

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

Assyr Abdulle <sup>\*</sup>      Giacomo Garegnani <sup>\*</sup>      Grigorios A. Pavliotis <sup>†</sup>  
Andrew M. Stuart <sup>‡</sup>      Andrea Zanoni <sup>\*</sup>

**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 motivating 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)

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

---

<sup>\*</sup>Institute of Mathematics, École Polytechnique Fédérale de Lausanne

<sup>†</sup>Department of Mathematics, Imperial College London

<sup>‡</sup>Department of Computing and Mathematical Sciences, Caltech

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  $X_t^\varepsilon$ . Given  $N \in \mathbb{N}_{>0}$ , we assume 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

$$\mu(dy) = Z^{-1} e^{-p(y)/\sigma} dy, \quad \text{where} \quad Z = \int_0^L e^{-p(y)/\sigma} dy,$$

and where the function  $\Phi$  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.

**Failure without pre-processing: explain.**

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  $\mu^\varepsilon(A \mid X_{0:T}^\varepsilon)$ . The following result holds.

**Theorem 1.1.** *Under assumption 2.1 and if  $T = \varepsilon^{-\gamma}$  for  $\gamma > 0$ , then the posterior distribution  $\mu^\varepsilon(A \mid X_{0:T}^\varepsilon) = \mathcal{N}(\bar{A}_T^\varepsilon, \sigma_T^2)$  satisfies*

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \bar{A}_T^\varepsilon &= \alpha, \\ \lim_{\varepsilon \rightarrow 0} \sigma_T^2 &= 0. \end{aligned}$$

*Proof.* Proposition 2.2 guarantees that

$$\lim_{\varepsilon \rightarrow 0} \left| \bar{A}_T - \hat{A}(X_{0:T}^\varepsilon) \right| = 0,$$

and that  $\sigma_T^2 \rightarrow 0$  for  $\varepsilon \rightarrow 0$ . Moreover, [2, Theorem 3.4] yields

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

which completes the proof. □

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

## 2 Posterior formal computations

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 the normalization constant,  $p_0$  is the density of  $\mu_0$ , 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  $M: \mathbb{R} \rightarrow \mathbb{R}^{N \times N}$  whose entries are given by

$$M_{ij} = \frac{1}{T} \int_0^T V'_i(X_t) V'_j(X_t) dt, \quad i, j = 1, \dots, N,$$

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

$$h_i = \frac{1}{T} \int_0^T V'_i(X_t) dX_t, \quad i = 1, \dots, N.$$

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

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

Completing the squares in the log-posterior density, it is possible to show that the posterior is a Gaussian  $\mu(A \mid X_{0:T}) = \mathcal{N}(\bar{A}_T, C_T)$ , whose precision matrix and mean are formally given by

$$\begin{aligned} C_T^{-1} &= C_0^{-1} + TM, \\ C_T^{-1} \bar{A}_T &= C_0^{-1} A_0 - Th. \end{aligned}$$

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

$$\hat{A}(X_{0:T}) = -M^{-1}h.$$

We now introduce regularity assumptions on the SDE.

*Assumption 2.1.* The potentials  $p$  and  $V$  are such that

- (i) for all  $T > 0$ , the symmetric matrix  $M$  is positive definite and its minimum eigenvalue satisfies  $\lambda_{\min}(M) \geq \bar{\lambda} > 0$ ,
- (ii) **other assumptions to be added**

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

**Proposition 2.2.** *Given a continuous stochastic process  $X_t$  which is ergodic with invariant measure  $\mu^\infty$  such that  $\mathbb{E}^{\mu^\infty}(V'_i(\cdot)) \leq C$  for some constant  $C > 0$  and for all  $i = 1, \dots, N$ , the posterior  $\mu(A \mid X_{0:T})$  contracts to the limit of  $\hat{A}(X_{0:T})$  for  $T \rightarrow \infty$ .*

