

Model Misspecification And Uncertainty Quantification For Drift Estimation In Multiscale Diffusion Processes

Assyr Abdulle ^{*} Giacomo Garegnani ^{*} Grigorios A. Pavliotis [‡]
 Andrew M. Stuart [†]

Abstract. *Copy-paste of SIAM UQ abstract, modify.* We present a novel technique for estimating the drift function of a diffusion process possessing two separated time scales. Our aim is fitting a homogenized diffusion model to a continuous sample path coming from the full multiscale process, thus dealing with an issue of model misspecification. We consider a Bayesian framework and study the asymptotic limit of posterior distributions over the drift function. In this setting, we show on the one hand that if the continuous multiscale data are not pre-processed, then the posterior distribution concentrates asymptotically on the wrong value of the drift function. On the other hand, we show that data can be treated ahead of the inference procedure in order to obtain the desired posterior. In particular, we prove that there exists a family of transformations which are linear on the space of continuous sample paths and which, when applied to multiscale data, allow the posterior distribution to be asymptotically correct. We present a series of numerical examples on test cases which corroborate our theoretical findings.

AMS subject classifications.

Keywords.

1 Introduction

Add blabla introduction. In [2] (*is there other literature?*), the authors prove that inference of the parameters of a homogenized model has to be performed carefully. In this work, the analysis contained in [2] is widened with respect to the following aspects:

- 1) a more general form of the drift function is considered, thus allowing a more flexible framework for applications and hinting possible extensions to the infinite-dimensional case,
- 2) the inference procedure is reinterpreted from a Bayesian perspective, which guarantees more complete uncertainty quantification on the inference result. Moreover, given the nature of the problem, posterior distributions follow a Gaussian law which can be analytically determined, thus guaranteeing computationally fast inference,
- 3) we extend the sub-sampling technique introduced in [2], which can be applied to discrete sequences, by introducing theoretical tools which allow the treatment of continuous streams of data.

Let $\varepsilon > 0$ and let us consider the one-dimensional multiscale stochastic differential equation (SDE)

$$dX_t^\varepsilon = -V'(X_t^\varepsilon) dt - \frac{1}{\varepsilon} p' \left(\frac{X_t^\varepsilon}{\varepsilon} \right) + \sqrt{2\sigma} dW_t, \quad (1.1)$$

where $\sigma > 0$ and W_t is a standard one-dimensional Brownian motion. The functions $V, p: \mathbb{R} \rightarrow \mathbb{R}$ are slow and fast potentials driving the dynamics of the solution X_t^ε . Given $N \in \mathbb{N}_{>0}$, we assume

^{*}Institute of Mathematics, École Polytechnique Fédérale de Lausanne, {assyr.abdulle, giacomo.garegnani}@epfl.ch

[†]Department of Computing and Mathematical Sciences, Caltech, Pasadena, astuart@caltech.edu

[‡]Department of Mathematics, Imperial College London, g.pavliotis@imperial.ac.uk

the slow potential to be of the form

$$V(x) = \sum_{i=1}^N \alpha_i V_i(x),$$

for coefficients $\alpha_i \in \mathbb{R}$, $i = 1, \dots, N$, and smooth functions $V_i: \mathbb{R} \rightarrow \mathbb{R}$. Moreover, we assume p to be smooth and periodic of period L . Theory of homogenization [1] guarantees the existence of an SDE of the form

$$dX_t^0 = -KV'(X) dt + \sqrt{2K\sigma} dW_t, \quad (1.2)$$

where W_t is the same Brownian motion and where the fast dynamics have been eliminated, such that $X_t^\varepsilon \rightarrow X_t^0$ for $\varepsilon \rightarrow 0$ in law as random variables in $\mathcal{C}^0((0, T), \mathbb{R})$. In the following, we will denote $A_i := K\alpha_i$. The coefficient K is given by the formula

$$K = \int_0^L (1 + \Phi'(y))^2 \mu(dy),$$

with

$$\mu(dy) = \frac{1}{Z} \exp\left(-\frac{p'(y)}{\sigma}\right) dy, \quad Z = \int_0^L \exp\left(-\frac{p'(y)}{\sigma}\right) dy,$$

and where the function Φ is the solution of the elliptic partial differential equation

$$-p'(y)\Phi'(y) + \sigma\Phi''(y) = p''(y), \quad 0 \leq y \leq L,$$

endowed with periodic boundary conditions.

