ESAIM: PS 29 (2025) 400–449
https://doi.org/10.1051/ps/2025012

ESAIM: Probability and Statistics
www.esaim-ps.org

NONPARAMETRIC MOMENT METHOD FOR SCALAR MCKEAN–VLASOV STOCHASTIC DIFFERENTIAL EQUATIONS

F. Comte[1,*], V. Genon-Catalot[1] and C. Larédo[2]

Abstract. We study the nonparametric estimation of both the potential and the interaction terms of a scalar McKean–Vlasov stochastic differential equation (SDE) in stationary regime from a continuous observation on a time interval $[0, T]$, with asymptotic framework $T \to +\infty$. To estimate the two functions, we consider the observation of four *i.i.d.* sample paths. The observation of two sample paths could be enough at the cost of much more computations. Estimators of the potential and the interaction functions are built using a combination of a moment method and a projection method on sieves. The potential and the interaction term do not belong to $\mathbb{L}^2(\mathbb{R})$, so we define a specific risk fitted to this estimation problem and obtain a bound for it. A nonparametric estimator of the invariant density also is proposed. The method is implemented on simulated data for several examples of McKean–Vlasov SDEs and a model selection procedure is experimented.

Mathematics Subject Classification. 62G05, 62M99, 60J60.

Received November 20, 2024. Accepted June 2, 2025.

1. Introduction

McKean–Vlasov processes are described by a stochastic differential equation where the coefficients depend on the distribution of the solution itself. They are also called nonlinear or self-stabilizing processes. These processes have been extensively studied since their first description by McKean [1], and appear in various applications: for the modelling of granular media in statistical physics [2], in neurosciences (see *e.g.* [3, 4]), for population dynamics and ecology [5, 6], for epidemics dynamics [7, 8] and in finance (see *e.g.* [9] and the references therein).

A large number of contributions is devoted to probabilistic properties of these models. Existence and uniqueness of solutions under several different sets of assumptions can be found in *e.g.*, [10]–[17]. More references are given below.

We are interested in a one dimensional McKean–Vlasov process of the form

$$dX(t) = -[b(X(t)) + \phi \star \mu_t(X(t))]\mathrm{d}t + \sigma dW(t), \quad X(0) = \eta \sim \nu, \tag{1.1}$$

where b, ϕ are deterministic functions, $\star$ denotes the convolution product, μ_t is the law of $X(t)$, W is a Brownian motion and η is a random variable independent of W. Such process can be obtained as the limit of a system of N

Keywords and phrases: Invariant density, interaction function, Hermite basis, McKean-Vlasov stochastic differential equation, moment method, projection estimators, nonparametric drift estimation.

[1] Université Paris Cité, MAP5, UMR 8145 CNRS, 75006, France.
[2] Université Paris Cité, LPSM UMR 8001 CNRS, 75006, France.

[*] Corresponding author: fabienne.comte@u-paris.fr

interacting particles as N tends to infinity (*propagation of chaos*). More precisely, consider the N-dimensional process given by the stochastic differential system:

$$\begin{aligned} dX_i^N(t) &= -b(X_i^N(t))\mathrm{d}t - \frac{1}{N}\sum_{j=1}^{N}\phi(X_i^N(t)-X_j^N(t))\mathrm{d}t + \sigma dW_i(t), \\ &= -b(X_i^N(t))\mathrm{d}t - \phi\star\mu_t^N(X_i^N(t))\mathrm{d}t + \sigma dW_i(t), \quad X_i^N(0)=X_0^i, i=1,\dots,N, \end{aligned} \tag{1.2}$$

where $\mu_t^N = N^{-1}\sum_{j=1}^N \delta_{X_j^N(t)}$ is the empirical measure associated with $(X_j^N(t), j=1,\dots,N)$, $(W_i, i=1\dots,N)$ are N independent Brownian motions, $X_0^i, i=1,\dots,N$ are independent and identically distributed (*i.i.d.*) random variables, independent of $(W_j, j=1,\dots,N)$. The function ϕ describes the interaction between the N particles.

For any fixed integer p, the process $(X_i^N(t), i=1,\dots,p)$ converges in distribution as N tends to infinity to p *i.i.d.* McKean–Vlasov processes $(X^{(i)}(t))$, given by

$$dX^{(i)}(t) = -[b(X^{(i)}(t)) + \phi\star\mu_t(X^{(i)}(t))]\mathrm{d}t + \sigma dW_i(t), \quad X^{(i)}(0)=X_0^i, i=1,\dots,p.$$

Thus, the process $(X^{(i)}(t))$, *i.e.* the mean-field limit, provides a good approximation of the behaviour of the i-th particle of the system (see *e.g.*, [18]–[22]). Moreover, existence and uniqueness of invariant distributions for (1.1) are studied in [14, 15, 19],[21]–[25]).

Statistical inference for interacting particle systems and their mean-field limits has started more recently. On the one hand, several papers deal with inference based on the observation on a fixed time interval of an interacting particle system with asymptotic properties as the number N of particles tends to infinity (see *e.g.* parametric inference in [9], [26]–[33], semiparametric inference in [34], and nonparametric inference in [35, 36]), [37, 38].

On the other hand, several papers consider inference based on the observation of the mean-field limit process (1.1) on a time interval $[0,T]$, that is inference for the limiting process of one typical particle for large N. In [39] and [40], estimation of unknown parameters in the potential term b and the interaction term ϕ are studied under the asymptotic framework $\sigma\to 0$. More recently, statistical inference has been investigated under the assumption of stationarity for (1.1) when the process is observed in the stationary regime with asymptotic framework $T\to+\infty$. In [41], the case of $b=0$ and ϕ equal to an odd polynomial is studied and the estimation of the coefficients of the polynomial is treated by a pseudo-likelihood approach. In [42], general functions b and ϕ depending on unknown parameters are considered and a different pseudo-likelihood function is proposed. Pavliotis and Zanoni [43] use a moment method to estimate the coefficients of polynomial functions b and ϕ.

In here, our aim is the nonparametric estimation of the functions b,ϕ from the continuous observation of a process distributed as (1.1) on a time interval $[0,T]$ when the process is in stationary regime and T tends to infinity. When $\phi\equiv 0$ (usual diffusion process), nonparametric inference for the function b, based on the observation of one sample path on a time interval $[0,T]$, when the process is in stationary regime and T is large, is an extensively developped subject. We can quote [44, 45], see also the book [46], (asymptotic equivalence for scalar diffusions, [47], for multidimensional diffusions, [48]), [49]–[53].

In the case of the McKean–Vlasov process (1.1) with $\phi\neq 0$ observed in the stationary regime, the joint nonparametric estimation of (b,ϕ) with general functions (b,ϕ) has never been investigated up to our knowledge. There are differences from the case of a usual diffusion. Indeed, although there are clear assumptions (see below) for existence and uniqueness of an invariant distribution, this distribution is not explicitely given as a function of b and ϕ. It is solution of an implicit equation. The convolution term depending on the unknown marginal distribution of the process introduces new difficulties for the nonparametric estimation of b and ϕ linked with identifiability problems. For these reasons, the observation of only one sample path is not enough to infer both b and ϕ. One possibility is to split the sampling interval into $[0,T/2]$ and $[T/2,T]$, use the first half to estimate the unknown stationary density and use the second half for estimation of (b,ϕ). This is done in Genon-Catalot and

Larédo [42] for parametric inference on (b, ϕ). However, this requires the estimation of the stationary density and induces bias terms due to this estimation.

In what follows, the novelty is that we assume that the observation is composed of four independent copies $(X^{(\ell)}(t), t \in [0, T], \ell = 1, 2, 3, 4)$ in stationary regime and our asymptotic framework is $T \to +\infty$. Actually, only two copies are sufficient but four copies avoid cumbersome and tedious additional computations. We stress that the use of two independent trajectories circumvents the estimation step of the stationary distribution in our procedure. Our estimation method is a moment method inspired by the work in parametric setting of Pavliotis and Zanoni [43].

In Section 2, we present our set of assumptions ensuring that equation (1.1) admits a unique strong solution and that the model has a unique invariant distribution $\rho(x)\mathrm{d}x$. We give useful properties of the density ρ and of the process (1.1) in the stationary regime. Proposition 2.4 states a result on the infinitesimal generator of (1.1) often used in the proofs. In Section 3, the principle of the estimation method is described. It is based on properties of the infinitesimal generator of (1.1), and it combines a moment method with a projection method on finite dimensional subspaces of $\mathbb{L}^2(\mathbb{R}, \mathrm{d}x) := \mathbb{L}^2(\mathrm{d}x)$. Nonparametric estimators of b and ϕ based on the observation of four *i.i.d.* sample paths $(X^{(i)}(t), t \in [0, T])$ distributed as (1.1) in the stationary regime are built. Defining an adequate risk for the estimators is an additional difficulty since b, ϕ are not square integrable in our framework. We propose a definition of the risk well fitted to the problem which is defined as follows. For a couple (u, v) of functions, we set, when defined

$$\|(u, v)\|_V^2 := \int (u(x) + v \star \rho(x))^2 \rho(x)\mathrm{d}x. \tag{1.3}$$

The risk of an estimator $(\widehat{b}, \widehat{\phi})$ is defined by

$$\mathbb{E}(\|(\widehat{b} - b, \widehat{\phi} - \phi)\|_V^2). \tag{1.4}$$

Theorem 3.3 gives the risk bound of our projection estimators on a fixed space. The choice of the loss function measured with the norm $\|(u, v)\|_V^2$ is discussed and justified by the result of our Theorem 3.3. Section 4 is devoted to the nonparametric estimation of the invariant density by projection method . The estimator exhibits a parametric rate. In Section 5, we present numerical simulation results on several models for the estimation of the invariant density and of b and ϕ. In the latter case, we experiment a model selection procedure. We detail the simulation method of the sample paths. We observe that the invariant density estimator performs quite well. Clearly, estimating b and ϕ is more difficult. We compare the estimator obtained by model selection with an oracle and observe that most of the time the selection method performs as well as the oracle. Section 6 contains concluding remarks and Section 7 (Supplementary material) is devoted to proofs.

2. Assumptions and preliminaries

In the sequel, the notation $\lesssim$ means $\leq$ up to a constant. Let us set

$$B(x) = \int_0^x b(y)\mathrm{d}y, \quad \Phi(x) = \int_0^x \phi(y)\mathrm{d}y.$$

The function b is the potential term, the function ϕ is the interaction function. We assume that $b \not\equiv 0$ and that these functions satisfy:

- **[H1]** The function Φ is even. The functions B and Φ are C^2 and there exist constants K and λ such that

$$\forall x \in \mathbb{R}, \quad B''(x) \geq K \geq 0, \ \Phi''(x) \geq \lambda, \quad K + \lambda > 0.$$

- **[H2]** The functions $B' = b, \Phi' = \phi$ are locally Lipschitz with polynomial growth, *i.e.* there exist $c > 0$, $\ell \in \mathbb{N}^* = \{1, 2, 3, \ldots\}$ such that

$$\forall x, y \in \mathbb{R},\ |b(x) - b(y)| + |\phi(x) - \phi(y)| \leq c|x - y|(1 + |x|^\ell + |y|^\ell).$$

- **[H3]** The functions b, b', ϕ, ϕ' have ℓ polynomial growth: there exists a constant c such that

$$|b(x)| + |b'(x)| + |\phi(x)| + |\phi'(x)| \leq c(1 + |x|^\ell).$$

According to [14, 16, 21, 22, 25, 37], under assumptions **[H1]**–**[H3]**, equation (1.1) admits a unique solution $((X(t), \mu_t(x)\mathrm{d}x)), t \geq 0)$, where $\mu_t(x)\mathrm{d}x = \mathcal{L}(X(t))$ is the marginal distribution of $X(t)$. Model (1.1) admits a unique invariant distribution ρ such that $\int_{\mathbb{R}} x^2 \rho(x)\mathrm{d}x < +\infty$. Note that our set of assumptions is slightly more general than the one given in [42], as Φ is not assumed to be convex. But the important point is that we have $K + \lambda > 0$.

If the initial variable $X(0)$ of (1.1) follows the invariant distribution ρ, then, for all t, $\mathcal{L}(X(t)) = \mathcal{L}(X(0))$. Thus, we assume

- **[H4]** The initial variable $X(0)$ follows the invariant distribution ρ.

Under **[H1]**–**[H3]**, the invariant distribution has density ρ given as the solution of the implicit equation

$$\rho(x) = \frac{1}{M} \exp\left\{-\frac{2}{\sigma^2}\int_0^x [b(u) + \phi \star \rho(u)]\mathrm{d}u\right\}, \tag{2.1}$$

where $M = \int_{\mathbb{R}} \exp\{-\frac{2}{\sigma^2}\int_0^y [b(u) + \phi \star \rho(u)]\mathrm{d}u\}\mathrm{d}y$. As $(\Phi \star \rho)' = \phi \star \rho$, $\int_0^x \phi \star \rho(u)\mathrm{d}u = \Phi \star \rho(x) - \Phi \star \rho(0)$,

$$\rho(x) \propto \exp\left[-\frac{2}{\sigma^2}(B(x) + \Phi \star \rho(x))\right]. \tag{2.2}$$

Note that if $b \equiv 0$, there is a one-parameter family of invariant distributions (see *e.g.* [36, 54]). The invariant distribution is not unique unless its expectation is specified. We exclude this case here.

Contrary to classical stochastic differential equations, the invariant distribution is not explicit. Nevertheless, we can prove:

Proposition 2.1. *Under* **[H1]**–**[H3]**,

$$\rho(x) \lesssim \exp\left[-\frac{(K+\lambda)}{\sigma^2}\left(x + \frac{b(0) + \phi \star \rho(0)}{(K+\lambda)}\right)^2\right], \tag{2.3}$$

Moreover, for $0 \leq z \leq x$ and for $x \leq z \leq 0$,

$$\frac{\rho(x)}{\rho(z)} \leq \exp\left[C(x - z)\right] \quad \textit{where} \quad C = -(2/\sigma^2)(b(0) + \phi \star \rho(0)). \tag{2.4}$$

Inequality (2.3) is proved in [42] and [37]. Proposition 2.1 has the obvious consequence that the invariant distribution has moments of any orders and (2.3) implies that, for all $k \in \mathbb{R}$ and all $\varepsilon > 0$,

$$\int_{\mathbb{R}} \exp(kx)(\rho(x))^\varepsilon \mathrm{d}x < +\infty. \tag{2.5}$$

Note that, by **[H3]**, $\phi \star \rho$ has polynomial growth as ρ has moments of any order.

Under **[H1]**–**[H4]**, the initial variable η follows the invariant distribution $\rho(x)\mathrm{d}x$, which implies that the distribution $\mathcal{L}(X(t)) = \mu_t(\mathrm{d}x)$ satisfies

$$\forall t \geq 0,\ \mu_t(\mathrm{d}x) = \rho(x)\mathrm{d}x.$$

Therefore, equation (1.1) becomes:

$$dX(t) = -(b(X(t)) + \phi \star \rho(X(t)))\mathrm{d}t + \sigma dW(t), \quad X(0) = \eta \sim \rho(x)\mathrm{d}x. \tag{2.6}$$

As stressed in Genon-Catalot and Larédo (2023b), the following result holds and has important consequences:

Proposition 2.2. *Assume* **[H1]**–**[H4]** *and consider the stochastic differential equation*

$$dY(t) = -(b(Y(t)) + \phi \star \rho(Y(t)))\mathrm{d}t + \sigma dW(t).$$

Then $(Y(t))$ is a positive recurrent diffusion with stationary density given by (2.1). If $Y(0) \sim \rho(x)\mathrm{d}x$, it is ergodic. Moreover,
- If $Y(0) \neq X(0)$, $(Y(t)) \not\equiv (X(t))$.
- If $Y(0) = X(0) = \eta \sim \rho(x)\mathrm{d}x$, then $X(t) = Y(t)$ for all $t \geq 0$.

The result simply derives from the uniqueness of solutions. Thus, in the stationary regime, the process $(X(t))$ given by (2.6) is identical to the classical SDE $(Y(t))$ in the stationary regime and is ergodic.

The infinitesimal generator of the SDE (2.6) is given by:

$$Lg = \frac{\sigma^2}{2} g'' - (b + \phi \star \rho)g' = \frac{\sigma^2}{2\rho}(g'\rho)'. \tag{2.7}$$

The operator L acts on $\mathbb{L}^2(\rho(x)\mathrm{d}x) := \mathbb{L}^2(\rho)$ and is defined on the domain $\mathcal{D}$,

$$\mathcal{D} = \{g \in \mathbb{L}^2(\rho),\ \ g' \text{ abs. continuous},\ Lg \in \mathbb{L}^2(\rho), \lim_{|x|\to\infty} g'(x)\rho(x) = 0\}. \tag{2.8}$$

(see *e.g.* [55]).

Below, we use that, for all function $\Psi \in \mathcal{D}$, $\int_{\mathbb{R}} L\Psi(x)\rho(x)\mathrm{d}x = 0$. This relation obviously holds as, by (2.7)–(2.8),

$$\int_{\mathbb{R}} L\Psi(x)\rho(x)\mathrm{d}x = \int_{\mathbb{R}} \frac{\sigma^2}{2\rho(x)}(\Psi'\rho)'(x)\rho(x)\mathrm{d}x = \frac{\sigma^2}{2}\int_{\mathbb{R}}(\Psi'\rho)'(x)\mathrm{d}x = [\Psi'\rho]_{-\infty}^{+\infty} = 0.$$

In order to apply our estimation method, we need the following lemma.

Lemma 2.3. *Consider a function ψ such that $\psi \in \mathbb{L}^2(\mathrm{d}x)$, ψ is absolutely continuous and bounded and ψ' has polynomial growth. Then, $x \mapsto \int_0^x \psi(t)\mathrm{d}t$ and $x \mapsto \int_0^x \psi \star \rho(t)\mathrm{d}t$ both belong to $\mathcal{D}$.*

Note that, if a function h has exponential growth, *i.e.* there exists a constant $a > 0$ such that, for all $u \in \mathbb{R}$, $|h(u)| \leq a\mathrm{e}^{au}$, then by Proposition 2.1 (see (2.3) and (2.5)), $h \in \mathbb{L}^2(\rho)$. Now, we can state a key property for the statistical strategy.

Proposition 2.4. *Assume* **[H1]**–**[H4]** *and that h has exponential growth. Set $h_c = h - \int_{\mathbb{R}} h(x)\rho(x)\mathrm{d}x$ and define the function Γ_h by $\Gamma_h(x) = \int_0^x (\Gamma_h)'(y)\mathrm{d}y$ where*

$$(\Gamma_h)'(y) = \frac{2}{\sigma^2\rho(y)} \int_{-\infty}^{y} h_c(u)\rho(u)\mathrm{d}u = -\frac{2}{\sigma^2\rho(y)} \int_{y}^{+\infty} h_c(u)\rho(u)\mathrm{d}u \tag{2.9}$$

Then, $\Gamma_h \in \mathcal{D}$, $L\Gamma_h = h_c$, and $(\Gamma_h)' \in \mathbb{L}^2(\rho)$. Moreover, for $(X(t))$ satisfying (2.6),

$$\int_0^T h_c(X(s))\mathrm{d}s = \sigma \int_0^T (\Gamma_h)'(X(s))dW(s) + \Gamma_h(X(0)) - \Gamma_h(X(T)), \tag{2.10}$$

$$\mathrm{Var}\Big(\frac{1}{T}\int_0^T h_c(X(s))\mathrm{d}s\Big) \leq \frac{2\sigma^2}{T} \int [(\Gamma_h)'(x)]^2\rho(x)\mathrm{d}x + \frac{8}{T^2}\int \Gamma_h^2(x)\rho(x)\mathrm{d}x. \tag{2.11}$$

Note that relation (2.9) holds since $\int_{\mathbb{R}} h_c(x)\rho(x)\mathrm{d}x = 0$. The relation (2.10) with the explicit function $(\Gamma_h)'$ given by (2.9) is especially important and used several times in our proofs (Props. 3.2, 3.4, 3.5, 4.1).

3. Estimation method for the potential and the interaction functions

3.1. Notations

• For h a function, we set $\|h\|^2 = \int_{\mathbb{R}} h^2(x)\mathrm{d}x$, $\mathbb{L}^2(\mathrm{d}x) = \{h, \|h\|^2 < +\infty\}$, $\|h\|_\rho^2 = \int_{\mathbb{R}} h^2(x)\rho(x)\mathrm{d}x$, $\mathbb{L}^2(\rho) = \{h, \|h\|_\rho^2 < +\infty\}$.
• For $\mathbf{x}$ a vector in $\mathbb{R}^d$, $\|\mathbf{x}\|_{\mathbb{R}^d}$ denotes the Euclidean norm of the vector.
• For a matrix M, $M^\perp$ denotes its transpose. For M a square matrix, $\mathrm{Tr}(M)$ is the trace of M, and $\|M\|_{\mathrm{op}}$ is the square-root of the largest eigenvalue of $MM^\perp$. If M is symmetric, $\|M\|_{\mathrm{op}}$ is simply equal to $\sup_i\{|\lambda_i|\}$ where the λ_i are the eigenvalues of M.
• $\mathcal{M}_{r,s}(\mathbb{R})$ denotes the set of $(r \times s)$ matrices.
• Below, vectors are denoted using bold letters and coordinates or functions are denoted with usual letters (not bold).

For $(x_0, \dots, x_{m-1}, y_0, \dots, y_{p-1}) \in \mathbb{R}^{m+p}$, we denote

$$(\mathbf{x}, \mathbf{y})_{m,p} = (x_0, \dots, x_{m-1}, y_0, \dots, y_{p-1})^\perp = \begin{pmatrix} \mathbf{x}_m \\ \mathbf{y}_p \end{pmatrix} \tag{3.1}$$

with $\mathbf{x}_m = (x_0, \dots, x_{m-1})^\perp, \mathbf{y}_p = (y_0, \dots, y_{p-1})^\perp$.
• For a couple (u, v) of functions, we set, when defined, $\|(u, v)\|_V^2 = \|u + v \star \rho\|_\rho^2$.

3.2. Principle of the method

Let ψ be a function satisfying the assumptions of Lemma 2.3 and set $\Psi(x) = \int_0^x \psi(y)\mathrm{d}y$. Then, following Pavliotis and Zanoni [43], we notice that, from the definition of ρ and (2.7)–(2.8), $\int_{\mathbb{R}} L\Psi(x)\rho(x)\mathrm{d}x = 0$, so that

$$\int_{\mathbb{R}} \psi(x)b(x)\rho(x)\mathrm{d}x + \int_{\mathbb{R}} \psi(x)\phi \star \rho(x)\rho(x)\mathrm{d}x = \frac{1}{2}\int_{\mathbb{R}} \psi'(x)\sigma^2\rho(x)\mathrm{d}x. \tag{3.2}$$

To estimate b and ϕ, we must take into account that the functions b, ϕ and $\phi \star \rho$ do not belong to $\mathbb{L}^2(\mathrm{d}x)$. By **[H3]**, these three functions have polynomial growth and therefore, as ρ has moments of any order, belong to $\mathbb{L}^2(\rho)$. So we proceed as follows.