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

$$C_T = T^{-1} (M^{-1} - Q^{-1}),$$

where

$$Q = M + TMC_0M.$$

The eigenvalues of  $Q$  satisfy for all  $i = 1, \dots, N$

$$\lambda_i(Q) = \lambda_i(M) + T\lambda_i(MC_0M) \geq \bar{\lambda} + T\lambda_{\min}(C_0),$$

and therefore we obtain

$$\lambda_{\max}(Q^{-1}) \leq \frac{1}{\bar{\lambda} + T\lambda_{\min}(C_0)}. \quad (2.1)$$

Similarly, we have a bound for the maximum eigenvalue of  $M^{-1}$  given by

$$\lambda_{\max}(M^{-1}) \leq \frac{1}{\bar{\lambda}}, \quad (2.2)$$

which implies that

$$\text{tr } C_T \leq \frac{N}{T} \left( \frac{1}{\bar{\lambda}} + \frac{1}{\bar{\lambda} + T\lambda_{\min}(C_0)} \right),$$

and which therefore allows us to conclude that

$$\lim_{T \rightarrow \infty} \|C_T\|_F = 0.$$

We now consider the mean. Replacing the expression of the maximum likelihood estimator, we obtain

$$\left\| m_T - \hat{A}(X_{0:T}) \right\|_2 = T^{-1} \left\| M^{-1}C_0^{-1}A_0 - Q^{-1}(C_0^{-1}A_0 - h) \right\|_2.$$

Let us remark that (2.1) implies  $\|Q^{-1}\|_F \rightarrow 0$  for  $T \rightarrow \infty$  and (2.2) implies  $\|M^{-1}\|_F \leq C$  for some  $C > 0$  independently of  $T$ . Moreover, the ergodic theorem guarantees that  $\|h\|_2 \leq C$  for  $T$  sufficiently big, and therefore the Cauchy-Schwarz and the triangle inequalities imply

$$\lim_{T \rightarrow \infty} \left\| m_T - \hat{A}(X_{0:T}) \right\|_2 = 0,$$

which proves the desired result.  $\square$

*Remark 2.3.* Proposition 2.2 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 A filtering approach

In the previous section, we have shown that posterior distributions over the drift coefficient of the homogenized equation are asymptotically incorrect if multiscale data are replaced into the expression of the likelihood. Hence, the need of pre-processing the data is highlighted. In particular, we consider a filtering approach. Let  $k: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$  be a kernel function and consider the process  $Z^\varepsilon := (Z_t^{\varepsilon,k}, 0 \leq t \leq T)$  defined by the weighted average

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

where the equality is component-wise in  $\mathbb{R}^d$ . Let  $\beta, \delta > 0$  and let us consider a family of exponential kernel functions defined as

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

where  $C_\beta$  is a normalizing constant given by

$$C_\beta = \beta \Gamma(1/\beta)^{-1},$$

and where  $\Gamma(\cdot)$  is the gamma function. In the following, only in case  $\beta = 1$  a rigorous analysis is carried on. Nonetheless, numerical experiments show that for higher values of  $\beta$  the performances of estimators computed employing the filter are more robust and qualitatively better.

Given a trajectory  $X^\varepsilon := (X_t^\varepsilon, 0 \leq t \leq T)$ , it is relatively inexpensive to compute  $Z^\varepsilon$  from a computational standpoint. In particular, the process  $Z^\varepsilon$  is the convolution of the kernel with the process  $X^\varepsilon$ . Hence, computational tools based on the Fast Fourier Transform (FFT) exist and allow to compute  $Z^\varepsilon$  fast component-wise.

