} = 0;$$

2. *for any* $\delta > 0$,

$$\lim_{N\to\infty} \sup_n \mu^{(n)}\{x\in K\colon \|x-P_N(x)\|_K \geqslant \delta\} = 0,$$

*where* $P_N$ *is the projection operator with range* $\mathrm{span}\{e_1, \ldots, e_N\}$.

In some applications it is convenient to consider stochastic differential equations which are driven by a cylindrical Brownian motion rather than a Hilbert space valued Brownian motion. We close this section by giving the definition of a cylindrical Brownian motion and its connection with a Hilbert space valued Brownian motion. Recall that $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, \theta)$ is a probability space with an increasing family of right continuous $\theta$-complete sigma fields $\{\mathscr{F}_t\}_{0\leqslant t\leqslant T}$.

DEFINITION 2.6. A family $\{B(t, h)\colon 0\leqslant t\leqslant T, h\in H\}$ of random variables is said to be an *$\mathscr{F}_t$-cylindrical Brownian motion* if

(i) for every $h\in H$, $\|h\| = 1$, $\{B(t, h), \mathscr{F}_t\}_{0\leqslant t\leqslant T}$ is a standard Wiener process,

(ii) for every $0\leqslant t\leqslant T$, $\alpha_1, \alpha_2\in \boldsymbol{R}$ and $f_1, f_2\in H$,

$$B(t, \alpha_1 f_1+\alpha_2 f_2) = \alpha_1 B(t, f_1)+\alpha_2 B(t, f_2) \text{ a.s.}$$

Let $(H_1, \langle\cdot,\cdot\rangle_1)$ be a Hilbert space such that $H_1\supset H$ and the identity map $i\colon H\to H_1$ is Hilbert–Schmidt. Obviously, $H_1$ is not uniquely determined. Observe that the Hilbert–Schmidt embedding implies that if $\{e_i\}_{i=1}^{\infty}$ and $\{f_k\}_{k=1}^{\infty}$ are CONS in $H$ and $H_1$, respectively, then

$$\sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \langle e_i, f_k\rangle_1^2 < \infty. \tag{2.3}$$

Now let $\{e_i\}_{i=1}^{\infty}$ be a CONS in $H$ and define $\beta_j(t) \doteq B(t, e_j)$. Then from (2.3) we infer that the sequence $\{\sum_{j=1}^{n} e_j\beta_j(t)\}$ converges, in probability, in $H_1$ as $n\to\infty$. Furthermore, there is a trace class operator $Q_1$ on $H_1$ such that

$$W^*(t) \doteq \sum_{j=1}^{\infty} e_j\beta_j(t) \tag{2.4}$$

is a $Q_1$-Wiener process on $H_1$. The choice of the Hilbert space $H_1$ is immaterial in the sense that, for all such extensions $Q_1^{1/2}(H_1) = H$ and for $u\in H$, $\|u\| = \|Q_1^{-1/2}u\|_1$. Therefore, we can assume without loss of generality that $Q_1$ is strictly positive. We refer the reader to [5], Section 4.3, for proofs of these statements and further details. The following elementary lemma shows that one can always go from a cylindrical Brownian motion to a Hilbert space valued Wiener process in a measurable way.

LEMMA 2.7. *Let* $B(\cdot,\cdot)$ *be a cylindrical Brownian motion as above. Let* $X$ *be a random variable which is measurable with respect to* $\sigma\{B(s, h)\colon 0\leqslant s\leqslant T,$

*$h \in H\}$. Let the Hilbert space $H_1$ and an $H_1$-valued Wiener process $W^*(\cdot)$ be as above. Then there exists a measurable map $f$: $C([0, T]: H_1) \to \boldsymbol{R}$ such that $\theta \circ X^{-1} = \theta \circ f(W^*)^{-1}$.*

Proof. Note that $\mathscr{B}(C([0, T]: H_1))$ [the Borel $\sigma$-field on $C([0, T]: H_1)$] is precisely the sigma field $\sigma\{\langle \pi_t(\cdot), h\rangle_1: t \in [0, T], h \in H_1\}$, where $\pi_t$: $C([0, T]: H_1) \to H_1$ is defined as $\pi_t(x) \doteq x(t)$. Thus to prove the lemma it suffices to show that $\sigma\{\langle W^*(t), h\rangle_1: t \in [0, T], h \in H_1\}$ equals $\sigma\{B(t, h): h \in H, 0 \leqslant t \leqslant T\}$. The last statement is an immediate consequence of (2.4) and the observation that if $\{f_k\}_{k=1}^{\infty}$ is a CONS of eigenvectors of $Q_1$ with eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$, then for every $h \in H$,

$$B(t, h) = \sum_{j=1}^{\infty} \langle h, f_j\rangle \langle W(t), f_j\rangle_1 \text{ a.s. } \blacksquare$$

**3. The representation theorem.** This section is devoted to the proof of the representation theorem. For a bounded operator $A$ on $H$ let $\|A\|_{\text{op}}$ denote its operator norm. We begin with the following lemma:

LEMMA 3.1. *Let $\{v^{(n)}\}$ be a sequence of elements of $\mathscr{A}$* (cf. (2.1)). *Assume that*

$$M \doteq \sup_n \int_0^T E\|v^{(n)}(s)\|_0^2 \, ds < \infty. \tag{3.1}$$

*Then the sequence $\{\int_0^{\cdot} v^{(n)}(s)\,ds\}$ is tight in $\mathscr{C}([0, T]: H)$.*

Proof. For $0 \leqslant t \leqslant T$ define $X^{(n)}(t) \doteq \int_0^t v^{(n)}(s)\,ds$. The Cauchy–Schwarz inequality and the observation that $\|h\| \leqslant Q\|_{\text{op}}^{1/2} \|h\|_0$ for $h \in H_0$ yield that, for $\{\tau_n\}$ and $\{\delta_n\}$ as in Theorem 2.4,

$$\|X^{(n)}(\tau_n + \delta_n) - X^{(n)}(\tau_n)\| \leqslant \sqrt{\delta_n}\,\|Q\|_{\text{op}}^{1/2} \Big(\int_0^T \|v^{(n)}(s)\|_0^2\, ds\Big)^{1/2}.$$

Thus, by (3.1), $\|X^{(n)}(\tau_n + \delta_n) - X^{(n)}(\tau_n)\|$ converges to 0 in $L^2(\theta)$. It now suffices, in view of Theorem 2.4, to show that for each $t \in [0, T]$ the sequence $\{X^{(n)}(t)\}$ is tight in $H$. We will verify conditions 1 and 2 of Theorem 2.5 for the measures induced by $\{X^{(n)}\}$. Let $\{e_j\}$ be a CONS of eigenvectors as in the proof of Theorem 2.3. In order to verify condition 1, it suffices to note that, for $A > 0$ and $n, i \in \boldsymbol{N}_0$,

$$\theta\{|\langle X^{(n)}, e_i\rangle| > A\} \leqslant \frac{TM}{A^2}.$$

For condition 2 observe that

$$\|X^{(n)}(t) - P_N(X^{(n)}(t))\|^2 = \sum_{j=N+1}^{\infty} \Big\langle \int_0^t v^{(n)}(s)\,ds,\, e_j\Big\rangle^2.$$

Denoting $Q^{-1/2} v^{(n)}(s)$ by $\phi^{(n)}(s)$, we can rewrite the right-hand side of the last equality as $\sum_{j=N+1}^{\infty} \langle \int_0^t \phi^{(n)}(s)\,ds,\ Q^{1/2} e_j \rangle^2$. The Cauchy–Schwarz inequality shows that this last expression can be at most $T \int_0^T \|\phi^{(n)}(s)\|^2\,ds \sum_{j=N}^{\infty} \lambda_i$. Observing finally that $\|\phi^{(n)}(s)\| = \|v^{(n)}(s)\|_0$ and recalling (3.1), we can verify condition 2 applying Chebyshev's inequality. ∎

The following lemma will be used in some of the tightness arguments in Sections 3 and 4.

