

MODERATE DEVIATION PRINCIPLE FOR MULTISCALE SYSTEMS DRIVEN BY FRACTIONAL BROWNIAN MOTION

S. BOURGUIN, T. DANG, AND K. SPILIOPOULOS

ABSTRACT. In this paper we study the moderate deviations principle (MDP) for slow-fast stochastic dynamical systems where the slow motion is governed by small fractional Brownian motion (fBm) with Hurst parameter $H \in (1/2, 1)$. We derive conditions on the moderate deviations scaling and on the Hurst parameter H under which the MDP holds. In addition, we show that in typical situations the resulting action functional is discontinuous in H at $H = 1/2$, suggesting that the tail behavior of stochastic dynamical systems perturbed by fBm can have different characteristics than the tail behavior of such systems that are perturbed by standard Brownian motion.

1. INTRODUCTION

The goal of this paper is to study the asymptotic behavior, in the moderate deviations regime, of the following system of slow-fast dynamics

$$(1) \quad \begin{aligned} dX_t^\epsilon &= g(X_t^\epsilon, Y_t^\epsilon)dt + \sqrt{\epsilon}f(X_t^\epsilon, Y_t^\epsilon)dW_t^H, & X_0^\epsilon &= x_0 \\ dY_t^\epsilon &= \frac{1}{\epsilon}c(Y_t^\epsilon) + \frac{1}{\sqrt{\epsilon}}\sigma(Y_t^\epsilon)dB_t, & Y_0^\epsilon &= y_0. \end{aligned}$$

Here ϵ is a small parameter that goes to zero. We assume that $t \in [0, 1]$ and $(X^\epsilon, Y^\epsilon) \in \mathbb{R}^n \times \mathbb{R}^d$. Also, B is a standard m -dimensional Brownian motion, while W^H is a p -dimensional fractional Brownian motion (fBm) with Hurst parameter $H \in (1/2, 1)$ independent of B . As is known, if $H = 1/2$ then $W^{1/2}$ will be a standard Brownian motion. Moreover, the integral with respect to W^H is a pathwise Riemann–Stieltjes integral and is commonly known as a Young integral (see Appendix A for a brief introduction).

Since, $1/\epsilon \uparrow \infty$ as $\epsilon \downarrow 0$, we expect that under the appropriate conditions, the distribution of Y^ϵ will be converging to its invariant distribution, while the equation that X^ϵ satisfies can be viewed as a perturbation of a dynamical system by small multiplicative noise of magnitude $\sqrt{\epsilon}$. We can think of X^ϵ as the slow component and of Y^ϵ as the fast component. Model (1) is a prototypical dynamical system that exhibits multiple characteristic scales in time and is perturbed by small noise to account for imperfect information and to capture random phenomena. Such systems arise naturally as models in a great variety of applied fields, including physics, chemistry, biology, neuroscience, meteorology, and mathematical finance, to name a few.

The novelty of this paper lies in the consideration of the tail behavior of (1) in the case where $H \neq 1/2$. In the case of $H = 1/2$, i.e., when both the X^ϵ and Y^ϵ components are driven by Brownian motions, the asymptotic behavior of systems like (1) have been extensively studied in the literature. We refer the interested reader to [BF77, DS12, Fre78, FS99, FW12, Gui03, KL99, MS17, LS90, PV01, PV05, Spi14, Spi13], which contain results on related typical averaging dynamics, central limit theorems, moderate and large deviations. Choosing the noise that perturbs the system to be standard Brownian motion, we embed the Markov property and semimartingale structure of the standard Brownian motion in the

BOSTON UNIVERSITY, DEPARTMENT OF MATHEMATICS AND STATISTICS, 111 CUMMINGTON MALL, BOSTON, MA 02215, USA

E-mail address: bourguin@math.bu.edu, ycloud77@bu.edu, kspiliop@math.bu.edu.

2010 *Mathematics Subject Classification.* 60F10, 60G22, 60H10, 60H07.

Key words and phrases. Fractional Brownian motion, multiscale processes, small noise, moderate deviations.

S.B. was partially supported by the Simons Foundation Award 635136. K.S. was partially supported by the National Science Foundation (DMS 1550918, DMS 2107856) and Simons Foundation Award 672441.

system. However, many physical dynamical system exhibit long-range dependence or a particular sort of self-similarity that may not be amenable to accurate description by a model driven by standard Brownian noise.

One way to account for this issue, is to perturb the dynamical system by fractional Brownian motion. Such practice has seen growing interest in literature, for example the references [BFG⁺19, BS20, Che03, CR98a, Fuk17, FZ17a, GJRS18, HJL19, SV03b] to name a few. However, the corresponding literature for multiscale systems like (1) in the case of perturbation by fractional Brownian motion is quite sparse and still in its infancy. We refer the interested reader to the very recent papers of [BGS21, HL20, PIX20] for results concerning typical averaging behavior, homogenization, and fluctuations corrections for multiscale models like (1) under different sets of assumptions for the model coefficients. As discussed in these papers, replacing Brownian motion by fractional Brownian motion creates a number of issues that need to be overcome. These issues are mainly related to the partial loss of the Markovian structure as well as to the proper averaging of the integral with respect to the fractional Brownian motion W^H which originates from the interaction of ergodicity and fBm.

The intent of this paper is to study the tail behavior of X^ϵ in (1) as $\epsilon \downarrow 0$ in the moderate deviation setting. To be more precise, letting $h(\epsilon) \rightarrow \infty$ such that $\sqrt{\epsilon}h(\epsilon) \rightarrow 0$ and defining $\bar{X}_t = \lim_{\epsilon \rightarrow 0} X_t^\epsilon$ (the limit in the appropriate sense), we define the moderate deviations process

$$(2) \quad \eta_t^\epsilon = \frac{X_t^\epsilon - \bar{X}_t}{\sqrt{\epsilon}h(\epsilon)}.$$

Moderate deviations for X^ϵ refers to large deviations for η^ϵ . In fact the scaling by $\sqrt{\epsilon}h(\epsilon)$ implies that moderate deviations is in the regime between central limit theorem (corresponding to the choice $h(\epsilon) = 1$) and large deviations (corresponding to the choice $h(\epsilon) = 1/\sqrt{\epsilon}$). Moderate deviations for systems like (1) and for $H = 1/2$, i.e., when both slow and fast components are driven by standard Brownian motions, have been considered in [Gui03, MS17]. An interesting conclusion of our results for the case $H \neq 1/2$, which will be discussed in Remark 7 is that the resulting action functional is not continuous in H at $H = 1/2$.

In order to study the moderate deviations principle for X^ϵ , we shall follow the weak convergence method of [DE11]. The core of this approach lies in the use of a variational representation of exponential functionals of the driving noise (W^H, B) , see [DE11, Zha09]. In our case, such a representation leads to a representation of the exponential functional of the moderate deviations process η^ϵ that appears in the Laplace principle (which is equivalent to the moderate deviations) as a variational infimum of a family of controlled moderate deviations processes $\eta^{\epsilon, w^\epsilon}$ together with a quadratic cost over a suitable family of stochastic controls w^ϵ . To be more precise, letting a be a bounded Borel function on $C([0, 1]; \mathbb{R}^n)$, we have the representation

$$(3) \quad -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] = \inf_{w^\epsilon \in \mathcal{S}} \mathbb{E}\left[\frac{1}{2} \|w^\epsilon\|_{\mathcal{S}}^2 + a(\eta^{\epsilon, w^\epsilon})\right],$$

where \mathcal{S} denotes the Cameron-Martin space associated with the process $\{(W_t^H, B_t) : t \in [0, 1]\}$ (see (42)) and the controlled deviations process $\eta^{\epsilon, w^\epsilon}$ is defined by

$$(4) \quad \eta_t^{\epsilon, w^\epsilon} = \frac{X_t^{\epsilon, w^\epsilon} - \bar{X}_t}{\sqrt{\epsilon}h(\epsilon)}$$

with the controlled processes $(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon})$ defined by (14).

Essentially, proving the moderate deviations principle for X^ϵ amounts to finding the limit as $\epsilon \rightarrow 0$ to (3). When $H \neq 1/2$, i.e., when the standard Brownian motion in the slow component is replaced by fBm, a number of additional technical issues come up and the standard methodology needs to be modified. After we introduce proper notation, we explain in Remark 9 of Section 3 one of the core ideas that allow us to study the $H \neq 1/2$ case in a way that naturally extends the $H = 1/2$ setting.

The rest of the paper is organized as follows. In Section 2 we establish necessary notation, go over our assumptions and present the main result, Theorem 1, on the moderate deviations principle with

an explicit representation of the action functional, as well as a corollary of the aforementioned theorem. Section 3 contains the details of the weak convergence approach for the problem at hand, introduces the appropriate controlled processes and presents Theorem 2 which has a variational representation of the moderate deviations action functional. Theorem 1 can be viewed as a direct consequence of Theorem 2. In Section 3 we also go over one of the main ideas that essentially unlock the computation for the case $H \neq 1/2$, in a way that naturally extends the standard $H = 1/2$ framework, see Remark 9. Section 4 contains examples that demonstrate our theoretical results.

Section 5 contains the proof of Theorem 2 and consequently of Theorem 1 as well. In particular, in Section 5 we prove tightness of the appropriate controlled processes and occupational measures introduced in Section 3, we identify their weak limit which then allows to prove the limit Laplace principle lower and the upper bound of (3). The proof of the Laplace upper bound leads to the exact representation of Theorem 1. Next, Section 6 provides the proof of Corollary 1. Section 7 discusses avenues for future work on this topic.

Appendix A recalls several aspects of fBm and necessary results on pathwise stochastic integration with respect to fBm used in this paper. In Appendix A, we also discuss the Cameron-Martin space of fBm and prove associated results that are potentially of independent interest as well. Appendix B recalls regularity results of [PV01, PV05] on Poisson equations, proves necessary a-priori bounds of the slow-fast process (X^ϵ, Y^ϵ) in (1) as well as necessary a-priori estimates on $\eta^{\epsilon, w^\epsilon}$ from (4) that allows to establish the necessary tightness results in Section 3.

2. NOTATION, CONDITIONS AND MAIN RESULTS

In this section, we introduce some notation, present the main assumptions we make, and state our main results. We work with a canonical probability space (Ω, \mathcal{F}, P) equipped with a filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ satisfying the usual conditions (namely, $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is right continuous and \mathcal{F}_0 contains all P -negligible sets).

We will denote by $A : B$ the Frobenius inner product $\Sigma_{i,j}[a_{i,j} \cdot b_{i,j}]$ of matrices $A = (a_{i,j})$ and $B = (b_{i,j})$. We will use single bars $|\cdot|$ to denote the Frobenius (or Euclidean) norm of a matrix and double bars $\|\cdot\|$ to denote the operator norm. For $\alpha \in (0, 1)$, $|\cdot|_\alpha$ is the standard Hölder semi-norm, i.e.

$$|h|_\alpha = \sup_{0 \leq s \neq t \leq 1} \frac{|h(s) - h(t)|}{|s - t|}.$$

In addition, for a given sets A, B and $i, j \in \mathbb{N}$ and $\zeta > 0$, $\mathcal{C}^{i,j+\zeta}(A \times B)$ is the space of functions with i bounded derivatives in x and j derivatives in y , with all partial derivatives being ζ -Hölder continuous with respect to y , uniformly in x .

2.1. Conditions. We start by stating the assumptions we make on the coefficients of Y^ϵ ensuring its ergodicity. Note that these assumptions are satisfied in the context of the multi-scale models studied in [BGS21, Theorem 1] and [HL20, Theorem A].

Condition H1.

- $c(y) = -\Gamma y + \zeta(y)$, for which Γ be a $d \times d$ positive matrix with bounded entries and $\zeta(y)$ is a uniformly Lipschitz function with Lipschitz coefficient L_ζ . Moreover, $\zeta(y) \leq C|y|$ and $\langle (\Gamma - L_\zeta I)\xi, \xi \rangle \geq \gamma_0 |\xi|^2$ for some $\gamma_0 > 0$.
- $c(y), \sigma(y)$ have first and second derivatives that are α -Hölder continuous for some $\alpha > 0$.
- $\sigma(y)\sigma^\top(y)$ is uniformly continuous, bounded and non-degenerate.
- There are positive constants β_1, β_2 such that $0 < \beta_1 \leq \frac{\langle \sigma(y)\sigma^\top(y)y, y \rangle}{|y|^2} \leq \beta_2, \forall y \in \mathbb{R}^d \setminus \{0\}$.

Remark 1. Condition H1 guarantees that the fast process Y^ϵ has a unique invariant measure, which we denote by $\mu(dy)$ in the sequel.

Denote by \mathcal{L} the normalized infinitesimal generator of the fast motion Y^ϵ (with respect to which averaging is being performed). It is given by

$$(5) \quad \mathcal{L}F(y) = \nabla_y F(y)^\top c(y) + \frac{1}{2}\sigma(y)\sigma^\top(y) : \nabla_y^2 F(y).$$

Set $\mathcal{Y} = \mathbb{R}^d$. For any function $G(x, y)$, define the averaged function \bar{G} by

$$\bar{G}(x) = \int_{\mathcal{Y}} G(x, y)\mu(dy).$$

In particular, the averaging of the drift term g in the slow motion X^ϵ with respect to μ will be given by

$$\bar{g}(x) = \int_{\mathcal{Y}} g(x, y)\mu(dy).$$

Remark 2. Under the growth assumption on g and its derivatives in either the upcoming Condition H2-A or H2-B, Theorem 4 implies that the partial differential equation

$$(6) \quad \mathcal{L}\phi(x, y) = g(x, y) - \bar{g}(x), \quad \int_{\mathcal{Y}} \phi(x, y)\mu(dy) = 0$$

has a unique, twice differentiable solution (that we denote by $\phi(x, y)$ in the sequel) in the class of functions that grow at most polynomially in $|y|$.

Finally, we provide two different sets of assumptions on the coefficients of X^ϵ , each of which is based on the available averaging results for X^ϵ appearing in [BGS21] and [HL20], respectively. Depending on the specific multi-scale model at hand, one may choose to work with one set of assumptions or the other.

Condition H2-A. The assumptions below relate to the setting of [HL20].

- $f(x, y)$ and $g(x, y)$ are uniformly bounded with bounded first and second partial derivatives.
- There exists β in $[0, 1]$ such that $\beta + H > 1$ and $h(\epsilon)^{-1}\epsilon^{-\frac{\beta}{2}} \rightarrow 0$ as $\epsilon \rightarrow 0$.

Condition H2-B. The assumptions below relate to the setting of [BGS21]. We assume that there are constants D_f, D_g, M_f, M_k in $[0, 1]$ and α in $(0, 1]$ such that

- $g \in \mathcal{C}^{2,\alpha}(\mathbb{R}^n, \mathcal{Y})$.
- $f = f(y)$ and g satisfy the growth assumption $|f(y)| \leq C(1 + |y|^{D_f})$ and $|g(x, y) + \nabla_x g(x, y) + \nabla_x^2 g(x, y)| \leq C(1 + |y|^{D_g})$.
- D_f and D_g are related via $0 \leq D_f + D_g < 1$.
- $f(y)$ and $\nabla_x \phi(x, y)f(y)$ are respectively M_f and M_k -Hölder continuous, where $\phi(x, y)$ is defined at (6). Moreover, we have $\min\left\{\frac{M_f}{2} + H, \frac{M_k}{2} + H\right\} > 1$.
- $h(\epsilon)^{-1}\epsilon^{-\frac{M_f}{2}} \rightarrow 0$ as $\epsilon \rightarrow 0$.

Remark 3. Conditions H2-A and H2-B relate to the averaging results in [HL20] and [BGS21] that state that the slow motion X^ϵ converges in probability, as ϵ goes to 0, to a deterministic limit \bar{X} defined to be the solution to the integral equation

$$d\bar{X}_t = \bar{g}(\bar{X}_t)dt, \quad \bar{X}_0 = x_0.$$

In addition, we need to assume uniqueness of a strong solution. Without having to refer to it again, this assumption is always in effect in this paper.

Condition H3. The stochastic differential equation at (1) has a unique strong solution.

Remark 4. We direct readers to [GN08, MS11, dSEE18] for existence and uniqueness of solutions to stochastic differential equations like (1).

Finally, define the operator $Q_{\bar{X}}^H$ by

$$(7) \quad Q_{\bar{X}}^H = \bar{f}(\bar{X}) \dot{K}_H \left(\bar{f}(\bar{X}) \dot{K}_H \right)^* + \int_{\mathcal{Y}} \nabla_y \phi(\bar{X}, y) \sigma(y) (\nabla_y \phi(\bar{X}, y) \sigma(y))^{\top} \mu(dy),$$

where μ is the invariant measure defined in Remark 1, \dot{K}_H is the operator (related to the fractional Brownian motion) defined in (41) (see Appendix A.3). Let $h \in L^2([0, 1]; \mathbb{R}^n)$ then the operator $\bar{f}(\bar{X}) \dot{K}_H \left(\bar{f}(\bar{X}) \dot{K}_H \right)^*$ admits the explicit representation

$$\begin{aligned} \left[\bar{f}(\bar{X}) \dot{K}_H \left(\bar{f}(\bar{X}) \dot{K}_H \right)^* h \right] (t) &= c_H^2 \bar{f}(\bar{X}_t) t^{H-1/2} \\ &\quad \int_0^t (t-z)^{H-3/2} z^{1-2H} \int_z^1 (s-z)^{H-3/2} s^{H-1/2} \bar{f}(\bar{X}_s)^{\top} h(s) ds dz \end{aligned}$$

such that the constant c_H equals $(H(2H-1)/\beta(2-2H, H-1/2))^{1/2}$, where $\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the standard beta function.

Remark 5. In both this paper and the paper [BGS21], the latter of which provides averaging results on multiscale models like (1), one needs to bound terms that are Young integrals. However, each paper uses a different bounding technique which leads to different assumptions on (1). For instance, the authors of [BGS21] use the maximal inequality in their Lemma 1 to bound

$$(8) \quad \int_0^t f(Y_s^{\epsilon}) dW_s^H,$$

an integral term which appears in (1). Having the kernel $f(Y^{\epsilon})$ independent from the driving process W^H simplifies the application of the maximal inequality and yields the necessary bound on the integral (8). For this paper, we instead have to bound

$$(9) \quad \int_0^t f(Y_s^{\epsilon, w^{\epsilon}}) dW_s^H,$$

an integral term which appears in (14). In this case, the control process w^{ϵ} is dependent on W^H , implying the kernel $f(Y_s^{\epsilon, w^{\epsilon}})$ is dependent on W^H as well. This lack of independence between the kernel and the driving process leads us to substitute the Young-Loéve inequality for the maximal inequality in order to bound (9). The Young-Loéve inequality for Young integrals requires some kind of uniform Hölder continuity of the kernel, which explains why we impose certain uniform Hölder continuity condition on the coefficients of (1), an assumption not made in [BGS21].

2.2. Main results. The weak convergence approach to large deviations developed in [DE11] states that the large deviations principle for η_t^{ϵ} is equivalent to the Laplace principle which states that for any bounded continuous function $a: C([0, 1]; \mathbb{R}^n) \rightarrow \mathbb{R}$, there exists a rate function (also called action functional) $S^H: C([0, 1]; \mathbb{R}^n) \rightarrow \mathbb{R}$ that satisfies

$$\lim_{\epsilon \rightarrow \infty} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta_t^{\epsilon}))] = \inf_{\Phi \in C([0, 1]; \mathbb{R}^n)} \{S^H(\Phi) + a(\Phi)\}.$$

In this paper, we prove that the above Laplace principle holds and our main result, Theorem 1, identifies that rate function $S^H(\Phi)$ explicitly. The statement of this theorem is given below.

Theorem 1. *Let Conditions H1 and either H2-A or H2-B be satisfied. Moreover, assume that the operator $Q_{\bar{X}}^H$ defined in (7) is invertible. Then, the process $\{X^{\epsilon}: \epsilon > 0\}$ satisfies the moderate deviations principle, with the action functional $S^H(\Phi)$ given by*

$$S^H(\Phi) = \int_0^1 \left(\dot{\Phi}_s - \nabla_x \bar{g}(\bar{X}_s) \Phi_s \right)^{\top} (Q_{\bar{X}_s}^H)^{-1} \left(\dot{\Phi}_s - \nabla_x \bar{g}(\bar{X}_s) \Phi_s \right) ds$$

if $\Phi \in C([0, 1]; \mathbb{R}^n)$ is absolutely continuous, and ∞ otherwise.