#### 3.1 Ergodic properties of the filter

In the following, we consider the filter with  $\beta = 1$ , i.e.,

$$k(t, s) = \frac{1}{\delta} e^{-\frac{t-s}{\delta}},$$

since in this case it holds

$$dZ_t^\varepsilon = k(0)X_t^\varepsilon dt + \int_0^t \partial_t k(t-s)X_s^\varepsilon ds dt = \frac{1}{\delta} (X_t^\varepsilon - Z_t^\varepsilon) dt.$$

We will treat analytically two different cases. First, the case where the state  $X_t^\varepsilon$  is a vector of  $\mathbb{R}^d$  and we wish to find the

Considering the processes  $X^\varepsilon$  and  $Z^\varepsilon$  together, we obtain a system of stochastic differential equations

$$\begin{aligned} dX_t^\varepsilon &= - \left( \alpha \nabla_x V(X_t^\varepsilon) - \frac{1}{\varepsilon} \nabla_x p \left( \frac{X_t^\varepsilon}{\varepsilon} \right) \right) dt + \sqrt{2\sigma} dW_t, \\ dZ_t^\varepsilon &= \frac{1}{\delta} (X_t^\varepsilon - Z_t^\varepsilon) dt. \end{aligned}$$

Let  $\rho^\varepsilon = \rho^\varepsilon(x, z)$  be the stationary distribution of the joint process  $(X^\varepsilon, Z^\varepsilon)^\top$ . Then  $\rho^\varepsilon$  is the solution of the  $2d$ -dimensional stationary Fokker–Planck equation

$$\begin{aligned} \sigma \Delta_x \rho^\varepsilon(x, z) &+ \left( \alpha \Delta_x V(x) + \frac{1}{\varepsilon^2} \Delta_x p \left( \frac{x}{\varepsilon} \right) + \frac{1}{\delta} \right) \rho^\varepsilon(x, z) \\ &+ \left( \alpha \nabla_x V(x) + \frac{1}{\varepsilon} \nabla_x p \left( \frac{x}{\varepsilon} \right) \right) \cdot \nabla_x \rho^\varepsilon(x, z) + \frac{1}{\delta} (z - x) \cdot \nabla_z \rho^\varepsilon(x, z) = 0. \end{aligned} \tag{3.1}$$

*Example 3.1.* A closed form solution of (3.1) can be obtained in a simple homogenized case. Let  $p(y) = 0$  and the parameters  $\alpha, \sigma$  be replaced respectively by  $A$  and  $\Sigma$ . Then, if  $V(x) = x^2/2$ , equation (3.1) has the analytical solution

$$\rho^0(x, z) = \frac{1}{A_{\rho^0}} \exp \left( -\frac{A}{\Sigma} \frac{x^2}{2} - \frac{1}{\delta \Sigma} \frac{(x - (1 + A\delta)z)^2}{2} \right),$$

where

$$A_{\rho^0} = \int_{\mathbb{R}} \int_{\mathbb{R}} \exp \left( -\frac{A}{\Sigma} \frac{x^2}{2} - \frac{1}{\delta \Sigma} \frac{(x - (1 + A\delta)z)^2}{2} \right) dx dz.$$

This is the density of a multivariate normal distribution  $\mathcal{N}(0, \Gamma)$ , where the covariance matrix is given by

$$\Gamma = \frac{\Sigma}{A(1 + A\delta)} \begin{pmatrix} 1 + A\delta & 1 \\ 1 & 1 \end{pmatrix}.$$

Let us remark that this distribution can be obtained from direct computations involving Gaussian processes. In particular, it is known that  $X_t \sim \mathcal{GP}(\mu_t, \mathcal{C}(t, s))$ , where

$$\mu_t = X_0 e^{-At}, \quad \mathcal{C}(t, s) = \frac{\Sigma}{A} \left( e^{-A|t-s|} - e^{-A(t+s)} \right).$$

The basic properties of Gaussian processes imply that  $Z_t$  is a Gaussian process, and that the couple  $(X_t, Z_t)^\top$  is a Gaussian process, too, whose mean and covariance are computable explicitly.

In a general case, it is not possible to find an explicit solution to (3.1). Nevertheless, it is possible to show some relevant properties of the solution itself, which are summarized in the following Lemma.

