SIAM J. CONTROL OPTIM.
Vol. 60, No. 1, pp. 385–409

SOME LARGE DEVIATION ASYMPTOTICS IN SMALL NOISE FILTERING PROBLEMS*

ANUGU SUMITH REDDY†, AMARJIT BUDHIRAJA‡, AND AMIT APTE§

Abstract. We consider nonlinear filters for diffusion processes when the observation and signal noises are small and of the same order. As the noise intensities approach zero, the nonlinear filter can be approximated by a certain variational problem that is closely related to Mortensen's optimization problem [R. Mortensen, *J. Optim. Theory Appl.*, 2 (1968), pp. 386–394]. This approximation result can be made precise through a certain Laplace asymptotic formula. In this work we study probabilities of deviations of true filtering estimates from that obtained by solving the variational problem. Our main result gives a large deviation principle for Laplace functionals whose typical asymptotic behavior is described by Mortensen-type variational problems. Proofs rely on stochastic control representations for positive functionals of Brownian motions and Laplace asymptotics of the Kallianpur–Striebel formula.

Key words. Laplace asymptotics, large deviation principle, nonlinear filtering, small observation and signal noise, minimum energy estimate, 4DVAR

AMS subject classifications. 60F10, 60G35, 93E11

DOI. 10.1137/21M142616X

1. Introduction.

In this work we study certain large deviation asymptotics for nonlinear filtering problems with small signal and observation noise. As the noise in the signal and observation processes vanishes, the filtering problem can formally be replaced by a variational problem, and one may approximate the filtering estimates (namely suitable conditional probabilities or expectations) by solutions of certain deterministic optimization problems. However, due to randomness there will be occasional large deviations of the true nonlinear filter estimates from the variational problem solutions. The main goal of this work is to investigate the probabilities of such deviations by establishing a suitable large deviation principle. Large deviations and related asymptotic problems in the context of small noise nonlinear filtering have been investigated, under different settings, in many works [15, 13, 2, 16, 21, 3, 24, 18, 19, 11, 22, 1]. We summarize the main results of these works and their relation to the current work at the end of this section.

In order to describe our results precisely, we begin by introducing the filtering model that we study. We consider a signal process X^ε given as the solution of the d-dimensional stochastic differential equation (SDE)

$$dX^\varepsilon(t) = b(X^\varepsilon(t))dt + \varepsilon\sigma(X^\varepsilon(t))dW(t),\ X^\varepsilon(0) = x_0,\ 0 \le t \le T, \tag{1.1}$$

*Received by the editors June 10, 2021; accepted for publication (in revised form) November 28, 2021; published electronically February 10, 2022.
https://doi.org/10.1137/21M142616X
Funding: The first and third authors acknowledge support of the Department of Atomic Energy, Government of India, under projects 12-R&D-TFR-5.10-1100 and RTI4001. The second author acknowledges support from the NSF (DMS-1814894 and DMS-1853968). The second author is also grateful for the support from the Nelder Fellowship from Imperial College, London.
†Department of Mathematics, Indian Institute of Technology-Bombay, Mumbai, 400076, India (anugu.reddy@math.iitb.ac.in).
‡Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599 USA (amarjit@unc.edu).
§International Centre for Theoretical Sciences-Tata Institute of Fundamental Research, Bangalore, 560089, India (apte@icts.res.in).

and an m-dimensional observation process Y^ε governed by the equation

$$Y^\varepsilon(t) = \int_0^t h(X^\varepsilon(s))ds + \varepsilon B(t),\ 0 \le t \le T, \tag{1.2}$$

on some probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$. Here $\varepsilon \in (0,\infty)$ is a small parameter, $T \in (0,\infty)$ is some given finite time horizon, W and B are mutually independent standard Brownian motions in $\mathbb{R}^k$ and $\mathbb{R}^m$, respectively, $x_0 \in \mathbb{R}^d$ is the known deterministic initial condition of the signal, and the functions b, σ, and h are required to satisfy the following condition.

Assumption 1. The following hold.

(a) The functions b, σ, h from $\mathbb{R}^d \to \mathbb{R}^d$, $\mathbb{R}^d \to \mathbb{R}^{d\times k}$, $\mathbb{R}^d \to \mathbb{R}^m$ are Lipschitz: For some $c_{\text{lip}} \in (0,\infty)$

$$\|b(x) - b(y)\| + \|\sigma(x) - \sigma(y)\| + \|h(x) - h(y)\| \le c_{\text{lip}}\|x - y\| \text{ for all } x, y \in \mathbb{R}^d.$$

(b) The function σ is bounded: For some $c_\sigma \in (0,\infty)$ $\sup_{x\in\mathbb{R}^d} \|\sigma(x)\| \le c_\sigma$.

(c) The function h is twice continuously differentiable with bounded derivatives.

Note that under Assumption 1 there is a unique pathwise solution of (1.1), and the solution is a stochastic process with sample paths in $\mathcal{C}_d$ (the space of continuous functions from $[0,T]$ to $\mathbb{R}^d$ equipped with the uniform metric).

The filtering problem is concerned with the computation of the conditional expectations of the form

$$\bar{\mathbb{E}}\left[\phi(X^\varepsilon) \mid \mathcal{Y}_T^\varepsilon\right], \tag{1.3}$$

where $\mathcal{Y}_T^\varepsilon \doteq \sigma\{Y^\varepsilon(s) : 0 \le s \le T\}$ and $\phi : \mathcal{C}_d \to \mathbb{R}$ is a suitable map. The stochastic process with values in the space of probability measures on $\mathcal{C}_d$, given by $\bar{\mathbb{P}}\left[X^\varepsilon \in \cdot \mid \mathcal{Y}_T^\varepsilon\right]$, is usually referred to as the nonlinear filter.

In this work we are interested in the study of the asymptotic behavior of the nonlinear filter as $\varepsilon \to 0$. Denote by $\xi^* \in \mathcal{C}_d$ the unique solution of

$$d\xi^*(t) = b(\xi^*(t))dt,\ \ \xi^*(0) = x_0. \tag{1.4}$$

It can be shown that, under additional conditions (see the discussion in section 2), as $\varepsilon \to 0$,

$$\bar{\mathbb{P}}\left[X^\varepsilon \in \cdot \mid \mathcal{Y}_T^\varepsilon\right] \to \delta_{\xi^*},\ \text{in probability, under } \bar{\mathbb{P}}, \tag{1.5}$$

weakly. In particular, for Borel subsets A of $\mathcal{C}_d$ whose closure does not contain ξ^* one will have $\bar{\mathbb{P}}\left[X^\varepsilon \in A \mid \mathcal{Y}_T^\varepsilon\right] \to 0$ in probability as $\varepsilon \to 0$. We will refer to such sets A as *sets of nontypical state trajectories*. It is of interest to study the rate of decay of conditional probabilities of sets of nontypical state trajectories. As a special case of the results of the current paper (see Corollary 4.2), it will follow that for every real continuous and bounded function ϕ on $\mathcal{C}_d$, denoting

$$U^\varepsilon[\phi] \doteq \bar{\mathbb{E}}\left[\exp\left\{-\frac{1}{\varepsilon^2}\phi(X^\varepsilon)\right\} \mid \mathcal{Y}_T^\varepsilon\right], \tag{1.6}$$

$$-\varepsilon^2 \log U^\varepsilon[\phi] \xrightarrow{\bar{\mathbb{P}}} \inf_{\eta\in\mathcal{C}_d}\left[\phi(\eta) + \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta)\right], \tag{1.7}$$

where $\xrightarrow{\bar{\mathbb{P}}}$ denotes convergence in probability under $\bar{\mathbb{P}}$, and J is the rate function on $\mathcal{C}_d$ associated with the large deviation principle for $\{X^\varepsilon\}_{\varepsilon>0}$ (see section 2). From this convergence it follows using standard arguments (see, e.g., [6, Theorem 1.8]) that, for all Borel subsets A of $\mathcal{C}_d$,

(1.8)

$$\underline{\lim}^{\bar{\mathbb{P}}}_{\varepsilon\to 0}\varepsilon^2 \log \bar{\mathbb{P}}\,[X^\varepsilon \in A \mid \mathcal{Y}^\varepsilon_T] \geq -\inf_{\eta\in A^o}\left[\frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta)\right],$$

$$\overline{\lim}^{\bar{\mathbb{P}}}_{\varepsilon\to 0}\varepsilon^2 \log \bar{\mathbb{P}}\,[X^\varepsilon \in A \mid \mathcal{Y}^\varepsilon_T] \leq -\inf_{\eta\in \bar{A}}\left[\frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta)\right],$$

where for real random variables Z^ε and a constant $\alpha \in \mathbb{R}$ we say $\overline{\lim}^{\bar{\mathbb{P}}}_{\varepsilon\to 0} Z^\varepsilon \leq \alpha$ (resp., $\underline{\lim}^{\bar{\mathbb{P}}}_{\varepsilon\to 0} Z^\varepsilon \geq \alpha$) if $(Z^\varepsilon - \alpha)^+$ (resp., $(\alpha - Z^\varepsilon)^+$) converges to 0 in $\bar{\mathbb{P}}$-probability, and for a set A, A^o and $\bar{A}$ denote its interior and closure, respectively.

Thus the convergence in (1.7) gives information on rates of decay of conditional probabilities of sets of nontypical state trajectories. Formally, denoting the infimum in the above display as $S(\xi^*, A)$, we can write approximations for conditional probabilities:

$$\bar{\mathbb{P}}\,[X^\varepsilon \in A \mid \mathcal{Y}^\varepsilon_T] \approx \exp\left\{-\frac{1}{\varepsilon^2} S(\xi^*, A)\right\}. \tag{1.9}$$

However, due to stochastic fluctuations, one may find that for some "rogue" observation trajectories the conditional probabilities on the left side of (1.9) are quite different from the deterministic approximation on the right side of (1.9). In order to quantify the probabilities of observing such rogue observation trajectories that cause deviations from the bounds in (1.8), a natural approach is to study a large deviation principle for $\mathbb{R}$ valued random variables $\{-\varepsilon^2 \log U^\varepsilon[\phi]\}$ whose typical (law of large numbers) behavior is described by the right side of (1.7). Establishing such a large deviation principle is the goal of this work. Such a result gives information on decay rates of probabilities of the form

$$\bar{\mathbb{P}}\left\{\left|\varepsilon^2 \log \bar{\mathbb{P}}\,[X^\varepsilon \in A \mid \mathcal{Y}^\varepsilon_T] + \inf_{\eta\in A}\left[\frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta)\right]\right| > \delta\right\}$$

for suitable sets $A \in \mathcal{B}(\mathcal{C}_d)$ and $\delta > 0$. Our main result is Theorem 2.1, which gives a large deviation principle for $\{-\varepsilon^2 \log U^\varepsilon[\phi]\}$, for every continuous and bounded function ϕ on $\mathcal{C}_d$ with a rate function defined by the variational formula in (2.16)–(2.17).

Notation. The following notation and definitions will be used. For $p \in \mathbb{N}$ the Euclidean norm in $\mathbb{R}^p$ will be denoted as $\|.\|$, and the corresponding inner product will be written as $\langle\cdot,\cdot\rangle$. The space of finite positive measures (resp., probability measures) on a Polish space S will be denoted by $\mathcal{M}(S)$ (resp., $\mathcal{P}(S)$). The space of bounded measurable (resp., continuous and bounded) functions from $S \to \mathbb{R}$ will be denoted by $\mathrm{BM}(S)$ and $C_b(S)$, respectively. For $\phi \in \mathrm{BM}(S)$, $\|\phi\|_\infty \doteq \sup_{x\in S}|\phi(x)|$. For $\phi \in \mathrm{BM}(S)$ and $\mu \in \mathcal{M}(S)$, $\mu[\phi] \doteq \int \phi\, d\mu$. Borel σ-field on a Polish space S will be denoted as $\mathcal{B}(S)$. For $p \in \mathbb{N}$ and $T \in (0,\infty)$, $\mathcal{C}_{p,T}$ will denote the space of continuous functions from $[0,T]$ to $\mathbb{R}^p$ which is equipped with the supremum norm, defined as $\|f\|_{*,T} \doteq \sup_{0\leq t\leq T}\|f(t)\|$, $f \in \mathcal{C}_{p,T}$. Since $T \in (0,\infty)$ will be fixed in most of this work, frequently the subscript T in $\mathcal{C}_{p,T}$ and $\|f\|_{*,T}$ will be dropped.

We denote by $\mathcal{L}_p^2 \doteq L^2([0,T]:\mathbb{R}^p)$ the Hilbert space of square-integrable functions from $[0,T]$ to $\mathbb{R}^p$. By convention, the infimum over an empty set will be taken to be ∞. For random variables X_n, X with values in some Polish space S, convergence in distribution of X_n to X will be denoted as $X_n \Rightarrow X$. A function I from a Polish space S to $[0,\infty]$ is called a rate function if it has compact sublevel sets, namely the set $\{x \in S : I(x) \le m\}$ is a compact set of S for every $m \in (0,\infty)$. Given a function $a:(0,\infty)\to(0,\infty)$ such that $a(\varepsilon)\to\infty$ as $\varepsilon \to 0$, and a rate function I on a Polish space S, a collection $\{U^\varepsilon\}_{\varepsilon>0}$ of S valued random variables is said to satisfy a large deviation principle (LDP) with rate function I and speed $a(\varepsilon)$ if for every $\phi \in C_b(S)$

$$\lim_{\varepsilon\to 0} -a(\varepsilon)^{-1}\log \mathbb{E}\left[\exp\{-a(\varepsilon)\phi(U^\varepsilon)\}\right] = \inf_{x\in S}[I(x)+\phi(x)].$$

Relation with existing body of work. Denote by $\mathcal{C}_m^1$ the collection of absolutely continuous functions $y \in \mathcal{C}_m$ that satisfy $\int_0^T \|\dot{y}(s)\|^2 ds < \infty$. For $y \in \mathcal{C}_m^1$ define $I_y : \mathcal{C}_d \to [0,\infty]$ as

$$I_y(\eta) = \frac{1}{2}\int_0^T \|h(\eta(s)) - \dot{y}(s)\|^2 ds + J(\eta), \tag{1.10}$$

where J is the rate function of $\{X^\varepsilon\}$ defined in (2.9). The functional I_y was introduced in Mortensen [20] as the objective function in an optimization problem whose minima describe the most probable trajectory given the data in a nonlinear filtering problem in an appropriate asymptotic sense. This functional is also used in implementing the popular 4DVAR data assimilation algorithm (cf. [7, section 3.2], [12, Chapter 16]). Connection of the optimization problem associated with the objective function in (1.10) with the asymptotics of the classical small noise filtering problem has been studied by several authors [15, 14, 16]. We now describe this connection.

