A wavelet particle approximation for McKean–Vlasov and 2D-Navier–Stokes statistical solutions

Viet Chi Tran*

Université Paris X-Nanterre, Equipe Modal’X, batiment G, 200 avenue de la République, 92101 Nanterre Cedex, France

Received 21 March 2006; received in revised form 13 November 2006; accepted 17 April 2007
Available online 3 May 2007

Abstract

Letting the initial condition of a PDE be random is interesting when considering complex phenomena. For 2D-Navier–Stokes equations, it is for instance an attempt to take into account the turbulence arising with high velocities and low viscosities. The solutions of these PDEs are random and their laws are called statistical solutions.

We start by studying McKean–Vlasov equations with initial conditions parameterized by a real random variable \( \theta \), and link their weak measure solutions to the laws of nonlinear SDEs, for which the drift coefficients are expressed as conditional expectations in the diffusions’ laws given \( \theta \). We propose an original stochastic particle method to compute the first-order moments of the statistical solutions, obtained by approximating the conditional expectations by wavelet regression estimators. We establish a convergence rate that improves the ones obtained for existing methods with Nadaraya–Watson kernel estimators.

We then carry over these results to 2D-Navier–Stokes equations and compute some physical quantities of interest, like the mean velocity vector field. Numerical simulations illustrate the method and allow us to test its robustness.

© 2007 Elsevier B.V. All rights reserved.

MSC: primary 60K35; 42C40; secondary 76D06; 65C20; 65C35; 62G08

Keywords: Statistical solution; 2D-Navier–Stokes equation; McKean–Vlasov equation; Stochastic particle approximation; Wavelet regression estimator; Numerical discretization scheme of SDEs

1. Introduction

Partial differential equations (PDEs) with random initial conditions are interesting when considering complex phenomena or when introducing the notion of uncertainty in the
initial state. The solutions of these PDEs are random and their laws are called statistical solutions.

The 2D-Navier–Stokes equations, which model the velocity \( \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2 \) of homogeneous viscous incompressible fluids in the plane, have been among the equations motivating these developments. A random initial condition is then an attempt to take into account the turbulence arising with high velocities and low viscosities. Vishik and Fursikov [25], Constantin and Wu [5] have studied these equations with analytical tools.

In this article, we are interested in a probabilistic approach to these problems, which will allow us to develop stochastic numerical approximation schemes generalizing the work of Talay and Vaillant [22,24]. More precisely, we consider the following 2D-Navier–Stokes equation: \( P(\omega) \)-almost surely \( (P(\omega)-a.s.) \forall x = (x_1, x_2) \in \mathbb{R}^2, \forall t \in \mathbb{R}^+ \),

\[
\begin{cases}
\frac{\partial \mathbf{u}}{\partial t}(t, x, \omega) + (\mathbf{u} \cdot \nabla)\mathbf{u}(t, x, \omega) = v \Delta \mathbf{u}(t, x, \omega) - \nabla p \\
\nabla \cdot \mathbf{u}(t, x, \omega) = 0, \quad \mathbf{u}(0, x, \omega) = \hat{u}_0(x, \omega).
\end{cases}
\] (1.0.1)

The random initial condition is a random function \( \hat{\mathbf{u}}_0(x, \omega) \), depending on \( x \in \mathbb{R}^2 \) and on an alea \( \omega \) belonging to the probability space \((\Omega, \mathcal{F}, P)\). The pressure is denoted by \( p \) and the viscosity by \( v \). A probabilistic approach for 2D-Navier–Stokes equations with deterministic initial conditions, carried by Marchioro and Pulvirenti [14] and Méléard [15,16], relies on the 2D-vortex equations obtained by considering the curl of the velocity \( v = \text{curl}(\mathbf{u}) \). In the case of a random initial condition, we obtain: \( P(\omega) - a.s., \forall x \in \mathbb{R}^2, \forall t \in \mathbb{R}^+ \),

\[
\begin{cases}
\frac{\partial v}{\partial t}(t, x, \omega) = -(K * v \cdot \nabla)v(t, x, \omega) + v \Delta v(t, x, \omega), \\
v(0, x, \omega) = \hat{v}_0(x, \omega) = \text{curl}(\hat{\mathbf{u}}_0)(x, \omega),
\end{cases}
\] (1.0.2)

where \( K \) is the Biot and Savart kernel:

\[
\forall x = (x_1, x_2) \in \mathbb{R}^2, \quad K(x) = \frac{1}{2\pi|x|^2} (-x_2, x_1). 
\]

(1.0.3)