2 The Bayesian perspective

Let us denote by $A \in \mathbb{R}^N$ the vector of the coefficients $A_i := K\alpha_i$ appearing in the drift term of the homogenized SDE (1.2). In a Bayesian setting, our goal is to determine the posterior distribution $\mu(A \mid X_{0:T})$ given a continuous trajectory $X_{0:T} := (X_t, 0 \leq t \leq T)$. Choosing a Gaussian prior $\mu_0 = \mathcal{N}(A_0, C_0)$ on A , where $A_0 \in \mathbb{R}^N$ and $C_0 \in \mathbb{R}^{N \times N}$ is symmetric positive definite, the posterior distribution admits a density $p(A \mid X_{0:T})$ with respect to the Lebesgue measure which satisfies

$$p(A \mid X_{0:T}) = \frac{1}{Z} p(X_{0:T} \mid A) p_0(A),$$

where Z is a normalization constant, p_0 is the density of μ_0 , and where Girsanov's formula allows to write the likelihood as

$$p(X_{0:T} \mid A) = \exp \left\{ - \int_0^T \langle A, \mathbf{V}'(X_t) \rangle dX_t - \frac{1}{2} \int_0^T \langle A, \mathbf{V}'(X_t) \rangle^2 dt \right\}.$$

In the formula above, we denote by $\mathbf{V}: \mathbb{R} \rightarrow \mathbb{R}^N$ the vector-valued function whose i -th component is $\mathbf{V}_i = V_i$, and by \mathbf{V}' the vector of its derivatives, i.e., the Jacobian matrix. Moreover, we denote by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product in \mathbb{R}^N and by $\|\cdot\|_2$ the corresponding norm. The log-posterior density is then given by

$$\begin{aligned} \log p(A \mid X_{0:T}) &= -\log Z - \int_0^T \langle A, \mathbf{V}'(X_t) \rangle dX_t - \frac{1}{2} \int_0^T \langle A, \mathbf{V}'(X_t) \rangle^2 dt \\ &\quad - \frac{1}{2} \left\| C_0^{-1/2}(A - A_0) \right\|_2^2. \end{aligned}$$

Since the log-posterior density is quadratic in A , the posterior is Gaussian, and it is therefore sufficient to determine its mean and covariance to fully characterize it. We denote by m_T and

C_T the mean and covariance matrix, respectively. Let us consider the matrix-valued function $M_{\mathbf{V}}: \mathbb{R} \rightarrow \mathbb{R}^{N \times N}$ whose entries are given by

$$(M_{\mathbf{V}}(T))_{ij} = \int_0^T V_i'(X_t) V_j'(X_t) dt, \quad i, j = 1, \dots, N,$$

and the vector-valued function $I_{\mathbf{V}}: \mathbb{R} \rightarrow \mathbb{R}^N$ defined by

$$(I_{\mathbf{V}}(T))_i = \int_0^T V_i'(X_t) dX_t.$$

If we employ this notation, we can rewrite the log-posterior density as

$$\log p(A \mid X_{0:T}) = -\log Z - \langle A, I_{\mathbf{V}}(T) \rangle - \frac{1}{2} \langle A, M_{\mathbf{V}}(T) A \rangle - \frac{1}{2} \langle A - A_0, C_0^{-1} (A - A_0) \rangle.$$

Identifying the quadratic and linear terms in the log-posterior density, it is possible to show that the posterior covariance and mean are formally given by

$$\begin{aligned} C_T^{-1} &= C_0^{-1} + M_{\mathbf{V}}(T), \\ m_T &= C_T (C_0^{-1} A_0 - I_{\mathbf{V}}(T)). \end{aligned}$$

We are now interested in the limit of the posterior distribution for $T \rightarrow \infty$. Let us first define the maximum likelihood estimator (MLE) $\hat{A}(X_{0:T})$ of A , which is obtained by maximizing the log-likelihood function $\log p(X_{0:T} \mid A)$, and which is hence formally given by

$$\hat{A}(X_{0:T}) = -M_{\mathbf{V}}(T)^{-1} I_{\mathbf{V}}(T).$$

We now introduce regularity assumptions on the SDE.

Assumption 1. The potentials p and V are such that

- i) the symmetric matrix $M_{\mathbf{V}}(T)$ is positive definite for all $T > 0$,
- ii) there exist $a, b > 0$ such that $x V_i'(x) \geq -a + bx^2$, for all $x \in \mathbb{R}$ and for all $i = 1, \dots, N$,
- iii) **other assumptions to be added**

We can now state the main result for asymptotic convergence of the posterior distribution.