**Lemma 3.2.** *Let  $\rho^\varepsilon$  be the **unique solution** of (3.1) and let us write*

$$\rho^\varepsilon(x, z) = \varphi^\varepsilon(x) \psi^\varepsilon(z) R^\varepsilon(x, z), \quad (3.2)$$

where  $\varphi^\varepsilon$  and  $\psi^\varepsilon$  are the marginal densities, i.e.,

$$\varphi^\varepsilon(x) = \int_{\mathbb{R}} \rho^\varepsilon(x, z) dz, \quad \psi^\varepsilon(z) = \int_{\mathbb{R}} \rho^\varepsilon(x, z) dx.$$

Then, it holds

$$\varphi^\varepsilon(x) = \frac{1}{C_{\varphi^\varepsilon}} \exp \left( -\frac{\alpha}{\sigma} V(x) - \frac{1}{\sigma} p \left( \frac{x}{\varepsilon} \right) \right), \quad (3.3)$$

where

$$C_{\varphi^\varepsilon} = \int_{\mathbb{R}} \exp \left( -\frac{\alpha}{\sigma} V(x) - \frac{1}{\sigma} p \left( \frac{x}{\varepsilon} \right) \right) dx.$$

Moreover, it holds

$$\sigma \delta \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi^\varepsilon(x) \psi^\varepsilon(z) \nabla_z V(z) \cdot \nabla_x R^\varepsilon(x, z) dx dz = \mathbb{E}^{\rho^\varepsilon} [((X^\varepsilon)^2 - (Z^\varepsilon)^2) V''(Z^\varepsilon)].$$

*Proof.* Integrating equation (3.1) with respect to  $z$  we obtain the stationary Fokker–Planck equation for the process  $X^\varepsilon$ , i.e.

$$\sigma \Delta_x \varphi^\varepsilon(x) + \left( \alpha \Delta_x V(x) + \frac{1}{\varepsilon^2} \Delta_x p \left( \frac{x}{\varepsilon} \right) \right) \varphi^\varepsilon(x) + \left( \alpha \nabla_x V(x) + \frac{1}{\varepsilon} \nabla_x p \left( \frac{x}{\varepsilon} \right) \right) \cdot \nabla_x \varphi^\varepsilon(x) = 0, \quad (3.4)$$

whose solution is given by

$$\varphi^\varepsilon(x) = \frac{1}{C_{\varphi^\varepsilon}} \exp \left( -\frac{\alpha}{\sigma} V(x) - \frac{1}{\sigma} p \left( \frac{x}{\varepsilon} \right) \right),$$

and which proves (3.3). By integrating equation (3.1) with respect to  $x$  and integrating by parts we obtain

$$\psi^\varepsilon(z) + z \cdot \nabla_z \psi^\varepsilon(z) = \nabla_z \cdot \int_{\mathbb{R}^d} x \rho^\varepsilon(x, z) dx,$$

which can be written as

$$\nabla_z \cdot (z \psi^\varepsilon(z)) = \nabla_z \cdot \int_{\mathbb{R}^d} x \rho^\varepsilon(x, z) dx.$$

Now, since  $\psi^\varepsilon$  is the density of a probability distribution, this implies that

$$z \psi^\varepsilon(z) = \int_{\mathbb{R}^d} x \rho^\varepsilon(x, z) dx,$$

which, replacing the decomposition (3.2), yields

$$z = \int_{\mathbb{R}^d} x \varphi^\varepsilon(x) R^\varepsilon(x, z) dx. \quad (3.5)$$

Replacing the decomposition (3.2) in equation (3.1) and using equation (3.4) we obtain the following equality

$$\sigma \varphi^\varepsilon \psi^\varepsilon \Delta_x R^\varepsilon + \frac{1}{\delta} \varphi^\varepsilon \psi^\varepsilon R^\varepsilon + \sigma \psi^\varepsilon \nabla_x \varphi^\varepsilon \cdot \nabla_x R^\varepsilon + \frac{1}{\delta} \varphi^\varepsilon (z - x) \cdot (R^\varepsilon \nabla_z \psi^\varepsilon + \psi^\varepsilon \nabla_z R^\varepsilon) = 0. \quad (3.6)$$