The 2D-vortex equations generalize the McKean–Vlasov equations (the kernel \( K \) explodes at 0). McKean–Vlasov equations with random probability density initial conditions have been studied by Talay and Vaillant [22]. They generalized the probabilistic approach developed by Sznitman [20] and Méléard [17] to the case of a random initial condition and proposed two particle approximations to compute numerically the moments of the statistical solutions. They left, however, the case of 2D-vortex equations open. Our purpose is to provide original stochastic particle methods to compute the first-order moments of the statistical solutions of McKean–Vlasov and 2D-vortex equations. This will allow us to approximate physical quantities of interest for 2D-Navier–Stokes equations with random initial conditions. We start by studying the following McKean–Vlasov equation with random initial condition: \( P(\omega) - a.s., \forall x \in \mathbb{R}^n, \forall t \in \mathbb{R}^+ \),

\[
\begin{cases}
\frac{\partial}{\partial t} v(t, x, \omega) = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i} (u_{b,i}(t, x, \omega)v(t, x, \omega)) + \frac{\sigma^2}{2} \Delta v(t, x, \omega) \\
u_b(t, x, \omega) = \int_{\mathbb{R}^n} b(x, y)v(t, y, \omega)dy, \quad v(0, x, \omega) = \hat{v}_0(x, \omega)
\end{cases}
\] (1.0.4)
where \( b \) is Lipschitz continuous and bounded from \((\mathbb{R}^n)^2 \) to \( \mathbb{R}^n \) and \( \sigma \neq 0 \). The choice of a nonconstant diffusion coefficient with the form of \( \mu_b \) can be studied in a similar way, as considered in [23], but for the sake of simplicity it is not developed here.

**Assumption 1.1.** In the following, we assume that \( \hat{v}_0 \) has the following form:

\[
P(\,d\omega) - a.s., \quad \hat{v}_0(\,.,\omega) = v_0(\,.,\theta(\omega)), \quad \text{where:} \qquad (1.0.5)
\]

(i) \( v_0 \) is a deterministic bounded function on \( \mathbb{R}^n \times \mathbb{R} \) integrable in \( x \in \mathbb{R}^n \),

(ii) \( \theta \) is a real random variable with distribution function \( G \). Its law \( v_G \) is assumed absolutely continuous with respect to the Lebesgue measure on \( \mathbb{R} \) with a connected support \( \Theta \).

The law \( L(\hat{v}_0) \) of the initial condition is the image law of the probability measure \( v_G(\,d\alpha) \) on \( \mathbb{R} \) through \( v_0 \). This will allow us to simulate real random variables \( \theta \) instead of infinite-valued ones \( \hat{v}_0 \). The form \((1.0.5)\) was already assumed in [22], where the case of a discrete probability measure \( v_G(\,d\alpha) \) has also been treated.

In Section 2, we study Eq. \((1.0.4)\) with a probabilistic interpretation due to Talay and Vaillant [22], and link its weak solutions to the laws of nonlinear diffusions whose drift coefficients are expressed as conditional expectations in the law of the diffusions given \( \theta \). Then, we propose an original stochastic particle approximation obtained by replacing the unknown conditional expectations by wavelet regression estimators. Our main result, which gives the conditional expectations by wavelet regression estimators. Our main result, which gives the convergence rate of the method, is stated and explained in Theorem 2.16, and proved in Section 3. It involves an error term of order \( N^{-s/(2s+1)} \), where \( N \) is the number of particles and \( s \) is a parameter related to the regularities of \( v_0 \) and \( G \) that is precised in the sequel. We generalize our results to 2D-vortex equations in Section 4. Simulations carried in Section 5 allow us to simulate the random velocity vector field of a 2D-Navier–Stokes equation with a random initial condition, and to observe the robustness of our approximation on several test cases.

**Notation:** Let \( C^k_b(E,F) \) (resp. \( C^{k+\varepsilon}_b(E,F) \), \( B_b(E,F) \), \( L^p(E,F) \)) be the set of bounded functions from \( E \) to \( F \) of class \( C^k \) with continuous bounded partial derivatives up to order \( k \in \mathbb{N} \) (resp. of functions of \( C^k_b(E,F) \) whose order \( k \) derivatives are \( \varepsilon \)-Hölder continuous, of bounded measurable functions, of measurable functions \( f \) such that \( \int |f|^p < \infty \)). We denote by \( \|f\|_p \) and \( \|f\|_{\text{Lip}} \) the \( L^p \) and Lipschitz norms, when they exist.

We denote by \( \mathcal{P}(E) \) the set of probabilities on \( E \), embedded with the \( L^1 \)-Vaserstein metric:

\[
\forall \eta_1, \eta_2 \in \mathcal{P}(E), \quad W_1(\eta_1, \eta_2) = \inf \int_{E \times E} |x - y| \pi(\,dx,\,dy), \quad (1.0.6)
\]

the infimum being taken on all probability measures \( \pi \) with marginals \( \eta_1 \) and \( \eta_2 \). We denote by \( \mathcal{M}_S(E) \) the space of signed measures on \( E \) with a finite total variation norm.

\( T > 0 \) is an arbitrary finite time, and \( C > 0 \) is a constant that can change from line to line.