Remark 6. A sufficient condition for invertibility of $Q_{\bar{X}}^H$ is that for all x and non-zero $z \in \mathbb{R}^n$ and for all $y \in \mathbb{R}^d$, we have

$$(10) \quad \left\langle \int_{\mathcal{Y}} \nabla_y \phi(x, y) \sigma(y) (\nabla_y \phi(x, y) \sigma(y))^{\top} \mu(dy) z, z \right\rangle > 0.$$

The sufficiency of this condition is established in Lemma 9. In most situations, Condition (10) proves easier to verify than the invertibility of $Q_{\bar{X}}^H$ itself.

Remark 7. Let us now briefly compare the results in the $H = 1/2$ and $H \neq 1/2$ case. As can be seen from the results of [MS17], when $H = 1/2$, i.e., when the slow motion in (1) is driven by standard Brownian motion, the corresponding MDP is as in Theorem 1 but with $Q_{\bar{X}}^H$ defined in (7) replaced by

$$(11) \quad Q_{\bar{X}}^{1/2} = \int_{\mathcal{Y}} f(\bar{X}, y) f(\bar{X}, y)^{\top} \mu(dy) + \int_{\mathcal{Y}} \nabla_y \phi(\bar{X}, y) \sigma(y) (\nabla_y \phi(\bar{X}, y) \sigma(y))^{\top} \mu(dy).$$

It is interesting to note that the mapping $H \mapsto Q_{\bar{X}}^H$ is not, in general, continuous in H at $H = 1/2$. Indeed, if $H = 1/2$, then the discussion of Appendix A.3 shows that $\dot{K}_{1/2}$ the operator defined in (41), will be the identity operator, so one would actually expect that

$$\lim_{H \rightarrow 1/2} Q_{\bar{X}}^H = \int_{\mathcal{Y}} f(\bar{X}, y) \mu(dy) \int_{\mathcal{Y}} f(\bar{X}, y)^{\top} \mu(dy) + \int_{\mathcal{Y}} \nabla_y \phi(\bar{X}, y) \sigma(y) (\nabla_y \phi(\bar{X}, y) \sigma(y))^{\top} \mu(dy).$$

which is of course different from (11). Hence, we indeed have, unless of course $f(x, y) = f(x)$, that

$$Q_{\bar{X}}^{1/2} \neq \lim_{H \rightarrow 1/2} Q_{\bar{X}}^H.$$

This result also immediately says that in general there is no continuity of the mapping $H \mapsto S^H$ at $H = 1/2$.

The lack of continuity does not come as a surprise. It is related to the fact that when averaging integrals with respect to fBm with $H \in (1/2, 1)$, then one averages the integrand directly as opposed to the quadratic variation which is what happens when $H = 1/2$. We refer the interested reader to the recent papers [BGS21, HL20, LS20] for further discussion on this.

In certain circumstances, we can provide an explicit formula for $(Q_{\bar{X}}^H)^{-1}$ as the following corollary of Theorem 1 shows. The proof will be presented in Section 6.

Corollary 1. *Let Conditions H1 and either H2-A or H2-B be satisfied. Assume further that $1/2 < H < 3/4$, $g = g(x)$ and $f(\bar{X})$ is an invertible square matrix with a bounded inverse denoted by L . Then $Q_{\bar{X}}^H$ is invertible and the process $\{X^{\epsilon} : \epsilon > 0\}$ satisfies the moderate deviations principle as in Theorem 1. In particular, given $\Psi \in L^2([0, 1]; \mathbb{R}^n)$, $(Q_{\bar{X}}^H)^{-1}$ has the explicit form*

$$(Q_{\bar{X}}^H)^{-1} \Psi = c_H^{-2} \Gamma(H - 1/2)^{-2} L^{\top} t^{1/2-H} D_{1-}^{H-1/2} t^{2H-1} D_{0+}^{H-1/2} t^{1/2-H} L \Psi$$

or equivalently,

$$\begin{aligned} ((Q_{\bar{X}}^H)^{-1} \Psi)(t) &= c_H^{-2} \Gamma(H - 1/2)^{-2} \Gamma(3/2 - H)^{-2} L_t^{\top} t^{1/2-H} \\ &\quad \frac{d}{dt} \left[\int_t^1 (z-t)^{1/2-H} z^{2H-1} \frac{d}{dz} \left[\int_0^z (z-s)^{1/2-H} s^{1/2-H} L_s \Psi_s ds \right] dz \right]. \end{aligned}$$

3. THE CONTROLLED PROCESSES

The proof of the Laplace principle is based on a variational formula established in [Zha09, Theorem 3.2], which can be regarded as an abstract Wiener space counterpart of the stochastic control representation from [BD98, Theorem 3.1] for the classical Wiener space. Recall that \mathcal{S} denotes the Cameron-Martin space associated with the process $\{(W_t^H, B_t) : t \in [0, 1]\}$ defined in (42). Let a be a bounded Borel function on $C([0, 1]; \mathbb{R}^n)$. Then, the variational formula (applied to the framework of this paper) from [Zha09, Theorem 3.2] states that

$$\begin{aligned} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] &= \inf_{w^\epsilon \in \mathcal{S}} \mathbb{E}\left[\frac{1}{2} \|w^\epsilon\|_{\mathcal{S}} + a(\eta^{\epsilon, w^\epsilon})\right] \\ (12) \quad &= \inf_{w^\epsilon = (K_H \hat{u}^\epsilon, K_{1/2} \hat{v}^\epsilon) \in \mathcal{S}} \mathbb{E}\left[\frac{1}{2} \int_0^1 |\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 ds + a(\eta^{\epsilon, w^\epsilon})\right], \end{aligned}$$

where the controlled deviations process $\eta^{\epsilon, w^\epsilon}$ is defined by

$$(13) \quad \eta_t^{\epsilon, w^\epsilon} = \frac{X_t^{\epsilon, w^\epsilon} - \bar{X}_t}{\sqrt{\epsilon} h(\epsilon)}$$

and the controlled processes X^{ϵ, w^ϵ} and Y^{ϵ, w^ϵ} are defined by

$$\begin{aligned} (14) \quad X_t^{\epsilon, w^\epsilon} &= x_0 + \int_0^t g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) + \sqrt{\epsilon} h(\epsilon) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \\ &\quad + \int_0^t \sqrt{\epsilon} f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \\ Y_t^{\epsilon, w^\epsilon} &= y_0 + \int_0^t \frac{1}{\epsilon} c(Y_s^{\epsilon, w^\epsilon}) + \frac{h(\epsilon)}{\sqrt{\epsilon}} \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds + \int_0^t \frac{1}{\sqrt{\epsilon}} \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s. \end{aligned}$$

Note that, based on (13) and (14), we can rewrite $\eta^{\epsilon, w^\epsilon}$ in the form

$$\begin{aligned} (15) \quad \eta_t^{\epsilon, w^\epsilon} &= \int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} \left[g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s) \right] ds + \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \\ &\quad + \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H. \end{aligned}$$

Let $\mathcal{U} = \mathbb{R}^m$ and $\mathcal{V} = \mathbb{R}^p$. These are the spaces in which the control processes u^ϵ and v^ϵ take value in, respectively. Define $\theta(x, \eta, y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r) : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{U} \times \mathcal{U} \times \mathcal{V} \times \mathcal{V} \times [0, 1] \times [0, 1]$ by

$$(16) \quad \begin{aligned} \theta(x, \eta, y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r) &= (\nabla_y \phi(x, y^{(1)}) \sigma(y^{(1)}) v^{(1)} + \nabla_x \bar{g}(x) \eta) \\ &\quad + c_H f(x, y^{(1)})(s - r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbb{1}_{[0, s]}(r). \end{aligned}$$

Condition H1 and Theorem 4 guarantee that the function θ is bounded in x , affine in η , $u^{(2)}$ and $v^{(1)}$ and bounded polynomially in $|y|$.

Next, we introduce the occupation measure P^ϵ . Let A_1, A_2, B and Γ be Borel sets of $\mathcal{U}, \mathcal{V}, \mathcal{Y} = \mathbb{R}^d$ and $[0, 1]$, respectively. Let $(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon})$ solve (14). Associate with $(Y^{\epsilon, w^\epsilon}, \hat{u}^\epsilon, \dot{v}^\epsilon)$ a family of occupation measures P^ϵ defined by

$$(17) \quad P^\epsilon(A_1 \times A_2 \times B \times \Gamma) = \int_{\Gamma} \mathbb{1}_{A_1}(\hat{u}_s^\epsilon) \mathbb{1}_{A_2}(\dot{v}_s^\epsilon) \mathbb{1}_B(Y_s^{\epsilon, w^\epsilon}) ds.$$

Definition 1. Let $F(x, \eta, y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r) : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{U} \times \mathcal{U} \times \mathcal{V} \times \mathcal{V} \times [0, 1] \times [0, 1]$ be a function that has at most polynomial growth in $|y|$. Let \mathcal{L} be a second order elliptic partial differential operator and denote its domain by $\mathcal{D}(\mathcal{L})$. A pair $(\psi, P) \in C([0, 1]; \mathbb{R}^n) \times \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$ is called a viable pair with respect to (θ, \mathcal{L}) if

- The function $\psi \in C([0, 1]; \mathbb{R}^n)$ is absolutely continuous.

- The measure P is integrable in the sense that

$$\int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2 + |y|^2] P(dudvdyds) < \infty.$$

- For all $t \in [0, 1]$,

$$\begin{aligned} \psi_t = \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,t]^2} F(\bar{X}_s, \psi_s, y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r) \\ P \otimes P(du^{(1)}du^{(2)}dv^{(1)}dv^{(2)}dy^{(1)}dy^{(2)}dsdr). \end{aligned}$$

- For all $t \in [0, 1]$, it holds that

$$(18) \quad P(dudvdydt) = \nu_{y,t}(dudv)\mu(dy)dt,$$

where $\nu_{y,t}$ is a kernel on $\mathcal{U} \times \mathcal{V}$ dependent on $y \in \mathcal{Y}$ and $t \in [0, 1]$, while μ is the unique invariant measure associated with the operator \mathcal{L} .

In order to indicate that the pair (ψ, P) is viable with respect to (F, \mathcal{L}) , we write $(\psi, P) \in \mathcal{V}(F, \mathcal{L})$.

The controlled process (13) and the definition of viable pairs (Definition 1) will be used to prove the theorem below.

Theorem 2. *Let Conditions H1 and either H2-A or H2-B be satisfied. Then, the process $\{X^\epsilon: \epsilon > 0\}$ from (1) satisfies the moderate deviations principle, with the action functional $S^H(\Phi)$ given by*

$$(19) \quad S^H(\Phi) = \inf_{(\Phi, P) \in \mathcal{V}(\theta, \mathcal{L})} \left[\frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] P(dudvdyds) \right]$$

with the convention that the infimum over the empty set is ∞ .

Remark 8. As will be shown in the proof of Theorem 2, Theorem 1 follows directly from Theorem 2.

Remark 9. In this remark we discuss one of the key ideas that allows to naturally generalize the computations to the $H \neq 1/2$ case. In the course of the proof, we will need to handle terms of the form $\int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds$, where u^ϵ is the control process introduced in the beginning of this section. Roughly speaking, if $H = 1/2$ and P^ϵ is the occupational measure defined as in (17), then one has $\dot{u}^\epsilon = \hat{u}^\epsilon$ and thus

$$\int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds = \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,t]} f(X_s^{\epsilon, w^\epsilon}, y) u^\epsilon P^\epsilon(dudvdyds),$$

and then after establishing tightness of $(X^{\epsilon, w^\epsilon}, P^\epsilon)$ one can study its limit. This approach does not work exactly like that in the case where the Hurst parameter $H \neq 1/2$. In order to generalize this idea for the case $H \neq 1/2$, we first notice that one can write that

$$\frac{d}{ds} u_s^\epsilon = \frac{d}{ds} [K_H \hat{u}^\epsilon]_s,$$

where K_H is the operator associated to fBm, see Appendix A.3. With this observation at hand we then write

$$\begin{aligned} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds &= \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \frac{d}{ds} [K_H \hat{u}^\epsilon]_s ds \\ &= \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \left(c_H s^{H-1/2} \int_0^s (s-r)^{H-3/2} r^{1/2-H} \hat{u}_r^\epsilon dr \right) ds \\ &= c_H \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} f(X_s^{\epsilon, w^\epsilon}, y^{(1)}) (s-r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbb{1}_{[0,t]}(s) \mathbb{1}_{[0,s]}(r) \\ &\quad P^\epsilon \otimes P^\epsilon(du^{(1)}du^{(2)}dv^{(1)}dv^{(2)}dy^{(1)}dy^{(2)}dsdr), \end{aligned}$$

(20)

which is what allows us then to take limits. The details are in Section 5.

4. EXAMPLES

4.1. Fractional volatility model. In [Shi99, Chapter 4], the author collects various empirical studies which observe persistence or long memory phenomena in financial data such as financial indexes and currency cross rates, among others. This motivates us, for the first example of this paper, to consider the multiscale volatility model

$$(21) \quad \begin{aligned} dX_t^\epsilon &= Y_t^\epsilon dt + \sqrt{\epsilon} \tau dW_t^H, \\ dY_t^\epsilon &= \beta(\theta - Y_t^\epsilon) dt + v \sqrt{Y_t} dB_t. \end{aligned}$$

We assume $\tau > 0$ and β, θ, v are real constants such that $2\beta\theta \geq v^2$. (W^H, B) is an independent pair of one-dimensional fractional Brownian motion of Hurst parameter $H > 1/2$ and one-dimensional Brownian motion. X^ϵ is a (re-scaled) log pricing of a security with a perturbed fractional Brownian noise (in order to simulate the long memory effect). The fast volatility process Y^ϵ follows the Cox–Ingersoll–Ross model of interest rate and the assumption $2\beta\theta \geq v^2$ ensures that Y^ϵ is strictly positive. It is also worth noting that adding fractional Brownian noise to financial models to simulate long memory has been an increasingly common practice in literature, see [Che03, CR98b, FZ17b, HJL19, SV03a].

The stochastic differential equation of Y^ϵ in (21) has the Cox–Ingersoll–Ross process as its unique solution, which implies the process X^ϵ as an integral function of Y^ϵ plus a fractional Brownian noise term is well-defined. Moreover, in the context of the previous section, the invariant measure μ has the Gamma density ([FPSS11, Section 33.4])

$$\mu(y) = \frac{(2\beta/v^2)^{2\beta\theta/v^2}}{\Gamma(2\beta\theta/v^2)} y^{2\beta\theta/v^2-1} e^{-2\beta y/v^2}, \text{ for } y \geq 0.$$

Then according to [BGS21, Theorem 1], X^ϵ converges in probability in $C([0, 1])$ to

$$\bar{X}_t = \int_{\mathbb{R}} y \mu(dy) = \theta.$$

The Poisson equation at (6) has an unique solution $\phi(y)$ due to Theorem 4 and this solution satisfies $\phi'(y) = -\frac{1}{2\beta}$. This implies the operator Q^H defined at (36) is

$$Q^H = \tau^2 \dot{K}_H \dot{K}_H^* + \left(\frac{\tau}{2\beta} \right)^2,$$

which is invertible since $\tau > 0$ (see Lemma 9). Here $\dot{K}_H \dot{K}_H^*$ is the operator

$$[\dot{K}_H \dot{K}_H^* h](t) = c_H^2 t^{H-1/2} \int_0^t (t-z)^{H-3/2} z^{1-2H} \int_z^1 (s-z)^{H-3/2} s^{H-1/2} h(s) ds dz.$$

Let us now discuss moderate deviations of X^ϵ . We have already established Y^ϵ has an invariant measure μ , which makes Condition H1 redundant for the model (21). Moreover, if we use the notation of Section 1 then the equation of X^ϵ at (21) has $f = \tau$ and $(\frac{d}{dx}\phi(y))f = 0$. Therefore, based on the proofs of Lemma 12 and Lemma 15, Condition H2-B in this setting simplifies to $|f(y)| \leq C(1 + |y|^{D_f})$ and $0 \leq D_f < 1$, which is clearly satisfied for $f = \tau$. Then, as long as there exists $\beta \in [0, 1]$ such that $\beta + H > 1$ and $h(\epsilon)^{-1}\epsilon^{-\frac{\beta}{2}} \rightarrow 0$ as $\epsilon \rightarrow 0$, Theorem 1 asserts that for the moderate deviations process $(X^\epsilon - \bar{X})/h(\epsilon)\sqrt{\epsilon}$, its action functional, when finite, takes the form

$$S^H(\Phi) = \int_0^1 \dot{\Phi}_s (Q^H)^{-1} \dot{\Phi}_s ds.$$

4.2. Fractional Langevin equation. For the second example, we consider the multiscale model

$$(22) \quad \begin{aligned} dX_t^\epsilon &= (-Q'(Y_t^\epsilon) - V'(X_t^\epsilon)) dt + \sqrt{\epsilon} \sqrt{2D} dW_t^H, \\ dY_t^\epsilon &= -\frac{1}{\epsilon} Q'(Y_t^\epsilon) dt + \frac{1}{\sqrt{\epsilon}} \sqrt{2D} dB_t. \end{aligned}$$

The equation of X^ϵ can be viewed as a rescaled Langevin equation with a fractional Brownian noise. A simpler version of this fractional Langevin equation that does not contain a fast process Y^ϵ was studied in [AMP21, CKM03, GJR18] among others. We assume that

- \mathcal{Y} is the one-dimensional unit torus.
- There is a constant C such that $|Q'(y)| + |V'(x)| + |V''(x)| + |V'''(x)| \leq C(1 + |y|)$.
- $Q'(y)$ is Lipschitz.
- $V'''(x), Q'''(y)$ are continuous.
- D is a real non-zero constant.

Our assumption implies that $Q'(y), V'(x)$ are Lipschitz and that $|Q'(y)| + |V'(x)| \leq C(1 + |y|)$, so that there is a unique strong solution to (22) based on [GN08, Theorem 2.2].