In section 2 we will introduce a continuous map $\hat{\Lambda}_T^\varepsilon : \mathcal{C}_m \to \mathcal{P}(\mathcal{C}_d)$ such that $\hat{\Lambda}_T^\varepsilon(Y^\varepsilon) = \bar{\mathbb{P}}(X^\varepsilon \in \cdot \mid \mathcal{Y}_T^\varepsilon)$ a.s. In [15], Hijab established, under conditions (that include boundedness and smoothness of various coefficients functions), an LDP for the collection of probability measures (on $\mathcal{C}_d$) $\{\hat{\Lambda}_T^\varepsilon(y)\}_{\varepsilon>0}$ (with speed ε^{-2}), for a fixed y in $\mathcal{C}_m^1$, with rate function $\hat{I}_y : \mathcal{C}_d \to [0,\infty]$ given by

$$\hat{I}_y(\eta) = I_y(\eta) - \inf_{\hat{\eta}\in\mathcal{C}_d}\{I_y(\hat{\eta})\}. \tag{1.11}$$

In a related direction, Hijab's Ph.D. dissertation [14] studied asymptotics of the unnormalized conditional density and established, under conditions, an asymptotic formula of the form $q^\varepsilon(x,t) = \exp\left\{-\varepsilon^{-2}(W(x,t)+o(1))\right\}$, where $q^\varepsilon(x,t)$ denotes the solution of the Zakai equation associated with the nonlinear filter (cf. [17]). The deterministic function $W(x,t)$ coincides with Mortensen's (deterministic) minimum energy estimate [20] which is given as solution of a certain minimization problem related to the objective function $I_y(\eta)$. Results of Hijab were extended to random initial conditions in [16], once again assuming boundedness and smoothness of coefficients. In related work, the problem of constructing observers for dynamical systems as limits of stochastic nonlinear filters is studied in [2]. Heunis [13] studies a somewhat different asymptotic problem for small noise nonlinear filters. Specifically, it is shown in [13] that for every $\phi \in C_b(\mathcal{C}_d)$, $w \in \mathcal{C}_m$, and for any $\eta \in \mathcal{C}_d$ for which the map defined in (2.13) has a *unique* minimizer (at, say, η^*), $\hat{\Lambda}_T^\varepsilon\left(\int_0^\cdot h(\eta(s))ds + \varepsilon w\right)[\phi] \to \phi(\eta^*)$, as $\varepsilon \to 0$. This result and its connection to our work are further discussed in section 2. In particular, the statement in (1.5) follows readily upon using ideas similar to those in [13]. The work of Pardoux and Zeitouni [21] considers a one dimensional nonlinear filtering problem where the observation noise is small while the signal noise is $O(1)$ (specifically, the term $\varepsilon\sigma(X^{\varepsilon}(t))$ in (1.1) is replaced by 1). In this case the conditional distribution of $X(T)$ given $\mathcal{Y}_T^{\varepsilon}$ converges a.s. to a Dirac measure $\delta_{X(T)}$ as $\varepsilon \to 0$. The paper [21] proves a quenched LDP for this conditional distribution (regarded as a collection of probability measures on $\mathcal{C}_d$ parametrized by $X(T)(\omega)$) in $\mathcal{C}_d$. In a somewhat different direction, in a sequence of papers [24, 19, 18], the authors have studied asymptotics of the filtering problem under a small signal-to-noise ratio limit, under various types of model settings. In this case the nonlinear filter converges to the unconditional law of the signal and the authors establish large deviation principles characterizing probabilities of deviation of the filter from the above deterministic law. An analogous result in a correlated signal-observation noise case was studied in [3]. Finally, yet another type of large deviation problem in the context of nonlinear filtering (with correlated signal-observation noise) when the observation noise is $O(1)$ and the signal noise and drift are suitably small has been considered in a series of papers [11, 22, 1].

The closest connections of the current work are with [15] and [13]. Specifically, the asymptotic statements in (1.7) and (1.8) which follow as a special case of our results (see Corollary 4.2) are analogous to results in [15], except that instead of a fixed observation path we consider the actual observation process Y^{ε} (also we make substantially weaker assumptions on coefficients than [15]). However, our main interest is in an LDP for the convergence to the deterministic limit in (1.7); thus roughly speaking we are interested in quantifying the probability of deviations from the convergence statement in [15] (when a fixed observation path is replaced with the observation process Y^{ε}). This large deviation result, given in Theorem 2.1, is the main contribution of our work.

Proof idea. The proof of Theorem 2.1 is based on a variational representation for functionals of Brownian motions obtained in [4] (see also [5]); using this the proof of the LDP reduces to proving a key weak convergence result given in Lemma 4.1. The proof of Lemma 4.1 is the technical heart of this work. Important use is made of some key estimates obtained in [13] (see in particular Proposition 5.3). One of the key steps is to argue that terms of order ε^{-1} can be ignored in the exponent when studying Laplace asymptotics for the quantity on the left side of (3.6). This relies on several careful large deviation exponential estimates which are developed in section 5. Once Lemma 4.1 is available the proof of the LDP in Theorem 2.1 follows readily using the now well-developed weak convergence approach for the study of large deviation problems (cf. [6]).

Organization. It will be convenient to formulate the filtering problem on canonical path spaces and also to represent the nonlinear filter through a map given on the path space of the observation process. This formulation and our main result (Theorem 2.1) are given in section 2. The key idea in the proof of the LDP is a variational representation from [4]. A somewhat simplified version of this representation (cf. [6]) that is used in this work is presented in section 3. Section 4 presents a key lemma (Lemma 4.1) that is needed for implementing the weak convergence method for proving the large deviation result in Theorem 2.1. Section 5 is devoted to the proof of Lemma 4.1. Using this lemma, we complete the proof of Theorem 2.1 in section 6.

2. Setting and main result. Recall that X^{ε} has sample paths in $\mathcal{C}_d$. Similarly, the processes Y^{ε}, W, B have sample paths in $\mathcal{C}_m, \mathcal{C}_k, \mathcal{C}_m$, respectively. It will be convenient to formulate the filtering problem on suitable path spaces. Denote, for $p \in \mathbb{N}$, the standard Wiener measure on $(\mathcal{C}_p, \mathcal{B}(\mathcal{C}_p))$ as $\mathcal{W}_p$ and the Wiener measure with variance parameter ε^2 as $\mathcal{W}_p^\varepsilon$. Denote the canonical coordinate process on $(\mathcal{C}_k, \mathcal{B}(\mathcal{C}_k))$ as $\{\gamma(t) : 0 \le t \le T\}$ and consider the SDE on the probability space $(\mathcal{C}_k, \mathcal{B}(\mathcal{C}_k), \mathcal{W}_k)$,

$$dx^\varepsilon(t) = b(x^\varepsilon(t))dt + \varepsilon\sigma(x^\varepsilon(t))d\gamma(t),\ x^\varepsilon(0) = x_0,\ 0 \le t \le T.$$

From Assumption 1, the above SDE has a unique strong solution with sample paths in $\mathcal{C}_d$.

Consider the map $\mathcal{C}_k \to \Omega_x \doteq \mathcal{C}_d \times \mathcal{C}_k$ defined as $\omega \mapsto (x^\varepsilon(\omega), \gamma(\omega))$ and let $\mu^\varepsilon \doteq \mathcal{W}_k \circ (x^\varepsilon, \gamma)^{-1}$. Next, let $\Omega_y \doteq \mathcal{C}_m$ and consider the probability space

$$(\Omega, \mathcal{F}, \mathbb{Q}^\varepsilon) \doteq (\Omega_x \times \Omega_y, \mathcal{B}(\Omega_x) \otimes \mathcal{B}(\Omega_y), \mu^\varepsilon \otimes \mathcal{W}_m^\varepsilon).$$

Abusing notation, denote the coordinate maps on the above probability space as ξ, γ, ζ, namely

$$\xi(\omega) = \omega_1,\quad \gamma(\omega) = \omega_2,\quad \zeta(\omega) = \omega_3 \text{ for } \omega = ((\omega_1, \omega_2), \omega_3) \in \Omega_x \times \Omega_y.$$

We will frequently write $\xi(\omega)(s)$ as $\xi(s)$ for $(\omega, s) \in \Omega \times [0, T]$. Similar notational shorthand will be followed for other coordinate maps.

Note that, under $\mathbb{Q}^\varepsilon$, $\xi(0) = x_0$, γ and $\varepsilon^{-1}\zeta$ are independent standard Brownian motions in $\mathbb{R}^k$ and $\mathbb{R}^m$, respectively, and

$$\xi(t) = x_0 + \int_0^t b(\xi(s))ds + \varepsilon \int_0^t \sigma(\xi(s))d\gamma(s),\ 0 \le t \le T. \tag{2.1}$$

Define, for $\mathbb{Q}^\varepsilon$ a.e. $\omega = ((\omega_1, \omega_2), \omega_3)$, for $t \in [0, T]$,

$$L_t^\varepsilon(\omega) \doteq \exp\left\{\frac{1}{\varepsilon^2}\int_0^t \langle h(\xi(s)), d\zeta(s)\rangle - \frac{1}{2\varepsilon^2}\int_0^t \|h(\xi(s))\|^2 ds\right\}.$$

Note that, since under $\mathbb{Q}^\varepsilon$, $\varepsilon^{-1}\zeta$ is a standard Brownian martingale with respect to the filtration $\mathcal{F}_t^0 \doteq \sigma\{\gamma(s), \xi(s), \zeta(s) : 0 \le s \le t\}$, the first integral in the exponent is well defined as an Itô integral. From the independence of ξ and ζ under $\mathbb{Q}^\varepsilon$ and Assumption 1 it follows that L_t^ε is an $\{\mathcal{F}_t^0\}$-martingale under $\mathbb{Q}^\varepsilon$. Define a probability measure $\mathbb{P}^\varepsilon$ on $(\Omega, \mathcal{F})$ as $d\mathbb{P}^\varepsilon/d\mathbb{Q}^\varepsilon(\omega) \doteq L_T^\varepsilon(\omega)$ for $\mathbb{Q}^\varepsilon$ a.e. ω. Note that, by Girsanov's theorem, under $\mathbb{P}^\varepsilon$

$$\beta(t) \doteq \frac{1}{\varepsilon}\zeta(t) - \frac{1}{\varepsilon}\int_0^t h(\xi(s))ds,\ 0 \le t \le T, \tag{2.2}$$

is a standard m-dimensional Brownian motion which is independent of (ξ, γ). Rewriting the above equation as $\zeta(t) = \int_0^t h(\xi(s))ds + \varepsilon\beta(t)$, $0 \le t \le T$, we see that $\bar{\mathbb{P}} \circ (X^\varepsilon, Y^\varepsilon)^{-1} = \mathbb{P}^\varepsilon \circ (\xi, \zeta)^{-1}$. Next, for $\varepsilon > 0$, define $\Gamma_T^\varepsilon : \mathcal{C}_m \to \mathcal{M}(\mathcal{C}_d)$ as

$$\Gamma_T^\varepsilon(\omega_3)[A] \doteq \int_{\Omega_x} 1_A(\omega_1) L_T^\varepsilon((\omega_1, \omega_2), \omega_3) d\mu^\varepsilon(\omega_1, \omega_2),\ \omega_3 \in \mathcal{C}_m,\ A \in \mathcal{B}(\mathcal{C}_d). \tag{2.3}$$

The maps are well defined $\mathbb{P}^\varepsilon$-a.s., and using results of [8, 9, 10], one can obtain versions of these maps (denoted as $\hat{\Gamma}_T^\varepsilon$) which are continuous on $\mathcal{C}_m$. Also, define $\Lambda_T^\varepsilon : \mathcal{C}_m \to \mathcal{P}(\mathcal{C}_d)$ as

$$\Lambda_T^\varepsilon(\omega_3)[A] \doteq \frac{\Gamma_T^\varepsilon(\omega_3)[A]}{\Gamma_T^\varepsilon(\omega_3)[\mathcal{C}_d]},\ \mathbb{P}^\varepsilon\text{-a.e. } \omega_3 \in \mathcal{C}_m,\ A \in \mathcal{B}(\mathcal{C}_d). \tag{2.4}$$

Once again, for each $\varepsilon > 0$, this map is well defined $\mathbb{P}^\varepsilon$-a.s., and a continuous version of the map exists (which we denote as $\hat{\Lambda}^\varepsilon_T$) from [8, 9, 10]. Write, for $f \in \mathrm{BM}(\mathcal{C}_d)$,

(2.5)
$$\Gamma^\varepsilon_T(f,\omega_3) \doteq \int_{\mathcal{C}_d} f(\tilde{\omega})\Gamma^\varepsilon_T(\omega_3)[d\tilde{\omega}],\ \Lambda^\varepsilon_T(f,\omega_3) \doteq \int_{\mathcal{C}_d} f(\tilde{\omega})\Lambda^\varepsilon_T(\omega_3)[d\tilde{\omega}],\ \mathbb{P}^\varepsilon\text{-a.e. } \omega_3 \in \mathcal{C}_m.$$

Then with $(X^\varepsilon, Y^\varepsilon)$ as in (1.1)–(1.2), for $\phi \in \mathrm{BM}(\mathcal{C}_d)$

$$\bar{\mathbb{E}}\left[\phi(X^\varepsilon) \mid \mathcal{Y}^\varepsilon_T\right] = \Lambda^\varepsilon_T(\phi, Y^\varepsilon) \text{ a.s. } \bar{\mathbb{P}}. \tag{2.6}$$