2. A probabilistic approach to the computation of McKean–Vlasov statistical solutions’ intensities

2.1. Statistical solutions of McKean–Vlasov PDEs and their intensities

We consider the McKean–Vlasov PDE \((1.0.4)\) with a random initial condition. The probabilistic approach, which constitutes the frame of our study, relies on its weak form.
Definition 2.1. Let \( v_0, \theta \) be as in Assumption 1.1. The random variable \( (R_t(dx, \theta))_{t \in [0, T]} \in C([0, T], \mathcal{M}_S(\mathbb{R}^n)) \) is a weak measure-solution of (1.0.4) if \( P(\omega) - a.s., \forall \varphi \in C^2_b(\mathbb{R}^n, \mathbb{R}), \forall t \in [0, T], \)

\[
\begin{align*}
\left\{ \begin{array}{l}
\int_{\mathbb{R}^n} \varphi(x) R_t(dx, \theta) = \int_{\mathbb{R}^n} \varphi(x)v_0(x, \theta)dx \\
+ \int_0^t \int_{\mathbb{R}^n} \left( \sum_{i=1}^n u_{b,i}(s, x, \theta) \frac{\partial \varphi}{\partial x_i}(x) + \frac{\sigma^2}{2} \Delta \varphi(x) \right) R_s(dx, \theta) ds \\
u_b(t, x, \theta) = \int_{\mathbb{R}^n} b(x, y) R_t(dy, \theta).
\end{array} \right.
\end{align*}
\]

(2.1.1)

When \( P(\omega) - a.s. \) and for all \( t \in [0, T], \) the time marginals \( R_t(dx, \theta) \) admit densities \( v(t, x, \theta) \) with respect to the Lebesgue measure \( dx \) on \( \mathbb{R}^n, \) the family of these densities is called weak function-solution of the PDE.

We will say that \( R(\theta) \in \mathcal{M}_S(C([0, T], \mathbb{R}^n)) \) is a weak measure-solution of PDE (2.1.1) when its time projections constitute a weak measure-solution in the sense of Definition 2.1.

Let us now define the statistical solutions of the McKean–Vlasov problem:

Definition 2.2. When Eq. (2.1.1) admits a unique solution \( (R_t(dx, \theta))_{t \in [0, T]} \) (up to a null-set), the following map is \( \mathcal{L}(\hat{v}_0)-a.s. \) well defined:

\[
S : v_0(., a) \in L^1(\mathbb{R}^n, \mathbb{R}) \cap L^\infty(\mathbb{R}^n, \mathbb{R}) \mapsto (R_t(dx, a))_{t \in [0, T]} \in C\left([0, T], \mathcal{M}_S(\mathbb{R}^n)\right).
\]

The law \( m \in \mathcal{P}(C([0, T], \mathcal{M}_S(\mathbb{R}^n))) \) of the weak measure-solution \( (R_t(dx, \theta))_{t \in [0, T]} \) of (2.1.1) is called statistical solution of the McKean–Vlasov problem. Let us introduce:

\[
\tilde{\Phi} : a \in \Theta \mapsto v_0(., a) \in L^1(\mathbb{R}^n, \mathbb{R}) \cap L^\infty(\mathbb{R}^n, \mathbb{R}).
\]

Then, \( m \) is the image measure of \( \mathcal{L}(\hat{v}_0) \) through \( S, \) and the image measure of \( v_G \) through \( S \circ \tilde{\Phi}: \)

\[
m = \mathcal{L}(\hat{v}_0) \circ S^{-1} = v_G \circ (S \circ \tilde{\Phi})^{-1}.
\]

(2.1.2)

For \( t \in [0, T], \) the law \( m_t \in \mathcal{P}(\mathcal{M}_S(\mathbb{R}^n)) \) of \( R_t(dx, \theta) \) is the \( t \)-time projection of \( m, \) and is called spatial statistical solution at time \( t \) of the McKean–Vlasov equation.

Definition 2.3. The intensity \( I(m_t) \in \mathcal{M}_S(\mathbb{R}^n) \) of \( m_t \in \mathcal{P}(\mathcal{M}_S(\mathbb{R}^n)) \) is given by:

\[
\forall f \in B_0(\mathbb{R}^n, \mathbb{R}), \langle I(m_t), f \rangle = \int_{\mathcal{M}_S(\mathbb{R}^n)} \langle R_t, f \rangle m_t(dR) = \int_{\mathbb{R}} \langle (S \circ \tilde{\Phi}(a))_t, f \rangle v_G(da).
\]

(2.1.3)

Our aim is to approximate (2.1.3) by using a stochastic particle method with wavelets.

Remark 2.4. When there exists a unique weak function-solution to PDE (2.1.1), and when the initial mean energy is finite \( (\int_\mathbb{R} \| v_0(., a) \|_2^2 v_G(da) < \infty), \) the intensity \( I(m_t) \) can be linked to the first moment of the statistical solution as defined in [22]: \( \forall t \in [0, T], \exists M_1(t) \in L^2(\mathbb{R}^n, \mathbb{R}), \forall f \in L^2(\mathbb{R}^n, \mathbb{R}), \langle I(m_t), f \rangle = \langle M_1(t), f \rangle_{L^2(\mathbb{R}^n, \mathbb{R})}. \)
2.2. Probabilistic approach to the problem

Assumption 2.5. We assume that:

(A1) The function \( b \) is a bounded Lipschitz continuous function,
(A2) The coefficient \( \sigma \) is non zero,
(A3) \( P(\omega)\)-a.s., \( v_0(.,\theta) \in L^1(\mathbb{R}^n, \mathbb{R}) \cap L^\infty(\mathbb{R}^n, \mathbb{R}) \), and \( \exists A > 0 \), \( P(\omega)\)-a.s., \( \|v_0(.,\theta)\|_1 \leq A \), and \( \|v_0(.,\theta)\|_\infty \leq A \),
(A4) \( P(\omega)\)-a.s., \( \int x^2 v_0(x,\theta) \, dx < \infty \).