We now multiply the equation above by  $x$  and integrate with respect to  $x$ . Let us consider some simplifications explicitly. First, an integration by parts yields

$$\sigma \psi^\varepsilon \int_{\mathbb{R}} x \varphi^\varepsilon \partial_{xx}^2 R^\varepsilon dx + \sigma \psi^\varepsilon \int_{\mathbb{R}} x (\varphi^\varepsilon)' \partial_x R^\varepsilon dx = -\sigma \psi^\varepsilon \int_{\mathbb{R}} \varphi^\varepsilon \partial_x R^\varepsilon dx. \quad (3.7)$$

Then, applying (3.5), we have

$$\frac{1}{\delta} \psi^\varepsilon \int_{\mathbb{R}} x \varphi^\varepsilon R^\varepsilon dx = \frac{1}{\delta} z \psi^\varepsilon. \quad (3.8)$$

Moreover, again applying (3.5), we can compute

$$\frac{1}{\delta} (\psi^\varepsilon)' \int_{\mathbb{R}} (z - x) x \varphi^\varepsilon R^\varepsilon dx = \frac{1}{\delta} (\psi^\varepsilon)' \left( z^2 - \int_{\mathbb{R}} x^2 \varphi^\varepsilon R^\varepsilon dx \right). \quad (3.9)$$

Finally, we compute the last term always applying (3.5) obtaining

$$\frac{1}{\delta} \psi^\varepsilon \int_{\mathbb{R}} (z - x) x \varphi^\varepsilon \partial_z R^\varepsilon dx = \frac{1}{\delta} \psi^\varepsilon \left( z - \int_{\mathbb{R}} x^2 \varphi^\varepsilon \partial_z R^\varepsilon dx \right). \quad (3.10)$$

Replacing the equalities (3.7), (3.8), (3.9) and (3.10) into (3.6) we obtain

$$(\psi^\varepsilon)' \left( z^2 - \int_{\mathbb{R}} x^2 \varphi^\varepsilon R^\varepsilon dx \right) + \psi^\varepsilon \left( 2z - \int_{\mathbb{R}} x^2 \varphi^\varepsilon \partial_z R^\varepsilon dx - \delta \sigma \int_{\mathbb{R}} \varphi^\varepsilon \partial_x R^\varepsilon dx \right) = 0.$$

We rewrite the equality above as

$$\begin{aligned} \delta \sigma \psi^\varepsilon(z) \int_{\mathbb{R}} \varphi^\varepsilon(x) \partial_x R^\varepsilon(x, z) dx &= (\psi^\varepsilon)'(z) z^2 - (\psi^\varepsilon)'(z) \int_{\mathbb{R}} x^2 \varphi^\varepsilon(x) R^\varepsilon(x, z) dx \\ &\quad + 2 \psi^\varepsilon(x) z - \psi^\varepsilon \int_{\mathbb{R}} x^2 \varphi^\varepsilon(x) \partial_z R^\varepsilon(x, z) dx. \end{aligned}$$

Multiplying by  $V'(z)$ , integrating with respect to  $z$  and integrating by parts, we obtain

$$\begin{aligned} \sigma \delta \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^\varepsilon(x) \psi^\varepsilon(z) \partial_x R^\varepsilon(x, z) dx dz &= \int_{\mathbb{R}} \int_{\mathbb{R}} (x^2 - z^2) V''(z) \rho_\varepsilon(x, z) dx dz \\ &= \mathbb{E}^{\rho^\varepsilon} [(X^\varepsilon)^2 - (Z^\varepsilon)^2] V''(Z^\varepsilon), \end{aligned}$$

which is the desired result.

**Multi-dimensional:** First

$$\iint (\sigma \varphi^\varepsilon \psi^\varepsilon \Delta_x R^\varepsilon + \sigma \psi^\varepsilon \nabla_x \varphi^\varepsilon \cdot \nabla_x R^\varepsilon) (x \cdot \nabla_z V(z)) \, dx \, dz = -\sigma \iint \varphi \psi \nabla_z V(z) \cdot \nabla_x R \, dx \, dz.$$

Second □

*Remark 3.3.* It could be possible to find a closed-form solution to (3.1) in a more general framework than Example 3.1. Nonetheless, an explicit solution is not necessary for our argument.

