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 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)

$$\mathrm{d}X_t^\varepsilon = -V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right) + \sqrt{2\sigma}\,\mathrm{d}W_t, \tag{1.1}$$

where $\sigma > 0$ and W_t is a standard one-dimensional Brownian motion. The functions $V, p\colon \mathbb{R} \to \mathbb{R}$ 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 $V_i \colon \mathbb{R} \to \mathbb{R}$. 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

$$\mathrm{d}X_t^0 = -KV'(X)\,\mathrm{d}t + \sqrt{2K\sigma}\,\mathrm{d}W_t, \tag{1.2}$$

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

$$K = \int_0^L (1 + \Phi'(y))^2\,\mu(\mathrm{d}y), \tag{1.3}$$

with

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

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

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

endowed with periodic boundary conditions.

2 Posterior formal computations

Let us denote by $A \in \mathbb{R}^N$ the vector of the coefficients $A_i \coloneqq K\alpha_i$ 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 $X_{0:T} \coloneqq (X_t, 0 \le t \le T)$. 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 μ_0, and where Girsanov's formula allows to write the likelihood as

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

In the formula above, we denote by $\mathbf{V} \colon \mathbb{R} \to \mathbb{R}^N$ 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

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

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 $M\colon \mathbb{R} \to \mathbb{R}^{N\times N}$ whose entries are given by

$$M_{ij} = \frac{1}{T}\int_0^T V_i'(X_t)\, V_j'(X_t)\,\mathrm{d}t, \quad i,j = 1,\dots,N,$$

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

$$h_i = \frac{1}{T}\int_0^T V_i'(X_t)\,\mathrm{d}X_t, \quad i = 1,\dots,N.$$

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

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

We are now interested in the limit of the posterior distribution for $T \to \infty$. Let us first define the maximum likelihood estimator (MLE) $\widehat{A}(X_{0:T})$ 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

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

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 μ^∞ 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 $\widehat{A}(X_{0:T})$ for $T \to \infty$.*

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

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

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

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

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

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

which implies that

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

and which therefore allows us to conclude that

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

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

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

Let us remark that (2.1) implies $\left\|Q^{-1}\right\|_F \to 0$ for $T \to \infty$ and (2.2) implies $\left\|M^{-1}\right\|_F \leq C$ 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

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

which proves the desired result. □

Remark 2.3. Proposition 2.2 guarantees that in the asymptotic limit of $T \to \infty$, 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 $N = 1$

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 $\mathcal{N}(\bar{A}_T, \sigma_T^2)$ where

$$\begin{aligned}
\bar{A}_T &= \frac{A_0}{1 + \sigma_0^2 \int_0^T V'(X_t)^2 \,\mathrm{d}t} - \frac{\sigma_0^2 \int_0^T V'(X_t)\,\mathrm{d}X_t}{1 + \sigma_0^2 \int_0^T V'(X_t)^2\,\mathrm{d}t}, \\
\sigma_T^{-2} &= \sigma_0^{-2} + \int_0^T V'(X_t)^2 \,\mathrm{d}t.
\end{aligned}$$

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 $X_{0:T}^\varepsilon \coloneqq (X_t^\varepsilon, 0 \leq t \leq T)$ 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 3.1. *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*

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

Proof. Proposition 2.2 guarantees that

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

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

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

which completes the proof. □

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

3.2 Success with pre-processing

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, let $k \colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ be a kernel function and consider the process $(Z_t^{\varepsilon,k}, 0 \le t \le T)$ defined by the weighted average

$$Z_t^{\varepsilon,k} := \int_0^t k^\varepsilon(t,s) X_s^\varepsilon \,\mathrm{d}s. \tag{3.1}$$

Substituting $Z_t^{\varepsilon,k}$ in the likelihood term, we obtain

$$\log p(A \mid Z_t^{\varepsilon,k}) = -A \int_0^T V'(Z_t^{\varepsilon,k}) \,\mathrm{d}Z_t^{\varepsilon,k} - \frac{A^2}{2} \int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t, \tag{3.2}$$

and the MLE of the drift coefficient based on transformed data is therefore given by

$$\widehat{A}_T(Z_{0:T}^{\varepsilon,k}) = -\frac{\int_0^T V'(Z_t^{\varepsilon,k}) \,\mathrm{d}Z_t^{\varepsilon,k}}{\int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t}. \tag{3.3}$$

In order to obtain a satisfactory estimator, the kernel k has to satisfy a series of conditions, which are listed in the following assumption.

Assumption 3.2. The kernel $k \colon \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ either satisfies

(i) (regularity) for all $s,t > 0$, $k^\varepsilon \in \mathcal{C}^0(\mathbb{R}^+ \times \mathbb{R}^+; \mathbb{R})$, $(t,s) \mapsto \partial_t k^\varepsilon(t,s) \in \mathcal{C}^0(\mathbb{R}^+ \times \mathbb{R}^+; \mathbb{R})$

(ii) (sign) for all $s,t > 0$, $k^\varepsilon(t,s) \ge 0$ and $\partial_t k^\varepsilon(t,s) \le 0$,

(iii) (divergence–free) for all $s,t > 0$, $\partial_t k^\varepsilon(t,s) + \partial_s k^\varepsilon(t,s) = 0$,

(iv) (normalization) there exists a function $\varphi(t,\varepsilon)$ such that $0 \le \varphi(t,\varepsilon) < 1$ for all $t, \varepsilon > 0$ and satisfying for all $p > 0$

$$\lim_{\varepsilon \to 0} \lim_{T \to \infty} \int_0^T \varphi(t,\varepsilon)^p \,\mathrm{d}t = 0,$$

such that

$$\int_0^t k^\varepsilon(t,s) \,\mathrm{d}s = 1 - \varphi(t,\varepsilon),$$

(v) (time-invariance) for $s,t > 0$ and any $r > 0$ it holds $k^\varepsilon(t+r, s+r) = k^\varepsilon(t,s)$. Moreover, there exists $\zeta \in (0,1)$ such that

$$k(t,t) \le C\varepsilon^{-\zeta},$$

for any $t > 0$.

