

Parametric inference for small variance and long time horizon McKean-Vlasov diffusion models

Valentine Genon-Catalot¹ and Catherine Larédo^{*2,3}

¹ *Université de Paris, CNRS, MAP5, UMR 8145, F-75006, Paris, FRANCE*

² *MaIAGE, INRAE, Université Paris-Saclay, 78350, Jouy-en-Josas, FRANCE*

³ *LPSM, UMR 80001, 75013, Paris, FRANCE*

e-mail: valentine.genon-catalot@mi.parisdescartes.fr; catherine.laredo@inrae.fr

Abstract: Let (X_t) be solution of a one-dimensional McKean-Vlasov stochastic differential equation with classical drift term $V(\alpha, x)$, self-stabilizing term $\Phi(\beta, x)$ and small noise amplitude ε . Our aim is to study the estimation of the unknown parameters α, β from a continuous observation of $(X_t, t \in [0, T])$ under the double asymptotic framework ε tends to 0 and T tends to infinity. After centering and normalization of the process, uniform bounds for moments with respect to $t \geq 0$ and ε are derived. We then build an explicit approximate log-likelihood leading to consistent and asymptotically Gaussian estimators, under the condition that $\varepsilon\sqrt{T}$ tends to 0, with original rates of convergence: the rate for the estimation of α is either ε^{-1} or $\sqrt{T}\varepsilon^{-1}$, the rate for the estimation of β is \sqrt{T} . Moreover, the estimators are asymptotically efficient.

MSC2020 subject classifications: Primary 60J60, 60J99, 62F12, 62M05.

Keywords and phrases: McKean-Vlasov stochastic differential equations, continuous observations, small noise, infinite time horizon, parametric inference, asymptotic properties of estimators.

Received January 2021.

Contents

1	Introduction	5812
2	Probabilistic properties	5814
2.1	Assumptions and recap of previous results	5814
2.2	Statement of probabilistic results	5816
3	Estimation when both ε tends to 0 and T tends to infinity	5817
3.1	Notations and assumptions	5817
3.2	Approximate likelihood	5818
3.3	Preliminary results	5819
3.4	Rates of convergence	5820
3.5	Asymptotic properties of estimators	5822
3.6	Asymptotic efficiency.	5824

*Corresponding author.

4	Concluding remarks	5825
5	Appendix	5826
5.1	Proofs of Section 2	5826
5.2	Proofs of Section 3	5831
	Acknowledgments	5851
	References	5851

1. Introduction

We consider here parametric inference for a dynamical model subject to three sources of forcing: the geometry of the state space is described by a potential term $V(\alpha, x)$, a Brownian motion with small noise allows to include small random perturbations and a self-stabilization term $\Phi(\beta, x)$. Such processes appear when describing the limit behaviour of a large population of interacting particles with an interaction function between the dynamical systems. More precisely, we study the one-dimensional process

$$dX_t = V(\alpha, X_t)dt - b(\theta, t, \varepsilon, X_t)dt + \varepsilon dW_t, \quad X_0 = x_0, \quad (1.1)$$

where (W_t) is a Wiener process, x_0 is deterministic and known,

$$b(\theta, t, \varepsilon, x) = \int_{\mathbb{R}} \Phi(\beta, x - y)u_t^{\theta, \varepsilon}(dy) = \mathbb{E}[\Phi(\beta, x - X_t^{\theta, \varepsilon, x_0})], \quad (1.2)$$

$u_t^{\theta, \varepsilon}(dy) := u_t^{\theta, \varepsilon, x_0}(dy)$ is the distribution of $X_t := X_t^{\theta, \varepsilon, x_0}$, $V : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $\Phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ are deterministic Borel known functions and $\theta = (\alpha, \beta) \in \Theta = \Theta_\alpha \times \Theta_\beta \subset \mathbb{R}^2$ is an unknown parameter. A solution of (1.1) is the couple $(X_t, u_t^{\theta, \varepsilon}(dy))_{t \geq 0}$ composed of the stochastic process (X_t) and the family of distributions $(u_t^{\theta, \varepsilon})$. The function $x \rightarrow b(\theta, t, \varepsilon, x)$ (see (1.2)) depends on θ, t, ε , the starting point x_0 and $u_t^{\theta, \varepsilon}$. When defined, the process (1.1) is a *time-inhomogeneous* Markov process known as self-stabilizing diffusion, nonlinear stochastic differential equation, or McKean-Vlasov stochastic differential equation.

These models were first described by [39] and arised in Statistical Physics for the modeling of granular media by interacting particle systems (see *e.g.* [7]). Due to their growing importance, many fundamental probabilistic tools for their study were developed later (see *e.g.* [18], [48] for a survey, [5, 6, 40, 43] and many others). [26] is concerned with small noise properties and large deviations results for these processes.

Except [31], the statistical inference for interacting particle systems remained unstudied for many years. Since 2010, the fields of application of these models progressively encompassed Statistical Physics and these equations were shown to describe collective and observable dynamics in other application fields such as Mathematical Biology (see *e.g.* [4, 44]), Epidemics Dynamics (see [3, 8, 17]), Finance (see [22] and references therein). More recently, several authors were

concerned by statistical studies based on the direct observation of large interacting particle systems (see [16, 22]).

The convergence as N tends to infinity of systems of N interacting particles has been investigated (*propagation of chaos*). One of the most important limiting processes is the class of McKean-Vlasov diffusion processes (see *e.g.* [43, 48]). Therefore, it is a worthwhile stochastic model to study from the statistical point of view.

Inference for stochastic differential equations (SDEs) ($\Phi(\beta, x) \equiv 0$) based on the observation of sample paths on a time interval $[0, T]$ has been widely investigated. Authors consider continuous or discrete observations, parametric or nonparametric inference under various asymptotic frameworks: small diffusion asymptotics on a fixed time interval; long time interval, especially for ergodic models; observation of n *i.i.d.* paths with large n . Among many studies, we refer first to several textbooks: [29, 30, 32, 33, 34]. Second, among the many papers on the topic, we can quote: [10, 11, 12, 13, 19, 20, 23, 24, 25, 28, 35, 47, 50, 51]. Moreover, these works have opened the field of inference for more complex stochastic differential equations: diffusions with jumps (see *e.g.* [2, 41, 46]), SDEs driven by Lévy processes (see *e.g.* [42]), diffusions with mixed effects (see *e.g.* [14, 15, 45]), stochastic partial differential equations (see *e.g.* [1, 9]).

For what concerns McKean-Vlasov models, in a previous paper, [21] has investigated the estimation of (α, β) based on the continuous observation on a fixed time interval $[0, T]$ of model (1.1) under the asymptotic framework $\varepsilon \rightarrow 0$. The assumptions on the model are those of [26]. In particular, it is assumed that $\Phi(\beta, \cdot)$ is increasing and odd. Then, it appears that, as $\varepsilon \rightarrow 0$ with fixed T , only α can be consistently estimated but not β . More information is needed for β . Assuming that n *i.i.d.* sample paths of model (1.1) are observed on the fixed interval $[0, T]$, that $\varepsilon \rightarrow 0$ and n tends to infinity, both parameters are estimated but they have different rates, $\sqrt{n\varepsilon^{-1}}$ for α , \sqrt{n} for β .

In the present paper, we consider one sample path (X_t) which is continuously observed on $[0, T]$ and study the parametric inference of (α, β) for the McKean-Vlasov model (1.1) in the new double asymptotic framework $\varepsilon \rightarrow 0, T \rightarrow \infty$. Up to our knowledge, this framework has not been studied before. As a side result, the inference for classical SDEs in this asymptotic framework is obtained.

In Section 2, we study probabilistic properties of the process (X_t) . Let $x_t(\alpha)$ be the solution of the ordinary differential equation

$$dx_t(\alpha) = V(\alpha, x_t(\alpha))dt, \quad x_0(\alpha) = x_0. \quad (1.3)$$

We prove that all the moments of $\varepsilon^{-1}(X_t - x_t(\alpha))$ are uniformly bounded in $t \geq 0, \varepsilon \leq 1$ (Theorem 2.2), that the Gaussian approximating process of (X_t) as $\varepsilon \rightarrow 0$ obtained in [21] holds on \mathbb{R}^+ and that the remainder terms of this approximation have moments uniformly bounded in $t \geq 0, \varepsilon \leq 1$ (Theorem 2.3). Moreover, Corollary 2.1 deals with the rate of the difference $b(\theta, t, X_t) - \Phi(\beta, X_t - x_t(\alpha))$ as ε tends to 0. To prove these results, important assumptions are that $\Phi(\beta, \cdot)$ is increasing and odd, that $V(\alpha, \cdot)$ is decreasing and that there

is a unique $x^*(\alpha)$ such that $V(\alpha, x^*(\alpha)) = 0$ which is an attractive point for (1.3). In Section 3, we define an approximate log-likelihood and study its properties together with the associated estimators as $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$. Two cases have to be distinguished. Either Case (1), $x^*(\alpha)$ depends on α , or Case (2), $x^*(\alpha)$ does not depend on α . In the two cases, we obtain that the estimators of (α, β) are consistent and have different rates of convergence. In Case (1), the estimator of α is asymptotically Gaussian with the fast rate $\sqrt{T}\varepsilon^{-1}$ while in Case (2), its rate is ε^{-1} . In both cases, the parameter β is estimated at rate \sqrt{T} (Theorems 3.1-3.2-3.3). Finally, we prove the asymptotic efficiency of the estimators by means of an asymptotic equivalence of experiments property. Section 4 gives some concluding remarks. Proofs are gathered in Section 5. Throughout the paper, we assume that $\varepsilon \leq 1$.

2. Probabilistic properties

2.1. Assumptions and recap of previous results

We consider the following assumptions:

- [H0] For all α, β , the functions $x \rightarrow V(\alpha, x)$ and $x \rightarrow \Phi(\beta, x)$ are locally Lipschitz.
- [H1] Either, $\Phi(\beta, \cdot) \equiv 0$ for all β , or for all β the function $x \rightarrow \Phi(\beta, x)$ is odd, increasing and grows at most polynomially: there exist $K(\beta) > 0$ and $r(\beta) \in \mathbb{N}$ such that $|\Phi(\beta, x) - \Phi(\beta, y)| \leq |x - y|(K(\beta) + |x|^{r(\beta)} + |y|^{r(\beta)})$, $x, y \in \mathbb{R}$.
- [H2-k] The functions $x \rightarrow V(\alpha, x)$ and $x \rightarrow \Phi(\beta, x)$ have continuous partial derivatives up to order k and these derivatives have polynomial growth: for all α, β , and all $i, i \leq k$, there exist constants $k(\alpha) > 0, k(\beta) > 0$ and integers $\gamma(\alpha) \geq 0, \gamma(\beta) \geq 0$, such that

$$\forall x \in \mathbb{R}, \quad \left| \frac{\partial^i V}{\partial x^i}(\alpha, x) \right| \leq k(\alpha)(1 + |x|^{\gamma(\alpha)}), \quad \left| \frac{\partial^i \Phi}{\partial x^i}(\beta, x) \right| \leq k(\beta)(1 + |x|^{\gamma(\beta)}).$$

- [H3] For all α , the function $x \rightarrow V(\alpha, x)$ is continuously differentiable and there exists $K_V(\alpha) > 0$ such that for all $x \in \mathbb{R}$, $\frac{\partial V}{\partial x}(\alpha, x) \leq -K_V(\alpha)$.
- [H4] There exists $x^*(\alpha)$ such that $V(\alpha, x^*(\alpha)) = 0$.

Note that the case $\Phi(\beta, x) \equiv 0$ corresponds to a classical stochastic differential equation.

Let us recall some results of [26] where Equation (1.1) is studied in the more general case of X_0 a random variable, independent of (W_t) with distribution μ . Under [H0]-[H1] and [H3], if $\mathbb{E}X_0^{8q^2} < +\infty$ where $q = [(r(\beta)/2) + 1]$, then, for all θ , there exists a drift term $b(\theta, t, \varepsilon, x) = b^\mu(\theta, t, \varepsilon, x)$ such that (1.1) admits a unique strong solution $(X_t = X_t^{\theta, \varepsilon, \mu})$ satisfying $b(\theta, t, \varepsilon, x) = \int_{\mathbb{R}} \Phi(\beta, x - y) u_t^{\theta, \varepsilon, \mu}(dy)$ and X is the unique strong solution of (1.1). Moreover, for all $n \in \{1, \dots, 4q^2\}$, whenever $\mathbb{E}X_0^{2n} < +\infty$, $\sup_{t \geq 0} \mathbb{E}X_t^{2n} < +\infty$. We assume here $X_0 = x_0$ deterministic, which yields that $\forall n \in \mathbb{N}$, $\sup_{t \geq 0} \mathbb{E}X_t^{2n} < +\infty$. Under [H3], $x^*(\alpha)$ in

[H4] is the unique value such that $V(\alpha, x^*(\alpha)) = 0$. Under [H0], [H3]-[H4], $x_t(\alpha)$ converges as $t \rightarrow +\infty$ to the fixed point $x^*(\alpha)$ with exponential rate (see also below Proposition 3.1). Assumptions [H1] and [H3]-[H4] are especially used in the proofs of Theorems 2.2, 2.3 and Corollary 2.1 to obtain uniform bounds.

In a previous paper, we have studied the process (X_t) on a fixed time interval $[0, T]$. Let us recall the results that we need in the sequel. First, properties of continuity and differentiability of $b(\theta, t, \varepsilon, x)$ defined in (1.2) with respect to ε and x at $(\theta, t, 0, x)$ can be derived from the assumptions.

Lemma 2.1 ([21]). *Assume [H0]-[H1], [H2-2], [H3] and that $X_0 = x_0$ is deterministic. Then,*

- (i) *For all $\theta, t \geq 0$, $(\varepsilon, x) \rightarrow b(\theta, t, \varepsilon, x)$ is continuously differentiable on $\mathbb{R}^+ \times \mathbb{R}$.*
- (ii) *$\lim_{\varepsilon \rightarrow 0} b(\theta, t, \varepsilon, x) = \Phi(\beta, x - x_t(\alpha))$.*
- (iii) *At $\varepsilon = 0$, $\frac{\partial b}{\partial \varepsilon}(\theta, t, 0, x) = 0$ and $\frac{\partial b}{\partial x}(\theta, t, 0, x) = \frac{\partial \Phi}{\partial x}(\beta, x - x_t(\alpha))$.*

Property (ii) is also proved in [26]. This property is strengthened in Corollary 2.1 yielding uniform closeness in t .

Next, the asymptotic properties of (X_t) on a fixed time interval $[0, T]$ as $\varepsilon \rightarrow 0$ have been studied. As ε tends to 0, (X_t) converges uniformly in probability on $[0, T]$ to $x_t(\alpha)$ defined in (1.3). Due to [H1], $\Phi(\beta, 0) = 0$, so that $x_t(\alpha)$ corresponds to the solution of (1.1) with $\varepsilon = 0$. Moreover, setting

$$a(\theta, t) = \frac{\partial V}{\partial x}(\alpha, x_t(\alpha)) - \frac{\partial \Phi}{\partial x}(\beta, 0), \tag{2.1}$$

define $(g_t(\theta))$ the Ornstein-Uhlenbeck process

$$dg_t(\theta) = a(\theta, t)g_t(\theta)dt + dW_t, \quad g_0(\theta) = 0. \tag{2.2}$$

Note that $\frac{\partial \Phi}{\partial x}(\beta, 0) \geq 0$ so that, under [H3], $a(\theta, t) \leq -(K_V(\alpha) + \frac{\partial \Phi}{\partial x}(\beta, 0)) < 0$. Then, the following expansion of (X_t) with respect to ε holds.

Theorem 2.1 ([21]). *Assume [H0], [H1] and [H2-3], then*

$$X_t = x_t(\alpha) + \varepsilon g_t(\theta) + \varepsilon^2 R_t^\varepsilon(\theta), \tag{2.3}$$

where the remainder term $R_t^\varepsilon(\theta)$ has moments uniformly bounded on $[0, T]$.

Equation (2.2) can be solved

$$g_t(\theta) = \int_0^t \exp\left(\int_s^t a(\theta, u)du\right)dW_s = \int_0^t e^{A(\theta, t) - A(\theta, s)}dW_s, \text{ where} \tag{2.4}$$

$$A(\theta, t) = \int_0^t a(\theta, u)du. \tag{2.5}$$

To illustrate the results, consider the following explicit example.

Example 1. Let $V(\alpha, x) = -\alpha x$, $\Phi(\beta, x) = \beta x$ with $\alpha > 0, \beta \geq 0$. We have $b(\theta, t, \varepsilon, x) = \beta(x - \mathbb{E}_\theta(X_t))$, and Equation (1.1) writes:

$$dX_t = -\alpha X_t dt - \beta(X_t - \mathbb{E}_\theta(X_t))dt + \varepsilon dW_t, \quad X_0 = x_0.$$

We easily check that $\mathbb{E}_\theta(X_t) = x_0 e^{-\alpha t}$ and (1.1) can be solved explicitly:

$$X_t = x_0 e^{-\alpha t} + \varepsilon e^{-(\alpha+\beta)t} \int_0^t e^{(\alpha+\beta)s} dW_s. \quad (2.6)$$

The remainder term $R_t^\varepsilon(\theta)$ is here equal to 0.

2.2. Statement of probabilistic results

Under the assumptions of Section 2.1, we can extend the results of [21] and prove uniform bounds on \mathbb{R}^+ , thanks to [H1] and [H3]-[H4].

Theorem 2.2. *Let (X_t) denote the solution of (1.1) and $x_t(\alpha)$ of (1.3).*

(i) *Assume [H0]-[H1], [H3]. Then, for all $n \geq 1$, there exists a constant $\delta(\alpha, n)$ such that*

$$\forall \varepsilon \in (0, 1], \quad \forall t \geq 0, \quad \mathbb{E}_\theta \left(\frac{X_t - x_t(\alpha)}{\varepsilon} \right)^{2n} \leq \delta(\alpha, n).$$

(ii) *If moreover [H2-2] and [H4] hold, there exists a constant $\delta(\alpha) > 0$ such that,*

$$\forall \varepsilon \in (0, 1] \quad \forall t \geq 0, \quad \varepsilon^{-2} |\mathbb{E}_\theta(X_t - x_t(\alpha))| \leq \delta(\alpha).$$

In the special case where $V(\alpha, x) = -\alpha x$, $\mathbb{E}_\theta(X_t) = x_t(\alpha)$.

The following comment is important for the sequel.

Comment 2.1. *The bounds $\delta(\alpha, n)$ and $\delta(\alpha)$ depend on θ only through α . From the proofs, we have that $\delta(\alpha, n) = (nK_V^{-1}(\alpha))^n$ and that $\delta(\alpha)$ is a function of $B(\alpha), k(\alpha), \gamma(\alpha), K_V^{-1}(\alpha)$, where $B(\alpha) = \sup_{t \geq 0} |x_t(\alpha)| < +\infty$. These bounds increase in each of their variables (see (5.6) and (5.9)). Thus, if these quantities are upper bounded by constants independent of α , the bounds of Theorem 2.2 are uniform in α . This property is used in Section 3 in the proofs of consistency for the estimators and of the asymptotic sufficiency property.*

Note that, under [H0], [H3]-[H4], we easily check that $(x_t(\alpha) - x^*(\alpha))^2 \leq (x_0 - x^*(\alpha))^2 \exp(-2K_V(\alpha)t)$. Therefore $x_t(\alpha)$ converges as $t \rightarrow +\infty$ to $x^*(\alpha)$ with exponential rate. It follows immediately from Theorem 2.2 that, under [H0]-[H1], [H3]-[H4], $X_t \rightarrow x^*(\alpha)$ in probability as $t \rightarrow +\infty$ and $\varepsilon \rightarrow 0$. The Dirac measure $\delta_{x^*(\alpha)}$ appears as the limit of the distribution of (X_t) as $t \rightarrow +\infty$ and $\varepsilon \rightarrow 0$.

We also have that the remainder term $R_t^\varepsilon(\theta)$ defined in (2.3) has moments uniformly bounded on \mathbb{R}^+ .

Theorem 2.3. *Under [H0]-[H1], [H2-3], [H3]-[H4], the expansion $X_t = x_t(\alpha) + \varepsilon g_t(\theta) + \varepsilon^2 R_t^\varepsilon(\theta)$ holds on \mathbb{R}^+ and $R_t^\varepsilon(\theta)$ satisfies*

$$\sup_{t \geq 0, \varepsilon \in (0, 1]} \mathbb{E}_\theta |R_t^\varepsilon(\theta)| = O(1) \text{ and } \forall p \geq 1, \quad \sup_{t \geq 0, \varepsilon \in (0, 1]} \mathbb{E}_\theta (R_t^\varepsilon(\theta) - \mathbb{E}_\theta R_t^\varepsilon(\theta))^{2p} = O(1).$$

Note that $\mathbb{E}_\theta g_t^2(\theta)$ is uniformly bounded on \mathbb{R}^+ . Indeed, using the explicit expression of $g_t(\theta)$ given in (2.4) and the property that, under [H3], for $s \leq t$, $A(\theta, t) - A(\theta, s) \leq -K_V(\alpha)(t - s)$, we get that

$$\mathbb{E}_\theta g_t^2(\theta) = \int_0^t \exp [2(A(\theta, t) - A(\theta, s))] ds \leq (2K_V(\alpha))^{-1}.$$

Define the difference (see Lemma 2.1, (ii))

$$D(\theta, t, \varepsilon, x) = b(\theta, t, \varepsilon, x) - \Phi(\beta, x - x_t(\alpha)). \tag{2.7}$$

The following corollary dealing with $D(\theta, t, \varepsilon, X_t)$ is a crucial tool for the statistical study. As for $R_t^\varepsilon(\theta)$, uniform bounds hold for $D(\theta, t, \varepsilon, X_t)$.