Proposition 2.1. *Given a trajectory $X_{0:T}$ for T arbitrarily large and under Assumption 1, the posterior mean and covariance satisfy*

$$\begin{aligned} \lim_{T \rightarrow \infty} \|m_T - \hat{A}(X_{0:T})\|_2 &= 0, \\ \lim_{T \rightarrow \infty} \|C_T\|_F &= 0, \end{aligned}$$

i.e., the posterior tends in distribution to the Dirac delta centred in the limit of the MLE.

Proof. Let us first consider the covariance matrix. Hua's identity yields

$$C_T = M_{\mathbf{V}}(T)^{-1} - Q_T^{-1},$$

where

$$Q_T = M_{\mathbf{V}}(T) + M_{\mathbf{V}}(T) C_0 M_{\mathbf{V}}(T).$$

Let us now consider the matrix $M_{\mathbf{V}}(T)$. We now consider the mean. We have

$$\|m_T - \hat{A}(X_{0:T})\|_2 = \|M_{\mathbf{V}}(T)^{-1} C_0^{-1} A_0 - Q_T^{-1} (C_0^{-1} A_0 - I_{\mathbf{V}}(T))\|_2.$$

to be completed □

Remark 1. Proposition 2.1 guarantees that in the asymptotic limit of $T \rightarrow \infty$, it is equivalent to consider the Bayesian approach and the maximum likelihood approach, and can therefore be interpreted as a consistency result for both approaches. Nonetheless, in this Gaussian framework the Bayesian approach provides richer information on the inference result with a negligible additional cost.

3 Convergence analysis for a single coefficient

Let us consider $N = 1$ and write $A_1 = A$ and $V_1 = V$, respectively. The posterior distribution of A given a trajectory $(X_t, 0 \leq t \leq T)$ is in this case a Gaussian $\mathcal{N}(m_T, \sigma_T^2)$ where

$$m_T = \frac{A_0}{1 + \sigma_0^2 \int_0^T V'(X_t)^2 dt} - \frac{\sigma_0^2 \int_0^T V'(X_t) dX_t}{1 + \sigma_0^2 \int_0^T V'(X_t)^2 dt},$$

$$\sigma_T^{-2} = \sigma_0^{-2} + \int_0^T V'(X_t)^2 dt.$$

We will in the following sections analyse the convergence of the posterior, both in case data from the multiscale process is not treated and in case they are pre-processed.

3.1 Failure without pre-processing

Let us consider a trajectory $X_{0:T}^\varepsilon := (X_t^\varepsilon, 0 \leq t \leq T)$ coming from the multiscale equation (1.1), and the corresponding posterior distribution over the parameter A , which we denote by $\mu^\varepsilon(A, X_{0:T}^\varepsilon)$. The following result holds.

Theorem 3.1. *Under assumption 1 and if $T = \varepsilon^{-\gamma}$ for $\gamma > 0$, then the posterior distribution $\mu^\varepsilon(A | X_{0:T}^\varepsilon) = \mathcal{N}(m_T, \sigma_T^2)$ satisfies*

$$\lim_{T \rightarrow \infty} m_T = \alpha,$$

$$\lim_{T \rightarrow \infty} \sigma_T^2 = 0.$$

Proof. Proposition 2.1 guarantees that

$$\lim_{T \rightarrow \infty} |m_T - \hat{A}(X_{0:T}^\varepsilon)| = 0,$$

and that $\sigma_T^2 \rightarrow 0$ for $T \rightarrow \infty$. Moreover, [2, Theorem xx] yields

$$\lim_{T \rightarrow \infty} \hat{A}(X_{0:T}^\varepsilon) = \alpha,$$

which completes the proof. \square

The result above implies that the posterior distribution over the drift coefficient concentrates asymptotically on an undesired value.

3.2 Success with pre-processing

Let $Z_t^{\varepsilon,k}$ be defined as

$$Z_t^{\varepsilon,k} := \int_0^t k^\varepsilon(t, s) X_s^\varepsilon ds,$$

where $k^\varepsilon(t, s)$ is a kernel satisfying the following assumption.