LEMMA 3.2. *Let $\{v^{(n)}\}$ be a sequence of elements of $\mathscr{A}$. Assume there is $M < \infty$ such that*

$$\sup_n \int_0^T \|v^{(n)}(s)\|_0^2\,ds \leqslant M \quad a.s.$$

*Suppose further that $v^{(n)}$ converges in distribution to $v$ with respect to the weak topology on $L^2([0, T]: H_0)$. Then $\int_0^{\cdot} v^{(n)}(s)\,ds$ converges in distribution to $\int_0^{\cdot} v(s)\,ds$ in $\mathscr{C}([0, T]: H)$.*

Proof. For $N \in N_0$ define

$$S_N \doteq \{u \in L^2([0, T]: H_0): \int_0^T \|u(s)\|_0^2\,ds \leqslant N\}. \tag{3.2}$$

One can endow $S_N$ with the weak topology, in which case it is a Polish space (cf. [9]). The lemma then follows immediately by observing that the map $\tau: S_M \to \mathscr{C}([0, T]: H)$ defined by $\tau(u) \doteq \int_0^{\cdot} u(s)\,ds$ is continuous. ∎

The following lemma concerning measurable selections will be used in the proof of the main theorem below.

LEMMA 3.3. *Let $E_1, E_2$ be Polish spaces and let $f: E_1 \times E_2 \to \boldsymbol{R}$ be a bounded continuous function. Let $K$ be a compact set in $E_2$. For each $x \in E_1$ define the sets*

$$\Gamma_x^1 \doteq \{y \in K: \inf_{y_0 \in K} f(x, y_0) = f(x, y)\},$$

$$\Gamma_x^2 \doteq \{y \in K: \sup_{y_0 \in K} f(x, y_0) = f(x, y)\}.$$

*Then for $i = 1, 2$ there exist Borel measurable functions $g_i: E_1 \to E_2$ such that $g_i(x) \in \Gamma_x^i$ for all $x \in E_1$.*

Proof. Let $x_n$ be a sequence in $E_1$ converging to $\bar{x}$. For each $n \in N_0$ and $i = 1, 2$ let $y_n^i \in \Gamma_{x_n}^i$. In view of Corollary 10.3 of [12] it suffices to show that $\{y_n^i\}$ has a limit point in $\Gamma_{\bar{x}}^i$. Let $\bar{y}^i$ be a limit point of $\{y_n^i\}$. The result now is an immediate consequence of the fact that for each $n$ both

$$\inf_{y_0 \in K} f(x_n, y_0) - f(x_n, y_n^1) \quad \text{and} \quad \sup_{y_0 \in K} f(x_n, y_0) - f(x_n, y_n^2)$$

equal zero and that the maps

$$(x, y) \to f(x, y) - \inf_{y_0 \in K} f(x, y_0) \quad \text{and} \quad (x, y) \to f(x, y) - \sup_{y_0 \in K} f(x, y_0)$$

are continuous. ∎

For probability measures $\theta_1, \theta_2$ on $(\Omega, \mathscr{F})$ we define the relative entropy of $\theta_1$ with respect to $\theta_2$ by

$$R(\theta_1 \| \theta_2) \doteq \int_\Omega \left( \log \frac{d\theta_1}{d\theta_2}(\omega) \right) \theta_1(d\omega)$$

whenever $\theta_1$ is absolutely continuous with respect to $\theta_2$ and $\log(d\theta_1/d\theta_2)$ is $\theta_1$-integrable. In all other cases, set $R(\theta_1 \| \theta_2) \doteq \infty$. Define

$$\mathscr{A}_N \doteq \{v \in \mathscr{A}: v(\omega) \in S_N \ \theta\text{-a.s.}\}. \tag{3.3}$$

LEMMA 3.4. *Let $\{f^{(n)}\}$ be a uniformly bounded sequence of real-valued measurable functions on $\mathscr{C}([0, T]: H)$ converging to $f$ a.s. $\theta$. Then*

$$\inf_{v \in \mathscr{A}_N} E\left( \frac{1}{2} \int_0^T \|v(s)\|_0^2 + f^{(n)}\left(W + \int_0^\cdot v(s)\, ds\right) \right) \tag{3.4}$$

*converges to*

$$\inf_{v \in \mathscr{A}_N} E\left( \frac{1}{2} \int_0^T \|v(s)\|_0^2 + f\left(W + \int_0^\cdot v(s)\, ds\right) \right) \quad \textit{as } n \to \infty. \tag{3.5}$$

Proof. Let $\varepsilon > 0$ be arbitrary. For each $n \in N_0$ pick an element $v^{(n),\varepsilon}$ of $\mathscr{A}_N$ such that

$$E\left( \frac{1}{2} \int_0^T \|v^{(n),\varepsilon}(s)\|_0^2 + f^{(n)}\left(W + \int_0^\cdot v^{(n),\varepsilon}(s)\, ds\right) \right)$$

is at most $\varepsilon$ larger than the infimum in (3.4). Since $\{v^{(n),\varepsilon}, n \in N_0\}$ is tight in $S_N$, we can pick a subsequence (relabeled by $n$) along which $(v^{(n),\varepsilon}, W)$ converges weakly to $(v^\varepsilon, W)$. Using Lemma 3.2 we see that $W + \int_0^\cdot v^{(n),\varepsilon}(s)\, ds$ converges weakly as elements of $\mathscr{C}([0, T]: H)$ to $W + \int_0^\cdot v^\varepsilon(s)\, ds$.