### 3.1.1 Convergence of the invariant measure

It is known that the invariant distribution of the process  $X^\varepsilon$  converges weakly to the invariant distribution of the process  $X^0$  solution to (1.2). We now consider the random variable  $(X^\varepsilon, Z^\varepsilon)^\top$  and present an analogous convergence result for the couple.

**Lemma 3.4.** *Let  $\mu^\varepsilon$  be the invariant measure of the couple  $(X^\varepsilon, Z^\varepsilon)^\top$ . Then, the measure  $\mu^\varepsilon$  converges weakly to the measure  $\mu^0(dx, dz) = \rho^0(x, z) \, dx \, dz$ , whose density  $\rho^0$  is the unique solution of the Fokker–Planck equation*

$$\Sigma \partial_{xx}^2 \rho^0(x, z) + \left( AV''(x) + \frac{1}{\delta} \right) \rho^0(x, z) + AV'(x) \partial_x \rho^0(x, z) + \frac{1}{\delta} (z - x) \partial_z \rho^0(x, z) = 0, \quad (3.11)$$

where  $A$  and  $\Sigma$  are the coefficients of the homogenized equation (1.2).

*Proof.* □

We now present an analogous result to Lemma 3.2 for the limit distribution.

**Corollary 3.5.** *Let  $\rho^0$  be the **unique solution** of (3.11) and let us write*

$$\rho^0(x, z) = \varphi^0(x) \psi^0(z) R^0(x, z),$$

where  $\varphi^0$  and  $\psi^0$  are the marginal densities, i.e.,

$$\varphi^0(x) = \int_{\mathbb{R}} \rho^0(x, z) \, dz, \quad \psi^0(z) = \int_{\mathbb{R}} \rho^0(x, z) \, dx.$$

Then, if  $A$  and  $\Sigma$  are the coefficients of the homogenized equation (1.2), it holds

$$\varphi^0(x) = \frac{1}{C_{\varphi^0}} \exp \left( -\frac{A}{\Sigma} V(x) \right), \quad \text{where} \quad C_{\varphi^0} = \int_{\mathbb{R}} \exp \left( -\frac{A}{\Sigma} V(x) \right) \, dx.$$

Moreover, it holds

$$\Sigma \delta \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^0(x) \psi^0(z) \partial_x R^0(x, z) \, dx \, dz = \mathbb{E}^{\rho^0} [((X^0)^2 - (Z^0)^2) V''(Z^0)].$$

*Proof.* The proof is directly obtained from Lemma 3.2 replacing  $p(y) = 0$  and  $\alpha, \sigma$  by  $A, \Sigma$  respectively. □

## 3.2 The estimator

The estimator of the homogenized drift coefficient  $A$  we analyse is given by the formula

$$\widehat{A}^\varepsilon(T) = -\frac{\int_0^T V'(Z_t^\varepsilon) \, dX_t^\varepsilon}{\int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) \, dt}. \quad (3.12)$$



**Theorem 3.6.** Let  $\widehat{A}^\varepsilon(T)$  be defined in (3.12). Then, if **Assumptions**,

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \widehat{A}^\varepsilon(T) = A, \quad \text{a.s.}, \quad (3.13)$$

where  $A$  is the drift coefficient of the homogenized equation (1.1). Moreover, the following central limit theorem holds

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} \sqrt{T} \left( A - \widehat{A}^\varepsilon(T) \right) = \Lambda, \quad \text{in law},$$

where  $\Lambda \sim \mathcal{N}(0, \gamma^2)$  and where

$$\gamma = \sqrt{2\sigma} \frac{\sqrt{\mathbb{E}^{\rho^0} [V'(Z)^2]}}{\mathbb{E}^{\rho^0} [V'(Z)V'(X)]}.$$