Also,

$$\mathbb{E}_{\mathbb{P}^\varepsilon}\left[\phi(\xi) \mid \sigma\{\zeta(s) : 0 \le s \le T\}\right] = \Lambda^\varepsilon_T(\phi, \zeta) \text{ a.s. } \mathbb{P}^\varepsilon. \tag{2.7}$$

Equation (2.6) (or (2.7)) is known as the Kallianpur–Striebel formula, where $\mathbb{E}_{\mathbb{P}^\varepsilon}$ denotes the expectation under the probability measure $\mathbb{P}^\varepsilon$, and

$$\bar{\mathbb{P}} \circ (X^\varepsilon, Y^\varepsilon, W, B, \Lambda^\varepsilon_T(\phi, Y^\varepsilon))^{-1} = \mathbb{P}^\varepsilon \circ (\xi, \zeta, \gamma, \beta, \Lambda^\varepsilon_T(\phi, \zeta))^{-1}. \tag{2.8}$$

Let, for $\xi_0 \in \mathcal{C}_d$,

$$J(\xi_0) \doteq \inf_{\varphi \in \mathcal{U}(\xi_0)} \left[\frac{1}{2}\int_0^T \|\varphi(t)\|^2 dt\right], \tag{2.9}$$

where $\mathcal{U}(\xi_0)$ is the collection of all φ in $\mathcal{L}^2_k$ such that

$$\xi_0(t) = x_0 + \int_0^t b(\xi_0(s))ds + \int_0^t \sigma(\xi_0(s))\varphi(s)ds,\ t \in [0,T]. \tag{2.10}$$

Note that, by Assumption 1, for every $\varphi \in \mathcal{L}^2_k$ there is a unique solution of (2.10). By classical results of Freidlin and Wentzell (see, e.g., [6, Theorem 10.6]) the collection $\{X^\varepsilon\}$ of $\mathcal{C}_d$ valued random variables satisfies an LDP with rate function J and speed ε^{-2}; namely, for all $F \in C_b(\mathcal{C}_d)$

$$\lim_{\varepsilon\to 0} -\varepsilon^2 \log \int_{\Omega_x} \exp\left\{-\frac{1}{\varepsilon^2}F(\hat{\xi})\right\} d\mu^\varepsilon = \inf_{\xi_0 \in \mathcal{C}_d} \left[F(\xi_0) + J(\xi_0)\right], \tag{2.11}$$

where we denote the first coordinate process on Ω_x by $\hat{\xi}$, i.e., $\hat{\xi}(\omega) = \omega_1$ for $\omega = (\omega_1, \omega_2) \in \Omega_x = \mathcal{C}_d \times \mathcal{C}_k$. In [13] it is shown that for every $w \in \mathcal{C}_m$, and a given $\eta \in \mathcal{C}_d$, the probability measure

$$\hat{\Lambda}^\varepsilon_T\left(\int_0^\cdot h(\eta(s))ds + \varepsilon w(\cdot)\right) \to \delta_{\eta^*} \tag{2.12}$$

weakly *if* the map

$$\tilde{\eta} \mapsto J(\tilde{\eta}) + \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\tilde{\eta}(s))\|^2 ds \tag{2.13}$$

attains its infimum over $\mathcal{C}_d$ *uniquely* at η^*, where we recall that $\hat{\Lambda}^\varepsilon_T$ is the continuous version of Λ^ε_T. We remark that [13] assumes in addition to Assumption 1 that h and b are bounded, but an examination of the proof shows (see calculations in section 5) that these conditions can be replaced by linear growth conditions that are implied by Assumption 1.

Recall the function $\xi^* \in \mathcal{C}_d$ from (1.4). Then using ideas similar to those in [13], under Assumption 1, and assuming in addition that either $\sigma\sigma^\dagger$ is positive definite or h is a one-to-one function, it follows that

$$\Lambda_T^\varepsilon \to \delta_{\xi^*}, \text{ in probability, under } \mathbb{P}^\varepsilon, \tag{2.14}$$

weakly, as $\varepsilon \to 0$. This is a consequence of the fact that when $\eta = \xi^*$ the map in (2.13) achieves its minimum (which is 0) uniquely at ξ^*.

As a consequence of the results of the current paper (see Corollary 4.2), one can show the Laplace asymptotic formula in (1.7). Recall from the discussion in the introduction that the convergence in (1.7) gives information on asymptotics of conditional probabilities of sets of nontypical state trajectories. In order to quantify the decay rate of probabilities of observing rare observation trajectories that cause deviations from the deterministic variational quantity in (1.7), we will establish an LDP for $\{-\varepsilon^2 \log U^\varepsilon[\phi]\}$ defined in (1.6).

We now present the rate function associated with this LDP.

Define the map $H : \mathcal{C}_d \times \mathcal{C}_d \times \mathcal{L}_m^2 \to \mathbb{R}_+$ as

$$H(\eta, \tilde{\eta}, \psi) \doteq \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\tilde{\eta}(s)) - \psi(s)\|^2 ds. \tag{2.15}$$

Also, for $\varphi \in \mathcal{L}_k^2$, let ξ_0^φ be given as the unique solution of (2.10).

We now introduce the rate function that will govern the large deviation asymptotics of $-\varepsilon^2 \log U^\varepsilon[\phi]$.

Fix $\phi \in C_b(\mathcal{C}_d)$ and define $I^\phi : \mathbb{R} \to [0, \infty]$ as

$$I^\phi(z) = \inf_{(\varphi,\psi)\in\mathcal{S}(z)} \left[\frac{1}{2}\int_0^T \|\varphi(t)\|^2 dt + \frac{1}{2}\int_0^T \|\psi(t)\|^2 dt\right], \tag{2.16}$$

where $\mathcal{S}(z)$ is the collection of all (φ, ψ) in $\mathcal{L}_k^2 \times \mathcal{L}_m^2$ such that

$$\inf_{\eta\in\mathcal{C}_d} [H(\eta, \xi_0^\varphi, \psi) + \phi(\eta) + J(\eta)] - \inf_{\eta\in\mathcal{C}_d} [H(\eta, \xi_0^\varphi, \psi) + J(\eta)] = z. \tag{2.17}$$

The following is the main result of the work.

THEOREM 2.1. *Suppose that Assumption 1 is satisfied. Then for every $\phi \in C_b(\mathcal{C}_d)$, the collection $\{-\varepsilon^2 \log U^\varepsilon[\phi]\}$ satisfies an LDP on $\mathbb{R}$ with rate function I^ϕ and speed ε^{-2}.*

3. A variational representation.

Fix $\phi \in C_b(\mathcal{C}_d)$. Recall the functional $U^\varepsilon[\phi]$ from (1.6). From (2.6), note that one can write $U^\varepsilon[\phi]$ as

$$U^\varepsilon[\phi] = \Lambda_T^\varepsilon\left(\exp\{-\varepsilon^{-2}\phi(\cdot)\}, Y^\varepsilon\right),$$

whose distribution under $\bar{\mathbb{P}}$ is the same as the distribution of $\Lambda_T^\varepsilon\left(\exp\{-\varepsilon^{-2}\phi(\cdot)\}, \zeta\right)$ under $\mathbb{P}^\varepsilon$. Let

$$V^\varepsilon[\phi] = -\varepsilon^2 \log \Lambda_T^\varepsilon\left(\exp\{-\varepsilon^{-2}\phi(\cdot)\}, \zeta\right).$$

Using this equality of laws and the equivalence between LDPs and Laplace principles (see, e.g., [6, Theorems 1.5 and 1.8]), in order to prove Theorem 2.1 it suffices to show that I^ϕ has compact sublevel sets, i.e.,

$$\text{for every } m \in \mathbb{R}_+, \{z \in \mathbb{R} : I^\phi(z) \leq m\} \text{ is compact,} \tag{3.1}$$

and, for every $G \in C_b(\mathbb{R})$,

$$\lim_{\varepsilon\to 0} -\varepsilon^2 \log \mathbb{E}_{\mathbb{P}^\varepsilon}\left[\exp\left\{-\varepsilon^{-2} G(V^\varepsilon[\phi])\right\}\right] = \inf_{z\in\mathbb{R}}\{G(z) + I^\phi(z)\}. \tag{3.2}$$

The proof of the identity in (3.2) will use a variational representation for nonnegative functionals of Brownian motions given by Boué and Dupuis [4]. We now use this representation to give a variational formula for the left side of the above equation. Let $\mathcal{F}_t$ denote the $\mathbb{P}^\varepsilon$-completion of $\mathcal{F}_t^0$ and denote by $\mathcal{A}^k$ (resp., $\mathcal{A}^m$) the collection of all $\{\mathcal{F}_t\}$-progressively measurable $\mathbb{R}^k$ (resp., $\mathbb{R}^m$) valued processes g such that for some $M = M(g) \in (0,\infty)$, $\int_0^T \|g(s)\|^2 ds \le M$ a.s.

For $(u,v) \in \mathcal{A}^k \times \mathcal{A}^m$, let ξ^u solve, on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}^\varepsilon)$,

$$\xi^u(t) = x_0 + \int_0^t b(\xi^u(s))ds + \varepsilon\int_0^t \sigma(\xi^u(s))d\gamma(s) + \int_0^t \sigma(\xi^u(s))u(s)ds. \tag{3.3}$$

Also define

$$\zeta^{u,v}(t) = \int_0^t h(\xi^u(s))ds + \varepsilon\beta(t) + \int_0^t v(s)ds,\ 0 \le t \le T. \tag{3.4}$$

The reader should note the difference between the trajectory ξ_0^φ, which (as stated below (2.15)) is defined as a solution to (2.10), and the trajectory ξ^u introduced above. The former is given as the solution of a controlled ordinary differential equation (with control φ), while the latter is the solution of a controlled SDE (with control u). In particular, on setting $\gamma \equiv 0$ and $u = \varphi$, the latter reduces to the former. Occasionally, to emphasize the dependence of above processes on ε we will write $(\xi^u, \zeta^{u,v})$ as $(\xi^{\varepsilon,u}, \zeta^{\varepsilon,u,v})$.

Now let

$$\bar{V}^{\varepsilon,u,v}[\phi] \doteq -\varepsilon^2 \log \Lambda_T^\varepsilon\left(\exp\{-\varepsilon^{-2}\phi(\cdot)\}, \zeta^{\varepsilon,u,v}\right). \tag{3.5}$$

When clear from the context we will drop (u,v,ϕ) from the notation in $\bar{V}^{\varepsilon,u,v}[\phi]$ and simply write $\bar{V}^\varepsilon$. Then it follows from [4] (cf. [6, Theorem 8.3]) that

$$\begin{aligned} &-\varepsilon^2 \log \mathbb{E}_{\mathbb{P}^\varepsilon}\left[\exp\left\{-\varepsilon^{-2}G(V^\varepsilon[\phi])\right\}\right] \\ &= \inf_{(u,v)\in\mathcal{A}^k\times\mathcal{A}^m} \mathbb{E}_{\mathbb{P}^\varepsilon}\left[G(\bar{V}^{\varepsilon,u,v}[\phi]) + \frac{1}{2}\int_0^T (\|u(s)\|^2 + \|v(s)\|^2)ds\right]. \end{aligned} \tag{3.6}$$

Indeed, since $V^\varepsilon[\phi]$ is a measurable functional of (γ,β), i.e., $V^\varepsilon[\phi] = \mathcal{G}^\varepsilon(\gamma,\beta)$, for some measurable functional $\mathcal{G}^\varepsilon : \mathcal{C}_k \times \mathcal{C}_m \to \mathbb{R}$, $G \circ \mathcal{G}^\varepsilon$ is a bounded measurable functional on $\mathcal{C}_k \times \mathcal{C}_m$, and we can apply [6, Theorem 8.3], with $\mathcal{R}$ there taken to be $\mathcal{A}_b$. We note that the latter result is stated for a general Hilbert space valued Brownian motion, while here we apply it for a finite dimensional Brownian motion, and so, in particular, in the notation of [6, section 8.1], $\mathcal{H} = \mathcal{H}_0 = \mathbb{R}^{k+m}$ and Λ is the identity operator.

4. A key lemma.

For $M \in (0,\infty)$, let

$$S_M \doteq \left\{(\varphi,\psi) \in \mathcal{L}_k^2 \times \mathcal{L}_m^2 : \int_0^T (\|\varphi(s)\|^2 + \|\psi(s)\|^2)ds \le M\right\}.$$

We equip S_M with the weak topology under which $(\varphi_n,\psi_n) \to (\varphi,\psi)$ as $n \to \infty$ if and only if for all $(f,g) \in \mathcal{L}_k^2 \times \mathcal{L}_m^2$

$$\int_0^T [\langle\varphi_n(s), f(s)\rangle + \langle\psi_n(s), g(s)\rangle]ds \to \int_0^T [\langle\varphi(s), f(s)\rangle + \langle\psi(s), g(s)\rangle]ds$$

as $n \to \infty$. This topology can be metrized so that S_M is a compact metric space.

Recall $\phi \in C_b(\mathcal{C}_d)$ in the statement of Theorem 2.1. For $(\varphi, \psi) \in \mathcal{L}_k^2 \times \mathcal{L}_m^2$ define

$$V_0^{\varphi,\psi}[\phi] \doteq \inf_{\eta \in \mathcal{C}_d} [H(\eta, \xi_0^\varphi, \psi) + \phi(\eta) + J(\eta)] - \inf_{\eta \in \mathcal{C}_d} [H(\eta, \xi_0^\varphi, \psi) + J(\eta)]. \tag{4.1}$$

Note that with this notation $\mathcal{S}(z)$ (introduced below (2.16)) is the collection of all (φ, ψ) in $\mathcal{L}_k^2 \times \mathcal{L}_m^2$ such that $V_0^{\varphi,\psi}[\phi] = z$. When (φ, ψ) are $\mathcal{L}_k^2 \times \mathcal{L}_m^2$ valued random variables, we will denote the random variable $V_0^{\varphi,\psi}[\phi](\omega) \doteq V_0^{\varphi(\omega),\psi(\omega)}[\phi]$, $\omega \in \Omega$, once more as $V_0^{\varphi,\psi}[\phi]$.