We next claim that

$$E\left(f^{(n)}\left(W + \int_0^\cdot v^{(n),\varepsilon}(s)\, ds\right)\right) \to E\left(f\left(W + \int_0^\cdot v^\varepsilon(s)\, ds\right)\right).$$

This is a consequence of [1], Lemma 2.8 (b), which states that for the last display to hold it is sufficient that the relative entropies

$$R\left(\mathscr{L}_\theta\left(W + \int_0^\cdot v^{(n),\varepsilon}(s)\, ds\right) \| \mathscr{L}_\theta(W)\right)$$

be uniformly bounded in $n$, where $\mathscr{L}_\theta(W)$ and $\mathscr{L}_\theta\left(W + \int_0^\cdot v^{(n),\varepsilon}(s)\, ds\right)$ denote the measures induced on $C([0, T]: H)$ by $W$ and $W+\int_0^{\cdot} v^{(n),\varepsilon}(s)\,ds$, respectively. But this is immediate by Theorem 2.2, since these relative entropies equal

$$E\int_0^T \|v^{(n),\varepsilon}(s)\|_0^2\,ds \leqslant N.$$

Using the weak convergence of $v^{(n),\varepsilon}$ to $v^\varepsilon$ and Fatou's lemma, we obtain

$$\liminf_{n\to\infty} E\left(\frac{1}{2}\int_0^T \|v^{(n),\varepsilon}(s)\|_0^2\,ds\right) \geqslant E\left(\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2\,ds\right).$$

Thus the limit inferior, as $n\to\infty$, of the expression in (3.4) is at least the expression in (3.5).

For the reverse inequality, pick an element $v^\varepsilon$ of $\mathscr{A}_N$ such that

$$E\left(\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2\,ds + f\left(W+\int_0^{\cdot} v^\varepsilon(s)ds\right)\right) \tag{3.6}$$

is at most $\varepsilon$ larger than the infimum in (3.5). Clearly,

$$E\left(\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2\,ds + f^{(n)}\left(W+\int_0^{\cdot} v^\varepsilon(s)ds\right)\right)$$

is at least the infimum in (3.4). As $n\to\infty$, this quantity converges to the expression in (3.6). Thus the limit superior, as $n\to\infty$, of the expression in (3.4) is at most the expression in (3.5). This proves the reverse inequality, and hence the lemma. ■

LEMMA 3.5. *Let $f$ be a bounded continuous function mapping $\mathscr{C}([0, T]: H)$ into $\mathbf{R}$.*

1. *Let $\tilde{v}\in\mathscr{A}$ be such that*

$$E\left(\exp\left\{\int_0^T \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\,ds\right\}\right) = 1,$$

*define $\tilde{W}(t) \doteq W(t)-\int_0^t \tilde{v}(s)\,ds$, and let $E^{\tilde{v}}$ denote expectation with respect to the measure $\gamma^{\tilde{v}}$ defined by*

$$d\gamma^{\tilde{v}} \doteq \exp\left\{\int_0^T \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\,ds\right\} d\theta.$$

*Let $v_0\in\mathscr{A}$ be an elementary process, and assume there is $M_0\in(0,\infty)$ such that $\|v_0(s)\|_0 \leqslant M_0$ for all $s\in[0, T]$ a.s. Then for every $\varepsilon>0$ there exist elementary processes $v_1, v_2\in\mathscr{A}^W$ such that $\|v_i(s)\|_0 \leqslant M_0$ for $i=1, 2$ and all $s\in[0, T]$, and*

$$E\left(\frac{1}{2}\int_0^T \|v_1(s)\|_0^2\,ds + f\left(W+\int_0^{\cdot} v_1(s)\,ds\right)\right) - \varepsilon \tag{3.7}$$

$$\leqslant E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|v_0(s)\|_0^2\,ds + f\left(\tilde{W} + \int_0^{\cdot} v_0(s)\,ds\right)\right)$$

$$\leqslant E\left(\frac{1}{2}\int_0^T \|v_2(s)\|_0^2\,ds + f\left(W + \int_0^{\cdot} v_2(s)\,ds\right)\right) + \varepsilon.$$

2. *Let $\mathscr{A}^{(b)}$ denote the subclass of $\mathscr{A}$ consisting of bounded elementary processes. Then*

$$\inf_{\tilde{v}\in\mathscr{A}^{(b)}} E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\,ds + f\left(\tilde{W} + \int_0^{\cdot} \tilde{v}(s)\,ds\right)\right)$$
$$= \inf_{v\in\mathscr{A}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\,ds + f\left(W + \int_0^{\cdot} v(s)\,ds\right)\right).$$

Proof. For the proof of part 1 we will use Lemma 3.3. We will only show the first inequality in (3.7) since the proof of the second inequality is similar, save that the corresponding supremization part of Lemma 3.3 is used instead. Suppose that the elementary process $v_0$ takes the form

$$v_0(s,\omega) \doteq X_0(\omega)\mathscr{I}_{\{0\}}(s) + \sum_{j=1}^{l} X_j(\omega)\mathscr{I}_{(t_j,t_{j+1}]}(s),$$

where $(s,\omega)\in[0,T]\times\Omega$, $0 = t_1 \leqslant t_2 \leqslant \ldots \leqslant t_{l+1} = T$ and $X_j$ are $H_0$-valued $\mathscr{F}_{t_j}$-measurable random variables satisfying $\|X_j(\omega)\|_0 \leqslant M_0$ a.s. for all $j\in\{0,\ldots,l\}$, and $\mathscr{I}$ denotes the indicator function. Define $F_1\colon H_0^{\otimes l+1}\to \boldsymbol{R}$ by

$$F_1(x_0,\ldots,x_l) \doteq \frac{1}{2}\sum_{i=0}^{l-1}(t_{i+1}-t_i)\|x_i\|_0^2,$$

so that

$$F_1\left(X_0(\omega),\ldots,X_l(\omega)\right) = \frac{1}{2}\int_0^T \|v_0(s)\|_0^2\,ds.$$

For $j=1,\ldots,l$ define measurable maps $\tilde{Z}_j$ from $\Omega$ to $\mathscr{M}_j \doteq C([0,t_{j+1}-t_j],H)$ by

$$\tilde{Z}_j(\omega)(s) \doteq \tilde{W}(\omega)(s+t_j) - \tilde{W}(\omega)(t_j), \quad 0\leqslant s\leqslant t_{j+1}-t_j.$$

From the continuity and boundedness of the map $f$ it follows that there exists a continuous bounded map $F_2\colon \left(H_0^{\otimes l+1}\times\left(\prod_{j=1}^{l}\mathscr{M}_j\right)\right)\to\boldsymbol{R}$ such that

$$f\left(\tilde{W} + \int_0^{\cdot} v_0(s)\,ds\right) = F_2(X_j,\ 0\leqslant j\leqslant l;\ \tilde{Z}_j,\ 1\leqslant j\leqslant l) \text{ a.s.}$$

For $1\leqslant i\leqslant l$, let $\boldsymbol{X}_i$ and $\tilde{\boldsymbol{Z}}_i$ denote the vectors $(X_0,\ldots,X_i)$ and $(\tilde{Z}_1,\ldots,\tilde{Z}_i)$, respectively. With this notation

$$(3.8)\qquad E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|v_0(s)\|_0^2\,ds + f\big(\tilde{W} + \int_0^{\bullet} v_0(s)\,ds\big)\right) = E^{\tilde{v}}\big(F_1(X_l) + F_2(X_l, \tilde{Z}_l)\big).$$