Next, we consider averaging of X^ϵ . Since \mathcal{Y} is the unit torus, Condition 3 in [BGS21, Theorem 1] is not needed for ergodicity of Y^ϵ . Condition 1 in [BGS21, Theorem 1] is met by our second and fourth assumptions for (22) above. Thus, we conclude X^ϵ converges in probability in $C([0, 1])$ to

$$\bar{X}_t = \int_0^t \left(- \int_{\mathcal{Y}} Q'(y) \mu(dy) - V'(\bar{X}_s) \right) ds$$

where, according to [PS07], the invariant measure μ is the Gibbs measure

$$\mu(dy) = \frac{1}{Z} e^{-Q(y)/D}, \quad Z = \int_{\mathcal{Y}} e^{-Q(y)/D} dy.$$

The Poisson equation at (6) becomes

$$-Q'(y)\phi'(y) + D\phi''(y) = \bar{Q}' - Q'(y), \quad \int_{\mathcal{Y}} \phi(y) \mu(dy) = 0$$

such that $\bar{Q}' = \int_{\mathcal{Y}} Q'(y) \mu(dy)$. Its solution satisfies

$$\phi'(y) = \frac{\bar{Q}'}{D} e^{Q(y)/D} \int_0^y e^{-Q(\xi)/D} d\xi + M e^{Q(y)/D} + 1$$

where the constant M is

$$M = - \left(\frac{\bar{Q}'}{D} \int_{\mathcal{Y}} e^{-Q(y)/D} \int_0^y e^{Q(\rho)/D} \int_0^\xi e^{-Q(\xi)/D} d\xi d\rho dy + \int_{\mathcal{Y}} y e^{-Q(y)/D} dy \right) \left(\int_{\mathcal{Y}} e^{-Q(y)/D} \int_0^y e^{Q(\xi)/D} d\xi dy \right)^{-1}.$$

At this point, we can consider moderate deviations of X^ϵ . In the notation of the previous section, we have $g(x, y) = -Q'(y) - V'(x)$, $c(y) = -Q'(y)$, $f = \sigma = \sqrt{2D}$. Since \mathcal{Y} is the unit torus, the first recurrence assumption in Condition H1 and $D_f + D_g < 1$ in Condition H2-B are no longer needed. In addition, the fact that $\frac{d}{dx}\phi(y) = 0$ makes redundant the assumption $M_k/2 + H > 1$ in Condition H2-B. Then the rest of Conditions H1 and H2-B are satisfied by (22). In particular, we have $g(x, y) \in \mathcal{C}^{2,\alpha}(\mathbb{R}^d \times \mathcal{Y})$ since $V''(x), V'''(x)$ are bounded. Next, the operator Q^H in (7) becomes

$$Q^H = 2D\dot{K}_H \dot{K}_H^* + \frac{2D}{Z} \int_{\mathcal{Y}} e^{-Q(y)/D} \left(\frac{\bar{Q}'}{D} e^{Q(y)/D} \int_0^y e^{-Q(\xi)/D} d\xi + M e^{Q(y)/D} + 1 \right)^2 dy$$

where $\dot{K}_H \dot{K}_H^*$ is

$$\begin{aligned} [\dot{K}_H \dot{K}_H^* h](t) &= c_H^2 t^{H-1/2} \\ &\int_0^t (t-z)^{H-3/2} z^{1-2H} \int_z^1 (s-z)^{H-3/2} s^{H-1/2} h(s) ds dz. \end{aligned}$$

Thus, under the condition that Q^H is invertible and there exists $\beta \in [0, 1]$ such that $\beta + H > 1$ and $h(\epsilon)^{-1}\epsilon^{-\frac{\beta}{2}} \rightarrow 0$ as $\epsilon \rightarrow 0$, Theorem 1 says for the moderate deviations process $(X^\epsilon - \bar{X})/h(\epsilon)\sqrt{\epsilon}$, its action functional, when finite, is

$$S^H(\Phi) = \int_0^1 \left(\dot{\Phi}_s + V''(\bar{X}_s)\Phi_s \right) (Q^H)^{-1} \left(\dot{\Phi}_s + V''(\bar{X}_s)\Phi_s \right) ds.$$

Remark 10. Here we compare the result above to the moderate deviations of X^ϵ in

$$(23) \quad \begin{aligned} dX_t^\epsilon &= (-Q'(Y_t^\epsilon) - V'(X_t^\epsilon))dt + \sqrt{\epsilon}\sqrt{2D}dW_t, \\ dY_t^\epsilon &= -\frac{1}{\epsilon}Q'(Y_t^\epsilon)dt + \frac{1}{\sqrt{\epsilon}}\sqrt{2D}dB_t. \end{aligned}$$

We assume W is a Brownian motion independent from B and \mathcal{Y} is the one-dimensional unit torus. Under appropriate conditions, [MS17, Theorem 2.1] says the moderate deviations action functional of (23), when finite, is

$$S^{1/2}(\Phi) = \int_0^1 \left(\dot{\Phi}_s + V''(\bar{X}_s)\Phi_s \right) (Q^{1/2})^{-1} \left(\dot{\Phi}_s + V''(\bar{X}_s)\Phi_s \right) ds.$$

such that

$$Q^{1/2} = 2D + \frac{2D}{Z} \int_y e^{-Q(y)/D} \left(\frac{\bar{Q}'}{D} e^{Q(y)/D} \int_0^y e^{-Q(\xi)/D} d\xi + M e^{Q(y)/D} + 1 \right)^2 dy.$$

Notice in this particular situation, we have continuity of the mapping $H \mapsto S^H$ at $H = 1/2$ (see Remark 7).

5. PROOF OF THEOREM 2

The proof of Theorem 2 will be divided into five subsections. In Subsections 5.1 and 5.2, we prove tightness and convergence of the pair $(\eta^{\epsilon, w^\epsilon}, P^\epsilon)$, respectively. In Subsection 5.3, we prove the Laplace principle lower bound. In Subsection 5.4, we prove that the level sets of $S(\cdot)$ are compact. Finally, in Subsection 5.5, we prove the Laplace principle upper bound and the representation formula of Theorem 1. The main additional work that needs to be done due to the effect of the fBm is seen in the bounds that we need in order to prove tightness (see also Appendix B) and in the proof of the upper bound in Subsection 5.5.

5.1. Proof of tightness. The main result of this section is the following proposition on tightness.

Proposition 1. *Let Conditions H1 and either H2-A or H2-B be satisfied. Consider any family $\{w^\epsilon: \epsilon > 0\}$ of controls in \mathcal{S} satisfying, for some $N < \infty$,*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 |\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 ds < N$$

almost surely. Then, the family $\{(\eta^{\epsilon, w^\epsilon}, P^\epsilon): \epsilon > 0\}$ is tight.

The proof of Proposition 1 will be divided into two parts which are the subject of Subsections 5.1.1 and 5.1.2.

5.1.1. Tightness of $\{P^\epsilon: \epsilon > 0\}$ in $\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$. The argument for tightness is similar to the argument for tightness in [HSS19]. As a first step, we claim that

$$g(P) = \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]} [u]^2 + [v]^2 + [y]^2 P(dudvdydt).$$

is a tightness function from $\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$ to $\mathbb{R} \cup \infty$. Since g is bounded from below, it is sufficient to show that for every $k \in \mathbb{N}$, the level sets

$$L_k = \{P \in \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]): g(P) \leq k\}$$

are relatively compact. For $\epsilon > 0$, let M be a positive constant large enough so that $k/M < \epsilon$ and define

$$A^\epsilon = \{(u, v, y, t) \in \mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1] : |f(u, v, y, t)| > M\}.$$

By Chebyshev's inequality,

$$\begin{aligned} \sup_{P \in L_k} P(A^\epsilon) &\leq \frac{1}{M} \int_{\{(u, v, y, t) \in \mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1] : |f(u, v, y, t)| \geq k\}} |f(u, v, y, t)| P(dudvdyydt) \\ &\leq \frac{g(P)}{M} < \frac{k}{M} < \epsilon. \end{aligned}$$

Therefore, we get

$$\sup_{P \in L_k} P((\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]) \setminus A^\epsilon) > 1 - \epsilon.$$

Since $(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]) \setminus A^\epsilon$ is also compact, this implies that L_k is a tight set of measures and g is a tightness function on $\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$.

For the second step, define $G : \mathcal{P}(\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])) \rightarrow \mathbb{R} \cup \infty$ by

$$G(\nu) = \int_{\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])} g(x) \nu(dx).$$

Then, according to [DE11, Theorem A.3.17], G is a tightness function on $\mathcal{P}(\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]))$. Moreover, the same theorem states that $\{P^\epsilon : \epsilon > 0\}$ is a tight family in $\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$ as long as

$$\sup_{\epsilon > 0} G(\mathcal{L}(P^\epsilon)) < \infty,$$

which is equivalent to

$$\sup_{\epsilon > 0} \mathbb{E}[g(P^\epsilon)] < \infty.$$

As the above holds by Lemma 10 and Lemma 11, we get that indeed $\{P^\epsilon : \epsilon > 0\}$ is tight in $\mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1])$.

5.1.2. Tightness of $\{\eta^{\epsilon, w^\epsilon} : \epsilon > 0\}$ in $C([0, 1]; \mathbb{R}^n)$. Let $\omega_f(\delta) = \sup_{|s-t| < \delta} |f(s) - f(t)|$ be the modulus of continuity of a function f in $C([0, 1]; \mathbb{R}^n)$. According to [Bil13, Theorem 7.3], the family $\{\eta^{\epsilon, w^\epsilon} : \epsilon > 0\}$ is tight in $C([0, 1]; \mathbb{R}^n)$ if and only if

- For each positive δ , there exist an $a, \delta_0 > 0$ such that $P\left(\left|\eta_0^{\epsilon, w^\epsilon}\right| \geq a\right) \leq \delta$ for $\epsilon \leq \delta_0$.
- For all $a > 0$, $\lim_{\delta \rightarrow 0} \limsup_{\epsilon \rightarrow 0} P(\omega_{\eta^{\epsilon, w^\epsilon}}(\delta) \geq a) = 0$.

We only need to check the second condition above since the first condition is automatically true as $\eta_0^{\epsilon, w^\epsilon} = 0$. Recall from (15) that $\eta^{\epsilon, w^\epsilon}$ is given by

$$\begin{aligned} \eta_t^{\epsilon, w^\epsilon} &= \frac{1}{\sqrt{\epsilon} h(\epsilon)} \int_0^t \left[g\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) - \bar{g}(\bar{X}_s) \right] ds + \int_0^t f\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) \dot{u}_s^\epsilon ds \\ &\quad + \frac{1}{h(\epsilon)} \int_0^t f\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) dW_s^H. \end{aligned}$$

A combination of the Poisson equation stated in (6) and Itô's formula yields

$$\int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} \left[g\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) - \bar{g}\left(X_s^{\epsilon, w^\epsilon}\right) \right] ds = \int_0^t \nabla_y \phi\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) \sigma\left(Y_s^{\epsilon, w^\epsilon}\right) \dot{v}_s^\epsilon ds + R_1^\epsilon(t),$$

where

$$\begin{aligned}
R_1^\epsilon(t) = & -\frac{\sqrt{\epsilon}}{h(\epsilon)} \left(\phi(X_t^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}) - \phi(x_0, y_0) \right) + \frac{\sqrt{\epsilon}}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \\
& + \epsilon \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \\
& + \frac{1}{h(\epsilon)} \int_0^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \\
(24) \quad & + \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H.
\end{aligned}$$

Therefore, we can rewrite $\eta^{\epsilon, w^\epsilon}$ as

$$\begin{aligned}
\eta_t^{\epsilon, w^\epsilon} = & \int_0^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds + \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \\
& + \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H + \frac{1}{\sqrt{\epsilon} h(\epsilon)} \int_0^t [\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)] ds + R_1^\epsilon(t) \\
(25) \quad & = D_1^\epsilon(t) + D_2^\epsilon(t) + D_3^\epsilon(t) + D_4^\epsilon(t) + R_1^\epsilon(t).
\end{aligned}$$

In combination with Markov's inequality, Lemma 12 implies tightness of $\{D_1^\epsilon: \epsilon > 0\}$ and $\{D_2^\epsilon: \epsilon > 0\}$. Lemma 15 implies tightness of $\{D_3^\epsilon: \epsilon > 0\}$ and Lemma 16 implies tightness of $\{D_4^\epsilon: \epsilon > 0\}$. It remains to prove the tightness of $\{R_1^\epsilon: \epsilon > 0\}$. The estimates at (43), (44) combined with Lemma 11 and the fact that $\frac{\sqrt{\epsilon}}{h(\epsilon)} \rightarrow 0$ imply that the first term in equation (24) converges to 0 in probability, which implies tightness in $C([0, 1]; \mathbb{R}^n)$. Furthermore, tightness of the remaining integral terms in (24) is implied by Markov's inequality, Lemma 12 and Lemma 15. This shows that $\{R_1^\epsilon: \epsilon > 0\}$ is tight and hence that $\{\eta^{\epsilon, w^\epsilon}: \epsilon > 0\}$ is indeed tight in $C([0, 1]; \mathbb{R}^n)$.

5.2. Proof of existence of a viable pair. In the previous subsection, we have proved that the family of processes $\{(\eta^{\epsilon, w^\epsilon}, P^\epsilon): \epsilon > 0\}$ is tight (see Proposition 1). It follows that for any subsequence of ϵ converging to 0, there exists a subsubsequence of $\{(\eta^{\epsilon, w^\epsilon}, P^\epsilon): \epsilon > 0\}$ which is convergent in distribution to some limit $(\bar{\eta}, \bar{P})$. The goal of this subsection is to show that $(\bar{\eta}, \bar{P})$ is a viable pair with respect to (θ, \mathcal{L}) according to Definition 1 (where \mathcal{L} is the generator defined in (5)).

By the Skorokhod Representation Theorem, we may assume that $\eta^{\epsilon, w^\epsilon}$ converges to $\bar{\eta}$ almost surely along any subsequence. This will allow us to obtain an equation satisfied by $\bar{\eta}$ since we can study the almost sure limits of each individual summand in the representation of $\eta^{\epsilon, w^\epsilon}$ we had obtained in (25). Recall that we had

$$\begin{aligned}
\eta_t^{\epsilon, w^\epsilon} = & \int_0^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds + \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \\
(26) \quad & + \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H + \frac{1}{\sqrt{\epsilon} h(\epsilon)} \int_0^t [\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)] ds + R_1^\epsilon(t).
\end{aligned}$$

Note that we can write the term before last as

$$(27) \quad \int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} [\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)] ds = \int_0^t \nabla_x \bar{g}(\bar{X}_s) \eta_s^{\epsilon, w^\epsilon} ds + R_2^\epsilon(t),$$

where the remainder term $R_2^\epsilon(t)$ is given by

$$R_2^\epsilon(t) = \frac{1}{2} \int_0^t \nabla_x^2 \bar{g}(\zeta_s) \eta_s^{\epsilon, w^\epsilon} (X_s^{\epsilon, w^\epsilon} - \bar{X}_s)^2 ds$$

with ζ_s being a point in between $X_s^{\epsilon, w^\epsilon}$ and \bar{X}_s . Under either Condition H2-A or Condition H2-B, $\nabla_x^2 \bar{g}(x)$ is bounded, so that we can write

$$(28) \quad R_2^\epsilon(t) \leq \int_0^1 |\eta_s^{\epsilon, w^\epsilon}| |X_s^{\epsilon, w^\epsilon} - \bar{X}_s| ds.$$

Lemma 17 assesses the convergence to zero of $R_2^\epsilon(t)$, and [DS12, Lemma 3.2] addresses the convergence of all the other terms at play in (26) and (27) except for one, namely

$$A^{\epsilon, w^\epsilon}(t) = \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds.$$

In order to deal with this last term, let us introduce the term

$$B^{\epsilon, w^\epsilon}(t) = \int_0^t f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds.$$

and prove that it has the same limit as $A^{\epsilon, w^\epsilon}(t)$. First, note that if we assume that Condition H2-B holds, the assumption that f does not depend on x implies that $A^{\epsilon, w^\epsilon}(t) = B^{\epsilon, w^\epsilon}(t)$ even before taking limits. If instead we assume that Condition H2-A holds, we can use the Lipschitz continuity of $f(x, y)$ to write

$$|A^{\epsilon, w^\epsilon}(t) - B^{\epsilon, w^\epsilon}(t)| \leq \int_0^t |X_s^{\epsilon, w^\epsilon} - \bar{X}_s| |\dot{u}_s^\epsilon| ds \leq \sup_{0 \leq s \leq 1} |X_s^{\epsilon, w^\epsilon} - \bar{X}_s| \int_0^t |\dot{u}_s^\epsilon| ds.$$

Proposition 10 and the fact that X^{ϵ, w^ϵ} converges to \bar{X} in probability then imply that

$$(29) \quad |A^{\epsilon, w^\epsilon}(t) - B^{\epsilon, w^\epsilon}(t)| \rightarrow 0$$

almost surely as ϵ goes to 0. Therefore identifying the limit of $A^{\epsilon, w^\epsilon}(t)$ is the same as identifying the weak limit of $B^{\epsilon, w^\epsilon}(t)$. Using the definition of our occupation measures given by (17) and Lemma 6, we can rewrite $B^{\epsilon, w^\epsilon}(t)$ as

$$\begin{aligned} B^{\epsilon, w^\epsilon}(t) &= \int_0^t f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \left(c_H s^{H-1/2} \int_0^s (s-r)^{H-3/2} r^{1/2-H} \hat{u}_r dr \right) ds \\ &= c_H \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} f(\bar{X}_s, y) (s-r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbf{1}_{[0,t]}(s) \mathbf{1}_{[0,s]}(r) \\ &\quad P \otimes P(du^{(1)} du^{(2)} dv^{(1)} dv^{(2)} dy^{(1)} dy^{(2)} ds dr). \end{aligned}$$

In order to somewhat compactify notation, let us introduce, for any $t \in [0, 1]$, the function

$$\begin{aligned} k\left(y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r\right) \\ = c_H f(\bar{X}_s, y) (s-r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbf{1}_{[0,t]}(s) \mathbf{1}_{[0,s]}(r) \end{aligned}$$

as well as, for $0 < \zeta < s$, the sequence

$$\begin{aligned} k^\zeta\left(y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r\right) \\ = c_H k\left(y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r\right) \mathbf{1}_{[\zeta, s-\zeta]}(r). \end{aligned}$$

With these definitions at hand, we can state the following convergence lemma.

Lemma 1. *Assume Conditions H1 and either H2-A or H2-B hold. Then, one has that*

- (i) $\left| \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k^\zeta dP^\epsilon \otimes dP^\epsilon - \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k dP^\epsilon \otimes dP^\epsilon \right| \rightarrow 0 \text{ a.s. as } \zeta \rightarrow 0;$
- (ii) $\left| \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k^\zeta dP^\epsilon \otimes dP^\epsilon - \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k^\zeta d\bar{P} \otimes d\bar{P} \right| \rightarrow 0 \text{ a.s. as } \epsilon \rightarrow 0;$
- (iii) $\left| \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k^\zeta d\bar{P} \otimes d\bar{P} - \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,1]^2} k d\bar{P} \otimes d\bar{P} \right| \rightarrow 0 \text{ a.s. as } \zeta \rightarrow 0;$

Proof. We first prove part (i). Since $k^\zeta \rightarrow k$ pointwise as $\zeta \rightarrow 0$, all we need to prove is that the function k is integrable with respect to $P^\epsilon \otimes P^\epsilon$ on $\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0, 1]^2$ as then, the Dominated Convergence Theorem applies and yields the desired limit. We have for some constant $C < \infty$

$$\begin{aligned} & \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0, 1]^2} k dP^\epsilon \otimes dP^\epsilon \\ & \leq c_H \int_0^1 \left| f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \right| s^{H-1/2} \int_0^s (s-r)^{H-3/2} r^{1/2-H} |\hat{u}_r^\epsilon| dr ds \\ & \leq \int_0^1 \left| f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \right| \left[\dot{K}_H |\hat{u}^\epsilon| \right]_s ds \\ & \leq C \sqrt{\int_0^1 \left| f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \right|^2 ds}, \end{aligned}$$

where the second inequality follows by Lemma 6 and the last inequality is a consequence of Hölder's inequality and Proposition 10. Now, the boundedness of f under Condition H2-A or the sublinear growth of f under Condition H2-B together with Lemma 11 yield

$$\int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0, 1]^2} k dP^\epsilon \otimes dP^\epsilon \leq C \sqrt{\int_0^1 \left| f(\bar{X}_s, Y_s^{\epsilon, w^\epsilon}) \right|^2 ds} < \infty.$$

Part (iii) is proven in the exact same way. Part (ii) is a consequence of the weak convergence of P^ϵ to \bar{P} and the uniform integrability of $\{P^\epsilon \otimes P^\epsilon : \epsilon > 0\}$ (implied by the second point of Definition 1). \square

The above results allow us to obtain an explicit representation of the limit points $(\bar{\eta}, \bar{P})$, which is the object of the following proposition.

Proposition 2. *Let $(\bar{\eta}, \bar{P})$ be a limit point of $\{(\eta^{\epsilon, w^\epsilon}, P^\epsilon) : \epsilon > 0\}$. Under Conditions H1 and either H2-A or H2-B, it holds that*

$$\begin{aligned} \bar{\eta}_t &= \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]} [\nabla_y \phi(\bar{X}_s, y) \sigma(y) v + \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s \mathbf{1}_{[0, t]}(s)] d\bar{P} \\ &+ c_H \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0, 1]^2} f(\bar{X}_s, y^{(1)}) (s-r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbf{1}_{[0, t]}(s) \mathbf{1}_{[0, s]}(r) d\bar{P} \otimes d\bar{P}. \end{aligned}$$

Proof. As pointed out earlier, we can consider the limiting behavior of each individual summands in the representation (26) of $\eta^{\epsilon, w^\epsilon}$. First, under Conditions H1 and either H2-A or H2-B, Lemma 1 and (29) guarantee that

$$\lim_{\epsilon \rightarrow 0} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds = c_H \int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0, 1]^2} f(\bar{X}_s, y) (s-r)^{H-3/2} s^{H-1/2} r^{1/2-H} u^{(2)} \mathbf{1}_{[0, t]}(s) \mathbf{1}_{[0, s]}(r) d\bar{P} \otimes d\bar{P}.$$

Next, we consider the Young integral terms. Under Conditions H1 and H2-A, part (i) of Lemma 15 implies that, as $\epsilon \rightarrow 0$,

$$\mathbb{E} \left[\sup_{0 \leq i \leq 1} \left| \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq Ch(\epsilon)^{-1} \epsilon^{-\frac{\beta}{2}} \rightarrow 0,$$

and

$$\mathbb{E} \left[\left| \sup_{t \in [0, 1]} \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq Ch(\epsilon)^{-1} \epsilon^{\frac{1}{2}} \rightarrow 0.$$

Likewise, under Conditions H1 and H2-B, part (ii) of Lemma 15 implies that as $\epsilon \rightarrow 0$,

$$\mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{1}{h(\epsilon)} \int_0^t f(Y_s^{\epsilon,w^\epsilon}) dW_s^H \right| \right] \leq Ch(\epsilon)^{-1} \epsilon^{-\frac{M_f}{2}} \rightarrow 0,$$

and

$$\mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon,w^\epsilon}, Y_s^{\epsilon,w^\epsilon}) f(Y_s^{\epsilon,w^\epsilon}) dW_s^H \right| \right] \leq Ch(\epsilon)^{-1} \epsilon^{1-\frac{M_k}{2}} \rightarrow 0.$$

Consequently, under Conditions H1 and either H2-A or H2-B, we get

$$\lim_{\epsilon \rightarrow 0} \frac{1}{h(\epsilon)} \int_0^t f(Y_s^{\epsilon,w^\epsilon}) dW_s^H = 0$$

and

$$\lim_{\epsilon \rightarrow 0} \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon,w^\epsilon}, Y_s^{\epsilon,w^\epsilon}) f(Y_s^{\epsilon,w^\epsilon}) dW_s^H = 0.$$

For the limits of the remaining terms in the representation (26) of $\eta^{\epsilon,w^\epsilon}$, we refer to [DS12, Lemma 3.2] in which these terms have already been addressed. \square

The following proposition asserts that the invariant measure of Y^ϵ and the Lebesgue measure are among the marginals of \bar{P} .