(vi) (decay) there exists a function $\psi(t,\varepsilon)$ satisfying for all $p > 0$

$$\lim_{\varepsilon \to 0} \lim_{T \to \infty} \int_0^T \psi(t,\varepsilon)^p = 0,$$

such that it holds

$$\int_0^t (t-s)^p\, k^\varepsilon(t,s)\,\mathrm{d}s = C_p \psi(t,\varepsilon)^p.$$

where $C_p > 0$ is uniformly bounded with respect to ε and t. Moreover,

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\int_0^T tk(t,0)\,\mathrm{d}t = 0.$$

We now give an example of kernels satisfying the assumptions above.

Example 3.3. Let $0<\varepsilon<1$, $\delta = \varepsilon^\zeta$ for $\zeta \in (0,1)$, $\beta \geq 1$ 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 all β there exists C_β such that (i)–(vi) are all true. Let us consider $\beta = 1$, for which we have $C_\beta = 1$. Assumptions (i), (ii) and (iii) are trivially satisfied. Concerning the normalization assumption (iv), we get

$$\frac{1}{\delta}\int_0^t e^{-(t-s)/\delta}\,\mathrm{d}s = 1 - e^{-t/\delta},$$

so that $\varphi(t,\varepsilon) = e^{-t/\delta}$. Moreover

$$\int_0^T e^{-t/\delta}\,\mathrm{d}t = \delta\left(1 - e^{-T/\delta}\right),$$

showing that (iv) is satisfied. Since $k^\varepsilon(t,s) = k^\varepsilon(t-s)$, time-invariance is guaranteed. Moreover, $k(t,t) = \delta^{-1}$ for all t, which verifies (v). For (vi), we have

$$\frac{1}{\delta}\int_0^t (t-s)^p e^{-(t-s)/\delta}\,\mathrm{d}s = \delta^p\,\gamma(p+1, t/\delta),$$

where γ is the lower incomplete Gamma function, i.e., for all $t>0$

$$\gamma(p+1,t) := \int_0^t s^p e^{-s}\,\mathrm{d}s.$$

Hence, noticing that for all t and δ it holds $\gamma(p+1,t/\delta) \leq \Gamma(p+1)$, where Γ is the (complete) Gamma function, we can choose $\psi(t,\varepsilon) = \delta$ and $C_p = \gamma(p+1,t/\delta)$. Finally,

$$\frac{1}{\delta}\int_0^T te^{-t/\delta}\,\mathrm{d}t = \delta - e^{-T/\delta}(T+\delta),$$

which shows that (vi) holds. For $\beta = 2$ and choosing $C_2 = 2/\sqrt{\pi}$ one can similarly verify that all the assumptions hold.

Remark 3.4. In the following, we will employ the equality

$$\int_0^t (t-s)\partial_t k^\varepsilon(t,s)\,\mathrm{d}s = -1 + \varphi(t,\varepsilon) + tk(t,0), \tag{3.4}$$

which is simply implied by assumptions (iii) and (iv). Moreover, for $p \geq 1$ we will employ the equality

$$\int_0^t (t-s)^p \partial_t k^\varepsilon(t,s)\,\mathrm{d}s = k^\varepsilon(t,0)t^p - p\,C_{p-1}\psi(t,\varepsilon)^{p-1},$$

which is a consequence of (iii) and (vi).

We can now introduce the main result of this section.

Theorem 3.5. *Under Assumption 2.1, 3.2 and 3.12 and given the Gaussian prior $\mu_0 = \mathcal{N}(A_0, \sigma_0^2)$, let $\mu^{\varepsilon,k} = \mathcal{N}(\bar{A}_T^{\varepsilon,k}, \sigma_T^2)$ be the posterior distribution obtained employing the likelihood function* (3.2). *Then, if $T = \varepsilon^{-\gamma}$ for $\gamma > 0$, it holds*

$$\lim_{\varepsilon\to 0} \bar{A}_T^{\varepsilon,k} = A,$$
$$\lim_{\varepsilon\to 0} \sigma_T^2 = 0,$$

i.e., the posterior concentrates on the drift coefficient of the homogenized process.

3.2.1 Estimates on the transformed process

It is useful 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

$$
\begin{aligned}
\mathrm{d}X_t^\varepsilon &= -\alpha V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(Y_t^\varepsilon\right) + \sqrt{2\sigma}\,\mathrm{d}W_t, \\
\mathrm{d}Y_t^\varepsilon &= -\frac{\alpha}{\varepsilon}V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon^2}p'\left(Y_t^\varepsilon\right) + \sqrt{\frac{2\sigma}{\varepsilon^2}}\,\mathrm{d}W_t.
\end{aligned}
$$

We now present some estimates involving the process $Z_t^{\varepsilon,k}$.