Consider two orthonormal bases of $\mathbb{L}^2(\mathrm{d}x)$, $(\varphi_j)_{j\geq 0}, (\theta_\ell)_{\ell\geq 0}$ composed of bounded derivable functions, with derivatives having polynomial growth, thus satisfying the assumptions of Lemma 2.3. Note that this excludes compactly supported bases whose functions are not differentiable on $\mathbb{R}$. We denote for $\tau = \varphi$ and $\tau = \theta$:

$$L_\tau(q) := \sup_{x\in\mathbb{R}} \sum_{k=0}^{q-1} \tau_k^2(x), \tag{3.3}$$

and assume that $L_\varphi(m) < +\infty, L_\theta(p) < +\infty$ for all fixed $m \geq 1, p \geq 1$. The quantity $L_\tau(q)$ only depends on the space generated by $(\tau_0, \ldots, \tau_{q-1})$. Indeed,

$$L_\tau(q) = \sup_{h\in \mathrm{span}(\tau_0,\ldots,\tau_{q-1})} \|h\|_\infty^2 / \|h\|^2.$$

It is finite for classical examples of bases. Generally, for all q, $L_\tau(q) \leq c_\tau q^\alpha$ with c_τ, α constants linked with the basis (see Sect. 3.7). As ρ is bounded, for all j, φ_j belongs to $\mathbb{L}^2(\rho)$. For $m \geq 1$, the functions $\varphi_0, \ldots, \varphi_{m-1}$ are linearly independent and generate a m-dimensional space

$$S_m = \mathrm{Vect}(\varphi_0, \ldots, \varphi_{m-1})$$

which is a subspace of $\mathbb{L}^2(\rho)$. Analogously, we define the p-dimensional subspace of $\mathbb{L}^2(\rho)$

$$\Sigma_p = \mathrm{Vect}(\theta_0, \ldots, \theta_{p-1}).$$

Recall that, by definition (1.3), for a couple (u, v) of functions, we have set, when defined, $\|(u, v)\|_V^2 = \|u + v \star \rho\|_\rho^2$. Note that $\|(u, v)\|_V^2$ is finite for $(u, v) = (b, \phi)$ and for $(u, v) \in S_m \times \Sigma_p$. Indeed for $(u, v) \in S_m \times \Sigma_p$, $\|u\|_\rho^2 \leq \|\rho\|_\infty \|u\|^2$ and by the Young inequality, $\|v \star \rho\| \leq \|\rho\|_1 \|v\| = \|v\|$, so that $v \star \rho$ belongs to $\mathbb{L}^2(\mathrm{d}x) \subset \mathbb{L}^2(\rho)$. Note also that $(u, v) \to \|(u, v)\|_V$ is only a semi-norm on $\mathbb{L}^2(\rho) \times \mathbb{L}^2(\rho)$. Indeed, $\|(u, v)\|_V = 0$ is equivalent to $u(x) + v \star \rho(x) = 0$ for all $x \in \mathbb{R}$. But this does not always implies $u \equiv 0, v \equiv 0$. Nevertheless, we denote $\langle (u_1, v_1), (u_2, v_2)\rangle_V$ the scalar product associated to the semi-norm $\|.\|_V$.
Now, we define the V-orthogonal projection of (b, ϕ) on $S_m \times \Sigma_p$:

$$(b_m^V, \phi_p^V) = \arg \min_{(u,v)\in S_m\times\Sigma_p} \|(b - u, \phi - v)\|_V^2. \tag{3.4}$$

Setting $b_m^V = \sum_{k=0}^{m-1} \beta_k \varphi_k$ and $\phi_p^V = \sum_{\ell=0}^{p-1} c_\ell \theta_\ell$, we obtain

$$\mathbf{V}_{m,p}(\mathbf{b}, \mathbf{c})_{m,p} = \mathbf{Z}_{m,p} \tag{3.5}$$

using the notation

$$(\mathbf{b}, \mathbf{c})_{m,p} = (\beta_0, \ldots, \beta_{m-1}, c_0, \ldots, c_{p-1})^\perp = \begin{pmatrix} \mathbf{b}_m \\ \mathbf{c}_p \end{pmatrix} \in \mathcal{M}_{m+p,1}(\mathbb{R})$$

with $\mathbf{b}_m = (\beta_0, \ldots, \beta_{m-1})^\perp$ and $\mathbf{c}_p = (c_0, \ldots, c_{p-1})^\perp$, and where $\mathbf{V}_{m,p}$ is the $(m+p) \times (m+p)$ matrix, element of $\mathcal{M}_{m+p,m+p}(\mathbb{R})$,

$$\mathbf{V}_{m,p} = \begin{pmatrix} \mathbf{V}^1 = \left(\int \varphi_j \varphi_k \rho\right)_{0 \le j,k \le m-1} & \mathbf{V}^{1,2} = \left(\int \theta_\ell \star \rho\, \varphi_k \rho\right)_{\substack{0 \le k \le m-1 \\ 0 \le \ell \le p-1}} \\ \mathbf{V}^{2,1} = \left(\int \theta_r \star \rho\, \varphi_j \rho\right)_{\substack{0 \le r \le p-1, \\ 0 \le j \le m-1}} & \mathbf{V}^2 = \left(\int \theta_\ell \star \rho \;\; \theta_r \star \rho\, \rho\right)_{0 \le \ell, r \le p-1} \end{pmatrix}, \tag{3.6}$$

where $\mathbf{V}^{2,1} = (\mathbf{V}^{1,2})^\perp$,

$$\mathbf{Z}_{m,p} = \begin{pmatrix} \left(\int (b + \phi \star \rho)\varphi_k \rho\right)_{0 \le k \le m-1} \\ \left(\int (b + \phi \star \rho)(\theta_\ell \star \rho)\rho\right)_{0 \le \ell \le p-1} \end{pmatrix} \in \mathcal{M}_{m+p,1}(\mathbb{R}). \tag{3.7}$$

As the functions $\psi = \varphi_0, \ldots, \varphi_{m-1},\ \theta_0 \star \rho, \ldots, \theta_{p-1} \star \rho$ satisfy the assumptions of Lemma 2.3, we may apply formula (3.2) to these functions and we get, using that $(\theta_\ell \star \rho)' = \theta'_\ell \star \rho$,

$$\mathbf{Z}_{m,p} = \frac{\sigma^2}{2} \begin{pmatrix} \left(\int \varphi'_k\, \rho\right)_{0 \le k \le m-1} \\ \left(\int \theta'_\ell \star \rho\, \rho\right)_{0 \le \ell \le p-1} \end{pmatrix} \in \mathcal{M}_{m+p,1}(\mathbb{R}). \tag{3.8}$$

All the integrals in $\mathbf{V}_{m,p}, \mathbf{Z}_{m,p}$ are on $\mathbb{R}$ and with respect to the Lebesgue measure $\mathrm{d}x$.

3.3. Theoretical constraints and identifiability

The vector $(\mathbf{b}, \mathbf{c})_{m,p}$ in (3.5) is uniquely defined if and only if the matrix $\mathbf{V}_{m,p}$ is invertible.
We emphasize that the matrix $\mathbf{V}_{m,p}$ is symmetric and nonnegative. Indeed,

$$(\mathbf{x}, \mathbf{y})_{m,p}^\perp \mathbf{V}_{m,p} (\mathbf{x}, \mathbf{y})_{m,p} = \int_{\mathbb{R}} \left(\sum_{j=0}^{m-1} x_j \varphi_j(x) + \sum_{\ell=0}^{p-1} y_\ell \theta_\ell \star \rho(x) \right)^2 \rho(x) \mathrm{d}x \ge 0.$$

Note that

$$(\mathbf{x}, \mathbf{y})_{m,p}^\perp \mathbf{V}_{m,p} (\mathbf{x}, \mathbf{y})_{m,p} = \|(u,v)\|_V^2 \quad \text{for} \quad u = \sum_{j=0}^{m-1} x_j \varphi_j, v = \sum_{\ell=0}^{p-1} y_\ell \theta_l.$$

As, by (2.2), the support of ρ is $\mathbb{R}$, $\mathbf{V}_{m,p}$ is invertible if:

$$\left\{ \forall x \in \mathbb{R}, \sum_{j=0}^{m-1} x_j \varphi_j(x) + \sum_{\ell=0}^{p-1} y_\ell \theta_\ell \star \rho(x) = 0 \right\} \Rightarrow (\mathbf{x}, \mathbf{y})_{m,p} = (\mathbf{0}, \mathbf{0})_{m,p} \tag{3.9}$$

Condition (3.9) is an identifiability constraint linked with the choice of the bases. In Section 3.4 hereafter, we propose a basis for φ_j and θ_ℓ satisfying Condition (3.9), so that $(\mathbf{b}, \mathbf{c})_{m,p}$ in (3.5) is well defined for this basis.

Nevertheless, (3.9) is not enough for our purpose and we need to reinforce the identifiability constraint and set the following assumption:

- **[H5]** There exists a constant $f_0 > 0$ fixed and not depending on m, p, such that

$$\forall \mathbf{x} \in \mathbb{R}^{m+p}, \quad \|\mathbf{x}\|_{\mathbf{R}^{m+p}} = 1, \quad \mathbf{x}^{\perp} \mathbf{V}_{m,p} \mathbf{x} \geq f_0 > 0.$$

Assumption **[H5]** ensures that, for all (m, p), the eigenvalues of $\mathbf{V}_{m,p}$ are all larger than f_0. As a consequence, under **[H5]**, there exists a positive constant s_0 independent of m, p such that

$$\|\mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}} \leq s_0.$$

Clearly $s_0 = 1/f_0$ suits.

Assumption [**H5**] may seem difficult to check. However let us stress that, in complex statistical problems, it often happens that identifiability assumptions are difficult or impossible to check. It is for instance the case when the identifiability conditions concern random terms (see *e.g.* parametric estimation of a diffusion coefficient from discrete data within a fixed time interval). Pavliotis and Zanoni [43], for interacting particle systems, have also assumptions impossible to check. They propose, as a way to get through, to check numerically on the data whether the method can be applied. Analogously, in our case, our method works well on simulated data provided that (m, p) are not too large.

3.4. Example of basis satisfying the identifiability assumption (3.9)

Our method requires bases composed of functions for which Lemma 1 can be applied. This excludes compactly supported bases whose functions are not derivable on $\mathbb{R}$. This is why we propose the Hermite basis, whose functions are $\mathbb{R}$-supported and regular, both for φ_j and θ_ℓ.

The Hermite polynomial and the Hermite function of order j are given, for $j \geq 0$, by:

$$H_j(x) = (-1)^j \mathrm{e}^{x^2} \frac{d^j}{\mathrm{d}x^j}(\mathrm{e}^{-x^2}), \quad h_j(x) = c_j H_j(x) \mathrm{e}^{-x^2/2}, \quad c_j = \left(2^j j! \sqrt{\pi}\right)^{-1/2} \tag{3.10}$$

The sequence $(h_j, j \geq 0)$ is an orthonormal basis of $\mathbb{L}^2(\mathrm{d}x)$. Hermite functions with odd (resp. even) index are odd (resp. even) and $h_j^*(x) = \sqrt{2\pi} i^j h_j(x)$, where h_j^* is the Fourier transform of h_j. Moreover (see Abramowitz and Stegun (1964, 22.14.17), [56, 57], p. 242, [64]), $\|h_j\|_\infty \leq \Phi_0, \Phi_0 \simeq 1/\pi^{1/4} \simeq 0.7511$, so that the basis $(\varphi_j = h_j)$ is constituted of bounded functions which satisfy the assumptions of Lemma 2.3.

Moreover, the following result holds.

Proposition 3.1. *The Hermite basis satisfies the identifiability condition (3.9).*

Lastly, it is proved in Lemma 1 of [58] that reminding of definition (3.3), it holds

$$L_h(m) \leq C\sqrt{m}, \tag{3.11}$$

which implies the quantity is finite.

3.5. Definition of the estimators

To define a nonparametric estimator of b, ϕ, we estimate the functions b_m^V, ϕ_p^V, *i.e.* we build estimators $\widehat{\beta}_0, \ldots, \widehat{\beta}_{m-1}$ and $\widehat{c}_0, \ldots, \widehat{c}_{p-1}$ of the coefficients $\beta_0, \ldots, \beta_{m-1}$ and $c_0, \ldots, c_{p-1}$ and set

$$\widehat{b}_m(x) = \sum_{j=0}^{m-1} \widehat{\beta}_j \varphi_j(x), \quad \widehat{\phi}_p(x) = \sum_{\ell=0}^{p-1} \widehat{c}_\ell \theta_\ell(x), \quad x \in \mathbb{R}, \tag{3.12}$$

with $\widehat{\mathbf{b}}_m = (\widehat{\beta}_0, \ldots, \widehat{\beta}_{m-1})^{\perp}$ and $\widehat{\mathbf{c}}_p = (\widehat{c}_0, \ldots, \widehat{c}_{p-1})^{\perp}$.

We assume that we have at our disposal four independent paths of (2.6), $(X^{(i)}(t))_{t\in[0,T]}$, for $i=1,2,3,4$, *i.e.*

$$dX^{(i)}(t) = -(b(X^{(i)}(t)) + \phi\star\rho(X^{(i)}t)))\mathrm{d}t + \sigma dW_i(t), \quad X^{(i)}(0) = \eta_i \sim \rho(x)\mathrm{d}x,$$

where (W_i) are independent Brownian motions, (η_i) are independent random variables, independent of the Brownian motions (W_i). We indicate below why we could restrict us to two trajectories.

For $m,p\geq 0$, we define an empirical version of the theoretical matrix $\mathbf{V}_{m,p}$ and of the vector vector $\mathbf{Z}_{m,p}$, denoted by $\widehat{\mathbf{V}}_{m,p}$ and $\widehat{\mathbf{Z}}_{m,p}$. The matrix $\widehat{\mathbf{V}}_{m,p}$ is built using the two sample paths $(X^{(i)}(t))_{t\in[0,T]}$, for $i=1,2$ and the vector $\widehat{\mathbf{Z}}_{m,p}$ is built using the two sample paths $(X^{(i)}(t))_{t\in[0,T]}$, for $i=3,4$. In this way, the estimators $\widehat{\mathbf{V}}_{m,p}$ and $\widehat{\mathbf{Z}}_{m,p}$ are independent. We could use only two sample paths to estimate both, but this independence simplifies considerably the study.

Let us now define $\widehat{\mathbf{V}}_{m,p}$ and $\widehat{\mathbf{Z}}_{m,p}$. For this, we introduce the following notation. For a function h, let us set

$$\overline{h}(x) = \frac{1}{T}\int_0^T h(x - X^{(2)}(t))\mathrm{d}t, \quad \overline{\overline{h}}(x) = \frac{1}{T}\int_0^T h(x - X^{(4)}(t))\mathrm{d}t.$$

Note that $\mathbb{E}\overline{h}(x) = \mathbb{E}\overline{\overline{h}}(x) = h\star\rho(x)$ and $h\star\rho(X^{(1)}(s))$ can be estimated by $\overline{h}(X^{(1)}(s)) = \frac{1}{T}\int_0^T h(X^{(1)}(s) - X^{(2)}(t))\mathrm{d}t$, using two trajectories. Although $\overline{h}(x), \overline{\overline{h}}(x)$ are random and depend respectively on $X^{(2)}$ and $X^{(4)}$, we omit this dependence in the notation for the sake of simplicity. We set

$$\widehat{\mathbf{V}}_{m,p} = \begin{pmatrix} \widehat{\mathbf{V}}^1 & \widehat{\mathbf{V}}^{1,2} \\ (\widehat{\mathbf{V}}^{1,2})^{\perp} & \widehat{\mathbf{V}}^2 \end{pmatrix}$$

with

$$\begin{aligned}
\widehat{\mathbf{V}}^1 &= \left(\frac{1}{T}\int_0^T \varphi_j(X^{(1)}(s))\varphi_k(X^{(1)}(s))\mathrm{d}s\right)_{0\leq j,k\leq m-1} \\
\widehat{\mathbf{V}}^{1,2} &= \left(\frac{1}{T}\int_0^T \overline{\theta}_\ell(X^{(1)}(s))\,\varphi_k(X^{(1)}(s))\mathrm{d}s\right)_{0\leq k\leq m-1, 0\leq \ell\leq p-1} \\
\widehat{\mathbf{V}}^2 &= \left(\frac{1}{T}\int_0^T \overline{\theta}_\ell(X^{(1)}(s))\,\overline{\theta}_r(X^{(1)}(s))\mathrm{d}s\right)_{0\leq \ell,r\leq p-1}.
\end{aligned}$$

The matrix $\widehat{\mathbf{V}}_{m,p}$ is built using the sample paths $(X^{(i)}(t))_{t\in[0,T]}$, for $i=1,2$ as for $\ell=0,\dots,p-1$,

$$\overline{\theta}_\ell(X^{(1)}(s)) = \frac{1}{T}\int_0^T \theta_\ell(X^{(1)}(s) - X^{(2)}(t))\mathrm{d}t.$$

We stress here is that we must use two independent sample paths to estimate the elements of $\mathbf{V}_{m,p}$. This is due to the fact that ρ is unknown.

Next, we set

$$\widehat{\mathbf{Z}}_{m,p} = \frac{\sigma^2}{2}\begin{pmatrix} \left(\frac{1}{T}\int_0^T \varphi_k'(X^{(3)}(s))\mathrm{d}s\right)_{0\le k\le m-1} \\ \left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s\right)_{0\le \ell\le p-1}\end{pmatrix},$$

which depends on the sample paths $(X^{(i)}(t))_{t\in[0,T]}$, for $i=3,4$, as, for $\ell=0,\dots,p-1$,

$$\overline{\overline{\theta'}}_\ell(X^{(3)}(s)) = \frac{1}{T}\int_0^T \theta_\ell'(X^{(3)}(s)-X^{(4)}(t))\mathrm{d}t.$$

Here also, two independent trajectories are required to estimate the elements of $\mathbf{Z}_{m,p}$. We could have used the same $X^{(1)}, X^{(2)}$ as for $\mathbf{V}_{m,p}$. But, the independence of the estimators built with four independent trajectories simplifies considerably proofs (see the Remark at the end of the proof of Them. 3.3).

Moments strategies often bring unbiased estimators. Here, due to the use of the functions $\overline{h}(x), \overline{\overline{h}}(x)$ which are unbiased estimators of $h\star\rho$, it is almost the case, and we can prove:

Proposition 3.2. *Under Assumptions [H1]–[H4], it holds that*

$$\mathbb{E}(\widehat{\mathbf{Z}}_{m,p}) = \mathbf{Z}_{m,p}, \quad \mathbb{E}(\widehat{\mathbf{V}}_{m,p}) = \mathbf{V}_{m,p} + \begin{pmatrix} \mathbf{0}_{m\times m} & \mathbf{0}_{m\times p} \\ \mathbf{0}_{p\times m} & O\left(\frac{1}{T}\right)\mathbb{1}_{p\times p}\end{pmatrix},$$

where $\mathbb{1}_{p\times p}$ denote the $p\times p$ matrix with all coefficients equal to 1.

Thus $\widehat{\mathbf{Z}}_{m,p}$ is unbiased and $\widehat{\mathbf{V}}_{m,p}$ is asymptotically unbiaised when T grows to infinity.

Provided that $\widehat{\mathbf{V}}_{m,p}$ is invertible, using notation (3.1), relation (3.5) suggests to define the estimator of $(\mathbf{b},\mathbf{c})_{m,p}$ by

$$\widehat{(\mathbf{b},\mathbf{c})}_{m,p} = \widehat{\mathbf{V}}_{m,p}^{-1}\widehat{\mathbf{Z}}_{m,p} = \begin{pmatrix}\widehat{\mathbf{b}}_m \\ \widehat{\mathbf{c}}_p\end{pmatrix}.$$

Although under **[H5]**, the theoretical matrix $\mathbf{V}_{m,p}$ is invertible with minimum eigenvalue away from 0, this does not guaranteee the invertibility of $\widehat{\mathbf{V}}_{m,p}$. Therefore, to get a proper definition of our estimator, we propose the standard cutoff strategy leading to replace $\widehat{\mathbf{V}}_{m,p}^{-1}$ by

$$\widetilde{\mathbf{V}}_{m,p}^{-1} := \widehat{\mathbf{V}}_{m,p}^{-1}\mathbf{1}_{\Delta_{m,p}} \quad \text{with} \quad \Delta_{m,p} = \{\|\widehat{\mathbf{V}}_{m,p}^{-1}\|_{\mathrm{op}} \le 2s_0\}. \tag{3.13}$$

Then we set

$$\widetilde{(\mathbf{b},\mathbf{c})}_{m,p} = \widetilde{\mathbf{V}}_{m,p}^{-1}\widehat{\mathbf{Z}}_{m,p} = \begin{pmatrix}\widetilde{\mathbf{b}}_m \\ \widetilde{\mathbf{c}}_p\end{pmatrix}. \tag{3.14}$$

where

$$\widetilde{\mathbf{b}}_m = (\widetilde{\beta}_0,\dots,\widetilde{\beta}_{m-1})^\perp, \quad \widetilde{\mathbf{c}}_p = (\widetilde{c}_0,\dots,\widetilde{c}_{p-1})^\perp.$$

The estimator of (b, ϕ) is thus given by

$$\widetilde{b}_m(x) = \sum_{j=0}^{m-1} \widetilde{\beta}_j \varphi_j(x), \quad \widetilde{\phi}_p(x) = \sum_{\ell=0}^{p-1} \widetilde{c}_\ell \theta_\ell(x), \quad x \in \mathbb{R}.$$

By construction, $\widetilde{\mathbf{V}}_{m,p}^{-1}$ and $\widehat{\mathbf{Z}}_{m,p}$ are independent. The risk of the estimator $(\widetilde{b}_m, \widetilde{\phi}_p)$ is defined by: $\mathbb{E}(\|(\widetilde{b}_m - b, \widetilde{\phi}_p - \phi)\|_V^2)$, (see (1.3), (1.4)).

3.6. Study of estimator risk

In the sequel, we use the following notations.

Set $(b, \phi)(x) = (b(x), \phi(x))$, and $(b, \phi)_{m,p}(x) = (b_m^V(x), \phi_p^V(x))$ where b_m^V and ϕ_p^V given by (3.4). Analogously, define

$$\widetilde{(b, \phi)}_{m,p}(x) = (\widetilde{b}_m(x), \widetilde{\phi}_p(x)),$$

and recall that (see (1.3)) $\|(b, \phi)\|_V^2 = \|(b + \phi \star \rho\|_\rho^2$.

For $(\mathbf{x}, \mathbf{y})_{m,p} = \begin{pmatrix} \mathbf{x}_m \\ \mathbf{y}_p \end{pmatrix}$ a vector of $\mathbb{R}^m \times \mathbb{R}^p$, define the square norm associated with $\mathbf{V}_{m,p}$ defined by (3.6), by

$$\begin{aligned} \|(\mathbf{x}, \mathbf{y})_{m,p}\|_{\mathbf{V}_{m,p}}^2 := (\mathbf{x}, \mathbf{y})_{m,p}^\perp \mathbf{V}_{m,p} (\mathbf{x}, \mathbf{y})_{m,p} &= \int_{\mathbb{R}} \left[\sum_{j=0}^{m-1} x_j \varphi_j(x) + \sum_{\ell=0}^{p-1} y_\ell \theta_\ell \star \rho(x) \right]^2 \rho(x) \mathrm{d}x \\ &\leq 2\|\rho\|_\infty \left[\int_{\mathbb{R}} \left(\sum_{j=0}^{m-1} x_j \varphi_j(x) \right)^2 \mathrm{d}x + \int_{\mathbb{R}} \left(\sum_{\ell=0}^{p-1} y_\ell \theta_\ell \star \rho(x) \right)^2 \mathrm{d}x \right]. \end{aligned}$$

According to Young's inequality, we have

$$\|\sum_{\ell=0}^{p-1} y_\ell \theta_\ell \star \rho\|^2 \leq \|\rho\|_1^2 \|\sum_{\ell=0}^{p-1} y_\ell \theta_\ell\|^2 = \|\sum_{\ell=0}^{p-1} y_\ell \theta_\ell\|^2.$$

Now using that the functions $(\varphi_j)_{0 \leq j \leq m-1}$ and $(\theta_\ell)_{0 \leq \ell \leq p-1}$ are orthonormal in $\mathbb{L}^2(\mathrm{d}x)$, we get

$$\|(\mathbf{x}, \mathbf{y})_{m,p}\|_{\mathbf{V}_{m,p}}^2 \leq 2\|\rho\|_\infty \|(\mathbf{x}, \mathbf{y})_{m,p}\|_{\mathbb{R}^{m+p}}^2. \tag{3.15}$$

Now, we can link the V-norm of a couple of functions to the $\|.\|_{\mathbf{V}_{m,p}}$−norm of a vector of $\mathbb{R}^m \times \mathbb{R}^p$. Indeed, consider two functions $\omega = \sum_{j=0}^{m-1} \omega_j \varphi_j \in S_m$ and $\zeta = \sum_{\ell=0}^{p-1} \zeta_\ell \theta_\ell \in \Sigma_p$ with coefficients respectively $\boldsymbol{\omega}_m = (\omega_j, 0 \leq j \leq m-1)^\perp$ and $\boldsymbol{\zeta}_p = (\zeta_\ell, 0 \leq \ell \leq p-1)^\perp$. We have

$$\|(\omega, \zeta)\|_V^2 = (\boldsymbol{\omega}, \boldsymbol{\zeta})_{m,p}^\perp \mathbf{V}_{m,p} (\boldsymbol{\omega}, \boldsymbol{\zeta})_{m,p} = \|(\boldsymbol{\omega}, \boldsymbol{\zeta})_{m,p}\|_{\mathbf{V}_{m,p}}^2.$$

Thus, using (3.15), we have

$$\|(\omega, \zeta)\|_V^2 \leq 2\,\|\rho\|_\infty \|(\omega, \zeta)\|^2$$

where $\|(\omega, \zeta)\|^2 = \sum_{j=0}^{m-1} \omega_j^2 + \sum_{\ell=0}^{p-1} \theta_\ell^2 = \|(\boldsymbol{\omega}, \boldsymbol{\zeta})_{m,p}\|_{\mathbb{R}^{m+p}}^2$.

The following theorem gives a bound for the risk of the estimator $\widetilde{(b,\phi)}_{m,p}(x) = (\widetilde{b}_m(x), \widetilde{\phi}_p(x))$ with fixed m, p.