Assumption 2. The kernel $k: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ either satisfies for all $T > 0$, $t, s \in (0, T)$ and $\alpha \geq 1/2$

- i) (regularity) $k^\varepsilon(t, s), \partial_t k^\varepsilon(t, s) \in \mathcal{C}^0(\mathbb{R}^+ \times \mathbb{R}^+; \mathbb{R})$
- ii) (divergence-free) $\partial_t k^\varepsilon(t, s) + \partial_s k^\varepsilon(t, s) = 0$,
- iii) (normalization) there exists a function $\psi(t, \varepsilon)$ satisfying for all $p \geq 0$ and for a constant $C > 0$ independent of ε

$$\lim_{t \rightarrow \infty, \varepsilon \rightarrow 0} \psi(t, \varepsilon) = 0, \quad \int_0^T |\psi(t, \varepsilon)|^p dt \leq C,$$

such that k^ε satisfies

$$\int_0^t (t-s) \partial_t k^\varepsilon(t, s) ds = -1 + \psi(t, \varepsilon),$$

$$\text{iv) (limit) } \lim_{t \rightarrow \infty} k^\varepsilon(t, 0) = 0.$$

or it can be written as

$$k^\varepsilon(t, s) = \chi_{[t-\delta, t]}(s),$$

where $\chi_{[t', t]}$ is the indicator function of the interval $[t', t]$ for $t' < t$ and $\delta = \varepsilon^\zeta$ for $\zeta \in (0, 1)$.

Example 1. Let $\delta = \varepsilon^\zeta$ for $\zeta \in (0, 1)$, $\beta > 0$ and

$$k^\varepsilon(t, s) = C_\beta \delta^{-1/\beta} e^{-(t-s)^\beta/\delta},$$

where C_β is a normalizing constant. It is possible to verify that for each β there exists C_β such that i)–iv) are all verified. For example, for $\beta = 1$ we have $C_1 = 1$, whereas for $\beta = 2$ we have $C_2 = 2/\sqrt{\pi}$.

Remark 2. The assumptions on the kernel given above could be relaxed without spoiling the theoretical results presented in the following. For example, condition ii) could be relaxed to

$$\lim_{t \rightarrow \infty} \int_0^t |\partial_t k^\varepsilon(t, s) + \partial_s k^\varepsilon(t, s)| = 0,$$

which is trivially implied by (ii)).

Due to assumption (i)) we can apply Leibniz's integral rule and obtain the following representation

$$\frac{dZ_t^{\varepsilon, k}}{dt} = k^\varepsilon(t, t) X_t^\varepsilon + \int_0^t \partial_t k^\varepsilon(t, s) X_s^\varepsilon ds.$$

Adding and subtracting X_t^ε inside the integral yields

$$\frac{dZ_t^{\varepsilon, k}}{dt} = \int_0^t \partial_t k^\varepsilon(t, s) (X_s^\varepsilon - X_t^\varepsilon) ds + \left(k^\varepsilon(t, t) + \int_0^t \partial_t k^\varepsilon(t, s) ds \right) X_t^\varepsilon.$$

Due to (ii)), we have

$$\begin{aligned} k^\varepsilon(t, t) + \int_0^t \partial_t k^\varepsilon(t, s) ds &= k^\varepsilon(t, 0) + \int_0^t (\partial_s k^\varepsilon(t, s) + \partial_t k^\varepsilon(t, s)) ds \\ &= k^\varepsilon(t, 0). \end{aligned}$$

Therefore,

$$\frac{dZ_t^{\varepsilon, k}}{dt} = \int_0^t \partial_t k^\varepsilon(t, s) (X_s^\varepsilon - X_t^\varepsilon) ds + k^\varepsilon(t, 0) X_t^\varepsilon.$$

Then

$$\begin{aligned} \frac{dZ_t^{\varepsilon, k}}{dt} &= \alpha \int_0^t \int_s^t \partial_t k^\varepsilon(t, s) V_0'(X_r^\varepsilon) (1 + \Phi'(Y_r^\varepsilon)) dr ds \\ &\quad - \sqrt{2\sigma} \int_0^t \int_s^t \partial_t k^\varepsilon(t, s) (1 + \Phi'(Y_r^\varepsilon)) dW_r ds \\ &\quad + \varepsilon \int_0^t \partial_t k^\varepsilon(t, s) (\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)) ds + k^\varepsilon(t, 0) X_t^\varepsilon \end{aligned}$$