Proposition 3. *Recall that μ denotes the unique invariant measure associated with the generator \mathcal{L} defined in (5). Under Condition H1, we have the decomposition*

$$\bar{P}(dudvdydt) = \nu_{y,t}(dudv)\mu(dy)dt,$$

where $\nu_{y,t}$ is a kernel on $\mathcal{U} \times \mathcal{V}$ dependent on $y \in \mathcal{Y}$ and $t \in [0,1]$.

Proof. Let F be an element of a dense subset of $C^2(\mathcal{Y})$ that consists of bounded functions with bounded first and second derivatives. By Ito's formula, we have

$$(30) \quad \begin{aligned} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,t]} \mathcal{L}F(y) dP^\epsilon &= \epsilon \left(F(Y_t^{\epsilon,w^\epsilon}) - F(y_0) \right) \\ &\quad - \sqrt{\epsilon} \int_0^t (\nabla_y F)^\top(Y_s^{\epsilon,w^\epsilon}) \sigma(Y_s^{\epsilon,w^\epsilon}) dB_s - \sqrt{\epsilon} h(\epsilon) \int_0^t (\nabla_y F)^\top(Y_s^{\epsilon,w^\epsilon}) \sigma(Y_s^{\epsilon,w^\epsilon}) \dot{v}_s^\epsilon ds \end{aligned}$$

Let us consider each individual term on the right-hand side of the above equation. The first term converges to 0 given that F is bounded. For the second term, an application of the Burkholder-Davis-Gundy inequality yields

$$\sqrt{\epsilon} \mathbb{E} \left[\left| \int_0^t (\nabla_y F)^\top(Y_s^{\epsilon,w^\epsilon}) \sigma(Y_s^{\epsilon,w^\epsilon}) dB_s \right| \right] \leq C \sqrt{\epsilon} \sqrt{\mathbb{E} \left[\int_0^1 |\sigma(Y_s^{\epsilon,w^\epsilon})|^2 ds \right]},$$

which converges to zero due to the boundedness of $\sigma(y)\sigma^\top(y)$ in Condition H1. Similarly, by Lemma 10, we have

$$\begin{aligned} &\left| \sqrt{\epsilon} h(\epsilon) \int_0^t (\nabla_y F)^\top(Y_s^{\epsilon,w^\epsilon}) \sigma(Y_s^{\epsilon,w^\epsilon}) \dot{v}_s^\epsilon ds \right| \\ &\leq \sqrt{\epsilon} h(\epsilon) \sqrt{\int_0^t \left| (\nabla_y F)^\top(Y_s^{\epsilon,w^\epsilon}) \sigma(Y_s^{\epsilon,w^\epsilon}) \sigma^\top(Y_s^{\epsilon,w^\epsilon}) \nabla_y F(Y_s^{\epsilon,w^\epsilon}) \right| ds \int_0^t |\dot{v}_s^\epsilon|^2 ds} \\ &\leq C \sqrt{\epsilon} h(\epsilon). \end{aligned}$$

Hence, (30) becomes

$$(31) \quad \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,t]} \mathcal{L}F(y) d\bar{P} = 0.$$

Moreover, it is immediate to see that $P^\epsilon(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, t]) = t$, which implies that the last marginal of \bar{P} is the Lebesgue measure. In other words, \bar{P} is of the form $\bar{P}(dudvdydt) = \nu_{t,y}(dudv)m(dy)dt$. Moreover, since \mathcal{L} is independent of the control (u, v) , (31) implies that

$$\int_{\mathcal{Y}} \mathcal{L}F(y)m(dy) = 0,$$

which implies that $m(dy)$ is the unique invariant measure $\mu(dy)$ associated with \mathcal{L} . \square

The next proposition asserts that the pair $(\bar{\eta}, \bar{P})$ is indeed a viable pair with respect to (θ, \mathcal{L}) , which was what this subsection was aimed at proving.

Proposition 4. *The pair $(\bar{\eta}, \bar{P})$ is a viable pair with respect to (θ, \mathcal{L}) , where θ is the function defined in (16) and \mathcal{L} is the generator defined in (5).*

Proof. Lemmas 10 and 11 together with Fatou's lemma ensure that \bar{P} satisfies the first property in Definition 1. The following two properties in Definition 1 have been established in Proposition 2 and Proposition 3. \square

5.3. Proof of the Laplace principle lower bound. The Laplace principle lower bound immediately derives from Fatou's lemma and Proposition 2, which is shown in the following proposition.

Proposition 5. *Assume Conditions H1 and either H2-A or H2-B are satisfied. Then, for all bounded and continuous mappings $a: C([0, 1]; \mathbb{R}^n) \rightarrow \mathbb{R}$, the following Laplace principle lower bound holds.*

$$\liminf_{\epsilon \rightarrow 0} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] \geq \inf_{\Phi \in C([0, 1]; \mathbb{R}^n)} S^H(\Phi) + a(\Phi),$$

where the rate function S^H is defined at (19).

Proof. We can write

$$\begin{aligned} \liminf_{\epsilon \rightarrow 0} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] &\geq \liminf_{\epsilon \rightarrow 0} \mathbb{E}\left[\frac{1}{2} \int_0^1 \left[|\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2\right] ds + a(\eta^{\epsilon, w^\epsilon})\right] - \delta \\ &= \liminf_{\epsilon \rightarrow 0} \mathbb{E}\left[\frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]} \left[|u|^2 + |v|^2\right] dP^\epsilon + a(\eta^{\epsilon, w^\epsilon})\right] \\ &\geq \mathbb{E}\left[\frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0, 1]} \left[|u|^2 + |v|^2\right] d\bar{P} + a(\bar{\eta})\right] \\ &\geq \inf_{\Phi \in C([0, 1]; \mathbb{R}^n)} S^H(\Phi) + a(\Phi). \end{aligned}$$

The first inequality comes from the variational formula (12). The second line is a direct application of the definition of the occupation measure P^ϵ . The third line follows from Fatou's lemma and the convergence result in Proposition 2. Finally, the last line is a consequence of Proposition 4. \square

5.4. Proof of the compactness of the level sets of $S^H(\cdot)$. We need to show that, for each $k \in \mathbb{R}$, the level sets of S^H given by

$$L_k = \{\Phi \in C([0, 1]; \mathbb{R}^n) : S^H(\Phi) \leq k\}, \quad k < \infty.$$

are compact subsets of $C([0, 1]; \mathbb{R}^n)$, which indicates that S^H is a good rate function. We will actually show that, for any $k \in \mathbb{R}$, L_k is relatively compact and closed. We start with relative compactness, which the following lemma addresses.

Lemma 2. *Let $\{(\Phi_n, P_n) : n \in \mathbb{N}\}$ be a sequence such that for every $n \in \mathbb{N}$, $(\Phi_n, P_n) \in \mathcal{V}(\theta, \mathcal{L})$ and $\Phi_n \in L_k$. Assuming Conditions H1 and either H2-A or H2-B are satisfied, this sequence is relatively compact in $C([0, 1]; \mathbb{R}^n)$.*

Proof. We can show that the family $\{P_n: n \in \mathbb{N}\}$ is relatively compact in the same way as in Subsection 5.1.1 where we proved the tightness of $\{P^\epsilon: \epsilon > 0\}$. To show the relative compactness of $\{\Phi_n: n \in \mathbb{N}\}$, it is sufficient to verify that

$$\lim_{\delta \rightarrow 0} \sup_{\Phi \in L_k} \omega_\Phi(\delta) = \lim_{\delta \rightarrow 0} \sup_{\Phi \in L_k} \sup_{|t-r| \leq \delta} |\Phi(t) - \Phi(r)| = 0.$$

By Proposition 7, the fact that $(\Phi_n, P_n) \in \mathcal{V}(\theta, \mathcal{L})$ implies there exists a pair of ordinary controls $(u, v) \in L^2(\mathcal{Y}^2 \times [0, 1]^2; \mathbb{R}^m \times \mathbb{R}^p)$ such that

$$\begin{aligned} \Phi_t &= \int_{\mathcal{Y} \times [0, t]} \nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y) \mu(dy) ds + \int_0^t \nabla_x \bar{g}(\bar{X}_s) \Phi_s ds \\ &\quad + \int_{\mathcal{Y} \times [0, t]} f(\bar{X}_s, y) (\dot{K}_H u)(s, y) \mu(dy) ds. \end{aligned}$$

Then,

$$\begin{aligned} \Phi_t - \Phi_r &= \int_{\mathcal{Y} \times [r, t]} \nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y) \mu(dy) ds + \int_0^t \nabla_x \bar{g}(\bar{X}_s) \Phi_s ds \\ &\quad + \int_{\mathcal{Y} \times [r, t]} f(\bar{X}_s, y) (\dot{K}_H u)(s, y) \mu(dy) ds \\ &= A_1 + A_2 + A_3. \end{aligned}$$

The term A_1 can be estimated by

$$\begin{aligned} |A_1| &\leq \sqrt{\int_{\mathcal{Y} \times [r, t]} |\nabla_y \phi(\bar{X}_s, y) \sigma(y)|^2 \mu(dy) ds} \sqrt{\int_{\mathcal{Y} \times [0, 1]} |v(s, y)|^2 \mu(dy) ds} \\ &\leq C \sqrt{\int_{\mathcal{Y} \times [r, t]} |y|^{2D_g} \mu(dy) ds} \\ &\leq C \sqrt{|t-r|}. \end{aligned}$$

Similarly, we have

$$|A_2| \leq C \int_r^t |\Phi_s| ds \leq C |t-r|,$$

which is immediate as $\Phi \in C([0, 1]; \mathbb{R}^n)$ is bounded. For the final term A_3 , we apply Proposition 10 to get

$$\begin{aligned} |A_3| &\leq \sqrt{\int_{\mathcal{Y} \times [r, t]} |y|^{2D_f} \mu(dy) ds} \sqrt{\int_{\mathcal{Y} \times [0, 1]} |(\dot{K}_H u)(s, y)|^2 \mu(dy) ds} \\ &\leq C \sqrt{|t-r|}. \end{aligned}$$

Combining the previous estimates leads to

$$|\Phi_t - \Phi_r| \leq C(|t-r| + \sqrt{|t-r|}),$$

which completes the proof. \square

The next step is to prove that the limit of a sequence of viable pairs is a viable pair. This is the object of the next lemma.

Lemma 3. *Let $\{(\Phi_n, P_n): n \in \mathbb{N}\}$ be a sequence such that for every $n \in \mathbb{N}$, $(\Phi_n, P_n) \in \mathcal{V}(\theta, \mathcal{L})$ and $\Phi_n \in L_k$. Furthermore, assume that the sequence $\{(\Phi_n, P_n): n \in \mathbb{N}\}$ converges to a limit (Φ, P) . Assuming Conditions H1 and either H2-A or H2-B are satisfied, we have $(\Phi, P) \in \mathcal{V}(\theta, \mathcal{L})$.*

Proof. Using Fatou's lemma, we can write

$$\int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP \leq \liminf_{n \rightarrow \infty} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP_n \leq k,$$

so that the second criterion in Definition 1 is satisfied. The third and fourth criteria can be proved in a similar but simpler manner as in the proofs of Propositions 2 and 3. \square

The final step is to prove that the map S^H is lower semicontinuous, which is done in the lemma below.

Lemma 4. *Assume Conditions H1 and either H2-A or H2-B hold. Then, $S^H(\Phi)$ is lower semicontinuous, which is equivalent to the statement that the level sets of S^H are closed in $C([0,1]; \mathbb{R}^n)$.*

Proof. Let Φ_n converge to Φ in $C([0,1]; \mathbb{R}^n)$. We will show

$$\liminf_{n \rightarrow \infty} S^H(\Phi_n) \geq S^H(\Phi).$$

When $S^H(\Phi_n) < \infty$, there exists P_n such that $(\Phi_n, P_n) \in \mathcal{V}(\theta, \mathcal{L})$ and

$$S^H(\Phi_n) \geq \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP_n - \frac{1}{n}.$$

By Lemma 2, we can consider a subsequence along which (Φ_n, P_n) converges to (Φ, P) . Moreover, Lemma 3 guarantees $(\Phi, P) \in \mathcal{V}(\theta, \mathcal{L})$. Consequently,

$$\begin{aligned} \liminf_{n \rightarrow \infty} S^H(\Phi_n) &= \liminf_{n \rightarrow \infty} \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP_n - \frac{1}{n} \\ &\geq \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP \\ &\geq \frac{1}{2} \inf_{(\Phi, P) \in \mathcal{V}} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] dP = S^H(\Phi), \end{aligned}$$

which concludes the proof. \square

We can now combine the preceding results in order to state the following proposition, which was the object of this subsection.

Proposition 6. *Assume Conditions H1 and either H2-A or H2-B are satisfied. Then, for every $k \in \mathbb{R}$, L_k is a compact subset of $C([0,1]; \mathbb{R}^n)$.*

5.5. Proof of the Laplace principle upper bound and representation formula. We begin by introducing an alternate representation of $S^H(\Phi)$. By the definition of viable pairs (see Definition 1) and that of $S^H(\Phi)$ (see (19)), we can write, for any $\Phi \in C([0,1]; \mathbb{R}^n)$,

$$\begin{aligned} S^H(\Phi) &= \inf_{(\Phi, P) \in \mathcal{V}(\theta, \mathcal{L})} \left[\frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] P(dudvdyds) \right] \\ &= L^r(\Phi, \bar{X}), \end{aligned}$$

where

$$L^r(\Phi, \bar{X}) = \inf_{P \in \mathcal{A}_\Phi^r} \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] P(dudvdydt).$$

The set \mathcal{A}_Φ^r consists of elements $P \in \mathcal{P}(\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1])$ for which the decomposition (18) holds and such that

$$\int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2 + |y|^2] P(dudvdyds) < \infty$$

and (recalling the definition of the function θ given in (16))

$$\int_{\mathcal{U}^2 \times \mathcal{V}^2 \times \mathcal{Y}^2 \times [0,t]^2} \theta\left(x, \Phi_s, y^{(1)}, y^{(2)}, u^{(1)}, u^{(2)}, v^{(1)}, v^{(2)}, s, r\right) dP \otimes dP = \Phi_t.$$

Now, for any $\Phi \in C([0, 1]; \mathbb{R}^n)$, let us define

$$(32) \quad L^o(\Phi, \bar{X}) = \inf_{w \in \mathcal{A}_\Phi^o} \frac{1}{2} \int_{\mathcal{Y} \times [0,1]} |w(t, y)|^2 \mu(dy) dt,$$

where the set \mathcal{A}_Φ^o consists of elements $w = (u, v) : \mathcal{Y}^2 \times [0, 1]^2 \rightarrow \mathbb{R}^{m+p}$ of \mathcal{S} such that, for any $t \in [0, 1]$,

$$(33) \quad \int_{\mathcal{Y} \times [0, t]} \left[\nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y) + \nabla_x \bar{g}(\bar{X}_s) \Phi_s + \bar{f}(\bar{X}_s) \left(\dot{K}_H u \right)(s, y) \right] \mu(dy) ds = \Phi_t$$

and

$$\int_{\mathcal{Y} \times [0,1]} \left[|u(t, y)|^2 + |v(t, y)|^2 \right] \mu(dy) dt < \infty.$$

Our claim is that one actually has that $L^r(\Phi, \bar{X}) = L^o(\Phi, \bar{X})$, which will provide us with the representation of $S^H(\Phi)$ we need to derive the upper bound of the Laplace principle. The equivalence between these two control systems is the object of the following proposition.

Proposition 7. *Under Conditions H1 and either H2-A or H2-B, it holds that $L^r(\Phi, \bar{X}) = L^o(\Phi, \bar{X})$.*

Proof. Let us first show $L^r(\Phi, \bar{X}) \geq L^o(\Phi, \bar{X})$. Choose any $P \in \mathcal{A}_\Phi^r$. Then, by definition of \mathcal{A}_Φ^r , the decomposition $P(dudvdydt) = \nu_{t,y}(dudv)\mu(dy)dt$ holds. This allows us to define an element $w = (u, v)$ with

$$(34) \quad u(t, y) = \int_{\mathcal{U} \times \mathcal{V}} v \nu_{t,y}(dudv) \quad \text{and} \quad v(t, y) = \int_{\mathcal{U} \times \mathcal{V}} v \nu_{t,y}(dudv).$$

We claim that $w \in \mathcal{A}_\Phi^o$. Jensen's inequality and the decomposition of P imply

$$(35) \quad \begin{aligned} \int_{\mathcal{Y} \times [0,1]} \left[|u(t, y)|^2 + |v(t, y)|^2 \right] \mu(dy) dt &= \int_{\mathcal{Y} \times [0,1]} \left[\left(\int_{\mathcal{U} \times \mathcal{V}} z_1 \nu_{t,y}(dudv) \right)^2 \right. \\ &\quad \left. + \left(\int_{\mathcal{U} \times \mathcal{V}} z_2 \nu_{t,y}(dudv) \right)^2 + \left(\int_{\mathcal{U} \times \mathcal{V}} v \nu_{t,y}(dudv) \right)^2 \right] \mu(dy) dt \\ &\leq \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} \left[|u|^2 + |v|^2 \right] \nu_{t,y}(dudv) \mu(dy) dt \\ &= \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} \left[|u|^2 + |v|^2 \right] P(dudvdydt) < \infty. \end{aligned}$$

Hence, the last property in the definition of \mathcal{A}_Φ^o is satisfied and based on (34), so is the first one. This shows that $w^{(1)}(t, y) \in \mathcal{A}_\Phi^o$. Furthermore, (35) yields

$$\begin{aligned} L^r(\Phi, \bar{X}) &= \inf_{P \in \mathcal{A}_\Phi^r} \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} \left[|u|^2 + |v|^2 \right] P(dudvdydt) \\ &\geq \inf_{P \in \mathcal{A}_\Phi^r} \frac{1}{2} \int_{\mathcal{Y} \times [0,1]} \left[|u(t, y)|^2 + |v(t, y)|^2 \right] \mu(dy) dt \\ &\geq \inf_{w \in \mathcal{A}_\Phi^o} \frac{1}{2} \int_{\mathcal{Y} \times [0,1]} |w(t, y)|^2 \mu(dy) dt = L^o(\Phi, \bar{X}). \end{aligned}$$

It remains to prove that $L^r(\Phi, \bar{X}) \leq L^o(\Phi, \bar{X})$. To this end, choose any $w = (u, v) \in \mathcal{A}_\Phi^o$ and construct a measure $P \in \mathcal{A}_\Phi^r$ according to

$$P(dudvdydt) = \delta_{u(t,y)}(du) \delta_{v(t,y)}(dv) \mu(dy) dt.$$

Checking that P satisfies all the needed properties to belong to \mathcal{A}_Φ^r is similar to what was done above. We hence have a set $\mathcal{B}^r \subseteq \mathcal{A}_\Phi^r$ that corresponds to the set \mathcal{A}_Φ^o , from which we deduce that

$$\begin{aligned} L^o(\Phi, \bar{X}) &= \inf_{(u,v) \in \mathcal{A}_\Phi^o} \frac{1}{2} \int_{\mathcal{Y} \times [0,1]} [|u(t,y)|^2 + |v(t,y)|^2] \mu(dy) dt \\ &\geq \inf_{P \in \mathcal{A}_\Phi^r} \frac{1}{2} \int_{\mathcal{U} \times \mathcal{V} \times \mathcal{Y} \times [0,1]} [|u|^2 + |v|^2] P(dudvdydt) = L^r(\Phi, \bar{X}), \end{aligned}$$

which concludes the proof. \square

The next step is to derive an explicit expression of $L^o(\Phi, \bar{X})$. The statement of this expression requires us to introduce some linear maps. For a given $x \in C([0,1]; \mathbb{R}^n)$, let $\pi_x : L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m) \rightarrow L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^n)$ and $\rho_x : L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^p) \rightarrow L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^n)$ be two operators defined by

$$\pi_x u(t) = \int_{\mathcal{Y}} \bar{f}(x) \dot{K}_H u(t, y) \mu(dy), \quad \rho_x v(t) = \int_{\mathcal{Y}} \nabla_y \phi(x, y) \sigma(y) v(t, y) \mu(dy).$$

Under either Condition H2-A or H2-B, $\bar{f}(x)$ is bounded. This fact and Proposition 10 yield

$$\|\pi_x u\|_{L^2([0,1]; \mathbb{R}^n)} \leq C \sqrt{\int_{\mathcal{Y} \times [0,1]} |\dot{K}_H u(t, y)|^2 \mu(dy) dt} \leq C \|u\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m)},$$

so that π_x is bounded. The operator ρ_x is also bounded via Lemma 11 and the estimates (43), (44). Therefore, the Hilbert adjoints π_x^* and ρ_x^* are well-defined and given by $\pi_x^* h = \dot{K}_H^* (\bar{f})^\top(x) h$ and $\rho_x^* h = \sigma(\cdot)^\top (\nabla_y \phi)^\top(x, \cdot) h$, respectively. It follows from these facts that

$$\Sigma_x(u, v) = \pi_x u + \rho_x v$$

is also a bounded operator. Thus, its Hilbert adjoint Σ_x^* exists and is given by $\Sigma_x^* h = (\pi_x^* h, \rho_x^* h)$. With these definitions at hand, let us finally define the operator Q_x^H from $L^2([0,1]; \mathbb{R}^n)$ to itself by

$$(36) \quad Q_x^H = \Sigma_x \Sigma_x^* = \bar{f}(x) \dot{K}_H \left(\bar{f}(x) \dot{K}_H \right)^* + \int_{\mathcal{Y}} (\nabla_y \phi(x, y) \sigma(y)) (\nabla_y \phi(x, y) \sigma(y))^\top \mu(dy).$$

such that for $h \in L^2([0,1]; \mathbb{R}^n)$,

$$\begin{aligned} \left[\bar{f}(\bar{X}) \dot{K}_H \left(\bar{f}(\bar{X}) \dot{K}_H \right)^* h \right] (t) &= c_H^2 \bar{f}(\bar{X}_t) t^{H-1/2} \\ &\quad \int_0^t (t-z)^{H-3/2} z^{1-2H} \int_z^1 (s-z)^{H-3/2} s^{H-1/2} \bar{f}(\bar{X}_s)^\top h(s) ds dz. \end{aligned}$$

We are now ready to present the explicit expression of $L^o(\Phi, \bar{X})$, which is the object of the next proposition.