Corollary 2.1. *Assume [H0]-[H1], [H2-3], [H3]-[H4]. Then $D(\theta, t, \varepsilon, X_t)$ defined in (2.7) satisfies,*

$$\sup_{t \in \mathbb{R}^+, \varepsilon \in (0,1]} \varepsilon^{-2} |\mathbb{E}_\theta D(\theta, t, \varepsilon, X_t)| = O(1), \tag{2.8}$$

$$\forall p \geq 1, \sup_{t \in \mathbb{R}^+, \varepsilon \in (0,1]} \varepsilon^{-6p} \mathbb{E}_\theta (D(\theta, t, \varepsilon, X_t) - \mathbb{E}_\theta D(\theta, t, \varepsilon, X_t))^{2p} = O(1). \tag{2.9}$$

Comment 2.2. *The constants $O(1)$ in Theorem 2.3 and Corollary 2.1 are independent of θ if the constants $k(\alpha), k(\beta), \gamma(\alpha), \gamma(\beta), K_V^{-1}(\alpha), B(\alpha)$ are upper bounded independently of θ (see Comment 2.1).*

3. Estimation when both ε tends to 0 and T tends to infinity

We have previously obtained that α can be estimated, while β cannot be estimated on a fixed time interval $[0, T]$ (see [21]). Therefore, to estimate both parameters, we have to combine two asymptotic frameworks. In [21], we considered n *i.i.d.* paths of process (1.1) with $\varepsilon \rightarrow 0$ and $n \rightarrow +\infty$. We investigate here the inference based on the continuous observation on $[0, T]$ of one sample path ruled by (1.1) under the double asymptotic framework of $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$.

3.1. Notations and assumptions

As it is usual in statistics, we consider the canonical space associated with the observation of $(X_t, t \in [0, T])$, $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \in [0, T]), \mathbb{P}_\theta)$, where $\Omega = C([0, T])$ is the space of continuous real-valued functions defined on $[0, T]$ endowed with the Borel σ -field associated with the uniform convergence on $[0, T]$, $(X_t, t \in [0, T])$ is the canonical process $(X_t(\omega) = \omega(t))$, $(\mathcal{F}_t, t \in [0, T])$ is the canonical filtration and $\mathbb{P}_\theta^{\varepsilon, T} = \mathbb{P}_\theta$ the distribution of (1.1) on $C([0, T])$. We denote \mathbb{E}_θ the expectation with respect to \mathbb{P}_θ .

In this section, we study the estimation of (α, β) from a continuous observation $(X_t, t \in [0, T])$ and, in addition to [H0]-[H1], [H3]-[H4], we assume

- [S0] The parameter set is $\Theta = \Theta_\alpha \times \Theta_\beta$ where $\Theta_\alpha, \Theta_\beta$ are bounded closed intervals. The true value of the parameter $\theta_0 = (\alpha_0, \beta_0)$ belongs to $\mathring{\Theta}$.

- [S1] The function $(\alpha, x) \rightarrow V(\alpha, x)$ (resp. $(\beta, x) \rightarrow \Phi(\beta, x)$) is defined and continuous on $U_\alpha \times \mathbb{R}$ (resp. $U_\beta \times \mathbb{R}$), and all the derivatives

$$(\alpha, x) \rightarrow \frac{\partial^{i+j} V}{\partial x^i \partial \alpha^j}(\alpha, x), \quad (\beta, x) \rightarrow \frac{\partial^{i+j} \Phi}{\partial x^i \partial \beta^j}(\beta, x)$$

exist, are continuous on $U_\alpha \times \mathbb{R}$ (resp. $U_\beta \times \mathbb{R}$), where U_α, U_β are open intervals containing respectively $\Theta_\alpha, \Theta_\beta$, with polynomial growth with respect to x : there exist $K > 0$ and $k \in \mathbb{N}$ such that for all $i, j \geq 0$,

$$\forall (\alpha, \beta) \in \Theta, \forall x \in \mathbb{R}, \left| \frac{\partial^{i+j} V}{\partial x^i \partial \alpha^j}(\alpha, x) \right| + \left| \frac{\partial^{i+j} \Phi}{\partial x^i \partial \beta^j}(\beta, x) \right| \leq K(1 + |x|^k).$$

- [S2] There exists K_V such that: $\forall \alpha \in \Theta_\alpha, K_V(\alpha) \geq K_V > 0$ (see [H3]).
- [S3] $\{x^*(\alpha) = x^*(\alpha_0)\} \Rightarrow \{\alpha = \alpha_0\}$.
- [S4] $x_0 \neq x^*$ and $\{s \rightarrow V(\alpha, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)) \equiv 0$ and $s \rightarrow x_s(\alpha) - x_s(\alpha_0) \equiv 0\} \Rightarrow \{\alpha = \alpha_0\}$.
- [S5] For all $\beta \in \Theta_\beta, \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \neq 0$.
- [S6] $\{\frac{\partial \Phi}{\partial x}(\beta, 0) = \frac{\partial \Phi}{\partial x}(\beta_0, 0)\} \Rightarrow \{\beta = \beta_0\}$.

Assumption [S0] is standard in parametric inference and used only for consistency. Assuming the existence of derivatives of any order is not necessary but it simplifies the exposure. The uniformity of the constants K, k in [S1] is only required for the consistency part. As $\Theta_\alpha, \Theta_\beta$ are supposed to be compact, this is not a strong assumption. Assumption [S2] is crucial for the statistical results since the uniform bounds on \mathbb{R}^+ of the moments explicitly depend on $K_V^{-1}(\alpha)$ (see Comment 2.1 after Theorem 2.2). Assumption [S3] (resp. [S4]) is an identifiability assumption for α associated with the case where the fixed point $x^*(\alpha)$ depends (resp. does not depend) on α . Assumption [S5] ensures that \sqrt{T} is the convergence rate for β . Finally [S6] is an identifiability assumption for β .

Let us state some consequences of these statistical assumptions on the results of Section 2. By the relation $V(\alpha, x^*(\alpha)) = 0$, the function $\alpha \rightarrow x^*(\alpha)$ is continuous so, as Θ_α is compact, $\sup_{\alpha \in \Theta_\alpha} |x^*(\alpha)| = A < +\infty$. Therefore, under [S2], $\sup_{\alpha \in \Theta_\alpha} \sup_{t \geq 0} |x_t(\alpha)| = B < +\infty$. In view of Comments 2.1 and 2.2, under [S1]-[S2] all the bounds of Theorems 2.2, 2.3 and Corollary 2.1 are not only uniform in t, ε but also in θ .

3.2. Approximate likelihood

First note that $b(\theta, s, \varepsilon, x)$ is a non random term given in (1.2) so that the Girsanov theorem holds and the loglikelihood associated with the observation of $(X_t, t \in [0, T])$ is $\ell_{\varepsilon, T}(\theta)$ with

$$\varepsilon^2 \ell_{\varepsilon, T}(\theta) = \int_0^T (V(\alpha, X_s) - b(\theta, s, \varepsilon, X_s)) dX_s - \frac{1}{2} \int_0^T (V(\alpha, X_s) - b(\theta, s, \varepsilon, X_s))^2 ds.$$

It contains the term $b(\theta, s, \varepsilon, X_s)$ which is involved for the estimation of θ . However, for small ε , $b(\theta, t, \varepsilon, x)$ is close to $\Phi(\beta, x - x_t(\alpha))$ (see Lemma 2.1, (ii)).

Therefore, as in [21], we replace $b(\theta, s, \varepsilon, x)$ by $\Phi(\beta, x - x_t(\alpha))$ and consider the process (ξ_t)

$$d\xi_t = [V(\alpha, \xi_t) - \Phi(\beta, \xi_t - x_t(\alpha))]dt + \varepsilon dW_t, \quad \xi_0 = x_0. \tag{3.1}$$

From [26], this process is close to (X_t) for small ε and was used for getting large deviations results. Here, we derive an approximate log-likelihood for (X_t) by plugging the observation (X_t) in the exact log-likelihood of (ξ_t) :

$$\Lambda_{\varepsilon,T}(\theta) = \frac{1}{\varepsilon^2} \int_0^T H(\theta, s, X_s) dX_s - \frac{1}{2\varepsilon^2} \int_0^T H^2(\theta, s, X_s) ds, \quad \text{with} \tag{3.2}$$

$$H(\theta, s, x) = V(\alpha, x) - \Phi(\beta, x - x_s(\alpha)). \tag{3.3}$$

Clearly, this approximate log-likelihood is easier to study.

Let us consider the estimators $(\hat{\alpha}_{\varepsilon,T}, \hat{\beta}_{\varepsilon,T})$ associated with $\Lambda_{\varepsilon,T}(\theta)$ defined as any solution of

$$(\hat{\alpha}_{\varepsilon,T}, \hat{\beta}_{\varepsilon,T}) = \arg \max_{(\alpha, \beta) \in \Theta_\alpha \times \Theta_\beta} \Lambda_{\varepsilon,T}(\alpha, \beta). \tag{3.4}$$

The general set-up for such statistical problems is to study first the gradient vector and the Hessian matrix of $\theta \rightarrow \Lambda_{\varepsilon,T}(\theta)$, *i.e.* its first and second partial derivatives with respect to α and β . This yields the right normalizations to get the associated convergence results. We obtain different rates for the two partial derivatives $\frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta)$, $\frac{\partial \Lambda_{\varepsilon,T}}{\partial \beta}(\theta)$ and a diagonal rate matrix for the Hessian (Theorems 3.1-3.2) so that the inference for (α, β) is a two-rate statistical problem. Such a situation often arises for various kinds of observations in statistics of stochastic processes (see [23, 32, 47]). This yields a specific difficulty for proving consistency and asymptotic normality of estimators (3.4). The sketch of proof relies on three steps. Prove that

1. $\hat{\alpha}_{\varepsilon,T}$ is consistent.
2. the sequence $(\hat{\alpha}_{\varepsilon,T} - \alpha_0)$, when suitably normalized, is \mathbb{P}_{θ_0} -tight.
3. $\hat{\beta}_{\varepsilon,T}$ is consistent.

Once the consistency is proved, asymptotic normality can be obtained from Theorems 3.1-3.2.

3.3. Preliminary results

Let us set

$$\ell(\alpha) = -\frac{\partial V}{\partial x}(\alpha, x^*(\alpha)); \quad \ell(\alpha, \beta) = \ell(\alpha) + \frac{\partial \Phi}{\partial x}(\beta, 0). \tag{3.5}$$

Note that $\ell(\alpha) \geq K_V(\alpha) \geq K_V > 0$ and $\ell(\alpha, \beta) \geq \ell(\alpha)$.

Proposition 3.1. *Assume [H4], [S1], [S2]. Then $(x_t(\alpha), \frac{\partial x_t}{\partial \alpha}(\alpha, t), \frac{\partial^2 x_t}{\partial \alpha^2}(\alpha, t))$ converges to $(x^*(\alpha), \frac{dx^*}{d\alpha}(\alpha), \frac{d^2 x^*}{d\alpha^2}(\alpha))$ exponentially fast with rate $\exp(-\ell(\alpha)t)$ as t tends to infinity.*

We also need to specify the asymptotic behaviour of functionals of the time inhomogeneous process $(g_t(\theta))$ defined by (2.2) or (2.4).

Proposition 3.2. *Assume [H1], [H4], [S1], [S2]. Then, as $T \rightarrow \infty$,*

- (i) $\frac{1}{T} \int_0^T [g_t(\theta)]^2 dt \rightarrow_{\mathbb{L}^2(\mathbb{P}_\theta)} [2\ell(\alpha, \beta)]^{-1}$,
- (ii) $\frac{1}{T} \int_0^T g_t(\theta) dt \rightarrow_{\mathbb{P}_\theta} 0$,
- (iii) *If the function $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is continuous and satisfies $\lim_{t \rightarrow +\infty} h(t) = 0$, then $\frac{1}{\sqrt{T}} \int_0^T g_t(\theta) h(t) dt \rightarrow_{\mathbb{P}_\theta} 0$.*

3.4. Rates of convergence

We may now consider the joint estimation of (α, β) and start by studying the convergence of the gradient vector and the Hessian matrix of $\theta \rightarrow \Lambda_{\varepsilon, T}(\theta)$.

The estimation of α varies according to the property that $x^*(\alpha)$ depends on α or not. As, for all α , $V(\alpha, x^*(\alpha)) \equiv 0$, under [S1],

$$\frac{d}{d\alpha}(V(\alpha, x^*(\alpha))) = \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) + \frac{\partial V}{\partial x}(\alpha, x^*(\alpha)) \frac{dx^*}{d\alpha}(\alpha) \equiv 0. \quad (3.6)$$

By [S2], $\frac{\partial V}{\partial x}(\alpha, x^*(\alpha)) = -\ell(\alpha) \neq 0$. Consequently, we distinguish the two cases:

- (1) $\frac{dx^*}{d\alpha}(\alpha) \neq 0 \Leftrightarrow \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) \neq 0$.
- (2) $\frac{dx^*}{d\alpha}(\alpha) \equiv 0 \Leftrightarrow \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) \equiv 0$: $x^*(\alpha) = x^*$ does not depend on α .

Let us remark that Example 1 presented in Section 2.1 belongs to Case (2). According to these two cases, we set

$$D_{\varepsilon, T}^{(1)} = \begin{pmatrix} \frac{\varepsilon}{\sqrt{T}} & 0 \\ 0 & \frac{1}{\sqrt{T}} \end{pmatrix}, D_{\varepsilon, T}^{(2)} = \begin{pmatrix} \varepsilon & 0 \\ 0 & \frac{1}{\sqrt{T}} \end{pmatrix}, \mathcal{J}_{\varepsilon, T}(\theta) = - \begin{pmatrix} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha^2}(\theta) & \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta \partial \alpha}(\theta) \\ \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta \partial \alpha}(\theta) & \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta) \end{pmatrix}.$$

Theorem 3.1. Case (1) ($\frac{dx^*}{d\alpha}(\alpha) \neq 0$). *Assume [H1], [H4], [S1], [S2] and [S5]. Then, if $\varepsilon \rightarrow 0, T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$, the following holds: under \mathbb{P}_θ , using (3.5) for the definition of $\ell(\alpha), \ell(\alpha, \beta)$,*

$$D_{\varepsilon, T}^{(1)} \begin{pmatrix} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\theta) \\ \frac{\partial \Lambda_{\varepsilon, T}}{\partial \beta}(\theta) \end{pmatrix} = \begin{pmatrix} \frac{\varepsilon}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\theta) \\ \frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \beta}(\theta) \end{pmatrix} \rightarrow_{\mathcal{L}} \mathcal{N}_2(\mathbf{0}, \mathcal{J}^{(1)}(\theta)), \quad (3.7)$$

where

$$\mathcal{J}^{(1)}(\theta) = \begin{pmatrix} \ell^2(\alpha, \beta) \left(\frac{\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha))}{\ell(\alpha)} \right)^2 & 0 \\ 0 & \frac{\left(\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \right)^2}{2\ell(\alpha, \beta)} \end{pmatrix}.$$

Moreover, the matrix $-D_{\varepsilon, T}^{(1)} \mathcal{J}_{\varepsilon, T}(\theta) D_{\varepsilon, T}^{(1)} = \mathcal{J}^{(1)}(\theta) + o_P(1)$.

The term $o_P(1) = o_{P_\theta}(1)$ is uniform with respect to θ . Assumptions [S2], [S5] ensure that $D_{\varepsilon,T}^{(1)}$ is the right rate matrix and that $\mathcal{J}^{(1)}(\theta)$ is non singular.

Theorem 3.2. Case (2) ($\forall \alpha, x^*(\alpha) = x^*$). Assume [H1], [H4], [S1], [S2], [S5] and $x_0 \neq x^*$. If $\varepsilon \rightarrow 0, T \rightarrow +\infty$ with $\varepsilon\sqrt{T} \rightarrow 0$, then under \mathbb{P}_θ ,

$$D_{\varepsilon,T}^{(2)} \begin{pmatrix} \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) \\ \frac{\partial \Lambda_{\varepsilon,T}}{\partial \beta}(\theta) \end{pmatrix} = \begin{pmatrix} \varepsilon \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) \\ \frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon,T}}{\partial \beta}(\theta) \end{pmatrix} \rightarrow_{\mathcal{L}} \mathcal{N}_2(\mathbf{0}, \mathcal{J}^{(2)}(\theta)),$$

where $\ell(\alpha, \beta)$ is defined in (3.5) and

$$\mathcal{J}^{(2)}(\theta) = \begin{pmatrix} \int_0^{+\infty} [\frac{\partial V}{\partial \alpha}(\alpha, x_s(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \frac{\partial x_s}{\partial \alpha}(\alpha, s)]^2 ds & 0 \\ 0 & \frac{(\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0))^2}{2\ell(\alpha, \beta)} \end{pmatrix}.$$

Moreover, the matrix $-D_{\varepsilon,T}^{(2)} \mathcal{J}_{\varepsilon,T}(\theta) D_{\varepsilon,T}^{(2)} = \mathcal{J}^{(2)}(\theta) + o_P(1)$.

The term $o_P(1) = o_{P_\theta}(1)$ is uniform with respect to θ . In Theorem 3.2, the additional condition $x_0 \neq x^*$ appears as a minimal assumption. Indeed, since x^* does not depend on α , $x_0 = x^*$ implies that, for all α and all $s \geq 0$, $x_s(\alpha) = x^*$ and using Assumption [S2] and (3.6), the term $\mathcal{J}^{(2)}(\theta)_{11} = 0$. Now, if $x_0 \neq x^*$, the integrand in $\mathcal{J}^{(2)}(\theta)_{11}$ tends to 0 as s tends to ∞ . This convergence is exponential (see Proposition 3.1), so that $\mathcal{J}^{(2)}(\theta)_{11}$ is finite. Assumptions [S2], [S5] ensure that $D_{\varepsilon,T}^{(2)}$ is the right rate matrix and that $\mathcal{J}^{(2)}(\theta)$ is non singular. We stress that Theorems 3.1 and 3.2 show that the estimation of α and β have different rates of convergence. While in both cases, β is estimated at rate \sqrt{T} , according to the assumptions α is estimated at rate $\sqrt{T}\varepsilon^{-1}$ or ε^{-1} .

We can check that these rates hold also for α when $\Phi(\beta, \cdot) \equiv 0$ (i.e. for classical stochastic differential equations), Assumption [S5] being required only for β . This yields the corollary stated below.

Corollary 3.1. Assume that $\Phi(\beta, \cdot) \equiv 0$ (classical stochastic differential equation) and that [H4], [S1] and [S2] hold. The contrast $\Lambda_{\varepsilon,T}(\theta)$ is equal to the exact log-likelihood $\ell_{\varepsilon,T}(\alpha)$ (it depends only on α). Then, if $\varepsilon \rightarrow 0, T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$, the following holds:

If $\frac{dx^*}{d\alpha}(\alpha) \neq 0$, under \mathbb{P}_α ,

$$\frac{\varepsilon}{\sqrt{T}} \ell'_T(\alpha) \rightarrow_{\mathcal{D}} \mathcal{N} \left(0, \left[\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) \right]^2 \right), \quad \frac{\varepsilon^2}{T} \ell''_T(\alpha) \rightarrow_{\mathbb{P}} - \left[\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) \right]^2.$$

If $\frac{dx^*}{d\alpha}(\alpha) \equiv 0$ and $x_0 \neq x^*$, under \mathbb{P}_α ,

$$\varepsilon \ell'_T(\alpha) \rightarrow_{\mathcal{D}} \mathcal{N} \left(0, \int_0^{+\infty} \left[\frac{\partial V}{\partial \alpha}(\alpha, x_s(\alpha)) \right]^2 ds \right),$$

$$\varepsilon^2 \ell_T''(\alpha) \rightarrow_{\mathbb{P}} - \int_0^{+\infty} \left[\frac{\partial V}{\partial \alpha}(\alpha, x_s(\alpha)) \right]^2 ds.$$

Up to our knowledge, these statistical results are also new for classical stochastic differential equations.