A probabilistic interpretation of McKean–Vlasov equations with initial conditions that are random probability densities has been given in Talay and Vaillant [22]. We extend it to the case where \( v_0 \) is a bounded integrable function satisfying (A3) and (A4). To this purpose, we follow Jourdain [10] and introduce the bounded random function \( h \) defined by: \( P(\omega)\)-a.s., \( \forall x \in \mathbb{R}^n \),

\[
h(x, \theta) = v_0(x, \theta)\|v_0(.,\theta)\|_1/\|v_0(.,\theta)\|_1 = \text{sign}(v_0(x, \theta))\|v_0(.,\theta)\|_1 \text{ or } 0,
\]

(2.2.1)

with the convention \( 0/0 = 0 \). Then \( P(\omega)\)-a.s., \( \forall x \in \mathbb{R}^n \), \( v_0(x, \theta) = h(x, \theta) p_0(x, \theta) \) where \( p_0 \) is the probability density:

\[
p_0(x, \theta) = \|v_0(x, \theta)\|/\|v_0(.,\theta)\|_1.
\]

(2.2.2)

For a probability transition measure \( Q(dy,a) \) on \( \mathcal{C}([0,T], \mathbb{R}^n) \), measurable in \( a \in \mathbb{R} \), let us introduce the following family \( (\tilde{Q}_t(dx, a))_{t \geq 0} \) of weighted signed transition measures on \( \mathbb{R}^n \), measurable in \( a \): \( P(\omega)\)-a.s., \( \forall B \) Borel subset of \( \mathbb{R}^n \), \( \forall t \in [0,T] \),

\[
\tilde{Q}_t(B, \theta) = \int_{\mathcal{C}([0,T], \mathbb{R}^n)} 1_B(y(t))h(y(0), \theta) Q(dy, \theta).
\]

(2.2.3)

Remark 2.6. If \( P(\omega)\)-a.s., and for all \( t \in [0,T] \), \( Q_t(dx, \theta) \) is absolutely continuous with respect to the Lebesgue measure on \( \mathbb{R}^n \), so is \( \tilde{Q}_t(dx, \theta) P(\omega) \) a.s.

The probabilistic approach consists in looking for Markovian processes whose time marginals are associated with the weak PDE (2.1.1) through (2.2.3). Here, the process of interest is given by the nonlinear stochastic differential equation (SDE) of Theorem 2.7. We state existence and uniqueness results for this SDE, and then for the evolution equation (2.1.1).

Notice the two different sources of randomness associated with the initial condition, through the random variable \( \theta \), and with the Brownian motion of the probabilistic approach.

Theorem 2.7. Suppose that Assumptions 1.1 and 2.5 are satisfied. Let \((W_t)_{t \in [0,T]}\) be a Brownian motion on \( \mathbb{R}^n \), and \((X_0(a))_{a \in \mathbb{R}}\) be a family of random variables such that \( a \mapsto X_0(a) \) is measurable and such that \( P(\omega)\)-a.s., \( \mathcal{L}(X_0(\theta)) = p_0(x, \theta) dx \), where \( p_0 \) is associated with \( v_0 \) by (2.2.2). We assume that \((W_t)_{t \in [0,T]}\), \( \theta \) and \((X_0(a))_{a \in \mathbb{R}}\) are independent. Then, pathwise existence and uniqueness hold for: \( P(\omega)\)-a.s., \( \forall t \in [0,T] \),

\[
\begin{aligned}
\mathrm{d}X_t(\theta) &= u_b(t, X_t(\theta), \theta) \, \mathrm{d}t + \sigma \, \mathrm{d}W_t \\
\mathcal{L}(dx, \theta) &= \mathcal{L}(X(\theta)) = \tilde{Q}(dx, \theta) \text{ is associated with } \tilde{Q}(dx, \theta) \text{ by (2.2.3)} \\
\mathcal{L}(\theta) &= v_G(dx), \mathcal{L}(X_0(\theta)) = p_0(x, \theta) dx, \quad u_b(t, x, \theta) = \int_{\mathbb{R}^n} b(x, y) \tilde{Q}_t(dy, \theta).
\end{aligned}
\]

(2.2.4)

Sketch of proof. The idea is that considering the conditioned diffusion given \( \theta \) leads us to the study of a diffusion with an initial condition of deterministic law.
Notice first that the law $\mathcal{L}(X_t(\theta))$ depends continuously on the initial condition $X_0(\theta)$ (see Kunita [12]), which is itself measurable in $\theta$. The conditional law $\mathcal{L}(X_t(\theta) \mid \theta)$ is therefore well defined. Since $\theta$ is independent of the Brownian motion $(W_t)_{t \in [0,T]}$ and of the family $(X_0(a))_{a \in \mathbb{R}}$,

$$\mathcal{L}(X_t(\theta) \mid \theta = a) = \mathcal{L}(X_t(a)),$$  \hspace{1cm} (2.2.5)

where $(X_t(a))_{t \in [0,T]}$ is the solution of the following SDE:

$$\begin{align*}
    \frac{dX_t(a)}{dt} &= u_b(t, X_t(a), a)dt + \sigma dW_t \\
    Q(dx, a) &= \mathcal{L}(X(a)), \quad Q(dx, a) \text{ is associated with } Q(dx, a) \text{ by (2.2.3)} \\
    \mathcal{L}(X_0(a)) &= p_0(x, a)dx, \quad u_b(t, x, a) = \int_{\mathbb{R}^n} b(x, y)\tilde{Q}_t(dy, a).
\end{align*}$$  \hspace{1cm} (2.2.6)

For a given realization $a$ of $\theta$, the proof is an adaptation of the proofs in [20,17] (see [23], Chapter 8 for detailed proofs).