Lemma 3.6. *The process $Z_t^{\varepsilon,k}$ defined in* (3.1) *satisfies, under assumptions (i)–(vi)*

$$
\mathrm{d}Z_t^{\varepsilon,k} = k^\varepsilon(t,0)X_t^\varepsilon\,\mathrm{d}t + \int_0^t \partial_t k^\varepsilon(t,s)(X_s^\varepsilon - X_t^\varepsilon)\,\mathrm{d}s\,\mathrm{d}t.
$$

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

$$
\frac{\mathrm{d}Z_t^{\varepsilon,k}}{\mathrm{d}t} = k^\varepsilon(t,t)X_t^\varepsilon + \int_0^t \partial_t k^\varepsilon(t,s)X_s^\varepsilon\,\mathrm{d}s.
$$

Adding and subtracting X_t^ε inside the integral yields

$$
\frac{\mathrm{d}Z_t^{\varepsilon,k}}{\mathrm{d}t} = \int_0^t \partial_t k^\varepsilon(t,s)(X_s^\varepsilon - X_t^\varepsilon)\,\mathrm{d}s + \left(k^\varepsilon(t,t) + \int_0^t \partial_t k^\varepsilon(t,s)\,\mathrm{d}s\right)X_t^\varepsilon.
$$

Due to assumption (iii), we have

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

which implies the desired result. □

We now estimate the deviation of the process $Z_t^{\varepsilon,k}$ from the original process X_t^ε.

Lemma 3.7. *Under Assumption 2.1 and Assumption 3.2 and if X_0^ε is distributed following its invariant distribution μ^ε, then it holds for all $p \geq 1$*

$$
\mathbb{E}^{\mu^\varepsilon}\left|Z_t^{\varepsilon,k}\right|^p \leq C, \tag{3.5}
$$

where C is a constant independent of ε. Moreover, it holds

$$
\mathbb{E}^{\mu^\varepsilon}\left|Z_t^{\varepsilon,k} - X_t^\varepsilon\right|^p \leq C\left(\psi(t,\varepsilon)^p + \psi(t,\varepsilon)^{p/2} + \varphi(t,\varepsilon)^p + \varepsilon^p\right) \tag{3.6}
$$

Proof. Let us first consider (3.5) and remark that

$$
\int_0^t \frac{k(t,s)}{1-\varphi(t,\varepsilon)}\,\mathrm{d}s = 1,
$$

which implies that $K_t^\varepsilon(\mathrm{d}s) := k(t,s)/(1-\varphi(t,\varepsilon))\,\mathrm{d}s$ is a probability measure on $(0,t)$. Hence, replacing the definition of $Z_t^{\varepsilon,k}$ and applying Jensen's inequality, we have for all $p \geq 1$

$$
\begin{aligned}
\mathbb{E}^{\mu^\varepsilon}\left|Z_t^{\varepsilon,k}\right|^p &= (1-\varphi(t,\varepsilon))^p\,\mathbb{E}^{\mu^\varepsilon}\left|\int_0^t X_s^\varepsilon\,K_t^\varepsilon(\mathrm{d}s)\right|^p \\
&\leq \int_0^t \mathbb{E}^{\mu^\varepsilon}\left|X_s^\varepsilon\right|^p K_t^\varepsilon(\mathrm{d}s) \\
&\leq C\int_0^t K_t^\varepsilon(\mathrm{d}s) = C,
\end{aligned}
$$

where we exploited the fact that $(1-\varphi(t,\varepsilon)) \leq 1$, and that $\mathbb{E}^{\mu^\varepsilon} |X(t)|^p \leq C$ for all $p \geq 1$ (see [2, Corollary 5.4]). We now consider (3.6). Adding and subtracting X_t^ε inside the integral and applying Jensen's inequality as above, it holds

$$
\begin{aligned}
\mathbb{E}^{\mu^\varepsilon} \left| Z_t^{\varepsilon,k} - X_t^\varepsilon \right|^p &\leq C\, \mathbb{E}^{\mu^\varepsilon} \left| \int_0^t (X_s^\varepsilon - X_t^\varepsilon) k(t,s) \,\mathrm{d}s \right|^p + C\varphi(t,\varepsilon)^p\, \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon|^p \\
&\leq C \int_0^t \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon - X_s^\varepsilon|^p \; K_t^\varepsilon(\mathrm{d}s) + C\varphi(t,\varepsilon)^p\, \mathbb{E}^{\mu^\varepsilon} |X_t^\varepsilon|^p =: I_1 + I_2,
\end{aligned}
$$

where C depends only on p. Let us now consider I_1. Due to [2, Lemma 6.1], we have

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

which, combined with (vi), implies

$$
I_1 \leq C(\psi(t,\varepsilon)^p + \psi(t,\varepsilon)^{p/2} + \varepsilon^p).
$$

Finally the boundedness of $\mathbb{E}^{\mu^\varepsilon} |X(t)|^p$ yields

$$
I_2 \leq C\varphi(t,\varepsilon)^p,
$$

which implies the desired result. $\square$

Lemma 3.8. *Under Assumption 3.2, the process $Z_t^{\varepsilon,k}$ admits the representation*

$$
\begin{aligned}
\frac{\mathrm{d}Z_t^{\varepsilon,k}}{\mathrm{d}t} =& \alpha \int_0^t \int_s^t \partial_t k^\varepsilon(t,s)\, V'(X_r^\varepsilon) \left(1 + \Phi'(Y_r^\varepsilon)\right) \mathrm{d}r \,\mathrm{d}s \\
&- \sqrt{2\sigma} \int_0^t \int_s^t \partial_t k^\varepsilon(t,s) \big(1 + \Phi'(Y_r^\varepsilon)\big) \,\mathrm{d}W_r \,\mathrm{d}s \\
&+ \varepsilon \int_0^t \partial_t k^\varepsilon(t,s) \big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big) \,\mathrm{d}s + k^\varepsilon(t,0) X_t^\varepsilon.
\end{aligned}
$$