3.5. Asymptotic properties of estimators

Let us first study the consistency properties of $(\hat{\alpha}_{\varepsilon, T}, \hat{\beta}_{\varepsilon, T})$ defined in (3.4). Since we are in a two-rate inference case, we first have to study, preliminary to the statistical scheme detailed in Section 3.1, the limits under \mathbb{P}_{θ_0} of the two suitably normalized quantities, $\Lambda_{\varepsilon, T}(\alpha, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta)$, for all $(\alpha, \beta) \in \Theta$ and $\Lambda_{\varepsilon, T}(\alpha_0, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta_0)$ for all $\beta \in \Theta_{\beta}$.

Define the three functions using (3.5),

$$\Lambda_1^{(1)}(\alpha, \alpha_0, \beta) = -\frac{1}{2} (V(\alpha, x^*(\alpha_0)) - \Phi(\beta, x^*(\alpha_0) - x^*(\alpha)))^2, \quad (3.8)$$

$$\Lambda_1^{(2)}(\alpha, \alpha_0, \beta) = -\frac{1}{2} \int_0^{+\infty} [V(\alpha, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)) - \Phi(\beta, x_s(\alpha_0) - x_s(\alpha))]^2 ds, \quad (3.9)$$

$$\Lambda_2(\alpha_0, \beta, \beta_0) = -\frac{1}{2} \left(\frac{\partial \Phi}{\partial x}(\beta, 0) - \frac{\partial \Phi}{\partial x}(\beta_0, 0) \right)^2 \frac{1}{2\ell(\alpha_0, \beta_0)}. \quad (3.10)$$

Lemma 3.1. Assume [H1], [H4], [S0], [S1], [S2]. Then, as $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$, the following holds in probability under \mathbb{P}_{θ_0} :

(i) **Case (1)** ($\frac{dx^*}{d\alpha}(\alpha) \neq 0$). Uniformly with respect to $(\alpha, \beta) \in \Theta_{\alpha} \times \Theta_{\beta}$,

$$\frac{\varepsilon^2}{T} (\Lambda_{\varepsilon, T}(\alpha, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta)) \rightarrow \Lambda_1^{(1)}(\alpha, \alpha_0, \beta).$$

(ii) **Case (2)** ($\frac{dx^*}{d\alpha}(\alpha) \equiv 0$). Uniformly with respect to $(\alpha, \beta) \in \Theta_{\alpha} \times \Theta_{\beta}$,

$$\varepsilon^2 (\Lambda_{\varepsilon, T}(\alpha, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta)) \rightarrow \Lambda_1^{(2)}(\alpha, \alpha_0, \beta).$$

(iii) **Both cases.** Uniformly with respect to $\beta \in \Theta_{\beta}$,

$$\frac{1}{T} (\Lambda_{\varepsilon, T}(\alpha_0, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta_0)) \rightarrow \Lambda_2(\alpha_0, \beta, \beta_0).$$

Let us determine the identifiability assumptions associated with Lemma 3.1 which ensure that the above limits are nul if and only if $\alpha = \alpha_0, \beta = \beta_0$.

For β , the statement (iii) yields that the identifiability assumption [S6] is straightforward in both cases since $\ell(\alpha_0, \beta_0) > 0$ under [S2].

Consider now successively the two cases for α .

Case (1): Assume that $\forall \beta, \Lambda_1^{(1)}(\alpha, \alpha_0, \beta) = 0$. This implies

$$\forall \beta, V(\alpha, x^*(\alpha_0)) = \Phi(\beta, x^*(\alpha_0) - x^*(\alpha)).$$

The left-hand side depends on α , the right-hand side depends on α and β . As $\Phi(\beta, \cdot)$ is an increasing function, this yields that

$$V(\alpha, x^*(\alpha_0)) = 0 \quad \text{and} \quad \Phi(\beta, x^*(\alpha_0) - x^*(\alpha)) = 0. \quad (3.11)$$

Since $\Phi(\beta, x) = 0$ implies $x = 0$, the last equality implies $x^*(\alpha) = x^*(\alpha_0)$. This yields that the identifiability assumption for α is here [S3].

Consider now the case of standard SDE ($\Phi(\beta, \cdot) \equiv 0$).

The condition $\{V(\alpha, x^*(\alpha_0)) = 0\}$ only remains in (3.11). By the uniqueness of the fixed point, this implies that $x^*(\alpha) = x^*(\alpha_0)$, also leading to [S3].

Case (2) (for all $\alpha, x^*(\alpha) = x^*$): If $x_0 = x^*, x_s(\alpha_0) = x_s(\alpha) = x^*$ for all $s \geq 0$. Thus, $\Lambda_1^{(2)}(\alpha, \alpha_0, \beta) = 0$ leading to the additional assumption $x_0 \neq x^*$.

Assume now that $x_0 \neq x^*$. The term under the integral in (3.9) converges to 0 exponentially fast (see Proposition 3.1). Hence, $\Lambda_1^{(2)}(\alpha, \alpha_0, \beta)$ is well defined and finite. This leads to [S4].

In the case $\Phi(\beta, \cdot) \equiv 0$, then [S4] has to be changed into:

- [S4b]: $x_0 \neq x^*$ and $\{s \rightarrow V(\alpha, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)) \equiv 0\} \Rightarrow \{\alpha = \alpha_0\}$.

Theorem 3.3. Assume [H1], [H4], [S0]-[S2], [S5]-[S6] and that $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$.

Case (1) ($\frac{dx^*}{d\alpha}(\alpha) \not\equiv 0$). Assume moreover that [S3] holds. Then $(\hat{\alpha}_{\varepsilon,T}, \hat{\beta}_{\varepsilon,T})$ is consistent and, under \mathbb{P}_{θ_0} ,

$$\left(\begin{array}{c} \frac{\sqrt{T}}{\varepsilon}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) \\ \sqrt{T}(\hat{\beta}_{\varepsilon,T} - \beta_0) \end{array} \right) \rightarrow_{\mathcal{L}} \mathcal{N}_2(0, [\mathcal{J}^{(1)}(\theta_0)]^{-1}), \text{ with } \mathcal{J}^{(1)}(\theta) \text{ defined in Theorem 3.1.}$$

Case (2) ($\forall \alpha, x^*(\alpha) = x^*$). Assume moreover that [S4] holds. Then $(\hat{\alpha}_{\varepsilon,T}, \hat{\beta}_{\varepsilon,T})$ is consistent and, under \mathbb{P}_{θ_0} ,

$$\left(\begin{array}{c} \frac{1}{\varepsilon}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) \\ \sqrt{T}(\hat{\beta}_{\varepsilon,T} - \beta_0) \end{array} \right) \rightarrow_{\mathcal{L}} \mathcal{N}_2(0, [\mathcal{J}^{(2)}(\theta_0)]^{-1}), \text{ with } \mathcal{J}^{(2)}(\theta) \text{ defined in Theorem 3.2.}$$

Let us consider several simple examples that illustrate these results.

Example 1 (continued): Let $V(\alpha, x) = -\alpha x, \Phi(\beta, x) = \beta x, \alpha > 0, \beta > 0$.

As $x^*(\alpha) = x^* = 0$, we are in Case (2). The contrast is equal to the exact log-likelihood, [S5] is satisfied and as $\varepsilon \rightarrow 0, T \rightarrow +\infty$ with $\varepsilon\sqrt{T} \rightarrow 0$, applying Theorem 3.2 yields that, under \mathbb{P}_{θ} ,

$$\left(\begin{array}{c} \varepsilon \frac{\partial \ell_{\varepsilon,T}(\theta)}{\partial \alpha} \\ \frac{1}{\sqrt{T}} \frac{\partial \ell_{\varepsilon,T}(\theta)}{\partial \beta} \end{array} \right) \rightarrow_{\mathcal{L}} \mathcal{N}_2(\mathbf{0}, \mathcal{J}^{(2)}(\theta)) \text{ with}$$

$$\mathcal{J}^{(2)}(\theta) = \begin{pmatrix} x_0^2 \int_0^\infty (1+s\beta)^2 e^{-2\alpha s} ds & 0 \\ 0 & \frac{1}{2(\alpha+\beta)} \end{pmatrix}.$$

The functions $\Lambda_1^{(2)}(\alpha, \alpha_0, \beta)$ and $\Lambda_2(\alpha_0, \beta, \beta_0)$ are explicit.

$$\Lambda_1^{(2)}(\alpha, \alpha_0, \beta) = -x_0^2 \frac{(\alpha + \beta)^2 + \alpha_0 \alpha}{4\alpha_0 \alpha (\alpha_0 + \alpha)} (\alpha - \alpha_0)^2, \quad \Lambda_2(\alpha_0, \beta, \beta_0) = -\frac{(\beta - \beta_0)^2}{4(\alpha_0 + \beta_0)}.$$

Hence, the two identifiability assumptions [S4] and [S6] are satisfied and

$$\begin{pmatrix} \frac{1}{\varepsilon}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) \\ \sqrt{T}(\hat{\beta}_{\varepsilon,T} - \beta_0) \end{pmatrix} \rightarrow_{\mathcal{L}} \mathcal{N}_2(0, \mathcal{J}^{(2)}(\theta_0)^{-1}).$$

Example 2 : Consider the slightly different case $V(\alpha, x) = -\alpha x + 1$, $\Phi(\beta, x) = \beta x$, $\alpha > 0, \beta > 0$. Then $x^*(\alpha) = \alpha^{-1}$ and $\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) = -\alpha^{-1} \neq 0$. We are in Case **(1)**. Applying Theorem 3.1 to $\Lambda_{\varepsilon,T}$, which is the exact log-likelihood, yields

$$\begin{pmatrix} \frac{\varepsilon}{\sqrt{T}} \frac{\partial \ell_{\varepsilon,T}}{\partial \alpha}(\theta) \\ \frac{1}{\sqrt{T}} \frac{\partial \ell_{\varepsilon,T}}{\partial \beta}(\theta) \end{pmatrix} \rightarrow_{\mathcal{L}} \mathcal{N}_2(\mathbf{0}, \mathcal{J}^{(1)}(\theta)) \text{ with } \mathcal{J}^{(1)}(\theta) = \begin{pmatrix} \frac{(\alpha+\beta)^2}{\alpha^4} & 0 \\ 0 & \frac{1}{2(\alpha+\beta)} \end{pmatrix}.$$

Assumptions [S3], [S5] and [S6] are satisfied,

$$\Lambda_1^{(1)}(\alpha, \alpha_0, \beta) = -\frac{(\alpha + \beta)^2}{2\alpha_0^2\alpha^2}(\alpha - \alpha_0)^2, \quad \Lambda_2(\alpha_0, \beta, \beta_0) = -\frac{(\beta - \beta_0)^2}{4(\alpha_0 + \beta_0)}$$

and

$$\begin{pmatrix} \frac{\sqrt{T}}{\varepsilon}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) \\ \sqrt{T}(\hat{\beta}_{\varepsilon,T} - \beta_0) \end{pmatrix} \rightarrow_{\mathcal{L}} \mathcal{N}_2(0, \mathcal{J}^{(1)}(\theta_0)^{-1}).$$

Example 3: Let $V(\alpha, x) = -\alpha x$ and Φ satisfying [H1], [S1], [S5], [S6]. We are in Case **(2)**, $x^*(\alpha) = x^* = 0$, $\ell(\alpha) = \alpha$ and

$$\begin{aligned} \Lambda_1^{(2)}(\alpha, \alpha_0, \beta) &= -\frac{1}{2} \int_0^\infty [x_0(\alpha_0 - \alpha)e^{-\alpha_0 s} - \Phi(\beta, x_0(e^{-\alpha_0 s} - e^{-\alpha s}))]^2 ds \\ \Lambda_2(\alpha_0, \beta, \beta_0) &= -\frac{1}{2} \left(\frac{\partial \Phi}{\partial x}(\beta, 0) - \frac{\partial \Phi}{\partial x}(\beta_0, 0) \right)^2 \times \frac{1}{2(\alpha_0 + \frac{\partial \Phi}{\partial x}(\beta_0, 0))}. \end{aligned} \quad (3.12)$$

We have $\Lambda_1^{(2)}(\alpha, \alpha_0, \beta) = 0$ iff $x_0 = x^*$ or $\alpha = \alpha_0$ and $\Lambda_2(\alpha_0, \beta, \beta_0) = 0$ iff $\beta = \beta_0$. Theorem 3.3 **(2)** yields the consistency and the asymptotic normality.

3.6. Asymptotic efficiency.

The approximate loglikelihood (3.2) corresponds to the exact loglikelihood of the process (ξ_t) defined in (3.1) where we have plugged the observed process (X_t) . It is natural to wonder whether information was lost in this approach. As in [21], we rely on the theory of asymptotic equivalence of experiments (see *e.g.* [37, 38]). Recall that $\mathbb{P}_\theta^{\varepsilon,T}$ denotes the distribution of (X_t) satisfying (1.1) with $X_0 = x_0$ on $(\Omega = C([0, T]), \mathcal{F}, (\mathcal{F}_t, t \in [0, T]))$ (see the notations introduced in Section 3), and let $\mathbb{Q}_\theta^{\varepsilon,T}$ be the distribution of (ξ_t) satisfying (3.1) with $\xi_0 = x_0$. Set $\Theta = \Theta_\alpha \times \Theta_\beta$ and consider the two statistical experiments

$$\mathcal{E}^{\varepsilon,T} = (\Omega, \mathcal{F}, (\mathbb{P}_\theta^{\varepsilon,T})_{\theta \in \Theta}) \quad \text{and} \quad \mathcal{G}^{\varepsilon,T} = (\Omega, \mathcal{F}, (\mathbb{Q}_\theta^{\varepsilon,T})_{\theta \in \Theta}).$$

Let $\Delta(\mathcal{E}^{\varepsilon,T}, \mathcal{G}^{\varepsilon,T})$ be their Le Cam deficiency distance. As the experiments are defined on the same probability space, we have the upper bound:

$$\Delta(\mathcal{E}^{\varepsilon,T}, \mathcal{G}^{\varepsilon,T}) \leq \Delta_0(\mathcal{E}^{\varepsilon,T}, \mathcal{G}^{\varepsilon,T}) = \sup_{\theta \in \Theta} \|\mathbb{P}_\theta^{\varepsilon,T} - \mathbb{Q}_\theta^{\varepsilon,T}\|_{TV},$$

with $\|\cdot\|_{TV}$ the total variation distance. By the Pinsker inequality (see *e.g.* [49]),

$$\|\mathbb{P} - \mathbb{Q}\|_{TV} \leq \sqrt{K(\mathbb{P}, \mathbb{Q})/2},$$

where $K(\mathbb{P}, \mathbb{Q})$ is the Kullback-Leibler divergence of \mathbb{P} with respect to \mathbb{Q} . Set $p(\theta, s, X_s) = V(\alpha, X_s) - b(\theta, \varepsilon, s, X_s)$, the Girsanov formula yields, using (3.3),

$$\varepsilon^2 \log \frac{d\mathbb{P}_\theta^{\varepsilon,T}}{d\mathbb{Q}_\theta^{\varepsilon,T}} = \int_0^T (p(\theta, s, X_s) - H(\theta, s, X_s)) dX_s - \frac{1}{2} \int_0^T (p^2(\theta, s, X_s) - H^2(\theta, s, X_s)) ds.$$

Under $\mathbb{P}_\theta^{\varepsilon,T}$, $dX_s - p(\theta, s, X_s) ds = \varepsilon dW_s$ and $p(\theta, s, X_s) - H(\theta, s, X_s) = D_s$ with $D_s = D(\theta, s, \varepsilon, X_s)$ defined in (2.7). Hence,

$$K(\mathbb{P}_\theta^{\varepsilon,T}, \mathbb{Q}_\theta^{\varepsilon,T}) = \frac{1}{2\varepsilon^2} \mathbb{E}_{\mathbb{P}_\theta^{\varepsilon,T}} \int_0^T D_s^2 ds.$$

Now, we have

$$\varepsilon^{-2} \mathbb{E}_{\mathbb{P}_\theta^{\varepsilon,T}} \int_0^T D_s^2 ds \leq \frac{2}{\varepsilon^2} \left(\int_0^T \mathbb{E}_{\mathbb{P}_\theta^{\varepsilon,T}} ((D_s - \mathbb{E}_{\mathbb{P}_\theta^{\varepsilon,T}} D_s)^2) ds + \int_0^T (\mathbb{E}_{\mathbb{P}_\theta^{\varepsilon,T}} D_s)^2 ds \right).$$

Using Corollary 2.1 yields that under $\mathbb{P}_\theta^{\varepsilon,T}$,

$$K(\mathbb{P}_\theta^{\varepsilon,T}, \mathbb{Q}_\theta^{\varepsilon,T}) \leq \varepsilon^2 T C(\theta),$$

where, by Comment 2.1 and Assumptions [S0]-[S2], $C(\theta)$ is uniformly bounded on Θ . Thus, the Le Cam deficiency distance $\Delta(\mathcal{E}^{\varepsilon,T}, \mathcal{G}^{\varepsilon,T})$ between the two experiments converges to 0 as $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$ under the condition $\varepsilon\sqrt{T} \rightarrow 0$, implying that they are asymptotically equivalent.

This proves the asymptotic efficiency of our method and of the estimators.

4. Concluding remarks

In this paper, we consider the one-dimensional McKean-Vlasov process (X_t) given by (1.1) with small noise ε , under assumptions ensuring existence and uniqueness of solutions. We are interested in the statistical estimation of the unknown parameters α, β present in the classical drift term $V(\alpha, x)$ and in the self-stabilizing term $\Phi(\beta, x)$. In a previous paper ([21]), we have shown that, on the basis of one trajectory continuously observed on a time interval $[0, T]$, while it is possible to estimate consistently α as ε tends to 0, it is not possible to estimate β if T is kept fixed. This is why in this paper, we consider the double

asymptotic framework $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$. This requires some additional assumptions on the model ([H3]-[H4]). In particular, we assume that there is a unique $x^*(\alpha)$ such that $V(\alpha, x^*(\alpha)) = 0$ and this value is an attractive point for the ordinary differential equation $x_t(\alpha) = x_0 + \int_0^t V(\alpha, x_s(\alpha)) ds$. We stress that this double asymptotic framework has never been studied even for classical stochastic differential equations (corresponding to $\Phi(\beta, \cdot) \equiv 0$).

In a first part, we study probabilistic properties of the process (X_t) . We prove that all the moments of $\varepsilon^{-1}(X_t - x_t(\alpha))$ are uniformly bounded in $t \geq 0, \varepsilon \leq 1$, that the Gaussian approximating process of (X_t) as $\varepsilon \rightarrow 0$ obtained in [21] holds on \mathbb{R}^+ and that the remainder terms of this approximation have moments uniformly bounded in $t \geq 0, \varepsilon \leq 1$.

In a second part, we define a contrast (approximate loglikelihood) and prove the consistency and asymptotic normality of the corresponding maximum contrast estimators as $\varepsilon \rightarrow 0$ and $T \rightarrow +\infty$ in such a way that $\varepsilon\sqrt{T} \rightarrow 0$. For the estimation of α , two cases have to be distinguished. Either Case **(1)**, $\frac{dx^*}{d\alpha}(\alpha) \neq 0$ or Case **(2)**, $\frac{dx^*}{d\alpha}(\alpha) \equiv 0$. In Case **(1)**, the estimator of α is asymptotically Gaussian with the fast rate $\sqrt{T}\varepsilon^{-1}$ while in Case **(2)**, its rate is ε^{-1} . In both cases, the parameter β is estimated at rate \sqrt{T} . This confirms the fact that a double asymptotic is needed for estimating both α and β on the basis of one trajectory. We also prove the asymptotic efficiency of our estimators by means of an asymptotic equivalence of experiments property.

Extensions of this work could be to consider multidimensional Mc-Kean-Vlasov models of the more general form (see *e.g.* [43, 48]):

$$dX_t = b(\theta, t, X_t, u_t^{\theta, c})dt + \varepsilon\sigma(c, t, X_t, u_t^{\theta, c})dW_t.$$

where θ, c are unknown parameters, $u_t^{\theta, c}$ is the distribution of X_t .

Another direction would be to study, for discrete observations of McKean-Vlasov diffusions, the estimation of both parameters in the drift and in the diffusion coefficient as in [23, 25].