**Proposition 2.8.** There exists a unique weak function-solution of (2.1.1) in $\mathcal{H}$:

$$\mathcal{H} = \left\{ (v_t(x))_{t \in [0,T]} \mid \forall t \in [0, T], \, v_t \in L^1(\mathbb{R}^n, \mathbb{R}) \cap L^{\infty}(\mathbb{R}^n, \mathbb{R}), \sup_{t \in [0,T]} \max(\|v_t\|_1, \|v_t\|_{\infty}) \leq A \right\}. \hspace{1cm} (2.2.7)$$

**Sketch of proof.** Let $(X_t(\theta))_{t \in [0,T]}$ be the unique pathwise solution of SDE (2.2.6), and let $\tilde{Q}(dx, \theta) \in \mathcal{C}([0, T], \mathcal{M}_\mathcal{S}(\mathbb{R}^n))$ be the measure associated with $\mathcal{L}(X(\theta))$ by (2.2.3). Using Itô’s formula to compute $\mathbb{E}(h(X_0(\theta), \theta)\varphi(X_t(\theta)))$, for $\varphi \in C^2_b(\mathbb{R}^n, \mathbb{R})$ and $h$ defined in (2.2.3), gives that $\tilde{Q}(dx, \theta)$ is a weak measure-solution of PDE (1.0.4). From Girsanov’s Theorem, $\mathcal{L}(X_t(\theta))$ admits for every $t \in [0, T]$ a density with respect to the Lebesgue measure on $\mathbb{R}^n$. Remark 2.6 concludes the proof of the existence of a weak function-solution. Following Mélaéard [16], the family of densities of $(\tilde{Q}_t(dx, \theta))_{t \in [0,T]}$ belongs $P(dx) - a.s.$ to $\mathcal{H}$. The uniqueness of the weak solution $(v(t, .., \theta))_{t \in [0,T]}$ of PDE (2.1.1) in $\mathcal{H}$ is proved by mean of a mild equation similar to the one introduced in [16].

**Remark 2.9.** When the initial condition is a signed measure, we require the boundness of $v_0$ for the uniqueness result in Proposition 2.8. If it is a nonnegative finite measure, we give in [23] a proof based on a Theorem of Bhatt and Karandikar [2] that relaxes the boundness of $v_0$.

### 2.3. Particle approximations of the intensity of spatial statistical solutions

Using the probabilistic approach of Section 2.2, we can now reformulate the intensity of the spatial statistical solutions defined in (2.1.3) with the nonlinear diffusions (2.2.4) and (2.2.6):

**Corollary 2.10.** $\forall t \in [0, T], \, \forall f \in \mathcal{B}_b(\mathbb{R}^n, \mathbb{R}),$

$$\langle I(m_t), f \rangle = \int_{\mathbb{R}} \mathbb{E}(h(X_0(a), a) f(X_t(a))) \nu_G(da)$$

$$= \mathbb{E}(h(X_0(\theta), \theta) f(X_t(\theta))).$$  \hspace{1cm} (2.3.1)
We now review three computable approximations of these quantities. The first of them relies on the first equality of (2.3.1), while the second and third ones use the second one. The two first approximations have been studied in [22]. The third one is our original approximation.

We will introduce Euler schemes. The discretization step is \( \Delta t = T/K \), for \( K \in \mathbb{N}^* \), which defines \( K + 1 \) discretization times \( t_k = k \Delta t \) for \( k = 0 \) to \( K \).

**Remark 2.11.** Until Section 4, we will assume, for the sake of simplicity, that \( n = 1 \).

### 2.3.1. Existing particle approximations and problematics

**Method 1:** A first idea is to use the results for McKean–Vlasov equations with deterministic initial conditions. For \( a \in \Theta \), we can approximate the expectation \( \mathbb{E}(h(X_0(a), a) f(X_T(a))) \) under the integral in the second term of (2.3.1) by computing a mean over interacting particles. The latter are simulated by replacing the unknown law \( \mathcal{L}(X_i(a)) \) appearing in the coefficient \( u_b(t, x, a) \) by the empirical law of the particle system. The convergence of the empirical law to \( \mathcal{L}(X_i(a)) \) is known as the propagation of chaos (see [17] or [20]). Once we have done this, we evaluate in turn the integral in \( a \) with a Monte-Carlo approximation.

The corresponding approximation error can be fairly obtained and is in \( 1/\sqrt{N_1} + 1/\sqrt{N_2} + \Delta t \), where \( N_1 \) is the number of simulated \( \theta_i \) (Monte-Carlo step), and \( N_2 \) is the number of particles computed for each realization of \( \theta \) (particle approximation step). This rate is the one of the Central Limit Theorem, which is the best we can hope for with stochastic methods. However, as pointed out in [22], there is a two-step imbricated simulation procedure which is numerically very expensive. Vaillant ([24], section 3.7 p. 79) showed that its complexity is of order \( O(N_1N_2^2/\Delta t) \).