We recall that every probability measure on a Polish space is tight. This implies there is a compact set $K_0 \subset H_0$ such that

$$E^{\tilde{v}}\big(F_1(X_l) + F_2(X_l, \tilde{Z}_l)\big) \geqslant E^{\tilde{v}}\big(\mathcal{I}_{K_0^{\otimes l+1}}(X_l)\big(F_1(X_l) + F_2(X_l, \tilde{Z}_l)\big)\big) - \varepsilon/[4(l+1)].$$

Since $(\tilde{W}(t), \mathcal{F}_t)$ is a Wiener process under $\gamma^{\tilde{v}}$, if $0 \leqslant u_1 \leqslant u_2 \leqslant T$, then $\tilde{W}(u_2) - \tilde{W}(u_1)$ is independent of $\mathcal{F}_{u_1}$. Therefore, $\tilde{Z}_j$ is independent of $(X_j, \tilde{Z}_{j-1})$ under $\gamma^{\tilde{v}}$. Let $\mu_j$ denote the standard Wiener measure on $\mathcal{M}_j$ and let $F_2^{(1)}$ be the real-valued continuous map on $\big(H_0^{\otimes l+1} \times (\prod_{j=1}^{l-1} \mathcal{M}_j)\big)$ obtained by integrating out $\tilde{Z}_l$ from $F_2$, i.e.,

$$F_2^{(1)}(y) \doteq \int F_2(y, z)\,\mu_l(dz), \qquad \text{where } y \in \Big(H_0^{\otimes l+1} \times \big(\prod_{j=1}^{l-1} \mathcal{M}_j\big)\Big).$$

Recalling that $\|X_l\|_0 \leqslant M_0$ a.s. and applying Lemma 3.3 with

$$E_2 \doteq H_0, \qquad E_1 \doteq \Big(H_0^{\otimes l} \times \big(\prod_{j=1}^{l-1} \mathcal{M}_j\big)\Big),$$

$$K \doteq K_0 \cap \{x \in H_0 \colon \|x\|_0 \leqslant M_0\} \qquad \text{and} \qquad f \doteq F_1 + F_2^{(1)},$$

we infer that there exists a measurable function

$$h\colon \Big(H_0^{\otimes l} \times \big(\prod_{j=1}^{l-1} \mathcal{M}_j\big)\Big) \to H_0$$

satisfying $\|h(\cdot)\|_0 \leqslant M_0$ such that the right-hand side of (3.8) is bounded from below by

$$E^{\tilde{v}}\big(F_1(X_{l-1}, h(X_{l-1}, \tilde{Z}_{l-1})) + F_2^{(1)}(X_{l-1}, h(X_{l-1}, \tilde{Z}_{l-1}), \tilde{Z}_{l-1})\big) - \varepsilon/[2(l+1)].$$

By subtracting an additional $\varepsilon/2(l+1)$ from this lower bound, we can take $h$ to be a continuous map via an application of [8], Theorem V.16a, and the dominated convergence theorem. We now iterate the above procedure $l$ times to obtain the following inequality:

$$E^{\tilde{v}}\big(F_1(X_l) + F_2(X_l, \tilde{Z}_l)\big) \geqslant E^{\tilde{v}}\big(F_1(\Gamma(\tilde{Z}_l)) + F_2(\Gamma(\tilde{Z}_l), \tilde{Z}_l)\big) - \varepsilon,$$

where

- $\Gamma\colon \prod_{j=1}^{l} \mathcal{M}_j \to H_0^{\otimes l+1}$ is continuous,
- $\Gamma(z_l)$ can be written $\big(\Gamma_0, \Gamma_1(z_1), \ldots, \Gamma_l(z_l)\big)$, where $z_i \doteq (z_1, \ldots, z_i) \in \prod_{j=1}^{i} \mathcal{M}_j$,
- $\Gamma_0$ is a non-random element of $H_0$ bounded in norm by $M_0$,
- for $i = 1, \ldots, l$, $\Gamma_i\colon \prod_{j=1}^{i} \mathcal{M}_j \to H_0$ satisfies $\|\Gamma_i(\boldsymbol{u})\|_0 \leqslant M_0$ for $\boldsymbol{u} \in \prod_{j=1}^{i} \mathcal{M}_j$.

Now define for $j = 1, \ldots, l$ measurable maps $Z_j$ from $\Omega$ to $\mathcal{M}_j$ by

$$Z_j(\omega)(s) \doteq W(\omega)(s) - W(\omega)(t_j), \qquad t_j \leqslant s \leqslant t_{j+1},$$

and let $\boldsymbol{Z}_i \doteq (Z_1, \ldots, Z_i)$ for $i \in \{1, \ldots, l\}$. Finally, define

$$\bar{v}(s, \omega) \doteq \Gamma_0 \mathscr{I}_{\{0\}}(s) + \sum_{j=1}^{l} \Gamma_j(\boldsymbol{Z}_j(\omega)) \mathscr{I}_{(t_j, t_{j+1}]}(s).$$

Clearly, $\bar{v}(s)$ is an elementary process in $\mathscr{A}^W$ satisfying $\|\bar{v}(s)\|_0 \leqslant M_0$ for each $s \in [0, T]$ and

$$E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|v_0(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} v_0(s)\, ds\right)\right) \geqslant E\left(\frac{1}{2}\int_0^T \|\bar{v}(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} \bar{v}(s)\, ds\right)\right) - \varepsilon.$$

This proves part 1.

We turn now to part 2 of the lemma. Taking $\tilde{v}$ in (3.7) to be a bounded elementary process and $v_0 = \tilde{v}$, we obtain

$$\text{(3.9)} \qquad \inf_{\tilde{v} \in \mathscr{A}^{(b)}} E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} \tilde{v}(s)\, ds\right)\right)$$
$$\geqslant \inf_{v \in \mathscr{A}^{W,(b)}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right),$$

where $\mathscr{A}^{W,(b)}$ is the subclass of $\mathscr{A}^W$ of bounded elementary processes. Since elements of $\mathscr{A}^{W,(b)}$ are piecewise constant, for every $v \in \mathscr{A}^{W,(b)}$ we can construct $\tilde{v} \in \mathscr{A}^{(b)}$ via a recursive conditioning argument so that

$$\text{(3.10)} \qquad E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} \tilde{v}(s)\, ds\right)\right)$$
$$= E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

Combining (3.9), (3.10), we have

$$\text{(3.11)} \qquad \inf_{\tilde{v} \in \mathscr{A}^{(b)}} E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^{\cdot} \tilde{v}(s)\, ds\right)\right)$$
$$= \inf_{v \in \mathscr{A}^{W,(b)}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

Next, taking $\tilde{v} \equiv 0$ in (3.7) and observing that $\mathscr{A}^{W,(b)} \subset \mathscr{A}^{(b)}$, we obtain

$$\text{(3.12)} \qquad \inf_{v \in \mathscr{A}^{W,(b)}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right)$$
$$= \inf_{v \in \mathscr{A}^{(b)}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