Theorem 3.3. *Assume that Assumptions* **[H1]**–**[H5]** *hold and that* $L_\varphi(m) + L_\theta(p) \le T$, *with* $L_\varphi(m)$ *and* $L_\theta(p)$ *defined in (3.3). Then*

$$\mathbb{E}\left(\|\widetilde{(b,\phi)}_{m,p} - (b,\phi)\|_V^2\right) \le \inf_{(u,v)\in S_m\times\Sigma_p} \|(b-u,\phi-v)\|_V^2 + C\,\frac{L_\varphi(m)+L_\theta(p)}{T},$$

where C *is a constant depending on* b, ϕ, ρ, s_0.

Let us comment this result. Our loss function is non standard. Let us give some explanations about it. If we consider the usual risk $\mathbb{E}(\|\widetilde{b}_m - b\|_\rho^2 + \|\widetilde{\phi}_p - \phi\|_\rho^2)$, a choice that we first investigated, then the bound we obtain involves the sum of three different bias terms:

$$\|b_m^\rho - b\|_\rho^2 + \|\phi_p^\rho - \phi\|_\rho^2 + \|(\phi_p^\rho - \phi)\star\rho\,\rho\|^2$$

where b_m^ρ (resp. ϕ_p^ρ) is the $\mathbb{L}^2(\rho)$-projection of b on S_m (resp. of ϕ on Σ_p). We do not know if the third term tends to 0 when p grows to infinity. This phenomenon is inherent to the problem. Indeed, it is worth stressing that [34], in the case $b = 0$, have a similar bias term composed of two terms, one of which may not tend to zero.

This is why we proposed to choose another loss function. The choice of our specific loss function induces a coherent result in Theorem 1: the risk of the estimator measured in norm $\|.\|_V$ is bounded by a variance term + a bias term measured in the same norm $\|.\|_V$. The bias term is decreasing as m, p increase; but, analogously, it is true that we do not know if it tends to zero. This confirms that we are facing a specific problem with this model.

The risk bound in Theorem 3.3 shows an usual decomposition into the bias term $\inf_{(u,v)\in S_m\times\Sigma_p}\|(b-u,\phi-v)\|_V^2$ and a variance term $(L_\varphi(m)+L_\theta(p))/T$ (see (3.3)) which increases with m, p. This decomposition may be used to realize the square bias-variance compromise provided that the rate of the bias term can be assessed on specific regularity spaces.

We need the following Propositions 3.4, 3.5 and Lemma 3.6 to prove Theorem 3.3. We state them to make clear the steps of the study of the risk.

Proposition 3.4. *Under* **[H1]**–**[H4]**, *it holds*

$$\mathbb{E}[\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2] \lesssim \frac{L_\varphi(m)+L_\theta(p)}{T}.$$

Proposition 3.5. *Under* **[H1]**–**[H4]**, $\mathbb{E}\left[\mathrm{Tr}(\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p})^2\right] \lesssim \dfrac{L_\varphi(m)}{T}$.

Therefore, we have

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right) \le \mathbb{E}\left[\mathrm{Tr}(\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p})^2\right] \lesssim \frac{L_\varphi(m)}{T}.$$

Note that this bound does not depend on p.

Lemma 3.6. *Under* **[H1]**–**[H5]**,

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\right) \le (40+1/4)s_0^4\,\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right)$$

Consequently,

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\right)\lesssim\frac{L_\varphi(m)}{T}.$$

Remark 3.7. Following ideas in Genon-Catalot and Larédo (2023a,b), [41, 42], we might have proposed $T^{-1}\int_0^T\varphi_\ell\star\widehat{\rho}_{[T]}(X^{(1)}(s)\varphi_r\star\widehat{\rho}_{[T]}(X^{(1)}(s)\mathrm{d}s$ instead of the term in $\widehat{\mathbf{V}}_2$, with $\widehat{\rho}_{[T]}$ computed with the path $X^{(2)}$ as in Section 4, formula (4.1) ($[T]$ denotes the integer part of T). It could be used in the definition of $\widehat{\mathbf{Z}}_{m,p}$ also, but would yield additional bias, compared to our strategy. In other words $\bar{\bar{h}}(x)$ is an unbiased estimator of $h\star\rho(x)$ which is not the case of $h\star\widehat{\rho}_{[T]}(x)$.

3.7. Risk bound when using the Hermite basis

Let us discuss the respective orders of the bias and variance terms in the bound obtained in Theorem 3.3, when considering the Hermite basis described in Section 3.4.

The variance term order depends on the bases.

In the case we use Hermite functions $(h_0,h_1,\dots,h_{m-1})$ to span S_m and $(h_1,h_3,\dots,h_{2p-1})$ to span Σ_p (recall that ϕ is odd), by (3.11), the variance term has order

$$\frac{L_\varphi(m)+L_\theta(p)}{T}\le C\frac{\sqrt{m}+\sqrt{p}}{T}.$$

Let us now discuss the square bias term. If the functions b and ϕ were square integrable on $\mathbb{R}$, we would upper bound the square bias by

$$\|\rho\|_\infty\left(\inf_{u\in S_m}\|b-u\|^2+\inf_{v\in\Sigma_p}\|\phi-v\|^2\right)=\|\rho\|_\infty\left(\sum_{j\ge m}\langle b,\varphi_j\rangle^2+\sum_{\ell\ge p}\langle\phi,\theta_{2\ell+1}\rangle^2\right)$$

For b and ϕ in Sobolev-Hermite spaces with regularity s_1 and s_2 (see *e.g.* Comte and Genon-Catalot (2020), Sect. 3.4, [59]), the resulting order would be $m^{-s_1}+p^{-s_2}$ and the resulting bias rate would be $T^{-2s_\star/(2s_\star+1)}$ where $s_\star=\min(s_1,s_2)$. This is the optimal rate for estimating a function with regularity $s_\star$. Thus, our bias term is meaningful.

However, the order of our specific bias term taking into account that $b,\phi,\phi\star\rho$ are in $\mathbb{L}^2(\rho)$ and not in $\mathbb{L}^2(\mathrm{d}x)$ would require the definition of specific regularity spaces to assess the rate of

$$\inf_{u\in S_m}\|b-u\|_\rho^2+\inf_{v\in\Sigma_p}\|(\phi-v)\star\rho\|_\rho^2.$$

4. Estimation of the stationary density

As $\rho\in\mathbb{L}^2(\mathrm{d}x)$, we consider an orthonormal basis of $\mathbb{L}^2(\mathrm{d}x)$, still denoted $(\varphi_j)_{j\ge0}$ and write $\rho=\sum_{j\ge0}a_j\varphi_j$ with $a_j=\langle\varphi_j,\rho\rangle=\int_{\mathbb{R}}\varphi_j(x)\,\rho(x)\mathrm{d}x$. For an integer D, we consider the estimator

$$\widehat{\rho}_D=\sum_{j=0}^{D-1}\widehat{a}_j\varphi_j,\quad \widehat{a}_j=\frac{1}{T}\int_0^T\varphi_j(X(s))\mathrm{d}s. \tag{4.1}$$

We can prove the following result

Proposition 4.1. *Under Assumptions* **[H1]**–**[H4]**,

$$\mathbb{E}\left[\|\widehat{\rho}_D - \rho\|^2\right] \leq \|\rho - \rho_D\|^2 + \frac{C}{T},$$

where $\rho_D = \sum_{j=0}^{D-1} a_j \varphi_j$, C *is a positive constant (not depending on* D, T*) and* $\|.\|$ *is the* $\mathbb{L}^2$*-norm.*

It follows from Proposition 4.1 that the variance of the estimator does not depend on D. Therefore, taking D as large as possible will make the bias term negligible under weak regularity conditions, and ρ can be estimated with (parametric) rate $1/T$. In Comte and Merlevède [66], the same result is obtained under specific conditions, mainly the so-called Castellana and Leadbetter condition which is difficult to check in general. The interesting point here is that the parametric rate is obtained without such specific condition in our framework.

Remark 4.2. In most cases, as mentioned above, the invariant density has no explicit expression. However, in the special case where b is odd and ϕ is linear, ρ is explicit. We follow Hermann and Tugaut (2010) to explain this case. Let us assume that b is odd and $\phi(x) = cx$. Then, $\phi \star \rho(x) = c(x-m)$ where m is the expectation of ρ. The invariant density ρ is given by

$$\rho(x) = [K(m)]^{-1} \exp\left[-2\sigma^{-2}(B(x) + \frac{c}{2}(x-m)^2)\right]$$

where $[K(m)]^{-1}$ is the norming constant and m is solution of

$$\int_{\mathbb{R}} x \exp\left[-2\sigma^{-2}(B(x) + \frac{c}{2}(x-m)^2)\right]\mathrm{d}x = m \int_{\mathbb{R}} \exp\left[-2\sigma^{-2}(B(x) + \frac{c}{2}(x-m)^2)\right]\mathrm{d}x = mK(m)$$

As B is even, this equation has an obvious solution $m = 0$. As the invariant distribution is unique, we find

$$\rho(x) = \exp\left[-2\sigma^{-2}(B(x) + \frac{c}{2}x^2)\right]\left[\int_{\mathbb{R}} \exp\left[-2\sigma^{-2}(B(y) + \frac{c}{2}y^2)\right]\mathrm{d}y\right]^{-1}.$$

Moreover, equation (2.6) is equal to

$$dX(t) = -[b(X(t)) + cX(t)]\mathrm{d}t + \sigma dW(t), \quad X(0) \sim \rho(x)\mathrm{d}x.$$

Models 1,2,3 presented below in Section 5 correspond to such cases and we use them to evaluate the performances of the invariant density estimator.

5. Simulation experiments

Numerical simulations of processes given by mean-field limits is not a simple task. A relatively easy solution to get approximate sample paths is to simulate the system of particles for large N since we know that for any fixed p, the processes $(X_i^N(t), i = 1, \ldots, p)$, where $(X_i^N(t), i = 1, \ldots, N)$ is given by (1.2), converge in distribution as N tends to infinity to p *i.i.d.* McKean–Vlasov processes $(X^{(i)}(t))$, given by

$$dX^{(i)}(t) = -[b(X^{(i)}(t)) + \phi \star \mu_t(X^{(i)}(t))]\mathrm{d}t + \sigma dW_i(t), \quad X^{(i)}(0) = X_0^i, i = 1, \ldots, p.$$

Moreover, the particle system $(X_i^N(t), i = 1, \ldots, N)$ as a N-dimensional diffusion process, admits an invariant density, say $\nu^N(x_1, \ldots, x_N)$ and each marginal $\nu_i^N(x_i)$ defines a distribution which converges weakly to $\rho(x_i)\mathrm{d}x_i$ (see *e.g.* [21, 22]. Therefore, we simulated the system of particles on $[-10, T]$, left out the interval $[-10, 0]$ and picked out four trajectories on $[0, T]$.

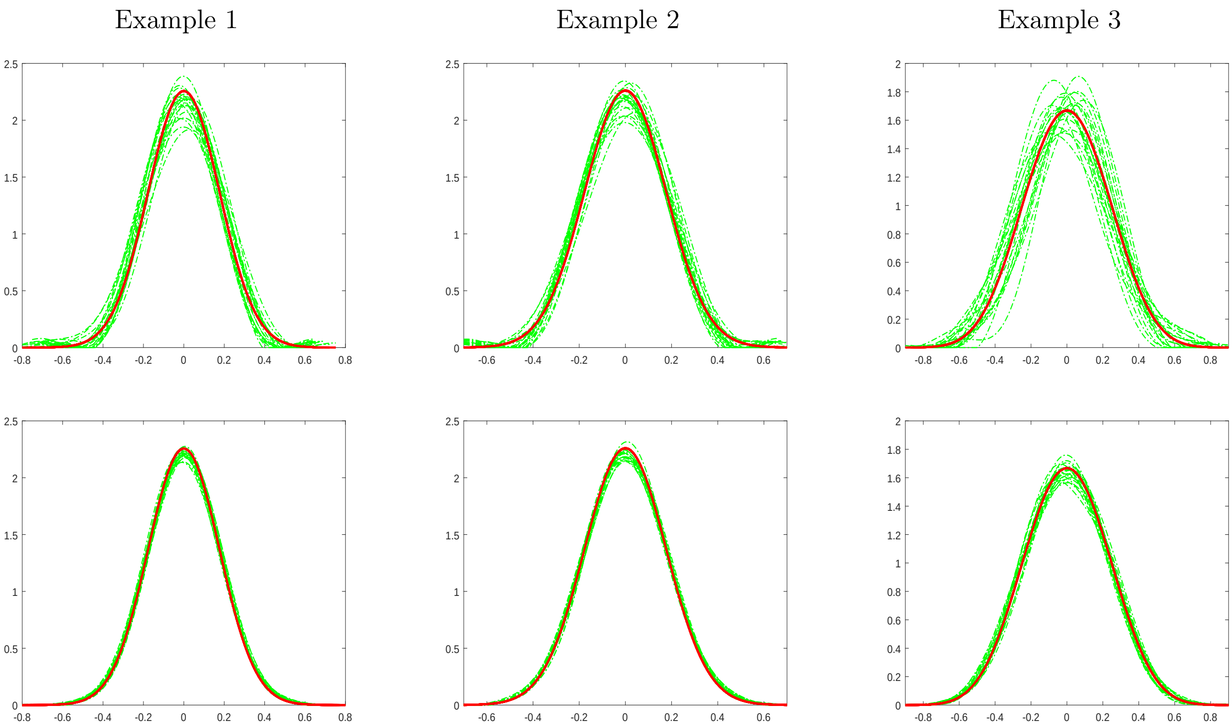

FIGURE 1. Density estimation for Examples 1 to 3. 20 estimated curves in dotted green, first line $T = 200$ ($D = 70$), last line $T = 2000$ ($D = 225$). In all cases, the true is in bold red.

We implemented the method by simulating $N = 800$ particules following equation (1.2) with Euler-type discretization scheme with step $\Delta = 0.1$. The initial value is taken equal to 0 and we exclude the 100 first values of the process to reach approximately the stationary regime. The values of T presented below are $T = 200$ and $T = 2000$, involving $n = 2000$ and $n = 20\,000$ generated $X_i^N(j\Delta)$, $j = 1, \dots, n$ and $i = 1, \dots, N$. In all cases, we took $\sigma = 0.25$. The estimation uses the sample paths $X_i^N(j\Delta), j = 1, \dots, n$ for $i = 1, \dots, 4$.

First, we present three models for which we can compute the true density ρ so that we can illustrate the performance of the density estimator described in Section 4 (see Rem. 4.2). With $b = 0.5, c = 0.5$, we consider

1. $b(x) = bx$, $\phi(x) = cx$, $\rho(x)$ is a $\mathcal{N}(0, \sigma^2/[2(b + c)])$,
2. $b(x) = b\tanh(x)$, $\phi(x) = cx$, $\rho(x) = c_1 \exp(-cx^2/\sigma^2)/(\cosh(x))^{2b/\sigma^2}$,
3. $b(x) = bx^3$, $\phi(x) = cx$, $\rho(x) = c_2 \exp\left[-(bx^4/2 + cx^2)/\sigma^2\right]$.

The constants c_1 and c_2 are computed numerically so that the function ρ integrates to one. We compute estimators corresponding to the Hermite basis $(\varphi_j = h_j)_{0\le j\le D-1}$ (see (3.10)). The dimension is chosen equal to $D = 5[\sqrt{T}]$, that is 70 for $T = 200$ and 225 for $T = 2000$. Figure 1 shows 20 estimates of the invariant density in dotted green and the true one in bold red. The choice of the parameters b, c seems important as it influences the simulation results. Note that choosing D too large ($D = T$, *e.g.*) implies numerical problems. We can see on Figure 1 that the method works very well and is improved when T grows. We observe that it is more difficult to estimate the invariant density in Example 3: the estimation of ρ is much better for $T = 2000$ ($D = 225$) than for $T = 200$ ($D = 70$). The three densities, though different, have very similar forms.

Figure 2 presents the estimation of b and ϕ for Example 2. Our protocol of estimation is detailed below. Two other models are illustrated hereafter where b is not odd and/or ϕ is not linear:

4. $b(x) = 2(x - 1)^3 + 2x$, $\phi(x) = 4\tanh(x)$,

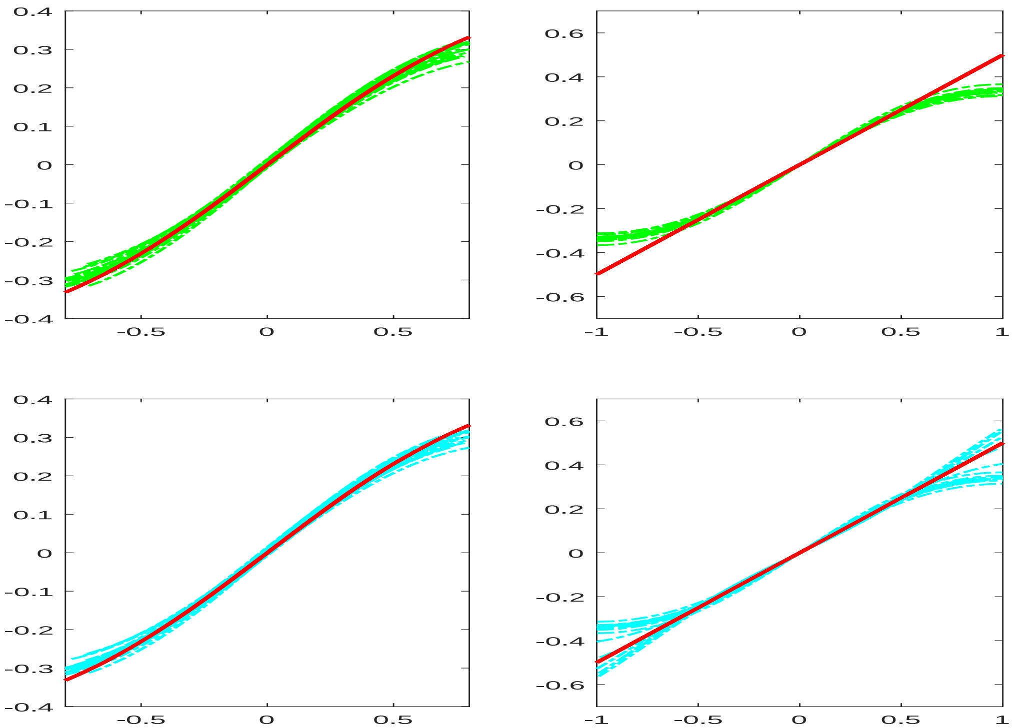

FIGURE 2. Example 2, $T = 2000$. 20 estimated b (left) and ϕ (right), all selected by penalization (top green) or as oracles (bottom cyan). The true functions are in bold red.

5. $b(x) = 2x$, $\phi(x) = 2(x + x^3)$.

To compute the estimators of b, we use the Hermite basis $(\varphi_j)_{0\le j\le m-1}$ for $m = 1, \ldots, 6$ (see Sect. 3.4). To compute the estimators of ϕ, we need to take into account that ϕ is odd. So we consider the basis $(\theta_\ell)_{0\le \ell\le p-1} = (h_{2\ell+1})_{0\le \ell\le p-1}$, as Hermite functions of odd index are odd functions. We consider dimensions $m \in \{1, \ldots, m_{\max} = 6\}$ and $p \in \{1, \ldots, p_{\max} = 3\}$. Larger proposals systematically involve estimators with an obvious variance effect. The fact that we consider small dimensions replaces the cutoff of $\widehat{\mathbf{V}}_{m,p}^{-1}$.

For examples 1 to 3, we could have simulated by Euler scheme four independent sample paths of $dX(t) = -(b(X(t)) + cX(t))dt + \sigma dB(t)$ instead of the particles system. We found that the estimation results were of the same type, so we kept the particles system which is suitable for all models.

Though we have no theoretical result concerning an adaptive choice for the dimensions m, p, we propose a criterion following the standard method to select data-driven dimensions. The selection of (m, p) is done by choosing the couple which minimizes

$$Crit(m,p) = -\widehat{(\mathbf{b},\mathbf{c})}_{m,p}^{\perp}\widehat{\mathbf{Z}}_{m,p} + \kappa \times \frac{\sqrt{m} + 4\sqrt{p}}{T},$$

where $\widehat{(\mathbf{b},\mathbf{c})}_{m,p}^{\perp}\widehat{\mathbf{Z}}_{m,p} = \|\widehat{(\mathbf{b},\mathbf{c})}_{m,p}\|^2_{\widehat{\mathbf{V}}_{m,p}} = \widehat{\mathbf{Z}}_{m,p}^{\perp}\widehat{\mathbf{V}}_{m,p}^{-1}\widehat{\mathbf{Z}}_{m,p}$. The criterion is inspired by an estimation of the empirical bias, as usually performed for least-squares contrast minimization, and a second term which has the order of the variance of the estimator. The constant κ is roughly calibrated and chosen here as $\kappa = 5$.

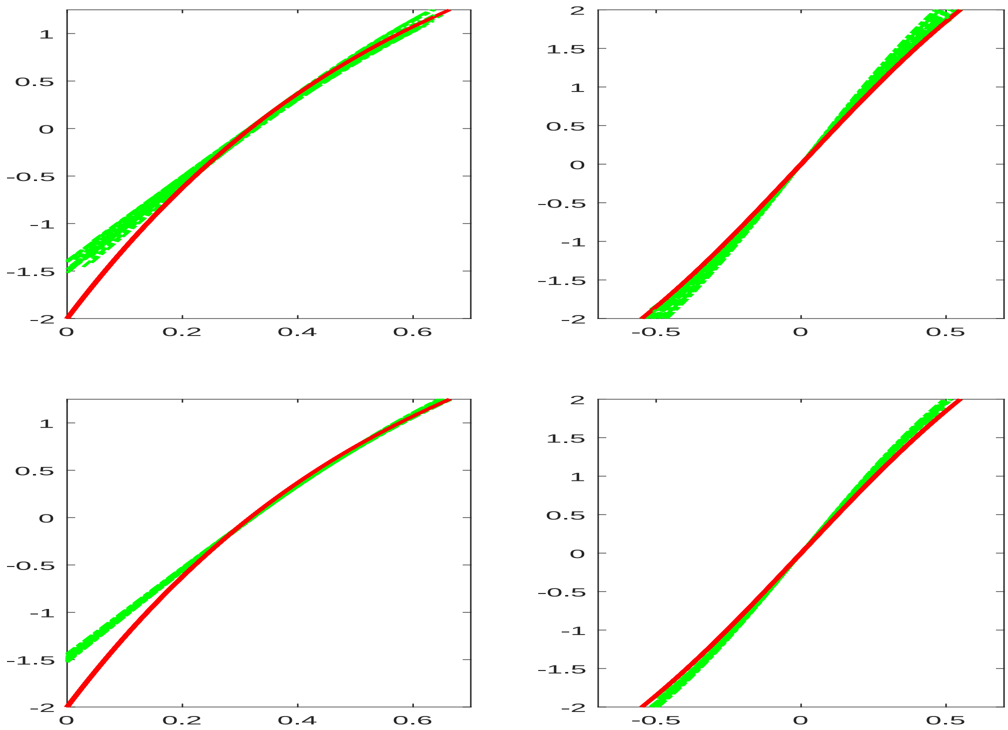

FIGURE 3. Example 4, estimation of b (left) and ϕ (right) for $T = 200$ (top) and $T = 2000$ (bottom). Dimensions selected by penalization. In all cases, the true is in bold red.

Figure 2 displays the results obtained for Example 2: we simulated 20 repetitions of independent particle systems up to $T = 2000$. The two top pictures show the 20 estimators for b (left), ϕ (right) obtained using our proposed selection device. The two pictures below present the corresponding oracle estimators. By "oracle", we mean that we compute the $\mathbb{L}^2$-distance between each estimator of our list and the true function (b or ϕ), and select the best one. This knowledge is clearly unavailable in practice and the oracle is only dedicated to bring a benchmark to evaluate the penalization strategy. We can see that the selection method performs very well and that, most of the time, it behaves as well as the oracle.

Figure 3 illustrates, for Example 4, the improvement which occurs when increasing T from 200 to 2000, for the estimation of b and ϕ: it is slightly disappointing, in this case, but does exist.

Figure 4 concerns Example 5 simulated up to $T = 200$. The two pictures on top display all the proposals for estimating b (left) and ϕ(right) with all the couples (m, p) of the list. We observe that some of these functions are obviously very bad estimators. This shows that the selection of a relevant estimator is crucial. In the two pictures below, we have plotted the estimator selected by model selection (dotted green), and the oracle (dotted black). While these estimators are the same for b (left), there is a noticeable difference for ϕ (right). Note that, for most generated paths, they coincide. As Hermite bases are parsimonious, small values of m and p perform well. The choice $m = 1$ is generally not good, and the most selected couple for (m, p) is $(2, 1)$.

To conclude, while the performance of the invariant density is really nice, we observe that estimating b and ϕ is obviously a more difficult problem. In any case, the choice of the functions b, ϕ is important because too small or too large ranges of values for the process can generate numerical difficulties. Lastly, we used four particles, following the procedure described in the theoretical part, but we also verified that using only two particles does not deteriorate the results.