**Methods 2 and 3:** In order to avoid imbricated simulations, we follow [22] and approximate directly \( \mathbb{E}(h(X_0(\theta), \theta) f(X_T(\theta))) \) in (2.3.1) by computing a mean over \( N \) interacting particles \( (\theta_i, \tilde{X}^{i,N}(\theta_i))_{i \in [1,N]} \), whose laws are expected to be close to \( \mathcal{L}(\theta, X(\theta)) \). Since

\[
u_b(t_k, x, a) = \mathbb{E}\left(h(X_0(\theta), \theta) b(x, X_{t_k}(\theta)) \right| \theta = a),
\]

(we will choose for \( (t, x) \mapsto \mathbb{E}(h(X_0(\theta), \theta) b(x, X_{t_k}(\theta))) \) a continuous modification of the conditional expectation process) Eq. (2.2.4) can be rewritten as: \( P(d\omega)-a.s., \forall t \in [0, T] \),

\[
\begin{align*}
\mathrm{d}X_t(\theta) &= \mathbb{E}(h(X_0(\theta), \theta) b(x, X_T(\theta)))|_{x=X_t(\theta)} \mathrm{d}t + \sigma \mathrm{d}W_t \\
\mathcal{L}(X_0(\theta), \theta) &= p_0(x, a) \mathrm{d}x \nu_G(da).
\end{align*}
\]

The particles \( \tilde{X}^{i,N}_T(\theta_i) \) are then obtained by replacing the unknown drift coefficient (2.3.2) by a regression estimator \( \tilde{u}_b(t, x, a) \) (instead of an empirical mean in the case where the conditional expectation is an expectation). For \( i \in [1, N] \), we simulate a random variable \( \theta_i \) of law \( \nu_G(da) \) and associate to it a single particle, defined by its initial condition \( \tilde{X}^{i,N}_0(\theta_i) \) of law \( p_0(x, \theta_i) \mathrm{d}x \) and by its path: \( \forall k \in [0, K - 1] \),

\[
\tilde{X}^{i,N}_{t_{k+1}}(\theta_i) = \tilde{X}^{i,N}_{t_k}(\theta_i) + \tilde{u}_b\left(t_k, \tilde{X}^{i,N}_{t_k}(\theta_i), \theta_i\right) \Delta t + \sigma \left(W^{i}_{t_{k+1}} - W^{i}_{t_k}\right).
\]

The approximation of \( \langle I(m_T), f \rangle \) is then:

\[
\langle I(m_T), f \rangle \simeq \frac{1}{N} \sum_{i=1}^{N} h\left(\tilde{X}^{i,N}_M(\theta_i), \theta_i\right) f\left(\tilde{X}^{i,N}_T(\theta_i)\right).
\]

The two following methods use this approach, but with different regression estimators.
Method 2: The particle method with random weights, introduced by [22] in the case of a probability density initial condition, is based on the use of Nadaraya–Watson regression estimators to compute $\hat{u}_b$: \( \forall i \in [1, N], \forall k \in [0, K], \)

\[
\hat{u}_b(t_k, \tilde{x}^{i,N}_k(\theta_i), \theta_i) = \sum_{j=1}^{N} \frac{H_N(\theta_i - \theta_j)}{\sum_{i=1}^{N} \frac{H_N(\theta_i - \theta_j)}{H_N(\theta_i - \theta_l)}} h \left( \tilde{x}^{j,N}_0(\theta_j), \theta_j \right) b \left( \tilde{x}^{i,N}_k(\theta_i), \tilde{x}^{j,N}_k(\theta_j) \right)
\]

where \( H_N = H(./h_N)/h_N \) for a given window \( h_N \) and a given Parzen–Rosenblatt kernel \( H \) on \( \mathbb{R} \) (see Bosq and Lecoutre [3]. A possible choice for \( H \) is the standard Gaussian density function).

The accuracy of this method is computed in [22,24], and gives a rate of order \( N^{-1/8} \). This is not as accurate as with Method 1, because the convergence rates of non-parametric regression estimators are slower than the ones of empirical means. However, simulations are faster and less demanding in memory. The complexity of the algorithm is in \( O\left(N^2/\Delta t\right) \).

2.3.2. Method 3: The wavelet particle approximation

We now propose a third method, based on (2.3.4) with the choice of wavelet regression estimators for $\hat{u}_b$. We first introduce and comment the assumptions under which we will work; then, we present the wavelet regression estimators that we will use.

Let us introduce the mapping:

\[
\Phi : a \in \mathbb{R} \mapsto p_0(., a) \in L^1(\mathbb{R}, \mathbb{R}) \cap L^\infty(\mathbb{R}, \mathbb{R}),
\]

(2.3.6)

where the probability density \( p_0 \) is associated with \( v_0 \) by (2.2.2). Notice that by Assumption 1.1, the distribution function \( G \) of \( \theta \) defines a bijection from the interior of \( \Theta \) into \( ]0, 1[ \).