Proposition 8. *Assume Conditions H1 and either H2-A or H2-B, and further that the operator $Q_{\bar{X}}^H$ is invertible. Then the ordinary control problem (32) has a finite minimum cost if and only if Φ is absolutely continuous and $\dot{\Phi}$ (defined a.e.) is square integrable. In this case, the solution is given by*

$$L^o(\Phi, \bar{X}) = \int_0^1 \left(\dot{\Phi}_s - \nabla_x \bar{g}(\bar{X}_s) \Phi_s \right)^\top (Q_{\bar{X}_s}^H)^{-1} \left(\dot{\Phi}_s - \nabla_x \bar{g}(\bar{X}_s) \Phi_s \right) ds$$

and is achieved for the optimal control

$$(\bar{u}, \bar{v}) = \left(\pi^*(Q_{\bar{X}}^H)^{-1} \left(\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi \right), \rho^*(Q_{\bar{X}}^H)^{-1} \left(\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi \right) \right).$$

Proof. In one direction, let us assume (32) has a finite minimum cost then Equation 33 is satisfied for some $(u, v) \in A_\Phi^o$. Hence, Φ is absolutely continuous. Furthermore, by Cauchy-Schwarz inequality,

$$(37) \quad \begin{aligned} \|\dot{\Phi}\|_{L^2([0,1];\mathbb{R}^n)} &= \int_0^1 \left| \int_{\mathcal{Y}} [\nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y) + \nabla_x \bar{g}(\bar{X}_s) \Phi_s + \bar{f}(\bar{X}_s) (\dot{K}_H u)(s, y)] \mu(dy) \right|^2 ds \\ &\leq \int_{\mathcal{Y} \times [0,1]} |\nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y)|^2 + |\nabla_x \bar{g}(\bar{X}_s) \Phi_s|^2 + |\bar{f}(\bar{X}_s) (\dot{K}_H u)(s, y)|^2 \mu(dy) ds \end{aligned}$$

To see that $\dot{\Phi}$ is square integrable, we study the right hand side of this inequality. $\nabla_y \phi(\bar{X}_s, y) \sigma(y) v(s, y) \in L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^n)$ due to Lemma 11, Remark 12 and boundedness of $\sigma(y) \sigma(y)^T$ in Condition H1. Next, we have $u \in L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m)$, $v \in L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^p)$ and it follows from Proposition 10 that $\dot{K}_H u \in L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m)$. This along with the fact that $\bar{f}(x), \nabla_x \bar{g}(x)$ is bounded under either Condition H2-A or H2-B, and the fact that Φ is bounded since its derivative exists a.e. in $[0, 1]$, implies the remaining quantities on the right side of (37) are in $L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^n)$. Therefore, we can conclude $\dot{\Phi}$ is square integrable.

In the other direction, let us assume that Φ is absolutely continuous and that the a.e. defined derivative is square integrable. Then $\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi$ is also square integrable. Then, we can construct a control (\bar{u}, \bar{v}) as in the statement of the proposition so that the set A_Φ^o associated with the ordinary control problem (32) is non-empty and therefore, the minimum cost L^o is finite. This settles the first claim of this proposition.

Next, let us derive an explicit formula for the minimum cost. When Φ is absolutely continuous and its a.e. defined derivative is square integrable, we have

$$\begin{aligned} \|(\bar{u}, \bar{v})\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m \times \mathbb{R}^p)}^2 &= \|\bar{u}\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m)}^2 + \|\bar{v}\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^p)}^2 \\ &= \left\langle (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi), \pi \pi^* (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)} \\ &\quad + \left\langle (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi), \rho \rho^* (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)} \\ &= \left\langle (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi), Q_{\bar{X}}^H (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)} \\ &= \left\langle \dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi, (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)}. \end{aligned}$$

Since we also know that $(\bar{u}, \bar{v}) \in A_\Phi^o$, this implies

$$L^o(\Phi, \bar{X}, \Phi) \leq \left\langle \dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi, (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)}.$$

Furthermore, by Lemma 8 and the fact that $\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \eta = \Sigma_{\bar{X}}(u, v)$, we can write

$$\begin{aligned} L^o(\Phi, \bar{X}) &\geq \inf_{(u, v) \in A_\Phi^o} \|\Sigma_{\bar{X}}^* (Q_{\bar{X}}^H)^{-1} \Sigma(u, v)\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m \times \mathbb{R}^p)}^2 \\ &= \left\| \Sigma_{\bar{X}}^* (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\|_{L^2(\mathcal{Y} \times [0,1]; \mathbb{R}^m \times \mathbb{R}^p)}^2 \\ &= \left\langle (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi), \Sigma_{\bar{X}} \Sigma_{\bar{X}}^* (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)} \\ &= \left\langle \dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi, (Q_{\bar{X}}^H)^{-1} (\dot{\Phi} - \nabla_x \bar{g}(\bar{X}) \Phi) \right\rangle_{L^2([0,1]; \mathbb{R}^n)}. \end{aligned}$$

Thus, the minimum cost of the ordinary control problem (32) is the quantity in the last line. \square

We are now ready to prove the Laplace principle upper bound, which is the object of the next proposition.

Proposition 9. Assume Conditions H1 and either H2-A or H2-B, and further that the operator Q_X^H is invertible. Then the following Laplace principle upper bound holds.

$$\limsup_{\epsilon \rightarrow 0} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] \leq \inf_{\Phi \in C([0,1]; \mathbb{R}^n)} S^H(\Phi) + a(\Phi),$$

where the function S is the one defined at (19).

Proof. We can assume, without loss of generality, that $\inf_{\Phi \in C([0,1]; \mathbb{R}^n)} S^H(\Phi) < \infty$, so that for any $\zeta > 0$, there exists an element $\Phi_0 \in C([0,1]; \mathbb{R}^n)$ for which

$$S^H(\Phi_0) + h(\Phi_0) \leq \inf_{\Phi \in C([0,1]; \mathbb{R}^n)} (S^H(\Phi) + h(\Phi)) + \zeta.$$

Let us also define

$$\tilde{w}(\phi, x, y, \eta) = (\tilde{u}(\phi, x, y, \eta), \tilde{v}(\phi, x, y, \eta)) = \left(\pi_x^* Q_x^{-1} \left(\dot{\phi} - \nabla_x \bar{g}(x) \eta \right), \rho_x^* Q_x^{-1} \left(\dot{\phi} - \nabla_x \bar{g}(x) \eta \right) \right)$$

and

$$w_0 = \left(\tilde{u}(\Phi_0, X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}, \eta^{\epsilon, w^\epsilon}), \tilde{v}(\Phi_0, X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}, \eta^{\epsilon, w^\epsilon}) \right).$$

We can then substitute w_0 into the control variable of equation (25) and take the limit of η^{ϵ, w_0} as $\epsilon \rightarrow 0$. This procedure is the same as the one that was carried out in Proposition 2 and after which we obtained

$$\begin{aligned} \bar{\eta}_t &= \int_0^t \left[\rho \left(\rho^*(Q_X^H)^{-1} \left(\dot{\Phi}_0(s) - \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s \right) \right) + \pi \left(\pi^*(Q_X^H)^{-1} \left(\dot{\Phi}_0(s) - \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s \right) \right) \right] ds \\ &\quad + \int_0^t \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s ds \\ &= \int_0^t Q_X^H (Q_X^H)^{-1} \left(\dot{\Phi}_0(s) - \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s \right) ds + \int_0^t \nabla_x \bar{g}(\bar{X}_s) \bar{\eta}_s ds \\ (38) \quad &= \Phi_0(t). \end{aligned}$$

In addition, we have

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \mathbb{E} \left[\frac{1}{2} \int_0^1 \left| \tilde{u}(\Phi_0(s), X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}, \eta_s^{\epsilon, w^\epsilon}) \right|^2 + \left| \tilde{v}(\Phi_0(s), X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}, \eta_s^{\epsilon, w^\epsilon}) \right|^2 ds \right] \\ = \mathbb{E} \left[\frac{1}{2} \int_{\mathcal{Y} \times [0,1]} \left| \tilde{u}(\Phi_0(s), \bar{X}_s, y, \bar{\eta}_s) \right|^2 + \left| \tilde{v}(\Phi_0(s), \bar{X}_s, y, \bar{\eta}_s) \right|^2 \mu(dy) ds \right] \\ (39) \quad = S^H(\Phi_0), \end{aligned}$$

where the last equality is a consequence of Propositions 7, 8 and (38). Therefore,

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0} -\frac{1}{h^2(\epsilon)} \ln \mathbb{E}[\exp(-h^2(\epsilon)a(\eta^\epsilon))] &= \limsup_{\epsilon \rightarrow 0} \inf_{w^\epsilon \in \mathcal{S}} \mathbb{E} \left[\frac{1}{2} \int_0^1 |\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 ds + a(\eta^{\epsilon, w^\epsilon}) \right] \\ &\leq \limsup_{\epsilon \rightarrow 0} \mathbb{E} \left[\frac{1}{2} \int_0^1 \left| \tilde{u}(\Phi_0(s), X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}, \eta_s^{\epsilon, w^\epsilon}) \right|^2 \right. \\ &\quad \left. + \left| \tilde{v}(\Phi_0(s), X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}, \eta_s^{\epsilon, w^\epsilon}) \right|^2 ds + a(\eta_s^{\epsilon, w_0}) \right] \\ &= \mathbb{E} \left[\frac{1}{2} \int_{\mathcal{Y} \times [0,1]} \left| \tilde{u}(\Phi_0(s), \bar{X}_s, y, \bar{\eta}_s) \right|^2 + \left| \tilde{v}(\Phi_0(s), \bar{X}_s, y, \bar{\eta}_s) \right|^2 \mu(dy) ds + a(\bar{\eta}_s) \right] \\ &= S^H(\Phi_0) + a(\Phi_0) \leq \inf_{\Phi \in C([0,1]; \mathbb{R}^n)} (S^H(\Phi) + h(\Phi)) + \zeta, \end{aligned}$$

where the first equality is due to the variational formula (12), the second line is due to the choice of a particular control, and the last two equalities are consequences of (38) and (39). Finally, the fact that ζ can be chosen arbitrarily yields the desired Laplace principle upper bound. \square

6. PROOF OF COROLLARY 1

First, observe that given any $\Psi \in L^2([0, 1]; \mathbb{R}^n)$, the quantities

$$D_1 = t^{H-1/2} \bar{f}(\bar{X})^\top \Psi, \quad D_2 = t^{1-2H} I_{1-}^{H-1/2} t^{H-1/2} \bar{f}(\bar{X})^\top \Psi$$

are in $L^2([0, 1]; \mathbb{R}^n)$. D_1 is square-integrable because $H > 1/2$ and \bar{f} is bounded under either Condition H2-A or H2-B. Regarding D_2 , notice that the assumptions of Lemma 5 are satisfied with $p = 2, \alpha = H - 1/2, \beta = 1 - 2H$ and $\gamma = H - 1/2$ for values of H in the range $(1/2, 3/4)$. Then the operator $t^{1-2H} I_{1-}^{H-1/2} t^{H-1/2}$ is bounded in $L^2([0, 1]; \mathbb{R}^n)$, which implies D_2 is square-integrable.

Next, $g = g(x)$ implies $\nabla_y \phi(x, y) = 0$, where $\phi(x, y)$ is defined in (6). Thus,

$$Q_{\bar{X}}^H = \bar{f}(\bar{X}) \dot{K}_H \dot{K}_H^* \bar{f}(\bar{X})^\top.$$

We also know

$$\dot{K}_H = c_H \Gamma(H - 1/2) t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H}, \quad \dot{K}_H^* = c_H \Gamma(H - 1/2) t^{1/2-H} I_{1-}^{H-1/2} t^{H-1/2}$$

so that

$$Q_{\bar{X}}^H = c_H^2 \Gamma(H - 1/2)^2 \bar{f}(\bar{X}) t^{H-1/2} I_{0+}^{H-1/2} t^{1-2H} I_{1-}^{H-1/2} t^{H-1/2} \bar{f}(\bar{X})^\top.$$

Recalling that $Lf(\bar{X}) = f(\bar{X})L = I$, at this point we want to show

$$(40) \quad W = c_H^{-2} \Gamma(H - 1/2)^{-2} L^\top t^{1/2-H} D_{1-}^{H-1/2} t^{2H-1} D_{0+}^{H-1/2} t^{1/2-H} L$$

is the left inverse of $Q_{\bar{X}}^H$. [KST06, Lemma 2.4] says given any $h \in L^2([0, 1]; \mathbb{R}^n)$, we have

$$D_{0+}^{H-1/2} I_{0+}^{H-1/2} h = h, \quad D_{1-}^{H-1/2} I_{1-}^{H-1/2} h = h.$$

This combined with the fact that $D_1, D_2 \in L^2([0, 1]; \mathbb{R}^n)$ implies for any $\Psi \in L^2([0, 1]; \mathbb{R}^n)$,

$$\begin{aligned} W Q_{\bar{X}}^H \Psi &= \left(c_H^{-2} \Gamma(H - 1/2)^{-2} L^\top t^{1/2-H} D_{1-}^{H-1/2} t^{2H-1} D_{0+}^{H-1/2} t^{1/2-H} L \right) \\ &\quad \left(c_H^2 \Gamma(H - 1/2)^2 \bar{f}(\bar{X}) t^{H-1/2} I_{0+}^{H-1/2} t^{1-2H} I_{1-}^{H-1/2} t^{H-1/2} \bar{f}(\bar{X})^\top \Psi \right) \\ &= L^\top t^{1/2-H} D_{1-}^{H-1/2} t^{2H-1} \left(D_{0+}^{H-1/2} I_{0+}^{H-1/2} \right) t^{1-2H} I_{1-}^{H-1/2} t^{H-1/2} \bar{f}(\bar{X})^\top \Psi \\ &= L^\top t^{1/2-H} \left(D_{1-}^{H-1/2} I_{1-}^{H-1/2} \right) t^{H-1/2} \bar{f}(\bar{X})^\top \Psi = \Psi. \end{aligned}$$

Therefore, W is the left inverse of $Q_{\bar{X}}^H$ and $\ker Q_{\bar{X}}^H = \{0\}$. Moreover, we know $Q_{\bar{X}}^H$ is self-adjoint, hence

$$\text{Im } Q_{\bar{X}}^H = \left[\ker(Q_{\bar{X}}^H)^* \right]^\perp = \left[\ker Q_{\bar{X}}^H \right]^\perp = \{0\}^\perp = L^2([0, 1]; \mathbb{R}^n).$$

It follows that $Q_{\bar{X}}^H$ is bijective. It is also bounded in $L^2([0, 1]; \mathbb{R}^n)$ via Proposition 10, so we can conclude it has a bounded inverse by the inverse mapping theorem.

Finally, the inverse of $Q_{\bar{X}}^H$ must coincide with the left inverse W at (40) and by using the formula for fractional derivatives in Appendix A, we get the second equation for $(Q_{\bar{X}}^H)^{-1}$ in the statement of this lemma.

7. CONCLUSIONS AND FUTURE WORK

In this paper we established the moderate deviations principle for slow-fast systems of the form (1) where the slow component is driven by fractional Brownian motion. There are many interesting potential directions for future work on this topic.

In this paper, the fast motion is driven by standard Brownian motion and is independent of the slow component. This was done in order to focus on the effect of fBm on the tail behavior of the slow component. If the fast component was driven by fBm as well, then one would first need to understand the proper ergodic behavior of the fast process, an issue still not fully resolved, see though [LS20] for some preliminary results in special cases. Feedback from the slow process into the fast process would

also mean interaction of the ergodic behavior of the fast process with the fBm driving the slow process, see [HL20] for partial preliminary results in this direction.

Another interesting direction would be to include “unbounded homogenization” terms in the slow component as done for similar systems driven by standard Brownian motion, see [Spi14].

Lastly, establishing the MDP opens the door to the construction of provably-efficient accelerated Monte Carlo methods, like importance sampling, for the estimation of rare event probabilities. See [SM20] for related work in the case where $H = 1/2$.

We plan to explore these avenues in future works on this topic.

APPENDIX A. FRACTIONAL BROWNIAN MOTION AND PATHWISE STOCHASTIC INTEGRATION

A.1. Fractional Brownian motion: definition and main properties. A fractional Brownian motion (fBm) is a centered Gaussian process $W^H = \{W_t^H : t \in [0, 1]\} \subset L^2(\Omega)$, characterized by its covariance function

$$R_H(t, s) = E(W_t^H W_s^H) = \frac{1}{2} \left(s^{2H} + t^{2H} - |t - s|^{2H} \right).$$

It is straightforward to verify that increments of fBm are stationary. The parameter $H \in (0, 1)$ is usually referred to as the Hurst exponent, Hurst parameter, or Hurst index.

By Kolmogorov’s continuity criterion, such a process admits a modification with continuous sample paths, and we always choose to work with such. In this case one may show in fact that almost every sample path is locally Hölder continuous of any order strictly less than H . It is this sense in which it is often said that the value of H determines the regularity of the sample paths.

Note that when $H = \frac{1}{2}$, the covariance function is $R_{\frac{1}{2}}(t, s) = t \wedge s$. Thus, one sees that $W^{\frac{1}{2}}$ is a standard Brownian motion, and in particular that its disjoint increments are independent. In contrast to this, when $H \neq \frac{1}{2}$, nontrivial increments are not independent. In particular, when $H > \frac{1}{2}$, the process exhibits long-range dependence.

Note moreover that when $H \neq \frac{1}{2}$, the fractional Brownian motion is not a semimartingale, and the usual Itô calculus therefore does not apply.

Another noteworthy property of fractional Brownian motion is that it is self-similar in the sense that, for any constant $a > 0$, the processes $\{W_t^H : t \in [0, 1]\}$ and $\{a^{-H} W_{at}^H : t \in [0, 1]\}$ have the same distribution.