Now, let $v \in \mathscr{A}$ be such that $E\{\int_0^T \|v(s)\|_0^2\, ds\} < \infty$. Choose a sequence $\{v^n: n \in N_0\}$ in $\mathscr{A}$ such that each $v^n$ is a bounded elementary process,

$$\lim_{n\to\infty} E\left(\int_0^T \|v^{(n)}(s) - v(s)\|_0^2\, ds\right) = 0,$$

$$\text{and} \quad \sup_{n\in N_0} E\left(\int_0^T \|v^{(n)}(s)\|_0^2\, ds\right) < 1 + E\left(\int_0^T \|v(s)\|_0^2\, ds\right).$$

Clearly, $\int_0^t (v^n(s) - v(s))\, ds$ converges to zero in probability for each $t \in [0, T]$. Also, an application of Lemma 3.1 shows that $\{\int_0^{\cdot} (v^n(s) - v(s))\, ds\}$ is tight in $\mathscr{C}([0, T]: H)$. Thus $(W, \int_0^{\cdot} v^n(s)\, ds)$ converges weakly to $(W, \int_0^{\cdot} v(s)\, ds)$, and since $f$ is continuous, we have

$$\lim_{n\to\infty} E\left(\frac{1}{2}\int_0^T \|v^n(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v^n(s)\, ds\right)\right) = E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

Using $\mathscr{A} \subset \mathscr{A}^{(b)}$, we prove that

$$\text{(3.13)} \qquad \inf_{v\in\mathscr{A}^{(b)}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right)$$
$$= \inf_{v\in\mathscr{A}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

The proof of part 2 is completed by combining (3.11), (3.12), and (3.13). ∎

We now present the main result of this section. Though in the theorem we take $f$ to be a bounded function, it can be shown (as in [1]) that the representation continues to hold if $f$ is bounded from above.

THEOREM 3.6. *Let $f$ be a bounded, Borel measurable function mapping $\mathscr{C}([0, T]: H)$ into $\boldsymbol{R}$. Then*

$$\text{(3.14)} \qquad -\log E \exp\{-f(W)\} = \inf_{v\in\mathscr{A}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

Proof. We claim that it suffices to prove the result for $f$ that are continuous. To see this, let $\{f^{(n)}\}$ be a sequence of real-valued continuous functions on $\mathscr{C}([0, T], H)$ such that $\sup_{x,n} |f^{(n)}(x)| \leqslant \sup_x |f(x)|$, and $f^{(n)}$ converges to $f$ $\theta$-a.s. Applying the dominated convergence theorem we obtain

$$-\log E \exp\{-f^{(n)}(W)\} \to -\log E \exp\{-f(W)\}.$$

For $\mathscr{B} \subset \mathscr{A}$ and $g$ a bounded, Borel measurable function mapping $\mathscr{C}([0, T]: H)$ into $\boldsymbol{R}$, define

$$\Lambda(\mathscr{B}, g) \doteq \inf_{v\in\mathscr{B}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + g\left(W + \int_0^{\cdot} v(s)\, ds\right)\right).$$

To prove the claim, we must show that $\Lambda(\mathscr{A}, f^{(n)})$ converges to $\Lambda(\mathscr{A}, f)$

as $n \to \infty$. Let

$$K \doteq \sup_x |f(x)| \quad \text{and} \quad \mathscr{C} \doteq \{v \in \mathscr{A}\colon E(\int_0^T \|v(s)\|_0^2\, ds) \leqslant 4K\}.$$

Then, clearly, $\Lambda(\mathscr{A}, f^{(n)})$ equals $\Lambda(\mathscr{C}, f^{(n)})$ and $\Lambda(\mathscr{A}, f)$ equals $\Lambda(\mathscr{C}, f)$. Let $\varepsilon > 0$ be arbitrary. Choose $N \in N_0$ such that $4K^2/N \leqslant \varepsilon/2$. Fix $v \in \mathscr{C}$ and define the stopping time

$$\tau_N \doteq \inf\{s \in [0, T]\colon \int_0^s \|v(s)\|_0^2 \geqslant N\} \wedge T.$$

Recall that

$$\mathscr{A}_N \doteq \{v \in \mathscr{A}\colon \int_0^T \|v(s)\|_0^2\, ds \leqslant N \ \ \theta\text{-a.s.}\}.$$

Let $v_N \in \mathscr{A}_N$ be defined by $v_N(s) \doteq v(s)\mathscr{I}_{[0,\tau_N]}(s)$, where $\mathscr{I}$ denotes the indicator function. We observe that

$$\begin{aligned}\Lambda(\mathscr{A}_N, f^{(n)}) &\leqslant E\left(\frac{1}{2}\int_0^T \|v_N(s)\|_0^2\, ds + f^{(n)}\left(W + \int_0^\cdot v_N(s)\, ds\right)\right)\\ &\leqslant E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\, ds + f^{(n)}\left(W + \int_0^\cdot v(s)\, ds\right)\right) + \varepsilon,\end{aligned}$$

where the inequality in the second line follows since $v \in \mathscr{C}$ implies that the probability of the set $\{\tau_N < T\}$ is at most $4K/N$. Taking the infimum over all $v \in \mathscr{C}$ in the inequality above we have

$$\Lambda(\mathscr{A}, f^{(n)}) \leqslant \Lambda(\mathscr{A}_N, f^{(n)}) \leqslant \Lambda(\mathscr{A}, f^{(n)}) + \varepsilon.$$

Exactly the same argument with $f^{(n)}$ replaced by $f$ gives

$$\Lambda(\mathscr{A}, f) \leqslant \Lambda(\mathscr{A}_N, f) \leqslant \Lambda(\mathscr{A}, f) + \varepsilon.$$

Finally, an application of Lemma 3.4 shows that $\Lambda(\mathscr{A}_N, f^{(n)})$ converges to $\Lambda(\mathscr{A}_N, f)$ as $n \to \infty$. This proves the claim.

Henceforth we will assume that $f$ is continuous. We prove that the left-hand side of (3.14) is bounded from above and below by the right-hand side.