5. Appendix

5.1. Proofs of Section 2

Proof of Theorem 2.2. Let

$$\zeta_t = (X_t - x_t(\alpha))/\varepsilon \quad \text{with distribution } \nu_t^\varepsilon(dz). \quad (5.1)$$

We have

$$d\zeta_t = \varepsilon^{-1}(V(\alpha, X_t) - V(\alpha, x_t(\alpha)))dt - \frac{1}{\varepsilon} \left(\int \Phi(\beta, X_t - y) u_t^{\theta, \varepsilon}(dy) \right) dt + dW_t.$$

Hence,

$$\zeta_t^2 = 2 \int_0^t \zeta_s dW_s + 2 \int_0^t \zeta_s \varepsilon^{-1} [V(\alpha, X_s) - V(\alpha, x_s(\alpha))] ds$$

$$- 2\varepsilon^{-1} \int_0^t \zeta_s \left(\int \Phi(\beta, X_s - y) u_s^{\theta, \varepsilon}(dy) \right) ds + t.$$

This implies, setting $m_2^\varepsilon(t) := \mathbb{E}_\theta \zeta_t^2$,

$$\begin{aligned} m_2^\varepsilon(t) &= 2 \int_0^t \mathbb{E}_\theta \left[\zeta_s \frac{1}{\varepsilon} (V(\alpha, X_s) - V(\alpha, x_s(\alpha))) \right] ds \\ &\quad - \frac{2}{\varepsilon} \int_0^t \mathbb{E}_\theta \left[\zeta_s \int \Phi(\beta, X_s - y) u_s^{\theta, \varepsilon}(dy) \right] ds + t. \end{aligned} \tag{5.2}$$

Using (5.1), $\Phi(\beta, X_s - y) = \Phi(\beta, X_s - x_s - (y - x_s)) = \Phi(\beta, \varepsilon(\zeta_s - \frac{y - x_s}{\varepsilon}))$. Hence,

$$\begin{aligned} 2\mathbb{E}_\theta \left(\zeta_s \varepsilon^{-1} \int \Phi(\beta, X_s - y) u_s^{\theta, \varepsilon}(dy) \right) &= 2\varepsilon^{-1} \int z \int \Phi(\beta, \varepsilon(z - z')) \varepsilon \nu_s^\varepsilon(dz') \nu_s^\varepsilon(dz) \\ &= \int (z - z') \int \Phi(\beta, \varepsilon(z - z')) \nu_s^\varepsilon(dz) \nu_s^\varepsilon(dz') \geq 0. \end{aligned}$$

Differentiating (5.2) and using [H3], we get $(m_2^\varepsilon)'(t) \leq -2K_V(\alpha)m_2^\varepsilon(t) + 1$.

Now, we can use the following property which holds for $f(\cdot)$ a $C^1(\mathbb{R}^+, \mathbb{R})$ function: If there exists $\ell > 0$ such that

$$\{t \geq 0, f(t) > \ell\} \subset \{t \geq 0, f'(t) < 0\} \quad \text{then} \quad \sup_{t \geq 0} f(t) \leq f(0) \vee \ell. \tag{5.3}$$

Thus, choosing $\ell = \frac{1}{2K_V(\alpha)}$ yields, since $m_2^\varepsilon(0) = 0$,

$$\sup_{t \geq 0} m_2^\varepsilon(t) \leq \frac{1}{2K_V(\alpha)}. \tag{5.4}$$

Let us now study $m_{2n}^\varepsilon(t) := \mathbb{E}_\theta \zeta_t^{2n}$. We have

$$\zeta_t^{2n} = 2n \int_0^t \zeta_s^{2n-1} d\zeta_s + n(2n - 1) \int_0^t \zeta_s^{2n-2} ds. \tag{5.5}$$

Analogously, for $n \geq 1$, using that $\Phi(\beta, \cdot)$ is odd,

$$\begin{aligned} &2n \mathbb{E}_\theta \left(\zeta_s^{2n-1} \frac{1}{\varepsilon} \int \Phi(\beta, X_s - y) u_s^{\theta, \varepsilon}(dy) \right) \\ &= 2n\varepsilon^{-1} \int z^{2n-1} \int \Phi(\beta, \varepsilon(z - z')) \varepsilon \nu_s^\varepsilon(dz) \nu_s^\varepsilon(dz') \\ &= n \int (z^{2n-1} - z'^{2n-1}) \int \Phi(\beta, \varepsilon(z - z')) \nu_s^\varepsilon(dz) \nu_s^\varepsilon(dz') \geq 0. \end{aligned}$$

The first term of $d\zeta_s$ in (5.5) satisfies under [H3],

$$\begin{aligned} \mathbb{E}_\theta \left[\zeta_s^{2n-1} \frac{1}{\varepsilon} (V(\alpha, X_s) - V(\alpha, x_s(\alpha))) \right] &= \mathbb{E}_\theta \left[\zeta_s^{2n-2} \zeta_s \frac{1}{\varepsilon} (V(\alpha, X_s) - V(x_s(\alpha))) \right] \\ &\leq -K_V(\alpha) \mathbb{E}_\theta \zeta_s^{2n}. \end{aligned}$$

Therefore, applying the Hölder inequality to $f(x) = x^{1-1/n}$, we get

$$\begin{aligned} (m_{2n}^\varepsilon(t))' &\leq -2nK_V(\alpha) m_{2n}^\varepsilon(t) + n(2n-1)m_{2n-2}^\varepsilon(t) \\ &\leq -2nK_V(\alpha) m_{2n}^\varepsilon(t) + n(2n-1)(m_{2n}^\varepsilon(t))^{1-1/n}. \end{aligned}$$

Choosing $\delta(\alpha, n) = (\frac{n}{K_V(\alpha)})^n$, we have that, for $x \geq \delta(\alpha, n)$, $-2nK_V(\alpha)x + n(2n-1)x^{1-1/n} < 0$. Thus, as $m_{2n}^\varepsilon(0) = 0$, applying (5.3) yields

$$\sup_{t \geq 0} m_{2n}^\varepsilon(t) \leq \delta(\alpha, n), \text{ with } \delta(\alpha, n) \text{ independent of } \varepsilon, \beta. \tag{5.6}$$

It remains to study $\mathbb{E}_\theta(X_t - x_t(\alpha))$. We have,

$$\mathbb{E}_\theta(X_t - x_t(\alpha)) = \int_0^t \mathbb{E}_\theta(V(\alpha, X_s) - V(\alpha, x_s(\alpha))) ds - \int_0^t \mathbb{E}_\theta b(\theta, s, \varepsilon, X_s) ds.$$

Let (\bar{X}_s) be an independent copy of (X_s) . Then,

$$\mathbb{E}_\theta b(\theta, s, \varepsilon, X_s) = \mathbb{E}_\theta \int \Phi(\beta, X_s - y) u_s^{\theta, \varepsilon}(dy) = \mathbb{E}_\theta(\Phi(\beta, X_s - \bar{X}_s)) = 0, \tag{5.7}$$

since $\Phi(\beta, \cdot)$ is odd and since the distribution of $X_s - \bar{X}_s$ is symmetric. Now, a Taylor expansion at $x_s(\alpha)$ yields

$$\begin{aligned} \mathbb{E}_\theta(V(\alpha, X_s) - V(\alpha, x_s(\alpha))) &= \mathbb{E}_\theta(X_s - x_s(\alpha)) \frac{\partial V}{\partial x}(\alpha, x_s(\alpha)) + R_s, \text{ with} \\ R_s &= \int_0^1 (1-u) \mathbb{E}_\theta \left((X_s - x_s(\alpha))^2 \frac{\partial^2 V}{\partial x^2}(\alpha, x_s(\alpha) + u(X_s - x_s(\alpha))) \right) du. \end{aligned}$$

Therefore,

$$f(\theta, t) := \mathbb{E}_\theta X_t - x_t(\alpha) = \int_0^t (\mathbb{E}_\theta X_s - x_s(\alpha)) \frac{\partial V}{\partial x}(\alpha, x_s(\alpha)) ds + \int_0^t R_s ds.$$

Differentiating with respect to t yields $\frac{\partial f}{\partial t}(\theta, t) = \frac{\partial V}{\partial x}(\alpha, x_t(\alpha)) f(\theta, t) + R_t$. Consequently, using that $f(\theta, 0) = 0$, we get

$$f(\theta, t) = \int_0^t R_s \exp \left(\int_s^t \frac{\partial V}{\partial x}(\alpha, x_u(\alpha)) du \right) ds. \tag{5.8}$$

Using [H2-2], $|R_s| \leq k(\alpha) \mathbb{E}_\theta \left((X_s - x_s(\alpha))^2 (1 + |x_s(\alpha)|^{\gamma(\alpha)} + |X_s - x_s(\alpha)|^{\gamma(\alpha)}) \right)$. Under [H3], [H4], $x_t(\alpha)$ is uniformly bounded on \mathbb{R}^+ by $B(\alpha)$ (see Comment 2.1). Using the first part, $\mathbb{E}_\theta(X_t - x_t(\alpha))^2 \leq \frac{\varepsilon^2}{2K_V(\alpha)}$. By the Hölder inequality,

$$\mathbb{E}_\theta |X_t - x_t(\alpha)|^{2+\gamma(\alpha)} \leq \varepsilon^{2+\gamma(\alpha)} (m_{2+2\gamma(\alpha)}^\varepsilon(t))^{1-\frac{\gamma(\alpha)}{2+2\gamma(\alpha)}}.$$

Therefore, $|R_s| \leq \varepsilon^2 C(\alpha)$ where

$C(\alpha) = \frac{k(\alpha)}{2K_V(\alpha)} (1 + B(\alpha)^{\gamma(\alpha)} + 2(1 + \gamma(\alpha)) (\frac{1+\gamma(\alpha)}{K_V(\alpha)})^{\frac{\gamma(\alpha)}{1+\gamma(\alpha)}})$ is independent of t, ε . Hence, $\mathbb{E}_\theta X_t - x_t(\alpha) = f(\theta, t)$ satisfies

$$|f(\theta, t)| \leq \varepsilon^2 C(\alpha) \int_0^t e^{-K_V(\alpha)(t-s)} ds \leq \frac{C(\alpha)}{K_V(\alpha)} \varepsilon^2. \quad \square \tag{5.9}$$

If $V(\alpha, x) = -\alpha x$, $\mathbb{E}_\theta X_t = x_0 - \alpha \int_0^t \mathbb{E}_\theta X_s ds$. Thus, $\mathbb{E}_\theta X_t = x_0 e^{-\alpha t} = x_t(\alpha)$.

Proof of Theorem 2.3. Recall that $x \rightarrow \Phi(\beta, x)$ is odd so that the odd derivatives are odd and therefore null at $x = 0$.

By (2.3), we have $R_t^\varepsilon(\theta) = \varepsilon^{-2}(X_t - x_t(\alpha) - \varepsilon g_t(\theta))$. Therefore, using (1.1), (1.3), (2.1) and (2.2),

$$\begin{aligned} dR_t^\varepsilon(\theta) &= \frac{1}{\varepsilon^2} (V(\alpha, X_t) - V(\alpha, x_t(\alpha)) - b(\theta, t, \varepsilon, X_t) - \varepsilon a(\theta, t)g_t(\theta)) dt \\ &= \frac{1}{\varepsilon^2} [(X_t - x_t(\alpha))a(\theta, t) - \varepsilon a(\theta, t)g_t(\theta)] dt + \nu(\theta, t, \varepsilon, X_t)dt \\ &= a(\theta, t)R_t^\varepsilon(\theta)dt + \nu(\theta, t, \varepsilon, X_t)dt, \quad R_0^\varepsilon(\theta) = 0, \text{ where} \end{aligned}$$

$$\nu(\theta, t, \varepsilon, X_t) = T_1(t) + T_2(t), \text{ with}$$

$$T_1(t) = \varepsilon^{-2} (V(\alpha, X_t) - V(\alpha, x_t(\alpha)) - (X_t - x_t(\alpha)) \frac{\partial V}{\partial x}(\alpha, x_t(\alpha))),$$

$$T_2(t) = -\varepsilon^{-2} \left(\int \Phi(\beta, X_t - y) u_t^{\theta, \varepsilon}(dy) - \frac{\partial \Phi}{\partial x}(\beta, 0)(X_t - x_t(\alpha)) \right). \quad (5.10)$$

The equation satisfied by $R_t^\varepsilon(\theta)$ can be solved. We get, using (2.1) and (2.4),

$$R_t^\varepsilon(\theta) = \int_0^t \nu(\theta, s, \varepsilon, X_s) \exp\left(\int_s^t a(\theta, u)du\right) ds. \quad (5.11)$$

Let us first study $T_1(t)$. A Taylor expansion at point $x_t(\alpha)$ yields, using [H2-2],

$$T_1(t) = \varepsilon^{-2} (X_t - x_t(\alpha))^2 \int_0^1 (1 - u) \frac{\partial^2 V}{\partial x^2}(\alpha, x_t(\alpha) + u(X_t - x_t(\alpha))) du, \quad (5.12)$$

$$|T_1(t)| \leq k(\alpha) \varepsilon^{-2} (X_t - x_t(\alpha))^2 (1 + |x_t(\alpha)|^{\gamma(\alpha)} + |X_t - x_t(\alpha)|^{\gamma(\alpha)}).$$

Therefore, since $x_t(\alpha)$ is uniformly bounded, applying Theorem 2.2 yields that, for all $p \geq 1$,

$$\mathbb{E}_\theta |T_1(t)|^{2p} = O(1) \text{ uniformly on } t \geq 0, \varepsilon > 0.$$

For $T_2(t)$, we have $-\varepsilon^2 T_2(t) = \int (\Phi(\beta, X_t - y) - \frac{\partial \Phi}{\partial x}(\beta, 0)(X_t - x_t(\alpha))) u_t^{\theta, \varepsilon}(dy)$.

A Taylor expansion at point 0 yields, noting that $\frac{\partial^2 \Phi}{\partial x^2}(\beta, 0) = 0$,

$$\Phi(\beta, X_t - y) - \frac{\partial \Phi}{\partial x}(\beta, 0)(X_t - x_t(\alpha)) = \frac{\partial \Phi}{\partial x}(\beta, 0)(x_t(\alpha) - y) + \rho_1(X_t, y),$$

where

$$\rho_1(X_t, y) = \frac{1}{2} (X_t - y)^3 \int_0^1 (1 - u)^2 \frac{\partial^3 \Phi}{\partial x^3}(\beta, u(X_t - y)) du. \quad (5.13)$$

Therefore, $T_2(t) = -\varepsilon^{-2} \frac{\partial \Phi}{\partial x}(\beta, 0)(x_t(\alpha) - \mathbb{E}_\theta X_t) - \varepsilon^{-2} \int \rho_1(X_t, y) u_t^{\theta, \varepsilon}(dy) = T_{21}(t) + T_{22}(t)$.

Let us study first $\mathbb{E}_\theta T_2(t) = T_{21}(t) + \mathbb{E}_\theta T_{22}(t)$. For the second term, we can write, for \bar{X}_t an independent copy of X_t ,

$$\mathbb{E}_\theta \int \rho_1(X_t, y) u_t^{\theta, \varepsilon}(dy) = \frac{1}{2} \mathbb{E}_\theta \left((X_t - \bar{X}_t)^3 \int_0^1 (1-u)^2 \frac{\partial^3 \Phi}{\partial x^3}(\beta, u(X_t - \bar{X}_t)) du \right).$$

Under [H2-3], $x \rightarrow x^3 \frac{\partial^3 \Phi}{\partial x^3}(\beta, ux)$ is well defined and odd so that

$$\mathbb{E}_\theta \int \rho_1(X_t, y) u_t^{\theta, \varepsilon}(dy) = \mathbb{E}_\theta(\rho_1(X_t, \bar{X}_t)) = 0. \quad (5.14)$$

Therefore, $\mathbb{E}_\theta T_{22}(t) = 0$. For $T_{21}(t)$ which is deterministic, applying Theorem 2.2 (ii) yields

$$|T_{21}(t)| \leq \delta(\alpha) \frac{\partial \Phi}{\partial x}(\beta, 0) = O(1) \quad \text{uniformly on } t \geq 0, \varepsilon > 0.$$

Therefore $|\mathbb{E}_\theta T_2(t)| = |T_{21}(t)|$ is also uniformly bounded for $t \geq 0, \varepsilon > 0$. Consider $T_2(t) - \mathbb{E}_\theta T_2(t) = T_{22}(t)$. By (5.14), $T_{22}(t) = -\varepsilon^{-2} \int \rho_1(X_t, y) u_t^{\theta, \varepsilon}(dy)$. Hence, if (\bar{X}_t) is an independent copy of (X_t) ,

$$\mathbb{E}_\theta(T_2(t) - \mathbb{E}_\theta T_2(t))^{2p} = \varepsilon^{-4p} \mathbb{E}_\theta(\rho_1(X_t, \bar{X}_t)^{2p}).$$

Now, by [H2-3] and (5.13),

$$\begin{aligned} \mathbb{E}_\theta(\rho_1(X_t, \bar{X}_t)^{2p}) &= 2^{-2p} \mathbb{E}_\theta \left((X_t - \bar{X}_t)^{6p} \left(\int_0^1 (1-u)^2 \frac{\partial^3 \Phi}{\partial x^3}(\beta, u(X_t - \bar{X}_t)) du \right)^{2p} \right) \\ &\leq 2^{-2p} \mathbb{E}_\theta \left(|X_t - \bar{X}_t|^{6p} (k(\beta)(1 + |X_t - \bar{X}_t|^{\gamma(\beta)})^{2p}) \right) \\ &\leq k^{2p}(\beta) 2^{-1} \mathbb{E}_\theta \left(|X_t - \bar{X}_t|^{6p} (1 + |X_t - \bar{X}_t|^{2p\gamma(\beta)}) \right). \end{aligned}$$

By splitting $X_t - \bar{X}_t$ into $X_t - x_t(\alpha) + x_t(\alpha) - \bar{X}_t$ we get that

$$\begin{aligned} \mathbb{E}_\theta(\rho_1(X_t, \bar{X}_t)^{2p}) &\leq k^{2p}(\beta) \mathbb{E}_\theta(2^{6p-1} (X_t - x_t(\alpha))^{6p}) \\ &\quad + k^{2p}(\beta) \mathbb{E}_\theta(2^{6p+2p\gamma(\beta)-1} (X_t - x_t(\alpha))^{6p+2p\gamma(\beta)}) \leq C_p(\alpha, \beta) \varepsilon^{6p}, \end{aligned}$$

where $C_p(\alpha, \beta)$ depends on $p, k(\beta)$ and $K_V^{-1}(\alpha)$. Applying Theorem 2.2 yields that, uniformly on $t > 0, \varepsilon > 0$,

$$\mathbb{E}_\theta(T_2(t) - \mathbb{E}_\theta T_2(t))^{2p} \leq \varepsilon^{2p} C_p'(\alpha, \beta).$$

Joining these inequalities there exist constants $\delta(\alpha, \beta), \delta_p(\alpha, \beta)$ such that for all $t \geq 0, \varepsilon > 0$,

$$\mathbb{E}_\theta |\nu(\theta, t, \varepsilon, X_t)| \leq \delta(\alpha, \beta); \quad \mathbb{E}_\theta (\nu(\theta, t, \varepsilon, X_t) - \mathbb{E}_\theta \nu(\theta, t, \varepsilon, X_t))^{2p} \leq \delta_p(\alpha, \beta).$$

Now, using (3.5), (2.5) and [H3], $\int_s^t a(\theta, u) du = A(\theta, t) - A(\theta, s) \leq -K(\theta)(t-s)$ with

$$K(\theta) = K(\alpha, \beta) = K_V(\alpha) + \frac{\partial \Phi}{\partial x}(\beta, 0) > 0. \quad (5.15)$$

Therefore (5.11) yields that

$$\mathbb{E}_\theta |R_t^\varepsilon(\theta)| \leq \int_0^t \mathbb{E}_\theta |\nu(\theta, s, \varepsilon, X_s)| e^{-K(\theta)(t-s)} ds \leq \frac{\delta(\alpha, \beta)}{K(\theta)}.$$

Consider now $\mathbb{E}_\theta (R_t^\varepsilon(\theta) - \mathbb{E}_\theta R_t^\varepsilon(\theta))^{2p}$. Equation (5.11) yields

$$\begin{aligned} (R_t^\varepsilon(\theta) - \mathbb{E}_\theta R_t^\varepsilon(\theta))^{2p} &\leq \left(\int_0^t (\nu(\theta, t, \varepsilon, X_t) - \mathbb{E}_\theta \nu(\theta, t, \varepsilon, X_t))^{2p} e^{p(A(\theta, t) - A(\theta, s))} ds \right) \\ &\quad \times \left(\int_0^t e^{\frac{p}{2p-1}(A(\theta, t) - A(\theta, s))} ds \right)^{2p-1}. \end{aligned}$$

Using the inequality for $A(\theta, s)$ yields that $M_p = \mathbb{E}_\theta (R_t^\varepsilon(\theta) - \mathbb{E}_\theta R_t^\varepsilon(\theta))^{2p}$ satisfies

$$M_p \leq \left(\frac{2p-1}{2pK(\theta)} \right)^{2p-1} \int_0^t \mathbb{E}_\theta (\nu(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta \nu(\theta, s, \varepsilon, X_s))^{2p} e^{-pK(\theta)(t-s)} ds.$$

Therefore, this expectation is uniformly bounded on $t \geq 0, \varepsilon > 0$. \square

Proof of Corollary 2.1 . We have that

$$D(\theta, t, \varepsilon, X_t) = \int (\Phi(\beta, X_t - y) - \Phi(\beta, X_t - x_t(\alpha))) u_t^{\theta, \varepsilon}(dy).$$

Similarly to the study $T_2(t)$, a Taylor expansion of $\Phi(\beta, \cdot)$ yields, using (5.13),

$$\Phi(\beta, X_t - y) - \Phi(\beta, X_t - x_t(\alpha)) = \frac{\partial \Phi}{\partial x}(\beta, 0)(x_t(\alpha) - y) + \rho_1(X_t, y) - \rho_2(X_t),$$

with

$$\rho_2(X_t) = \frac{1}{2}(X_t - x_t(\alpha))^3 \int_0^1 (1-u)^2 \frac{\partial^3 \Phi}{\partial x^3}(\beta, u(X_t - x_t(\alpha))) du. \tag{5.16}$$

Therefore, $D(\theta, t, \varepsilon, X_t) = \frac{\partial \Phi}{\partial x}(\beta, 0)(x_t(\alpha) - \mathbb{E}_\theta X_t) + \int \rho_1(X_t, y) u_t^{\theta, \varepsilon}(dy) - \rho_2(X_t)$. Using (5.14),

$$\mathbb{E}_\theta D(\theta, t, \varepsilon, X_t) = \frac{\partial \Phi}{\partial x}(\beta, 0)(x_t(\alpha) - \mathbb{E}_\theta X_t) - \mathbb{E}_\theta \rho_2(X_t).$$

By Theorem 2.2, $\mathbb{E}_\theta |\rho_2(X_t)| \lesssim \varepsilon^3 O(1)$. This yields (2.8). Moreover, as for the upper bound of $T_2(t)$, $\mathbb{E}_\theta |\rho_1(X_t, \bar{X}_t)|^{2p} \lesssim \mathbb{E}_\theta |X_t - \bar{X}_t|^{6p} \lesssim \varepsilon^{6p}$.