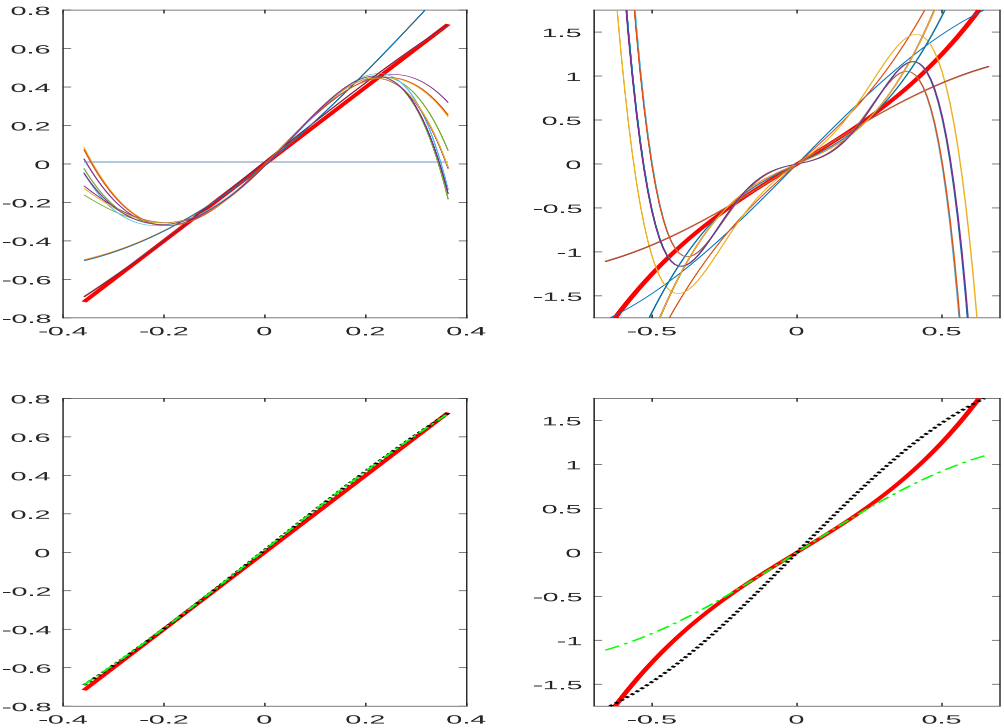

FIGURE 4. Example 5, $T = 200$. First line: the 6x3 proposals for b (left) and the 3x6 proposals for ϕ (right). Second line: $\widehat{b}_2$ (left) and $\widehat{\phi}_1$ (right) selected by penalization (green) or oracle (dotted black). In all cases, the true is in bold red.

6. Concluding remarks

6.1. Summary and remaining questions

In this paper, we consider the nonparametric estimation of the functions b, ϕ for model (1.1) based on the continuous observation on a time interval $[0, T]$ of four *i.i.d.* sample paths in stationary regime. The asymptotic framework is $T \to +\infty$. Due to the presence of the interaction term in equation (1.1) which contains a convolution with the unknown marginal law of the process, the problem is completely different from the case of a usual diffusion ($\phi \equiv 0$) and the observation of only one sample path is not enough to estimate both b and ϕ. The novelty here is to use independent paths to get estimators of terms such as $h \star \rho$ (where ρ is the invariant density) together with other crucial quantities. Actually, two sample paths could be enough but induces much more tedious computations. The estimation method is inspired by a paper of Pavliotis and Zanoni [43] for interacting particle systems in a parametric setting. To build the estimators, we use a combination of a moment method and a projection method. Defining a specific risk for the estimators fitted to the problem, we obtain a bound for the risk of our projection estimators involving as usual a variance term and a square bias term. The order of the variance term is precised for the Hermite basis. The rate of the square bias term, which is very specific to the problem, is not simple to evaluate. We sketch a discussion on this point. Note that no benchmark is available for comparison. We also provide a nonparametric estimator of the invariant density of the process which has the parametric rate without any condition such as Castellana and Leadbetter's one. The method is implemented on simulated data using an empirical model selection criterion, and performs reasonably well.

Remaining questions are worth of further research. It would be of interest to propose a theoretical background for assessing the rate of the bias term and for the selection procedure.

Moreover, the estimation procedure is based on the estimation of the vector $\mathbf{Z}_{m,p}$ suggested by formula (3.8), using the moment relation (3.2). But the vector $\mathbf{Z}_{m,p}$ is also given by (3.7) and could have been estimated by

$$\widehat{\widehat{\mathbf{Z}}}_{m,p} = \left(\left(\frac{1}{T}\int_0^T \varphi_k(X^{(3)}(s))dX^{(3)}(s)\right)^{\perp}_{0\leq k\leq m-1}, \left(\frac{1}{T}\int_0^T \overline{\overline{\theta}}_\ell(X^{(3)}(s))dX^{(3)}(s)\right)^{\perp}_{0\leq \ell\leq p-1}\right)^{\perp}$$

keeping the same matrix $\widehat{\mathbf{V}}_{m,p}$. The other estimator $\widetilde{\mathbf{V}}_{m,p}^{-1}\widehat{\widehat{\mathbf{Z}}}_{m,p}$ requires another study which is worth being done.

6.2. Discrete observations

Considering discrete observations is a natural expansion of this study as in practice, only discrete time observations are available. Therefore, it is worth of interest to study the same estimation problem based on discrete observations $(X^{(i)}(j\Delta), j = 1\dots,n, i = 1,\dots,4)$ with $\Delta = \Delta_n$ tending to 0 and $n\Delta_n$ tending to infinity (high frequency framework). One can consider the discretized versions of our estimators and this is actually what is done in the simulation section. The extension of the estimation results would be to the price of more computations and certainly a specific study. We note that in [62], it is proved that results concerning the nonparametric estimation of the invariant density based on discrete observations cannot be deduced from the continuous case but have to be studied explicitly, at least in dimension $\mathrm{d} \geq 2$. An additional term due to the discretization is present in the variance term. Therefore, nonparametric estimation of ergodic McKean–Vlasov SDE based on discrete observations should raise new questions and its investigation is an important topic.

6.3. Multidimensional McKean–Vlasov models

The case of a multidimensional McKean–Vlasov model is certainly of the utmost importance especially for applications. If the process evolves in $\mathbb{R}^d$, we have to estimate a drift $b = (b^1,\dots,b^d)$ and an interaction function $\phi = (\phi^1,\dots,\phi^d)$ which both are composed of d functions from $\mathbb{R}^d$ to $\mathbb{R}$. The specific difficulty for estimation of mutivariate functions by projection is the fact that, for each function from $\mathbb{R}^d$ to $\mathbb{R}$, we have to estimate a hypermatrix $(\beta_{j_1,j_2,\dots,j_d}, j_1 \leq m_1,\dots,j_\mathrm{d} \leq m_d)$ of coefficients, instead of a vector for univariate functions. The theory can be extended in a rather natural way, see [60], for the estimation of multivariate regression functions, or [61] for drift estimation of inhomogeneous diffusions. However, numerical implementation gets quickly more difficult to handle.

7. Proofs

7.1. Proof of Proposition 2.1

We have $b'(x) \geq K, \phi'(x) \geq \lambda$. Thus, $(\phi\star\rho)'(x) = \int \phi'(x-y)\rho(y)\mathrm{d}y \geq \lambda$. Therefore, for $x \geq 0$,

$$b(x) \geq Kx + b(0), \quad \phi\star\rho(x) \geq \lambda x + \phi\star\rho(0).$$

This implies, for $0 \leq z \leq x$,

$$\int_z^x [b(u) + \phi\star\rho(u)]\mathrm{d}u \geq (K+\lambda)(x^2-z^2)/2 + (b(0)+\phi\star\rho(0))(x-z).$$

Thus,

$$\frac{\rho(x)}{\rho(z)} \leq \exp\left[-\frac{(K+\lambda)}{\sigma^2}(x^2-z^2) + C(x-z)\right], \quad C = -(2/\sigma^2)(b(0)+\phi\star\rho(0)).$$

Analogously, for $x \le z \le 0$,

$$\frac{\rho(x)}{\rho(z)} \le \exp\,[-\frac{(K+\lambda)}{\sigma^2}(z^2-x^2)+C(x-z)].$$

Therefore, for all x,

$$\rho(x) \lesssim \exp\,[-\frac{(K+\lambda)}{\sigma^2}x^2+Cx] \propto \exp\left[-\frac{(K+\lambda)}{\sigma^2}\left(x-\frac{C\sigma^2}{2(K+\lambda)}\right)^2\right],$$

Moreover, for $0 \le z \le x$ or $x \le z \le 0$,

$$\frac{\rho(x)}{\rho(z)} \le \exp\,[C(x-z)]. \quad \square$$

7.2. Proof of Lemma 2.3

Set $\Psi(x) = \int_0^x \psi(t)\mathrm{d}t$. We prove that Ψ and $\int_0^{\cdot} \psi \star \rho$ both belong to $\mathcal{D}$ defined by (2.8). We have $\int_{\mathbb{R}} \Psi^2(x)\rho(x)\mathrm{d}x = \int_0^{+\infty} ... + \int_{-\infty}^0 ...$, where

$$\begin{aligned}\int_0^{+\infty} \Psi^2(x)\rho(x)\mathrm{d}x &\le \int_0^{+\infty} x\int_0^x \psi^2(t)\mathrm{d}t\rho(x)\mathrm{d}x = \int_{0<t<x} x\psi^2(t)\rho(x)dxdt \\ &= \int_0^{+\infty}\psi^2(t)(\int_t^{+\infty} x\rho(x)\mathrm{d}x)\mathrm{d}t \le \int_0^{+\infty}\psi^2(t)\mathrm{d}t\int_0^{+\infty} x\rho(x)\mathrm{d}x < +\infty.\end{aligned}$$

We proceed analogously for the integral on $(-\infty,0)$ and thus $\Psi \in \mathbb{L}^2(\rho)$.
To prove that $\int_0^x \psi \star \rho(t)\mathrm{d}t$ belongs to $\mathbb{L}^2(\rho)$, it is therefore enough to prove that $\psi \star \rho$ belongs to $\mathbb{L}^2(\mathrm{d}x)$. By the Young inequality,

$$\int_{\mathbb{R}}(\psi\star\rho)^2(x)\mathrm{d}x \le (\int_{\mathbb{R}}\rho(x)\mathrm{d}x)^2\int_{\mathbb{R}}\psi^2(x)\mathrm{d}x < +\infty.$$

Next, we must prove that $\psi(x)\rho(x)$ and $\psi\star\rho(x)\rho(x)$ tend to 0 as x tends to $+\infty$ and $-\infty$. This holds as ψ is bounded by assumption. And $\psi\star\rho$ is also bounded as

$$|\psi\star\rho(x)| \le \left(\int_{\mathbb{R}}\psi^2(t)\mathrm{d}t\int_{\mathbb{R}}\rho^2(y)\mathrm{d}y\right)^{1/2} \le \left(\|\rho\|_\infty\int_{\mathbb{R}}\psi^2(t)\mathrm{d}t\right)^{1/2}.$$

Now, we prove that $L\Psi$ and $L(\int_0^{\cdot}(\psi\star\rho))$ belong to $\mathbb{L}^2(\rho)$. We have $L\Psi = \frac{\sigma^2}{2}\psi' - (b+\phi\star\rho)\psi$. As ψ' has polynomial growth, $\int_{\mathbb{R}}[\psi'(x)]^2\rho(x)\mathrm{d}x < +\infty$ by Proposition 2.1. By our assumptions, $b = V'$ and $\phi\star\rho$ have ℓ polynomial growth. So, it is enough to check that $x \mapsto (1+x^\ell)\psi(x)$ belongs to $\mathbb{L}^2(\rho)$ which holds since ψ is bounded and ρ has moments of any order.

We have $L(\int_0^{\cdot}(\psi\star\rho)) = \frac{\sigma^2}{2}\psi'\star\rho - (b+\phi\star\rho)\psi\star\rho$. We use that $\psi\star\rho$ is bounded and $b+\phi\star\rho$ has polynomial growth to deduce that $(b+\phi\star\rho)\psi\star\rho$ belongs to $\mathbb{L}^2(\rho)$. Next, as ψ' has polynomial growth, $\psi'\star\rho$ has also polynomial growth. This implies that $\psi'\star\rho$ belongs to $\mathbb{L}^2(\rho)$. $\square$

7.3. Useful Lemma

We state here a useful Lemma.

Lemma 7.1. *Assume [H1]–[H4] and consider a function h such that there exists a positive constant a such that $\forall u \in \mathbb{R}, |h(u)| \leq a\mathrm{e}^{au}$. The following integrals are finite and bounded by constants depending on ρ and h only:*

(1) $I_- = \int_{-\infty}^{0} \left(\int_{-\infty}^{y} h(u)\rho(u)\mathrm{d}u \right)^2 \frac{\mathrm{d}y}{\rho(y)}, \quad I_+ = \int_0^{+\infty} \left(\int_u^{+\infty} h(u)\rho(u)\mathrm{d}u \right)^2 \frac{\mathrm{d}y}{\rho(y)}.$

(2) $J_- = \int_{x<y<0} \frac{|x|\rho(x)}{\rho^2(y)} \left(\int_{-\infty}^{y} h(u)\rho(u)\mathrm{d}u \right)^2 dxdy, J_+ = \int_{0<y<x} \frac{x\rho(x)}{\rho^2(y)} \left(\int_y^{+\infty} h(u)\rho(u)\mathrm{d}u \right)^2 dxdy.$

(3) $\int_{-\infty}^{0} (\int_{-\infty}^{x} \rho^2(y)\mathrm{d}y) \frac{\mathrm{d}x}{\rho(x)}, \quad \int_0^{+\infty} (\int_x^{+\infty} \rho^2(y)\mathrm{d}y) \frac{\mathrm{d}x}{\rho(x)}.$

(4) $\int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \frac{\mathrm{d}u}{\rho^2(u)} \int_u^{+\infty} \rho^2(y)\mathrm{d}y, \int_{-\infty}^{0} |x|\rho(x)\mathrm{d}x \int_x^0 \frac{\mathrm{d}u}{\rho^2(u)} \int_{-\infty}^{u} \rho^2(y)\mathrm{d}y.$

(5) $\int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \frac{\mathrm{d}u}{\rho^2(u)} \left(\int_u^{+\infty} \rho(y)\mathrm{d}y \right)^2, \int_{-\infty}^{0} |x|\rho(x)\mathrm{d}x \int_x^0 \frac{\mathrm{d}u}{\rho^2(u)} \left(\int_{-\infty}^{u} \rho(y)\mathrm{d}y \right)^2$

Proof of Lemma 7.1. We start with the first item. It is enough to look at I_-, the other one being analogous. We have:

$$I_- = 2 \int_{u<v<y<0} \frac{\mathrm{d}y}{\rho(y)} h(u)\rho(u)h(v)\rho(v)dudv.$$

On the integration set,

$$\frac{\rho(u)\rho(v)}{\rho(y)} = [\rho(u)\rho(v)]^{1/2} [\frac{\rho(u)}{\rho(y)}]^{1/2} [\frac{\rho(v)}{\rho(y)}]^{1/2} \leq [\rho(u)\rho(v)]^{1/2} \mathrm{e}^{(C/2)(u+v)-Cy}.$$

Thus,

$$\begin{aligned}
I_- &\leq 2 \int_{u<v<y<0} h(u)h(v)[\rho(u)\rho(v)]^{1/2} \mathrm{e}^{(C/2)(u+v)-Cy} dydudv \\
&= 2 \int_{u<v<0} h(u)h(v)[\rho(u)\rho(v)]^{1/2} \mathrm{e}^{(C/2)(u+v)} (\int_v^0 \mathrm{e}^{-Cy}\mathrm{d}y) dudv \\
&= \frac{2}{C} \int_{u<v<0} [\rho(u)\rho(v)]^{1/2} \mathrm{e}^{(C/2)u} (\mathrm{e}^{(C/2)v} - \mathrm{e}^{-(C/2)v}) h(u)h(v) dudv \\
&\leq \frac{2}{C} \int_{u<0} [\rho(u)]^{1/2} |h(u)| \mathrm{e}^{(C/2)u} \mathrm{d}u \int_{v<0} |(\mathrm{e}^{(C/2)v} - \mathrm{e}^{-(C/2)v}) h(v)| [\rho(v)]^{1/2} \mathrm{d}v \\
&< +\infty
\end{aligned}$$

by Proposition 2.1 and the assumption on h.

We prove the second item. We have:

$$J_- = 2 \int_{x<y<0, u<v<y} |x| \frac{\rho(x)\rho(u)\rho(v)}{\rho^2(y)} h(u)h(v) dxdydudv.$$

We can write:

$$\frac{\rho(x)\rho(u)\rho(v)}{\rho^2(y)} = [\rho(x)]^{1/3}[\frac{\rho(x)}{\rho(y)}]^{2/3}[\frac{\rho(u)}{\rho(y)}]^{2/3}[\frac{\rho(v)}{\rho(y)}]^{2/3}[\rho(u)\rho(v)]^{1/3}.$$

On the integration set, we have:

$$\frac{\rho(x)\rho(u)\rho(v)}{\rho^2(y)} \leq [\rho(x)\rho(u)\rho(v)]^{1/3}\mathrm{e}^{(2/3)C(x-y)}\mathrm{e}^{(2/3)C(u-y)}\mathrm{e}^{(2/3)C(v-y)}.$$

Thus,

$$J_- \leq \int_{x<0,u<v} |x||h(u)h(v)|[\rho(x)\rho(u)\rho(v)]^{1/3}\mathrm{e}^{(2C/3)(x+u+v)}(\int_{\sup(x,v)}^{+\infty} \mathrm{e}^{-2Cy}\mathrm{d}y)dxdudv < +\infty,$$

by Proposition 2.1 and the assumption on h.

We treat (3). By Proposition 2.1, and Inequality (2.4), we get straightforwardly that

$$\int_0^{+\infty} \frac{\mathrm{d}x}{\rho(x)} \int_x^{+\infty} \rho^2(y)\mathrm{d}y = \int_0^{+\infty} \rho^2(y)\left(\int_0^y \frac{\mathrm{d}x}{\rho(x)}\right)\mathrm{d}y < +\infty.$$

For (4), we have:

$$\int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \mathrm{d}u \frac{1}{\rho^2(u)} \int_u^{+\infty} \rho^2(y)\mathrm{d}y = \int_{x>0,0<u<x,0<u<y} x\rho(x)\frac{\rho^2(y)}{\rho^2(u)}dxdudy.$$

We can write:

$$\rho(x)\frac{\rho^2(y)}{\rho^2(u)} = [\rho(x)]^{1/3}[\frac{\rho(x)}{\rho(u)}]^{2/3}[\frac{\rho(y)}{\rho(u)}]^{4/3}[\rho(y)]^{2/3}$$

Therefore, using (2.4) again,

$$\begin{aligned}&\int_{x>0,0<u<x,0<u<y} x\rho(x)\frac{\rho^2(y)}{\rho^2(u)}dxdudy\\ &\leq \int_{x>0,0<u<x,0<u<y} x[\rho(x)]^{1/3}[\rho(y)]^{2/3}\exp\left[(2/3)C(x-u)\right]\exp\left[(4/3)C(y-u)\right]dxdudy\\ &\int_{x>0,y>0} x[\rho(x)]^{1/3}[\rho(y)]^{2/3}\exp\left[(2/3)Cx+(4/3)Cy\right]\int_0^{x\wedge y}\exp\left(-2u\right)\mathrm{d}udxdy < +\infty.\end{aligned}$$

For (5),

$$\int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \frac{\mathrm{d}u}{\rho^2(u)}\left[\int_u^{+\infty}\rho(y)\mathrm{d}y\right]^2 = \int \mathbf{1}_{0<u<x}\mathbf{1}_{u<y,u<y'}\frac{x\rho(x)\rho(y)\rho(y')}{\rho^2(u)}dxdydy'\mathrm{d}u.$$

As above, on the integration set,

$$\begin{aligned}\rho(x)\frac{\rho(y)\rho(y')}{\rho^2(u)} &= [\rho(x)]^{1/2}[\frac{\rho(x)}{\rho(u)}]^{1/2}[\frac{\rho^{3/4}(y)\rho^{3/4}(y')}{\rho^{3/2}(u)}][\rho(y)\rho(y')]^{1/4} \\ &\leq [\rho(x)]^{1/2}[\rho(y)]^{1/4}[\rho(y')]^{1/4}\mathrm{e}^{(C/2)x+(3C/4)(y+y')}\mathrm{e}^{-(5C/4)u},\end{aligned}$$

and conclude that the first integral of (5) is finite. We proceed analogously for the second one. □

7.4. Proof of Proposition 2.4

Recall that h having exponential growth, belongs to $\mathbb{L}^2(\rho)$ as noted before the proposition. Let us find $\Gamma_h = g$ such that $g \in \mathcal{D}$ and $Lg = h_c$. Using (2.7) and (2.8), we obtain, taking into account that we want $\lim_{|y|\to\infty} g'(y)\rho(y) = 0$,

$$g'(y) = \frac{2}{\sigma^2\rho(y)}\int_{-\infty}^{y} h_c(u)\rho(u)\mathrm{d}u = -\frac{2}{\sigma^2\rho(y)}\int_{y}^{+\infty} h_c(u)\rho(u)\mathrm{d}u,$$

where the equality above comes from the fact that $\int_{-\infty}^{+\infty} h_c(u)\mathrm{d}u = 0$. We choose $g(x) = \int_0^x g'(u)\mathrm{d}u$ which gives $Lg = h_c$. We must now prove

$$\mathcal{I} = \int_{\mathbb{R}}[g'(y)]^2\rho(y)\mathrm{d}y < +\infty, \quad \text{and} \quad \mathcal{J} = \int_{\mathbb{R}} g^2(x)\rho(x)\mathrm{d}x < +\infty. \tag{7.1}$$

We can split $\mathcal{I}$ into $\mathcal{I} = \frac{4}{\sigma^4}(\mathcal{I}_+ + \mathcal{I}_-)$ with

$$\mathcal{I}_- = \int_{-\infty}^{0}\frac{\mathrm{d}y}{\rho(y)}\left(\int_{-\infty}^{y} h_c(u)\rho(u)\mathrm{d}u\right)^2, \quad \mathcal{I}_+ = \int_{0}^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int_{y}^{+\infty} h_c(u)\rho(u)\mathrm{d}u\right)^2.$$

It is enough to look at the first integral, the other one being analogous. We have, using $h_c = h - \int h\rho$,

$$\mathcal{I}_- \leq 2\int_{-\infty}^{0}\frac{\mathrm{d}y}{\rho(y)}\left(\int_{-\infty}^{y} h(u)\rho(u)\mathrm{d}u\right)^2 + 2\int_{-\infty}^{0}\frac{\mathrm{d}y}{\rho(y)}\left(\int_{-\infty}^{y}\rho(u)\mathrm{d}u\right)^2\left(\int h(u)\rho(u)\mathrm{d}u\right)^2.$$

These integrals are finite by Lemma 7.1 (with $h(u) = 1$ for the second one).

Now, we also split $\mathcal{J} = \mathcal{J}_+ + \mathcal{J}_-$, with

$$\mathcal{J}_+ = \int_0^{+\infty}[g(x)]^2\rho(x)\mathrm{d}x, \quad J_- = \int_{-\infty}^{0}[g(x)]^2\rho(x)\mathrm{d}x,$$

and only treat $\mathcal{J}_-$. We have

$$\begin{aligned}\mathcal{J}_- &\leq \int_{-\infty}^{0}|x|\rho(x)\int_{x}^{0}[g'(y)]^2 dy\mathrm{d}x = \int_{x<y<0}|x|[g'(y)]^2\rho(x)dxdy \\ &\leq \frac{8}{\sigma^4}\int_{x<y<0}|x|\frac{\mathrm{d}y}{\rho^2(y)}\left(\int_{-\infty}^{y} h(u)\rho(u)\mathrm{d}u\right)^2\rho(x)dxdy \\ &+ \frac{8}{\sigma^4}\int_{x<y<0}|x|\frac{\mathrm{d}y}{\rho^2(y)}\left(\int_{-\infty}^{y}\rho(u)\mathrm{d}u\right)^2\left(\int h(z)\rho(z)dz\right)^2 \\ &:= \mathcal{J}_-^{(1)} + \mathcal{J}_-^{(2)} \times \left(\int h(z)\rho(z)\mathrm{d}z\right)^2\end{aligned}$$

These integrals are finite by Lemma 7.1.

Thus, (7.1) is satisfied and $\Gamma_h = g$ belongs to $\mathcal{D}$ and satisfies $L\Gamma_h = h_c$. Then (2.10) is obtained thanks to the Ito formula, and (2.11) is deduced using that $X(t)$ is stationary. □

7.5. Proof of Proposition 3.1

Replacing in (3.9) φ_j by h_j and θ_ℓ by $h_{2\ell-1}$, and applying Fourier transform yields

$$\forall x \in \mathbb{R},\ \sum_{j=0}^{m-1} y_j i^j h_j(x) + \sum_{\ell=0}^{p-1} y_{\ell+m} i^{2\ell+1} h_{2\ell+1}(x)\rho^*(x) = 0. \tag{7.2}$$

Multiplying by $\mathrm{e}^{x^2/2}$ gives $P(x) + Q(x)\rho^*(x) = 0$, $\forall x \in \mathbb{R}$, where P and Q are polynomials.