Proof. An application of the Itô formula shows that the process X_t^ε satisfies for all $s,t > 0$

$$
\begin{aligned}
X_t^\varepsilon - X_s^\varepsilon =& -\alpha \int_s^t V'(X_r^\varepsilon) \big(1 + \Phi'(Y_r^\varepsilon)\big) \,\mathrm{d}r \\
&+ \sqrt{2\sigma} \int_s^t \big(1 + \Phi'(Y_r^\varepsilon)\big) \,\mathrm{d}W_r - \varepsilon \big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big),
\end{aligned}
$$

see e.g. [2, Equation (5.8)]. Replacing the equality above into the decomposition given by Lemma 3.6 yields the desired result. $\square$

3.2.2 Ergodic properties of the transformed process

In this section we consider the transformed process and in particular the properties of the long time mean

$$
\mathcal{M}_T^{\varepsilon,k}(f) \coloneqq \frac{1}{T} \int_0^T f(Z_t^{\varepsilon,k}) \,\mathrm{d}t,
$$

where $f \colon \mathbb{R} \to \mathbb{R}$ is a smooth function. In the same spirit, we adopt the notation

$$
\mathcal{M}_T^{\varepsilon}(f) \coloneqq \frac{1}{T} \int_0^T f(X_t^\varepsilon) \,\mathrm{d}t,
$$

for the solution of (1.1). We can prove a result on long-time averages of the transformed process $Z_t^{\varepsilon,k}$.

Lemma 3.9. *Under Assumption 2.1 and Assumption 3.2, let $f\colon \mathbb{R}\to\mathbb{R}$ be Lipschitz continuous and polynomially bounded. Then, it holds in probability*

$$\lim_{T\to\infty} \mathcal{M}_T^{\varepsilon,k} = \mathbb{E}^{\mu^\varepsilon}(f) + \beta(\varepsilon),$$

where the function β satisfies

$$\lim_{\varepsilon\to 0} |\beta(\varepsilon)| = 0,$$

in probability.

Proof. We decompose

$$\lim_{T\to\infty} \mathcal{M}_T^{\varepsilon,k}(f) = \lim_{T\to\infty} \mathcal{M}_T^{\varepsilon}(f) + \lim_{T\to\infty} (\mathcal{M}_T^{\varepsilon,k}(f) - \mathcal{M}_T^{\varepsilon}(f)).$$

For the first term, the ergodic theorem yields

$$\lim_{T\to\infty} \mathcal{M}_T^{\varepsilon}(f) = \mathbb{E}^{\mu^\varepsilon}(f),$$

which holds almost surely, and therefore in probability. Let us denote

$$\beta(\varepsilon) \coloneqq \lim_{T\to\infty} (\mathcal{M}_T^{\varepsilon,k}(f) - \mathcal{M}_T^{\varepsilon}(f)).$$

Due to Lemma 3.7, we have for a constant $C > 0$ independent of ε, which changes its value from line to line

$$\begin{aligned}
\mathbb{E}^{\mu^\varepsilon} \left| \mathcal{M}_T^{\varepsilon,k}(f) - \mathcal{M}_T^{\varepsilon}(f) \right| &\leq \frac{1}{T} \int_0^T \mathbb{E}^{\mu^\varepsilon} \left| f(Z_t^{\varepsilon,k}) - f(X_t^\varepsilon) \right| \,\mathrm{d}t \\
&\leq \frac{C}{T} \int_0^T \mathbb{E}^{\mu^\varepsilon} \left| Z_t^{\varepsilon,k} - X_t^\varepsilon \right| \,\mathrm{d}t \\
&\leq \frac{C}{T} \int_0^T \left(\psi(t,\varepsilon) + \psi(t,\varepsilon)^{1/2} + \varphi(t,\varepsilon) + \varepsilon \right) \mathrm{d}t.
\end{aligned}$$

Let us remark that for f polynomially bounded

$$\mathbb{E}^{\mu^\varepsilon} |f(X_t^\varepsilon)| \leq C,$$

and likewise for $Z_t^{\varepsilon,k}$ due to Lemma 3.7. Hence,

$$\mathbb{E}^{\mu^\varepsilon} \left| \mathcal{M}_T^{\varepsilon,k}(f) - \mathcal{M}_T^{\varepsilon}(f) \right| \leq \frac{C}{T} \int_0^T \left(\mathbb{E}^{\mu^\varepsilon} \left| Z_t^{\varepsilon,k} \right| + \mathbb{E}^{\mu^\varepsilon} \left| X_t^\varepsilon \right| \right) \mathrm{d}t \leq 2C.$$

Therefore, by Assumption 3.2 and applying the dominated convergence theorem, we obtain

$$\lim_{\varepsilon\to 0} \mathbb{E}^{\mu^\varepsilon} |\beta(\varepsilon)| \leq \lim_{\varepsilon\to 0} \lim_{T\to\infty} \frac{C}{T} \int_0^T \left(\psi(t,\varepsilon) + \psi(t,\varepsilon)^{1/2} + \varphi(t,\varepsilon) + \varepsilon \right) \mathrm{d}t = 0.$$

which shows that the absolute value of $\beta(\varepsilon)$ tends to zero in L^1, and therefore in probability, and which concludes the proof. □

Remark 3.10. The result presented in Lemma 3.9 above does not imply ergodicity of the process $Z_t^{\varepsilon,k}$. It would be indeed relevant to find a modified measure $\mu^{\varepsilon,k}$ such that

$$\lim_{T\to\infty} \mathcal{M}_T^{\varepsilon,k}(f) = \mathbb{E}^{\mu^{\varepsilon,k}}(f),$$

almost surely. We suspect that such a measure, if it existed, could be proved to be arbitrarily close to the invariant measure μ^ε of the process X^ε. Nevertheless, the convergence in probability shown in Lemma 3.9 is sufficient for our analysis to hold.

3.2.3 Further estimates

In this section, we present estimates which are helpful in the proof of the main result.