By Theorem 2.2, uniformly on $t > 0$, $\mathbb{E}_\theta |\rho_2(X_t)|^{2p} \lesssim \mathbb{E}_\theta |X_t - x_t(\alpha)|^{6p} \leq \varepsilon^{6p} O(1)$. Joining these two inequalities, we get (2.9). \square

5.2. Proofs of Section 3

We start with two preliminary propositions useful for the inference.

Proof of Proposition 3.1.

Set $x_1(t) = x_t(\alpha), x_2(t) = \frac{\partial x_t}{\partial \alpha}(\alpha, t), x_3(t) = \frac{\partial^2 x_t}{\partial \alpha^2}(\alpha, t)$ and $\mathbf{x}(t) = [x_1(t) x_2(t) x_3(t)]'$. Then, $\mathbf{x}(t)$ is solution of the ordinary differential equation

$$d\mathbf{x}(t) = B(\mathbf{x}(t))dt, \quad B(\mathbf{x}(t)) = [B_1(\mathbf{x}(t)) \ B_2(\mathbf{x}(t)) \ B_3(\mathbf{x}(t))]' \quad (5.17)$$

where $B_1(\mathbf{x}) = V(\alpha, x_1), B_2(\mathbf{x}) = \frac{\partial V}{\partial \alpha}(\alpha, x_1) + \frac{\partial V}{\partial x}(\alpha, x_1)x_2$ and $B_3(\mathbf{x}) = \frac{\partial^2 V}{\partial \alpha^2}(\alpha, x_1) + [\frac{\partial V}{\partial x}(\alpha, x_1) + 2\frac{\partial^2 V}{\partial x \partial \alpha}(\alpha, x_1)]x_2 + \frac{\partial V}{\partial x}(\alpha, x_1)x_2^2 + \frac{\partial V}{\partial x}(\alpha, x_1)x_3$. We easily check that $B(\mathbf{x}^*) = 0$ for $\mathbf{x}^* = [x_1^* x_2^* x_3^*]'$ with, using (3.5) for $\ell(\alpha)$,

$$x_1^* = x^*(\alpha), \quad x_2^* = \frac{1}{\ell(\alpha)} \frac{\partial V}{\partial \alpha}(\alpha, x_1^*), \quad (5.18)$$

$$x_3^* = \frac{1}{\ell(\alpha)} \left(\frac{\partial^2 V}{\partial \alpha^2} + x_2^* \left[\frac{\partial V}{\partial x} + 2 \frac{\partial^2 V}{\partial x \partial \alpha} \right] + (x_2^*)^2 \frac{\partial V}{\partial x} \right) (\alpha, x_1^*). \quad (5.19)$$

We have to compute $DB(\mathbf{x}^*) = [\frac{\partial B_i}{\partial x_j}(\mathbf{x}^*)]_{1 \leq i, j \leq 3}$ to check if this point is asymptotically stable. The matrix $DB(\mathbf{x})$ is triangular with diagonal elements equal to $-\ell(\alpha) < 0$. Thus, the eigenvalues of $DB(\mathbf{x}^*)$ are negative which implies that \mathbf{x}^* is asymptotically stable for (5.17). Thus $\mathbf{x}(t)$ converges as $t \rightarrow +\infty$ to \mathbf{x}^* with exponential rate $\exp(-\ell(\alpha)t)$ (see *e.g.* [27]). \square

Note that $\alpha \rightarrow x^*(\alpha)$ is C^∞ on U_α . As Θ_α is compact, $\frac{dx^*}{d\alpha}(\alpha)$ and $\frac{d^2x^*}{d\alpha^2}(\alpha)$ are uniformly bounded on Θ_α as well as $x^*(\alpha)$.

Proof of Proposition 3.2.

Proof of (i): Consider the process (g_t) such that $dg_t = -\lambda g_t dt + dW_t, g_t(0) = 0$ with $\lambda > 0$. It is standard that $g_t = \exp(-\lambda t) \int_0^t \exp(\lambda s) dW_s$ and that (g_t) is a positive recurrent diffusion with invariant distribution $\mathcal{N}(0, 1/(2\lambda))$. By the ergodic theorem, $\frac{1}{T} \int_0^T g_s^2 ds$ converges a.s. to $(1/2\lambda)$. This implies, by the central limit theorem for martingales, that $\frac{1}{\sqrt{T}} \int_0^T g_s dW_s$ converges in distribution to $\mathcal{N}(0, (1/2\lambda))$. Moreover, one easily gets

$$\mathbb{E}\left(\frac{1}{T} \int_0^T g_s^2 ds\right) = \frac{1}{T} \int_0^T \mathbb{E}g_s^2 ds = \frac{1}{2\lambda} + o(1).$$

We can also compute

$$\mathbb{E}\frac{1}{T^2} \left(\int_0^T g_s^2 ds \right)^2 = \frac{1}{T^2} \int_{[0, T]^2} \mathbb{E}(g_s^2 g_t^2) ds dt.$$

If $\mathcal{L}(X, Y) = \mathcal{N}_2(0, \Sigma)$ with $\Sigma = \begin{pmatrix} \sigma^2 & \rho \\ \rho & \tau^2 \end{pmatrix}$. Then $\mathbb{E}(X^2 Y^2) = \sigma^2 \tau^2 + 2\rho^2$.

Applying this property to the centered Gaussian process (g_t) , we get that $\mathbb{E}(g_s^2 g_t^2) = 2\text{cov}^2(g_s, g_t) + \mathbb{E}g_s^2 \mathbb{E}g_t^2$. Therefore,

$$\mathbb{E}\frac{1}{T^2} \left(\int_0^T g_s^2 ds \right)^2 = \left(\frac{1}{T} \int_0^T \mathbb{E}g_s^2 ds \right)^2 + C_T(\lambda). \quad (5.20)$$

where

$$C_T(\lambda) = \frac{2}{T^2} \int_{[0,T]^2} \text{cov}^2(g_s, g_t) ds dt = \frac{4}{T^2} \int_{0 \leq s \leq t \leq T} \text{cov}^2(g_s, g_t) ds dt. \tag{5.21}$$

For $s \leq t$, $\text{cov}(g_s, g_t) = \int_0^s e^{-\lambda(t-u+s-u)} du = \frac{1}{2\lambda} (e^{\lambda(s-t)} - e^{\lambda(s+t)})$.
 By elementary computations, we see that $C_T(\lambda) = \frac{1}{T}O(1)$ so that

$$\mathbb{E} \left(\frac{1}{T} \int_0^T g_s^2 ds - \frac{1}{T} \int_0^T \mathbb{E} g_s^2 ds \right)^2 \rightarrow 0.$$

With this direct calculus, we have obtained that $\frac{1}{T} \int_0^T g_s^2 ds \rightarrow_{\mathbb{L}^2} \frac{1}{2\lambda}$.

We rely on this approach to prove Proposition 3.2 for the process $g_t(\theta)$. Using (2.1), (2.5) and (5.15), we have that under [H3], for $u \leq t$,

$$A(\theta, t) - A(\theta, u) \leq -K(\theta)(t - u). \tag{5.22}$$

Moreover, by (2.4), $g_t(\theta) = \int_0^t e^{A(\theta,t)-A(\theta,s)} dW_s$.

Equations (5.20)-(5.21) hold for $g_t(\theta)$,

$$\mathbb{E}_\theta \frac{1}{T^2} \left(\int_0^T [g_s(\theta)]^2 ds \right)^2 = \left(\frac{1}{T} \int_0^T \mathbb{E}_\theta [g_s(\theta)]^2 ds \right)^2 + \tilde{C}_T(\theta) \quad \text{with}$$

$\tilde{C}_T(\theta) = \frac{4}{T^2} \int_{0 \leq s \leq t \leq T} \text{cov}_\theta^2(g_s(\theta), g_t(\theta)) ds dt$. For $s \leq t$, using (5.15) and (5.22)

$$\text{cov}_\theta(g_s(\theta), g_t(\theta)) = \int_0^s e^{A(\theta,t)+A(\theta,s)-2A(\theta,u)} du \leq \int_0^s e^{-K(\theta)(t-u+s-u)} du.$$

Therefore $\tilde{C}_T(\theta) \leq C_T(K(\theta))$. Finally, using (5.20)-(5.21),

$$\mathbb{E}_\theta \left(\frac{1}{T} \int_0^T [g_s(\theta)]^2 ds - \frac{1}{T} \int_0^T \mathbb{E}_\theta [g_s(\theta)]^2 ds \right)^2 \leq C_T(K(\theta)) = \frac{1}{T}O(1).$$

Thus,

$$\frac{1}{T} \int_0^T [g_s(\theta)]^2 ds - \frac{1}{T} \int_0^T \mathbb{E}_\theta [g_s(\theta)]^2 ds \rightarrow_{\mathbb{L}^2} 0. \tag{5.23}$$

Now, the function $t \rightarrow \frac{\partial V}{\partial x}(\alpha, x_t(\alpha))$ is continuous. Therefore, under [H3]-[H4], as $t \rightarrow +\infty$, $x_t(\alpha) \rightarrow x^*(\alpha)$, and $\frac{\partial V}{\partial x}(\alpha, x_t(\alpha)) \rightarrow \frac{\partial V}{\partial x}(\alpha, x^*(\alpha)) = -\ell(\alpha) \leq -K_V(\alpha) < 0$. Therefore,

$$\forall h > 0, \exists t_0 > 0, \forall t \geq t_0, -\ell(\alpha) - h < \frac{\partial V}{\partial x}(\alpha, x_t(\alpha)) < -\ell(\alpha) + h.$$

It follows, using (3.5), that for all t, s such that $t \geq s \geq t_0$,

$$-(\ell(\alpha, \beta) + h)(t - s) \leq A(\theta, t) - A(\theta, s) \leq -(\ell(\alpha, \beta) - h)(t - s). \tag{5.24}$$

Choose $h > 0$ such that $\ell(\alpha, \beta) - h > 0$. We have, using (2.4), $\mathbb{E}_\theta(g_t(\theta)^2) = e^{2A(\theta,t)} \int_0^t e^{-2A(\theta,s)} ds$. Hence, $\mathbb{E}_\theta(T^{-1} \int_0^T [g_t(\theta)]^2 dt) = T_1 + T_2 + T_3$ where

$$\begin{aligned} T_1 &= \frac{1}{T} \int_{0 \leq s \leq t \leq t_0} e^{2(A(\theta,t) - A(\theta,s))} ds dt, \quad T_3 = \frac{1}{T} \int_{t_0 \leq s \leq t \leq T} e^{2(A(\theta,t) - A(\theta,s))} ds dt, \\ T_2 &= \frac{1}{T} \int_{0 \leq s \leq t_0, t_0 \leq t \leq T} e^{2(A(\theta,t) - A(\theta,s))} ds dt. \end{aligned}$$

As T tends to infinity, $T_1 = o(1)$. For T_2 we have,

$$T_2 = \frac{1}{T} \int_0^{t_0} e^{-2(A(\theta,s) - A(\theta,t_0))} ds \times \int_{t_0}^T e^{2(A(\theta,t) - A(\theta,t_0))} dt.$$

Now, using (5.24),

$$\int_{t_0}^T e^{2(A(\theta,t) - A(\theta,t_0))} dt \leq \int_{t_0}^T e^{-2(\ell(\alpha,\beta) - h)(t - t_0)} dt \leq \frac{1}{2(\ell(\alpha, \beta) - h)}.$$

Therefore $0 \leq T_2 \leq \frac{1}{T} O(1)$ and $T_2 \rightarrow 0$ as $T \rightarrow \infty$. Consider now T_3 :

$$\begin{aligned} T_3 &\leq \frac{1}{T} \int_{t_0}^T ds \int_s^T e^{-2(\ell(\alpha,\beta) - h)(t - s)} dt \\ &= \frac{1}{2(\ell(\alpha, \beta) - h)T} \left(T - t_0 - \frac{1 - e^{-2(\ell(\alpha,\beta) - h)(T - t_0)}}{2(\ell(\alpha, \beta) - h)} \right). \end{aligned}$$

Therefore, $\overline{\lim}_{T \rightarrow +\infty} T_3 \leq \frac{1}{2(\ell(\alpha,\beta) - h)}$. Analogously, $\underline{\lim}_{T \rightarrow +\infty} T_3 \geq \frac{1}{2(\ell(\alpha,\beta) + h)}$. Therefore, $T_3 \rightarrow \frac{1}{2(\ell(\alpha,\beta))}$ so that

$$\mathbb{E}_\theta \left(\frac{1}{T} \int_0^T [g_t(\theta)]^2 dt \right) \rightarrow \frac{1}{2\ell(\alpha, \beta)} \text{ as } T \rightarrow \infty.$$

Using (5.23), the first item is proved.

Proof of (ii): Let $Z_T = \int_0^T g_t(\theta) dt$. Using (2.4) and interchanging the order of integrations yields:

$$Z_T = \int_0^T g_t(\theta) dt = \int_0^T e^{A(\theta,t)} \int_0^t e^{-A(\theta,s)} dW_s dt = \int_0^T e^{-A(\theta,s)} dW_s \int_s^T e^{A(\theta,t)} dt.$$

Therefore, Z_T is centered and, using (5.22)

$$\begin{aligned} \mathbb{E} Z_T^2 &= \int_0^T e^{-2A(\theta,s)} ds \left(\int_s^T e^{A(\theta,t)} dt \right)^2 = \int_0^T ds \left(\int_s^T e^{A(\theta,t) - A(\theta,s)} dt \right)^2 \\ &\leq \int_0^T ds \left(\int_s^T e^{-K(\theta)(t-s)} dt \right)^2 \leq \int_0^T ds \left(\frac{1 - e^{-K(\theta)(T-s)}}{K(\theta)} \right)^2 \leq \frac{T}{K^2(\theta)}. \end{aligned}$$

Therefore, we find that $\mathbb{E}Z_T^2 \lesssim T$ and $T^{-1}Z_T$ tends to 0 in probability as T tends to infinity.

Proof of (iii): As $\lim_{t \rightarrow +\infty} h(t) = 0$, for all $h > 0$, there exists $T_0 > 0$ such that for all $T \geq T_0$, $|h(t)| \leq h$. So, we split

$$\begin{aligned} \frac{1}{\sqrt{T}} \int_0^T h(s)g_s(\theta)ds &= \frac{1}{\sqrt{T}} \int_0^{T_0} h(s)g_s(\theta)ds + \frac{1}{\sqrt{T}} \int_{T_0}^T h(s)g_s(\theta)ds \\ &= o_P(1) + \frac{1}{\sqrt{T}}Z(T_0, T), \text{ with} \end{aligned}$$

$$\begin{aligned} Z(T_0, T) &= \int_{T_0}^T h(s)e^{A(\theta,s)} \left(\int_0^{T_0} e^{-A(\theta,u)}dW_u + \int_{T_0}^s e^{-A(\theta,u)}dW_u \right) ds \\ &= \int_0^{T_0} e^{-A(\theta,u)}dW_u \int_{T_0}^T h(s)e^{A(\theta,s)} ds \\ &+ \int_{T_0}^T e^{-A(\theta,u)}dW_u \int_u^T h(s)e^{A(\theta,s)} ds = Z_{T,1} + Z_{T,2}. \end{aligned}$$

For the first term of $Z(T_0, T)$, $Z_{T,1}$, using (5.22) yields

$$\left| \int_{T_0}^T h(s)e^{A(\theta,s)} ds \right| \leq h \int_{T_0}^T e^{-K(\theta)s} ds \leq \frac{h}{K(\theta)} e^{-K(\theta)T_0} = hO_P(1).$$

Hence $\mathbb{E}(Z_{T,1})^2 = h^2O(1)$. For the second term of $Z(T_0, T)$, we write

$$\begin{aligned} \mathbb{E}(Z_{T,2})^2 &= \int_{T_0}^T e^{-2A(\theta,u)} du \left(\int_u^T h(s)e^{A(\theta,s)} ds \right)^2 \\ &= \int_{T_0}^T du \left(\int_u^T h(s)e^{A(\theta,s)-A(\theta,u)} ds \right)^2 \\ &\leq h^2 \int_{T_0}^T du \left(\int_u^T e^{-K(\theta)(s-u)} ds \right)^2 \leq (T - T_0)h^2 \frac{1}{K(\theta)^2}. \end{aligned}$$

Therefore, for all $T \geq T_0$, $\frac{1}{T}\mathbb{E}(Z(T_0, T))^2 \lesssim \frac{h^2}{T} + h^2$. Hence, $\lim_{T \rightarrow +\infty} \frac{1}{\sqrt{T}} \int_0^T g_s(\theta)h(s)ds = 0$. \square

Proof of Theorem 3.1. Recall that $H(\theta, s, x) = V(\alpha, x) - \Phi(\beta, x - x_s(\alpha))$. Thus, using (2.7),

$$dX_s = \varepsilon dW_s + H(\theta, s, X_s)ds - D(\theta, s, \varepsilon, X_s)ds. \tag{5.25}$$

The derivatives of H with respect to the parameters are given by:

$$\frac{\partial H}{\partial \alpha}(\theta, s, X_s) = \frac{\partial V}{\partial \alpha}(\alpha, X_s) + \frac{\partial \Phi}{\partial x}(\beta, X_s - x_s(\alpha)) \frac{\partial x_s}{\partial \alpha}(\alpha, s),$$

$$\begin{aligned}
\frac{\partial H}{\partial \beta}(\theta, s, X_s) &= -\frac{\partial \Phi}{\partial \beta}(\beta, X_s - x_s(\alpha)) \\
\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) &= \frac{\partial^2 V}{\partial \alpha^2}(\alpha, X_s) + \frac{\partial \Phi}{\partial x}(\beta, X_s - x_s(\alpha)) \frac{\partial^2 x_s}{\partial \alpha^2}(\alpha, s) \\
&\quad - \frac{\partial^2 \Phi}{\partial x^2}(\beta, X_s - x_s(\alpha)) \left(\frac{\partial x_s}{\partial \alpha}(\alpha, s) \right)^2 \\
\frac{\partial^2 H}{\partial \beta^2}(\theta, s, X_s) &= -\frac{\partial^2 \Phi}{\partial \beta^2}(\beta, X_s - x_s(\alpha)), \\
\frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s) &= \frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, X_s - x_s(\alpha)) \frac{\partial x_s}{\partial \alpha}(\alpha, s).
\end{aligned}$$

Note that for the convergence in distribution stated in Theorem 3.1, it is enough to prove that

$$\frac{\varepsilon}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\theta) = \ell(\alpha, \beta) \frac{\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) W_T}{\ell(\alpha)} \frac{1}{\sqrt{T}} + o_P(1), \quad (5.26)$$

$$\frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \beta}(\theta) = -\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s + o_P(1). \quad (5.27)$$

The bracket of the two stochastic integrals above is equal, up to a constant, to $T^{-1} \int_0^T g_s(\theta) ds$ and tends to 0 as T tends to infinity by Proposition 3.2.

Moreover, as $T^{-1} \int_0^T [g_s(\theta)]^2 ds$ tends to $[2\ell(\alpha, \beta)]^{-1}$, by the central limit theorem for martingales, $\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s$ converges in distribution to $\mathcal{N}(0, [2\ell(\alpha, \beta)]^{-1})$. The proof of (5.26)-(5.27) relies on the following Lemma:

Lemma 5.1. *Let $F(\theta, s, x)$ a continuous function on $\Theta \times \mathbb{R}^+ \times \mathbb{R}$, differentiable with respect to x and assume that there exist $C > 0$ and a nonnegative integer c such that, $\forall \theta \in \theta, \forall s \geq 0$,*

$$|F(\theta, s, x)| \leq C(1 + |x|^c) \quad \text{and} \quad \left| \frac{\partial F}{\partial x}(\theta, s, x) \right| \leq C(1 + |x|^c). \quad (5.28)$$

Then, for $T \geq 1, \varepsilon \leq 1, D(\theta, s, \varepsilon, x)$ given in (2.7), the following holds.