Let us moreover assume that B and Φ are $\mathcal{C}^\infty$ and let $f^{(n)}(x)$ be the n-th derivative of f. Inequality (2.3) yields that ρ belongs to the Schwarz class $\mathcal{S}(\mathbb{R}) = \{f \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R}),\ \text{s.t.} \forall m \geq 1, \forall n \geq 0, f^{(n)}(x) = O(x^{-m}) \text{ at } \infty\}$. Therefore $\rho^\star$ also belongs to $\mathcal{S}(\mathbb{R})$. Letting $x \to \infty$ in (7.2) yields that $P(x) \to 0$, which implies that the term with highest degree of $P(x)$, *i.e.* of $y_{m-1} i^{m-1} H_{m-1}(x)$, tends to 0. Therefore $y_{m-1} = 0$. Iterating the procedure yields successively that $y_{m-1}, y_{m-2}, \ldots, y_0$ are all equal to 0. Equation (7.2) reduces to the equation: $\forall x \in \mathbb{R},\ \sum_{\ell=0}^{p-1} y_{\ell+m} i^{2\ell+1} h_{2\ell+1}(x) = 0$. The functions h_j being orthogonal, this implies that $y_{\ell+m} = 0$ for $\ell = 0, \ldots, p-1$. The result is thus proved. □

7.6. Proof of Proposition 3.2

Obviously, $\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi_k'(X^{(3)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \varphi_k'(x)\rho(x)\mathrm{d}x$. For the term $\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s$, the expectation is computed in two steps by using the independence of $X^{(3)}$ and $X^{(4)}$.

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \mathbb{E}(\overline{\overline{\theta'}}_\ell(x))\rho(x)\mathrm{d}x = \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x.$$

Gathering both terms yields that $\mathbb{E}(\widehat{\mathbf{Z}}_{m,p}) = \mathbf{Z}_{m,p}$.

Concerning $\widehat{\mathbf{V}}_{m,p}$, it is clear that

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi_j(X^{(1)}(s))\varphi_k(X^{(1)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \varphi_j(x)\varphi_k(x)\rho(x)\mathrm{d}x$$

and with the same two-step computation as previously, we have

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \bar{\theta}_\ell(X^{(1)}(s))\,\varphi_k(X^{(1)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \mathbb{E}(\bar{\theta}_\ell(x))\varphi_k(x)\rho(x)\mathrm{d}x = \int_{\mathbb{R}} \theta_\ell \star \rho(x)\varphi_k(x)\mathrm{d}x.$$

Now we prove that

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \bar{\theta}_\ell(X^{(1)}(s))\,\bar{\theta}_r(X^{(1)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \theta_\ell \star \rho(x)\theta_r \star \rho(x)\rho(x)\mathrm{d}x + O(\frac{1}{T}).$$

First, we split:

$$\frac{1}{T}\int_0^T \bar{\theta}_\ell(X^{(1)}(s))\,\bar{\theta}_r(X^{(1)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta_\ell \star \rho(x)\theta_r \star \rho(x)\rho(x)\mathrm{d}x = A_1(\ell,r) + A_2(\ell,r) + A_3(\ell,r) + A_4(\ell,r),$$

where

$$A_1(\ell,r) = T^{-1}\int_0^T (\bar\theta_\ell(X^{(1)}(s)) - \theta_\ell\star\rho(X^{(1)}(s)))(\bar\theta_r(X^{(1)}(s)) - \theta_r\star\rho(X^{(1)}(s)))\mathrm{d}s.$$
$$A_2(\ell,r) = T^{-1}\int_0^T (\bar\theta_\ell(X^{(1)}(s)) - \theta_\ell\star\rho(X^{(1)}(s)))\theta_r\star\rho(X^{(1)}(s))\mathrm{d}s, \quad A_3(\ell,r) = A_2(r,\ell),$$
$$A_4(\ell,r) = T^{-1}\int_0^T \theta_\ell\star\rho(X^{(1)}(s))\theta_r\star\rho(X^{(1)}(s))\mathrm{d}s - \int_{\mathbb{R}}\theta_\ell\star\rho(x)\theta_r\star\rho(x)\rho(x)\mathrm{d}x,$$

The term $A_4(\ell,r)$ is centered as $\mathbb{E}(\theta_\ell\star\rho(X^{(1)}(s))\theta_r\star\rho(X^{(1)}(s)) = \int_{\mathbb{R}}\theta_\ell\star\rho(x)\theta_r\star\rho(x)\rho(x)\mathrm{d}x$. By the two-step computation already used, the terms $A_2(\ell,r), A_3(r,\ell)$ are centered too by the independence of $X^{(1)}$ and $X^{(2)}$, as

$$\mathbb{E}^{X^{(1)}(s)}(\bar\theta_\ell(X^{(1)}(s)) - \theta_\ell\star\rho(X^{(1)}(s))) = 0.$$

For the term $A_1(\ell,r)$, we have:

$$\mathbb{E}A_1(\ell,r) = \int_{\mathbb{R}}\rho(u)\mathrm{d}u\mathbb{E}\left[(\bar\theta_\ell(u) - \theta_\ell\star\rho(u))(\bar\theta_r(u) - \theta_r\star\rho(u))\right];$$

We intend to use Proposition 2.4. Set $\theta_{\ell,u}(x) = \theta_\ell(u-x)$ and note that $\theta^c_{\ell,u}(x) = \theta_{\ell,u}(x) - \int_{\mathbb{R}}\theta_{\ell,u}(x)\rho(x)\mathrm{d}x = \theta_\ell(u-x) - \theta_\ell\star\rho(u)$. Now with $\Gamma_{\theta_{\ell,u}}$ such that $L\Gamma_{\theta_{\ell,u}} = \theta^c_{\ell,u}$, it holds that

$$\begin{aligned}\frac{1}{T}\int_0^T \theta^c_{\ell,u}(X^{(2)}(t))\mathrm{d}t &= \bar\theta_\ell(u) - \theta_\ell\star\rho(u)\\ &= \frac{\sigma}{T}\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s) + \frac{1}{T}\left(\Gamma_{\theta_{\ell,u}}(X^{(2)}(0)) - \Gamma_{\theta_{\ell,u}}(X^{(2)}(T))\right).\end{aligned} \tag{7.3}$$

Consequently,

$$\mathbb{E}\left[(\bar\theta_\ell(u) - \theta_\ell\star\rho(u))(\bar\theta_r(u) - \theta_r\star\rho(u))\right] = T^{(1)}_{\ell,r}(u) + T^{(2)}_{\ell,r}(u) + T^{(3)}_{\ell,r}(u) + T^{(4)}_{\ell,r}(u)$$

with

$$T^{(1)}_{\ell,r}(u) = \frac{\sigma^2}{T^2}\mathbb{E}\left(\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))\Gamma'_{\theta_{r,u}}(X^{(2)}(s))\mathrm{d}s\right) = \frac{\sigma^2}{T}\int \Gamma'_{\theta_{\ell,u}}(y)\Gamma'_{\theta_{r,u}}(y)\rho(y)\mathrm{d}y$$
$$T^{(2)}_{\ell,r}(u) = \frac{1}{T^2}\mathbb{E}\left[\left(\Gamma_{\theta_{\ell,u}}(X^{(2)}(0)) - \Gamma_{\theta_{\ell,u}}(X^{(2)}(T))\right)\left(\Gamma_{\theta_{r,u}}(X^{(2)}(0)) - \Gamma_{\theta_{r,u}}(X^{(2)}(T))\right)\right]$$
$$T^{(3)}_{\ell,r}(u) = \frac{\sigma}{T^2}\mathbb{E}\left[\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s)\left(\Gamma_{\theta_{r,u}}(X^{(2)}(0)) - \Gamma_{\theta_{r,u}}(X^{(2)}(T))\right)\right]$$
$$T^{(4)}_{\ell,r}(u) = T^{(3)}_{r,\ell}(u)$$

By the Cauchy-Schwarz inequality, we get:

$$T^{(2)}_{\ell,r}(u) \le \frac{2}{T^2}\left(\int_{\mathbb{R}} \Gamma^2_{\theta_{\ell,u}}(y)\rho(y)\mathrm{d}y \int_{\mathbb{R}} \Gamma^2_{\theta_{r,u}}(y)\rho(y)\mathrm{d}y\right)^{1/2},$$

$$T^{(3)}_{\ell,r}(u) \le \frac{2\sigma}{T^{3/2}}\left(\int_{\mathbb{R}} (\Gamma'_{\theta_{\ell,u}}(y))^2\rho(y)\mathrm{d}y \int_{\mathbb{R}} \Gamma^2_{\theta_{r,u}}(y)\rho(y)\mathrm{d}y\right)^{1/2}.$$

Next, we must integrate all these terms w.r.t. $\rho(u)\mathrm{d}u$.
Using the Cauchy-Schwarz inequality for the second and the third inequalities leads to:

$$\int_{\mathbb{R}} \rho(u)T^{(1)}_{\ell,r}(u)\mathrm{d}u = \frac{\sigma^2}{T}\int_{\mathbb{R}^2} \Gamma'_{\theta_{\ell,u}}(y)\Gamma'_{\theta_{r,u}}(y)\rho(y)\rho(u)dudy,$$

$$\int_{\mathbb{R}} \rho(u)T^{(2)}_{\ell,r}(u)\mathrm{d}u \le \frac{\sigma^4}{T^2}\left(\int_{\mathbb{R}^2} \Gamma^2_{\theta_{\ell,u}}(y)\rho(y)\rho(u)dudy \int_{\mathbb{R}^2} \Gamma^2_{\theta_{r,u}}(y)\rho(y)\rho(u)dudy\right)^{1/2}$$

$$\int_{\mathbb{R}} \rho(u)T^{(3)}_{\ell,r}(u)\mathrm{d}u \le \frac{2\sigma}{T^{3/2}}\left(\int_{\mathbb{R}^2} [\Gamma'_{\theta_{\ell,u}}(y)]^2\rho(y)\rho(u)dudy \int_{\mathbb{R}^2} \Gamma^2_{\theta_{r,u}}(y)\rho(y)\rho(u)dudy\right)^{1/2}.$$

Computations analogous to the ones of Lemma 7.1 allow to prove that the integrals above are finite. We only treat the first one. Recall that

$$\Gamma'_{\theta_{\ell,u}}(y) = \frac{2}{\sigma^2\rho(y)}\int_{-\infty}^{y} \theta^c_{\ell,u}(z)\rho(z)\mathrm{d}z = -\frac{2}{\sigma^2\rho(y)}\int_{y}^{+\infty} \theta^c_{\ell,u}(z)\rho(z)\mathrm{d}z.$$

Therefore, we split

$$T^{(1)}_{\ell,r}(u) = \frac{\sigma^2}{T}\left(\int_{y<0}\dots + \int_{y>0}\dots\right).$$

As θ_ℓ is bounded, $\theta^c_{\ell,u}(.)$ is bounded too, so

$$\int_{y<0}\dots \lesssim \int_{y<0}\frac{\mathrm{d}y}{\rho(y)}\int_{z<y,z'<y}\rho(z)\rho(z')dzdz' \propto \int_{y<0}\frac{\mathrm{d}y}{\rho(y)}\left(\int_{z<y}\rho(z)\mathrm{d}z\right)^2$$

which is finite by Lemma 7.1.
Gathering the four terms, we get that

$$\mathbb{E}\left(\frac{1}{T}\int(\bar{\theta}_\ell(X^{(1)}(s)) - \theta_\ell\star\rho(X^{(1)}(s)))(\bar{\theta}_r(X^{(1)}(s)) - \theta_r\star\rho(X^{(1)}(s)))\mathrm{d}s\right)$$
$$= \frac{\sigma^2}{T}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \Gamma'_{\theta_{\ell,u}}(y)\Gamma'_{\theta_{r,u}}(y)\rho(y)\mathrm{d}y\right)\rho(u)\mathrm{d}u + \frac{1}{T^{3/2}}O(1).$$

Finally,

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \bar{\theta}_\ell(X^{(1)}(s))\bar{\theta}_r(X^{(1)}(s))\mathrm{d}s\right) = \int_{\mathbb{R}} \theta_\ell\star\rho(x)\,\theta_r\star\rho(x)\,\rho(x)\mathrm{d}x$$
$$+\frac{\sigma^2}{T}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \Gamma'_{\theta_{\ell,u}}(y)\Gamma'_{\theta_{r,u}}(y)\rho(y)\mathrm{d}y\right)\rho(u)\mathrm{d}u + \frac{1}{T^{3/2}}O(1).$$

This states a more precise result and ends the proof. □

7.7. Proof of Theorem 3.3

We have

$$\begin{aligned}\|\widetilde{(b,\phi)}_{m,p}-(b,\phi)\|_V^2 &= \|\widehat{(b,\phi)}_{m,p}-(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}}+\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c}\\ &= \left(\|\widehat{(b,\phi)}_{m,p}-(b,\phi)_{m,p}\|_V^2+\|(b,\phi)_{m,p}-(b,\phi)\|_V^2\right)\mathbf{1}_{\Delta_{m,p}}+\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c}\end{aligned}$$

where we use Pythagoras theorem: $(b,\phi)_{m,p}-(b,\phi)$ is orthogonal to $S_m\times\Sigma_p$ w.r.t the V-scalar product and thus V-orthogonal to $\widehat{(b,\phi)}_{m,p}-(b,\phi)_{m,p}$. Therefore

$$\|\widetilde{(b,\phi)}_{m,p}-(b,\phi)\|_V^2\le\|\widehat{(b,\phi)}_{m,p}-(b,\phi)_{m,p}\|_V^2\mathbf{1}_{\Delta_{m,p}}+\|(b,\phi)_{m,p}-(b,\phi)\|_V^2+\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c}$$

The middle rhs term is the squared bias term. Let us study the two others. We write, on $\Delta_{m,p}$,

$$\begin{aligned}&\|\widehat{(b,\phi)}_{m,p}-(b,\phi)_{m,p}\|_V^2=\|\widehat{(\mathbf{b},\mathbf{c})}_{m,p}-(\mathbf{b},\mathbf{c})_{m,p}\|_{\mathbf{V}_{m,p}}^2=\|\widehat{\mathbf{V}}_{m,p}^{-1}\widehat{\mathbf{Z}}_{m,p}-\mathbf{V}_{m,p}^{-1}\mathbf{Z}_{m,p}\|_{\mathbf{V}_{m,p}}^2\\ &=\|(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})+\mathbf{V}_{m,p}^{-1}(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})+(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})\mathbf{Z}_{m,p}\|_{\mathbf{V}_{m,p}}^2\\ &\le 3\|(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})\|_{\mathbf{V}_{m,p}}^2+3\|\mathbf{V}_{m,p}^{-1}(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})\|_{\mathbf{V}_{m,p}}^2+3\|(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})\mathbf{Z}_{m,p}\|_{\mathbf{V}_{m,p}}^2\end{aligned}$$

We set

$$\begin{aligned}T_1&:=\|(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})\|_{\mathbf{V}_{m,p}}^2\mathbf{1}_{\Delta_{m,p}}\\ T_2&:=\|\mathbf{V}_{m,p}^{-1}(\widehat{\mathbf{Z}}_{m,p}-\mathbf{Z}_{m,p})\|_{\mathbf{V}_{m,p}}^2\mathbf{1}_{\Delta_{m,p}}\\ T_3&:=\|(\widehat{\mathbf{V}}_{m,p}^{-1}-\mathbf{V}_{m,p}^{-1})\mathbf{Z}_{m,p}\|_{\mathbf{V}_{m,p}}^2\mathbf{1}_{\Delta_{m,p}}\\ T_4&:=\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c}\end{aligned}$$

We have

$$\mathbb{E}(\|\widetilde{(b,\phi)}_{m,p}-(b,\phi)\|_V^2)\le\|(b,\phi)_{m,p}-(b,\phi)\|_V^2+3\mathbb{E}(T_1+T_2+T_3)+\mathbb{E}(T_4).$$

The proof of Theorem 3.3 is structured in a study of the terms, $\mathbb{E}(T_i)$ for $i=1,\dots,4$ and relies on the results stated in Propositions 3.4 , 3.5 and Lemma 3.6.

• Study of $T_4=\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c}$. For this term, we have

Lemma 7.2. *Under [H5], it holds*

$$\mathbb{E}(T_4)=\mathbb{E}(\|(b,\phi)\|_V^2\mathbf{1}_{\Delta_{m,p}^c})\le(4+1/4)s_0^2\,\|(b,\phi)\|_V^2\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}-\mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right).\qquad(7.4)$$

The bound for $\mathbb{E}(T_4)$ follows then from Proposition 3.5,

$$\mathbb{E}(T_4)\lesssim s_0^2\|(b,\phi)\|_V^2\frac{L_\varphi(m)}{T}$$

Proof of Lemma 7.2. Let

$$\Omega_{m,p} = \left\{ \|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\mathrm{op}} < \frac{1}{2} \right\}. \tag{7.5}$$

We write that

$$\mathbb{P}(\Delta_{m,p}^c) = \mathbb{P}(\Delta_{m,p}^c \cap \Omega_{m,p}) + \mathbb{P}(\Delta_{m,p}^c \cap \Omega_{m,p}^c).$$

We have

$$\mathbb{P}(\Omega_{m,p}^c) \le 4\mathbb{E}\left(\|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\mathrm{op}}^2\right) \le 4\|\mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right).$$

Therefore, using [H5],

$$\mathbb{P}(\Delta_{m,p}^c \cap \Omega_{m,p}^c) \le \mathbb{P}(\Omega_{m,p}^c) \le 4s_0^2\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right).$$

Next, on $\Delta_{m,p}^c$, due to [H5], it holds that $\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}} > s_0$ and

$$\mathbb{P}(\Delta_{m,p}^c \cap \Omega_{m,p}) \le \mathbb{E}\left(\mathbf{1}_{\{\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}} > s_0\}}\mathbf{1}_{\Omega_{m,p}}\right) \le \frac{1}{s_0^2}\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Omega_{m,p}}\right).$$

Using (7.17), we get

$$\mathbb{P}(\Delta_{m,p}^c \cap \Omega_{m,p}) \le \frac{s_0^2}{4}\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right).$$

Adding both bounds gives (7.4) and concludes the proof of Lemma 7.2. □

• Study of $T_3 = \|(\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1})\mathbf{Z}_{m,p}\|_{\mathbf{V}_{m,p}}^2\mathbf{1}_{\Delta_{m,p}}$.
Using (3.15) and Lemma 3.6 yields

$$\begin{aligned}\mathbb{E}(T_3) &\le 2\|\rho\|_\infty\mathbb{E}\left(\|(\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1})\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\right)\|\mathbf{Z}_{m,p}\|_{\mathbb{R}^{m,p}}^2 \\ &\le 80.5\, s_0^4\, \|\rho\|_\infty\ \mathbb{E}(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2)\|\mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2.\end{aligned} \tag{7.6}$$

Now, we can prove the following result:

Lemma 7.3. *We have*

$$\|\mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2 \le 2\|\rho\|_\infty\|b + \phi\star\rho\|_\rho^2 = 2\|\rho\|_\infty\|(b,\phi)\|_V^2.$$

Thus, by (7.6), Proposition 3.5 and Lemma 7.3, we obtain,

$$\mathbb{E}(T_3) \lesssim s_0^4\frac{L_\varphi(m)}{T}.$$

Proof of Lemma 7.3. First, we note that, using (3.2) we get

$$\begin{aligned}\|\mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2 = &\sum_{k=0}^{m-1}\left(\int_{\mathbb{R}}\varphi_k(x)[b(x) + \phi\star\rho(x)]\rho(x)\mathrm{d}x\right)^2 \\ &+ \sum_{\ell=0}^{p-1}\left(\int_{\mathbb{R}}\theta_\ell\star\rho(x)[b(x) + \phi\star\rho(x)]\rho(x)\mathrm{d}x\right)^2.\end{aligned}$$

Using again the relation $\int (u \star v)\, w = \int u\,(v_- \star w)$ with $v_-(x) = v(-x)$ yields

$$\int_{\mathbb{R}} \theta_\ell \star \rho(x)[(b(x) + \phi \star \rho(x)]\rho(x)\mathrm{d}x = \int_{\mathbb{R}} \theta_\ell(x)(b + \phi \star \rho)\rho \star \rho_-)(x)\mathrm{d}x.$$

Therefore, by the projection argument and the Young inequality ($\|u \star v\| \leq \|u\|_1 \|v\|$) as$\|\rho\|_1 = 1$, we get

$$\begin{aligned}\|\mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}} &\leq \int_{\mathbb{R}} (b(x) + \phi \star \rho(x))^2 \rho^2(x)\mathrm{d}x + \int_{\mathbb{R}} [(b(x) + \phi \star \rho(x))\rho \star \rho_-(x)]^2 dx \\ &\leq \int_{\mathbb{R}} (b(x) + \phi \star \rho(x))^2 \rho^2(x)\mathrm{d}x + \|\rho_-\|_1^2 \int_{\mathbb{R}} (b(x) + \phi \star \rho(x))^2 \rho^2(x)\mathrm{d}x \\ &\leq 2\|\rho\|_\infty \int_{\mathbb{R}} (b(x) + \phi \star \rho(x))^2 \rho(x)\mathrm{d}x = 2\|\rho\|_\infty \|(b, \phi)\|_V^2. \quad \square\end{aligned}$$

- Study of $\mathbb{E}(T_2)$. Using [H5] and (3.15) yields

$$\begin{aligned}\mathbb{E}(T_2) = \mathbb{E}\left(\|\mathbf{V}^{-1}_{m,p}(\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p})\mathbf{1}_{\Delta_{m,p}}\|^2_{\mathbf{V}_{m,p}}\right) &\leq \|\mathbf{V}^{-1}_{m,p}\|_{\mathrm{op}} \mathbb{E}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}} \mathbf{1}_{\Delta_{m,p}}\right) \\ &\leq s_0\, \mathbb{E}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}}\right).\end{aligned}$$

Applying Proposition 3.4 yields that

$$\mathbb{E}(T_2) \lesssim s_0 \frac{L_\varphi(m) + L_\theta(p)}{T}. \tag{7.7}$$

- Study of $T_1 = \|(\widehat{\mathbf{V}}^{-1}_{m,p} - \mathbf{V}^{-1}_{m,p})(\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p})\|^2_{\mathbf{V}_{m,p}} \mathbf{1}_{\Delta_{m,p}}$. By (3.15),

$$T_1 \leq 2\|\rho\|_\infty (\|\widehat{\mathbf{V}}^{-1}_{m,p} - \mathbf{V}^{-1}_{m,p}\|^2_{\mathrm{op}} \mathbf{1}_{\Delta_{m,p}})\, \|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}}.$$

As the two factors of the right-hand-side are independent

$$\mathbb{E}(T_1) \leq 2\|\rho\|_\infty \mathbb{E}\left(\|\widehat{\mathbf{V}}^{-1}_{m,p} - \mathbf{V}^{-1}_{m,p}\|^2_{\mathrm{op}} \mathbf{1}_{\Delta_{m,p}}\right) \mathbb{E}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}}\right).$$

Using Lemma 3.6 we get

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}^{-1}_{m,p} - \mathbf{V}^{-1}_{m,p}\|^2_{\mathrm{op}} \mathbf{1}_{\Delta_{m,p}}\right) \leq (40 + 1/4) s_0^4 \mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|^2_{\mathrm{op}} \mathbf{1}_{\Delta_{m,p}}\right),$$

$$\mathbb{E}(T_1) \leq 2(40 + 1/4) s_0^4 \|\rho\|_\infty \mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|^2_{\mathrm{op}} \mathbf{1}_{\Delta_{m,p}}\right) \mathbb{E}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|^2_{\mathbb{R}^{m+p}}\right).$$

By applying Proposition 3.4 and 3.5 and using $L_\varphi(m) \leq T$, we obtain

$$\mathbb{E}(T_1) \lesssim s_0^4 \frac{L_\varphi(m) + L_\theta(p)}{T}.$$

Gathering the five bounds ends the proof of Theorem 3.3.$\square$

Remark 7.4. The proof of Theorem 1 consists in bounding four terms. It is only in the first one

$$T_1 := \|(\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1})(\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p})\|_{\mathbf{V}_{m,p}}^2 \mathbf{1}_{\Delta_{m,p}}$$

that we use the fact that we have four trajectoires. Indeed, T_1 is bounded by

$$T_1 \leq 2\|\rho\|_\infty (\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2 \mathbf{1}_{\Delta_{m,p}})\, \|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2,$$

where $\mathbf{V}_{m,p}$ is defined in (20) and for $x \in \mathbb{R}^m, y \in \mathbb{R}^p$, $||(x,y)||_{\mathbf{V}_{m,p}}^2 = (x,y)^{\perp}\mathbf{V}_{m,p}(x,y)$.
Because of the use of four trajectories, the two terms of the rhs are independent and we have

$$\mathbb{E}(T_1) \leq 2\|\rho\|_\infty \mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2 \mathbf{1}_{\Delta_{m,p}}\right) \mathbb{E}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2\right).$$

If we only had two trajectories, then, the two terms would not be independent and we would have to use the Cauchy-Schwarz inequality to separate the expectations. Therefore, we would have to study:

$$\mathbb{E}^{1/2}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^4 \mathbf{1}_{\Delta_{m,p}}\right), \quad \mathbb{E}^{1/2}\left(\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^4\right).$$

This would lengthen proofs a lot.