Proof of the upper bound. From Proposition 1.4.2 of [10] it follows that

$$-\log E \exp\{-f(W)\} = \inf_{\gamma \in \mathscr{P}(\Omega)\colon \gamma \ll \theta} \{R(\gamma \| \theta) + E^\gamma(f(W))\}, \tag{3.15}$$

where $\mathscr{P}(\Omega)$ is the class of all probability measures on $(\Omega, \mathscr{F})$. Let $\tilde{v} \in \mathscr{A}$ be a bounded elementary process. Clearly, $(\int_0^t \langle \tilde{v}(s), dW(s)\rangle, \mathscr{F}_t)_{0 \leqslant t \leqslant T}$ is a real-

-valued continuous martingale with quadratic variation $\int_0^t \|\tilde{v}(s)\|_0^2 ds$ ([5], Section 4). The boundedness assumption also implies that the expectation $E(\exp\{\int_0^T \|\tilde{v}(s)\|_0^2 ds\})$ is finite, and therefore Proposition 5.12 of [18] yields

$$E\left(\exp\left\{\int_0^T \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds\right\}\right) = 1.$$

By Theorem 2.2, $d\gamma^{\tilde{v}} \doteq \exp\{\int_0^T \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds\}\, d\theta$ is a probability measure and, under $\gamma^{\tilde{v}}$, $\tilde{W}(t) \doteq W(t) - \int_0^t \tilde{v}(s)\, ds$ is a $Q$-Wiener process. The definition of the relative entropy function implies

$$R(\gamma^{\tilde{v}}\,\|\,\theta) \doteq E^{\tilde{v}}\left(\log\frac{d\gamma^{\tilde{v}}}{d\theta}\right) = E^{\tilde{v}}\left(\int_0^T \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds\right),$$

where $E^{\tilde{v}}$ denotes the expectation with respect to $\gamma^{\tilde{v}}$. The last expression equals $\frac{1}{2}E^{\tilde{v}}(\int_0^T \|\tilde{v}(s)\|_0^2 ds)$, and consequently

$$R(\gamma^{\tilde{v}}\,\|\,\theta) + E^{\tilde{v}}(f(W)) = E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds + f\left(\tilde{W} + \int_0^{\bullet} \tilde{v}(s)\, ds\right)\right).$$

It follows from (3.15) that

$$-\log E\exp\{-f(W)\} \leqslant E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds + f\left(\tilde{W} + \int_0^{\bullet} \tilde{v}(s)\, ds\right)\right)$$

if $\tilde{v}$ is a bounded elementary process. Therefore, an application of part 2 of Lemma 3.5 yields

$$\begin{aligned}-\log E\exp\{-f(W)\} &\leqslant \inf_{\tilde{v}\in\mathscr{A}^{(b)}} E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2 ds + f\left(\tilde{W} + \int_0^{\bullet} \tilde{v}(s)\, ds\right)\right)\\ &= \inf_{v\in\mathscr{A}} E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2 ds + f\left(W + \int_0^{\bullet} v(s)\, ds\right)\right).\end{aligned}$$

Proof of the lower bound. From Proposition 1.4.2 of [10] we have

$$-\log E\exp\{-f(W)\} = R(\gamma_0\,\|\,\theta) + E^{\gamma_0}(f(W)),$$

where $d\gamma_0/d\theta \doteq c\exp\{-f(W)\}$ a.s., and $c$ is the normalizing constant. Define

$$L(t) = E\left(\frac{d\gamma_0}{d\theta}\,\middle|\,\mathscr{G}_t\right).$$

Clearly, $(L(t), \mathscr{G}_t)_{0\leqslant t\leqslant T}$ is a right continuous martingale bounded from above and below by $\exp(2\|f\|_\infty)$ and $\exp(-2\|f\|_\infty)$, respectively. It follows from Theorem 2.3 that there exists $u\in\mathscr{A}^W$ such that for all $0\leqslant t\leqslant T$

$$L(t) = 1 + \int_0^t \langle u(s), dW(s)\rangle.$$

We can rewrite the last equality as

$$L(t) = 1 + \int_0^t \langle L(s)\tilde{v}(s), dW(s)\rangle,$$

where $\tilde{v}(t) \doteq u(t)/L(t)$. Since $L(t)$ is a real-valued continuous non-negative martingale with $L(0) \equiv 1$, we have ([16], Lemma 7.1.4)

$$L(t) = \exp\left(\int_0^t \langle \tilde{v}(s), dW(s)\rangle - \frac{1}{2}\int_0^t \|\tilde{v}(s)\|_0^2\, ds\right).$$

It follows from Theorem 2.2 that under $\gamma_0$

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

is a Brownian motion with covariance $Q$. Therefore

$$-\log E\exp\{-f(W)\} = E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^\cdot \tilde{v}(s)\, ds\right)\right). \tag{3.16}$$

As in the proof of part 2 of Lemma 3.5, we can approximate $\tilde{v}$ by a sequence $\{\tilde{v}^n, n \in N_0\}$ of bounded elementary processes in $\mathscr{A}$ such that

$$E^{\tilde{v}}\left(\int_0^T \|\tilde{v}^n(s)\|_0^2\, ds\right) \leqslant 1 + E^{\tilde{v}}\left(\int_0^T \|\tilde{v}(s)\|_0^2\, ds\right)$$

and

$$E^{\tilde{v}}\left(\int_0^T \|\tilde{v}^n(s) - \tilde{v}(s)\|_0^2\, ds\right) \to 0 \quad \text{as } n \to \infty.$$

It now follows, as in the proof of part 2 of Lemma 3.5, that

$$E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|\tilde{v}^n(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^\cdot \tilde{v}^n(s)\, ds\right)\right)$$

converges to the right-hand side of (3.16) as $n \to \infty$.

Let $\varepsilon \to 0$ be arbitrary. We have shown that there exists $M_0 < \infty$ and an elementary process $v_0 \in \mathscr{A}$ satisfying $\|v_0(s)\|_0 \leqslant M_0$ for all $s \in [0, T]$ such that

$$-\log E\exp\{-f(W)\} \geqslant E^{\tilde{v}}\left(\frac{1}{2}\int_0^T \|v_0(s)\|_0^2\, ds + f\left(\tilde{W} + \int_0^\cdot v_0(s)\, ds\right)\right) - \varepsilon.$$

The proof is now completed by applying part 1 of Lemma 3.5. ∎

As an immediate corollary to Theorem 3.6 and Lemma 2.7 we have the following representation theorem for a cylindrical Brownian motion. Define $\mathscr{A}^*$ to be the class of $H$-valued $\mathscr{F}_t$-predictable processes $\phi$, satisfying

$$\theta\left\{\int_0^T \|\phi(s)\|^2\, ds < \infty\right\} = 1.$$

COROLLARY 3.7. *Let $\{B(t, h): 0 \leqslant t \leqslant T,\ h \in H\}$ be an $\{\mathscr{F}_t\}$-cylindrical Brownian motion. Let $X$ be a bounded random variable which is measurable with respect to $\sigma\{B(s, h): h \in H,\ 0 \leqslant s \leqslant T\}$. Then*

$$-\log E \exp\{-X\} = \inf_{v \in \mathscr{A}^*} E\left(\frac{1}{2}\int_0^T \|v(s)\|^2 ds + f\left(W^* + \int_0^{\cdot} v(s)\, ds\right)\right),$$

*where $f$ and $W^*(\cdot)$ are related to $B(\cdot, \cdot)$, and $X$ is as in Lemma 2.7.*

**4. A large deviation principle.** Let $W(\cdot)$ be, as in Sections 2 and 3, an $H$-valued Wiener process with trace class covariance $Q$. Let $\mathscr{E}$ be a Polish space and for $\varepsilon > 0$ let $\mathscr{G}^\varepsilon$: $C([0, T]: H) \to \mathscr{E}$ be a measurable map. In this section we are interested in the large deviation principle for the family of random elements

$$X^\varepsilon \doteq \mathscr{G}^\varepsilon(W(\cdot)) \quad \text{as } \varepsilon \to 0. \tag{4.1}$$

As stated in the introduction, for Polish space valued random elements the Laplace principle and the large deviation principle are equivalent. We will show in this section that under appropriate conditions a Laplace principle holds for $\{X^\varepsilon\}$. This general result will be applied in the sequel [3]. We begin with the following definitions:

DEFINITION 4.1. A function $I$ mapping $\mathscr{E}$ to $[0, \infty]$ is called a *rate function* if for each $M < \infty$ the level set $\{x \in \mathscr{E}: I(x) \leqslant M\}$ is compact.