- (i) $\mathbb{E} \int_0^T (F(\theta, s, X_s) - F(\theta, s, x_s(\alpha)))^2 ds \leq C_1(\theta, F) T \varepsilon^2$.
- (ii) $\mathbb{E} \left| \int_0^T F(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds \right| \leq C_2(\theta, F) T \varepsilon^2$.
- (iii) If $\int_0^{+\infty} |F(\theta, s, x_s(\alpha))| ds < +\infty$, then

$$\mathbb{E} \left| \int_0^T F(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds \right| \leq C_3(\theta, F)(\varepsilon^2 + \varepsilon^3 T).$$

where the constants $C_i(\theta, F)$ only depend on F and θ .

Note that the functions $F(\theta, s, x) = H(\theta, s, x), \frac{\partial H}{\partial \alpha}(\theta, s, x), \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x)$ satisfy (5.28) under [S1] so that Lemma 5.1 holds for these functions.

We now start the proof of (5.26)-(5.27).

Derivative of the contrast with respect to α

Replacing dX_s by its expression, we get (see (2.7), (3.2), (3.3) and (5.25)):

$$\frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) = \frac{1}{\varepsilon^2} \left(\int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, X_s) dX_s - \int_0^T H(\theta, s, X_s) \frac{\partial H}{\partial \alpha}(\theta, s, X_s) ds \right).$$

Hence,

$$\frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) = \frac{1}{\varepsilon} \int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, X_s) dW_s - \frac{1}{\varepsilon^2} \int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds. \tag{5.29}$$

Let us define

$$\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) = \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \frac{dx^*}{d\alpha}(\alpha). \tag{5.30}$$

Then,

$$\begin{aligned} & \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \\ &= \frac{\partial V}{\partial \alpha}(\alpha, x_s(\alpha)) - \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \left(\frac{\partial x_s}{\partial \alpha}(\alpha, s) - \frac{dx^*}{d\alpha}(\alpha) \right). \end{aligned}$$

Therefore $\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha))$ is the limit of $\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha))$ as $s \rightarrow \infty$. Since we are in Case (1), (3.6) yields using (3.5) that $\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) = \ell(\alpha) \frac{dx^*}{d\alpha}(\alpha) \neq 0$ and

$$\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) = \ell(\alpha, \beta) \frac{dx^*}{d\alpha}(\alpha) = \frac{\ell(\alpha, \beta)}{\ell(\alpha)} \frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha)) \neq 0.$$

Consequently, we can write

$$\begin{aligned} \frac{\varepsilon}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) &= \frac{W_T}{\sqrt{T}} \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \\ &+ \frac{1}{\sqrt{T}} \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \right) dW_s \tag{5.31} \end{aligned}$$

$$+ \frac{1}{\sqrt{T}} \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) - \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right) dW_s \tag{5.32}$$

$$- \frac{1}{\varepsilon \sqrt{T}} \left(\int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds \right). \tag{5.33}$$

Using Proposition 3.1 and [S1], $\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha))$ converges exponentially fast to 0 so that $\int_0^{+\infty} \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \right)^2 ds < +\infty$.

To study (5.31), we split

$$\int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \right) dW_s$$

$$\begin{aligned}
&= \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) - \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right) dW_s \\
&\quad + \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, s, x^*(\alpha)) \right) dW_s.
\end{aligned}$$

Therefore, (5.31) is $O_P(1/\sqrt{T})$ and tends to 0. By Lemma 5.1 (i),

$$\frac{1}{T} \mathbb{E}_\theta \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) - \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right)^2 ds \lesssim \varepsilon^2,$$

so that (5.32) is $O_P(\varepsilon)$. Lemma 5.1 (ii) yields

$$\frac{1}{\varepsilon\sqrt{T}} \mathbb{E}_\theta \int_0^T \left| \frac{\partial H}{\partial \alpha}(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) \right| ds \lesssim \varepsilon\sqrt{T},$$

so that (5.33) is also $o_P(1)$ under the condition $\varepsilon\sqrt{T} \rightarrow 0$. So we find that

$$\frac{\varepsilon}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\theta) = \frac{W_T}{\sqrt{T}} \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) + o_P(1) = \ell(\alpha, \beta) \frac{\frac{\partial V}{\partial \alpha}(\alpha, x^*(\alpha))}{\ell(\alpha)} \frac{W_T}{\sqrt{T}} + o_P(1),$$

which gives (5.26).

Derivative of the contrast with respect to β

We have:

$$\begin{aligned}
\frac{\partial \Lambda_{\varepsilon, T}}{\partial \beta}(\theta) &= \frac{1}{\varepsilon^2} \int_0^T \frac{\partial H}{\partial \beta}(\theta, s, X_s) dX_s - \frac{1}{\varepsilon^2} \int_0^T H(\theta, s, X_s) \frac{\partial H}{\partial \beta}(\theta, s, X_s) ds \\
&= -\frac{1}{\varepsilon} \int_0^T \frac{\partial \Phi}{\partial \beta}(\beta, X_s - x_s(\alpha)) dW_s \tag{5.34}
\end{aligned}$$

$$\begin{aligned}
&+ \frac{1}{\varepsilon^2} \int_0^T \frac{\partial \Phi}{\partial \beta}(\beta, X_s - x_s(\alpha)) D(\theta, s, \varepsilon, X_s) ds \\
&:= T_1 + T_2. \tag{5.35}
\end{aligned}$$

Since $x \rightarrow \frac{\partial \Phi}{\partial \beta}(\beta, x)$ is an odd function, $\frac{\partial \Phi}{\partial \beta}(\beta, 0) = 0$, so

$$\frac{\partial \Phi}{\partial \beta}(\beta, x) = x \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) + x^2 \int_0^1 (1-u) \frac{\partial^3 \Phi}{\partial \beta \partial x^2}(\beta, ux) du. \tag{5.36}$$

Replacing x by $X_s - x_s(\alpha) = \varepsilon g_s(\theta) + \varepsilon^2 R_s^\varepsilon(\theta)$ yields that

$$T_1 = -\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \left(\int_0^T g_s(\theta) dW_s \right) - T_{11}, \text{ with}$$

$$T_{11} = \varepsilon \int_0^T R_s^\varepsilon(\theta) dW_s$$

$$- \varepsilon^{-1} \int_0^T (X_s - x_s(\alpha))^2 \int_0^1 (1-u) \frac{\partial^3 \Phi}{\partial \beta \partial x^2}(\beta, u(X_s - x_s(\alpha))) du dW_s.$$

Thus,

$$\frac{1}{\sqrt{T}} T_1 = - \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \left(\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s \right) - \frac{1}{\sqrt{T}} T_{11}.$$

We have, by Theorem 2.3,

$$\mathbb{E}_\theta \int_0^T (R_s^\varepsilon(\theta))^2 ds \leq 2\mathbb{E}_\theta \int_0^T (R_s^\varepsilon(\theta) - \mathbb{E}_\theta R_s^\varepsilon(\theta))^2 ds + 2 \int_0^T (\mathbb{E}_\theta R_s^\varepsilon(\theta))^2 ds \lesssim TO(1),$$

where $O(1)$ does not depend on T and ε . This implies

$$\frac{1}{T} \mathbb{E}_\theta \left(\varepsilon \int_0^T R_s^\varepsilon(\theta) dW_s \right)^2 = \frac{\varepsilon^2}{T} \mathbb{E}_\theta \int_0^T (R_s^\varepsilon(\theta))^2 ds \lesssim \frac{\varepsilon^2}{T} \times T = \varepsilon^2.$$

Then, using [S1],

$$\begin{aligned} \frac{1}{T} \mathbb{E}_\theta \left(\varepsilon^{-1} \int_0^T (X_s - x_s(\alpha))^2 \int_0^1 (1-u) \frac{\partial^3 \Phi}{\partial \beta \partial x^2}(\beta, u(X_s - x_s(\alpha))) du dW_s \right)^2 &\lesssim \\ \frac{1}{\varepsilon^2 T} \mathbb{E}_\theta \left(\int_0^T (X_s - x_s(\alpha))^4 (1 + (X_s - x_s(\alpha))^{2c}) ds \right) &\lesssim \frac{1}{\varepsilon^2 T} \times \varepsilon^4 T = \varepsilon^2. \end{aligned}$$

Therefore,

$$\frac{1}{\sqrt{T}} T_1 = - \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \left(\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s \right) + O_P(\varepsilon).$$

For T_2 , we have using (5.36),

$$\begin{aligned} T_2 &= \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \frac{1}{\varepsilon^2} \int_0^T (X_s - x_s(\alpha)) D(\theta, s, \varepsilon, X_s) ds + \\ &\frac{1}{\varepsilon^2} \int_0^T (X_s - x_s(\alpha))^2 \int_0^1 (1-u) \frac{\partial^3 \Phi}{\partial \beta \partial x^2}(\beta, u(X_s - x_s(\alpha))) du D(\theta, s, \varepsilon, X_s) ds. \end{aligned}$$

We split $D(\theta, s, \varepsilon, X_s) = \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) + D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)$ and use Corollary 2.1 and Theorem 2.2.

The main term of $|T_2|/\sqrt{T}$ is $\frac{1}{\varepsilon^2 \sqrt{T}} \left| \int_0^T (X_s - x_s(\alpha)) \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) ds \right|$. Taking the expectation of this term yields

$$\begin{aligned} \mathbb{E}_\theta \frac{1}{\varepsilon^2 \sqrt{T}} \left| \int_0^T (X_s - x_s(\alpha)) \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) ds \right| &\leq \\ \frac{1}{\sqrt{T}} \varepsilon^{-2} \sup_{s, \varepsilon} |\mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| \int_0^T \mathbb{E}_\theta |X_s - x_s(\alpha)| ds &\lesssim \frac{1}{\sqrt{T}} O(1) \varepsilon T. \end{aligned}$$

Finally, $T_2/\sqrt{T} = \varepsilon\sqrt{T}O_P(1)$.

Therefore,

$$\frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon,T}}{\partial \beta}(\theta) = -\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \left(\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s \right) + o_P(1).$$

This yields (5.27). Hence the first part of Theorem 3.1, is proved.

It remains to study the limit of

$$D_{\varepsilon,T}^{(1)} \mathcal{J}_{\varepsilon,T}(\theta) D_{\varepsilon,T}^{(1)} = - \begin{pmatrix} \frac{\varepsilon^2}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \alpha^2}(\theta) & \frac{\varepsilon}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \beta \partial \alpha}(\theta) \\ \frac{\varepsilon}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \beta \partial \alpha}(\theta) & \frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \beta^2}(\theta) \end{pmatrix}.$$

We have:

$$\begin{aligned} \frac{\varepsilon^2}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \alpha^2}(\theta) &= \frac{1}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) dX_s - \frac{1}{T} \int_0^T H(\theta, s, X_s) \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) ds \\ &\quad - \frac{1}{T} \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) \right)^2 ds = T_1 + T_2 + T_3 \quad \text{with} \end{aligned}$$

$T_1 = \frac{\varepsilon}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) dW_s$; $T_2 = -\frac{1}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds$; $T_3 = -\frac{1}{T} \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) \right)^2 ds$. For T_1 , we write

$$\begin{aligned} T_1 &= \frac{\varepsilon}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha)) dW_s \\ &\quad + \frac{\varepsilon}{T} \int_0^T \left(\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) - \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha)) \right) dW_s. \end{aligned}$$

Noting that $\frac{\partial^2 \Phi}{\partial x^2}(\beta, x)$ is odd,

$$\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha)) = \frac{\partial^2 V}{\partial \alpha^2}(\alpha, x_s(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \frac{\partial^2 x_s}{\partial \alpha^2}(\alpha, s).$$

This function is uniformly bounded thanks to Proposition 3.1. Therefore, using Lemma 5.1,

$$\mathbb{E}_\theta T_1^2 \lesssim \frac{\varepsilon^2}{T^2} \times (T + \varepsilon^2 T) = \frac{\varepsilon^2}{T} (1 + \varepsilon^2) = o(1).$$

By Lemma 5.1, $\mathbb{E}_\theta(|T_2|) \lesssim \frac{1}{T} \times \varepsilon^2 T = \varepsilon^2$.

For the last and main term T_3 , we write (see (5.30)):

$$\begin{aligned} T_3 &= -\frac{1}{T} \int_0^T \left[\left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) \right)^2 - \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right)^2 \right] ds \\ &\quad - \frac{1}{T} \int_0^T \left[\left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right)^2 - \left(\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \right)^2 \right] ds - \left(\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \right)^2. \end{aligned}$$

For the first term, we use Lemma 5.1 to prove that it is $o_P(1)$. For the second term, we use that $(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)))^2$ converges to $(\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)))^2$ with exponential rate and this implies that this second term is $o(1)$. Hence T_3 tends to $-\left(\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha))\right)^2$.

Joining these results, we have proved that $\frac{\varepsilon^2}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha^2}(\theta)$ tends to $-\left(\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha))\right)^2$.

Let us study $\frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta)$. Using that $\frac{\partial^2 H}{\partial \beta^2}(\theta, s, X_s) = -\frac{\partial^2 \Phi}{\partial \beta^2}(\beta, X_s - x_s(\alpha))$ yields

$$\begin{aligned} \frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta) &= -\frac{1}{\varepsilon T} \int_0^T \frac{\partial^2 \Phi}{\partial \beta^2}(\beta, X_s - x_s(\alpha)) dW_s \\ &\quad + \frac{1}{\varepsilon^2 T} \int_0^T \frac{\partial^2 \Phi}{\partial \beta^2}(\beta, X_s - x_s(\alpha)) D(\theta, s, \varepsilon, X_s) ds \\ &\quad - \frac{1}{\varepsilon^2 T} \int_0^T \left(\frac{\partial \Phi}{\partial \beta}(\theta, X_s - x_s(\alpha)) \right)^2 ds = S_1 + S_2 + S_3. \end{aligned}$$

The following relation is analogous to (5.36):

$$\frac{\partial^2 \Phi}{\partial \beta^2}(\beta, x) = x \frac{\partial^3 \Phi}{\partial \beta^2 \partial x}(\beta, 0) + x^2 \int_0^1 (1-u) \frac{\partial^4 \Phi}{\partial \beta^2 \partial x^2}(\beta, ux) du. \tag{5.37}$$

Substituting x by $X_s - x_s(\alpha) = \varepsilon g_s(\theta) + \varepsilon^2 R^\varepsilon(s)$, we get that the main term of S_1 is

$$S_{11} = -\frac{1}{\varepsilon T} \int_0^T (X_s - x_s(\alpha)) \frac{\partial^3 \Phi}{\partial \beta^2 \partial x}(\beta, 0) dW_s = O_P\left(\frac{1}{\sqrt{T}}\right),$$

as, using Proposition 3.2,(i), $\mathbb{E}_\theta S_{11}^2 = O_P(1/T)$. For S_2 , we split as previously $D(\theta, s, \varepsilon, X_s) = \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) + D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)$ and find that the main term of S_2 is

$$S_{22} = \frac{1}{\varepsilon^2 T} \int_0^T (X_s - x_s(\alpha)) \frac{\partial^3 \Phi}{\partial \beta^2 \partial x}(\beta, 0) \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) ds,$$

where $|\mathbb{E}_\theta S_{22}| \lesssim \varepsilon$ using Corollary 2.1 and Theorem 2.2.

The limit is obtained by S_3 whose main term is (see (5.36))

$$\begin{aligned} S_{33} &= -\frac{1}{\varepsilon^2 T} \int_0^T \left((X_s - x_s(\alpha)) \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \right)^2 ds \\ &= -\left(\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \right)^2 \frac{1}{T} \int_0^T g_s^2(\theta) ds + o_P(1). \end{aligned}$$

Therefore, Proposition 3.2, (i) yields that S_{33} tends to $-\left(\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0)\right)^2 / (2(\ell(\alpha, \beta)))$. Joining these results, we get that the same holds for $\frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta)$.

It remains to study the off diagonal term $(\varepsilon/T) \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \alpha \partial \beta}(\theta)$. We have

$$\begin{aligned} \frac{\varepsilon}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \alpha \partial \beta}(\theta) &= \frac{1}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s) dW_s - \frac{1}{\varepsilon T} \int_0^T D(\theta, s, \varepsilon, X_s) \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s) ds \\ &\quad - \frac{1}{\varepsilon T} \left(\int_0^T \frac{\partial H}{\partial \beta}(\theta, s, X_s) \frac{\partial H}{\partial \alpha}(\theta, s, X_s) ds \right) = T_1 + T_2 + T_3 \end{aligned}$$

where $\frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s)$, $\frac{\partial H}{\partial \beta}(\theta, s, X_s)$ are detailed above. As before, the main term of T_1 is $\frac{1}{T} \int_0^T \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, x_s(\alpha)) dW_s = \frac{1}{T} \int_0^T \frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, 0) \frac{\partial x_s}{\partial \alpha}(\alpha, s) dW_s$. Since $\frac{\partial x_s}{\partial \alpha}(\alpha, s)$ is uniformly bounded, $\mathbb{E}_\theta T_1^2 = \frac{1}{T} O(1)$ and $T_1 = O_P(\frac{1}{\sqrt{T}})$.

For T_2 , by Lemma 5.1, $|\mathbb{E}_\theta T_2| \lesssim \frac{1}{\varepsilon T} \varepsilon^2 T = \varepsilon$.

For T_3 , we have, using (5.36),

$$T_3 = -\frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \frac{1}{T} \int_0^T g_s(\theta) \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) ds + o_P(1).$$

Now, setting $h(s) = \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) - \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha))$, we have

$$\frac{1}{T} \int_0^T g_s(\theta) \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) ds = \frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) \frac{1}{T} \int_0^T g_s(\theta) ds + \frac{1}{T} \int_0^T g_s(\theta) h(s) ds.$$

Since $x_s(\alpha) \rightarrow x^*(\alpha)$, $h(s) \rightarrow 0$, Proposition 3.2 yields that both terms above converge to 0.

To conclude, we have obtained

$$\frac{\varepsilon}{T} \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \beta \partial \alpha}(\theta) = o_P(1).$$

The proof of Theorem 3.1 is now complete. \square

Proof of Theorem 3.2. Let us set

$$h(\theta, s) = \frac{\partial V}{\partial \alpha}(\alpha, x_s(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \frac{\partial x_s}{\partial \alpha}(\alpha, s) = \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)). \quad (5.38)$$

Here, for the convergence in distribution, it is enough to prove

$$\varepsilon \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\theta) = \int_0^T h(\theta, s) dW_s + o_P(1) \quad (5.39)$$

$$\frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon,T}}{\partial \beta}(\theta) = \frac{\partial^2 \Phi}{\partial \beta \partial x}(\beta, 0) \frac{1}{\sqrt{T}} \int_0^T g_s(\theta) dW_s + o_P(1). \quad (5.40)$$

Indeed, the bracket of the two stochastic integrals is equal, up to a constant, to $\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) h(\theta, s) ds$.

We are in Case (2): using (3.6), it corresponds to $\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) = 0$. Therefore, by Proposition 3.1, $h(\theta, s)$ converges exponentially fast to $\frac{\partial H}{\partial \alpha}(\theta, x^*(\alpha)) = 0$

and $\int_0^{+\infty} \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha))\right)^2 ds < +\infty$. Proposition 3.2 yields that $\frac{1}{\sqrt{T}} \int_0^T g_s(\theta)h(\theta, s)ds$ tends to 0.

Let us prove (5.39). We now have (see (5.29)):

$$\begin{aligned} \varepsilon \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\theta) &= \int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha))dW_s + \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) - \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right) dW_s \\ &\quad - \frac{1}{\varepsilon} \int_0^T \frac{\partial H}{\partial \alpha}(\theta, s, X_s)D(\theta, s, \varepsilon, X_s)ds = T_1 + T_2 + T_3. \end{aligned} \tag{5.41}$$

Since $\mathbb{E}_\theta(T_1^2) < \infty$, $T_1 \rightarrow \int_0^{+\infty} \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha))dW_s$ as $T \rightarrow \infty$.

By Lemma 5.1, $\mathbb{E}_\theta(T_2^2) \lesssim \varepsilon^2 T = o(1)$ under the condition $\varepsilon\sqrt{T} \rightarrow 0$.

As $\int_0^{+\infty} \left| \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right| ds < +\infty$, Lemma 5.1 (iii) yields that $\mathbb{E}|T_3| \lesssim \varepsilon + \varepsilon^2 T = o(1)$.

This achieves the proof of (5.39).

The study of $\frac{1}{\sqrt{T}} \frac{\partial \Lambda_{\varepsilon, T}}{\partial \beta}(\theta)$ is similar to its study in Theorem 3.1. The proof of (5.40) is complete.

Now we study the limit of $D_{\varepsilon, T}^{(2)} \mathcal{J}_{\varepsilon, T}(\theta) D_{\varepsilon, T}^{(2)} = - \begin{pmatrix} \varepsilon^2 \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha^2}(\theta) & \frac{\varepsilon}{\sqrt{T}} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta \partial \alpha}(\theta) \\ \frac{\varepsilon}{\sqrt{T}} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta \partial \alpha}(\theta) & \frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta) \end{pmatrix}$.