The self-similarity and long-memory properties of the fractional Brownian motion make it an interesting and suitable input noise in many models in various fields such as analysis of financial time series, hydrology, and telecommunications. However, in order to develop interesting models based on fractional Brownian motion, one needs an integration theory with respect to it, which we present in the next subsection.

A.2. Pathwise stochastic integration with respect to fractional Brownian motion.

Stochastic integrals with respect to fractional Brownian motion can be understood, when $H \geq 1/2$, as generalized Stieltjes integral as introduced in the work of Zähle [Zäh98]. Let $f \in L^1([a, b])$ and $\alpha > 0$. The left-sided and right-sided fractional Riemann-Liouville integrals of f of order α are defined for almost all $x \in [a, b]$ by

$$I_{a+}^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x - y)^{\alpha-1} f(y) dy$$

and

$$I_{b-}^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_x^b (y - x)^{\alpha-1} f(y) dy$$

respectively, where $\Gamma(\alpha)$ is the Euler gamma function. The following integration by parts formula holds

$$\int_a^b I_{a+}^\alpha f(x) g(x) dx = \int_a^b f(x) I_{b-}^\alpha g(x) dx$$

for $f \in L^p([a, b])$, $g \in L^q([a, b])$ such that $1/p + 1/q \leq 1 + \alpha$.

For $0 < \alpha < 1$, we can define the fractional derivatives

$$D_{a+}^\alpha f(x) = \frac{d}{dx} I_{a+}^{1-\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_a^x (x-t)^{-\alpha} f(t) dt$$

and

$$D_{b-}^\alpha f(x) = \frac{d}{dx} I_{b-}^{1-\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_x^b (t-x)^{-\alpha} f(t) dt$$

as long as the right hand sides are well-defined. Furthermore, if $f \in I_{a+}^\alpha(L^p([a, b]))$ (respectively $f \in I_{b-}^\alpha(L^p([a, b]))$) and $0 < \alpha < 1$ then the previous fractional derivatives admit the Weyl representation

$$D_{a+}^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(x-a)^\alpha} + \alpha \int_a^x \frac{f(x) - f(y)}{(x-y)^{\alpha-1}} dy \right) \mathbb{1}_{(a,b)}(x)$$

and

$$D_{b-}^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(b-x)^\alpha} + \alpha \int_x^b \frac{f(x) - f(y)}{(y-x)^{\alpha-1}} dy \right) \mathbb{1}_{(a,b)}(x),$$

respectively, for almost all $x \in [a, b]$. There is also the integration by parts formula

$$\int_a^b D_{a+}^\alpha f(x) g(x) dx = \int_a^b f(x) D_{b-}^\alpha g(x) dx$$

for $f \in I_{a+}^\alpha(L^p([a, b]))$, $g \in I_{b-}^\alpha(L^q([a, b]))$ such that $1/p + 1/q \leq 1 + \alpha$.

The upcoming lemma contains a useful technical result in [SKM⁺93].

Lemma 5. *Let $p \geq 1$ and $b > 0$. Then the operator $t^\beta I_{0+}^\alpha t^\gamma$ is bounded in $L^p([0, b])$ if $\alpha > 0$, $\alpha + \beta + \gamma = 0$ and $(\gamma + 1)p > 1$. Meanwhile, the operator $t^\beta I_{1-}^\alpha t^\gamma$ is bounded in $L^p([0, b])$ if $\alpha > 0$, $\alpha + \beta + \gamma = 0$ and $(\alpha + \gamma)p < 1$.*

Proof. This is a consequence of [SKM⁺93, (5.45') and (5.46')]. \square

We refer to [SKM⁺93] for more detailed properties of fractional operators.

Next, let $f(a+) = \lim_{x \searrow 0} f(a+x)$ and $g(b-) = \lim_{x \searrow 0} g(b-x)$ and define

$$\begin{aligned} f_{a+}(x) &= (f(x) - f(a+)) \mathbb{1}_{(a,b)}(x) \\ g_{b-}(x) &= (g(x) - g(b-)) \mathbb{1}_{(a,b)}(x). \end{aligned}$$

We recall from [Zäh98] the definition of generalized Stieltjes fractional integrals with respect to irregular functions (in the sense of which we view the stochastic integrals with respect to fractional Brownian motion appearing in this paper).

Definition 2 (Generalized Stieltjes integral). Suppose that f and g are functions such that $f(a+)$, $g(a+)$ and $g(b-)$ exist, $f_{a+} \in I_{a+}^\alpha(L^p([a, b]))$ and $g_{b-} \in I_{b-}^{1-\alpha}(L^p([a, b]))$ for some $p, q \geq 1$, $1/p + 1/q \leq 1$, $0 < \alpha < 1$. Then the integral of f with respect to g is defined by

$$\int_a^b f dg = (-1)^\alpha \int_a^b D_{a+}^\alpha f(x) D_{b-}^{1-\alpha} g_{b-}(x) dx + f(a+) (g(b-) - g(a+)).$$

Remark 11. If $\alpha p < 1$, under the assumptions of the preceding definition, we have that $f \in I_{a+}^\alpha(L^p([a, b]))$ and we can write

$$\int_a^b f dg = (-1)^\alpha \int_a^b D_{a+}^\alpha f_{a+}(x) D_{b-}^{1-\alpha} g_{b-}(x) dx.$$

In [Zäh98], it was further shown that if f and g are respectively λ and μ -Hölder continuous such that $\lambda + \mu > 1$, then the conditions for the generalized Stieltjes integral $\int_a^b f dg$ are satisfied for $p = q = \infty$ and $\alpha < \lambda, 1 - \alpha < \mu$. In particular, this class of generalized Stieltjes integrals with Hölder continuous f, g coincides with the class of Riemann–Stieltjes integrals studied in [You36] by Young. We note here that any Young integrals appearing in this paper are constructed from Hölder continuous paths of a fractional Brownian motion. Further details are given in [Zäh98, Section 5.1].

A.3. The Cameron-Martin space of fractional Brownian motion. Consider the deterministic kernel

$$K_H(t, s) = c_H s^{1/2-H} \left(\int_s^t (u-s)^{H-3/2} u^{H-1/2} du \right) \mathbf{1}_{\{t>s\}}$$

for which $c_H = (H(2H-1)/\beta(2-2H, H-1/2))^{1/2}$. Slightly abusing notation, we also write K_H for the integral operator

$$K_H g(s) = \int_0^s K_H(s, r) g(r) dr.$$

For $H \geq 1/2$, the operator K_H can be represented as

$$K_H g = c_H \Gamma(H-1/2) I_{0+}^{1-H} t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H} g.$$

Additionally, we denote by \dot{K}_H the “derivation” of the operator K_H , i.e.,

$$(41) \quad \dot{K}_H g = c_H \Gamma(H-1/2) t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H} g.$$

The Cameron-Martin space \mathcal{H}_H associated with W^H is

$$\mathcal{H}_H = \{K_H \hat{g} : \hat{g} \in L^2([0, 1]; \mathbb{R}^m)\},$$

equipped with the inner product $\langle g, f \rangle_{\mathcal{H}_H} = \langle \hat{g}, \hat{f} \rangle_{L^2([0, 1]; \mathbb{R}^m)}$. Note that later on, we will alternate between \hat{g} and $K_H^{-1} g$, which are equivalent ways of writing the same quantity.

In this paper, the noise process we consider in the slow-fast systems we study is of the form

$$\{(W^H(t), B(t)) : t \in [0, 1]\},$$

where B is a m -dimensional standard Brownian motion, W^H is a p -dimensional fractional Brownian motion of Hurst parameter H and they are independent. We will hence need to work with the Cameron-Martin space associated with the process (W^H, B) , which, based on the previous description, is defined to be the space \mathcal{S} given by

$$(42) \quad \{(K_H \hat{g}_1, K_{1/2} \hat{g}_2) : (\hat{g}_1, \hat{g}_2) \in L^2([0, 1]; \mathbb{R}^{m+p})\}.$$

As a Cameron-Martin space, \mathcal{S} is a Hilbert space equipped with the inner product given by

$$\langle (g_1, g_2), (f_1, f_2) \rangle_{\mathcal{S}} = \langle g_1, f_1 \rangle_{\mathcal{H}_H} + \langle g_2, f_2 \rangle_{\mathcal{H}_{1/2}}.$$

Let us now state an important fact regarding the differentiability of elements in \mathcal{H}_H when $H > 1/2$ which we will need throughout the paper.

Lemma 6. *If $H > 1/2$ and $u \in \mathcal{H}_H$ such that $u = K_H \hat{u}$, $\hat{u} \in L^2([0, 1]; \mathbb{R}^n)$, then we have*

$$\begin{aligned} \dot{u}(t) &= \dot{K}_H \hat{u}(t) = c_H \Gamma(H-1/2) t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H} \hat{u}(t) \\ &= c_H t^{H-1/2} \int_0^t (t-s)^{H-3/2} s^{1/2-H} \hat{u}_s ds, \end{aligned}$$

such that $c_H = (H(2H-1)/\beta(2-2H, H-1/2))^{1/2}$.

Meanwhile, if $H = 1/2$ and $u \in \mathcal{H}_{1/2}$, then $\dot{u}_t = \hat{u}_t$.

Proof. This is a direct consequence of formula (41). \square

The following is another important property of the operator \dot{K}_H .

Proposition 10. *The map \dot{K}_H as described in Lemma 6 is a bounded operator in $L^2([0, 1]; \mathbb{R}^n)$.*

Proof. The assumptions of Lemma 5 are satisfied for $p = 2, \alpha = H - 1/2, \beta = 0$ and $\gamma = 1/2 - H$, hence the operator $I_{0+}^{H-1/2} t^{1/2-H}$ is bounded in $L^2([0, 1]; \mathbb{R}^n)$. Since $\dot{K}_H = t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H}$ based on Lemma 6, this implies

$$\begin{aligned} \|\dot{K}_H f\|_{L^2([0, 1]; \mathbb{R}^n)} &= \|t^{H-1/2} I_{0+}^{H-1/2} t^{1/2-H} f\|_{L^2([0, 1]; \mathbb{R}^n)} \leq \|I_{0+}^{H-1/2} t^{1/2-H} f\|_{L^2([0, 1]; \mathbb{R}^n)} \\ &\leq C \|f\|_{L^2([0, 1]; \mathbb{R}^n)}. \end{aligned}$$

□

For more details about fractional Brownian motion, we refer the reader to the monographs [BHØZ08, Nua06].

A.4. Results related to Young integrals. The two results presented here provide us with a way of bounding Young integrals and with a version of change of variable formula for differential equations that contain Young integrals, respectively.

Lemma 7 (Young–Loéve's inequality). *Let f and g be respectively α and β -Hölder continuous, such that $\alpha + \beta > 1$. Then a.s one has*

$$\left| \int_r^t f_s dg_s - f_r (g_t - g_r) \right| \leq C |f|_\alpha |g|_\beta |t - r|^{\alpha + \beta}.$$

Moreover, assume f is bounded then

$$\left| \int_r^t f_s dg_s \right| \leq C |f|_\alpha |g|_\beta |t - r|^{\alpha + \beta} + |f|_\infty |t - r|^\beta \leq C |f|_\alpha |g|_\beta |t - r|^\beta.$$

Proof. Refer to [FV10, Proposition 6.4]. □

Theorem 3. *For $i = 1, \dots, m$, let $0 < \alpha_i < 1/2$, $f^i \in I_{0+}^{\alpha_i}(L^2([0, b]))$ be bounded and $g_{b-}^i \in I_{b-}^{1-\alpha_i}(L^2([0, b]))$, where the function g_{b-}^i is defined below Lemma 5. Moreover, assume $h = (h^1, \dots, h^m)$ such that*

$$h_t^i = h_0^i + \int_0^t f_s^i dg_s^i.$$

Then for any C^1 mapping $F : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^n$ such that $\frac{\partial F}{\partial x_i} \in C^1, i = 1, \dots, m$ and $r \leq t \leq T$, it holds that

$$F(h_t, t) - F(h_r, r) = \sum_{i=1}^m \int_r^t \frac{\partial F}{\partial x_i}(h_s, s) f_s^i dg_s^i + \int_r^t \frac{\partial F}{\partial s}(h_s, s) ds.$$

In particular, this change of variable formula applies to the special case when f_i and g_i are respectively λ_i and μ_i -Hölder continuous such that $\lambda_i + \mu_i > 1, i = 1, \dots, m$.

Proof. For the change of variable formula in the general case, we refer to [Zäh99, Theorem 5.2].

Now, let us consider the special case and assume there is a constant C such that $|f_i|_{\lambda_i}, |g_i|_{\mu_i} < C$ and $\lambda_i + \mu_i > 1$ for $1 \leq i \leq m$. Then one can choose α_i in the interval $(0, 1/2)$ such that $\lambda_i > \alpha_i$ and $\mu_i > 1 - \alpha_i$ for $1 \leq i \leq m$. Based on the previous fact, Lemmas 13.2 and 13.2' in [SKM⁺93] imply respectively that

$$f^i \in I_{0+}^{\alpha_i}(L^2([0, b])), \quad g_{b-}^i \in I_{b-}^{1-\alpha_i}(L^2([0, b])).$$

Moreover, f^i as Hölder continuous functions on $[0, b]$ are necessarily bounded. Consequently, the general change of variable formula covers this particular case. □

APPENDIX B. REGULARITY RESULTS AND OTHER TECHNICAL LEMMAS

This appendix gathers results related to Poisson equations as well as the technical lemmas required for the analysis of the control problems.

B.1. Results related to Poisson equations. The following theorem is a consequence of [PV01, Theorem 2] and [PV05, Theorem 3] for solutions of Poisson equations. Let \mathcal{L} be the infinitesimal generator defined in (5).

Theorem 4. *Recall $C^{2,\zeta}(\mathbb{R}^n \times \mathcal{Y})$ for some $\zeta > 0$ is the function space defined at the beginning of Section 2. Let $h \in C^{2,\zeta}(\mathbb{R}^n \times \mathcal{Y})$ such that*

$$\int_{\mathcal{Y}} h(x, y) \mu(dy) = 0$$

and that for some positive constants K and D_h ,

$$|h(x, y)| + |\nabla_x h(x, y)| + |\nabla_x^2 h(x, y)| \leq K(1 + |y|^{D_h})$$

uniformly with respect to x . Then, there is a unique solution to

$$\mathcal{L}u(x, y) = -h(x, y), \quad \int_{\mathcal{Y}} u(x, y) \mu(dy) = 0.$$

Moreover, $u(\cdot, y) \in C^2$, $\nabla_x^2 u \in C(\mathbb{R}^n \times \mathcal{Y})$ and there exists a positive constant M such that

$$|u(x, y)| + |\nabla_y u(x, y)| + |\nabla_x u(x, y)| + |\nabla_x^2 u(x, y)| + |\nabla_y \nabla_x u(x, y)| \leq M(1 + |y|^{D_h}).$$

Remark 12. Consider the Poisson equation in (6). Under Conditions H1 and H2-A, Theorem 4 states that there exists a positive constant C such that, uniformly,

$$(43) \quad |\phi(x, y)| + |\nabla_y \phi(x, y)| + |\nabla_x \phi(x, y)| + |\nabla_x^2 \phi(x, y)| + |\nabla_y \nabla_x \phi(x, y)| < C.$$

On the other hand, under Conditions H1 and H2-B, Theorem 4 states that there exists a positive constant C such that, uniformly with respect to x ,

$$(44) \quad |\phi(x, y)| + |\nabla_y \phi(x, y)| + |\nabla_x \phi(x, y)| + |\nabla_x^2 \phi(x, y)| + |\nabla_y \nabla_x \phi(x, y)| < C(1 + |y|^{D_g}).$$

B.2. Ancillary results related to the control problems. This subsection gathers all technical results related to the study of the control problems appearing throughout the paper.

Lemma 8 (Lemma 5.2 in [HSS19]). *Let H, H' be Hilbert spaces and $a : H \rightarrow H'$ be a bounded linear operator. Moreover, let $q = aa^*$ and q^{-1} be the inverse of q . Then for any $u \in H$,*

$$\|a^* q^{-1} a u\|_H \leq \|u\|_H.$$

Lemma 9. *Assume that for all x and non-zero $z \in \mathbb{R}^n$ and for all $y \in \mathbb{R}^d$,*

$$\left\langle \int_{\mathcal{Y}} \nabla_y \phi(x, y) \sigma(y) (\nabla_y \phi(x, y) \sigma(y))^{\top} \mu(dy) z, z \right\rangle > 0.$$

Then, the operator Q_x^H defined in (36) is invertible and its inverse $(Q_x^H)^{-1}$ is a bounded in $L^2([0, 1]; \mathbb{R}^n)$.

Proof. Using the operators π, π^*, ρ, ρ^* defined in Section 5.5, we have $Q_x^H h = (\pi \pi^* + \rho \rho^*) h$ with

$$\rho \rho^* h(t) = \int_{\mathcal{Y}} \nabla_y \phi(x, y) \sigma(y) (\nabla_y \phi(x, y) \sigma(y))^{\top} h(t, y) \mu(dy).$$

Furthermore, $\pi\pi^*, \rho\rho^*$ are positive and self-adjoint operators, which means that Q_x^H is also positive and self-adjoint. In addition, the fact that $Q_x^H \geq \rho\rho^*$ and Condition H2-A or H2-B imply that $(Q_x^H)^2 \geq (\rho\rho^*)^2 > 0$. This leads to

$$\begin{aligned} \inf_{\|h\|_{L^2([0,1];\mathbb{R}^n)}=1} \|Q_x^H h\|_{L^2([0,1];\mathbb{R}^n)} &= \inf_{\|h\|_{L^2([0,1];\mathbb{R}^n)}=1} \langle (Q_x^H)^2 h, h \rangle_{L^2([0,1];\mathbb{R}^n)} \\ &\geq \inf_{\|h\|_{L^2([0,1];\mathbb{R}^n)}=1} \langle (\rho\rho^*)^2 h, h \rangle_{L^2([0,1];\mathbb{R}^n)} > 0, \end{aligned}$$

so that Q_x^H is bounded from below and $\ker Q_x^H = \{0\}$. This combined with self-adjointness implies

$$\text{Im } Q_x^H = [\ker(Q_x^H)^*]^\perp = [\ker Q_x^H]^\perp = \{0\}^\perp = L^2([0, 1]; \mathbb{R}^n).$$

It follows that Q_x^H is bijective. The operator Q_x^H is also bounded in $L^2([0, 1]; \mathbb{R}^n)$ via Proposition 10, so we can conclude it has a bounded inverse by the inverse mapping theorem. \square

Lemma 10. *It can be assumed that there exists a finite constant N such that, almost surely, the control process w^ϵ appearing in the variational representation (12) satisfies*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} \leq N.$$

Proof. This is an immediate consequence of [Zha09, Theorem 3.2]. \square

Lemma 11. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 |\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 ds < N$$

for some finite constant N . Then, under Condition H1, it holds that for $\epsilon_0 > 0$ small enough,

$$\sup_{\epsilon < \epsilon_0} \mathbb{E} \left[\int_0^1 \left| Y_s^{\epsilon, w^\epsilon} \right|^2 ds \right] < C$$

for some constant $C > 0$, which further implies that

$$\mathbb{E} \left[\sup_{t \in [0, 1]} \left| Y_t^{\epsilon, w^\epsilon} \right| \right] \leq \frac{C}{\sqrt{\epsilon}}.$$

Proof. The first estimate was proven in [SM20, Lemma 3.1]. For the second estimate, the dissipative property of the drift coefficient of Y^{ϵ, w^ϵ} and Itô's formula yield

$$\begin{aligned} Y_t^{\epsilon, w^\epsilon} &= e^{-\frac{1}{\epsilon}\Gamma t} y_0 + \int_0^t \frac{1}{\epsilon} e^{-\frac{1}{\epsilon}(t-s)} \zeta(Y_s^{\epsilon, w^\epsilon}) ds + \int_0^t \frac{h(\epsilon)}{\sqrt{\epsilon}} e^{-\frac{1}{\epsilon}(t-s)} \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds \\ &\quad + \int_0^t \frac{1}{\sqrt{\epsilon}} e^{-\frac{1}{\epsilon}(t-s)} \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s. \end{aligned}$$

We then apply the Burkholder-Davis-Gundy inequality to the Itô integral term and Hölder's inequality to the Riemann integral terms to get

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, 1]} \left| Y_t^{\epsilon, w^\epsilon} \right| \right] &\leq \sup_{t \in [0, 1]} e^{-\frac{1}{\epsilon}\Gamma t} y_0 + \frac{1}{\epsilon} \sqrt{\int_0^t e^{-\frac{2}{\epsilon}\Gamma(t-s)} ds} \sqrt{\mathbb{E} \left[\int_0^1 |\zeta(Y_s^{\epsilon, w^\epsilon})|^2 ds \right]} \\ &\quad + \frac{h(\epsilon)}{\sqrt{\epsilon}} \sqrt{\int_0^t e^{-\frac{2}{\epsilon}\Gamma(t-s)} ds} \sqrt{\int_0^1 |\dot{v}_s^\epsilon|^2 ds} + \frac{1}{\sqrt{\epsilon}} \sqrt{\mathbb{E} \left[\int_0^1 e^{-\frac{2}{\epsilon}\Gamma(t-s)} |\sigma(Y_s^{\epsilon, w^\epsilon})|^2 ds \right]}. \end{aligned}$$

Since $\sigma(y)\sigma^T(y)$ is bounded and $\zeta(y)$ is sublinear, the first estimate of this lemma can be applied to the expression $\mathbb{E}\left[\int_0^1 |\zeta(Y_s^{\epsilon, w^\epsilon})|^2 ds\right]$. Then, the simple fact that $\int_r^t e^{-\frac{2}{\epsilon}\Gamma(t-s)} ds \leq \epsilon \int_0^\infty e^{-2\Gamma s} ds = \frac{\epsilon}{2\Gamma}$ implies that

$$\mathbb{E}\left[\sup_{t \in [0,1]} |Y_t^{\epsilon, w^\epsilon}|\right] \leq C\left(\frac{1}{\sqrt{\epsilon}} + h(\epsilon)\right) \leq \frac{C}{\sqrt{\epsilon}}.$$

□

Lemma 12. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 |\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 ds < N$$

for some finite constant N .