The following lemma will be the key to the proof of Theorem 2.1.

LEMMA 4.1. *Fix $M \in (0, \infty)$. Let $\{(u_n, v_n)\}$ be a sequence of S_M valued random variables such that $(u_n, v_n) \in \mathcal{A}^k \times \mathcal{A}^m$ for every n. Suppose that (u_n, v_n) converges in distribution to (u, v). Suppose ε_n is a sequence of positive reals converging to 0 as $n \to \infty$. Then $\bar{V}^{\varepsilon_n, u_n, v_n}[\phi] \to V_0^{u,v}[\phi]$, in distribution, as $n \to \infty$.*

As an immediate corollary of the lemma we have the following.

COROLLARY 4.2. *As $\varepsilon \to 0$,*

$$-\varepsilon^2 \log U^\varepsilon[\phi] \xrightarrow{\bar{\mathbb{P}}} \inf_{\eta \in \mathcal{C}_d} \left[\phi(\eta) + \frac{1}{2} \int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta) \right].$$

Proof. Note that $\bar{V}^{\varepsilon,0,0}[\phi]$ under $\mathbb{P}^\varepsilon$ has the same distribution as $-\varepsilon^2 \log U^\varepsilon[\phi]$ under $\bar{\mathbb{P}}$. From Lemma 4.1 it follows that $\bar{V}^{\varepsilon,0,0}[\phi]$, and thus $-\varepsilon^2 \log U^\varepsilon[\phi]$ converges in distribution to $V_0^{0,0}[\phi]$. Also note that

$$\begin{aligned} V_0^{0,0}[\phi] &= \inf_{\eta \in \mathcal{C}_d} \left[H(\eta, \xi_0^0, 0) + \phi(\eta) + J(\eta) \right] - \inf_{\eta \in \mathcal{C}_d} \left[H(\eta, \xi_0^0, 0) + J(\eta) \right] \\ &= \inf_{\eta \in \mathcal{C}_d} [H(\eta, \xi^*, 0) + \phi(\eta) + J(\eta)] - \inf_{\eta \in \mathcal{C}_d} [H(\eta, \xi^*, 0) + J(\eta)], \end{aligned}$$

where we used the fact that $\xi_0^0 = \xi^*$. Observe that

$$\inf_{\eta \in \mathcal{C}_d} \left[H(\eta, \xi_0^0, 0) + J(\eta) \right] = \inf_{\eta \in \mathcal{C}_d} \left[\frac{1}{2} \int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta) \right] = 0$$

since $\frac{1}{2} \int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta)$ evaluated at $\eta = \xi^*$ equals 0. Thus, recalling the definition of H, we have that $-\varepsilon^2 \log U^\varepsilon[\phi]$ converges in distribution (under $\bar{\mathbb{P}}$) to

$$V_0^{0,0}[\phi] = \inf_{\eta \in \mathcal{C}_d} \left[\phi(\eta) + \frac{1}{2} \int_0^T \|h(\eta(s)) - h(\xi^*(s))\|^2 ds + J(\eta) \right].$$

Since the right side above is nonrandom, we in fact have convergence in probability, which completes the proof. ∎

5. Proof of Lemma 4.1.

Let $(u, v) \in \mathcal{A}_k \times \mathcal{A}_m$. Define canonical coordinate processes on Ω_x as $\hat{\xi}(\tilde{\omega}) = \tilde{\omega}_1$ and $\hat{\gamma}(\tilde{\omega}) = \tilde{\omega}_2$, $\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2) \in \mathcal{C}_d \times \mathcal{C}_k$. Note that

$$\exp\left\{-\varepsilon^{-2} \bar{V}^{\varepsilon,u,v}[\phi]\right\} = \frac{\Gamma_T^\varepsilon\left(\exp\{-\varepsilon^{-2}\phi(\cdot)\}, \zeta^{u,v}\right)}{\Gamma_T^\varepsilon\left(1, \zeta^{\varepsilon,u,v}\right)},$$

and for $f \in C_b(\mathcal{C}_d)$, $\mathbb{P}^\varepsilon$-a.s., recall from (2.5) that

$$\Gamma^\varepsilon_T(f, \zeta^{\varepsilon,u,v}) = \int_{\Omega_x} f(\hat{\xi}(\tilde{\omega})) \exp\left\{\frac{1}{\varepsilon^2}\int_0^t \langle h(\hat{\xi}(\tilde{\omega})(s)), d\zeta^{u,v}(s)\rangle - \frac{1}{2\varepsilon^2}\int_0^t \|h(\hat{\xi}(\tilde{\omega}(s)))\|^2 ds\right\}\mu^\varepsilon(d\tilde{\omega}).$$

Suppressing $\tilde{\omega}$ in notation, we have

$$\frac{1}{\varepsilon^2}\int_0^t \langle h(\hat{\xi}(s)), d\zeta^{u,v}(s)\rangle - \frac{1}{2\varepsilon^2}\int_0^t \|h(\hat{\xi}(s))\|^2 ds = \frac{1}{\varepsilon}\int_0^T \langle h(\hat{\xi}(s)), d\beta(s)\rangle - \frac{1}{\varepsilon^2}H(\hat{\xi}, \xi^u, v) + \frac{1}{2\varepsilon^2}\int_0^T (\|h(\xi^u(s))\|^2 + \|v(s)\|^2)ds + \frac{1}{\varepsilon^2}\int_0^T h(\xi^u(s))\cdot v(s)ds.$$

Thus, letting

$$F(\tilde{\omega}, \beta) \doteq \int_0^T \langle h(\hat{\xi}(\tilde{\omega})(s)), d\beta(s)\rangle, \tag{5.1}$$

we can write

$$\exp\left\{-\varepsilon^{-2}\bar{V}^{\varepsilon,u,v}[\phi]\right\} = \frac{\int_{\Omega_x} \exp\left\{\frac{1}{\varepsilon}F(\tilde{\omega},\beta) - \frac{1}{\varepsilon^2}(\phi(\hat{\xi}(\tilde{\omega})) + H(\hat{\xi}(\tilde{\omega}), \xi^u, v))\right\}\mu^\varepsilon(d\tilde{\omega})}{\int_{\Omega_x} \exp\left\{\frac{1}{\varepsilon}F(\tilde{\omega},\beta) - \frac{1}{\varepsilon^2}H(\hat{\xi}(\tilde{\omega}), \xi^u, v)\right\}\mu^\varepsilon(d\tilde{\omega})}. \tag{5.2}$$

Let now $\varepsilon_n, u_n, v_n, u, v$ be as in the statement of Lemma 4.1. Using Assumption 1 it is immediate that

$$(u_n, v_n, \xi^{\varepsilon_n,u_n}, \zeta^{\varepsilon_n,u_n,v_n}, \beta) \Rightarrow (u, v, \xi^u_0, \zeta^{u,v}_0, \beta) \tag{5.3}$$

in $S_M \times \mathcal{C}_d \times \mathcal{C}_m \times \mathcal{C}_m$, where

$$\zeta^{u,v}_0(t) = \int_0^t h(\xi^u_0(s))ds + \int_0^t v(s)ds,\ t \in [0,T].$$

By appealing to the Skorohod representation theorem we can obtain, on some probability space $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$, random variables $(\tilde{u}_n, \tilde{v}_n, \tilde{\xi}^n, \tilde{\zeta}^n, \tilde{\beta}^n)$ with the same law as the random vector on the left side of (5.3) and $(\tilde{u}, \tilde{v}, \tilde{\xi}_0, \tilde{\zeta}_0, \tilde{\beta})$ with the same law as the vector on the right side of (5.3), such that

$$(\tilde{u}_n, \tilde{v}_n, \tilde{\xi}^n, \tilde{\zeta}^n, \tilde{\beta}^n) \to (\tilde{u}, \tilde{v}, \tilde{\xi}_0, \tilde{\zeta}_0, \tilde{\beta}),\ \mathbb{P}^*\text{-a.s.} \tag{5.4}$$

Henceforth, to simplify notation we will drop the $\tilde{\cdot}$ from the notation in the above vectors and denote the corresponding process $\bar{V}^{\varepsilon_n,u_n,v_n}[\phi]$ as $\bar{V}^n[\phi]$. Then, from (5.2), and the distributional equality noted above, it follows that

$$\exp\left\{-\varepsilon_n^{-2}\bar{V}^n[\phi]\right\} = \frac{\int_{\Omega_x} e^{\frac{1}{\varepsilon_n}F(\tilde{\omega},\beta^n) - \frac{1}{\varepsilon_n^2}\left(\phi(\hat{\xi}(\tilde{\omega})) + H(\hat{\xi}(\tilde{\omega}),\xi^n,v) - \int_0^T h(\hat{\xi}(\tilde{\omega})(s))\cdot(v^n(s)-v(s))ds\right)}\mu^{\varepsilon_n}(d\tilde{\omega})}{\int_{\Omega_x} e^{\frac{1}{\varepsilon_n}F(\tilde{\omega},\beta^n) - \frac{1}{\varepsilon_n^2}\left(H(\hat{\xi}(\tilde{\omega}),\xi^n,v) - \int_0^T h(\hat{\xi}(\tilde{\omega})(s))\cdot(v^n(s)-v(s))ds\right)}\mu^{\varepsilon_n}(d\tilde{\omega})}. \tag{5.5}$$

To prove Lemma 4.1 it now suffices to show that, for all $\phi \in C_b(\mathcal{C}_d)$, as $n \to \infty$,

$$\bar{\Upsilon}^n_1[\phi] \doteq -\varepsilon_n^{-2}\log\left[\int_{\Omega_x} e^{\frac{1}{\varepsilon_n}F(\tilde{\omega},\beta^n) - \frac{1}{\varepsilon_n^2}\left(\phi(\hat{\xi}(\tilde{\omega})) + H(\hat{\xi}(\tilde{\omega}),\xi^n,v) - \int_0^T h(\hat{\xi}(\tilde{\omega})(s))\cdot(v^n(s)-v(s))ds\right)}\mu^{\varepsilon_n}(d\tilde{\omega})\right] \to \inf_{\eta\in\mathcal{C}_d}\left[H(\eta,\xi_0,v) + \phi(\eta) + J(\eta)\right] \text{ a.s. } \mathbb{P}^*. \tag{5.6}$$

Define $\Delta_1^n : \mathcal{C}_d \times \Omega^* \to \mathbb{R}$ as

(5.7)
$$\begin{aligned}\Delta_1^n(\eta) &= H(\eta,\xi_0,v) - H(\eta,\xi^n,v) + \int_0^T h(\eta(s))\cdot(v^n(s)-v(s))ds \\ &= \int_0^T (h(\eta(s))-v(s))\cdot(h(\xi^n(s))-h(\xi_0(s)))\,ds \\ &\quad + \frac{1}{2}\int_0^T \left(\|h(\xi_0(s))\|^2 - \|h(\xi^n(s))\|^2\right)ds + \int_0^T h(\eta(s))\cdot(v^n(s)-v(s))\,ds.\end{aligned}$$

Then from the continuity of h and the a.s. convergence in (5.4), we see that for every $\eta \in \mathcal{C}_d$

$$\text{as } n\to\infty,\ \ \Delta_1^n(\eta)\to 0,\ \text{ a.s. } \mathbb{P}^*. \tag{5.8}$$

Furthermore, with $\Delta^n(\tilde{\omega},\omega^*) \doteq \Delta_1^n(\hat{\xi}(\tilde{\omega}),\omega^*)$,

$$\bar{\Upsilon}_1^n[\phi] = -\varepsilon_n^2 \log\left[\int_{\Omega_x} \exp\left\{\frac{1}{\varepsilon_n}F(\tilde{\omega},\beta^n) - \frac{1}{\varepsilon_n^2}\left(\phi(\hat{\xi}(\tilde{\omega})) + H(\hat{\xi}(\tilde{\omega}),\xi_0,v) - \Delta^n\right)\right\}\mu^{\varepsilon_n}(d\tilde{\omega})\right].$$

In order to prove (5.6) we will show

$$\limsup_{n\to\infty} \bar{\Upsilon}_1^n[\phi] \le \inf_{\eta\in\mathcal{C}_d}\left[H(\eta,\xi_0,v)+\phi(\eta)+J(\eta)\right]\ \text{ a.s. } \mathbb{P}^* \tag{5.9}$$

and

$$\liminf_{n\to\infty} \bar{\Upsilon}_1^n[\phi] \ge \inf_{\eta\in\mathcal{C}_d}\left[H(\eta,\xi_0,v)+\phi(\eta)+J(\eta)\right]\ \text{ a.s. } \mathbb{P}^*. \tag{5.10}$$

The fact that F can be neglected in the asymptotic formula follows the lines of [13]; however, since, unlike [13], we do not assume h is bounded and our functional of interest is different from the one considered in [13], we provide the details.

5.1. Brief outline of the proof of Lemma 4.1. The proof of Lemma 4.1 is long, and so, for the reader's convenience, we provide here an overview of the approach and an outline of the proof. As observed earlier in the section, in order to prove Lemma 4.1 it suffices to show (5.6), for which, in turn, it suffices to show (5.9) and (5.10).

5.1.1. Proving (5.10). This is done in section 5.2. The proof appears towards the end of that section. The first ingredient in its proof is Proposition 5.6, which says that one can ignore the $\frac{1}{\varepsilon_n}F(\tilde{\omega},\beta^n)$ term (O1) when establishing the bound (5.10). In particular, the proposition allows us to estimate the negative of the quantity on the left side of (5.10), namely, $\limsup_{n\to\infty} -\bar{\Upsilon}_1^n[\phi]$, by the sum of the two terms on the right side of (5.26). The second of the two terms is treated using Lemmas 5.7 and 5.8, which allow us to control the contribution from the $\frac{1}{\varepsilon_n^2}\Delta^n$ term (O2S), whereas the first term is treated using Lemma 5.2. Using these results the inequality in (5.10) is obtained readily, as shown at the end of section 5.2. Thus the key steps in the proof of (5.10) are Proposition 5.6, Lemma 5.7, Lemma 5.8, and Lemma 5.2. Proposition 5.6 is based on Proposition 5.5, which in turn is based on Lemmas 5.1, 5.2, and Proposition 5.3.