We have

$$\begin{aligned} \varepsilon^2 \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha^2}(\theta) &= \varepsilon \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s)dW_s - \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s)D(\theta, s, \varepsilon, X_s)ds \\ &\quad - \int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s) \right)^2 ds = T_1 + T_2 + T_3. \end{aligned}$$

For the first term, we write

$$T_1 = \varepsilon \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha))dW_s + \varepsilon \int_0^T \left(\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s) - \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha)) \right) dW_s.$$

We have that $\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha)) = \frac{\partial^2 V}{\partial \alpha^2}(\alpha, x_s(\alpha)) + \frac{\partial \Phi}{\partial x}(\beta, 0) \frac{\partial^2 x_s}{\partial \alpha^2}(\alpha, s)$, which is uniformly bounded on \mathbb{R}^+ . Thus,

$$\mathbb{E}_\theta \left(\varepsilon \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x_s(\alpha))dW_s \right)^2 \lesssim \varepsilon^2 T = o(1).$$

The second term of T_1 is ruled by Lemma 5.1 (i) and is $\varepsilon o_P(1)$. Next, $\mathbb{E}_\theta|T_2| \lesssim T\varepsilon^2$ by Lemma 5.1 (ii).

$$\mathbb{E}_\theta \left| \int_0^T \frac{\partial^2 H}{\partial \alpha^2}(\theta, s, X_s)D(\theta, s, \varepsilon, X_s)ds \right| \lesssim \varepsilon^2 T.$$

Finally, we can check, using Lemma 5.1 (i), that the main term of T_3 is, using (5.38), $\int_0^T h^2(\theta, s)ds$, where $h(\theta, s)$ converges exponentially fast to 0. Therefore, $\int_0^T \left(\frac{\partial H}{\partial \alpha}(\theta, s, X_s)\right)^2 ds \rightarrow \int_0^{+\infty} h^2(\theta, s)ds < +\infty$, so that

$$\varepsilon^2 \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha^2}(\theta) \rightarrow - \int_0^{+\infty} \left(\frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) \right)^2 ds.$$

The study of $\frac{1}{T} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta^2}(\theta)$ is the same as for Theorem 3.1. It remains to study

$$\begin{aligned} \frac{\varepsilon}{\sqrt{T}} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha \partial \beta}(\theta) &= \frac{1}{\sqrt{T}} \int_0^T \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s) dW_s \\ &\quad - \frac{1}{\varepsilon \sqrt{T}} \int_0^T D(\theta, s, \varepsilon, X_s) \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, X_s) ds \\ &\quad - \frac{1}{\varepsilon \sqrt{T}} \left(\int_0^T \frac{\partial H}{\partial \beta}(\theta, s, X_s) \frac{\partial H}{\partial \alpha}(\theta, s, X_s) ds \right) = T_1 + T_2 + T_3. \end{aligned}$$

The proof is essentially analogous to the study of $\frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \alpha \partial \beta}(\theta)$ in the previous theorem. We point out the differences.

The main term of T_1 is $\frac{1}{\sqrt{T}} \int_0^T \frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, 0) \frac{\partial x_s}{\partial \alpha}(\alpha, s) dW_s$.

We are in Case (2) so that $\frac{\partial x_s}{\partial \alpha}(\alpha, s)$ converges exponentially fast to $\frac{dx^*}{d\alpha}(\alpha) = 0$. Therefore, $\int_0^{+\infty} \left(\frac{\partial x_s}{\partial \alpha}(\alpha, s)\right)^2 ds < +\infty$. Consequently, $T_1 = \frac{1}{\sqrt{T}} O_P(1)$.

For T_2 , the main term is $\frac{1}{\varepsilon \sqrt{T}} \int_0^T D(\theta, s, \varepsilon, X_s) \frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, x_s(\alpha)) ds$.

Using that $\frac{\partial^2 H}{\partial \alpha \partial \beta}(\theta, s, x_s(\alpha)) = \frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, 0) \frac{\partial x_s}{\partial \alpha}(\alpha, s)$ is integrable, we get by Lemma 5.1 (iii),

$$\mathbb{E}_\theta[T_2] \lesssim \frac{1}{\varepsilon \sqrt{T}} (\varepsilon^2 + \varepsilon^3 T) = o(1).$$

It remains to study T_3 . Using (5.36) and (5.38),

$$\begin{aligned} T_3 &= -\frac{1}{\varepsilon \sqrt{T}} \left(\int_0^T \frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, 0) (X_s - x_s(\alpha)) \frac{\partial H}{\partial \alpha}(\theta, s, x_s(\alpha)) ds \right) + o_P(1) \\ &= -\frac{\partial^2 \Phi}{\partial x \partial \beta}(\beta, 0) \frac{1}{\sqrt{T}} \int_0^T g_s(\theta) h(\theta, s) + o_P(1). \end{aligned}$$

As $h(\theta, s) \rightarrow 0$ as $s \rightarrow \infty$, $\frac{1}{\sqrt{T}} \int_0^T g_s(\theta) h(\theta, s) ds = o_P(1)$ by Proposition 3.2.

Hence, $\frac{\varepsilon}{\sqrt{T}} \frac{\partial^2 \Lambda_{\varepsilon, T}}{\partial \beta \partial \alpha}(\theta) = o_P(1)$. So the proof of Theorem 3.2 is complete. \square

Proof of Lemma 5.1.

Proof of (i) A Taylor expansion yields:

$$F(\theta, s, X_s) - F(\theta, s, x_s(\alpha)) = (X_s - x_s(\alpha)) \int_0^1 \frac{\partial F}{\partial x}(\alpha, x_s(\alpha) + u(X_s - x_s(\alpha))) du.$$

Hence

$$\begin{aligned} & (F(\theta, s, X_s) - F(\theta, s, x_s(\alpha)))^2 \\ & \leq 3C^2 \varepsilon^2 \left(\frac{(X_s - x_s(\alpha))^2}{\varepsilon^2} (1 + \sup_{s \geq 0} |x_s(\alpha)|^{2c}) + \varepsilon^{2c} \frac{(X_s - x_s(\alpha))^{2+2c}}{\varepsilon^{2+2c}} \right). \end{aligned}$$

By Theorem 2.2, since $s \rightarrow x_s(\alpha)$ is uniformly bounded on \mathbb{R}^+ by $B(\alpha)$, we get

$$\begin{aligned} & \mathbb{E}_\theta \int_0^T (F(\theta, s, X_s) - F(\theta, s, x_s(\alpha)))^2 ds \\ & \leq 3C^2 \varepsilon^2 T (\delta(\alpha, 1)(1 + B^{2c}(\alpha)) + \varepsilon^{2c} \delta(\alpha, 1 + c)) = C_1(\alpha, F) \varepsilon^2 T. \end{aligned}$$

This achieves the proof of (i).

Proof of (ii) For the second inequality, we split

$$\mathbb{E}_\theta \int_0^T F(\theta, s, X_s) D(\theta, s, \varepsilon, X_s) ds = A_1 + A_2 + A_3 + A_4, \quad \text{with}$$

$$A_1 = \mathbb{E}_\theta \int_0^T F(\theta, s, x_s(\alpha)) \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) ds,$$

$$A_2 = \mathbb{E}_\theta \int_0^T F(\theta, s, x_s(\alpha)) (D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)) ds,$$

$$A_3 = \mathbb{E}_\theta \int_0^T (F(\theta, s, X_s) - F(\theta, s, x_s(\alpha))) \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s) ds,$$

$$A_4 = \mathbb{E}_\theta \int_0^T (F(\theta, s, X_s) - F(\theta, s, x_s(\alpha))) (D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)) ds.$$

Since $|x_s(\alpha)| \leq B(\alpha)$, $F(\theta, s, x) \leq C(1 + |x|^c)$, we get that $|F(\theta, s, x_s(\alpha))| \leq C(1 + B^c(\alpha)) = C(\alpha)$. Therefore, using Corollary 2.1

$$\begin{aligned} |A_1| & \leq \sup_{s \geq 0} |\mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| \int_0^T |F(\theta, s, x_s(\alpha))| ds \\ & \leq \varepsilon^2 T \times \left[\sup_{s \geq 0} |F(\theta, s, x_s(\alpha))| \right] \leq C(\alpha) \varepsilon^2 T \end{aligned}$$

$$|A_2| \leq \int_0^T |F(\theta, s, x_s(\alpha))| \mathbb{E}_\theta |D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| ds$$

$$\begin{aligned} &\leq \varepsilon^3 \left[\int_0^T |F(\theta, s, x_s(\alpha))| \left[\mathbb{E}_\theta \varepsilon^{-6} |D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)|^2 \right]^{1/2} ds \right. \\ &\lesssim \varepsilon^3 T. \end{aligned}$$

For A_3 , we have using (i),

$$\begin{aligned} |A_3| &\leq \sup_{s \geq 0} |\mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| \times \mathbb{E}_\theta \left[T \int_0^T |F(\theta, s, X_s) - F(\theta, s, x_s(\alpha))|^2 \right]^{1/2} ds \\ &\lesssim \varepsilon^2 \sqrt{T} \times (\varepsilon^2 T)^{1/2} \lesssim \varepsilon^3 T. \end{aligned}$$

For A_4 , we write:

$$|A_4| \leq \int_0^T \mathbb{E}_\theta [|F(\theta, s, X_s) - F(\theta, s, x_s(\alpha))| |D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)|] ds.$$

We apply the Cauchy-Schwarz inequality.

$$\text{Using (i), } \mathbb{E}_\theta |F(\theta, s, X_s) - F(\theta, s, x_s(\alpha))|^2 \lesssim [\mathbb{E}_\theta (X_s - x_s(\alpha))^2]^{1/2} \leq \varepsilon.$$

Therefore, by Theorem 2.2 and Corollary 2.1, $|A_4| \lesssim \varepsilon^4 T$.

Finally, joining these inequalities yields (ii).

Proof of (iii) Since $\int_0^\infty |F(\theta, s, x_s(\alpha))| ds < \infty$, we bound differently A_1 and A_2 .

$$|A_1| \leq \sup_{s \geq 0} |\mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| \int_0^T |F(\theta, s, x_s(\alpha))| ds \lesssim \varepsilon^2.$$

Analogously, for A_2 ,

$$\begin{aligned} |A_2| &\leq \left[\int_0^T |F(\theta, s, x_s(\alpha))| \mathbb{E}_\theta |D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)| ds \right. \\ &\leq \sup_{s \geq 0} \left(\mathbb{E}_\theta [|D(\theta, s, \varepsilon, X_s) - \mathbb{E}_\theta D(\theta, s, \varepsilon, X_s)|^2] \right)^{1/2} \int_0^{+\infty} |F(\theta, s, x_s(\alpha))| ds \\ &\lesssim \varepsilon^3. \end{aligned}$$

The terms A_3, A_4 are bounded as previously. Thus $|A_1 + A_2 + A_3 + A_4| \lesssim \varepsilon^2 + \varepsilon^3 T$.

It remains to look at the functions $H(\theta, s, x)$, $\frac{\partial H}{\partial \alpha}(\theta, s, x)$, $\frac{\partial^2 H}{\partial \alpha^2}(\theta, s, x)$. Using [S1]-[S2], as $B = \sup_{\alpha, t} |x_t(\alpha)| < +\infty$, we easily check (5.28) for $H(\theta, s, x)$. By [S2] and Proposition 3.1, $\sup_{\alpha, t} |\frac{\partial x_t}{\partial \alpha}(\alpha, t)| < +\infty$, $\sup_{\alpha, t} |\frac{\partial^2 x_t}{\partial \alpha^2}(\alpha, t)| < +\infty$. Therefore, we can check that (5.28) holds for the two other functions. \square

Proof of Lemma 3.1. We have to study under \mathbb{P}_{θ_0} :

$$\begin{aligned} \varepsilon^2 \Lambda_{\varepsilon, T}(\alpha, \beta) &= \int_0^T H(\theta, s, X_s) [(V(\alpha_0, X_s) ds - b(\theta_0, s, X_s)) ds + \varepsilon dW_s] \\ &\quad - \frac{1}{2} \int_0^T H^2(\theta, s, X_s) ds = -\frac{1}{2} \int_0^T (H(\theta, s, X_s) - V(\alpha_0, X_s))^2 ds \\ &\quad + \frac{1}{2} \int_0^T V^2(\alpha_0, X_s) ds + \sum_{i=1}^4 T_i \end{aligned}$$

where, using (3.3), (2.7)

$$\begin{aligned} T_1 &= \varepsilon \int_0^T H(\theta, s, X_s) dW_s; \quad T_2 = \int_0^T H(\theta, s, X_s) D(\theta_0, s, \varepsilon, X_s) ds \\ T_3 &= - \int_0^T (H(\theta, s, X_s) - H(\theta, s, x_s(\alpha_0))) \Phi(\beta_0, X_s - x_s(\alpha_0)) ds \\ T_4 &= - \int_0^T H(\theta, s, x_s(\alpha_0)) \Phi(\beta_0, X_s - x_s(\alpha_0)) ds. \end{aligned}$$

Let us consider the first term of $\varepsilon^2 \Lambda_{\varepsilon, T}(\alpha, \beta)$. It satisfies, using Lemma 5.1 (i) that, under the condition $\varepsilon \sqrt{T} \rightarrow 0$,

$$\int_0^T (H(\theta, s, X_s) - V(\alpha_0, X_s))^2 ds = \int_0^T (H(\theta, s, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)))^2 ds + o_P(1).$$

Now, define the limit of its integrand term as $s \rightarrow \infty$,

$$h^*(\alpha, \alpha_0, \beta) = V(\alpha, x^*(\alpha_0)) - \Phi(\beta, x^*(\alpha_0) - x^*(\alpha)). \tag{5.42}$$

The two cases pointed out in Section 3.4 occur here.

Case (1): $\forall \beta, h^*(\alpha, \alpha_0, \beta) \neq 0$ and

$$\frac{1}{T} \int_0^T (H(\theta, s, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)))^2 ds \rightarrow (h^*(\alpha, \alpha_0, \beta))^2.$$

Case (2): $\forall \beta, h^*(\alpha, \alpha_0, \beta) = 0$; $\int_0^\infty (H(\theta, s, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)))^2 ds < \infty$.

The second term satisfies $\int_0^T V^2(\alpha_0, X_s) ds = \int_0^T V^2(\alpha_0, x_s(\alpha_0)) ds + o_P(1)$ in both cases. This integral converges, as $T \rightarrow \infty$ to $\int_0^\infty V^2(\alpha_0, x_s(\alpha_0)) ds < \infty$.

Consider now the remainder terms T_i of $\varepsilon^2 \Lambda_{\varepsilon, T}(\alpha, \beta)$.

We have $\mathbb{E}_{\theta_0} T_1^2 = \varepsilon^2 \mathbb{E}_{\theta_0} \int_0^T [V(\alpha, X_s) - \Phi(\beta, X_s - x_s(\alpha))]^2 ds$. Using Lemma 5.1 (i) and similar tools detailed in the proof yields that $\mathbb{E}_{\theta_0} T_1^2 \lesssim \varepsilon^2 T$. Therefore, under the condition $\varepsilon \sqrt{T} = o(1)$, we find that $T_1 = o_P(1), T_2 = o_P(1)$. For T_3 , applying Lemma 5.1 (i) yields that

$$\mathbb{E}_{\theta_0} \int_0^T [H(\theta, s, X_s) - H(\theta, s, x_s(\alpha_0))] \Phi(\beta_0, X_s - x_s(\alpha_0))^2 ds \lesssim T \varepsilon^2$$

and $\mathbb{E}_{\theta_0} |T_3| = \varepsilon \sqrt{T} = o_P(1)$.

For T_4 , using Theorem 2.3, $\Phi(\beta_0, X_s - x_s(\alpha_0)) = \frac{\partial \Phi}{\partial x}(\beta, 0)(\varepsilon g_s(\theta_0) + \varepsilon^2 R_s^\varepsilon(\theta_0))$.

Therefore $T_4 = \varepsilon \frac{\partial \Phi}{\partial x}(\beta, 0) \int_0^T H(\theta, s, x_s(\alpha_0)) g_s(\theta_0) ds + o_P(1)$.

The limit, as $s \rightarrow \infty$ of $H(\theta, s, x_s(\alpha_0))$ is $h^*(\alpha_0, \alpha, \beta)$ defined in (5.42). Therefore, we have to study $\frac{T_4}{T}$ in Case (1) and T_4 in Case (2). We have

$$T_4 = \varepsilon \frac{\partial \Phi}{\partial x}(\beta, 0) [h^*(\alpha, \alpha_0, \beta) \int_0^T g_s(\theta_0) ds$$

$$+ \int_0^T (H(\theta, s, x_s(\alpha_0)) - h^*(\alpha, \alpha_0, \beta))g_s(\theta_0)ds].$$

Therefore, in Case **(1)**, by Proposition 3.2 (ii) and (iii), $\frac{T_4}{T} = \varepsilon o_P(1)$.

In Case **(2)**, for all β , $h^*(\alpha_0, \alpha, \beta) = 0$ and Proposition 3.2 (iii) yields that $T_4 = \varepsilon\sqrt{T}o_P(1) = o_P(1)$.

Consider now $\varepsilon^2\Lambda_{\varepsilon,T}(\alpha_0, \beta)$. Noting that $h^*(\alpha_0, \alpha_0, \beta) = 0$, $T_4 = o_P(1)$. Using that $\int_0^T [\Phi(\beta, X_s - x_s(\alpha))]^2 ds = O_P(\varepsilon^2 T)$, we get

$$\begin{aligned} \varepsilon^2\Lambda_{\varepsilon,T}(\alpha_0, \beta) &= -\frac{1}{2} \int_0^T [\Phi(\beta, X_s - x_s(\alpha))]^2 ds + \frac{1}{2} \int_0^T V^2(\alpha_0, X_s) ds + o_P(1) \\ &= \frac{1}{2} \int_0^T V^2(\alpha_0, X_s) ds + o_P(1). \end{aligned}$$

Joining these results yields that, using (5.42),

$$\text{Case (1): } \frac{\varepsilon^2}{T}(\Lambda_{\varepsilon,T}(\alpha, \beta) - \Lambda_{\varepsilon,T}(\alpha_0, \beta)) \rightarrow -\frac{1}{2}(h^*(\alpha, \alpha_0, \beta))^2 = \Lambda_1^{(1)}(\alpha, \alpha_0, \beta).$$

$$\begin{aligned} \text{Case (2): } \varepsilon^2(\Lambda_{\varepsilon,T}(\alpha, \beta) - \Lambda_{\varepsilon,T}(\alpha_0, \beta)) &\rightarrow \Lambda_1^{(2)}(\alpha, \alpha_0, \beta), \text{ with} \\ \Lambda_1^{(2)}(\alpha, \alpha_0, \beta) &= -\frac{1}{2} \int_0^{+\infty} [V(\alpha, x_s(\alpha_0)) - V(\alpha_0, x_s(\alpha_0)) - \Phi(\beta, x_s(\alpha_0) - x_s(\alpha))]^2 ds. \end{aligned}$$

The uniformity of the convergence is obtained using that $\Theta_\alpha, \Theta_\beta$ are compact sets, Assumptions [S1], [S2] and Comments 2.1, 2.2.

Finally, it remains to study $\frac{1}{T}(\Lambda_{\varepsilon,T}(\alpha_0, \beta) - \Lambda_{\varepsilon,T}(\alpha_0, \beta_0))$.

$$\begin{aligned} \varepsilon^2\Lambda_{\varepsilon,T}(\alpha_0, \beta) &= \int_0^T H(\alpha_0, \beta, s, X_s) [(H(\theta_0, s, X_s) - D(\theta_0, s, \varepsilon, X_s))ds + \varepsilon dW_s] \\ &\quad - \frac{1}{2} \int_0^T H^2(\alpha_0, \beta, s, X_s) ds = -\frac{1}{2} \int_0^T (H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s))^2 ds \\ &\quad + \varepsilon \int_0^T H(\alpha_0, \beta, s, X_s) dW_s - \int_0^T H(\alpha_0, \beta, s, X_s) D(\theta_0, s, \varepsilon, X_s) ds. \end{aligned}$$

Now, $H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s) = -(\Phi(\beta, X_s - x_s(\alpha_0)) - \Phi(\beta_0, X_s - x_s(\alpha_0)))$,

$$\Phi(\beta, X_s - x_s(\alpha_0)) = \frac{\partial\Phi}{\partial x}(\beta, 0)(X_s - x_s(\alpha_0)) + O_P(\varepsilon) = \varepsilon \frac{\partial\Phi}{\partial x}(\beta, 0)g_s(\theta_0) + \varepsilon^2 O_P(1).$$

Therefore,

$$\begin{aligned} \frac{1}{T}(\Lambda_{\varepsilon,T}(\alpha_0, \beta) - \Lambda_{\varepsilon,T}(\alpha_0, \beta_0)) &= -\frac{1}{2T} \int_0^T \left(\frac{\partial\Phi}{\partial x}(\beta, 0) - \frac{\partial\Phi}{\partial x}(\beta_0, 0) \right)^2 g_s^2(\theta_0) ds \\ &\quad + T_1 + T_2 + \varepsilon O_P(1), \end{aligned}$$

where $T_1 = \frac{1}{\varepsilon T} \int_0^T (H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s)) dW_s$,
 $T_2 = -\frac{1}{\varepsilon^2 T} \int_0^T (H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s)) D(\theta_0, s, \varepsilon, X_s) ds$.
 For T_1 , we have, using Theorem 2.2,

$$\begin{aligned} \mathbb{E}_{\theta_0} T_1^2 &= \frac{1}{\varepsilon^2 T^2} \mathbb{E}_{\theta_0} \int_0^T (H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s))^2 ds \\ &\lesssim \frac{1}{\varepsilon^2 T^2} T \sup \mathbb{E}_{\theta_0} ((X_s - x_s(\alpha_0))^2) \lesssim \frac{1}{T}. \end{aligned}$$

Therefore $T_1 = o_P(1)$.