7.8. Proof of Proposition 3.4

We have

$$\mathbb{E}[\|\widehat{\mathbf{Z}}_{m,p} - \mathbf{Z}_{m,p}\|_{\mathbb{R}^{m+p}}^2] = \frac{\sigma^4}{4}\mathbb{E}\left\{\sum_{k=0}^{m-1}\left(\frac{1}{T}\int_0^T \varphi_k'(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi_k'(x)\rho(x)\mathrm{d}x\right)^2 \right.$$
$$\left. + \sum_{\ell=0}^{p-1}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_k(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta_k' \star \rho(x)\,\rho(x)\mathrm{d}x\right)^2\right\}. \tag{7.8}$$

We prove two lemmas for each term of the r.h.s.

Lemma 7.5. *Assume that φ_k is a differentiable and bounded basis. Let $g_k := \Gamma_{\varphi_k'}$ be defined by $Lg_k = [\varphi_k' - \int_{\mathbb{R}} \varphi_k'(x)\rho(x)\mathrm{d}x]$ (see Prop. 2.4). Then,*

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi_k'(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi_k'(x)\rho(x)\mathrm{d}x\right)^2 \leq \frac{2\sigma^2}{T}\int_{\mathbb{R}}[g_k'(x)]^2\rho(x)\mathrm{d}x + \frac{8}{T^2}\int_{\mathbb{R}} g_k^2(x)\rho(x)\mathrm{d}x. \tag{7.9}$$

Moreover,

$$\sum_{k=0}^{m-1}\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi_k'(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi_k'(x)\rho(x)\mathrm{d}x\right)^2 \lesssim \frac{L_\varphi(m)}{T}.$$

Lemma 7.6. *Recall* $\overline{\overline{\theta'}}_\ell(x) = \frac{1}{T}\int_0^T \theta'_\ell(x - X^{(4)}(t))\mathrm{d}t$. *Let* $G_{x,\ell}$ *be defined by* $LG_{x,\ell}(y) = \theta'_\ell(x-y) - \theta'_\ell \star \rho(x)$ *and* h_ℓ *be defined by* $Lh_\ell = \theta'_\ell \star \rho - \int_{\mathbb{R}} \theta'_\ell \star \rho(x)\rho(x)\mathrm{d}x$. *Then,*

$$\begin{aligned}\mathbb{E}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta'_\ell \star \rho(x)\rho(x)\mathrm{d}x\right)^2 \leq{}& \frac{4\sigma^2}{T}\int_{\mathbb{R}^2}\left(G'_{x,k}(y)\right)^2\rho(x)\rho(y)dxdy \\ &+\frac{8}{T^2}\int_{\mathbb{R}^2}\left(G_{x,\ell}(y)\right)^2\rho(x)\rho(y)dxdy \\ &+\frac{4\sigma^2}{T}\int_{\mathbb{R}}[h'_\ell(x)]^2\rho(x)\mathrm{d}x + \frac{8}{T^2}\int_{\mathbb{R}}[h_\ell(x)]^2\rho(x)\mathrm{d}x.\end{aligned}$$

Moreover,

$$\sum_{\ell=0}^{p-1}\mathbb{E}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta'_\ell \star \rho(x)\rho(x)\mathrm{d}x\right)^2 \lesssim \frac{L_\theta(p)}{T}.$$

From the two Lemmas and formula (7.8), we get the result of Proposition 3.4. □

Proof of Lemma 7.5. By definition of L and $g_k := \Gamma_{\varphi'_k}$ and Proposition 2.4, we have

$$\begin{aligned}\frac{1}{T}\int_0^T \varphi'_k(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi'_k(x)\rho(x)\mathrm{d}x ={}& \frac{\sigma}{T}\int_0^T g'_k(X^{(3)}(s))dW_3(s) \\ &+\frac{1}{T}[g_k(X^{(3)}(0)) - g_k(X^{(3)}(T))],\end{aligned}$$

and consequently by (2.11),

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi'_k(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi'_k(x)\rho(x)\mathrm{d}x\right)^2 \leq \frac{2\sigma^2}{T}\int_{\mathbb{R}}[g'_k(x)]^2\rho(x)\mathrm{d}x + \frac{8}{T^2}\int_{\mathbb{R}} g_k^2(x)\rho(x)\mathrm{d}x.$$

This yields inequality (7.9).
Now we have

$$\begin{aligned}g'_k(x) &= \frac{2}{\sigma^2\rho(x)}\int_{-\infty}^x\left(\varphi'_k(y) - \int_{\mathbb{R}}\varphi'_k(z)\rho(z)\mathrm{d}z\right)\rho(y)\mathrm{d}y \\ &= -\frac{2}{\sigma^2\rho(x)}\int_x^{+\infty}\left(\varphi'_k(y) - \int_{\mathbb{R}}\varphi'_k(z)\rho(z)\mathrm{d}z\right)\rho(y)\mathrm{d}y.\end{aligned}$$

We split the term

$$\int_{\mathbb{R}}[g'_k(x)]^2\rho(x)\mathrm{d}x = \int_0^{+\infty}[g'_k(x)]^2\rho(x)\mathrm{d}x + \int_{-\infty}^0[g'_k(x)]^2\rho(x)\mathrm{d}x := \frac{4}{\sigma^4}(I_+ + I_-)$$

with

$$I_+ = \int_0^{+\infty} \frac{\mathrm{d}x}{\rho(x)} \left[\int_x^{+\infty} \left(\varphi_k'(y) - \int \varphi_k'(z)\rho(z)\mathrm{d}z\right)\rho(y)\mathrm{d}y\right]^2$$
$$\leq 2\int_0^{+\infty} \frac{\mathrm{d}x}{\rho(x)} \left(\int_x^{+\infty} \varphi_k'(y)\rho(y)\mathrm{d}y\right)^2 + 2\int_0^{+\infty} \frac{\mathrm{d}x}{\rho(x)} \left(\int_x^{+\infty} \rho(y)\mathrm{d}y \int \varphi_k'(z)\rho(z)\mathrm{d}z\right)^2$$

Now we write by integration by part,

$$\int_x^{+\infty} \varphi_k'(y)\rho(y)\mathrm{d}y = -\varphi_k(x)\rho(x) - \int_x^{+\infty} \varphi_k(y)\rho'(y)\mathrm{d}y$$

as φ_j is bounded and $\lim_{x\to\pm\infty}\rho(x) = 0$.

As a consequence for all $x > 0$, it holds

$$\sum_{k=0}^{m-1} \left(\int_0^{+\infty} \varphi_k'(y)\rho(y)\mathbf{1}_{[x,+\infty]}(y)\mathrm{d}y\right)^2 \leq 2\sum_{k=1}^{m-1}\left\{\varphi_k^2(x)\rho^2(x) + \left(\int_0^{+\infty}\varphi_k(y)\rho'(y)\mathbf{1}_{[x,+\infty]}(y)\mathrm{d}y\right)^2\right\}$$
$$\leq 2L_\varphi(m)\rho^2(x) + 2\int_0^{+\infty} [\rho'(y)]^2\mathbf{1}_{[x,+\infty]}(y)\mathrm{d}y$$

by using the usual projection argument. Thus

$$\sum_{k=0}^{m-1}\int_0^{+\infty} \frac{\mathrm{d}x}{\rho(x)} \left(\int_x^{+\infty} \varphi_k'(y)\rho(y)\mathrm{d}y\right)^2 \leq 2L_\varphi(m) + 2\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\int_x^{+\infty}[\rho'(y)]^2\mathrm{d}y$$

where, thanks to some adaptation of. Lemma 7.1, noting that $\rho'(y) = \rho(y)\times k(y)$ where $k(y)$ has polynomial growth, we can prove

$$\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\int_x^{+\infty}[\rho'(y)]^2 dy = \int_0^{+\infty}[\rho'(y)]^2\left(\int_0^y \frac{\mathrm{d}x}{\rho(x)}\right)\mathrm{d}y < +\infty.$$

The second term is easier and analogous.

The term I_- is treated analogously using the second formula for g_k'. Therefore,

$$\frac{\sigma^2}{T}\sum_{k=0}^{m-1}\int_{\mathbb{R}}[g_k'(x)]^2\rho(x)\mathrm{d}x \lesssim \frac{L_\varphi(m)}{T}.$$

Now, with $g_k(x) = \int_0^x g_k'(u)\mathrm{d}u$, we look at $\int_{\mathbb{R}} g_k^2(x)\rho(x)\mathrm{d}x$ and split it again into $\int_0^{+\infty}\ldots + \int_{-\infty}^0\ldots$. We use that for $x \geq 0$,

$$[g_k(x)]^2 = \left(\int_0^x g_k'(u)\mathrm{d}u\right)^2 \leq x\int_0^x [g_k'(u)]^2 du.$$

We can write

$$\int_0^{+\infty} g_k^2(x)\rho(x)\mathrm{d}x \leq \frac{2}{\sigma^4}\int_0^{+\infty} x\rho(x)\mathrm{d}x\int_0^x \mathrm{d}u\left(\frac{1}{\rho(u)}\int_u^{+\infty}\varphi_k'(y)\rho(y)\mathrm{d}y\right)^2$$
$$+\frac{8}{\sigma^4}\int_0^{+\infty} x\rho(x)\mathrm{d}x\int_0^x \mathrm{d}u\left(\frac{1}{\rho(u)}\int_u^{+\infty}\rho(y)\mathrm{d}y\int\varphi_k'(z)\rho(z)\mathrm{d}z\right)^2.$$

Then, with the same integration by part for $\int \varphi_k' \rho$ as previously, we get

$$\sum_{k=0}^{m-1} \int_0^{+\infty} g_k^2(x)\rho(x)\mathrm{d}x \leq \frac{16}{\sigma^4}\int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \mathrm{d}u \left[L_\varphi(m) + \frac{1}{\rho^2(u)}\int_u^{+\infty}[\rho'(y)]^2\mathrm{d}y\right]$$
$$+ \frac{16}{\sigma^4}(\int_{\mathbb{R}}[\rho'(z)]^2 dz) \int_0^{+\infty} x\rho(x)\mathrm{d}x \int_0^x \mathrm{d}u \frac{1}{\rho^2(u)}\left(\int_u^{+\infty}\rho(y)\mathrm{d}y\right)^2.$$

Therefore, by Lemma 7.1, all integrals being finite, we obtain the result. □

Proof of Lemma 7.6. We have

$$\overline{\overline{\theta'}}_\ell(X^{(3)}(s)) - \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x = \overline{\overline{\theta'}}_\ell(X^{(3)}(s)) - \theta_\ell' \star \rho(X^{(3)}(s)) + \theta_\ell' \star \rho(X^{(3)}(s)) - \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x.$$

Then, for all x,

$$\overline{\overline{\theta'}}_\ell(x) - \theta_\ell' \star \rho(x) = \frac{1}{T}\int_0^T \theta_\ell'(x - X^{(4)}(t))\mathrm{d}t - \theta_\ell' \star \rho(x)$$
$$= \frac{\sigma}{T}\int_0^T G_{x,\ell}'(X^{(4)}(t))dW_4(t) + \frac{1}{T}[G_{x,\ell}'(X^{(4)}(0)) - G_{x,\ell}'(X^{(4)}(T))].$$

And,

$$\frac{1}{T}\int_0^T \theta_\ell' \star \rho(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x = \frac{\sigma}{T}\int_0^T h_\ell'(X^{(3)}(s))dW_3(s)$$
$$+\frac{1}{T}[h_\ell(X^{(3)}(0)) - h_\ell(X^{(3)}(T))].$$

Thus,

$$\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x = \frac{1}{T}\int_0^T \mathrm{d}s\left(\frac{\sigma}{T}\int_0^T G_{X^{(3)}(s)),\ell}'(X^{(4)}(t))dW_3(t)\right)$$
$$+\frac{1}{T}\int_0^T \mathrm{d}s\left(\frac{1}{T}[G_{X^{(3)}(s)),\ell}'(X^{(4)}(0)) - G_{X^{(3)}(s)),\ell}'(X^{(4)}(T))]\right)$$
$$+\frac{\sigma}{T}\int_0^T h_\ell'(X^{(3)}(s))dW_3(s) + \frac{1}{T}[h_\ell(X^{(3)}(0)) - h_\ell(X^{(3)}(T))]$$

It follows that

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta_\ell' \star \rho(x)\rho(x)\mathrm{d}x\right)^2 \lesssim \frac{1}{T}\int_0^T \mathrm{d}s\left(\frac{\sigma^2}{T^2}\mathbb{E}\mathbb{E}^{X^{(3)}(s)}\int_0^T \mathrm{d}t[G_{X^{(3)}(s)),\ell}'(X^{(4)}(t))]^2\right)$$
$$+\frac{1}{T}\int_0^T \mathrm{d}s\frac{1}{T^2}\mathbb{E}\mathbb{E}^{X^{(3)}(s)}[G_{X^{(3)}(s)),\ell}'(X^{(4)}(0)) - G_{X^{(3)}(s)),\ell}'(X^{(4)}(T))]^2$$
$$+\frac{\sigma^2}{T^2}\int_0^T \mathbb{E}[h_\ell'(X^{(3)}(s))]^2 ds + \frac{1}{T^2}\mathbb{E}[h_\ell(X^{(3)}(0)) - h_\ell(X^{(3)}(T))]^2.$$

This yields

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \overline{\overline{\theta'}}_\ell(X^{(3)}(s))\mathrm{d}s - \int_{\mathbb{R}} \theta'_\ell \star \rho(x)\rho(x)\mathrm{d}x\right)^2 \lesssim \frac{1}{T}\int_{\mathbb{R}^2}[G'_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy$$
$$+\frac{1}{T^2}\int_{\mathbb{R}^2}[G_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy + \frac{1}{T}\int_{\mathbb{R}}[h'_\ell(x)]^2\rho(x)\mathrm{d}x + \frac{1}{T^2}\int_{\mathbb{R}}[h_\ell(x)]^2\rho(x)\mathrm{d}x.$$

Next, we have to study $\sum_{\ell=0}^{p-1}$ for each of the above terms following the lines of Lemma 7.5. By formula (2.9), we have for $y \geq 0$,

$$G'_{x,\ell}(y) = \frac{2}{\sigma^2\rho(y)}\int_y^{+\infty}\left(\theta'_\ell(x-z) - \theta'_\ell \star \rho(x)\right)\rho(z)\mathrm{d}z.$$

By integration by part, we obtain, for $y \geq 0$

$$G'_{x,\ell}(y) = \frac{2}{\sigma^2\rho(y)}\left\{\theta_\ell(x-y)\rho(y) + \int_y^{+\infty}\theta_\ell(x-z)\rho'(z)\mathrm{d}z - \theta_\ell \star \rho'(x)\int_y^{+\infty}\rho(z)\mathrm{d}z\right\}.$$

Thus

$$\sum_{\ell=0}^{p-1}[G'_{x,\ell}(y)]^2 = \frac{4}{\sigma^4}\sum_{\ell=0}^{p-1}\left\{\theta_\ell(x-y) + \frac{1}{\rho(y)}\int_y^{+\infty}\theta_\ell(x-z)\rho'(z)\mathrm{d}z - \frac{\theta_\ell\star\rho'(x)}{\rho(y)}\int_y^{+\infty}\rho(z)\mathrm{d}z\right\}^2$$
$$\leq \frac{12}{\sigma^4}\left(L_\theta(p) + \frac{1}{\rho^2(y)}\int_y^{+\infty}[\rho'(z)]^2dz + \int(\rho'(u))^2\mathrm{d}u\frac{\int_y^{+\infty}\rho(z)\mathrm{d}z}{\rho^2(y)}\right)$$

and

$$\sum_{\ell=0}^{p-1}\int\int_0^{+\infty}[G'_{x,\ell}(y)]^2\rho(y)\rho(x)dxdy \leq \frac{12}{\sigma^4}\left(L_\theta(p) + \int_0^{+\infty}\frac{1}{\rho(y)}\int_y^{+\infty}[\rho'(z)]^2dz\mathrm{d}y\right.$$
$$\left. + \int(\rho'(u))^2\mathrm{d}u\int_0^{+\infty}(\rho'(u))^2\mathrm{d}w\int_y^{+\infty}\rho(z)\mathrm{d}z\frac{\mathrm{d}y}{\rho(y)}\right).$$

Following the proof of Lemma 7.1, we prove that all integrals are finite. The same holds for the case $y \leq 0$. The order of the term is $L_\theta(p)$.

The terms $\int_{\mathbb{R}^2}[G_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy$ are treated in a similar way with an additional integration due to the bound $G^2_{x,\ell}(y) \leq |x|\int_{[0,x]}[G'_{x,\ell}(u)]^2du$. This yields the same order $L_\varphi(p)$ for the sum over ℓ.

Now, for the terms in h_ℓ, we note that

$$\theta'_\ell \star \rho - \int_{\mathbb{R}}\theta'_\ell \star \rho(x)\rho(x)\mathrm{d}x = \theta_\ell \star \rho' - \int_{\mathbb{R}}\theta_\ell \star \rho'(x)\rho(x)\mathrm{d}x.$$

By Formula (2.9), we get, for $x \geq 0$,

$$h'_\ell(x) = \frac{2}{\sigma^2\rho(x)}\int_x^{+\infty}\left(\theta_\ell \star \rho'(u) - \int_{\mathbb{R}}\theta_\ell \star \rho'(z)\rho(z)\mathrm{d}z\right)\rho(u)\mathrm{d}u$$

so that

$$\sum_{\ell=0}^{p-1}[h_\ell'(x)]^2 \leq \frac{4}{\sigma^4\rho^2(x)}\left(\int_x^{+\infty}\left(\theta_\ell\star\rho'(u)-\int_{\mathbb{R}}\theta_\ell\star\rho'(z)\rho(z)\mathrm{d}z\right)\rho(u)\mathrm{d}u\right)^2.$$

Now, note that for $x\geq 0$

$$\int_x^{+\infty}\theta_\ell\star\rho'(u)\rho(u)\mathrm{d}u=\int_{\mathbb{R}}\theta_\ell(v)\left(\int_{\mathbb{R}}\mathbf{1}_{u\geq x}\rho'(u-v)\rho(u)\mathrm{d}u\right)\mathrm{d}v$$

which yields

$$\sum_{\ell=0}^{p-1}\left(\int_x^{+\infty}\theta_\ell\star\rho'(u)\rho(u)\mathrm{d}u\right)^2\leq\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\mathbf{1}_{u\geq x}\rho'(u-v)\rho(u)\mathrm{d}u\right)^2\mathrm{d}v.$$

This implies that

$$\begin{aligned}\sum_{k=0}^{p-1}\int_0^{\infty}[h_\ell'(x)]^2\rho(x)\mathrm{d}x\leq\frac{8}{\sigma^4}\Bigg\{&\int_0^{+\infty}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\mathbf{1}_{u\geq x}\rho'(u-v)\rho(u)\mathrm{d}u\right)^2\mathrm{d}v\frac{\mathrm{d}x}{\rho(x)}\\&+\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\rho'(u-v)\rho(u)\mathrm{d}u\right)^2\mathrm{d}v\int_0^{+\infty}\left(\int_x^{+\infty}\rho(u)\mathrm{d}u\right)^2\frac{\mathrm{d}x}{\rho(x)}\Bigg\}.\end{aligned}$$

As a consequence, this term is bounded by $C(\rho)/\sigma^4$ where $C(\rho)$ is a constant depending on ρ only.

We proceed as done several times previously with $\sum_{\ell=0}^{p-1}\int_0^\infty h_k^2(x)\rho(x)\mathrm{d}x$, which yields the same order. □

7.9. Proof of Proposition 3.5

Recall that the matrices $\widehat{V}_{m,p}$ and $V_{m,p}$ are symmetric nonnegative. To obtain Proposition 3.5, we prove that

$$\mathbb{E}\left[\mathrm{Tr}(\widehat{V}_{m,p}-V_{m,p})^2\right]\lesssim\frac{L_\varphi(m)}{T}.$$

Let us recall that for a function h, $\overline{h}(x)=\frac{1}{T}\int_0^T h(x-X^{(2)}(s))\mathrm{d}s$. We have $\mathrm{Tr}(\widehat{V}_{m,p}-V_{m,p})^2=T_{11}+2T_{12}+T_{22}$ where

$$T_{11}=\mathrm{Tr}(\widehat{V}^1-V^1)^2)=\sum_{j,k=0}^{m-1}\left(\frac{1}{T}\int_0^T\varphi_j(X^{(1)}(s))\varphi_k(X^{(1)}(s))\mathrm{d}s-\int_{\mathbb{R}}\varphi_j(x)\varphi_k(x)\rho(x)\right)^2dx$$

$$\begin{aligned}T_{12}&=\mathrm{Tr}(\widehat{V}^{1,2}-V^{1,2})(\widehat{V}_{1,2}-V_{1,2})^\perp\\&=\sum_{k=0}^{m-1}\sum_{\ell=0}^{p-1}\left(\frac{1}{T}\int_0^T\overline{\theta}_\ell(X^{(1)}(s))\,\varphi_k(X^{(1)}(s))\mathrm{d}s-\int_{\mathbb{R}}\theta_\ell\star\rho(x)\varphi_k(x)\rho(x)\right)^2dx\end{aligned}$$

$$T_{22}=\mathrm{Tr}(\widehat{V}^2-V^2)^2)=\sum_{\ell,r=0}^{p-1}\left(\frac{1}{T}\int_0^T\overline{\theta}_\ell(X^{(1)}(s))\,\overline{\theta}_r(X^{(1)}(s))\mathrm{d}s-\int_{\mathbb{R}}\theta_\ell\star\rho(x)\,\theta_r\star\rho(x)\,\rho(x)\right)^2dx,$$

Lemma 7.7. *We have* $\mathbb{E}(T_{11})\lesssim\frac{L_\varphi(m)}{T}+\frac{L_\varphi(m)}{T^2}$

Proof of Lemma 7.7. Let $g_{jk}(x) = \Gamma_{\varphi_j\varphi_k}$ where Γ_h is defined in Proposition 2.4. We have:

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \varphi_j(X^{(1)}(s))\varphi_k(X^{(1)}(s))\mathrm{d}s - \int_{\mathbb{R}} \varphi_j(x)\varphi_k(x)\rho(x)\right)^2 dx \leq \frac{2\sigma^2}{T}\int_{\mathbb{R}}[g'_{jk}(x)]^2\rho(x)\mathrm{d}x + \frac{8}{T^2}\int_{\mathbb{R}} g_{jk}^2(x)\rho(x)\mathrm{d}x.$$

For each integral of the rhs, we write $\int = \int_{-\infty}^0 + \int_0^{+\infty}$ and use the appropriate expression of g_{jk} in each integral. Thus,

$$\frac{4}{\sigma^4}\int_0^{+\infty}[g'_{jk}(x)]^2\rho(x)\mathrm{d}x = \int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\left(\int_x^{+\infty}[\varphi_j(y)\varphi_k(y) - \int_{\mathbb{R}}\varphi_j(z)\varphi_k(z)\rho(z)\mathrm{d}z]\rho(y)\mathrm{d}y\right)^2$$
$$\leq 2\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\left[\left(\int_x^{+\infty}\varphi_j(y)\varphi_k(y)\rho(y)\mathrm{d}y\right)^2 + \left(\int_x^{+\infty}\rho(y)\mathrm{d}y\right)^2\left(\int_{\mathbb{R}}\varphi_j(z)\varphi_k(z)\rho(z)\mathrm{d}z\right)^2\right].$$

Now, we compute:

$$\sum_{j=0}^{m-1}\sum_{k=0}^{m-1}\left(\int_x^{+\infty}\varphi_j(y)\varphi_k(y)\rho(y)\mathrm{d}y\right)^2 \leq \sum_{j=0}^{m-1}\int_x^{+\infty}\varphi_j^2(y)\rho^2(y)\mathrm{d}y$$
$$\leq L_\varphi(m)\int_x^{+\infty}\rho^2(y)\mathrm{d}y.$$

We proceed analogously for the second term. This yields

$$\sum_{0\leq j,k\leq m-1}\int_0^{+\infty}[g'_{jk}(x)]^2\rho(x)\mathrm{d}x \leq c_1 L_\varphi(m),$$

with

$$\frac{\sigma^4}{4}c_1 = 2\times\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\int_x^{+\infty}\rho^2(y)\mathrm{d}y + 2\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\left(\int_x^{+\infty}\rho(y)\mathrm{d}y\right)^2\int\rho^2(x)\mathrm{d}x.$$

The integral on $(-\infty, 0)$ is treated analogously. This yields

$$\sum_{0\leq j,k\leq m-1}\int_{\mathbb{R}}[g'_{jk}(x)]^2\rho(x)\mathrm{d}x \leq C_1 L_\varphi(m).$$

Next,

$$\int_0^{+\infty}\rho(x)g_{jk}^2(x)\mathrm{d}x \leq \int_0^{+\infty}\rho(x)x\int_0^x[g'_{jk}(y)]^2dy = \int_{0<y<x}x\rho(x)[g'_{jk}(y)]^2dxdy.$$

We can compute as above

$$\rho^2(y)\sum_{j=0}^{m-1}\sum_{k=0}^{m-1}[g'_{jk}(y)]^2 \leq 2L_\varphi(m)\int_y^{+\infty}\rho^2(z)\mathrm{d}z + 2\left(\int_y^{+\infty}\rho(z)\mathrm{d}z\right)^2\int_{\mathbb{R}}\rho^2(u)\mathrm{d}u.$$

Thus,

$$\frac{\sigma^4}{4}\sum_{0\le j,k\le m-1}\int_0^{+\infty}\rho(x)g_{jk}^2(x)\mathrm{d}x \le 2L_\varphi(m)\int_{0<y<x,y<z} x\frac{\rho(x)\rho^2(z)}{\rho^2(y)}dxdydz$$
$$+4L_\varphi(m)\int_{0<y<x,y<z<z'} x\frac{\rho(x)\rho(z)\rho(z')}{\rho^2(y)}dxdydzdz'\int\rho^2(u)\mathrm{d}u,$$

where all the integrals are finite. Analogously, $\int_{-\infty}^0\rho(x)g_{jk}^2(x)\mathrm{d}x \lesssim L_\varphi(m)$. This concludes the proof. □

Lemma 7.8. $\mathbb{E}T_{12}\lesssim \frac{L_\varphi(m)}{T}+\frac{L_\varphi(m)}{T^2}+\frac{1}{T}+\frac{1}{T^2}$.