5.1.2. Proving (5.9). This inequality is proved in section 5.3. The proof relies on Proposition 5.5 (which was also used in the proof of (5.10)) and Lemma 5.9. The proof of Lemma 5.9 relies on Lemma 5.8 from section 5.2. The role of these results is to once more control the terms (O1) and (O2S) in a suitable manner. Using these results, the proof of (5.9) is completed in section 5.3, after the proof of Lemma 5.9.

5.2. Proof of (5.10). We begin with the following lemmas.

LEMMA 5.1. *For* $C \in (0,\infty)$, $\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\mathcal{C}_x} e^{C\varepsilon^{-2}\|\hat{\xi}(\tilde{\omega})\|_*}\mu^\varepsilon(d\tilde{\omega}) < \infty$.

Proof. Note that for $t \in [0,T]$, $\hat{\xi}(t) = x_0 + \int_0^t b(\hat{\xi}(s))ds + \varepsilon \int_0^t \sigma(\hat{\xi}(s))d\hat{\gamma}(s)$.

Let $M(t) \doteq \int_0^t \sigma(\hat{\xi}(s))d\hat{\gamma}(s)$. Then by an application of Gronwall's lemma, it suffices to show that $\limsup_{\varepsilon\to 0} \varepsilon^2 \log E_{\mu^\varepsilon} e^{C\varepsilon^{-1}\|M\|_*} < \infty$ where E_{μ^ε} is the expectation under the probability measure μ^ε. Since σ is bounded and under μ^ε, $\hat{\gamma}$ is a Brownian motion, there is $C_1 \in (0,\infty)$ such that $E_{\mu^\varepsilon} e^{C\varepsilon^{-1}\|M\|_*} \le C_1 e^{C_1\varepsilon^{-2}}$ for every $\varepsilon > 0$. The result follows. ∎

LEMMA 5.2. *Let for* $\varepsilon > 0$, $\bar{\mathcal{R}}^\varepsilon$, $\bar{A}^\varepsilon$ *be measurable maps from* $\mathcal{C}_d$ *to* $\mathbb{R}$ *such that*

$$\sup_{\varepsilon>0} \bar{\mathcal{R}}^\varepsilon(\eta) \le c_R(1+\|\eta\|_*), \quad \sup_{\varepsilon>0} |\bar{A}^\varepsilon(\eta)| \le c_A(1+\|\eta\|_*) \text{ for all } \eta \in \mathcal{C}_d. \tag{5.11}$$

Then

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega}) \le \limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega}) \tag{5.12}$$

and for every $c_0 \in (0,\infty)$

$$\limsup_{M\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\{\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))\ge M\}} e^{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))+\varepsilon^{-2}c_0(1+\|\hat{\xi}(\tilde{\omega})\|_*)}\mu^\varepsilon(d\tilde{\omega}) = -\infty. \tag{5.13}$$

Proof. For $M \in (0,\infty)$, let $A^\varepsilon_M \doteq \bar{A}^\varepsilon \wedge M$. Then

$$\int_{\Omega_x} \exp\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega})$$
$$\le \int_{\Omega_x} \exp\{\varepsilon^{-1}A^\varepsilon_M(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega})$$
$$+ \int_{\Omega_x} \exp\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}1_{\{\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))\ge M\}}\mu^\varepsilon(d\tilde{\omega}).$$

Thus

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega})$$
$$\le \max\Big\{\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-1}A^\varepsilon_M(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega}),$$
$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}1_{\{\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))\ge M\}}\mu^\varepsilon(d\tilde{\omega})\Big\}.$$

Since

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-1}A^\varepsilon_M(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega})$$
$$= \limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\{\varepsilon^{-2}\bar{\mathcal{R}}^\varepsilon(\hat{\xi}(\tilde{\omega}))\}\mu^\varepsilon(d\tilde{\omega}),$$

in order to prove the lemma it suffices to show (5.13) for every $c_0 \in (0,\infty)$. Fix $\varepsilon \in (0,1)$. Using the fact that, on the set $\{\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) \geq M\}$, $\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) \leq \varepsilon^{-2}(\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))-M) + \varepsilon^{-1}M$, and the bound in (5.11), we see that

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \int_{\Omega_x} \exp\left\{\varepsilon^{-1}\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega})) + \varepsilon^{-2}c_0(1+\|\hat{\xi}(\tilde{\omega})\|_*)\right\} 1_{\{\bar{A}^\varepsilon(\hat{\xi}(\tilde{\omega}))\geq M\}}\mu^\varepsilon(d\tilde{\omega})$$
$$\leq -M + \limsup_{\varepsilon\to 0}\varepsilon^2 \log \int_{\Omega_x} \exp\left\{\varepsilon^{-2}(c_A+c_0)(1+\|\hat{\xi}(\tilde{\omega})\|_*)\right\}\mu^\varepsilon(d\tilde{\omega}).$$

Equation (5.13) now follows on applying Lemma 5.1 and taking $M\to\infty$. ∎

Note that, by Itô's formula,

$$\begin{aligned}
F(\tilde{\omega},\beta^n) &= \int_0^T \langle h(\hat{\xi}(\tilde{\omega})(s)), d\beta^n(s)\rangle \\
&= \langle h(\hat{\xi}(\tilde{\omega})(T)), \beta^n(T)\rangle - \sum_{l=1}^m \int_0^T \beta_l^n(s)\langle \nabla h_l(\hat{\xi}(\tilde{\omega})(s)), b(\hat{\xi}(\tilde{\omega})(s))\rangle ds \\
&- \frac{\varepsilon^2}{2}\sum_{i,j=1}^k \sum_{l=1}^m \int_0^T \beta_l^n(s)(\sigma\sigma^\dagger)_{ij}(\hat{\xi}(\tilde{\omega})(s))\frac{\partial^2 h_l}{\partial x_i \partial x_j}(\hat{\xi}(\tilde{\omega})(s))ds \\
&- \sum_{l=1}^m \int_0^T \beta_l^n(s)\left\langle \nabla h_l(\hat{\xi}(\tilde{\omega})(s)), \left(d\hat{\xi}(\tilde{\omega})(s) - b(\hat{\xi}(\tilde{\omega})(s))ds\right)\right\rangle \\
&= A_T(\hat{\xi}(\tilde{\omega}),\beta^n) + K_T(\hat{\xi}(\tilde{\omega}),\beta^n),
\end{aligned}$$

where, $\mathbb{P}^\varepsilon$-a.s., $K_T(\xi,\beta) \doteq -\sum_{l=1}^m \int_0^T \beta_l(s)\langle \nabla h_l(\xi(s)), (d\xi(s) - b(\xi(s))ds)\rangle$ and $A_T(\xi,\beta) \doteq \int_0^T \langle h(\xi(s)), d\beta(s)\rangle - K_T(\xi,\beta)$.

The following result is taken from Heunis [13] (cf. page 940 therein).

Proposition 5.3 (Heunis [13]). *The maps K_T and A_T are measurable and continuous, respectively, from $\mathcal{C}_d \times \mathcal{C}_m$ to $\mathbb{R}$, and there are $c_1, c_2 \in (0,\infty)$ such that for all $x>0$, $n\geq 1$,*

$$\mu^{\varepsilon_n}(\tilde{\omega} : |K_T(\hat{\xi}(\tilde{\omega}),\beta^n)| > x) \leq 2\exp\left\{-c_1\frac{x^2}{\varepsilon_n^2(1+\|\beta^n\|_*^2)}\right\}, \quad a.s.\ \mathbb{P}^*$$

and

$$|A_T(\hat{\xi}(\tilde{\omega}),\beta^n)| \leq c_2(1+\|\hat{\xi}(\tilde{\omega})\|_* + \|\beta^n\|_*) \ a.s.\ \mu^{\varepsilon_n}\otimes\mathbb{P}^*. \tag{5.14}$$

For $(\tilde{\omega},\omega^*) \in \Omega_x \times \Omega^*$, define $G^n(\tilde{\omega},\omega^*) \doteq -\phi(\hat{\xi}(\tilde{\omega})) - H(\hat{\xi}(\tilde{\omega}),\xi_0(\omega^*),v(\omega^*)) + \Delta^n(\tilde{\omega},\omega^*)$.

Remark 5.4. In the rest of this section, ω^* (chosen from a full $\mathbb{P}^*$-measure set) is fixed. We will therefore suppress the dependence on ω^* (wherever it does not cause confusion) to keep the expressions concise. For example, we will write $G^n(\tilde{\omega},\omega^*)$ and $\beta^n(\omega^*)$ as $G^n(\tilde{\omega})$ and β^n, respectively.

Proposition 5.5. *For any $\delta\in(0,\infty)$ and $\mathbb{P}^*$ a.e. ω^**

$$\limsup_{n\to\infty}\varepsilon_n^2 \log\int_{K^+\cup K^-} e^{\varepsilon_n^{-2}G^n(\tilde{\omega})+\varepsilon_n^{-1}F(\tilde{\omega},\beta^n)}\mu^{\varepsilon_n}(d\tilde{\omega}) = -\infty,$$

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^-} e^{\varepsilon_n^{-2} G^n(\tilde\omega) + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)} \mu^{\varepsilon_n}(d\tilde\omega) = -\infty, \tag{5.15}$$

where $K^- \doteq \{\varepsilon_n K_T(\hat\xi(\tilde\omega),\beta^n) < -\delta\}$ *and* $K^+ \doteq \{\varepsilon_n K_T(\hat\xi(\tilde\omega),\beta^n) > \delta\}$.

Proof. Note that, on the set K^-, $\varepsilon_n^{-2} G^n(\tilde\omega) + \varepsilon_n^{-1} F(\tilde\omega,\beta^n) \le \varepsilon_n^{-2}(G^n(\tilde\omega) - \delta) + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)$. Also note that, using the linear growth of h, one can find a measurable map $\theta : \Omega^* \to \mathbb{R}_+$ such that

$$G^n(\tilde\omega) \le \theta(\omega^*)(1 + \|\hat\xi(\tilde\omega)\|_*) \text{ for all } \tilde\omega \in \Omega_x,\ \mathbb{P}^* \text{ a.e. } \omega^*. \tag{5.16}$$

We will write $\theta(\omega^*)$ as θ, from now on. Using these observations, we have

$$\int_{K^-} e^{\varepsilon_n^{-2} G^n(\tilde\omega) + \varepsilon_n^{-1} F(\tilde\omega,\beta^n)} \mu^{\varepsilon_n}(d\tilde\omega) \le e^{\varepsilon_n^{-2}(\theta-\delta)} \int_{K^-} e^{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_* + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)} \mu^{\varepsilon_n}(d\tilde\omega).$$

Next, for every $M \in (0,\infty)$

$$\begin{aligned}
&\int_{K^-} e^{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_* + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)} \mu^{\varepsilon_n}(d\tilde\omega) \\
&\quad \le \int_{K^-} e^{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_* + \varepsilon_n^{-1} M} \mu^{\varepsilon_n}(d\tilde\omega) + \int_{K^- \cap l_M} e^{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_* + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)} \mu^{\varepsilon_n}(d\tilde\omega),
\end{aligned} \tag{5.17}$$

where $l_M \doteq \{A_T(\hat\xi(\tilde\omega),\beta^n) \ge M\}$. We now consider the two terms in the above display separately. For the first term, from the Cauchy–Schwarz inequality,

$$\int_{K^-} \exp\{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_*\} \mu^{\varepsilon_n}(d\tilde\omega) \le \left[\int_{\Omega_x} \exp\{2\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_*\} \mu^{\varepsilon_n}(d\tilde\omega)\right]^{1/2} \left[\mu^{\varepsilon_n}(K^-)\right]^{1/2},$$

and therefore

$$\begin{aligned}
&\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^-} \exp\{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_*\} \mu^{\varepsilon_n}(d\tilde\omega) \\
&\le \limsup_{n\to\infty} \frac{\varepsilon_n^2}{2} \log \int_{\Omega_x} \exp\{2\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_*\} \mu^{\varepsilon_n}(d\tilde\omega) + \limsup_{n\to\infty} \frac{\varepsilon_n^2}{2} \log \mu^{\varepsilon_n}(K^-) \\
&\le \limsup_{n\to\infty} \frac{\varepsilon_n^2}{2} \log \int_{\Omega_x} \exp\{2\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_*\} \mu^{\varepsilon_n}(d\tilde\omega) - c_1 \frac{\delta^2}{2\varepsilon_n^2(1 + \|\beta^n\|_*^2)} = -\infty,
\end{aligned}$$

where in the next to last line we have used Proposition 5.3 and in the last line we have appealed to Lemma 5.1 and the fact that $\sup_n \|\beta^n\|_* < \infty$ $\mathbb{P}^*$-a.s.

For the second term on the right side in (5.17), we have from Lemma 5.2 (see (5.13)) and (5.14) that

$$\limsup_{M\to\infty} \limsup_{n\to\infty} \varepsilon_n^2 \log \int_{l_M} e^{\varepsilon_n^{-2}\theta \|\hat\xi(\tilde\omega)\|_* + \varepsilon_n^{-1} A_T(\hat\xi(\tilde\omega),\beta^n)} \mu^{\varepsilon}(d\tilde\omega) = -\infty.$$

Using the last two displays in (5.17) and combining with (4) we have (5.15) and

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^-} \exp\{\varepsilon_n^{-2} G^n(\tilde\omega) + \varepsilon_n^{-1} F(\tilde\omega,\beta^n)\} \mu^{\varepsilon_n}(d\tilde\omega) = -\infty.$$

Next, from [13, Proposition 4.7], it follows that

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^+} \exp\{2\varepsilon_n^{-1} K_T(\hat{\xi}(\tilde{\omega}), \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) = -\infty.$$

Now using the Cauchy–Schwarz inequality and arguing as before, we see that

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^+} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) = -\infty. \qquad \square$$

Proposition 5.6 below is a useful tool which will be used to see that one can ignore the term involving F in the definition of $\Upsilon_n^1[\varphi]$ when establishing the bound at (5.11).