Assumption 2.12. 1. The law \( v_G(da) \) of \( \theta \) fulfills Assumption 1.1 and satisfies a Logarithmic Sobolev Inequality on \( \mathbb{R} \): \( \exists c_0 > 0, \forall f \in C^1(\mathbb{R}, \mathbb{R}), \)

\[
\langle v_G, f^2 \log f^2 \rangle - \langle v_G, f^2 \rangle \log(\langle v_G, f^2 \rangle) \leq c_0 \langle v_G, |\nabla f|^2 \rangle.
\]

2. Assumption 2.5 are satisfied, and additionally, \( \exists \varepsilon \in ]0, 1[ \), \( b \in C^1 + \varepsilon(\mathbb{R}, \mathbb{R}), \)

3. The map \( a \in \Theta \mapsto \|v_0(., a)\|_1 \in \mathbb{R} \) is \( L_0 \)-Lipschitz continuous.

4. \( \exists L_1 > 0, \forall a_1, a_2 \in \Theta, \mathbb{P}\{(v_0(X_0(a_1), a_1) > 0) \triangle \{v_0(X_0(a_2), a_2) > 0\}) \leq L_1|a_1 - a_2|, \)

where \( A \triangle B = (A \cup B) \setminus (A \cap B) \).

5. The map \( \Phi \) defined in (2.3.6) (resp. \( \Phi \circ G^{-1} \)) is Lipschitz continuous (resp. \( s \)-Hölder continuous, with \( s > 1/2 \)) with respect to the \( L^1 \)-Vaserstein metrics (1.0.6):

\[
\exists L_2 > 0, \forall a_1, a_2 \in \Theta, \mathbb{W}_1(p_0(x, a_1)dx, p_0(x, a_2)dx) \leq L_2|a_2 - a_1|.
\]

(resp. \( \forall a_1, a_2 \in [0, 1], \mathbb{W}_1(p_0(x, G^{-1}(a_1))dx, p_0(x, G^{-1}(a_2))dx) \leq L_2|a_2 - a_1|^s \).

6. \( \exists \xi_0 > 0, \exists C_2 > 0, \forall \xi > \xi_0, \int_{\mathbb{R}} \int_{\mathbb{R}} 1_{|x|>|\alpha|} p_0(x, a)dx \log(da) \leq C_1e^{-C_2 \xi}. \)

7. \( \phi \) and \( \psi \) are a father and a mother wavelets, generating a Multi-Resolution Analysis (MRA, see [9]), compactly supported and Lipschitz continuous. They thus belong to \( L^2(\mathbb{R}, \mathbb{R}) \cap L^1(\mathbb{R}, \mathbb{R}) \), and satisfy: \( \int (1 + |\alpha|) |\phi(\alpha)|d\alpha < \infty \), and \( \int (1 + |\alpha|) |\psi(\alpha)|d\alpha < \infty \).

Points 3, 4 and 5 state that the initial measure \( v_0(x, a)dx \) varies regularly with the parameter \( a \in \Theta \), in terms of total variation norm, sign and mass repartitions. Point 3 is trivially satisfied when \( \|v_0(., a)\|_1 \) is independent of \( a \). Point 4 is straightforward when \( P(d\omega) \)-a.s., \( v_0(., \theta) \) is a nonnegative function. These points are satisfied for probability densities.
Let us briefly recall some facts about wavelets (for an introduction, see for instance Härdle et al. [9]). We define the descendants of $\psi$ by

$$\forall \alpha \in \mathbb{R}, \; \psi_I(\alpha) = 2^{h/2} \psi(2^h \alpha - I_2)$$

where $I = (I_1, I_2)$ is a double index with $I_1 \in \mathbb{N}$ and $I_2 \in \mathbb{Z}$. To simplify notation, $\phi$ is often written $\psi_{-1,0}$, and we define for any $I_2 \in \mathbb{Z}$,

$$\forall \alpha \in \mathbb{R}, \; \psi_{-1I_2}(\alpha) = \phi(\alpha - I_2).$$

(2.3.8)

The following properties will be very useful. A proof is given in [23].

**Proposition 2.13.** Let $(\psi_I)_I$ be a MRA on $\mathbb{R}$ generated by compactly supported wavelets.

(i) $\forall \alpha \in \mathbb{R}, \; \forall I_1 \geq -1, \; \text{card}\{I_2 \in \mathbb{Z} \mid \alpha \in \text{supp}(\psi_{I_1,I_2})\} < \infty$ and does not depend on $I_1$.

(ii) Let $\mathcal{I}$ be a compact interval of $\mathbb{R}$. $\forall I_1 \geq -1, \; \text{card}\{I_2 \in \mathbb{Z} \mid \text{supp}(\psi_{I_1,I_2}) \cap \mathcal{I} \neq \emptyset\} \leq C 2^{|I_1|}$ where $C$ does not depend on $I_1$.