Lemma 3.11. *Under Assumption 2.1 and Assumption 3.2, it holds in law*

$$\alpha \int_s^t V'(X_r^\varepsilon)\left(1+\Phi'(Y_r^\varepsilon)\right) \mathrm{d}r = A(t-s)V'(Z_t^{\varepsilon,k}) + R(\varepsilon, t-s),$$

where for all $p \geq 1$ it holds

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

and where

$$\vartheta(t,\varepsilon) = \psi(t,\varepsilon) + \psi(t,\varepsilon)^{1/2} + \varphi(t,\varepsilon) + \varepsilon.$$

Proof. Let us denote $\Psi_t^\varepsilon \coloneqq 1 + \Phi'(Y_t^\varepsilon)$. Then

$$\begin{aligned}
\mathbb{E}^{\mu^\varepsilon} |R(\varepsilon, t-s)|^p &= \mathbb{E}^{\mu^\varepsilon} \left| \int_s^t \alpha V'(X_r^\varepsilon)\Psi_r^\varepsilon \,\mathrm{d}r - (t-s)AV'(Z_t^{\varepsilon,k}) \right|^p \\
&\leq C\,\mathbb{E}^{\mu^\varepsilon} \left| V'(Z_t^{\varepsilon,k}) \int_s^t \left(\alpha\Psi_r^\varepsilon - A\right)\,\mathrm{d}r \right|^p \\
&\quad + C\,\mathbb{E}^{\mu^\varepsilon} \left| \int_s^t \alpha\left(V'(X_r^\varepsilon) - V'(Z_t^{\varepsilon,k})\right)\Psi_r^\varepsilon \,\mathrm{d}r \right|^p \eqqcolon I_1 + I_2,
\end{aligned}$$

where C is a constant depending only on p. Let us first consider the term I_1. It is possible to show [2, Lemma 5.6], that

$$I_1^{1/p} \leq C(\varepsilon^2 + \varepsilon(t-s) + \varepsilon(t-s)^{1/2}).$$

Let us now consider I_2. Under Assumption 2.1 and applying Hölder's inequality we obtain

$$I_2 \leq C(t-s)^{p-1}\left(\int_s^t \mathbb{E}^{\mu^\varepsilon}\left|X_t^\varepsilon - Z_t^{\varepsilon,k}\right|^p \,\mathrm{d}r + \int_s^t \mathbb{E}^{\mu^\varepsilon}\left|X_t^\varepsilon - X_r^\varepsilon\right|^p \,\mathrm{d}r\right),$$

for a constant C depending only on p. An application of Lemma 3.7 and of [2, Lemma 6.1] finally yields

$$I_2^{1/p} \leq C(t-s)\left(\vartheta(t,\varepsilon) + (t-s) + (t-s)^{1/2}\right),$$

which, combined with the bound for I_1, gives the desired result. □

In order to prove the next result, it is necessary to introduce a further assumption on the kernel, which extends Assumption 3.2.(vi).

Assumption 3.12. There exists $\kappa \in (0,1)$ such that $k(t,s)$ satisfies for all $t > 0$

$$\int_0^t (t-s)^{-1/2} k(t,s)\,\mathrm{d}s \leq C\varepsilon^{-\kappa},$$

where C is a constant independent of ε and t.

Remark 3.13. The kernels belonging to the exponential family introduced in Example 3.3 satisfy the assumption above. For instance, choosing $\beta = 1$ we obtain

$$\frac{1}{\delta}\int_0^t (t-s)^{-1/2} e^{-(t-s)/\delta}\,\mathrm{d}s = \sqrt{\frac{\pi}{\delta}}\,\mathrm{erf}\left(\frac{t}{\delta}\right),$$

where erf$(\cdot)$ is the cumulative distribution function of a standard Gaussian random variable. Since $\delta = \varepsilon^{\zeta}$ with $\zeta \in (0,1)$ and $0 < \mathrm{erf}(x) \leq 1$ for all $x > 0$, Assumption 3.12 holds in this case with $C = \sqrt{\pi}$ and $\kappa = \zeta/2$. Moreover, Assumption 3.2.(iii) guarantees that

$$-\int_0^t (t-s)^{1/2} \partial_t k(t,s) \, \mathrm{d}s \leq C\varepsilon^{-\kappa},$$

which will be employed in the following Lemma.

Lemma 3.14. *Under Assumption 3.2 and Assumption 3.12, let $R(\varepsilon, t-s)$ be defined in Lemma 3.11. Then, it holds for all $T > 0$*

$$\mathbb{E}^{\mu^\varepsilon} \left| \int_0^t \partial_t k^\varepsilon(t,s) \, R(\varepsilon, t-s) \, \mathrm{d}s \right| \leq C \left(\varepsilon^{2-\zeta} + \varepsilon^{1-\kappa} + \vartheta(t,\varepsilon) + \psi(t,\varepsilon)^{1/2} + \psi(t,\varepsilon) \right),$$

where $\vartheta(t,\varepsilon)$ is defined in Lemma 3.11.