PROPOSITION 5.6. *For $\mathbb{P}^*$ a.e. ω^*,*

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} e^{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)} \mu^{\varepsilon_n}(d\tilde{\omega}) \le \limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} e^{\varepsilon_n^{-2} G^n(\tilde{\omega})} \mu^{\varepsilon_n}(d\tilde{\omega}).$$

Proof. Fix $\delta \in (0, \infty)$ and with K^+ and K^- defined as earlier, write

$$\begin{aligned}
&\int_{\Omega_x} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) \\
&= \int_{K^+} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) \\
&\quad + \int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}), \beta^n) \le \delta\}} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}).
\end{aligned}$$

From Proposition 5.5,

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^+} e^{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)} \mu^{\varepsilon_n}(d\tilde{\omega}) = -\infty. \tag{5.18}$$

Next note that

$$\begin{aligned}
&\int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}), \beta^n) \le \delta\}} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) \\
&\le \int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}), \beta^n) \le \delta\}} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \delta\varepsilon_n^{-2} + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}).
\end{aligned}$$

Now recalling (5.14) and (5.16) and applying (5.12), we get

$$\begin{aligned}
&\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}), \beta^n) \le \delta\}} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\}\mu^{\varepsilon_n}(d\tilde{\omega}) \\
&\le \delta + \limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} \exp\{\varepsilon_n^{-2} G^n(\tilde{\omega})\}\mu^{\varepsilon_n}(d\tilde{\omega}).
\end{aligned}$$

Since $\delta > 0$ is arbitrary, the result follows on combining the above with (5.18). $\square$

The proof of the following lemma follows along the lines of Varadhan's lemma (cf. [23, Theorem 2.6], [6, Theorem 1.18]). We provide details for the reader's convenience.

LEMMA 5.7. *Let $\{Z^\varepsilon\}_{\varepsilon>0}$ be random variables with values in a Polish space $(\mathcal{X}, d(\cdot,\cdot))$ that satisfies an LDP with rate function J and speed ε^{-2}. Let $\phi : \mathcal{X} \to \mathbb{R}$ be a continuous function bounded from above, namely $\sup_{x\in\mathcal{X}} \phi(x) < \infty$, and let $\{\psi^\varepsilon\}_{\varepsilon>0}$*

be a collection of real measurable maps on $\mathcal{X}$ *such that* $\sup_{\varepsilon>0}\sup_{x\in\mathcal{X}}|\psi^\varepsilon(x)|<\infty$. *Further suppose that for every* $\delta>0$ *and* $x\in\mathcal{X}$, *there exist* $\varepsilon_0(x)$, $\delta_1(x)\in(0,\infty)$ *such that* $|\psi^\varepsilon(y)|<\delta$ *for all* $d(x,y)<\delta_1(x)$ *and all* $0<\varepsilon<\varepsilon_0(x)$. *Then*

$$\lim_{\varepsilon\to 0}\varepsilon^2\log\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)]=\sup_{x\in\mathcal{X}}\left[\phi(x)-J(x)\right].$$

Proof. Define $R\doteq\sup_{x\in\mathcal{X}}(\phi(x)+\sup_{\varepsilon>0}|\psi^\varepsilon(x)|)$, $S\doteq\sup_{x\in\mathcal{X}}(\phi(x)-J(x))$, and $K\doteq\{x\in\mathcal{X}:J(x)\le|S|+R\}$. Since J is a rate function, K is compact in $\mathcal{X}$.

Fix $\delta\in(0,\infty)$. From the hypothesis of the lemma, for each $x\in\mathcal{X}$, there exist $\delta_1(x),\varepsilon_0(x)\in(0,\infty)$ such that $|\psi^\varepsilon(y)|<\delta$ for every $y\in B(x,\delta_1(x))$ and $\varepsilon\in(0,\varepsilon_0(x))$, where $B(z,\gamma)\doteq\{x\in\mathcal{X}:d(x,z)<\gamma\}$ is an open ball of radius γ in $\mathcal{X}$. Also, from the continuity of ϕ, for every $x\in\mathcal{X}$ there exists $\delta_2(x)\in(0,\infty)$ such that $|\phi(x)-\phi(y)|<\delta\ \forall y\in B(x,\delta_2(x))$. Next, from the lower semicontinuity of J, for every $x\in\mathcal{X}$, there exists $\delta_3(x)\in(0,\infty)$ such that

$$J(x)\le\inf_{y\in\overline{B(x,\delta_3(x))}}J(y)+\delta.$$

Let $\bar\delta(x)\doteq\min\{\delta_1(x),\delta_2(x),\delta_3(x)\}$. Now define an open cover $\cup_{x\in K}U(x)$ of K using the following open sets: $U(x)\doteq B(x,\bar\delta(x))$, $x\in K$.

Note that for any $x\in K$, $y\in U(x)$, and $\varepsilon<\varepsilon_0(x)$, we have

$$|\psi^\varepsilon(y)|<\delta,\ |\phi(x)-\phi(y)|<\delta,\text{ and } J(x)\le\inf_{z\in\overline{U(x)}}J(z)+\delta. \tag{5.19}$$

Since K is compact, there exist $N\in\mathbb{N}$ and $\{x_i\}_{i=1}^N\subset K$ such that $\{U_i\doteq U(x_i)\}_{i=1}^N$ cover K. For $i=1,\ldots,N$, we can find $0<\varepsilon(x_i)\le\varepsilon_0(x_i)$ such that with $\bar\varepsilon_0\doteq\min_{i=1,\ldots,N}\varepsilon(x_i)$, for every $\varepsilon<\bar\varepsilon_0$,

$$\mathbb{P}[Z^\varepsilon\in\overline{U_i}]\le\exp\left[\varepsilon^{-2}(-b_i+\delta)\right],\quad\mathbb{P}[Z^\varepsilon\in F]\le\exp\left[\varepsilon^{-2}(-\inf_{x\in F}J(x)+\delta)\right], \tag{5.20}$$

where $F\doteq\left(\cup_{i=1}^N U_i\right)^c$ and $b_i\doteq\inf_{x\in\overline{U_i}}J(x)$. Next note that

$$\begin{aligned}\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)]&\le\sum_{i=1}^N\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)1_{U_i}(Z^\varepsilon)]\\&\quad+\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)1_F(Z^\varepsilon)].\end{aligned}\tag{5.21}$$

Let $a_i\doteq\inf_{x\in\overline{U_i}}\phi(x)$. Then $|a_i-\phi(x)|<2\delta$ for $x\in U_i$. Thus, using (5.19) and (5.20)

$$\begin{aligned}&\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)1_{U_i}(Z^\varepsilon)]\\&\le(a_i-b_i+4\delta)\le\phi(x_i)-J(x_i)+5\delta\le\sup_{x\in\mathcal{X}}\left[\phi(x)-J(x)\right]+5\delta.\end{aligned}\tag{5.22}$$

Also

$$\begin{aligned}&\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)1_F]\le R-\inf_{x\in F}J(x)+\delta\\&\le-|S|+\delta\le\sup_{x\in\mathcal{X}}\left[\phi(x)-J(x)\right]+\delta,\end{aligned}\tag{5.23}$$

where the second inequality is a consequence of the observation that $F \subset K^c$. Since $\delta > 0$ is arbitrary, using (5.22) and (5.23) in (5.21) we now see that

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{E}[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)] \le \sup_{x\in\mathcal{X}}\left[\phi(x)-J(x)\right]. \tag{5.24}$$

For the lower bound, choose x_0 such that $\phi(x_0) - J(x_0) \ge S - \delta$. Let $\delta(x_0), \varepsilon(x_0) \in (0,\infty)$ be such that for all $x \in U \doteq B(x_0, \delta(x_0))$, $|\phi(x)-\phi(x_0)| < \delta$ and $|\psi^\varepsilon(x)| < \delta$ for $\varepsilon < \varepsilon(x_0)$. Then

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{E}\left[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)\right]\\
&\ge \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{E}\left[\exp\left(\varepsilon^{-2}\left\{\phi(Z^\varepsilon)+\psi^\varepsilon(Z^\varepsilon)\right\}\right)1_U(Z^\varepsilon)\right]\\
&\ge \phi(x_0) - 2\delta + \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbb{P}\left[Z^\varepsilon \in U\right]\\
&\ge \phi(x_0) - 2\delta - \inf_{x\in U} J(x) \ge \phi(x_0) - 2\delta - J(x_0) \ge \sup_{x\in\mathcal{X}}\left[\phi(x)-J(x)\right] - 3\delta.
\end{aligned}$$

Sending $\delta \to 0$, we have the lower bound, and combining with (5.24) the result follows. ∎

Recall the definition of Δ_1^n from (5.7). The term involving $\Delta^{n,Q}$ on the right side of the estimate in (5.26) below will be handled via an application of Lemma 5.7 (see (5.27)) by arguing from Lemma 5.8 that the term $\Delta^{n,Q}$ has the properties of the map ψ^ε (with $\varepsilon = 1/n$) in Lemma 5.7.

Lemma 5.8. *For $\mathbb{P}^*$ a.e. ω^* and every $\delta \in (0,\infty)$ and $\eta \in \mathcal{C}_d$ there exist $n_0 \in \mathbb{N}$ and $\delta_1 \in (0,\infty)$ such that $|\Delta_1^n(\tilde\eta,\omega^*)| < \delta$ whenever $\tilde\eta \in \mathcal{C}_d$, $\|\eta-\tilde\eta\|_* \le \delta_1$, and $n \ge n_0$.*

Proof. Consider ω^* in the set of full measure on which the convergence in (5.4) (and thus in (5.8)) holds. From (5.8), for any fixed $\delta \in (0,\infty)$ and $\eta \in \mathcal{C}_d$, we can find $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$, $|\Delta_1^n(\eta,\omega^*)| \le \frac{\delta}{2}$. Also, from continuity of h, we can find a $\delta_1 \in (0,\infty)$ such that for all $\tilde\eta \in \mathcal{C}_d$ with $\|\eta-\tilde\eta\|_* \le \delta_1$

$$\begin{aligned}
&\sup_{n\in\mathbb{N}} \int_0^T \|h(\eta(s)) - h(\tilde\eta(s))\|(\|h(\xi^n(s))\| + \|h(\xi_0(s))\|)ds \le \frac{\delta}{4},\\
&\sup_{n\in\mathbb{N}} \int_0^T \|h(\eta(s)) - h(\tilde\eta(s))\|(\|v^n(s)\| + \|v(s)\|)ds \le \frac{\delta}{4}.
\end{aligned}$$

Thus for all $n \ge n_0$ and $\tilde\eta \in \mathcal{C}_d$ with $\|\eta-\tilde\eta\|_* \le \delta_1$

$$\begin{aligned}
|\Delta_1^n(\tilde\eta)| &\le |\Delta_1^n(\tilde\eta) - \Delta_1^n(\eta)| + |\Delta_1^n(\eta)|\\
&\le \int_0^T \|h(\eta(s)) - h(\tilde\eta(s))\|(\|h(\xi^n(s))\| + \|h(\xi_0(s))\|)ds\\
&\quad + \int_0^T \|h(\eta(s)) - h(\tilde\eta(s))\|(\|v^n(s)\| + \|v(s)\|)ds + \frac{\delta}{2} \le \delta.
\end{aligned}$$
∎

We now complete the proof of (5.10).

Completing the proof of (5.10). Note that, from Proposition 5.6, $\mathbb{P}^*$-a.s.,

$$\begin{aligned}
\limsup_{n\to\infty} -\bar\Upsilon_1^n[\phi] &= \limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} \exp\{\varepsilon_n^{-2}G^n(\tilde\omega) + \varepsilon_n^{-1}F(\tilde\omega,\beta^n)\}\mu^{\varepsilon_n}(d\tilde\omega)\\
&\le \limsup_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} \exp\{\varepsilon_n^{-2}G^n(\tilde\omega)\}\mu^{\varepsilon_n}(d\tilde\omega).
\end{aligned} \tag{5.25}$$

For $Q \in (0,\infty)$, let $\Delta^{n,Q} \doteq (\Delta^n \wedge Q) \vee (-Q)$. We will again suppress the dependence on ω^*. Then

$$\begin{aligned}\int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde\omega)\right\}\mu^{\varepsilon_n}(d\tilde\omega) &\le \int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde\omega)\right\} 1_{\{|\Delta^n|\ge Q\}}\mu^{\varepsilon_n}(d\tilde\omega)\\ &+ \int_{\Omega_x} \exp\{\varepsilon_n^{-2}\big(-\phi(\hat\xi(\tilde\omega)) - H(\hat\xi(\tilde\omega),\xi_0,v) + \Delta^{n,Q}(\omega)\big)\}\mu^{\varepsilon_n}(d\tilde\omega).\end{aligned} \tag{5.26}$$

In order to treat the second term on the right side above, we will use Lemmas 5.7 and 5.8. Lemma 5.7 will be applied with ε replaced with ε_n and random variables Z^{ε_n} replaced by X^{ε_n} that are distributed as $\mu^{\varepsilon_n}\circ\hat\xi^{-1}$. We will use the result in (2.11) which gives an LDP for $\{X^{\varepsilon_n}\}$ (equivalently an LDP for the sequence of measures $\{\mu^{\varepsilon_n}\}$) with rate function J. The role of ϕ in Lemma 5.7 will be played by the map $-\phi(\cdot) - H(\cdot,\xi_0,v))$ for a given ω^*. Note that this is a continuous map on $\mathcal{C}_d$ which is bounded from above. Also the role of ψ^{ε_n} in Lemma 5.7 is played by the map $\eta \mapsto (\Delta_1^n(\eta)\wedge Q)\vee(-Q) \doteq \Delta_1^{n,Q}(\eta)$, which clearly satisfies $\sup_{n\ge1}\sup_{\eta\in\mathcal{C}_d}\Delta_1^{n,Q}(\eta) < \infty$, and by Lemma 5.8, for every $\delta\in(0,\infty)$ and $\eta\in\mathcal{C}_d$ there exist $n_0\in\mathbb{N}$ and $\delta_1\in(0,\infty)$ such that $|(\Delta_1^n(\tilde\eta)\wedge Q)\vee(-Q)| < \delta$ whenever $\tilde\eta\in\mathcal{C}_d$, $\|\eta-\tilde\eta\|_* \le \delta_1$, and $n\ge n_0$. Combining these observations, we now obtain from Lemma 5.7 that, for $\mathbb{P}^*$ a.e. ω^*

$$\begin{aligned}\limsup_{n\to\infty}\varepsilon_n^2\log\int_{\Omega_x}\exp\left\{\varepsilon_n^{-2}\left(-\phi(\hat\xi(\tilde\omega)) - H(\hat\xi(\tilde\omega),\xi_0,v) + \Delta^{n,Q}(\omega)\right)\right\}\mu^{\varepsilon_n}(d\tilde\omega)\\ \le -\inf_{\eta\in\mathcal{C}_d}\left[H(\eta,\xi_0,v)+\phi(\eta)+J(\eta)\right].\end{aligned} \tag{5.27}$$