Lemma 3.2. *Lemma text*

$$\alpha \int_s^t V_0'(X_r^\varepsilon) (1 + \Phi'(Y_r^\varepsilon)) dr = A(t-s) V_0'(Z_t^{\varepsilon, k}) + R(\varepsilon, t-s),$$

where for all $p \geq 1$ it holds

$$\left(\mathbb{E}^{\mu^\varepsilon} |R(\varepsilon, t)|^p \right)^{1/p} \leq C(\varepsilon^2 + \varepsilon t^{1/2} + \varepsilon t + t^{3/2}),$$

4 Convergence analysis for $N > 1$

5 “Old” computations which could be useful

In this section, we study the convergence with respect to the parameter ε of point estimates of the drift and the diffusion coefficients when the estimator is computed employing continuous data coming from the multiscale model.

5.1 Drift coefficient

Let $X^\varepsilon := (X_t^\varepsilon, 0 \leq t \leq T)$ be the solution of (1.1) and define $\mathcal{H}_\Delta(X^\varepsilon)$ as

$$\mathcal{H}_\Delta(X^\varepsilon)_t := \begin{cases} X_0, & t = 0, \\ \frac{1}{t} \int_0^t X_s ds, & 0 < t < \Delta, \\ \frac{1}{\Delta} \int_{t-\Delta}^t X_s ds, & \Delta \leq t \leq T, \end{cases}$$

with $\Delta > 0$. Let us denote for ease of notation, $Z_t^\varepsilon := \mathcal{H}_\Delta(X^\varepsilon)_t$. The maximum likelihood estimator of the drift coefficient is then

$$\hat{A}_{T,\Delta}(Z_t^\varepsilon) = - \frac{\int_0^T V'_0(Z_t^\varepsilon) dZ_t^\varepsilon}{\int_0^T V'_0(Z_t^\varepsilon)^2 dt}.$$

Let us remark that for $0 < t < \Delta$,

$$d(tZ_t^\varepsilon) = X_t dt,$$

which implies

$$dZ_t^\varepsilon = \frac{1}{t}(X_t^\varepsilon - Z_t^\varepsilon) dt.$$

For $\Delta \leq t \leq T$, instead

$$dZ_t^\varepsilon = \frac{1}{\Delta}(X_t^\varepsilon - X_{t-\Delta}^\varepsilon) dt.$$

We rewrite the estimator as

$$\hat{A}_{T,\Delta}(Z_t^\varepsilon) = - \frac{\int_0^\Delta V'_0(Z_t^\varepsilon) \frac{1}{t}(X_t^\varepsilon - Z_t^\varepsilon) dt}{\int_0^T V'_0(Z_t^\varepsilon)^2 dt} - \frac{\int_\Delta^T V'_0(Z_t^\varepsilon)(X_t^\varepsilon - X_{t-\Delta}^\varepsilon) dt}{\Delta \int_0^T V'_0(Z_t^\varepsilon)^2 dt}.$$

The goal of this section is proving the following result.

Theorem 5.1. *Under assumption **red assumptions**, if there exists $\zeta \in (0, 1)$ such that $\Delta = \varepsilon^\zeta$ and $\gamma > \zeta$ such that $T = \varepsilon^{-\gamma}$, it holds*

$$\lim_{\varepsilon \rightarrow 0} \hat{A}_{T,\Delta}(Z_t^\varepsilon) = A, \quad \text{in law.}$$

It is useful in the following to rewrite (1.1) as a system of two coupled SDEs. In particular, introducing the variable $Y_t^\varepsilon := X_t^\varepsilon/\varepsilon$, one has

$$\begin{aligned} dX_t^\varepsilon &= -\alpha V'_0(X_t^\varepsilon) dt - \frac{1}{\varepsilon} V'_1(Y_t^\varepsilon) + \sqrt{2\sigma} dW_t, \\ dY_t^\varepsilon &= -\frac{\alpha}{\varepsilon} V'_0(X_t^\varepsilon) dt - \frac{1}{\varepsilon^2} V'_1(Y_t^\varepsilon) + \sqrt{\frac{2\sigma}{\varepsilon^2}} dW_t. \end{aligned}$$

The analysis necessary to prove Theorem 5.1 is based on the expansion

$$\begin{aligned} X_t^\varepsilon - X_{t-\Delta}^\varepsilon &= -\alpha \int_{t-\Delta}^t V'_0(X_s^\varepsilon)(1 + \Phi'(Y_s^\varepsilon)) ds \\ &\quad + \sqrt{2\sigma} \int_{t-\Delta}^t (1 + \Phi'(Y_s^\varepsilon)) dW_s \\ &\quad - \varepsilon(\Phi(Y_t^\varepsilon) - \Phi(Y_{t-\Delta}^\varepsilon)), \end{aligned} \tag{5.1}$$

for $t \geq \Delta$ (see [2, Equation (5.8)]). The following lemma ensures that the process Z_t^ε has bounded moments.