For T_2 , set $F(X_s) = H(\alpha_0, \beta, s, X_s) - H(\theta_0, s, X_s)$.

Then, splitting $D(\theta_0, s, \varepsilon, X_s)$ as in the proof of Lemma 5.1,

$$\begin{aligned} \int_0^T F(X_s) D(\theta_0, s, \varepsilon, X_s) ds &= \int_0^T F(X_s) \mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s) \\ &+ \int_0^T F(X_s) (D(\theta_0, s, \varepsilon, X_s) - \mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s)) ds. \end{aligned}$$

Using that $\mathbb{E}_{\theta_0} \int_0^T |F(X_s)| ds \leq \sqrt{T} (\mathbb{E}_{\theta_0} \int_0^T F^2(X_s) ds)^{1/2} \leq \varepsilon T$, we get

$$\begin{aligned} \mathbb{E}_{\theta_0} \left| \int_0^T F(X_s) \mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s) ds \right| &\leq \sup_s |\mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s)|, \\ \mathbb{E}_{\theta_0} \int_0^T |F(X_s)| ds &\lesssim \varepsilon^3 T. \end{aligned}$$

Now, we have that

$$\mathbb{E}_{\theta_0} |F(X_s) (D(\theta_0, s, \varepsilon, X_s) - \mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s))| \leq \varepsilon^3 [\mathbb{E}_{\theta_0} (X_s - x_s(\alpha_0))^2]^{1/2} O(1).$$

Hence, $\mathbb{E}_{\theta_0} \left| \int_0^T F(X_s) (D(\theta_0, s, \varepsilon, X_s) - \mathbb{E}_{\theta_0} D(\theta_0, s, \varepsilon, X_s)) ds \right| \leq \varepsilon^4 T$.

These two inequalities yield that $T_2 = o_P(1)$ and finally, as $T \rightarrow \infty$,

$$\begin{aligned} \frac{1}{T} (\Lambda_{\varepsilon, T}(\alpha_0, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta_0)) &= -\frac{1}{2T} \int_0^T \left[\frac{\partial \Phi}{\partial x}(\beta, 0) - \frac{\partial \Phi}{\partial x}(\beta_0, 0) \right]^2 g_s^2(\theta_0) ds + o_P(1) \\ &\rightarrow -\frac{1}{2\ell(\alpha_0, \beta_0)} \left[\frac{\partial \Phi}{\partial x}(\beta, 0) - \frac{\partial \Phi}{\partial x}(\beta_0, 0) \right]^2 = \Lambda_2(\alpha_0, \beta, \beta_0). \end{aligned}$$

Moreover, we can prove that this convergence is uniform with respect to $\beta \in \Theta_\beta$. \square

Proof of Theorem 3.3. We just give here a sketch of the proof. To get (i), we prove the three steps 1-3 of [23], Section 4.4.1, that we have recalled at the beginning of Section 3.5.

Proof of 1. Case (1): since $(\varepsilon^2/T)(\Lambda_{\varepsilon, T}(\alpha, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta)) \rightarrow_{\mathbb{P}_{\theta_0}} \Lambda_1^{(1)}(\alpha, \alpha_0, \beta)$, uniformly with respect to (α, β) where $(\alpha, \beta) \rightarrow \Lambda_1^{(1)}(\alpha, \alpha_0, \beta)$ is continuous, < 0 , and $= 0$ iff $\alpha = \alpha_0$ implies the consistency of $\hat{\alpha}_{\varepsilon, T}$.

Analogously, in Case (2), since $\varepsilon^2(\Lambda_{\varepsilon, T}(\alpha, \beta) - \Lambda_{\varepsilon, T}(\alpha_0, \beta)) \rightarrow_{\mathbb{P}_{\theta_0}} \Lambda_1^{(2)}(\alpha, \alpha_0, \beta)$, uniformly with respect to (α, β) implies the consistency of $\hat{\alpha}_{\varepsilon, T}$.

Proof of 2. By (1), $\hat{\alpha}_{\varepsilon, T}$ is consistent thus $\mathbb{P}_{\theta_0}(\hat{\alpha}_{\varepsilon, T} \in \Theta_\alpha) \rightarrow 1$ as ε tends to 0. On the set $(\hat{\alpha}_{\varepsilon, T} \in \Theta_\alpha)$, we have:

$$0 = \frac{\partial \Lambda_{\varepsilon, T}}{\partial \alpha}(\hat{\alpha}_{\varepsilon, T}, \hat{\beta}_{\varepsilon, T}) = V_{\varepsilon, T} + (\hat{\alpha}_{\varepsilon, T} - \alpha_0) N_{\varepsilon, T}, \quad \text{where}$$

$$V_{\varepsilon,T} = \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\alpha_0, \hat{\beta}_{\varepsilon,T}), \quad N_{\varepsilon,T} = \int_0^1 \frac{\partial^2 \Lambda_{\varepsilon,T}}{\partial \alpha^2}(\alpha_0 + t(\hat{\alpha}_{\varepsilon,T} - \alpha_0), \hat{\beta}_{\varepsilon,T}) dt.$$

Thus, we have $\sqrt{T}\varepsilon^{-1}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) = -\frac{(\varepsilon/\sqrt{T})V_{\varepsilon,T}}{(\varepsilon^2/T)N_{\varepsilon,T}}$ for Case **(1)** and $\varepsilon^{-1}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) = -\frac{\varepsilon V_{\varepsilon,T}}{\varepsilon^2 N_{\varepsilon,T}}$ for Case **(2)**.

We must prove that $(\varepsilon/\sqrt{T})V_{\varepsilon,T}$ and $(\varepsilon^2/T)N_{\varepsilon,T}$ for Case **(1)**, $\varepsilon V_{\varepsilon,T}$ and $\varepsilon^2 N_{\varepsilon,T}$ for Case **(2)**, are tight under \mathbb{P}_{θ_0} . This can be done using the same tools as in Theorems 3.1 and 3.2, and using the assumption that $\frac{\partial \Phi}{\partial x}(\beta, 0)$ is uniformly bounded on Θ_β and that $\hat{\beta}_{\varepsilon,T} \in \Theta_\beta$.

Proof of 3. To obtain the consistency of $\hat{\beta}_{\varepsilon,T}$, it is enough to prove that:

$$\frac{1}{T}(\Lambda_{\varepsilon,T}(\hat{\alpha}_{\varepsilon,T}, \beta) - \Lambda_{\varepsilon,T}(\hat{\alpha}_{\varepsilon,T}, \beta_0)) \rightarrow \Lambda_2(\alpha_0, \beta, \beta_0) \quad (5.43)$$

uniformly in β .

Consider first Case **(1)**. Using (5.26), we have, setting $\alpha_u = \alpha_0 + u(\hat{\alpha}_{\varepsilon,T} - \alpha_0)$,

$$\begin{aligned} \Lambda_{\varepsilon,T}(\hat{\alpha}_{\varepsilon,T}, \beta) - \Lambda_{\varepsilon,T}(\hat{\alpha}_{\varepsilon,T}, \beta_0) &= (\Lambda_{\varepsilon,T}(\alpha_0, \beta) - \Lambda_{\varepsilon,T}(\alpha_0, \beta_0)) \\ &+ \frac{\sqrt{T}}{\varepsilon}(\hat{\alpha}_{\varepsilon,T} - \alpha_0) \frac{\varepsilon}{\sqrt{T}} R(\varepsilon, \theta, T), \end{aligned}$$

with

$$\begin{aligned} R(\varepsilon, \theta, T) &= \int_0^1 \left(\frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\alpha_u, \beta) - \frac{\partial \Lambda_{\varepsilon,T}}{\partial \alpha}(\alpha_u, \beta_0) \right) du \\ &= \frac{\sqrt{T}}{\varepsilon} \left(\int_0^1 \frac{(\ell(\alpha_u, \beta) - \ell(\alpha_u, \beta_0))}{\ell(\alpha_u)} \frac{\partial V}{\partial \alpha}(\alpha_u, x^*(\alpha_u)) du \frac{W_T}{\sqrt{T}} + o_P(1) \right). \end{aligned}$$

Now, since $\hat{\alpha}_{\varepsilon,T}$ is consistent, the integral term converges to a constant $C(\theta_0, \beta)$ which is bounded. Therefore $\frac{1}{T^{3/2}}\varepsilon R(\varepsilon, \theta, T) = \frac{W_T}{T^{3/2}}C(\theta_0, \beta) + \frac{1}{T}o_P(1) = o_P(1)$. Using now the tightness of $\varepsilon^{-1}\sqrt{T}(\hat{\alpha}_{\varepsilon,T} - \alpha_0)$ yields (5.43). The uniformity in β follows from the continuity of $\theta \rightarrow \ell(\theta)$.

For Case **(2)**, we use (5.38)-(5.39) and

$$R(\varepsilon, \theta, T) = \frac{1}{\varepsilon} \left(\int_0^T \int_0^1 (h(\alpha_u, \beta, s) - h(\alpha_u, \beta_0, s)) du dW_s + o_P(1) \right)$$

Therefore $\frac{1}{T}\varepsilon R(\varepsilon, \theta, T) = \frac{1}{T} \int_0^T dW_s \left(\int_0^1 (h(\alpha_u, \beta, s) - h(\alpha_u, \beta_0, s)) du \right) + \frac{1}{T}o_P(1)$. Since $\hat{\alpha}_{\varepsilon,T}$ is consistent, the integral term converges to a function $F(\theta_0, \beta, s)$ which is bounded uniformly in s .

Hence $\mathbb{E}_{\theta_0} \left(\frac{1}{T}\varepsilon R(\varepsilon, \theta, T) \right)^2 = \frac{1}{T^2} \int_0^T ds \mathbb{E}_{\theta_0} \left(\int_0^1 (h(\alpha_u, \beta, s) - h(\alpha_u, \beta_0, s)) du \right)^2 \lesssim \frac{1}{T}O(1)$. Hence, we get that (5.43) also holds in Case **(2)**.

Under the identifiability assumption for β , we get that in both cases the consistency of $\hat{\beta}_{\varepsilon,T}$.

The proof of the asymptotic normality follows, by standard tools from (i) and Theorems 3.1 and 3.2. \square

Acknowledgments

We thank the two referees for their helpful and detailed comments.

References

- [1] Altmeyer, R. and Reiss, M. (2021). Nonparametric estimation for linear SPDEs from local measurements. *Annals of Applied Probability*, **31** (1), 1-38. [MR4254472](#)
- [2] Amorino, C. and Gloter, A. (2020). Contrast function estimation for the drift parameter of ergodic jump diffusion process. *Scandinavian Journal of Statistics*, **47** (2), 279-246. [MR4157142](#)
- [3] Ball, F. and Sirl, D. (2020) Stochastic *SIR* in Structured Populations. *Stochastic Epidemic Models with Inference*, **Part II**, 123-240. Britton, T. and Pardoux, E., Editors. Lecture Notes in Mathematics, **2255**, Mathematical Biosciences Subseries, Springer. [MR4299426](#)
- [4] Baladron, J., Fasoli, D., Faugeras, O. and Touboul, J. (2012). Mean field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *The Journal of Mathematical Neuroscience*, **2**(1):10, 1-50. [MR2974499](#)
- [5] Benachour, S., Roynette, B., Talay, D. and Vallois, P. (1998a). Nonlinear self-stabilizing processes - I Existence, invariant probability, propagation of chaos. *Stochastic Processes and their Applications*, **75**, 173-201. [MR1632193](#)
- [6] Benachour, S., Roynette, B. and Vallois, P. (1998b). Nonlinear self-stabilizing processes - II Convergence to invariant probability. *Stochastic Processes and their Applications*, **75**, 203-224. [MR1632197](#)
- [7] Benedetto, D., Caglioti, E. and Pulverenti, M. (1997). A kinetic equation for granular media. *Mathematical Modelling and Numerical Analysis*, **31** (5), 615-641. [MR1471181](#)
- [8] Britton, T. and Pardoux, E., Editors (2020) *Stochastic Epidemic Models with Inference*. Lecture Notes in Mathematics, **2255**, Mathematical Biosciences Subseries, Springer. [MR4299427](#)
- [9] Cialenco, I. (2018). Statistical Inference for SPDEs: an Overview. *Statistical Inference for Stochastic Processes*, **21**, 309-329. [MR3824970](#)
- [10] Comte, F. and Genon-Catalot, V. (2020). Non parametric drift estimation for *i.i.d.* paths of stochastic differential equations. *The Annals of Statistics*, **48**, 3336-3365. [MR4185811](#)
- [11] Comte, F., Genon-Catalot, V. and Rozenholc, Y. (2007). Penalized non-parametric mean square estimation of the coefficients of diffusion processes. *Bernoulli*, **13**, 514-543. [MR2331262](#)
- [12] Dalalyan, A. (2005). Sharp adaptive estimation of the drift function for ergodic diffusions. *The Annals of Statistics*, **33**, 2507-2528. [MR2253093](#)
- [13] Dalalyan, A. and Reiss, M. (2007). Asymptotic statistical equivalence for ergodic diffusions: the multidimensional case. *Probability Theory and Related Fields*, **137**, 25-47. [MR2278451](#)

- [14] Delattre M., Genon-Catalot V. and Samson A. (2013). Maximum likelihood estimation for stochastic differential equations with random effects. *Scandinavian Journal of Statistics*, **40**, 322-343. [MR3066417](#)
- [15] Delattre M., Genon-Catalot V. and Larédo, C. (2018). Parametric inference for discrete observations of diffusion processes with mixed effects. *Stochastic processes and their Applications*, **128**, 1929-1957. [MR3797649](#)
- [16] Della Maestra, L. and Hoffmann, M. (2021). Nonparametric estimation for interacting particle systems: McKean-Vlasov models. *Probability Theory and Related Fields*, 1-63.
- [17] Forien, R. and Pardoux, E. (2019). Household epidemic models and McKean-Vlasov Poisson driven SDEs. *Preprint arXiv:1907.03001*.
- [18] Gärtner, J. (1988). On the McKean-Vlasov limit for interacting diffusions. *Mathematische Nachrichten*, **137**, 197-248. [MR0968996](#)
- [19] Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations. *Statistics*, **21**, 99-116. [MR1056065](#)
- [20] Genon-Catalot, V. and Larédo, C. (2014). Asymptotic equivalence of non-parametric diffusion and Euler scheme experiments. *The Annals of Statistics*, **42**, 1145-1165. [MR3224284](#)
- [21] Genon-Catalot, V. and Larédo, C. (2021). Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equation. *Stochastic Processes and Their Applications*, **142**, 513-548. [MR4324348](#)
- [22] Giesecke, K., Schwenkler, G. and Sirignano, J.A. (2020). Inference for large financial systems. *Mathematical Finance*, **30**, 3-46. [MR4067069](#)
- [23] Gloter, A. and Sørensen, M. (2009). Estimation for stochastic differential equations with a small diffusion coefficient. *Stochastic Processes and their Applications*, **119**, 679-699. [MR2500255](#)
- [24] Gobet, E., Hoffmann, M. and Reiss, M. (2004). Nonparametric estimation of scalar diffusions based on low frequency data. *The Annals of Statistics*, **32**, 2223-2253. [MR2102509](#)
- [25] Guy, R., Larédo, C. and Vergu, E. (2014). Parametric inference for discretely observed multidimensional diffusions with small diffusion coefficient. *Stochastic Processes and their Applications*, **124**, 51-80. [MR3131286](#)
- [26] Herrmann, S., Imkeller, P. and Peithmann, D. (2008). Large deviations and a Kramers' type law for self-stabilizing diffusions. *The Annals of Applied Probability*, **18**, 1379-1423. [MR2434175](#)
- [27] Hirsch, M.W. and Smale, S. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, New York, San Francisco, London. [MR0486784](#)
- [28] Hoffmann, M. (1999). Adaptive estimation in diffusion processes. *Stochastic Processes and their Applications*, **79**, 135-163. [MR1670522](#)
- [29] Höpfner, R. (2014). *Asymptotic Statistics with a View to Stochastic Processes*. Walter de Gruyter, Berlin/Boston. [MR3185373](#)
- [30] Iacus, S. M. (2010). *Simulation and Inference for Stochastic Differential Equations. With R examples*. Springer Series in Statistics, Springer. [MR2410254](#)

- [31] Kasonga, R.A. (1990). Maximum likelihood theory for large interacting systems. *SIAM Journal on Applied Mathematics*, **50**, 865-875. [MR1050917](#)
- [32] Kessler, M., Lindner, A. and Sørensen, M., Editors (2012). *Statistical Methods for Stochastic Differential Equations*. CRC press. Taylor & Francis Group, Boca Raton. [MR2975799](#)
- [33] Kutoyants, Y.A. (1984). *Parameter Estimation for Stochastic Processes*. Heldermann, Berlin. [MR0777685](#)
- [34] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. Springer Series in Statistics, Springer. [MR2144185](#)
- [35] Larédo, C. (1990). A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. *The Annals of Statistics*, **18**, 1158-1171. [MR1062703](#)
- [36] Larédo, C. (2020). Statistical Inference for Epidemic Processes in a Homogeneous Community. In *Stochastic Epidemic Models with Inference*, **Part IV** 363-472. Britton, T. and Pardoux, E., Editors. Lecture Notes in Mathematics, **2255**, Mathematical Biosciences Subseries, Springer. [MR4299436](#)
- [37] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics, Springer. [MR0856411](#)
- [38] Le Cam, L. and Yang, G. L. (2000). *Asymptotics in Statistics: Some Basic Concepts*. Second Edition. Springer Series in Statistics, Springer. [MR1784901](#)
- [39] McKean, H.P., Jr (1966). A class of Markov processes associated with non-linear parabolic equation. *Proceedings of the National Academy of Sciences, USA*, **56**, 1907-1911. [MR0221595](#)
- [40] Malrieu, F. (2003). Convergence to equilibrium for granular media equations and their Euler schemes. *Annals of Applied Probability*, **13**, 540-560. [MR1970276](#)
- [41] Masuda, H. (2007). Ergodicity and exponential beta-mixing for multidimensional diffusions with jumps. *Stochastic Processes and their Applications*, **117**, 35-56. [MR2287102](#)
- [42] Masuda, H. (2019). Non-Gaussian quasi-likelihood estimation of SDE driven by locally stable Lévy process. *Stochastic Processes and their Applications*, **129**, 1013-1059. [MR3913278](#)
- [43] Méléard, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations*, Lecture Notes in Mathematics, **1627**, 42-95, Springer. [MR1431299](#)
- [44] Molginer, A. and Edelstein-Keshet, L. (1999). A non-local model for a swarm. *Journal of Mathematical Biology*, **38**, 534-570. [MR1698215](#)
- [45] Piccini, U., de Gaetano, A. and Ditlevsen, S. (2010). Stochastic differential mixed-effects models. *Scandinavian Journal of Statistics*, **37**, 67-90. [MR2675940](#)
- [46] Schmisser, E. (2014). Non-parametric adaptive estimation of the drift for a jump diffusion process. *Stochastic Processes and their Applications*, **124**, 883-914. [MR3131317](#)
- [47] Sørensen, M. and Uchida, M. (2003). Small diffusion asymptotics for dis-

- cretely sampled stochastic differential equations. *Bernoulli*, **9**, 1051-1069. [MR2046817](#)
- [48] Sznitman, A.S. (1991). Topics in propagation of chaos. *Ecole d'été de probabilités de Saint-Flour XIX-1989*. Lecture Notes in Mathematics, **1464**, 165-251, Springer. [MR1108185](#)
- [49] Tsybakov, A. (2009). *Introduction to Nonparametric Estimation*, Springer Series in Statistics, Springer. [MR2724359](#)
- [50] Yoshida, N. (1992a). Estimation for diffusion processes from discrete observation. *Journal of Multivariate Analysis*, **41**, 220-242. [MR1172898](#)
- [51] Yoshida, N. (1992b). Asymptotic expansion for statistics related to small diffusions. *Journal of the Japan Statistical Society*, **22**, 139-159. [MR1212246](#)