(i) *Under Conditions H1 and H2-A, there exist constants C that change from line to line such that*

$$\begin{aligned} & \mathbb{E}\left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \right| \right] \leq C\rho, \\ & \mathbb{E}\left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \right|^2 \right] \leq C\rho \\ & \mathbb{E}\left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds \right|^2 \right] \leq C\rho, \\ & \mathbb{E}\left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^2 \right] \leq C\rho, \\ & \mathbb{E}\left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^2 \right] \leq C\rho. \end{aligned}$$

(ii) Under Conditions H1 and H2-B, there exist constants C that change from line to line such that for any q in $(1, \frac{1}{D_f + D_g}]$, we have

$$\begin{aligned} & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \right|^q \right] \leq C \rho^{q-1}, \\ & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \right|^{2q} \right] \leq C \rho^{q-1} \\ & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds \right|^{2q} \right] \leq C \rho^{q-1}, \\ & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^{2q} \right] \leq C \rho^{q-1}, \\ & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_0^t f(Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^{2q} \right] \leq C \rho^{q-1}. \end{aligned}$$

Proof. We start with part (i). The first estimate is straightforward due to the boundedness of $\nabla_x \phi(x, y)$ stated in (43) and the boundedness of $g(x, y)$ guaranteed by Condition H2-A. For the second estimate, we assume that $0 \leq r \leq t \leq 1$ and apply the Burkholder-Davis-Gundy inequality to obtain

$$\begin{aligned} & \mathbb{E} \left[\sup_{\substack{0 \leq r \leq t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \right|^2 \right] \\ & \leq \mathbb{E} \left[\left(\int_r^{r+\rho} \left| \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \right|^2 ds \right) \right] \leq C \rho. \end{aligned}$$

For the third estimate, we can write

$$\begin{aligned} & \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds \right|^2 \right] \\ & \leq \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left(\int_r^t \left| \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \right|^2 ds \int_0^1 |\dot{v}_s^\epsilon|^2 ds \right) \right] \\ & \leq C \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left(\int_r^t \left| \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \right|^2 ds \right) \right] \leq C \rho. \end{aligned}$$

The last inequality in part (i) is a consequence of the boundedness of $\sigma(y) \sigma^\top(y)$ in Condition H1 and the boundedness of $\nabla_x \phi(x, y)$ stated in (43) (requiring Condition H2-A). Finally, the two remaining estimates of part (i) are derived similarly to the previous one.

We continue with part (ii). For the first inequality, the sublinear growth of $\nabla_x \phi(x, y)$ in y stated at (44) (requiring Condition H2-B) and the sublinear growth of $g(x, y)$ in y from Condition H2-B imply for any

q in $\left(1, \frac{1}{D_g}\right]$,

$$\begin{aligned} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \right|^q \right] &\leq C \mathbb{E} \left[\left(\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \int_r^t 1 + |Y_s^{\epsilon, w^\epsilon}|^{2D_g} ds \right)^q \right] \\ &\leq C \rho^{q-1} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \int_r^t \left(1 + |Y_s^{\epsilon, w^\epsilon}|^{2qD_g} \right) ds \right] \leq C \rho^{q-1}, \end{aligned}$$

where the last inequality is due to Lemma 11. For the second estimate, assume that $0 \leq r \leq t \leq 1$. Then, the Burkholder-Davis-Gundy inequality combined with the sublinear growth of $\nabla_y \phi(x, y)$ in y (requiring Condition H2-B) and the boundedness of $\sigma(y)\sigma^\top(y)$ in Condition H1 imply that for any q in $\left(1, \frac{1}{D_g}\right]$,

$$\begin{aligned} \mathbb{E} \left[\sup_{\substack{0 \leq r \leq t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \right|^{2q} \right] &\leq \mathbb{E} \left[\left(\int_r^{r+\rho} \left| \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \right|^2 ds \right)^q \right] \\ &\leq C \rho^{q-1} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \int_r^t 1 + |Y_s^{\epsilon, w^\epsilon}|^{2qD_g} ds \right] \leq C \rho^{q-1}. \end{aligned}$$

The arguments for the three remaining estimates of part (ii) are similar, so we will handle one case only. The sublinear growth of $\nabla_x \phi(x, y)$ in y stated at (44) (requiring Condition H2-B) and sublinear growth of $f(y)$ in y in Condition H2-B imply that for any q in $\left(1, \frac{1}{D_f + D_g}\right]$,

$$\begin{aligned} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^{2q} \right] &\leq \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left(\int_r^t \left| \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) \right|^2 ds \right)^q \left(\int_0^1 |\dot{u}_s^\epsilon|^2 ds \right)^q \right] \\ &\leq C \rho^{q-1} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \int_r^t 1 + |Y_s^{\epsilon, w^\epsilon}|^{2q(D_f + D_g)} ds \right] \leq C \rho^{q-1}, \end{aligned}$$

where the last inequality is once again obtained using Lemma 11. \square

Lemma 13. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 [|\dot{u}_s^\epsilon|^2 + |\dot{v}_s^\epsilon|^2] ds < N$$

for some finite constant N . Under Condition H1, for $0 < \alpha \leq 1/2$, we have the almost sure Hölder estimate

$$|Y^{\epsilon, w^\epsilon}|_\alpha \leq \frac{C}{\sqrt{\epsilon}}.$$

Proof. Without loss of generality, let us assume $t > r$. The dissipative property of the drift coefficient of Y^{ϵ, w^ϵ} and Itô's formula yield

$$\begin{aligned} Y_t^{\epsilon, w^\epsilon} &= e^{-\frac{1}{\epsilon}\Gamma(t-r)}Y_r^{\epsilon, w^\epsilon} + \int_r^t \frac{1}{\epsilon}e^{-\frac{1}{\epsilon}(t-s)}\zeta(Y_s^{\epsilon, w^\epsilon})ds + \int_r^t \frac{h(\epsilon)}{\sqrt{\epsilon}}e^{-\frac{1}{\epsilon}(t-s)}\sigma(Y_s^{\epsilon, w^\epsilon})\dot{v}_s^\epsilon ds \\ &\quad + \int_r^t \frac{1}{\sqrt{\epsilon}}e^{-\frac{1}{\epsilon}(t-s)}\sigma(Y_s^{\epsilon, w^\epsilon})dB_s. \end{aligned}$$

Now, by subtracting $Y_r^{\epsilon, w^\epsilon}$ from both sides and applying Hölder's inequality along with the Burkholder-Davis-Gundy inequality, we get

$$\begin{aligned} \mathbb{E}\left[\left|Y_t^{\epsilon, w^\epsilon} - Y_r^{\epsilon, w^\epsilon}\right|\right] &\leq \left|e^{-\frac{1}{\epsilon}\Gamma(t-r)} - 1\right| \mathbb{E}\left[\left|Y_r^{\epsilon, w^\epsilon}\right|\right] + \frac{1}{\epsilon}\sqrt{\int_r^t e^{-\frac{2}{\epsilon}\Gamma(t-s)}ds} \mathbb{E}\left[\sqrt{\int_0^1 |\zeta(Y_s^{\epsilon, w^\epsilon})|^2 ds}\right] \\ &\quad + \frac{h(\epsilon)}{\sqrt{\epsilon}}\sqrt{\int_r^t e^{-\frac{2}{\epsilon}\Gamma(t-s)}ds} \mathbb{E}\left[\sqrt{\int_0^1 |\dot{v}_s^\epsilon|^2 ds}\right] \\ (45) \quad &\quad + \frac{1}{\sqrt{\epsilon}}\mathbb{E}\left[\sqrt{\int_r^t e^{-\frac{2}{\epsilon}\Gamma(t-s)}|\sigma(Y_s^{\epsilon, w^\epsilon})\sigma^T(Y_s^{\epsilon, w^\epsilon})|ds}\right]. \end{aligned}$$

To bound the first term on the right-hand side, we combine the second estimate in Lemma 11 and the fact that $e^{-\frac{1}{\epsilon}\Gamma(t-r)} - 1 = \frac{1}{\epsilon}\int_r^t e^{-\frac{1}{\epsilon}\Gamma(t-s)}ds \leq |t-r|$. For the second term, note that $\int_r^t e^{-\frac{2}{\epsilon}\Gamma(t-s)}ds = C\epsilon|t-r|$. Moreover, the sublinearity of $\zeta(y)$ and the first estimate in Lemma 11 yield a finite bound on the expression $\mathbb{E}\left[\sqrt{\int_0^1 |\zeta(Y_s^{\epsilon, w^\epsilon})|^2 ds}\right]$. The third term on the right-hand side of (45) can be treated similarly with the help of Lemma 10. Regarding the last term, recall that $\sigma(y)\sigma^T(y)$ is bounded in Condition H1. Thus, we have

$$\mathbb{E}\left[\sup_{0 \leq r, t \leq 1} \left|Y_t^{\epsilon, w^\epsilon} - Y_r^{\epsilon, w^\epsilon}\right|\right] \leq C\frac{1}{\sqrt{\epsilon}}|t-r|^{1/2}.$$

The Kolmogorov Continuity Theorem then yields the almost sure Hölder continuity of Y^{ϵ, w^ϵ} . \square

Lemma 14. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 \left[|\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2\right] ds < N$$

for some finite constant N . Under Conditions H1 and H2-A or H2-B, there exists a constant C and ϵ_0 small enough such that for $0 < \beta \leq 1/2$,

$$\mathbb{E}\left[\sup_{\epsilon < \epsilon_0} \left|X^{\epsilon, w^\epsilon}\right|_\beta\right] < C.$$

Proof. We begin by proving the result under Conditions H1 and H2-A. According to Condition H2-A, $f(x, y)$ is Lipschitz-continuous and bounded, so that $f(x, y)$ is also γ -Hölder continuous for $0 < \gamma \leq 1$. This further implies

$$\begin{aligned} \mathbb{E}\left[\left|f\left(X_t^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}\right) - f\left(X_r^{\epsilon, w^\epsilon}, Y_r^{\epsilon, w^\epsilon}\right)\right|\right] &\leq \mathbb{E}\left[\left|f\left(X_t^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}\right) - f\left(X_r^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}\right)\right|\right] \\ &\quad + \mathbb{E}\left[\left|f\left(X_r^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}\right) - \left(X_r^{\epsilon, w^\epsilon}, Y_r^{\epsilon, w^\epsilon}\right)\right|\right] \\ &\leq C\mathbb{E}\left[\left|X_t^{\epsilon, w^\epsilon} - X_r^{\epsilon, w^\epsilon}\right|\right] + \mathbb{E}\left[\left|Y_t^{\epsilon, w^\epsilon} - Y_r^{\epsilon, w^\epsilon}\right|^\gamma\right] \\ &\leq C\left(\mathbb{E}\left[\left|X_t^{\epsilon, w^\epsilon} - X_r^{\epsilon, w^\epsilon}\right|\right] + \epsilon^{-\frac{\gamma}{2}}|t-r|^\gamma\right), \end{aligned}$$

and hence that for $0 < \gamma \leq 1$,

$$\mathbb{E}\left[\left|f\left(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}\right)\right|_\gamma\right] \leq C\left(\mathbb{E}\left[\left|X^{\epsilon, w^\epsilon}\right|_\gamma\right] + \epsilon^{-\frac{\gamma}{2}}\right).$$

This last estimate, together with the Young-Loéve inequality in Lemma 7 imply that for $1 - H < \beta \leq 1$,

$$\begin{aligned} \mathbb{E}\left[\left|\int_r^t f\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) dW_s^H\right|\right] &\leq C\mathbb{E}\left[\left|f\left(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}\right)\right|_\beta\right] |t - r|^H \\ (46) \quad &\leq C\mathbb{E}\left[\left|X^{\epsilon, w^\epsilon}\right|_\beta + \epsilon^{-\frac{\beta}{2}}\right] |t - r|^H. \end{aligned}$$

Meanwhile, a similar estimate to the one stated in part (i) of Lemma 12 states that

$$\mathbb{E}\left[\left|\int_0^t f\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) \dot{u}_s^\epsilon ds\right|\right] \leq C |r - t|^{\frac{1}{2}}.$$

Moreover, boundedness of $g(x, y)$ in Condition H2-A yields

$$\mathbb{E}\left[\left|\int_r^t g\left(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}\right) ds\right|\right] \leq C |t - r|.$$

Thus, using the estimate $\mathbb{E}\left[\left|Y^{\epsilon, w^\epsilon}\right|_\alpha\right] \leq C \frac{1}{\sqrt{\epsilon}}, \alpha \leq \frac{1}{2}$ in Lemma 13 (which requires Condition H1), we can deduce that, for $1 - H < \beta \leq 1/2$,

$$\mathbb{E}\left[\left|X_t^{\epsilon, w^\epsilon} - X_r^{\epsilon, w^\epsilon}\right|\right] \leq C\left(\mathbb{E}\left[\sqrt{\epsilon} \left|X^{\epsilon, w^\epsilon}\right|_\beta\right] |t - r|^H + |t - r|^{\frac{1}{2}} + |t - r|\right),$$

and consequently,

$$\mathbb{E}\left[\left|X^{\epsilon, w^\epsilon}\right|_\beta\right] \leq C\left(\mathbb{E}\left[\sqrt{\epsilon} \left|X^{\epsilon, w^\epsilon}\right|_\beta\right] |t - r|^{H-\beta} + |t - r|^{\frac{1}{2}-\beta} + |t - r|^{1-\beta}\right).$$

Now, by choosing ϵ_0 small enough, we get $\mathbb{E}\left[\left|X^{\epsilon, w^\epsilon}\right|_\beta\right] \leq C$ for some constant C . Since for $0 < \beta_1 \leq \beta_2 \leq 1$, β_2 -Hölder continuity of X^{ϵ, w^ϵ} implies β_1 -Hölder continuity, the conclusion follows.

We now present a proof of the claim under Conditions H1 and H2-B. Under Condition H2-B, $f(y)$ is M_f -Hölder continuous while Y^{ϵ, w^ϵ} is $\frac{1}{2}$ -Hölder continuous by Lemma 13, so that

$$\mathbb{E}\left[\left|f\left(Y_t^{\epsilon, w^\epsilon}\right) - f\left(Y_r^{\epsilon, w^\epsilon}\right)\right|\right] \leq C \left|Y_t^{\epsilon, w^\epsilon} - Y_r^{\epsilon, w^\epsilon}\right|^{M_f} \leq C \epsilon^{-\frac{M_f}{2}} |t - r|^{\frac{M_f}{2}}$$

or equivalently

$$\mathbb{E}\left[\left|f\left(Y^{\epsilon, w^\epsilon}\right)\right|_{\frac{M_f}{2}}\right] \leq C \epsilon^{-\frac{M_f}{2}}.$$

Then, the Young-Loéve inequality in Lemma 7 implies that, for $1 - \frac{M_f}{2} < K \leq H$,

$$\begin{aligned} \mathbb{E}\left[\left|\int_r^t f\left(Y_s^{\epsilon, w^\epsilon}\right) dW_s^H\right|\right] &\leq \left|Y^{\epsilon, w^\epsilon}\right|_{\frac{M_f}{2}} |t - r|^{\frac{M_f}{2}+K} |W^H|_K + \left|f\left(Y_r^{\epsilon, w^\epsilon}\right)\right| |W_t^H - W_r^H| \\ &\leq C\left(\left|Y^{\epsilon, w^\epsilon}\right|_{\frac{M_f}{2}} |t - r|^{\frac{M_f}{2}+K} + \mathbb{E}\left[\sup_{t \in [0, 1]} \left|Y_t^{\epsilon, w^\epsilon}\right|^{D_f}\right] |t - r|^K\right) \\ (47) \quad &\leq C\left(\epsilon^{-\frac{M_f}{2}} + \epsilon^{-\frac{1}{2}}\right) |t - r|^K \leq C \epsilon^{-\frac{1}{2}} |t - r|^K, \end{aligned}$$

where the first inequality is obtained by Condition H2-B and the last inequality is a consequence of the estimate $\mathbb{E}\left[\left|Y^{\epsilon, w^\epsilon}\right|_\alpha\right] \leq C \frac{1}{\sqrt{\epsilon}}, \alpha \leq \frac{1}{2}$ in Lemma 13 (which requires Condition H1). Moreover, similar calculations to those performed in the proof of part (ii) of Lemma 12 yield that, for any q in $\left(1, \frac{1}{D_f}\right]$,

$$\mathbb{E}\left[\left|\int_r^t f\left(Y_s^{\epsilon, w^\epsilon}\right) \dot{u}_s^\epsilon ds\right|\right] \leq C |t - r|^{\frac{1}{2}-\frac{1}{2q}},$$

as well as

$$\mathbb{E} \left[\left| \int_r^t g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \right| \right] \leq C |t - r|^{\frac{1}{2}}.$$

Consequently, we have

$$\mathbb{E} \left[\left| X_t^{\epsilon, w^\epsilon} - X_r^{\epsilon, w^\epsilon} \right| \right] \leq C \left(|t - r|^K + \sqrt{\epsilon} h(\epsilon) |t - r|^{\frac{1}{2} - \frac{1}{2q}} + |t - r|^{\frac{1}{2}} \right).$$

By choosing ϵ_0 small enough and noting that $K > 1 - \frac{M_f}{2} \geq \frac{1}{2}$, we arrive at

$$\mathbb{E} \left[\sup_{\epsilon < \epsilon_0} \left| X^{\epsilon, w^\epsilon} \right|_\beta \right] < C$$

for $0 \leq \beta \leq \frac{1}{2}$. \square

Lemma 15. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 \left[|\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 \right] ds < N$$

for some finite constant N . Then, the following two assertions hold.

(i) Under Conditions H1 and H2-A, there exists a constant C such that for any β in $(1 - H, 1]$,

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq C \epsilon^{-\frac{\beta}{2}}$$

and

$$\mathbb{E} \left[\left| \sup_{t \in [0, 1]} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq C \epsilon^{-\frac{1}{2}}.$$

(ii) Under Conditions H1 and H2-B, there exists a constant C such that

$$\mathbb{E} \left[\sup_{t \in [0, 1]} \left| \int_0^t f(Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq C \epsilon^{-\frac{M_f}{2}}$$

and

$$\mathbb{E} \left[\sup_{t \in [0, 1]} \left| \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq C \epsilon^{-\frac{M_k}{2}}.$$

Proof. We begin by proving part (i). The first estimate is immediate based on Lemma 14 and the estimate at (46). Regarding the second estimate, the inequality at (43) and Condition H2-A imply that $\nabla_x \phi(x, y) f(x, y)$ is Lipschitz continuous. Hence, by the Young-Loeve inequality for $1 - H < \beta \leq 1$, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \\ \leq \mathbb{E} \left[\left| \nabla_x \phi(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}) f(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}) \right|_\beta \right] \\ \leq |\nabla_x \phi(x, y) f(x, y)|_{\text{Lip}} \mathbb{E} \left[\left| X^{\epsilon, w^\epsilon} \right|_\beta + \left| Y^{\epsilon, w^\epsilon} \right|_\beta \right] \\ \leq C \left(1 + \epsilon^{-\frac{1}{2}} \right), \end{aligned}$$

where the last inequality is a consequence of Lemmas 13 and 14.