*Proof.* Replacing the expression of  $dX_t^\varepsilon$  into (3.12), we get

$$\begin{aligned} \widehat{A}^\varepsilon(T) &= \frac{\alpha \int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) dt + \int_0^T V'(Z_t^\varepsilon) \frac{1}{\varepsilon} p' \left( \frac{X_t^\varepsilon}{\varepsilon} \right) dt - \sqrt{2\sigma} \int_0^T V'(Z_t^\varepsilon) dW_t}{\int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) dt} \\ &= \alpha + \frac{\int_0^T V'(Z_t^\varepsilon) \frac{1}{\varepsilon} p' \left( \frac{X_t^\varepsilon}{\varepsilon} \right) dt}{\int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) dt} - \sqrt{2\sigma} \frac{\int_0^T V'(Z_t^\varepsilon) dW_t}{\int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) dt} \\ &=: \alpha + I_1^\varepsilon(T) - I_2^\varepsilon(T). \end{aligned}$$

We study the terms  $I_1^\varepsilon$  and  $I_2^\varepsilon$  separately. The ergodic theorem applied to  $I_1^\varepsilon$  yields

$$\lim_{T \rightarrow \infty} I_1^\varepsilon(T) = \frac{\mathbb{E}^{\rho^\varepsilon} \left[ \frac{1}{\varepsilon} V'(Z^\varepsilon) p' \left( \frac{X^\varepsilon}{\varepsilon} \right) \right]}{\mathbb{E}^{\rho^\varepsilon} [V'(Z^\varepsilon) V'(X^\varepsilon)]}, \quad \text{a.s.} \quad (3.14)$$

Let us consider the numerator of the expression above. Due to Lemma 3.2 and integrating by parts, we have

$$\begin{aligned} \mathbb{E}^{\rho^\varepsilon} \left[ V'(Z^\varepsilon) \frac{1}{\varepsilon} p' \left( \frac{X^\varepsilon}{\varepsilon} \right) \right] &= \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \frac{1}{\varepsilon} p' \left( \frac{x}{\varepsilon} \right) \frac{1}{C_{\varphi^\varepsilon}} e^{-\frac{\alpha}{\sigma} V(x)} e^{-\frac{1}{\sigma} p(\frac{x}{\varepsilon})} \psi^\varepsilon(z) R^\varepsilon(x, z) dx dz \\ &= -\sigma \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{d}{dx} \left( e^{-\frac{1}{\sigma} p(\frac{x}{\varepsilon})} \right) \frac{1}{C_{\varphi^\varepsilon}} e^{-\frac{\alpha}{\sigma} V(x)} V'(z) \psi^\varepsilon(z) R^\varepsilon(x, z) dx dz \\ &= \sigma \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{C_{\varphi^\varepsilon}} e^{-\frac{1}{\sigma} p(\frac{x}{\varepsilon})} \partial_x \left( e^{-\frac{\alpha}{\sigma} V(x)} R^\varepsilon(x, z) \right) V'(z) \psi^\varepsilon(z) dx dz, \end{aligned}$$

which implies

$$\begin{aligned} \mathbb{E}^{\rho^\varepsilon} \left[ \frac{1}{\varepsilon} V'(Z^\varepsilon) p' \left( \frac{X^\varepsilon}{\varepsilon} \right) \right] &= -\alpha \int_{\mathbb{R}} \int_{\mathbb{R}} V'(x) V'(z) \rho^\varepsilon(x, z) dx dz \\ &\quad + \sigma \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^\varepsilon(x) \psi^\varepsilon(z) \partial_x R^\varepsilon(x, z) dx dz \\ &= -\alpha \mathbb{E}^{\rho^\varepsilon} [V'(Z^\varepsilon) V'(X^\varepsilon)] + \sigma \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^\varepsilon(x) \psi^\varepsilon(z) \partial_x R^\varepsilon(x, z) dx dz. \end{aligned}$$

Replacing the equality above into (3.14), we obtain

$$\lim_{T \rightarrow \infty} I_1^\varepsilon(T) = -\alpha + \frac{\sigma \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^\varepsilon(x) \psi^\varepsilon(z) \partial_x R^\varepsilon(x, z) dx dz}{\mathbb{E}^{\rho^\varepsilon} [V'(Z^\varepsilon) V'(X^\varepsilon)]}, \quad \text{a.s.}$$