Lemma 5.2. *The process Z_t^ε has bounded moments of all order, i.e., for all $p \geq 1$ and $t \geq 0$ it holds*

$$\mathbb{E}^{\mu^\varepsilon} |Z_t^\varepsilon|^p \leq C,$$

for $C > 0$ a constant uniform in $\varepsilon \rightarrow 0$.

Proof. The process X_t^ε has bounded moments (see [2, Corollary 5.4]), which implies the desired result with an application of the Hölder inequality. In fact, for $0 < t < \Delta$,

$$\begin{aligned} \mathbb{E}^{\mu^\varepsilon} |Z_t^\varepsilon|^p &\leq \frac{t^{p-1}}{t^p} \int_0^t \mathbb{E}^{\mu^\varepsilon} |X_s^\varepsilon|^p \, ds \\ &\leq t^{-1} \int_0^t C \, ds = C. \end{aligned}$$

For $\Delta \leq t \leq T$ the procedure is analogue. □

In the following lemma the difference between the processes X_t^ε and Z_t^ε is bounded.

Lemma 5.3. *Under assumptions **add assumptions***

$$\mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - Z_t^\varepsilon|^p \leq C(\Delta^p + \Delta^{p/2} + \varepsilon^p),$$

where $C > 0$ is a constant independent of Δ and ε .

Proof. By definition of Z_t^ε for $\Delta \leq t \leq T$ and applying Hölder's inequality we have

$$\begin{aligned} \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - Z_t^\varepsilon|^p &= \Delta^{-p} \mathbb{E}^{\mu^\varepsilon} \left| \int_{t-\Delta}^t (X_t^\varepsilon - X_s^\varepsilon) \, ds \right|^p \\ &\leq \Delta^{-1} \int_{t-\Delta}^t \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - X_s^\varepsilon|^p \, ds \end{aligned}$$

We can now apply [2, Lemma 6.1] to the integrand to obtain

$$\mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - Z_t^\varepsilon|^p \leq C \Delta^{-1} \int_{t-\Delta}^t (\Delta^p + \Delta^{p/2} + \varepsilon^p) \, ds,$$

which implies the desired result. The case $0 < t \leq \Delta$ can be proved analogously. □

Lemma 5.4 (See [2, Proposition 5.8]). *Under assumptions **add assumptions**, it holds in law*

$$\alpha \int_{t-\Delta}^t V_0'(X_s^\varepsilon) (1 + \Phi'(Y_s^\varepsilon)) \, ds = A \Delta V_0'(Z_t^\varepsilon) + R(\varepsilon, \Delta),$$

where for every $p > 0$ and if Δ and ε are sufficiently small, then

$$\left(\mathbb{E}^{\mu^\varepsilon} |R(\varepsilon, \Delta)|^p \right)^{1/p} \leq C(\varepsilon^2 + \Delta^{1/2} \varepsilon + \Delta^{3/2}),$$

where $C > 0$ is independent of ε and Δ .

Proof. Let us denote $\Psi(t) := 1 + \Phi'(Y_t^\varepsilon)$. Then

$$\begin{aligned} \mathbb{E}^{\mu^\varepsilon} |R(\varepsilon, \Delta)|^p &= \mathbb{E}^{\mu^\varepsilon} \left| \int_{t-\Delta}^t \alpha V_0'(X_s^\varepsilon) \Psi(s) \, ds - \Delta A V_0'(Z_t^\varepsilon) \right|^p \\ &\leq C \mathbb{E}^{\mu^\varepsilon} \left| V_0'(Z_t^\varepsilon) \int_{t-\Delta}^t (\alpha \Psi(s) - A) \, ds \right|^p \\ &\quad + C \mathbb{E}^{\mu^\varepsilon} \left| \int_{t-\Delta}^t \alpha (V_0'(X_t^\varepsilon) - V_0'(Z_t^\varepsilon)) \Psi(s) \, ds \right|^p. \end{aligned}$$

The result is then obtained following the proof of [2, Proposition 5.8] and replacing [2, Lemma 6.1] with Lemma 5.3, and [2, Corollary 4.1] with Lemma 5.2. \square

We can now prove Theorem 5.1.