Proof. Due to Lemma 3.11, Assumption 3.2.(ii) and in light of Remark 3.4 we have

$$\begin{aligned}
\mathbb{E}^{\mu^\varepsilon} \left| \int_0^t \partial_t k^\varepsilon(t,s) \, R(\varepsilon, t-s) \, \mathrm{d}s \right| &\leq -C \int_0^t \partial_t k^\varepsilon(t,s) \Big(\varepsilon^2 + (t-s)^{1/2}\varepsilon + (t-s)\, \vartheta(t,\varepsilon) \\
&\qquad + (t-s)^{3/2} + (t-s)^2 \Big) \, \mathrm{d}s \\
&\leq -C \int_0^t \partial_t k^\varepsilon(t,s) \Big(\varepsilon^2 + (t-s)^{1/2}\varepsilon + (t-s)\, \vartheta(t,\varepsilon) \Big) \, \mathrm{d}s \\
&\quad + \frac{3}{2} C_{\frac{1}{2}} \psi(t,\varepsilon)^{1/2} + 2C_1 \psi(t,\varepsilon) - (t^{3/2} + t^2) k^\varepsilon(t,0) \\
&\leq -C \int_0^t \partial_t k^\varepsilon(t,s) \Big(\varepsilon^2 + (t-s)^{1/2}\varepsilon + (t-s)\, \vartheta(t,\varepsilon) \Big) \, \mathrm{d}s \\
&\quad + C \Big(\psi(t,\varepsilon)^{1/2} + \psi(t,\varepsilon) \Big).
\end{aligned}$$

We now observe that due to the assumptions on k there exists a constant $C > 0$ such that

$$\begin{aligned}
-\int_0^t \partial_t k^\varepsilon(t,s) \varepsilon^2 \, \mathrm{d}s &= \varepsilon^2 \int_0^t \partial_s k^\varepsilon(t,s) \, \mathrm{d}s \\
&\leq \varepsilon^2 k^\varepsilon(t,t) \\
&\leq C\varepsilon^{2-\zeta}, \\
-\int_0^t \partial_t k^\varepsilon(t,s)(t-s)^{1/2}\varepsilon \, \mathrm{d}s &\leq C\varepsilon^{1-\kappa}, \\
-\vartheta(t,\varepsilon) \int_0^t \partial_t k^\varepsilon(t,s)(t-s) \, \mathrm{d}s &\leq \vartheta(t,\varepsilon).
\end{aligned}$$

These bounds yield the desired result. □

3.2.4 Estimation of the diffusion coefficient

In this section, we compute the quadratic variation of the process $Z_t^{\varepsilon,k}$, which, as we will show, can be employed as an estimator for the diffusion coefficient of the homogenized equation. Let us first report a result from [2], which will be employed in the following.

Lemma 3.15. *Under Assumption 2.1 and for ε sufficiently small it holds in law*

$$\sqrt{2\sigma} \int_s^t \left(1 + \Phi'(Y_r^\varepsilon) \right) \mathrm{d}W_r = \sqrt{2\Sigma(t-s)}\xi_{t,s} + S(\varepsilon),$$

where $\Sigma \coloneqq K\sigma$ and where K is given in (1.3)*, the random variables $\xi_{t,s} \sim \mathcal{N}(0,1)$ for all $t, s > 0$, and where for all $p > 0$ it holds*

$$\left(\mathbb{E}^{\mu^\varepsilon} \left|S(\varepsilon)\right|^p\right)^{1/p} \leq C\varepsilon^\kappa,$$

for any $\kappa \in (0, 1/2)$. Moreover, for $t, s, t', s' > 0$ it holds

$$\mathbb{E}\,\xi_{t,s}\xi_{t',s'} = \frac{1}{\sqrt{(t-s)(t'-s')}}(\min\{t,t'\} - \min\{s,t'\} - \min\{t,s'\} + \min\{s,s'\}).$$

Proof. The proof ought to be found in [2, Proposition 5.8]. Moreover, the process $\xi_{t,s}$ is defined as the normalized increment of a time-changed Brownian motion as

$$\xi_{t,s} = \frac{\widehat{W}_{2\Sigma t} - \widehat{W}_{2\Sigma s}}{\sqrt{2\Sigma(t-s)}},$$

which proves the claim for its covariance. □

Due to Lemma 3.8 and in light of Lemma 3.11 and Lemma 3.15, we have

$$\begin{aligned}
\frac{\mathrm{d}Z_t^{\varepsilon,k}}{\mathrm{d}t} &= A\int_0^t \partial_t k^\varepsilon(t,s)\left(V'(X_t^\varepsilon)\,(t-s) + R(\varepsilon, t-s)\right)\mathrm{d}s \\
&\quad - \sqrt{2\Sigma}\int_0^t \partial_t k^\varepsilon(t,s)\left(\sqrt{t-s}\,\xi_{t,s} + S(\varepsilon)\right)\mathrm{d}s \\
&\quad + \varepsilon\int_0^t \partial_t k^\varepsilon(t,s)\big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big)\,\mathrm{d}s + k^\varepsilon(t,0)X_t^\varepsilon \\
&= \int_0^t \partial_t k^\varepsilon(t,s)\left(\sqrt{2\Sigma(t-s)}\,\xi_{t,s} + \widetilde{R}(\varepsilon, t-s)\right)\mathrm{d}s + k^\varepsilon(t,0)X_t^\varepsilon,
\end{aligned}$$

where

$$\widetilde{R}(\varepsilon, t-s) \coloneqq AV'(X_t^\varepsilon)\,(t-s) + R(\varepsilon, t-s) - \sqrt{2\Sigma}S(\varepsilon) + \varepsilon\big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big).$$

We first study the behaviour of the quantity

$$I_1(t) \coloneqq \int_0^t \partial_t k^\varepsilon(t,s)\sqrt{2\Sigma(t-s)}\,\xi_{t,s}\,\mathrm{d}s,$$

which satisfies

$$Z_{t_{j+1}} - Z_{t_j} \approx \int_{t_j}^{t_{j+1}} I_1(t)\,\mathrm{d}t.$$

We have

$$\mathbb{E}\,I_1^2(t) = 2\Sigma\int_0^t\int_0^t \partial_s k^\varepsilon(t,s)\partial_r k^\varepsilon(t,r)\sqrt{(t-s)(t-r)}\,\mathbb{E}\,\xi_{t,s}\xi_{t,r}\,\mathrm{d}r\,\mathrm{d}s.$$

Let us remark that $\sqrt{t-s}\,\xi_{t,s} = \widetilde{W}(t) - \widetilde{W}(s)$, where $\widetilde{W}$ is a standard Brownian motion.