Proof of Lemma 7.8.

$$\frac{1}{T}\int_0^T\overline{\theta}_\ell(X^{(1)}(s))\,\varphi_k(X^{(1)}(s))\mathrm{d}s-\int_{\mathbb{R}}\theta_\ell\star\rho(x)\varphi_k(x)\rho(x)\mathrm{d}x = T_1(k,\ell)+T_2(k,\ell)$$
$$T_1(k,\ell)=\frac{1}{T}\int_0^T[\overline{\theta}_\ell(X^{(1)}(s))-\theta_\ell\star\rho(X^{(1)}(s))]\,\varphi_k(X^{(1)}(s))\mathrm{d}s$$
$$T_2(k,\ell)=\frac{1}{T}\int_0^T\theta_\ell\star\rho(X^{(1)}(s))\,\varphi_k(X^{(1)}(s))\mathrm{d}s-\int_{\mathbb{R}}\theta_\ell\star\rho(x)\varphi_k(x)\rho(x)\mathrm{d}x.$$

The second term is easier. Set $h_{\ell k}=\Gamma_{\varphi_k\theta_\ell\star\rho}$ so that

$$T_2(k,\ell)=\frac{\sigma}{T}\int_0^T h'_{\ell k}(X^{(1)}(s))dW_1(s))+\frac{1}{T}[h_{\ell k}(X^{(1)}(0))-h_{\ell k}(X^{(1)}(T))].$$

Therefore,

$$\mathbb{E}[T_2(k,\ell)]^2\lesssim\frac{1}{T}\int_{\mathbb{R}}t[h'_{\ell k}(x)]^2\rho(x)\mathrm{d}x+\frac{1}{T^2}\int_{\mathbb{R}}h_{\ell k}^2(x)\rho(x)\mathrm{d}x.$$

We must compute, for $x>0$, (and then for $x<0$)

$$\frac{\sigma^2}{4}\rho^2(x)\sum_{\ell=0}^{p-1}\sum_{k=0}^{m-1}\int[h'_{\ell k}(x)]^2\le 2\sum_{\ell=0}^{p-1}\sum_{k=0}^{m-1}\left(\int_x^{+\infty}\theta_\ell\star\rho(y)\varphi_k(y)\rho(y)\mathrm{d}y\right)^2$$
$$+2\sum_{\ell=0}^{p-1}\sum_{k=0}^{m-1}\left(\int_x^{+\infty}\rho(y)\mathrm{d}y\right)^2\left(\int_{\mathbb{R}}\theta_\ell\star\rho(y)\varphi_k(y)\rho(y)\mathrm{d}y\right)^2$$
$$\le 2\sum_{\ell=0}^{p-1}\int_x^{+\infty}[\theta_\ell\star\rho(y)]^2\rho^2(y)\mathrm{d}y+2\int_{\mathbb{R}}[\theta_\ell\star\rho(y)]^2\rho^2(y)\mathrm{d}y(\int_x^{+\infty}\rho(y)\mathrm{d}y)^2.$$

Now,

$$\sum_{\ell=0}^{p-1}\int_x^{+\infty}[\theta_\ell\star\rho(y)]^2\rho^2(y)\mathrm{d}y=\sum_{\ell=0}^{p-1}\int_x^{+\infty}\left(\int_{\mathbb{R}}\theta_\ell(z)\rho(y-z)\mathrm{d}z\right)^2\rho^2(y)\mathrm{d}y$$
$$\le\int_x^{+\infty}\int_{\mathbb{R}}\rho^2(y-z)\mathrm{d}z\rho^2(y)\mathrm{d}y=\int_x^{+\infty}\rho^2(y)\mathrm{d}y\int_{\mathbb{R}}\rho^2(u)\mathrm{d}u.$$

And

$$\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}[\theta_\ell\star\rho(y)]^2\rho^2(y)\mathrm{d}y\le\left(\int_{\mathbb{R}}\rho^2(y)\mathrm{d}y\right)^2.$$

Thus,

$$\begin{aligned}\sum_{\ell=0}^{p-1}\sum_{k=0}^{m-1}\int_0^{+\infty}[h'_{\ell k}(x)]^2\rho(x)\mathrm{d}x &\lesssim \int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\int_x^{+\infty}\rho^2(y)\mathrm{d}y\int_{\mathbb{R}}\rho^2(u)\mathrm{d}u\\ &+\left(\int_{\mathbb{R}}\rho^2(y)\mathrm{d}y\right)^2\int_0^{+\infty}\frac{\mathrm{d}x}{\rho(x)}\left(\int_x^{+\infty}\rho(y)\mathrm{d}y\right)^2.\end{aligned}$$

We proceed analogously for $\int_0^{+\infty}$.
The other term $\int_{\mathbb{R}}[h_{jk}(x)]^2\rho(x)\mathrm{d}x$ is treated analogously.
This means that

$$\mathbb{E}(\sum_{k,\ell}T_2^2(k,\ell))\lesssim\frac{1}{T}+\frac{1}{T^2}.$$

Let us now study $T_1(k,\ell)$. This term contains the two trajectories $X^{(1)}$ and $X^{(2)}$. For each x, set $\theta_\ell(x-z)=\theta_{\ell,x}(z)$ and $G_{x,\ell}=\Gamma_{\theta_{\ell,x}}$ so that

$$\bar\theta_\ell(x)-\theta_\ell\star\rho(x)=\frac{\sigma}{T}\int_0^T G'_{x,\ell}(X^{(2)}(t))dW_2(t)+\frac{1}{T}[G_{x,\ell}(X^{(2)}(0))-G_{x,\ell}(X^{(2)}(T))].$$

Recall that

$$G'_{x,\ell}(y)=\frac{2}{\sigma^2\rho(y)}\int_y^{+\infty}[\theta_\ell(x-z)-\theta_\ell\star\rho(x)]\rho(z)\mathrm{d}z=-\frac{2}{\sigma^2\rho(y)}\int_{-\infty}^{y}[\theta_\ell(x-z)-\theta_\ell\star\rho(x)]\rho(z)\mathrm{d}z$$

Thus,

$$\begin{aligned}T_1(k,\ell) &= \frac{1}{T}\int_0^T\varphi_k(X^{(1)}(s))\mathrm{d}s\left[\frac{\sigma}{T}\int_0^T G'_{X^{(1)}(s),\ell}(X^{(2)}(t))dW_2(t)\right]\\ &+ \frac{1}{T}\int_0^T\varphi_k(X^{(1)}(s))\mathrm{d}s\frac{1}{T}[G_{X^{(1)}(s),\ell}(X^{(2)}(0))-G_{X^{(1)}(s),\ell}(X^{(2)}(T))]\\ &:= S_1(k,\ell)+S_2(k,\ell).\end{aligned}$$

We have:

$$S_1(k,\ell)=\frac{1}{T}\int_0^T\left[\frac{1}{T}\int_0^T\varphi_k(X^{(1)}(s))G'_{X^{(1)}(s),\ell}(X^{(2)}(t))\mathrm{d}s\right]dW_2(t)$$

Using the independence of the two trajectories, we get

$$\begin{aligned}\mathbb{E}S_1^2(k,\ell) &= \frac{1}{T^4}\int_0^T dt\mathbb{E}\left(\int_0^T\varphi_k(X^{(1)}(s))G'_{X^{(1)}(s),\ell}(X^{(2)}(t))\mathrm{d}s\right)^2\\ &= \frac{1}{T}\int_{\mathbb{R}^2}\varphi_k^2(x)[G'_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy.\end{aligned}$$

Moreover, $\mathbb{E}S_2^2(k,\ell) \leq \frac{2}{T^2}\int_{\mathbb{R}^2}\varphi_k^2(x)[G_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy$. We have to compute $\sum_{k,\ell}$ of the two terms. We have

$$\begin{aligned}\sum_{\ell=0}^{p-1}\sum_{k=0}^{m-1}\int_{\mathbb{R}^2}\varphi_k^2(x)[G'_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy &\leq L_\varphi(m)\sum_{\ell=0}^{p-1}\int_{\mathbb{R}^2}[G'_{x,\ell}(y)]^2\rho(x)\rho(y)dxdy\\ &\lesssim L_\varphi(m)\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_0^{+\infty}[G'_{x,\ell}(y)]^2\rho(y)\mathrm{d}y]\\ &+ L_\varphi(m)\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_{-\infty}^{0}[G'_{x,\ell}(y)]^2\rho(y)\mathrm{d}y]\end{aligned}$$

We only treat the term containing $\int_0^{+\infty}$. We have:

$$\begin{aligned}\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_0^{+\infty}[G'_{x,\ell}(y)]^2\rho(y)\mathrm{d}y] &\lesssim \sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_0^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int_y^{+\infty}\theta_\ell(x-z)\rho(z)\mathrm{d}z\right)^2]\\ &+\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_0^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int\theta_\ell(x-z))\rho(z)\mathrm{d}z\right)^2\left(\int_y^{+\infty}\rho(z)\mathrm{d}z\right)^2]\\ &\lesssim\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x[\int_0^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int\theta_\ell(u)\mathbf{1}_{[y,+\infty)}(x-u)\rho(x-u)\mathrm{d}u\right)^2]\\ &+\sum_{\ell=0}^{p-1}\int_{\mathbb{R}}\rho(x)\mathrm{d}x\int_0^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int_y^{+\infty}\rho(z)\mathrm{d}z\right)^2\left(\int\theta_\ell(u)\rho(x-u)\mathrm{d}u\right)^2\\ &\lesssim\int\rho(x)\rho(x-u)\frac{\rho(x-u)}{\rho(y)}\mathbf{1}_{x-u>y>0})dxdydu\\ &+\int\rho(x)\rho^2(x-u)dxdu\int_0^{+\infty}\frac{\mathrm{d}y}{\rho(y)}\left(\int_y^{+\infty}\rho(z)\mathrm{d}z\right)^2\end{aligned}$$

Using the bound on ρ and a previous study, we find that the second term above is finite. For the first term, we have

$$\begin{aligned}&\int\rho(x)\rho(x-u)\frac{\rho(x-u)}{\rho(y)}\mathbf{1}_{x-u>y>0}dxdydu \leq \int\rho(x)\rho(x-u)\mathrm{e}^{C(x-u-y))}\mathbf{1}_{x-u>y>0}dxdydu\\ &=\int\rho(x)\rho(x-u)\mathrm{e}^{C(x-u)}\int_0^{x-u}\mathrm{e}^{(-Cy)}\mathrm{d}y\mathbf{1}_{x-u>0}dxdydu\\ &=\int\rho(x)\rho(x-u)\mathrm{e}^{C(x-u)}\mathbf{1}_{x-u>0}dxdu-\int\rho(x)\rho(x-u)\mathrm{e}^{-C(x-u)}\mathbf{1}_{x-u>0}dxdu\\ &=\int\mathrm{d}u\mathrm{e}^{-Cu}\int_u^{+\infty}\rho(x)\rho(x-u)\mathrm{e}^{Cx}\mathrm{d}x-\int\mathrm{d}u\mathrm{e}^{Cu}\int_u^{+\infty}\rho(x)\rho(x-u)\mathrm{e}^{-Cx}\mathrm{d}x\end{aligned}$$

The two terms of the sum are analogous. We look at the first one. For this, we compute, using Proposition 2.1 (where b is given):

$$\rho(x)\rho(x-u) \lesssim \exp\left(-\frac{K+\lambda}{\sigma^2}[(x+b)^2+(x-u+b)^2]\right)$$
$$\lesssim \exp\left(-\frac{K+\lambda}{\sigma^2}(u^2/2)\right)\exp\left(-\frac{2(K+\lambda)}{\sigma^2}[(x+b-(u/2))^2]\right)$$

Thus,

$$\int_u^{+\infty}\rho(x)\rho(x-u)\mathrm{e}^{Cx} \lesssim \exp\left(-\frac{K+\lambda}{\sigma^2}\frac{u^2}{2}\right)\int \mathrm{e}^{Cx}\exp\left(-\frac{2(K+\lambda)}{\sigma^2}[(x+b-(u/2))^2]\right)\mathrm{d}x$$
$$\lesssim \exp\left(-\frac{K+\lambda}{\sigma^2}\frac{u^2}{2}\right)\mathrm{e}^{C((u/2)-b)}\int \mathrm{e}^{Cz}\exp\left(-\frac{2(K+\lambda)z^2}{\sigma^2}\right)\mathrm{d}z$$
$$\lesssim \exp\left(-\frac{K+\lambda}{\sigma^2}\frac{u^2}{2}\right)\mathrm{e}^{C((u/2)-b)}.$$

Finally,

$$\int_0^{+\infty}\mathrm{d}u\mathrm{e}^{-Cu}\int_u^{+\infty}\rho(x)\rho(x-u)\mathrm{e}^{Cx}\mathrm{d}x \lesssim \int_0^{+\infty}\mathrm{d}u\exp\left(-\frac{K+\lambda}{\sigma^2}(u^2/2)\right)\mathrm{e}^{-C((u/2))} < +\infty.$$

This means that

$$\sum_{k,l}\mathbb{E}S_1^2(k,\ell) \lesssim \frac{L_\varphi(m)}{T}.$$

Now, we look at $\sum_{k,\ell}\mathbb{E}S_2^2(k,\ell)$. This brings no new difficulty and we can prove that this sum is bounded by a constant $\times\frac{L_\varphi(m)}{T^2}$. $\square$

Lemma 7.9. $\mathbb{E}(T_{22}) \lesssim \frac{1}{T}$

Proof of Lemma 7.9. We write the decomposition

$$T_{2,2} = \sum_{0\le\ell,r\le p-1}\left(A_1(\ell,r)+A_2(\ell,r)+A_3(\ell,r)+A_4(\ell,r)\right)^2$$

with

$$A_1(\ell,r) = \frac{1}{T}\int_0^T\left(\overline{\theta}_\ell(X^{(1)}(s))-\theta_\ell\star\rho(X^{(1)}(s))\right)\left(\overline{\theta}_r(X^{(1)}(s))-\theta_r\star\rho(X^{(1)}(s))\right)\mathrm{d}s$$
$$A_2(\ell,r) = \frac{1}{T}\int_0^T\left(\overline{\theta}_\ell(X^{(1)}(s))-\theta_\ell\star\rho(X^{(1)}(s))\right)\theta_r\star\rho(X^{(1)}(s))\mathrm{d}s$$
$$A_3(\ell,r) = A_2(r,\ell)$$
$$A_4(\ell,r) = \frac{1}{T}\int_0^T\theta_\ell\star\rho(X^{(1)}(s))\theta_r\star\rho(X^{(1)}(s))\mathrm{d}s - \int\theta_\ell\star\rho\,\theta_r\star\rho\,\rho.$$

We only treat the term $A_1(\ell,r)$ as the other ones are easier and analogous terms have been already treated previously.

Using that $[(1/T)\int_0^T \Phi(u)\mathrm{d}u]^2 \le (1/T)\int_0^T \Phi^2(u)\mathrm{d}u$, we have

$$\begin{aligned}\mathbb{E}\sum_{r,\ell} A_1^2(\ell,r) &\le \mathbb{E}\left\{\sum_{r,\ell}\int_{\mathbb{R}}\left[(\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u))(\overline{\theta}_r(u)-\theta_r\star\rho(u))\right]^2\rho(u)\mathrm{d}u\right\}\\ &\le \mathbb{E}\left\{\int_{\mathbb{R}}\left(\sum_{\ell=0}^{p-1}(\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u))^2\right)\left(\sum_{r=0}^{p-1}(\overline{\theta}_r(u)-\theta_r\star\rho(u))^2\right)\rho(u)\mathrm{d}u\right\}\\ &= \mathbb{E}\left\{\int_{\mathbb{R}}\left(\sum_{\ell=0}^{p-1}(\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u))^2\right)^2\rho(u)\mathrm{d}u\right\} := \mathbb{A}_1 \end{aligned}\tag{7.10}$$

We first work on the sums and write each as *e.g.* $\sum_{\ell=0}^{p-1}\langle\varphi_\ell,\Xi\rangle^2$ for some complicated and random Ξ, in order to upper bound the term by $\int\Xi^2$.

Recall that

$$\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u) = \frac{\sigma}{T}\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s) + \frac{1}{T}(\Gamma_{\theta_{\ell,u}}(X^{(2)}(0)) - \Gamma_{\theta_{\ell,u}}(X^{(2)}(T))),$$

with $\theta_{\ell,u}(z)=\theta_\ell(u-z)$, and

$$\begin{aligned}\Gamma'_{\theta_{\ell,u}}(y) &= \frac{2}{\sigma^2\rho(y)}\left(\int \mathbf{1}_{v\le y}\theta_\ell(u-v)\rho(v)\mathrm{d}v - \theta_\ell\star\rho(u)\int_{-\infty}^y\rho(v)\mathrm{d}v\right)\\ &= \frac{2}{\sigma^2\rho(y)}\left(\int \mathbf{1}_{u-z\le y}\theta_\ell(z)\rho(u-z)\mathrm{d}z - \theta_\ell\star\rho(u)\int_{-\infty}^y\rho(v)\mathrm{d}v\right) &(7.11)\\ &= \frac{2}{\sigma^2\rho(y)}\left(\int \mathbf{1}_{u-z\ge y}\theta_\ell(z)\rho(u-z)\mathrm{d}z - \theta_\ell\star\rho(u)\int_{y}^{+\infty}\rho(v)\mathrm{d}v\right). &(7.12)\end{aligned}$$

Therefore

$$\sum_{\ell=0}^{p-1}(\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u))^2 \le \frac{2}{T^2}(\sigma^2 B_1(u) + C_1(u)),$$

with

$$\begin{aligned} B_1(u) &= \sum_{\ell=0}^{p-1}\left(\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s)\right)^2\\ C_1(u) &= \sum_{\ell=0}^{p-1}(\Gamma_{\theta_{\ell,u}}(X^{(2)}(0)) - \Gamma_{\theta_{\ell,u}}(X^{(2)}(T)))^2\end{aligned}$$

Then,

$$\mathbb{A}_1 = \int_{\mathbb{R}}\rho(u)\mathrm{d}u\left(\sum_{\ell=0}^{p-1}(\overline{\theta}_\ell(u)-\theta_\ell\star\rho(u))^2\right)^2 \le \frac{8}{T^4}(\sigma^4\int_{\mathbb{R}}\rho(u)\mathrm{d}uB_1^2(u) + \int_{\mathbb{R}}\rho(u)duC_1^2(u)),\tag{7.13}$$

We work on $B_1(u)$ and split it into in order to use the two forms of $\Gamma'_{\varphi_{\ell,u}}(y)$. We have:

$$\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s) = \int_0^T \mathbf{1}_{(X^{(2)}(s)<0)}\Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s) + \int_0^T \mathbf{1}_{(X^{(2)}(s)\geq 0)}\Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s)$$

Now, we use (7.11) for the first stochastic integral:

$$\begin{aligned}
&\int_0^T \mathbf{1}_{(X^{(2)}(s)<0)}\Gamma'_{\theta_{\ell,u}}(X^{(2)}(s))dW_2(s)\\
&= \frac{2}{\sigma^2}\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\left(\int \mathbf{1}_{(u-z\leq X^{(2)}(s))}\theta_\ell(z)\rho(u-z)\mathrm{d}z\right)dW_2(s)\\
&\quad -\frac{2}{\sigma^2}\theta_\ell\star\rho(u)\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\left(\int_{-\infty}^{X^{(2)}(s)}\rho(v)\mathrm{d}v\right)dW_2(s)
\end{aligned}$$

By the stochastic Fubini theorem (see *e.g.* [63], Lem. 2.1), we interchange the ordinary and the stochastic integral. This yields:

$$\begin{aligned}
&\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\left(\int \mathbf{1}_{(u-z\leq X^{(2)}(s))}\theta_\ell(z)\rho(u-z)\mathrm{d}z\right)dW_2(s)\\
&= \int \theta_\ell(z)\rho(u-z)\left(\int_0^T \frac{\mathbf{1}_{(u-z\leq X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}dW_2(s)\right)\mathrm{d}z.
\end{aligned}$$

Therefore, using the above equality, we get:

$$\begin{aligned}
&\sum_\ell\left[\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\left(\int \mathbf{1}_{(u-z\leq X^{(2)}(s))}\theta_\ell(z)\rho(u-z)\mathrm{d}z\right)dW_2(s)\right]^2 \qquad (7.14)\\
&= \sum_\ell\left[\int \theta_\ell(z)\rho(u-z)\left(\int_0^T \frac{\mathbf{1}_{(u-z\leq X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}dW_2(s)\right)\mathrm{d}z\right]^2\\
&= \int \rho^2(u-z)\left(\int_0^T \frac{\mathbf{1}_{(u-z\leq X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}dW_2(s)\right)^2 dz\\
&= \int \rho^2(v)\mathbf{1}_{v<0}M_T^2(v)\mathrm{d}v
\end{aligned}$$

with

$$M_T(v) = \int_0^T \frac{\mathbf{1}_{(v\leq X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}dW_2(s). \qquad (7.15)$$

More simply,

$$\begin{aligned}
&\sum_\ell\left(\theta_\ell\star\rho(u)\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\int_{-\infty}^{X^{(2)}(s)}\rho(v)\mathrm{d}v\, dW_2(s)\right)^2\\
&= \int \rho^2(z)\mathrm{d}z\left(\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))}\int_{-\infty}^{X^{(2)}(s)}\rho(v)\mathrm{d}v\, dW_2(s)\right)^2
\end{aligned}$$

We proceed in the same way for the term containing $\mathbf{1}_{(X^{(2)}(s)<0)}$ using the second expression of $\Gamma'_{\theta_{\ell,u}}(y)$. So that by gathering all terms, we have obtained

$$B_1(u) = \sum_{\ell=0}^{p-1} \left(\int_0^T \Gamma'_{\theta_{\ell,u}}(X^{(2)}(s)) dW_2(s) \right)^2 \lesssim B_{11}(u) + B_{12}(u) + B_{13}(u) + B_{14}(u)$$

with

$$\begin{aligned}
B_{11}(u) &= \int \rho^2(u-z)\mathrm{d}z \left(\int_0^T \frac{\mathbf{1}_{(u-z\leq X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))} dW_2(s) \right)^2 = \int \rho^2(v)\mathbf{1}_{v<0} M_T^2(v)\mathrm{d}v \\
B_{12}(u) &= \int \rho^2(z)\mathrm{d}z \left(\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)<0)}}{\rho(X^{(2)}(s))} \int_{-\infty}^{X^{(2)}(s)} \rho(v)\mathrm{d}v \; dW_2(s) \right)^2 \\
B_{13}(u) &= \int \rho^2(u-z)\mathrm{d}z \left(\int_0^T \frac{\mathbf{1}_{(0\leq X^{(2)}(s)<u-z)}}{\rho(X^{(2)}(s))} dW_2(s) \right)^2 \\
B_{14}(u) &= \int \rho^2(z)\mathrm{d}z \left(\int_0^T \frac{\mathbf{1}_{(X^{(2)}(s)>0)}}{\rho(X^{(2)}(s))} \int_{X^{(2)}(s)}^{+\infty} \rho(v)\mathrm{d}v \; dW_2(s) \right)^2.
\end{aligned}$$