Proof of Theorem 5.1. Consider the decomposition (5.1). Denoting

$$J_t := \sqrt{2\sigma} \int_{t-\Delta}^t (1 + \Phi'(Y_s^\varepsilon)) dW_s,$$

we have due to Lemma 5.4 the equality in law

$$X_t^\varepsilon - X_{t-\Delta}^\varepsilon = -A\Delta V'(Z_t^\varepsilon) + J_t + \widehat{R}(\varepsilon, \Delta),$$

where, since $\zeta \in (0, 1)$, we have

$$\left(\mathbb{E}^{\mu^\varepsilon} \left| \widehat{R}(\varepsilon, \Delta) \right|^p \right)^{1/p} \leq C(\varepsilon + \varepsilon^{3\zeta/2})$$

Therefore, we have that the estimator satisfies

$$\begin{aligned} \widehat{A}_{T,\Delta}(Z_t^\varepsilon) &= A - A \frac{\int_0^\Delta V'_0(Z_t^\varepsilon)^2 dt}{\int_0^T V'_0(Z_t^\varepsilon)^2 dt} - \frac{\int_0^\Delta V'_0(Z_t^\varepsilon) \frac{1}{t} (X_t^\varepsilon - Z_t^\varepsilon) dt}{\int_0^T V'_0(Z_t^\varepsilon)^2 dt} \\ &\quad - \frac{\int_\Delta^T V'_0(Z_t^\varepsilon) J_t dt}{\Delta \int_0^T V'_0(Z_t^\varepsilon)^2 dt} - \frac{\widehat{R}(\varepsilon, \Delta) \int_\Delta^T V'_0(Z_t^\varepsilon) dt}{\Delta \int_0^T V'_0(Z_t^\varepsilon)^2 dt} \\ &=: A - I_1 - I_2 - I_3 - I_4, \end{aligned} \tag{5.2}$$

in law. Let us analyse the terms I_i , $i = 1, \dots, 4$ separately. Let us consider I_1 and multiply both the numerator and the denominator by $1/T$. Due to assumption **add assumption** and Lemma 5.2, we have

$$\frac{A}{T} \mathbb{E}^{\mu^\varepsilon} \left| \int_0^\Delta V'_0(Z_t^\varepsilon)^2 dt \right| \leq C\varepsilon^{\gamma+\zeta},$$

for a constant $C > 0$ independent of Δ and ε . Hence the numerator vanishes in L^1 and thus in law for $\varepsilon \rightarrow 0$. We split the denominator as

$$\frac{1}{T} \int_0^T V'_0(Z_t^\varepsilon)^2 dt = \frac{1}{T} \int_0^T V'_0(X_t^\varepsilon)^2 dt + \frac{1}{T} \int_0^T (V'_0(Z_t^\varepsilon)^2 - V'_0(X_t^\varepsilon)^2) dt$$

For the first term, we have by the ergodic theorem

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T V'_0(X_t^\varepsilon)^2 dt = \mathbb{E}^{\mu^\varepsilon} |V'_0|^2, \quad \text{a.s.}$$

For the second term, we have applying Cauchy-Schwarz's inequality and due to assumption **add assumption** and Lemma 5.3

$$\begin{aligned} \frac{1}{T} \mathbb{E}^{\mu^\varepsilon} \left| \int_0^T (V'_0(Z_t^\varepsilon)^2 - V'_0(X_t^\varepsilon)^2) dt \right| &\leq \frac{C}{T} \int_0^T \left(\mathbb{E}^{\mu^\varepsilon} |V'_0(Z_t^\varepsilon) - V'_0(X_t^\varepsilon)|^2 \right)^{1/2} dt \\ &\leq C \left(\Delta + \Delta^{1/2} + \varepsilon \right), \end{aligned}$$

which implies that the denominator tends to a finite value in probability for $\varepsilon \rightarrow 0$. Therefore, by Slutsky's theorem,

$$\lim_{\varepsilon \rightarrow 0} I_1 = 0, \quad \text{in law.}$$

Let us now consider I_2 and multiply numerator and denominator by $1/T$. The denominator is the same as I_1 , and therefore does not need to be treated further. The numerator can be bounded in L^1 as