3.2.5 Proof of Theorem 3.5

In view of Proposition 2.2, it suffices to prove that the MLE $\widehat{A}_T(Z_{0:T}^{\varepsilon,k})$ converges to the drift coefficient A of the homogenized diffusion. Rewriting (3.3) in view of Lemma 3.8 and denoting

$\Psi_t^\varepsilon \coloneqq 1 + \Phi'(Y_t^\varepsilon)$, one gets

$$\begin{aligned}\widehat{A}_T(Z_{0:T}^{\varepsilon,k}) = -\frac{1}{\int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t} \Bigg(&\alpha \int_0^T V'(Z_t^{\varepsilon,k}) \int_0^t \partial_t k^\varepsilon(t,s) \int_s^t V'(X_r^\varepsilon)\, \Psi_r^\varepsilon \,\mathrm{d}r \,\mathrm{d}s \,\mathrm{d}t \\ &- \sqrt{2\sigma} \int_0^T V'(Z_t^{\varepsilon,k}) \int_0^t \partial_t k^\varepsilon(t,s) \int_s^t \Psi_r^\varepsilon \,\mathrm{d}W_r \,\mathrm{d}s \,\mathrm{d}t \\ &+ \varepsilon \int_0^T \int_0^t \partial_t k^\varepsilon(t,s) \big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big) \,\mathrm{d}s \,\mathrm{d}t \\ &+ \int_0^T k^\varepsilon(t,0) X_t^\varepsilon \,\mathrm{d}t \Bigg) \eqqcolon I_1 + I_2 + I_3 + I_4.\end{aligned}$$

Before analysing the terms above singularly, let us remark that Lemma 3.9 yields

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t = \mathbb{E}^{\mu^\varepsilon} \left(V'(\cdot)^2\right) + \beta(\varepsilon),$$

in probability, and introduce the notation $C_V \coloneqq \mathbb{E}^{\mu^\varepsilon} \left(V'(\cdot)^2\right)$. Let us first consider I_1. Due to Lemma 3.11, we can rewrite in law

$$\begin{aligned}I_1 = -\frac{A}{\int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t} \Bigg(&\int_0^T V'(Z_t^{\varepsilon,k})^2 \int_0^t \partial_t k^\varepsilon(t,s)\, (t-s) \,\mathrm{d}s \,\mathrm{d}t \\ &+ \int_0^T V'(Z_t^{\varepsilon,k})^2 \int_0^t \partial_t k^\varepsilon(t,s)\, R(\varepsilon, t-s) \,\mathrm{d}s \,\mathrm{d}t \Bigg) \eqqcolon I_{1,1} + I_{1,2},\end{aligned}$$

where, applying (3.4) shows that $I_{1,1}$ satisfies

$$\begin{aligned}I_{1,1} &= \frac{A}{\int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t} \int_0^T V'(Z_t^{\varepsilon,k})^2 \left(1 - \varphi(t,\varepsilon) - t k^\varepsilon(t,0)\right) \,\mathrm{d}t \\ &= A - A\, \frac{\int_0^T V'(Z_t^{\varepsilon,k})^2 \left(\varphi(t,\varepsilon) + t k^\varepsilon(t,0)\right) \,\mathrm{d}t}{\int_0^T V'(Z_t^{\varepsilon,k})^2 \,\mathrm{d}t}.\end{aligned}$$

Regarding the second term above, dividing the numerator and the denominator by T, we have that the denominator tends a.s. to C_V for $\varepsilon \to 0$ and the numerator satisfies

$$\frac{1}{T} \mathbb{E}^{\mu^\varepsilon} \left| \int_0^T V'(Z_t^{\varepsilon,k})^2 \left(\varphi(t,\varepsilon) + t k^\varepsilon(t,0)\right) \,\mathrm{d}t \right| \leq C \varepsilon^\gamma \int_0^{\varepsilon^{-\gamma}} \left(\varphi(t,\varepsilon) + t k^\varepsilon(t,0)\right) \,\mathrm{d}t.$$

Hence, due to Assumption 3.2 and applying Slutsky's theorem we have

$$\lim_{\varepsilon\to 0} I_{1,1} = A,$$

in law. Since the limit A is a finite deterministic value, the same result holds in probability. The rest of the proof is devoted to showing that $I_{1,2}$ as well as I_2, I_3 and I_4 vanish for $\varepsilon \to 0$ in probability. Let us consider $I_{1,2}$. We divide the numerator and the denominator by T, so that the denominator converges a.s. to C_V for $\varepsilon \to 0$. For the numerator, we have due to Lemma 3.14

$$\begin{aligned}\frac{1}{T} \mathbb{E}^{\mu^\varepsilon} \left| \int_0^T V'(Z_t^{\varepsilon,k})^2 \int_0^t \partial_t k^\varepsilon(t,s)\, R(\varepsilon, t-s) \,\mathrm{d}s \,\mathrm{d}t \right| \leq \frac{C}{T} \int_0^T \Big(&\varepsilon^{2-\zeta} + \varepsilon^{1-\kappa} + \vartheta(t,\varepsilon) \\ &+ \psi(t,\varepsilon)^{1/2} + \psi(t,\varepsilon) \Big) \,\mathrm{d}s.\end{aligned}$$