Therefore, we have

$$\int \rho(u)B_1^2(u)\mathrm{d}u \lesssim 4\sum_{i=1}^{4} \int \rho(u)B_{1i}^2(u)\,\mathrm{d}u.$$

We only look at the first term:

$$\begin{aligned}
\int \rho(u)B_{11}^2(u)\,\mathrm{d}u &= \int \rho(u)\rho^2(v)\rho^2(v')\mathbf{1}_{v<0,v'<0} M_T^2(v)M_T^2(v') du dv dv' \\
&= \int \rho^2(v)\rho^2(v')\mathbf{1}_{v<0,v'<0} M_T^2(v)M_T^2(v') dv dv'.
\end{aligned}$$

We have to take the expectation of the above term. Hence, we deal with $\mathbb{E}(M_T^2(v)M_T^2(v'))$. Using the Cauchy-Schwarz and the Burkholder-Davis-Gundy inequalities yields

$$\mathbb{E}(M_T^2(v)M_T^2(v')) \leq (\mathbb{E}(M_T^4(v))\mathbb{E}(M_T^4(v')))^{1/2} \lesssim (\mathbb{E}(\langle M(v)\rangle_T^2)\mathbb{E}(\langle M(v')\rangle_T^2))^{1/2}.$$

We have (see (7.15)):

$$\begin{aligned}
\mathbb{E}(\langle M(v)\rangle_T^2) &= \mathbb{E}\left(\int_0^T \frac{\mathbf{1}_{(v<X^{(2)}(s)<0)}}{\rho^2(X^{(2)}(s))}\mathrm{d}s \right)^2 \leq T\int_0^T \mathbb{E}\left(\frac{1_{(v<X^{(2)}(s)<0)}}{\rho^4(X^{(2)}(s))} \right)\mathrm{d}s \\
&= T^2 \int \frac{1_{(v<x<0)}}{\rho^3(x)}\mathrm{d}x
\end{aligned}$$

Thus, we obtain:

$$\int \rho(u)B_{11}^2(u)\,\mathrm{d}u \lesssim T^2 \int \rho^2(v)\rho^2(v')\mathbf{1}_{v<0,v'<0}\left(\int \frac{1_{(v<x<0)}}{\rho^3(x)}\frac{1_{(v'<x'<0)}}{\rho^3(x')}dxdx'\right)^{1/2} dvdv'$$
$$= T^2\left[\int \rho^2(v)\mathbf{1}_{v<0}\left(\int \frac{1_{(v<x<0)}}{\rho^3(x)}\mathrm{d}x\right)^{1/2}\mathrm{d}v\right]^2.$$

We split $\rho^2(v) = \rho^{1/2}(v)\rho^{3/2}(v)$ and get:

$$\int \rho(u)B_{11}^2(u)\,\mathrm{d}u = T^2\left[\int \rho^{1/2}(v)\mathbf{1}_{v<0}\left(\int \rho^3(v)\frac{1_{(v<x<0)}}{\rho^3(x)}\mathrm{d}x\right)^{1/2}\mathrm{d}v\right]^2$$
$$\leq T^2\left[\int \rho^{1/2}(v)\mathbf{1}_{v<0}\exp\left[-(3/2)Cv\right]\left(\int_v^0 \exp(3Cx)\mathrm{d}x\right)^{1/2}\mathrm{d}v\right]^2$$
$$\lesssim T^2.$$

Dealing analogously with the other terms, we finally find (see (7.10) and (7.13)):

$$\mathbb{E}\mathbb{A}_1 \lesssim \frac{1}{T^2}.$$

The proof of Lemma 7.9 is complete. □

7.10. Proof of Lemma 3.6

We split the expectation along

$$\Omega_{m,p} = \left\{\|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\mathrm{op}} < \frac{1}{2}\right\}.$$

(see (7.5)) and its complement. On the complement, we have, by the definition of $\Delta_{m,p}$ (3.13),

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}^c}\right) \leq 2\mathbb{E}\left((\|\widehat{\mathbf{V}}_{m,p}^{-1}\|_{\mathrm{op}}^2 + \|\mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2)\mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}^c}\right)$$
$$\leq 10s_0^2\mathbb{P}(\Omega_{m,p}^c)$$

Moreover

$$\mathbb{P}(\Omega_{m,p}^c) \leq 4\mathbb{E}\left(\|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\mathrm{op}}^2\right) \leq 4\|\mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right).$$

Thus, gathering the last two inequalities, we get, using [H5],

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}^c}\right) \leq 40s_0^4\,\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\mathrm{op}}^2\right). \quad (7.16)$$

For the other term, we have with [H5],

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}}\right) \leq s_0^2\,\mathbb{E}\left(\|\mathbf{V}_{m,p}^{1/2}\widehat{\mathbf{V}}_{m,p}^{-1}\mathbf{V}_{m,p}^{1/2} - \mathbf{Id}_{m+p}\|_{\mathrm{op}}^2\mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}}\right).$$

Let $\mathbf{A} = \mathbf{Id}_{m+p}$ and $\mathbf{B} = \mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}$. Then we use the following theorem:

Theorem. [65] *Let* $\mathbf{A}$, $\mathbf{B}$ *be* $(m \times m)$ *matrices. If* $\mathbf{A}$ *is invertible and* $\|\mathbf{A}^{-1}\mathbf{B}\|_{\text{op}} < 1$, *then* $\tilde{\mathbf{A}} := \mathbf{A} + \mathbf{B}$ *is invertible and it holds*

$$\|\tilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_{\text{op}} \leq \frac{\|\mathbf{B}\|_{\text{op}}\|\mathbf{A}^{-1}\|_{\text{op}}^2}{1 - \|\mathbf{A}^{-1}\mathbf{B}\|_{\text{op}}}.$$

We have $\|\mathbf{A}^{-1}\mathbf{B}\|_{\text{op}} = \|\mathbf{B}\|_{\text{op}} < 1/2$ on $\Omega_{m,p}$, so that $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{B} = \mathbf{V}_{m,p}^{1/2}\widehat{\mathbf{V}}_{m,p}^{-1}\mathbf{V}_{m,p}^{1/2}$ and

$$\begin{aligned}
\|\widetilde{\mathbf{A}}^{-1} - \mathbf{A}^{-1}\|_{\text{op}} &= \|\mathbf{V}_{m,p}^{1/2}\widehat{\mathbf{V}}_{m,p}^{-1}\mathbf{V}_{m,p}^{1/2} - \mathbf{Id}_{m+p}\|_{\text{op}} \\
&\leq \frac{\|\mathbf{B}\|_{\text{op}}\|\mathbf{A}^{-1}\|_{\text{op}}^2}{1 - \|\mathbf{A}^{-1}\mathbf{B}\|_{\text{op}}} = \frac{\|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\text{op}}}{1 - \|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}^{-1}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\text{op}}} \\
&\leq \frac{1}{2}\|\mathbf{V}_{m,p}^{-1/2}\widehat{\mathbf{V}}_{m,p}\mathbf{V}_{m,p}^{-1/2} - \mathbf{Id}_{m+p}\|_{\text{op}} \\
&\leq \frac{\|\mathbf{V}_{m,p}^{-1}\|_{\text{op}}}{2}\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\text{op}}.
\end{aligned}$$

Therefore we obtain

$$\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\text{op}}^2 \mathbf{1}_{\Delta_{m,p}}\mathbf{1}_{\Omega_{m,p}}\right) \leq \mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p}^{-1} - \mathbf{V}_{m,p}^{-1}\|_{\text{op}}^2 \mathbf{1}_{\Omega_{m,p}}\right) \leq \frac{s_0^4}{4}\mathbb{E}\left(\|\widehat{\mathbf{V}}_{m,p} - \mathbf{V}_{m,p}\|_{\text{op}}^2\right). \tag{7.17}$$

By (7.16) and (7.17), we get Lemma 3.6. □

7.11. Proof of Proposition 4.1

Let $g_j := \Gamma_{\varphi_j}$ be such that $Lg_j = \varphi_j - \int \varphi_j\, \rho$, as defined in Proposition 2.4, which is such that by (2.9)

$$g_j'(x) = \frac{2}{\sigma^2\rho(x)}\int_{-\infty}^{x}[\varphi_j(y) - \int_{\mathbb{R}}\varphi_j\rho]\rho(y)\mathrm{d}y = -\frac{2}{\sigma^2\rho(x)}\int_{x}^{+\infty}[\varphi_j(y) - \int_{\mathbb{R}}\varphi_j\rho]\rho(y)\mathrm{d}y, \tag{7.18}$$

and by (2.10)

$$\int_0^T \left(\varphi_j(X(s)) - \int_{\mathbb{R}}\varphi_j\,\rho\right)\mathrm{d}s = \sigma\int_0^T g_j'(X(s))dW(s) + g_j(X(0)) - g_j(X(T)). \tag{7.19}$$

First by Pythagoras Theorem, the equality

$$\mathbb{E}\|\widehat{\rho}_m - \rho\|^2 = \|\rho - \rho_m\|^2 + \mathbb{E}\|\widehat{\rho}_m - \rho_m\|^2 \tag{7.20}$$

holds. The result will follow from a bound on the second rhs term, $\mathbb{E}\|\widehat{\rho}_m - \rho_m\|^2 = \sum_{j=0}^{m-1}\mathbb{E}[(\widehat{a}_j - a_j)^2]$ which is the variance term.

By formula (7.19), and Inequality (2.11) in Proposition 2.4, we get

$$\mathbb{E}(\|\widehat{\rho}_m - \rho_m\|^2) = \mathbb{E}\sum_{j=0}^{m-1}\left(\frac{1}{T}\int_0^T \varphi_j(X(s))\mathrm{d}s - \int_{\mathbb{R}}\varphi_j\rho\right)^2$$
$$\leq \frac{2}{T}\sum_{j=0}^{m-1}\left\{\sigma^2\int_{\mathbb{R}}(g_j')^2(x)\rho(x)\mathrm{d}x + \frac{4}{T}\int_{\mathbb{R}} g_j^2(x)\rho(x)\mathrm{d}x\right\} := \frac{2}{T}(\mathbb{T}_1 + \frac{1}{T}\mathbb{T}_2).$$

Now we use (7.18), and write

$$\mathbb{T}_1 = \sum_{j=0}^{m-1} I_j^+ + \sum_{j=0}^{m-1} I_j^-,$$

$$\text{where } I_j^- = \frac{4}{\sigma^4}\int_{-\infty}^0\left(\int_{-\infty}^x \varphi_k(y)\rho(y)\mathrm{d}y - (\int\varphi_j\rho)\int_{-\infty}^x\rho(y)\mathrm{d}y\right)^2\frac{\mathrm{d}x}{\rho(x)},$$

$$I_j^+ = \frac{4}{\sigma^4}\int_0^{+\infty}\left(\int_x^{+\infty}\varphi_k(y)\rho(y)\mathrm{d}y - (\int\varphi_j\rho)\int_x^{+\infty}\rho(y)\mathrm{d}y\right)^2\frac{\mathrm{d}x}{\rho(x)}.$$

Both terms are handled similarly, so we consider the first one only.

$$\sum_{j=0}^{m-1} I_j^+ \leq \frac{8}{\sigma^4}\int_0^{+\infty}\left\{\sum_{j=0}^{m-1}\left(\int_x^{+\infty}\varphi_j(y)\rho(y)\mathrm{d}y\right)^2 + \sum_{j=0}^{m-1}(\int\varphi_j\rho)^2(\int_x^{+\infty}\rho(y)\mathrm{d}y)^2\right\}\frac{\mathrm{d}x}{\rho(x)}$$
$$\leq \frac{8}{\sigma^4}\int_0^{+\infty}\left\{\int_x^{+\infty}\rho^2(y)\mathrm{d}y + (\int\rho^2)(\int_x^{+\infty}\rho(y)\mathrm{d}y)^2\right\}\frac{\mathrm{d}x}{\rho(x)}$$
$$\leq \frac{8}{\sigma^4}C(\rho)$$

where $C(\rho)$ is a finite constant coming from Lemma 7.1.

For $\mathbb{T}_2$, we write $\mathbb{T}_2 = \sum_{j=0}^{m-1}(\int_{-\infty}^0 g_j^2\rho + \int_0^{+\infty} g_j^2\rho)$ and consider only the last sum, both terms being again similar. For $x \geq 0$, we write that

$$g_j^2(x) = \left(\int_0^x g_j'(u)\mathrm{d}u\right)^2 \leq x\int_0^x (g_j')^2(u)\mathrm{d}u.$$

$$\text{Thus } \sum_{j=0}^{m-1}\int_0^{+\infty} g_j^2(x)\rho(x)\mathrm{d}x \leq \int_0^{+\infty} x\sum_{j=0}^{m-1}\left(\int_0^x (g_j')^2(u)\mathrm{d}u\right)\rho(x)\mathrm{d}x$$

Using (7.18), we get

$$\sum_{j=0}^{m-1}\int_0^{+\infty} g_j^2(x)\rho(x)\mathrm{d}x \leq \frac{8}{\sigma^4}\int_0^{+\infty}\left\{\int_x^{+\infty}\left[\int_u^{+\infty}\rho^2(y)\mathrm{d}y + (\int\rho^2)\left(\int_u^{+\infty}\rho(y)\mathrm{d}y\right)^2\right]\frac{\mathrm{d}u}{\rho^2(u)}\right\}x\rho(x)\mathrm{d}x.$$

This term is finite by Lemma 7.1. So $\mathbb{T}_2$ is bounded by say $C^\star(\rho)/\sigma^4$. This yields

$$\mathbb{E}(\|\widehat{\rho}_m - \rho_m\|^2) \leq \frac{C(\rho)}{\sigma^4}\frac{1}{T}(1+\frac{1}{T}).$$

Plugging this in (7.20) ends the proof of Proposition 4.1. □

Data availability statement

No new data/codes were created or analyzed in this study.

References

[1] H.P. McKean, Jr., A class of Markov processes associated with nonlinear parabolic equation. *PNAS* **6** (1966) 1907–1911.

[2] D. Benedetto, E. Caglioti and M. Pulverenti, A kinetic equation for granular media. *Math. Model. Numer. Anal.* **31** (1997) 615–641.

[3] J. Baladron, D. Fasoli, O. Faugeras and J. Touboul, Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons. *J. Math. Neurosci.* **2** (2012) 1–50.

[4] D.A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Statist. Phys.* **31** (1983) 29–85.

[5] L.A. Carrillo, Y.-P. Choi and M. Hauray, The derivation of swarming models: mean-field limit and Wasserstein distances, in *Collective Dynamics from Bacteria to Crowds*, edited by A. Muntean and F. Toschi. CISM International Centre for Mechanical Sciences, Vol. 553. Springer, Vienna (2014).

[6] A. Molginer and L. Edelstein-Keshet, A non-local model for a swarm. *J. Math. Biol.* **38** (1999) 534–570.

[7] F. Ball and D. Sirl, Stochastic SIR in structured populations. *Stochastic Epidemic Models Inference* **Part II** (2020) 123–240. Lecture Notes in Mathematics, Vol. 2255, edited by T. Britton and E. Pardoux. Mathematical Biosciences Subseries. Springer.

[8] R. Forien and E. Pardoux, Household epidemic models and McKean–Vlasov Poisson driven SDEs. *Ann. Appl. Probab.* **32** (2022) 1210–1233.

[9] K. Giesecke, G. Schwenkler and J.A. Sirignano, Inference for large financial systems. *Math. Finance* **30** (2020) 3–46.

[10] A.-S. Sznitman, Topics in propagation of chaos. *Ecole d'été de probabilités de Saint-Flour XIX-1989. Lecture Notes in Math.*, Vol. 1464 Springer, Berlin (1991) 165–251.

[11] S. Méléard, Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models, in Probabilistic Models for Nonlinear Partial Differential Equations, Lecture Notes in Mathematics, Vol. 1627. Springer (1996) 42–95.

[12] T. Funaki, A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrscheinl. verw. Gebiete* **67** (1984) 331–348.

[13] J. Gärtner, On the Mc-Kean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137** (1988) 197–248.

[14] S. Benachour, B. Roynette and P. Vallois, Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* **75** (1998a) 173–201.

[15] S. Benachour, B. Roynette and P. Vallois, Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl.* **75** (1998b) 203–224.

[16] S. Herrmann, P. Imkeller and D. Peithmann, Large deviations and a Kramers'type low for self-stabilizing diffusions. *Ann. Appl. Probab.* **18** (2008) 1379–1423.

[17] V.N. Kolokoltsov, Non-linear Markov Processes and Kinetic Equations, Vol. 182. Cambridge University Press (2010).

[18] F. Bolley, I. Gentil and A. Guillin, Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.* **208** (2013) 429–445.

[19] A. Eberle, A. Guillin and R. Zimmer, Quantitative Harris type theorems for diffusions and Mc-Kean–Vlasov processes. *Trans. Amer. Math. Soci.* (2019) DOI:10.1090/TRAN/7576.

[20] N. Fournier and A. Guillin, On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** (2015) 707–738.

[21] F. Malrieu, Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13** (2003) 540–560.

[22] P. Cattiaux, A. Guillin and F. Malrieu, Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** (2008) 19–40.

[23] O. Butkovsky, On ergodic properties of nonlinear Markov chains and Mc-Kean–Vlasov equations. *Theory Probab. Appl.* **58** (2014) 661–674.

[24] S. Herrmann and J. Tugaut, Non uniqueness of stationary measures for self-stabilizing diffusions. *Stochastic Process. Appl.* **120** (2010) 1215–1246.

[25] A. Durmus, A. Eberle, A. Guillin and R. Zimmer, An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** (2020) 5387–5398.

[26] R.A. Kasonga, Maximum likelihood theory for large interacting systems. *SIAM J. Appl. Math.* **50** (1990) 865–875.

[27] X. Chen, Maximum likelihood estimation of potential energy in interacting particle systems from single-trajectory data. *Electron. Commun. Probab.* **26** (2021) 1–13.

[28] Z. Li, F. Lu, M. Maggioni, S. Tang and C. Zhang, On the identifiability of interaction functions in systems of interacting particles. *Stochastic Process. Appl.* **132** (2021) 135–163.

[29] L. Sharrock, N. Kantas, P. Parpas and G.A. Pavliotis, Online parameter estimation for the McKean–Vlasov stochastic differential equation *Stochastic Process. Appl.* **162** (2023) 481–546.

[30] F. Lu, M. Maggioni and S. Tang, Learning interaction kernels in stochastic systems of interacting particles from multiple trajectories. *Found. Computat. Math.* **22** (2022) 1013–1067.

[31] C. Amorino, A. Heidari, V. Pilipauskaitė and M. Podolskij, Parameter estimation of discretely observed interacting particle systems. *Stochastic Process. Appl.* **163** (2023) 350–386.

[32] L. Della Maestra and M. Hoffmann, The LAN property for McKean–Vlasov models in a mean-field regime. *Stochastic Process. Appl.* **155** (2022b) 109–146.

[33] G.A. Pavliotis and A. Zanoni, Eigenfunction martingale estimators for interacting particle systems and their mean field limit. *SIAM J. Appl. Dyn. Syst.* **21** (2022) 2338–2370.

[34] D. Belomestny, V. Pilipauskaitė and M. Podolskij, Semiparametric estimation of McKean–Vlasov SDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **59** (2023) 79–96.

[35] L. Della Maestra and M. Hoffmann, Nonparametric estimation for interacting particle systems: McKean–Vlasov models. *Probab. Theory Related Fields* **182** (2022a) 551–613.

[36] F. Comte and V. Genon-Catalot, Nonparametric adaptive estimation for interacting particle systems. *Scand. J. Statist.* **50** (2023) 1716–1755.

[37] C. Amorino, D. Belomestny, V. Pilipauskaite, M. Podolskij and S.-Y. Zhou, Polynomial rates *via* deconvolution for nonparametric estimation in McKean–Vlasov SDEs. *ArXiv preprint* arXiv:2401.04667v2 (2024).

[38] D. Belomestny, M. Podolskij and S.Y. Zhou, On nonparametric estimation of the interaction function in particle system models. *ArXiv preprint* arXiv:2402.14419 (2024).

[39] V. Genon-Catalot and C. Larédo, Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations. *Stochastic Process. Appl.* **142** (2021a) 513–548.

[40] V. Genon-Catalot and C. Larédo, Parametric inference for small variance and long time horizon McKean–Vlasov diffusion models. *Electron. J. Statist.* **15** (2021b) 5811–5854.

[41] V. Genon-Catalot and C. Larédo, Inference for ergodic McKean−Vlasov stochastic differential equations with polynomial interactions. *Ann. Instit. H. Poincaré (B) Probab. Statist.* **60** (2024) 2668–2693.

[42] V. Genon-Catalot and C. Larédo, Parametric inference for ergodic McKean–Vlasov stochastic differential equations. *Bernoulli* **30** (2023b) 1971–1997.

[43] G.A. Pavliotis and A. Zanoni, A method of moments estimator for interacting particle systems and their mean field. *SIAM/ASA J. Uncertainty Quantif.* **12** (2024) 262–288.

[44] M. Hoffmann, Adaptive estimation in diffusion processes. *Stochastic Process. Appl.* **79** (1999) 135–163.

[45] F. Comte, V. Genon-Catalot and Y. Rozenholc, Penalized nonparametric mean square estimation of the coefficients of diffusion processes. *Bernoulli* **13** (2007) 514–543.

[46] M. Kessler, A. Lindner and M. Sørensen, editors, *Statistical Methods for Stochastic Differential Equations.* CRC Press. Taylor & Francis Group, Boca Raton (2012).

[47] A. Dalalyan and M. Reiss, Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** (2006) 248–282.

[48] A. Dalalyan and M. Reiss, Asymptotic statistical equivalence for ergodic diffusions: the multidimensional case. *Probab. Theory Related Fields* **137** (2007) 25–47.

[49] E. Schmisser, Penalized nonparametric drift estimation for a multidimensional diffusion process. *Statistics* **47** (2013) 61–84.

[50] Strauch, Sharp adaptive drift estimation for ergodic diffusions: the multivariate case. *Stochastic Process. Appl.* **125** (2015) 2562–2602.

[51] C. Strauch, Exact adaptive pointwise drift estimation for multidimensional ergodic diffusions. *Probab. Theory Related Fields* **164** (2016) 361–400.

[52] R. Nickl and K. Ray, Nonparametric statistical inference for drift vector fields of multi-dimensional diffusions. *Ann. Statist.* **48** (2020) 1383–1408.

[53] F. Comte and V. Genon-Catalot, Drift estimation on non-compact support for diffusion models. *Stochastic Process. Appl.* **134** (2021) 174–207.

[54] J. Carrillo, R.J. McCann and C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamer.* **19** (2003) 971–1018.

[55] V. Genon-Catalot, T. Jeantheau and C. Larédo, Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* **6** (2000) 1051–1079.

[56] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series, 55 For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. (1964).

[57] G. Szegö, Orthogonal Polynomials. 4th edn. American Mathematical Society, Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, R.I. (1975).

[58] F. Comte and C. Lacour, Non-compact estimation of the conditional density from direct or noisy data. *Ann. Inst. Henri Poincaré, Probab. Stat.* **59** (2023) 1463–1507.

[59] F. Comte and V. Genon-Catalot, Regression function estimation as a partly inverse problem. *Ann. Inst. Statist. Math.* **72** (2020) 1023–1054.

[60] F. Dussap, Nonparametric multiple regression by projection on non-compactly supported bases. *Ann. Inst. Statist. Math.* **75** (2023) 731–771.

[61] F. Comte and V. Genon-Catalot, New results for drift estimation in inhomogeneous stochastic differential equation. *J. Multivariate Anal.* **208** (2025) 105415.

[62] C. Amorino, A. Gloter, Estimation of the invariant density for discretely observed diffusion processes: impact of the sampling and of the asynchronicity. *Statistics* **57** (2023) 213–259.

[63] J.E. Hutton and P.I. Nelson, Interchanging the order of differentiation and stochastic integration. *Stochastic Process. Appl.* **18** (1984) 371–377.

[64] J. Indritz, An inequality for Hermite polynomials. *Proc. Amer. Math. Soc.* **12** (1961) 981–983.

[65] G.W. Stewart and J.-G. Sun, Matrix Perturbation Theory. Academic Press, Inc., Boston (1990).

[66] F. Comte and F. Merlevède, Super optimal rates for nonparametric density estimation via projection estimators. *Stochastic Process. Appl.* **115** (2005) 797–826.