DEFINITION 4.2. Let $I$ be a rate function on $\mathscr{E}$. A family $\{X^\varepsilon\}_{\varepsilon > 0}$ of $\mathscr{E}$-valued random elements is said to satisfy the *Laplace principle* on $\mathscr{E}$ with rate function $I$ if for all real-valued bounded and continuous functions $h$ on $\mathscr{E}$:

$$\lim_{\varepsilon \to 0} \varepsilon \log E\left\{\exp\left[-\frac{1}{\varepsilon} h(X^\varepsilon)\right]\right\} = -\inf_{x \in \mathscr{E}} \{h(x) + I(x)\}. \tag{4.2}$$

The set $S_N$ of bounded deterministic controls defined in (3.2) will play a central role in the proof of the Laplace principle. Also recall the definition of $\mathscr{A}_N$ given in (3.3).

We are now ready to formulate the main assumption on $\mathscr{G}^\varepsilon$ under which the Laplace principle holds.

ASSUMPTION 4.3. *There exists a measurable map $\mathscr{G}^0$: $C([0, T]: H) \to \mathscr{E}$ such that the following hold:*

(i) *Consider $M < \infty$ and a family $\{v^\varepsilon\} \subset \mathscr{A}_M$ such that $v^\varepsilon$ converges in distribution (as $S_M$-valued random elements) to $v$. Then $\mathscr{G}^\varepsilon\left(W(\cdot) + (\sqrt{\varepsilon})^{-1} \int_0^{\cdot} v^\varepsilon(s)\, ds\right)$ converges in distribution to $\mathscr{G}^0\left(\int_0^{\cdot} v(s)\, ds\right)$.*

(ii) *For every $M < \infty$ the set*

$$\Gamma_M \doteq \left\{\mathscr{G}^0\left(\int_0^{\cdot} v(s)\, ds\right): v \in S_M\right\}$$

*is a compact subset of $\mathscr{E}$.*

For each $f \in \mathscr{E}$ define

$$I(f) \doteq \inf_{\{v \in L^2([0,T]:H_0): f = \mathscr{G}^0(\int_0^{\cdot} v(s)ds)\}} \left\{ \frac{1}{2} \int_0^T \|v(s)\|_0^2 \, ds \right\}, \tag{4.3}$$

where the infimum over the empty set is taken to be $\infty$.

It can be shown that solutions of a wide class of stochastic dynamical systems driven by a Hilbert space valued Wiener process or a cylindrical Brownian motion can be written as $\{\mathscr{G}^\varepsilon(W(\cdot))\}$ with $\mathscr{G}^\varepsilon$ satisfying Assumption 4.3 (see [3]).

Observe that if $\{\mathscr{G}^\varepsilon\}$ satisfies Assumption 4.3, then $I$ is a rate function on $\mathscr{E}$. The following is the main theorem of this section:

THEOREM 4.4. *Let $X^\varepsilon$ be as in* (4.1). *Suppose that $\{\mathscr{G}^\varepsilon\}$ satisfies Assumption* 4.3. *Then the family $\{X^\varepsilon\}_{\varepsilon>0}$ satisfies the Laplace principle in $\mathscr{E}$ with rate function $I$ as defined in* (4.3).

Proof. In order to prove the theorem we must show that (4.2) holds for all real-valued bounded and continuous functions $h$ on $\mathscr{E}$.

Proof of the lower bound. From Theorem 3.6 we have

$$\begin{aligned} -\varepsilon \log E \left\{ \exp\left[ -\frac{1}{\varepsilon} h(X^\varepsilon) \right] \right\} &= \inf_{v \in \mathscr{A}} E\left( \frac{\varepsilon}{2} \int_0^T \|v(s)\|_0^2 \, ds + h \circ \mathscr{G}^\varepsilon \left(W(\cdot) + \int_0^{\cdot} v(s)\,ds\right)\right) \\ &= \inf_{v \in \mathscr{A}} E\left( \frac{1}{2} \int_0^T \|v(s)\|_0^2 \, ds + h \circ \mathscr{G}^\varepsilon \left( W(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^{\cdot} v(s)\,ds \right)\right). \end{aligned}$$

Fix $\delta > 0$. Then for every $\varepsilon > 0$ there exists $v^\varepsilon \in \mathscr{A}$ such that

$$\begin{aligned} &\inf_{v \in \mathscr{A}} E\left( \frac{1}{2} \int_0^T \|v(s)\|_0^2 \, ds + h \circ \mathscr{G}^\varepsilon \left( W(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^{\cdot} v(s)\,ds \right)\right) \\ &\qquad \geqslant E\left( \frac{1}{2} \int_0^T \|v^\varepsilon(s)\|_0^2 \, ds + h \circ \mathscr{G}^\varepsilon \left( W(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^{\cdot} v^\varepsilon(s)\,ds \right)\right) - \delta. \end{aligned} \tag{4.4}$$

We will prove that

$$\begin{aligned} &\liminf_{\varepsilon \to 0} E\left( \frac{1}{2} \int_0^T \|v^\varepsilon(s)\|_0^2 \, ds + h \circ \mathscr{G}^\varepsilon \left( W(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^{\cdot} v^\varepsilon(s)\,ds \right)\right) \\ &\qquad \geqslant \inf_{x \in \mathscr{E}} \{ I(x) + h(x) \}. \end{aligned} \tag{4.5}$$

We claim that in proving (4.5) we can assume without loss of generality that for all $\varepsilon > 0$ and a.s.

$$\int_0^T \|v^\varepsilon(s)\|_0^2\,ds \leqslant N \tag{4.6}$$

for some finite number $N$. To see this, observe that if $M \doteq \|h\|_\infty$, then

$$\sup_{\varepsilon>0} E\left(\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2\,ds\right) \leqslant 2M+\delta < \infty.$$

Now define stopping times

$$\tau_N^\varepsilon \doteq \inf\{t\in[0,\,T]:\ \int_0^t \|v^\varepsilon(s)\|_0^2\,ds \geqslant N\} \wedge T.$$

The processes $v^{\varepsilon,N}(s) \doteq v^\varepsilon(s)\mathscr{I}_{[0,\tau_N^\varepsilon]}(s)$ are in $\mathscr{A}$, $\mathscr{I}$ being as before the indicator function, and furthermore

$$\theta\{v^\varepsilon \neq v^{\varepsilon,N}\} \leqslant \theta\{\int_0^T \|v^\varepsilon(s)\|_0^2\,ds \geqslant N\} \leqslant \frac{2M+\delta}{N}.$$