Next, using the linear growth property of h $\sup_n|\Delta_1^n(\eta)| \le c_\Delta(\omega^*)(1+\|\eta\|_*)$, $\mathbb{P}^*$-a.s. for some measurable map $c_\Delta : \Omega^* \to \mathbb{R}_+$. Thus, using the boundedness of ϕ and the nonnegativity of H, we have

$$\begin{aligned}&\limsup_{Q\to\infty}\limsup_{n\to\infty}\varepsilon_n^2\log\int_{\Omega_x} e^{\varepsilon_n^{-2}G^n(\tilde\omega)}1_{\{|\Delta^n|\ge Q\}}\mu^{\varepsilon_n}(d\tilde\omega)\\ &\le \limsup_{Q\to\infty}\limsup_{n\to\infty}\varepsilon_n^2\log\int_{\Omega_x} e^{\varepsilon_n^{-2}(c_\Delta+\|\phi\|_\infty)(1+\|\hat\xi(\tilde\omega)\|_*)}1_{\{c_\Delta(1+\|\hat\xi(\tilde\omega)\|_*)\ge Q\}}\mu^{\varepsilon_n}(d\tilde\omega) = -\infty,\end{aligned}$$

where the last equality follows from Lemma 5.2 (see (5.13)). Using the last bound together with (5.27) in (5.26) and (5.25), we now have the inequality in (5.10).

5.3. Proof of (5.9). Recall (5.4). We begin with the following lemma.

LEMMA 5.9. *For $\mathbb{P}^*$ a.e. ω^**

$$\liminf_{n\to\infty}\varepsilon_n^2\log\int_{\Omega_x} e^{\varepsilon_n^{-2}G^n(\tilde\omega)+\varepsilon_n^{-1}A_T(\hat\xi(\tilde\omega),\beta^n)}\mu^{\varepsilon}(d\tilde\omega) \ge -\inf_{\eta\in\mathcal{C}_d}\left[H(\eta,\xi_0,v)+\phi(\eta)+J(\eta)\right].$$

Proof. Fix $\eta_0\in\mathcal{C}_d$ and $\delta\in(0,\infty)$. From continuity of ϕ on $\mathcal{C}_d$, of A_T on $\mathcal{C}_d\times\mathcal{C}_m$, and of $\eta\mapsto H(\eta,\xi_0,v)$ (for $\mathbb{P}^*$ a.e. ω^*) on $\mathcal{C}_d$, a.s. convergence of β^n to β, and Lemma 5.8, we can find, for $\mathbb{P}^*$ a.e. ω^*, a neighborhood G of η_0 and $n_1\in\mathbb{N}$ such that $\sup_{\tilde\eta\in G}|\Delta_1^n(\tilde\eta)| < \delta$ for all $n\ge n_1$,

$$\begin{aligned}\inf_{\tilde\eta\in G}A_T(\tilde\eta,\beta^n) &\ge A_T(\eta_0,\beta^n) - \delta \text{ for all } n\ge n_1,\\ \inf_{\tilde\eta\in G}[-\phi(\tilde\eta) - H(\tilde\eta,\xi_0,v] &\ge [-\phi(\eta_0) - H(\eta_0,\xi_0,v] - \delta.\end{aligned}$$

Observe that

$$\int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\right\} \mu^\varepsilon(d\tilde{\omega})$$
$$\geq \exp\left\{\varepsilon_n^{-2}[-\phi(\eta_0) - H(\eta_0, \xi_0, v - 2\delta)] + \varepsilon_n^{-1}(A_T(\eta_0, \beta^n) - \delta)\right\} \mu^\varepsilon(G),$$

Noting that $\sup_n |A_T(\eta_0, \beta^n)| < \infty$ $\mathbb{P}^*$-a.s. and applying the large deviation result from (2.11), we now have

$$\liminf_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\right\} \mu^\varepsilon(d\tilde{\omega})$$
$$\geq [-\phi(\eta_0) - H(\eta_0, \xi_0, v - 2\delta)] - \inf_{\tilde{\eta}\in G} J(\tilde{\eta})$$
$$\geq -\phi(\eta_0) - H(\eta_0, \xi_0, v - J(\eta_0)) - 2\delta.$$

Since $\delta \in (0, \infty)$ and $\eta_0 \in \mathcal{C}_d$ are arbitrary, the result follows. ∎

We now complete the proof of (5.9).

Completing the proof of (5.9). Fix $\delta \in (0, \infty)$. Then with K^- defined as earlier

$$\int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\right\} \mu^{\varepsilon_n}(d\tilde{\omega})$$
$$\geq \int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}),\beta^n)\geq -\delta\}} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} F(\tilde{\omega}, \beta^n)\right\} \mu^{\varepsilon_n}(d\tilde{\omega})$$
$$\geq \int_{\{\varepsilon_n K_T(\hat{\xi}(\tilde{\omega}),\beta^n)\geq -\delta\}} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1}(-\delta\varepsilon_n^{-1} + A_T(\hat{\xi}(\tilde{\omega}), \beta^n))\right\} \mu^{\varepsilon_n}(d\tilde{\omega})$$
$$= \int_{\Omega_x} \exp\left\{\varepsilon_n^{-2}(G^n(\tilde{\omega}) - \delta) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\right\} \mu^{\varepsilon_n}(d\tilde{\omega})$$
$$- \int_{K^-} \exp\left\{\varepsilon_n^{-2}(G^n(\tilde{\omega}) - \delta) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\right\} \mu^{\varepsilon_n}(d\tilde{\omega}).$$

From Proposition 5.5 (see (5.15))

$$\limsup_{n\to\infty} \varepsilon_n^2 \log \int_{K^-} e^{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)} \mu^{\varepsilon_n}(d\tilde{\omega}) = -\infty.$$

Thus to prove (5.9) it suffices to show that, $\mathbb{P}^*$-a.s.,

$$\liminf_{n\to\infty} \varepsilon_n^2 \log \int_{\Omega_x} \exp\left\{\varepsilon_n^{-2} G^n(\tilde{\omega}) + \varepsilon_n^{-1} A_T(\hat{\xi}(\tilde{\omega}), \beta^n)\right\} \mu^{\varepsilon_n}(d\tilde{\omega})$$
$$\geq -\inf_{\eta\in\mathcal{C}_d} [H(\eta, \xi_0, v) + \phi(\eta) + J(\eta)].$$

However, the above is an immediate consequence of Lemma 5.9. This proves (5.9).

Finally we complete the proof of Lemma 4.1.

Completing the proof of Lemma 4.1. As noted above (5.6), in order to prove Lemma 4.1 it suffices to show (5.6) for every $\phi \in C_b(\mathcal{C}_d)$. Also, for this it is enough to show (5.10) and (5.9). The inequality in (5.10) was shown in section 5.2, and the proof of the inequality in (5.9) was provided in section 5.3. Combining these we have Lemma 4.1.

6. Proof of Theorem 2.1. In order to prove the theorem it suffices to show (3.1) and (3.2). Proofs of (3.2) and (3.1) are in sections 6.1 and 6.2, respectively.

6.1. Proof of (3.2). Let $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be a sequence of positive reals such that $\varepsilon_n \to 0$ as $n \to \infty$. To show (3.2) it suffices to show that for every $G \in C_b(\mathbb{R})$

$$\liminf_{n\to\infty} -\varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] \geq \inf_{z\in\mathbb{R}} \{G(z) + I^\phi(z)\}, \tag{6.1}$$

$$\limsup_{n\to\infty} -\varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] \leq \inf_{z\in\mathbb{R}} \{G(z) + I^\phi(z)\}. \tag{6.2}$$

We begin with (6.1). Fix $\delta \in (0,1)$ and choose $(\tilde{u}_n, \tilde{v}_n) \in \mathcal{A}^k \times \mathcal{A}^m$ such that

$$\begin{aligned} &- \varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] \\ &\geq \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[G(\bar{V}^{\varepsilon_n,\tilde{u}_n,\tilde{v}_n}[\phi]) + \frac{1}{2}\int_0^T (\|\tilde{u}_n(s)\|^2 + \|\tilde{v}_n(s)\|^2) ds\right] - \delta. \end{aligned} \tag{6.3}$$

Note that

$$\sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\int_0^T (\|\tilde{u}_n(s)\|^2 + \|\tilde{v}_n(s)\|^2) ds\right] \leq 2(2\|G\|_\infty + 1) \doteq c_G. \tag{6.4}$$

Using a standard localization argument to modify $(\tilde{u}_n, \tilde{v}_n)$ (cf. [6, Theorem 3.17]) we can find $M \in (0,\infty)$ and $(u_n, v_n) \in \mathcal{A}^k \times \mathcal{A}^m$ that take values in S_M a.s. such that

$$\begin{aligned} &- \varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] \\ &\geq \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[G(\bar{V}^{\varepsilon_n,u_n,v_n}[\phi]) + \frac{1}{2}\int_0^T (\|u_n(s)\|^2 + \|v_n(s)\|^2) ds\right] - 2\delta. \end{aligned} \tag{6.5}$$

Note that $\{(u_n, v_n)\}$ is a sequence of S_M valued random variables, and since S_M is weakly compact, every subsequence of $\{(u_n, v_n)\}$ has a weakly convergent subsubsequence. It suffices to show (6.1) along such a subsubsequence which we denote once more as $\{n\}$. Denoting the limit as (u, v), given on some probability space $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$, we have from Lemma 4.1 that, as $n \to \infty$, $\bar{V}^{\varepsilon_n,u_n,v_n}[\phi] \to V_0^{u,v}[\phi]$, in distribution. Using the fact that $G \in C_b(\mathbb{R})$ and Fatou's lemma, we now have

$$\begin{aligned} &\liminf_{n\to\infty} \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[G(\bar{V}^{\varepsilon_n,u_n,v_n}[\phi]) + \frac{1}{2}\int_0^T (\|u_n(s)\|^2 + \|v_n(s)\|^2) ds\right] \\ &\geq \mathbb{E}_{\mathbb{P}^0} \left[G(V_0^{u,v}[\phi]) + \frac{1}{2}\int_0^T (\|u(s)\|^2 + \|v(s)\|^2) ds\right] \\ &\geq \mathbb{E}_{\mathbb{P}^0} \left[G(V_0^{u,v}[\phi]) + I^\phi(V_0^{u,v}[\phi])\right] \geq \inf_{z\in\mathbb{R}} [G(z) + I^\phi(z)], \end{aligned}$$

where the second inequality uses that, by definition, $(u, v) \in \mathcal{S}(V_0^{u,v}[\phi])$ a.s. Combining the above display with (6.5) and recalling that $\delta > 0$ is arbitrary, we have (6.1).

We now give the proof of (6.2). Fix $\delta \in (0,1)$ and let $z^* \in \mathbb{R}$ be such that

$$G(z^*) + I^\phi(z^*) \leq \inf_{z\in\mathbb{R}} [G(z) + I^\phi(z)] + \delta. \tag{6.6}$$

Now choose $(\varphi, \psi) \in \mathcal{S}(z^*)$ such that

$$\frac{1}{2}\int_0^T \|\varphi(t)\|^2 dt + \frac{1}{2}\int_0^T \|\psi(t)\|^2 dt \le I^\phi(z^*) + \delta. \tag{6.7}$$

Since $(\varphi, \psi) \in \mathcal{A}_k \times \mathcal{A}_m$ (as they are nonrandom and square-integrable), we have from (3.6) that, for every $n \in \mathbb{N}$,

$$\begin{aligned} &- \varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] \\ &\le \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[G(\bar{V}^{\varepsilon_n,\varphi,\psi}[\phi]) + \frac{1}{2}\int_0^T (\|\varphi(s)\|^2 + \|\psi(s)\|^2) ds\right]. \end{aligned} \tag{6.8}$$

Also, from Lemma 4.1, as $n \to \infty$, $\bar{V}^{\varepsilon_n,\varphi,\psi}[\phi] \to V_0^{\phi,\psi}[\phi]$, in distribution. Since $(\varphi, \psi) \in \mathcal{S}(z^*)$, (2.17) holds with z replaced with z^* and so $V_0^{\phi,\psi}[\phi] = z^*$. Thus sending $n \to \infty$ in (6.8), we have

$$\begin{aligned} \limsup_{n\to\infty} - \varepsilon_n^2 \log \mathbb{E}_{\mathbb{P}^{\varepsilon_n}} \left[\exp\left\{-\varepsilon_n^{-2} G(V^{\varepsilon_n}[\phi])\right\}\right] &\le G(z^*) + \frac{1}{2}\int_0^T (\|\varphi(s)\|^2 + \|\psi(s)\|^2) ds \\ &\le G(z^*) + I^\phi(z^*) + \delta \le \inf_{z\in\mathbb{R}} [G(z) + I^\phi(z)] + 2\delta, \end{aligned}$$

where the second inequality uses (6.7) while the third uses (6.6). Since $\delta > 0$ is arbitrary, we have (6.2), and, together with (6.1), this completes the proof of (3.2).