Due to Lemma 3.2, we therefore have

$$\lim_{T \rightarrow \infty} I_1^\varepsilon(T) = -\alpha + \frac{\mathbb{E}^{\rho^\varepsilon} [(X^2 - Z^2) V''(Z)]}{\delta \mathbb{E}^{\rho^\varepsilon} [V'(Z) V'(X)]}, \quad \text{a.s.}$$

We now pass to the limit as  $\varepsilon$  goes to zero and

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} I_1^\varepsilon(T) = -\alpha + \frac{\mathbb{E}^{\rho^0}[(X^2 - Z^2)V''(Z)]}{\delta \mathbb{E}^{\rho^0}[V'(Z)V'(X)]}, \quad \text{a.s.},$$

where the function  $\rho^0$  is the density of the limit invariant distribution for  $\varepsilon \rightarrow 0$ , as in Lemma 3.4. Due to Lemma 3.5, we have

$$\frac{1}{\delta} \mathbb{E}^{\rho^0}[(X^2 - Z^2)V''(Z)] = \Sigma \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^0(x) \psi^0(z) \partial_x R^0(x, z) \, dx \, dz,$$

and finally, an integration by parts yields

$$\begin{aligned} \mathbb{E}^{\rho^0}[V'(Z)V'(X)] &= \int_{\mathbb{R}} \int_{\mathbb{R}} V'(x)V'(z) \frac{1}{C_{\varphi^0}} e^{-\frac{A}{\Sigma}V(x)} \psi^0(z) R^0(x, z) \, dx \, dz \\ &= -\frac{\Sigma}{A} \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \frac{d}{dx} \left( \frac{1}{C_{\varphi^0}} e^{-\frac{A}{\Sigma}V(x)} \right) \psi^0(z) R^0(x, z) \, dx \, dz \\ &= \frac{\Sigma}{A} \int_{\mathbb{R}} \int_{\mathbb{R}} V'(z) \varphi^0(x) \psi^0(z) \partial_x R^0(x, z) \, dx \, dz. \end{aligned}$$

We can therefore conclude that

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} I_1(T) = -\alpha + A, \quad \text{a.s.} \quad (3.15)$$

We now consider the second term  $I_2$ , and rewrite it as

$$I_2^\varepsilon(T) = -\sqrt{2\sigma} \frac{\frac{1}{T} \int_0^T V'(Z_t^\varepsilon)^2 \, dt}{\frac{1}{T} \int_0^T V'(Z_t^\varepsilon) V'(X_t^\varepsilon) \, dt} \frac{\int_0^T V'(Z_t^\varepsilon) \, dW_t}{\int_0^T V'(Z_t^\varepsilon)^2 \, dt} =: -\sqrt{2\sigma} I_{2,1}^\varepsilon(T) I_{2,2}^\varepsilon(T).$$

The ergodic theorem yields

$$\lim_{T \rightarrow \infty} I_{2,1}^\varepsilon(T) = \frac{\mathbb{E}^{\rho^\varepsilon}[V'(Z^\varepsilon)^2]}{\mathbb{E}^{\rho^\varepsilon}[V'(Z^\varepsilon)V'(X^\varepsilon)]} =: M^\varepsilon,$$

where  $M^\varepsilon$  is bounded uniformly in  $\varepsilon$ . Moreover, **the strong law of large numbers for martingales implies**

$$\lim_{T \rightarrow \infty} I_{2,2}^\varepsilon(T) = 0, \quad \text{a.s.},$$

independently of  $\varepsilon$ . Therefore

$$\lim_{\varepsilon \rightarrow 0} \lim_{T \rightarrow \infty} I_2^\varepsilon(T) = 0, \quad \text{a.s.},$$

which, together with (3.15), proves (3.13).  $\square$

## 4 Numerical experiments

## 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.