This observation implies that the right-hand side of (4.4) is at most

$$E\left(\frac{1}{2}\int_0^T \|v^{\varepsilon,N}(s)\|_0^2\,ds + h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} v^{\varepsilon,N}(s)\,ds\right)\right) - \frac{2M(2M+\delta)}{N} - \delta.$$

Hence it suffices to prove (4.5) with $v^\varepsilon(s)$ replaced by $v^{\varepsilon,N}(s)$. This proves the claim. Henceforth we will assume that (4.6) holds. Pick a subsequence (relabeled by $\varepsilon$) along which $v^\varepsilon$ converges in distribution to $v$ as $S_N$-valued random elements. We now infer from Assumption 4.3 that

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} E\left(\frac{1}{2}\int_0^T \|v^\varepsilon(s)\|_0^2\,ds + h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} v^\varepsilon(s)\,ds\right)\right)\\
&\geqslant E\left(\frac{1}{2}\int_0^T \|v(s)\|_0^2\,ds + h\big(\mathscr{G}^0\big(\int_0^{\cdot} v(s)\,ds\big)\big)\right)\\
&\geqslant \inf_{\{(x,v)\in\mathscr{E}\times L^2([0,T]:H_0):x=\mathscr{G}^0(\int_0^{\cdot} v(s)ds)\}} \left\{\frac{1}{2}\int_0^T \|v(s)\|_0^2\,ds + h(x)\right\} \geqslant \inf_{x\in\mathscr{E}}\{I(x)+h(x)\}.
\end{aligned}$$

This completes the proof of the lower bound.

Proof of the upper bound. Since $h$ is bounded, $\inf_{x\in\mathscr{E}}\{I(x)+h(x)\} < \infty$. Let $\delta > 0$ be arbitrary, and let $x_0\in\mathscr{E}$ be such that

$$I(x_0)+h(x_0) \leqslant \inf_{x\in\mathscr{E}}\{I(x)+h(x)\} + \delta/2.$$

Choose $\tilde{v} \in L^2([0, T]: H_0)$ such that

$$\frac{1}{2}\int_0^T \|\tilde{v}(t)\|_0^2\,dt \leqslant I(x_0)+\delta/2 \qquad \text{and} \qquad x_0 = \mathscr{G}^0\Big(\int_0^{\cdot} \tilde{v}(s)\,ds\Big).$$

By Theorem 3.6, for bounded and continuous functions $h$

$$\begin{aligned}
(4.7)\qquad & \limsup_{\varepsilon\to 0} -\varepsilon \log E\big(\exp\{-h(X^\varepsilon)/\varepsilon\}\big)\\
&= \limsup_{\varepsilon\to 0} \inf_{v\in\mathscr{A}} E\left(\frac{1}{2}\int_0^T \|v(t)\|_0^2\,dt + h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} v(s)\,ds\right)\right)\\
&\leqslant \limsup_{\varepsilon\to 0} E\left(\frac{1}{2}\int_0^T \|\tilde{v}(t)\|_0^2\,dt + h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \tilde{v}(s)\,ds\right)\right)\\
&= \frac{1}{2}\int_0^T \|\tilde{v}(t)\|_0^2\,dt + \limsup_{\varepsilon\to 0} E\left(h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \tilde{v}(s)\,ds\right)\right)\\
&\leqslant I(x_0)+\delta/2 + \limsup_{\varepsilon\to 0} E\left(h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \tilde{v}(s)\,ds\right)\right).
\end{aligned}$$

Now by Assumption 4.3, as $\varepsilon \to 0$

$$E\left(h\circ\mathscr{G}^\varepsilon\left(W(\cdot)+\frac{1}{\sqrt{\varepsilon}}\int_0^{\cdot} \tilde{v}(s)\,ds\right)\right)$$

converges to $h\big(\mathscr{G}^0\big(\int_0^{\cdot} \tilde{v}(s)\,ds\big)\big) = h(x_0)$. Thus the expression in (4.7) can be at most

$$\inf_{x\in\mathscr{E}} \{I(x)+h(x)\} + \delta.$$

Since $\delta$ is arbitrary, the proof is complete. ■

## REFERENCES


[1] M. Boué and P. Dupuis, *A variational representation for certain functionals of Brownian motion*, Ann. Probab. 26, No. 4 (1998), pp. 1641–1659.

[2] W. Bryc, *Large deviations by asymptotic value method*, 1990.

[3] A. Budhiraja and P. Dupuis, *Large deviation properties of dynamical systems driven by infinite dimensional Brownian motion*, Preprint, 1999.

[4] P. Chow, *Large deviation problem for some parabolic Itô equations*, Comm. Pure Appl. Math. 45 (1992), pp. 97–120.

[5] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992.

[6] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Academic Press, San Diego, Calif., 1989.

[7] J.-D. Deuschel and D. Stroock, *Large Deviations*, Academic Press, San Diego, Calif., 1989.
[8] J. L. Doob, *Measure Theory*, Springer, New York 1994.
[9] N. Dunford and J. Schwartz, *Linear Operators. Parts I, II, III*, Interscience Publishers, Wiley, 1958.
[10] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley, 1997.
[11] R. S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Springer, New York 1985.
[12] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, 1986.
[13] M. I. Freidlin, *Random perturbations of reaction diffusion equations: the quasi-deterministic approach*, Trans. Amer. Math. Soc. 305 (1988), pp. 665–697.
[14] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York 1984.
[15] K. Itô, *Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces*, SIAM, Philadelphia, 1984.
[16] G. Kallianpur, *Stochastic Filtering Theory*, Springer, 1980.
[17] G. Kallianpur and J. Xiong, *Stochastic Differential Equations in Infinite Dimensional Spaces*, Institute of Mathematical Statistics, 1996.
[18] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, 1991.
[19] M. Metivier, *Semimartingales*, Walter de Gruyter, 1982.
[20] M. Metivier and J. Pellaumail, *Stochastic Integration*, New York 1980.
[21] S. Peszat, *Large deviation principle for stochastic evolution equations*, Probab. Theory Related Fields 98 (1994), pp. 113–136.
[22] R. Sowers, *Large deviations for a reaction diffusion equation with non-gaussian perturbations*, Ann. Probab. 20 (1992), pp. 504–537.
[23] S. R. S. Varadhan, *Large Deviations and Applications*, SIAM, Philadelphia, 1984.
[24] S. R. S. Varadhan, *Asymptotic probabilities and differential equations*, Comm. Pure Appl. Math. 19 (1966), pp. 261–286.
[25] J. B. Walsh, *An introduction to stochastic partial differential equations*, Ecole d'Eté de Probabilités de Saint-Flour XIV, Lecture Notes in Math. (1986), pp. 266–443.

Amarjit Budhiraja
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46656, USA

Paul Dupuis
Lefschetz Center for Dynamical Systems
Brown University
Providence, RI 02912, USA

*Received on 24.2.2000*