$$\frac{1}{T} \mathbb{E}^{\mu^\varepsilon} \left| \int_0^\Delta V'_0(Z_t^\varepsilon) \frac{1}{t} (X_t^\varepsilon - Z_t^\varepsilon) dt \right| \leq \frac{C}{\Delta T} \int_0^\Delta \frac{\Delta}{t} \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - Z_t^\varepsilon| dt,$$

which, since $Z_0^\varepsilon = X_0^\varepsilon$, vanishes for $\varepsilon \rightarrow 0$. Hence, an application of Slutsky's theorem yields

$$\lim_{\varepsilon \rightarrow 0} I_2 = 0, \quad \text{in law.}$$

We consider now I_3 , which can be rewritten as

$$\begin{aligned} I_3 &= \frac{1}{\sqrt{T\Delta}} \frac{\frac{1}{\sqrt{T\Delta}} \int_\Delta^T V'_0(Z_t^\varepsilon) J_t dt}{\frac{1}{T} \int_0^T V'_0(Z_t^\varepsilon)^2 dt} \\ &= \varepsilon^{(\gamma-\zeta)/2} \frac{\frac{1}{\sqrt{T\Delta}} \int_\Delta^T V'_0(Z_t^\varepsilon) J_t dt}{\frac{1}{T} \int_0^T V'_0(Z_t^\varepsilon)^2 dt} \end{aligned}$$

Let us remark that J_t is a martingale and that by Itô isometry

$$\mathbb{E}^{\mu^\varepsilon} |J_\Delta|^2 = 2\Sigma\Delta,$$

Therefore, we can apply the central limit theorem for martingales to the numerator and obtain the equality in law

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{1}{\sqrt{T\Delta}} \int_\Delta^T V'_0(Z_t^\varepsilon) J_t dt &= \frac{1}{\sqrt{\Delta}} \mathcal{N} \left(0, \mathbb{E}^{\mu^\varepsilon} \left(|V'_0(X_0^\varepsilon)|^2 |J_\Delta|^2 \right) \right) \\ &= C\mathcal{N}(0, 1). \end{aligned}$$

The denominator is the same as in I_2 and I_3 and tends in probability to a finite value. Hence, since by hypothesis $\gamma > \zeta$, we have

$$\lim_{\varepsilon \rightarrow 0} I_3 = 0, \quad \text{in law.}$$

For the last term I_4 , we have

$$I_4 = \frac{\varepsilon^{\gamma-\zeta} \widehat{R}(\varepsilon, \Delta) \int_\Delta^T V'_0(Z_t^\varepsilon) dt}{\frac{1}{T} \int_0^T V'_0(Z_t^\varepsilon)^2 dt}.$$

For the numerator, we have by the Cauchy–Schwarz inequality and due to Lemma 5.4

$$\begin{aligned} \varepsilon^{\gamma-\zeta} \mathbb{E}^{\mu^\varepsilon} \left| \widehat{R}(\varepsilon, \Delta) \int_\Delta^T V'_0(Z_t^\varepsilon) dt \right| &\leq \varepsilon^{\gamma-\zeta} \left(\mathbb{E}^{\mu^\varepsilon} \left| \widehat{R}(\varepsilon, \Delta) \right|^2 \right)^{1/2} \left(\mathbb{E}^{\mu^\varepsilon} \left| \int_\Delta^T V'_0(Z_t^\varepsilon) dt \right|^2 \right)^{1/2} \\ &\leq C \varepsilon^{\gamma-\zeta} (\varepsilon + \varepsilon^{3\zeta/2}) \varepsilon^{-\gamma} \\ &\leq C \left(\varepsilon^{1-\zeta} + \varepsilon^{\zeta/2} \right) \end{aligned}$$

which implies that, since the denominator is the same as before,

$$\lim_{\varepsilon \rightarrow 0} I_4 = 0, \quad \text{in law.}$$

The decomposition (5.2), together with the limits of I_i for $i = 1, \dots, 4$, prove the desired result. \square

6 Numerical experiments

6.1 Ornstein–Uhlenbeck process

6.2 Bistable potential

6.3 Real(istic) data?

References

- [1] A. BENSOUSSAN, J.-L. LIONS, AND G. PAPANICOLAOU, *Asymptotic analysis for periodic structures*, North-Holland Publishing Co., Amsterdam, 1978.
- [2] G. A. PAVLIOTIS AND A. M. STUART, *Parameter estimation for multiscale diffusions*, J. Stat. Phys., 127 (2007), pp. 741–781.