6.2. Proof of (3.1). Fix $\phi \in C_b(\mathcal{C}_d)$ and $M \in (0, \infty)$. Consider the set $\{z \in \mathbb{R} : I^\phi(z) \le M\} \doteq E_M$ and let $\{z_n\}_{n\in\mathbb{N}}$ be a sequence in this set. Since for each $n \in \mathbb{N}$, $I^\phi(z_n) \le M$, we can find $(\varphi_n, \psi_n) \in \mathcal{S}(z_n) \subset \mathcal{L}_k^2 \times \mathcal{L}_m^2$ such that

$$\frac{1}{2}\int_0^T (\|\varphi_n(s)\|^2 + \|\psi_n(s)\|^2) ds \le M + \frac{1}{n}. \tag{6.9}$$

Since $(\varphi_n, \psi_n) \in \mathcal{S}(z_n)$,

$$z_n \doteq V_0^{\varphi_n,\psi_n}[\phi] = \inf_{\eta\in\mathcal{C}_d} [H(\eta, \xi_0^{\varphi_n}, \psi_n) + \phi(\eta) + J(\eta)] - \inf_{\eta\in\mathcal{C}_d} [H(\eta, \xi_0^{\varphi_n}, \psi_n) + J(\eta)]. \tag{6.10}$$

Note that we can write

$$\begin{aligned} H(\eta, \xi_0^{\varphi_n}, \psi_n) &= \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi_0^{\varphi_n}(s)) - \psi_n(s)\|^2 ds \\ &= \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi_0^{\varphi_n}(s))\|^2 ds - \int_0^T [h(\eta(s)) - h(\xi_0^{\varphi_n}(s))] \cdot \psi_n(s) ds \\ &\quad + \frac{1}{2}\int_0^T \|\psi_n(s)\|^2 ds = \tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) + \frac{1}{2}\int_0^T \|\psi_n(s)\|^2 ds, \end{aligned}$$

where for $\eta, \tilde{\eta} \in \mathcal{C}_d$ and $\psi \in \mathcal{L}_m^2$

$$\tilde{H}(\eta, \tilde{\eta}, \psi) \doteq \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\tilde{\eta}(s))\|^2 ds - \int_0^T [h(\eta(s)) - h(\tilde{\eta}(s))] \cdot \psi(s) ds.$$

From (6.10) and the relation between H and $\tilde{H}$ it follows that

$$z_n = \inf_{\eta\in\mathcal{C}_d} \left[\tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) + \phi(\eta) + J(\eta)\right] - \inf_{\eta\in\mathcal{C}_d} \left[\tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) + J(\eta)\right]. \tag{6.11}$$

Also, from (6.9) it follows that $\{(\varphi_n, \psi_n)\}_{n\in\mathbb{N}} \subset S_{2(M+1)}$. Since $S_{2(M+1)}$ is compact, we can find a subsequence along which (φ_n, ψ_n) converges to some $(\varphi, \psi) \in S_{2(M+1)}$. In fact, from (6.9) and lower semicontinuity, $(\varphi, \psi) \in S_{2M}$. Define

(6.12)
$$\begin{aligned} z^* \doteq V_0^{\varphi,\psi}[\phi] &= \inf_{\eta\in\mathcal{C}_d}[H(\eta, \xi_0^\varphi, \psi) + \phi(\eta) + J(\eta)] - \inf_{\eta\in\mathcal{C}_d}[H(\eta, \xi_0^\varphi, \psi) + J(\eta)] \\ &= \inf_{\eta\in\mathcal{C}_d}\left[\tilde{H}(\eta, \xi_0^\varphi, \psi) + \phi(\eta) + J(\eta)\right] - \inf_{\eta\in\mathcal{C}_d}\left[\tilde{H}(\eta, \xi_0^\varphi, \psi) + J(\eta)\right]. \end{aligned}$$

In order to complete the proof of (3.1) it suffices to show that

$$\text{as } n \to \infty, \quad z_n \to z^*. \tag{6.13}$$

We first argue that in the infimum appearing in (the second line of) (6.12) and (6.11), $\{\eta \in \mathcal{C}_d\}$ can be replaced by $\{\eta \in K\}$ for some fixed compact set K. To see this, note that, with ξ^* as in (1.4),

$$\inf_{\eta\in\mathcal{C}_d}\left[\tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) + \phi(\eta) + J(\eta)\right] \le \tilde{H}(\xi^*, \xi_0^{\varphi_n}, \psi_n) + \|\phi\|_\infty + J(\xi^*).$$

Also, note that $J(\xi^*) = 0$ and

$$\begin{aligned} \tilde{H}(\xi^*, \xi_0^{\varphi_n}, \psi_n) &= \frac{1}{2}\int_0^T \|h(\xi^*(s)) - h(\xi_0^{\varphi_n}(s))\|^2 ds - \int_0^T [h(\xi^*(s)) - h(\xi_0^{\varphi_n}(s))] \cdot \psi_n(s) ds \\ &\le \int_0^T \|h(\xi^*(s)) - h(\xi_0^{\varphi_n}(s))\|^2 + \frac{1}{2}\int_0^T \|\psi_n(s)\|^2 ds \\ &\le 2T\|h(\xi^*(\cdot))\|_*^2 + 2\int_0^T \|h(\xi_0^{\varphi_n}(s))\|^2 ds + \frac{1}{2}\int_0^T \|\psi_n(s)\|^2 ds \\ &\le 2T\|h(\xi^*(\cdot))\|_*^2 + \kappa_1(M+1) \doteq \kappa_2, \end{aligned}$$

where $\kappa_1 \in (0, \infty)$ depends only on x_0, T, and the linear growth coefficients of h, b, σ. Thus, taking $\kappa_3 \doteq \kappa_2 + \|\phi\|_\infty + 1$, we see that the first infimum in (6.11) can be replaced by the infimum over the set $K_0^n \doteq \{\eta \in \mathcal{C}_d : \tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) + \phi(\eta) + J(\eta) \le \kappa_3\}$.

Using the relation $a \cdot b \ge -\frac{1}{4}\|a\|^2 - \|b\|^2$,

$$\begin{aligned} \tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) \ge \frac{1}{2}\int_0^T &\|h(\eta(s)) - h(\xi_0^{\varphi_n}(s))\|^2 \\ &- \frac{1}{4}\int_0^T \|h(\eta(s)) - h(\xi_0^{\varphi_n}(s))\|^2 - \int_0^T \|\psi_n(s)\|^2 ds \ge -2M. \end{aligned}$$

Thus, with $\kappa_4 \doteq \kappa_3 + \|\phi\|_\infty + 1 + 2M$, K_0^n is contained in the compact set $K \doteq \{\eta \in \mathcal{C}_d : J(\eta) \le \kappa_4\}$. Thus the first infimum in (6.11) can be replaced by the infimum over the set K. Similarly, the second infimum in (6.11) and both infima in (the second line of) (6.12) can be replaced by infima over the same compact set K. Note that if B_n, B are maps from $K \to \mathbb{R}$ such that $B_n \to B$ uniformly on compact sets, then $\inf_{\eta\in K}[B_n(\eta) + J(\eta)] \to \inf_{\eta\in K}[B(\eta) + J(\eta)]$. Thus, to complete the proof of (6.13) it suffices to show that

$$\text{as } n \to \infty, \quad \tilde{H}(\eta, \xi_0^{\varphi_n}, \psi_n) \to \tilde{H}(\eta, \xi_0^\varphi, \psi), \quad \text{uniformly for } \eta \in K. \tag{6.14}$$

For this note that from Assumption 1 and the convergence of $\varphi_n \to \varphi$ it follows that $\xi_0^{\varphi_n} \to \xi_0^\varphi$ in $\mathcal{C}_d$ as $n \to \infty$. Also, since K is compact, $\sup_{\eta\in K}\|\eta\|_* < \infty$. Combining these observations with the continuity and linear growth of h we have that, as $n \to \infty$,

$$\frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi_0^{\varphi_n}(s))\|^2 ds \to \frac{1}{2}\int_0^T \|h(\eta(s)) - h(\xi_0^{\varphi}(s))\|^2 ds \tag{6.15}$$

uniformly for $\eta \in K$. Also, writing

$$\int_0^T h(\xi_0^{\varphi_n}(s)) \cdot \psi_n(s) ds = \int_0^T [h(\xi_0^{\varphi_n}(s)) - h(\xi_0^{\varphi}(s))] \cdot \psi_n(s) ds + \int_0^T h(\xi_0^{\varphi}(s)) \cdot \psi_n(s) ds$$

and using the convergence $(\xi_0^{\varphi_n}, \psi_n) \to (\xi_0^{\varphi}, \psi)$, the bound in (6.9), and the Lipschitz property of h, we have that, as $n \to \infty$,

$$\int_0^T h(\xi_0^{\varphi_n}(s)) \cdot \psi_n(s) ds \to \int_0^T h(\xi_0^{\varphi}(s)) \cdot \psi(s) ds. \tag{6.16}$$

Finally, we claim that, as $n \to \infty$,

$$\int_0^T h(\eta(s)) \cdot \psi^n(s) ds \to \int_0^T h(\eta(s)) \cdot \psi(s) ds, \tag{6.17}$$

uniformly for $\eta \in K$. To show the claim, it suffices to show that if $\eta^n \to \eta$ in K, then

$$\int_0^T h(\eta^n(s)) \cdot \psi^n(s) ds \to \int_0^T h(\eta(s)) \cdot \psi(s) ds. \tag{6.18}$$

Write the right-hand side as

$$\int_0^T h(\eta^n(s)) \cdot \psi^n(s) ds = \int_0^T (h(\eta^n(s)) - h(\eta(s))) \cdot \psi^n(s) ds + \int_0^T h(\eta(s)) \cdot \psi^n(s) ds.$$

The convergence in (6.18) is now immediate from the above display on using the Lipschitz property of h, the bound in (6.9), and the convergence of (η^n, ψ^n) to (η, ψ), which proves the claim. Combining the convergence properties in (6.15), (6.16), and (6.17), we now have the statement in (6.14), which, as noted previously, proves (3.1).

Acknowledgments. Part of this research was carried out when AB was visiting the International Centre for Theoretical Sciences-TIFR and Imperial College, London, and he thanks his hosts Amit Apte and Dan Crisan, respectively, for their generous hospitality.

REFERENCES

[1] G. B. Arous and F. Castell, *Flow decomposition and large deviations*, J. Funct. Anal., 140 (1996), pp. 23–67.

[2] J. S. Baras, A. Bensoussan, and M. R. James, *Dynamic observers as asymptotic limits of recursive filters: Special cases*, SIAM J. Appl. Math., 48 (1988), pp. 1147–1158, https://doi.org/10.1137/0148068.

[3] A. Berkaoui, B. Djehiche, and Y. Ouknine, *Sur les grandes déviations en théorie de filtrage non linéaire*, Studia Math., 148 (2001), pp. 5–21.

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

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

[6] A. Budhiraja and P. Dupuis, *Analysis and Approximation of Rare Events: Representations and Weak Convergence Methods*, Probab. Theory Stoch. Model. 94, Springer, New York, 2019.

[7] A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen, *Data assimilation in the geosciences: An overview of methods, issues, and perspectives*, Wiley Interdiscip. Rev. Climate Change, 9 (2018), e535.

[8] J. Clark, *The design of robust approximations to the stochastic differential equations of nonlinear filtering*, in Communcation Systems and Random Process Theory, NATO Advanced Study Inst. Ser. E Appl. Sci. 25, Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, 1978, pp. 721–734.

[9] M. H. Davis, *On a multiplicative functional transformation arising in nonlinear filtering theory*, Z. Wahrsch. Verw. Gebiete, 54 (1980), pp. 125–139.

[10] M. H. Davis, *A pathwise solution of the equations of nonlinear filtering*, Theory Probab. Appl., 27 (1982), pp. 167–175.

[11] H. Doss, *Un nouveau principe de grandes déviations en théorie du filtrage non linéaire*, Ann. Inst. H. Poincaré Probab. Statist., 27 (1991), pp. 407–423.

[12] S. J. Fletcher, *Data Assimilation for the Geosciences: From Theory to Application*, Elsevier, New York, 2017.

[13] A. J. Heunis, *Nonlinear filtering of rare events with large signal-to-noise ratio*, J. Appl. Probab., 24 (1987), pp. 929–948.

[14] O. Hijab, *Minimum Energy Estimation*, Ph.D. thesis, University of California, Berkeley, Berkeley, CA, 1980.

[15] O. Hijab, *Asymptotic Bayesian estimation of a first order equation with small diffusion*, Ann. Probab., 12 (1984), pp. 890–902.

[16] M. R. James and J. S. Baras, *Nonlinear filtering and large deviations: A PDE-control theoretic approach*, Stochastics, 23 (1988), pp. 391–412.

[17] G. Kallianpur, *Stochastic Filtering Theory*, Appl. Math. 13, Springer, New York, 1980.

[18] V. Maroulas, X. Pan, and J. Xiong, *Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise*, Stochastic Process. Appl., 130 (2020), pp. 203–231.

[19] V. Maroulas and J. Xiong, *Large deviations for optimal filtering with fractional Brownian motion*, Stochastic Process. Appl., 123 (2013), pp. 2340–2352.

[20] R. Mortensen, *Maximum-likelihood recursive nonlinear filtering*, J. Optim. Theory Appl., 2 (1968), pp. 386–394.

[21] É. Pardoux and O. Zeitouni, *Quenched large deviations for one dimensional nonlinear filtering*, SIAM J. Control Optim., 43 (2005), pp. 1272–1297, https://doi.org/10.1137/S0363012903365032.

[22] T. J. Rabeherimanana, *Petites Perturbations de Systèmes Dynamiques et Algèbres de Lie Nilpotentes*, Ph.D. thesis, Université Sorbonne Paris Cité, Paris, France, 1992.

[23] D. W. Stroock, *An Introduction to the Theory of Large Deviations*, Universitext, Springer, New York, 1984.

[24] J. Xiong, *Large deviation principle for optimal filtering*, Appl. Math. Optim., 47 (2003), pp. 151–155.