We now proceed to the proof of part (ii). For the first estimate, we perform a similar calculation to the one that was done at (47) (this requires Conditions H1 and H2-B) and get

$$\mathbb{E} \left[\sup_{t \in [0,1]} \left| \int_0^t f(Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] \leq C \left(\left| Y^{\epsilon, w^\epsilon} \right|_{\frac{M_f}{2}} + \mathbb{E}[(y_0)^{D_f}] \right) \leq C \epsilon^{-\frac{M_f}{2}}.$$

Next, under Conditions H1 and H2-B, the M_k -Hölder continuity of $\nabla_x \phi(x, y) f(x)$ together with the estimates in Lemmas 13 and 14 yield

$$\begin{aligned} \mathbb{E} \left[\left| \nabla_x \phi(X_r^{\epsilon, w^\epsilon}, Y_r^{\epsilon, w^\epsilon}) f(Y_r^{\epsilon, w^\epsilon}) - \nabla_x \phi(X_t^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}) f(Y_t^{\epsilon, w^\epsilon}) \right| \right] \\ \leq \mathbb{E} \left[\left| X_r^{\epsilon, w^\epsilon} - X_t^{\epsilon, w^\epsilon} \right|_{\frac{M_k}{2}}^{M_k} \right] + \mathbb{E} \left[\left| Y_r^{\epsilon, w^\epsilon} - Y_t^{\epsilon, w^\epsilon} \right|_{\frac{M_k}{2}}^{M_k} \right] \\ \leq C \left(1 + \epsilon^{-\frac{M_k}{2}} \right) |r - t|^{\frac{M_k}{2}}, \end{aligned}$$

so that

$$\mathbb{E} \left[\left| \nabla_x \phi(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}) f(Y^{\epsilon, w^\epsilon}) \right|_{\frac{M_k}{2}} \right] \leq C \epsilon^{-\frac{M_k}{2}}.$$

Therefore, as $\frac{M_k}{2} + H > 1$ in Condition H2-B, we can apply the Young-Loéve inequality to obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(Y_s^{\epsilon, w^\epsilon}) dW_s^H \right| \right] &\leq C \left(|W^H|_H \mathbb{E} \left[\left| \nabla_x \phi(X^{\epsilon, w^\epsilon}, Y^{\epsilon, w^\epsilon}) f(Y^{\epsilon, w^\epsilon}) \right|_{\frac{M_k}{2}} \right] \right. \\ &\quad \left. + \mathbb{E}[|\nabla_x \phi(x_0, y_0) f(y_0)|] \right) \\ &\leq C \epsilon^{-\frac{M_k}{2}}. \end{aligned}$$

□

Lemma 16. *Assume $w^\epsilon \in \mathcal{S}$ is a control such that*

$$\sup_{\epsilon > 0} \|w^\epsilon\|_{\mathcal{S}} = \sup_{\epsilon > 0} \int_0^1 \left[|\hat{u}_s^\epsilon|^2 + |\hat{v}_s^\epsilon|^2 \right] ds < N$$

for some finite constant N . Under Conditions H1 and either H2-A or H2-B, there exists a constant C such that

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \eta_t^{\epsilon, w^\epsilon} \right|^2 \right] < C.$$

Furthermore, this implies for any $\rho > 0$,

$$\mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} \left(\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s) \right) ds \right| \right] \leq C \rho.$$

Proof. Under Condition H2-A or H2-B, $\nabla_x \bar{g}(x)$ is bounded. This fact, combined with equation (27) and the fact that X^{ϵ, w^ϵ} converges to \bar{X} in probability, implies that there exists some constant C such that

$$(48) \quad \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} \left(\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s) \right) ds \right|^2 \right] \leq C \int_0^1 \mathbb{E} \left[\sup_{0 \leq r \leq s} \left| \eta_r^{\epsilon, w^\epsilon} \right|^2 \right] ds.$$

In addition, based on equation (25), we have

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \eta_t^{\epsilon, w^\epsilon} \right|^2 \right] &\leq C \left(\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) \dot{v}_s^\epsilon ds \right|^2 \right] \right. \\
&\quad + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} (\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)) ds \right|^2 \right] \\
&\quad + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \\
(49) \quad &\quad \left. + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^2 \right] + \mathbb{E} \left[\sup_{0 \leq t \leq 1} |R_2^\epsilon(t)|^2 \right] \right),
\end{aligned}$$

with

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq t \leq 1} |R_2^\epsilon(t)|^2 \right] &\leq C \left(\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{\sqrt{\epsilon}}{h(\epsilon)} (\phi(X_t^{\epsilon, w^\epsilon}, Y_t^{\epsilon, w^\epsilon}) - \phi(x_0, y_0)) \right|^2 \right] \right. \\
&\quad + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{\sqrt{\epsilon}}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) g(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) ds \right|^2 \right] \\
&\quad + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \epsilon \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \dot{u}_s^\epsilon ds \right|^2 \right] \\
&\quad + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \\
&\quad \left. + \mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{1}{h(\epsilon)} \int_0^t \nabla_y \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) \sigma(Y_s^{\epsilon, w^\epsilon}) dB_s \right|^2 \right] \right).
\end{aligned}$$

We will estimate the terms on the right-hand side of (49), starting with those which contain Young integrals. Condition H2-A guarantees that there exists some β in $[0, 1]$ such that $\beta + H > 1$ and $h(\epsilon)^{-1} \epsilon^{-\frac{\beta}{2}} \rightarrow 0$ as $\epsilon \rightarrow 0$, so that part (i) of Lemma 15 (which requires Conditions H1 and H2-A) yields

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{1}{h(\epsilon)} \int_0^t f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \leq C h(\epsilon)^{-2} \epsilon^{-\beta} \rightarrow 0.$$

Part (i) of Lemma 15 also implies that

$$\mathbb{E} \left[\sup_{t \in [0, 1]} \left| \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \leq C \frac{1}{h(\epsilon)^2} \rightarrow 0.$$

Meanwhile, under Condition H2-B, we use part (ii) of Lemma 15 to get, as $\epsilon \rightarrow 0$,

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \frac{1}{h(\epsilon)} \int_0^t f(Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \leq C h(\epsilon)^{-2} \epsilon^{-M_f} \rightarrow 0$$

and

$$\mathbb{E} \left[\sup_{t \in [0, 1]} \left| \frac{\epsilon}{h(\epsilon)} \int_0^t \nabla_x \phi(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) f(X_s^{\epsilon, w^\epsilon}, Y_s^{\epsilon, w^\epsilon}) dW_s^H \right|^2 \right] \leq C h(\epsilon)^{-2} \epsilon^{2-M_k} \rightarrow 0.$$

The remaining terms on the right-hand side of (49), except the term

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} \left| \int_0^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} (\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)) ds \right|^2 \right],$$

are bounded by using Lemmas 11 and 12 (which require Conditions H1 and H2-B). Thus, it follows from the estimates at (48) and (49) that

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} |\eta_t^{\epsilon, w^\epsilon}|^2 \right] \leq C_1 + C_2 \int_0^1 \mathbb{E} \left[\sup_{0 \leq r \leq s} |\eta_r^{\epsilon, w^\epsilon}|^2 \right] ds.$$

An application of Gronwall's inequality then yields the first claim of our statement, which is

$$\mathbb{E} \left[\sup_{0 \leq t \leq 1} |\eta_t^{\epsilon, w^\epsilon}|^2 \right] \leq C.$$

For the second claim, we proceed similarly to the derivation of the estimate at (48). Then for $\rho > 0$,

$$\begin{aligned} \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \left| \int_r^t \frac{1}{\sqrt{\epsilon} h(\epsilon)} (\bar{g}(X_s^{\epsilon, w^\epsilon}) - \bar{g}(\bar{X}_s)) ds \right| \right] &\leq C \mathbb{E} \left[\sup_{\substack{0 \leq r, t \leq 1 \\ |r-t| < \rho}} \int_r^t |\eta_s^{\epsilon, w^\epsilon}| ds \right] \\ &\leq C \rho \mathbb{E} \left[\sup_{0 \leq t \leq 1} |\eta_t^{\epsilon, w^\epsilon}|^2 \right] \leq C \rho. \end{aligned}$$

□

Lemma 17. *Let R_2^ϵ be the remainder term that appears in equation (27). Under Conditions H1 and either H2-A or H2-B, it holds that $R_2^\epsilon \rightarrow 0$ in $C([0, 1]; \mathbb{R}^n)$ in probability along a subsequence.*

Proof. For the purpose of identifying the limit of R_2^ϵ , we invoke the Skorokhod Representation Theorem and assume that $X^{\epsilon, w^\epsilon} \rightarrow \bar{X}$ a.s. in $C([0, 1]; \mathbb{R}^n)$ as $\epsilon \rightarrow 0$. As \bar{X} is bounded under Condition H2-A or H2-B, the Dominated Convergence Theorem implies that

$$(50) \quad \lim_{\epsilon \rightarrow 0} \mathbb{E} \left[\sup_{0 \leq s \leq 1} |X_s^{\epsilon, w^\epsilon} - \bar{X}_s|^2 \right] = 0.$$

Now, we employ the bound (28) and get

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq s \leq 1} |R_2^\epsilon(s)| \right] &\leq \mathbb{E} \left[\int_0^1 |\nabla_x^2 \bar{g}|_\infty |\eta_s^{\epsilon, w^\epsilon}| |X_s^{\epsilon, w^\epsilon} - \bar{X}_s| ds \right] \\ &\leq C \mathbb{E} \left[\sup_{0 \leq s \leq 1} |\eta_s^{\epsilon, w^\epsilon}| \sup_{0 \leq s \leq 1} |X_s^{\epsilon, w^\epsilon} - \bar{X}_s| \right] \\ &\leq C \mathbb{E} \left[\sup_{0 \leq s \leq 1} |\eta_s^{\epsilon, w^\epsilon}|^2 \right] \mathbb{E} \left[\sup_{0 \leq s \leq 1} |X_s^{\epsilon, w^\epsilon} - \bar{X}_s|^2 \right] \\ &\leq C \mathbb{E} \left[\sup_{0 \leq s \leq 1} |X_s^{\epsilon, w^\epsilon} - \bar{X}_s|^2 \right], \end{aligned}$$

In particular, the second inequality is due to the boundedness of $\nabla_x^2 \bar{g}$ implied by either Condition H2-A or H2-B. The last inequality is a consequence of Lemma 16 (which requires Conditions H1 and either H2-A or H2-B). (50) then gives us the desired limit. □

REFERENCES

- [AMP21] Giacomo Ascione, Yuliya Mishura, and Enrica Pirozzi, *Fractional ornstein-uhlenbeck process with stochastic forcing, and its applications*, Methodology and Computing in Applied Probability **23** (2021), no. 1, 53–84.
- [BD98] Michelle Boué and Paul Dupuis, *A variational representation for certain functionals of brownian motion*, The Annals of Probability **26** (1998), no. 4, 1641–1659.
- [BF77] D. Baer and M. I. Freidlin, *Theorems on large deviations, and stability under random perturbations*, Dokl. Akad. Nauk SSSR **235** (1977), no. 2, 253–256. MR 0451366
- [BFG⁺19] C. Bayer, P. K. Friz, A. Gulisashvili, B. Horvath, and B. Stemper, *Short-time near-the-money skew in rough fractional volatility models*, Quant. Finance **19** (2019), no. 5, 779–798. MR 3939657
- [BGS21] Solesne Bourguin, Siragan Gailus, and Konstantinos Spiliopoulos, *Typical dynamics and fluctuation analysis of slow-fast systems driven by fractional brownian motion*, Stochastics and Dynamics **21** (2021), no. 07, 2150030.
- [BHØZ08] Francesca Biagini, Yaozhong Hu, Bernt Øksendal, and Tusheng Zhang, *Stochastic calculus for fractional brownian motion and applications*, Springer Science & Business Media, 2008.

- [Bil13] Patrick Billingsley, *Convergence of probability measures*, John Wiley & Sons, 2013.
- [BS20] Amarjit Budhiraja and Xiaoming Song, *Large deviation principles for stochastic dynamical systems with a fractional brownian noise*, arXiv preprint arXiv:2006.07683 (2020).
- [Che03] Patrick Cheridito, *Arbitrage in fractional Brownian motion models*, Finance Stoch. **7** (2003), no. 4, 533–553. MR 2014249
- [CKM03] Patrick Cheridito, Hideyuki Kawaguchi, and Makoto Maejima, *Fractional ornstein-uhlenbeck processes*, Electronic Journal of probability **8** (2003), 1–14.
- [CR98a] Fabienne Comte and Eric Renault, *Long memory in continuous-time stochastic volatility models*, Math. Finance **8** (1998), no. 4, 291–323. MR 1645101
- [CR98b] ———, *Long memory in continuous-time stochastic volatility models*, Mathematical finance **8** (1998), no. 4, 291–323.
- [DE11] Paul Dupuis and Richard S Ellis, *A weak convergence approach to the theory of large deviations*, John Wiley & Sons, 2011.
- [DS12] Paul Dupuis and Konstantinos Spiliopoulos, *Large deviations for multiscale diffusion via weak convergence methods*, Stochastic Processes and their Applications **122** (2012), no. 4, 1947–1987.
- [dSEE18] José Luís da Silva, Mohamed Erraoui, and El Hassan Essaky, *Mixed stochastic differential equations: Existence and uniqueness result*, Journal of Theoretical Probability **31** (2018), no. 2, 1119–1141.
- [FPSS11] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Sølna, *Multiscale stochastic volatility for equity, interest rate, and credit derivatives*, Cambridge University Press, 2011.
- [Fre78] Mark Iosifovich Freidlin, *The averaging principle and theorems on large deviations*, Russian mathematical surveys **33** (1978), no. 5, 117.
- [FS99] Mark I Freidlin and Richard B Sowers, *A comparison of homogenization and large deviations, with applications to wavefront propagation*, Stochastic processes and their applications **82** (1999), no. 1, 23–52.
- [Fuk17] Masaaki Fukasawa, *Short-time at-the-money skew and rough fractional volatility*, Quant. Finance **17** (2017), no. 2, 189–198. MR 3592946
- [FV10] Peter K. Friz and Nicolas B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications. MR 2604669
- [FW12] Mark I. Freidlin and Alexander D. Wentzell, *Random perturbations of dynamical systems*, third ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer, Heidelberg, 2012, Translated from the 1979 Russian original by Joseph Szücs. MR 2953753
- [FZ17a] Martin Forde and Hongzhong Zhang, *Asymptotics for rough stochastic volatility models*, SIAM J. Financial Math. **8** (2017), no. 1, 114–145. MR 3608743
- [FZ17b] ———, *Asymptotics for rough stochastic volatility models*, SIAM Journal on Financial Mathematics **8** (2017), no. 1, 114–145.
- [GJR18] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum, *Volatility is rough*, Quantitative finance **18** (2018), no. 6, 933–949.
- [GJRS18] Hamza Guennoun, Antoine Jacquier, Patrick Roome, and Fangwei Shi, *Asymptotic behavior of the fractional Heston model*, SIAM J. Financial Math. **9** (2018), no. 3, 1017–1045. MR 3836176
- [GN08] João Guerra and David Nualart, *Stochastic differential equations driven by fractional brownian motion and standard brownian motion*, Stochastic analysis and applications **26** (2008), no. 5, 1053–1075.
- [Gui03] Arnaud Guillin, *Averaging principle of sde with small diffusion: moderate deviations*, The Annals of Probability **31** (2003), no. 1, 413–443.
- [HJL19] Blanka Horvath, Antoine Jacquier, and Chloé Lacombe, *Asymptotic behaviour of randomised fractional volatility models*, J. Appl. Probab. **56** (2019), no. 2, 496–523. MR 3986948
- [HL20] Martin Hairer and Xue-Mei Li, *Averaging dynamics driven by fractional brownian motion*, The Annals of Probability **48** (2020), no. 4, 1826–1860.
- [HSS19] Wenqing Hu, Michael Salins, and Konstantinos Spiliopoulos, *Large deviations and averaging for systems of slow-fast stochastic reaction-diffusion equations*, Stochastics and Partial Differential Equations: Analysis and Computations **7** (2019), no. 4, 808–874.
- [KL99] FC Klebaner and R Liptser, *Moderate deviations for randomly perturbed dynamical systems*, Stochastic processes and their applications **80** (1999), no. 2, 157–176.
- [KST06] Anatoly A. Kilbas, Hari M. Srivastava, and Juan J. Trujillo, *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies, vol. 204, Elsevier Science B.V., Amsterdam, 2006. MR 2218073
- [LS90] Robert Liptser and Jordan Stoyanov, *Stochastic version of the averaging principle for diffusion type processes*, Stochastics: An International Journal of Probability and Stochastic Processes **32** (1990), no. 3-4, 145–163.
- [LS20] Xue-Mei Li and Julian Sieber, *Slow-fast systems with fractional environment and dynamics*, arXiv preprint arXiv:2012.01910 (2020).
- [MS11] Yuliia S Mishura and Georgiy M Shevchenko, *Existence and uniqueness of the solution of stochastic differential equation involving wiener process and fractional brownian motion with hurst index $h > 1/2$* , Communications in Statistics-Theory and Methods **40** (2011), no. 19-20, 3492–3508.
- [MS17] Matthew R Morse and Konstantinos Spiliopoulos, *Moderate deviations for systems of slow-fast diffusions*, Asymptotic Analysis **105** (2017), no. 3-4, 97–135.
- [Nua06] David Nualart, *The malliavin calculus and related topics*, vol. 1995, Springer, 2006.

- [PIX20] Bin Pei, Yuzuru Inahama, and Yong Xu, *Averaging principles for mixed fast-slow systems driven by fractional brownian motion*, arXiv preprint arXiv:2001.06945 (2020).
- [PS07] Grigorios A Pavliotis and AM Stuart, *Parameter estimation for multiscale diffusions*, Journal of Statistical Physics **127** (2007), no. 4, 741–781.
- [PV01] E Pardoux and Yu Veretennikov, *On the poisson equation and diffusion approximation. i*, The Annals of Probability **29** (2001), no. 3, 1061–1085.
- [PV05] Etienne Pardoux and A Yu Veretennikov, *On the poisson equation and diffusion approximation 3*, The Annals of Probability **33** (2005), no. 3, 1111–1133.
- [Shi99] Albert N Shiryaev, *Essentials of stochastic finance: facts, models, theory*, vol. 3, World scientific, 1999.
- [SKM⁺93] Stefan G Samko, Anatoly A Kilbas, Oleg I Marichev, et al., *Fractional integrals and derivatives*, vol. 1, Gordon and breach science publishers, Yverdon Yverdon-les-Bains, Switzerland, 1993.
- [SM20] Konstantinos Spiliopoulos and Matthew R Morse, *Importance sampling for slow-fast diffusions based on moderate deviations*, Multiscale Modeling & Simulation **18** (2020), no. 1, 315–350.
- [Spi13] Konstantinos Spiliopoulos, *Large deviations and importance sampling for systems of slow-fast motion*, Applied Mathematics & Optimization **67** (2013), no. 1, 123–161.
- [Spi14] ———, *Fluctuation analysis and short time asymptotics for multiple scales diffusion processes*, Stochastics and Dynamics **14** (2014), no. 03, 1350026.
- [SV03a] Tommi Sottinen and Esko Valkeila, *On arbitrage and replication in the fractional black-scholes pricing model*, Statistics & Decisions **21** (2003), no. 2, 93–108.
- [SV03b] ———, *On arbitrage and replication in the fractional Black-Scholes pricing model*, Statist. Decisions **21** (2003), no. 2, 93–107. MR 2000665
- [You36] L. C. Young, *An inequality of the Hölder type, connected with Stieltjes integration*, Acta Math. **67** (1936), no. 1, 251–282. MR 1555421
- [Zäh98] Martine Zähle, *Integration with respect to fractal functions and stochastic calculus. i*, Probability theory and related fields **111** (1998), no. 3, 333–374.
- [Zäh99] Martina Zähle, *On the link between fractional and stochastic calculus*, Stochastic dynamics (1999), 305–325.
- [Zha09] Xicheng Zhang, *A variational representation for random functionals on abstract wiener spaces*, Journal of Mathematics of Kyoto University **49** (2009), no. 3, 475–490.