Under the assumptions on the kernel and since $\kappa \in (0,1)$ and $\zeta \in (0,1)$, the numerator vanishes in the limit for $\varepsilon \to 0$. Hence, we conclude

$$\lim_{\varepsilon\to 0} I_{1,2} = 0,$$

in law, and therefore in probability. We now consider I_2. Let us introduce the notation

$$M_t := \int_0^t \int_s^t \partial_t k^\varepsilon(t,s)\Psi_r^\varepsilon \, \mathrm{d}W_r \, \mathrm{d}s,$$

and remark that M_t is a martingale, as for $t' < t$ and the natural filtration $\mathcal{F}_t$ generated by W_t we get

$$\begin{aligned}
\mathbb{E}\left(M_t \mid \mathcal{F}_{t'}\right) &= \int_0^{t'} \partial_t k^\varepsilon(t,s)\, \mathbb{E}\left(\int_s^t \Psi_r^\varepsilon \, \mathrm{d}W_r \mid \mathcal{F}_{t'}\right) \mathrm{d}s \\
&\quad + \int_{t'}^t \partial_t k^\varepsilon(t,s)\, \mathbb{E}\left(\int_s^t \Psi_r^\varepsilon \, \mathrm{d}W_r \mid \mathcal{F}_{t'}\right) \mathrm{d}s \\
&= \int_0^{t'} \partial_t k^\varepsilon(t,s) \int_s^{t'} \Psi_r^\varepsilon \, \mathrm{d}W_r \, \mathrm{d}s = M_{t'}.
\end{aligned}$$

Let us denote by Q_t the martingale

$$Q_t := \int_0^t \partial_t k^\varepsilon(t,s)\Psi_r^\varepsilon \, \mathrm{d}W_r,$$

whose quadratic variation is given by

$$\langle Q \rangle_t = \int_0^t \left(\partial_t k^\varepsilon(t,s)\Psi_r^\varepsilon\right)^2 \, \mathrm{d}r$$

so that by linearity and since $\mathrm{d}(tQ_t) = t\mathrm{d}(Q_t) + Q_t \mathrm{d}t$,

$$\begin{aligned}
M_t &= \int_0^t Q_t \, \mathrm{d}s - \int_0^t Q_s \, \mathrm{d}s \\
&= tQ_t - tQ_t + \int_0^t s \, \mathrm{d}Q_s = \int_0^t s \, \mathrm{d}Q_s.
\end{aligned}$$

The quadratic variation of M_t is therefore given by

$$\langle M \rangle_t = \int_0^t \left(s\, \partial_t k^\varepsilon(t,s)\, \Psi_s^\varepsilon\right)^2 \, \mathrm{d}s.$$

The martingale term is giving me headaches. Let us now consider I_3. Dividing the numerator and the denominator by T and considering Assumption 3.2, we have for the numerator

$$\begin{aligned}
\frac{1}{T}\, \mathbb{E}^{\mu^\varepsilon} \left| \varepsilon \int_0^T \int_0^t \partial_t k^\varepsilon(t,s)\big(\Phi(Y_t^\varepsilon) - \Phi(Y_s^\varepsilon)\big) \, \mathrm{d}s \, \mathrm{d}t \right| &\leq -\frac{C\varepsilon}{T} \int_0^T \int_0^t \partial_t k^\varepsilon(t,s) \, \mathrm{d}s \, \mathrm{d}t \\
&\leq \frac{C\varepsilon}{T} \int_0^T \left(k^\varepsilon(t,t) - k^\varepsilon(t,0)\right) \mathrm{d}t \\
&\leq C\varepsilon k^\varepsilon(T,T) \\
&\leq C\varepsilon^{1-\zeta}.
\end{aligned}$$

For the denominator we have convergence in probability towards C_V, which implies since $\zeta \in (0,1)$

$$\lim_{\varepsilon \to 0} I_3 = 0,$$

in probability. We conclude by considering I_4. Reasoning as above, and due to Lemma 3.7, we have

$$\frac{1}{T}\, \mathbb{E}^{\mu^\varepsilon} \left| \int_0^T k^\varepsilon(t,0) X_t^\varepsilon \, \mathrm{d}t \right| \leq \frac{C}{T} \int_0^T k^\varepsilon(t,0) \, \mathrm{d}t,$$

which, considering Assumption 3.2, implies that

$$\lim_{\varepsilon \to 0} I_4 = 0,$$

which concludes the proof. □

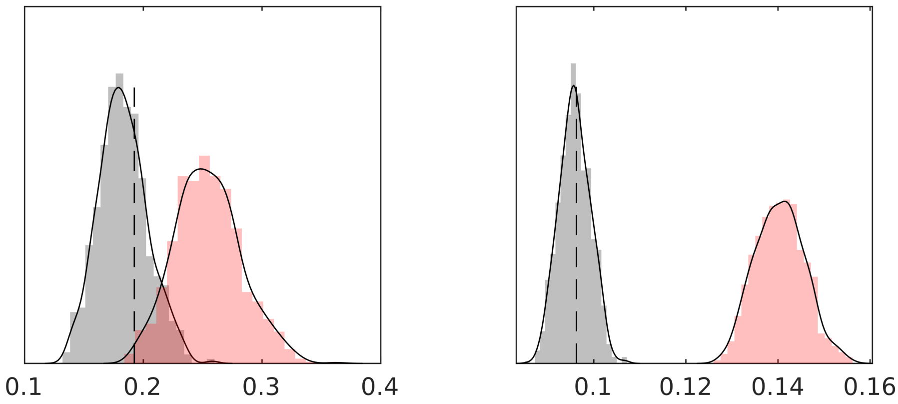

Figure 1: Results for the Ornstein–Uhlenbeck process

4 Convergence analysis for $N > 1$

5 Numerical experiments

5.1 Ornstein–Uhlenbeck process

Results in 1.

5.2 Bistable potential

5.3 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.