Recall that our aim is to approximate the conditional expectation (2.3.2) defining $u_b$. From $(A_1)$ and $(A_2)$, the map $\alpha \in [0,1] \mapsto u_b(t_k, x, G^{-1}(\alpha))$ is a bounded map, which can also be considered as a bounded map on $\mathbb{R}$ with support in $[0,1]$. It belongs to $L^2(\mathbb{R})$ and a wavelet expansion on the MRA $(\psi_I)_I$ is thus available for this function:

$$\forall \alpha \in [0,1], \; \forall k \in [0, K], \; \forall x \in \mathbb{R}, \; u_b(t_k, x, G^{-1}(\alpha)) = \sum_{I_1=-1}^{+\infty} \sum_{I_2 \in \mathbb{Z}} \beta^{(k)}(I_1)(x) \psi_{I_1}(\alpha).$$

This is equivalent to the expansion of $a \in \mathbb{R} \mapsto u_b(t, x, a)$ on the warped wavelet basis $(\psi_I \circ G)_I$:

$$\forall a \in \Theta, \; \forall k \in [0, K], \; \forall x \in \mathbb{R}, \; u_b(t_k, x, a) = \sum_{I_1=-1}^{+\infty} \sum_{I_2 \in \mathbb{Z}} \beta^{(k)}(I_1)(x) \psi_{I_1}(G(a)).$$

(2.3.9)

The coefficients $\beta^{(k)}(I_1)(x)$ can be expressed with respect to these two equivalent expansions:

$$\beta^{(k)}(I_1)(x) = \int_{[0,1]} \psi_I(\alpha) u_b(t_k, x, G^{-1}(\alpha)) d\alpha = \int_{\mathbb{R}} \psi_I(G(a)) u_b(t_k, x, a) v_G(da).$$

(2.3.10)

Using (2.3.9) and (2.3.10), we propose the following regression estimator $\hat{u}_b$ to approximate $u_b$. The sum in $I_1$ is truncated (the sum in $I_2$ is finite for every given $I_1$ from Proposition 2.13), and the integrals are replaced with their empirical counterparts: $\forall a \in \Theta, \; \forall k \in [0, K], \; \forall x \in \mathbb{R},$

$$\hat{u}_b(t_k, x, a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \hat{\beta}^{(k)}(I_1)(x) \psi_{I_1}(G_N(a)),$$

(2.3.11)

with:

$$\hat{\beta}^{(k)}(I_1)(x) = \sum_{j=1}^{N} \frac{1}{N} \psi_I(G_N(\theta_j)) h \left( \tilde{X}^{j,N}_0(\theta_j), \theta_j \right) b \left( x, \tilde{X}^{j,N}_k(\theta_j) \right),$$

(2.3.12)

where $I_1^N \in \mathbb{N}^*$ is the resolution level chosen such that:

$$2^{I_1^N} \sim N^{1/(2x+1)},$$

(2.3.13)

and where $G_N$ is the empirical distribution function of $(\theta_i)_{i \in [1, N]}$. 
Expression (2.3.12) is simpler when we rank the couples \((\theta_j, \tilde{X}^{j,N}(\theta_j))_{j \in [1,N]}\) in increasing order of \((\theta_j)_{j \in [1,N]}\). The ranked couples are denoted by \((\theta_{(j)}, \tilde{X}^{(j),N}(\theta_{(j)}))_{j \in [1,N]}\) and:

\[
\tilde{\beta}^{(k)}_I(x) = \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) h \left( \tilde{X}^{(j),N}_{0}(\theta_{(j)}), \theta_{(j)} \right) b \left( x, \tilde{X}^{(j),N}_{h_k}(\theta_{(j)}) \right). \tag{2.3.14}
\]

**Remark 2.14.** In this work, we have chosen to replace \(G\) by its empirical counterpart \(G_N\) in (2.3.11) and (2.3.14). This allows us to implement the numerical scheme even in cases where the distribution function \(G\) is unknown (for instance when the simulation procedure for the \(\theta_i\) is complicated, or when the initial condition is obtained from the data). There are also numerical advantages to using \(G_N\). (2.3.14) corresponds to the wavelet coefficient estimator of a regression computed on an equi-spaced design (i.e. the family \((\theta_j)_{j \in [1,N]}\) is replaced by the regular grid \((j/N)_{j \in [1,N]}\), which is often already implemented in statistical software.

When we plug the regression estimators (2.3.11) into the particle system defined in (2.3.4), with \(x = \tilde{X}^{i,N}_{h_k}(\theta_i)\) and \(a = \theta_i\), we obtain the particle system on which Method 3 is based:

**Definition 2.15.** Let \((\theta_i)_{i \in [1,N]}\) be i.i.d. variables of law \(v_G(da)\), let \((\tilde{X}^{i,N}_{0}(\theta_i))_{i \in [1,N]}\) be independent random variables of laws \(p_0(x, \theta_i)dx\), and let \(W = (W^1, \ldots, W^N)\) be a \(N\)-dimensional Brownian motion; all these variables being independent of each other. We define the following particle system with \(\tilde{\beta}^{(k)}_I(x)\) as in (2.3.12); \(\forall i \in [1,N], \forall k \in [0, K - 1]\),

\[
\tilde{X}^{i,N}_{h_{k+1}}(\theta_i) = \tilde{X}^{i,N}_{h_k}(\theta_i) + \sum_{I_1 = -1}^{I_N} \sum_{I_2 \in \mathbb{Z}} \tilde{\beta}^{(k)}_I(\tilde{X}^{i,N}_{h_k}(\theta_i)) \psi_I \left( G_N(\theta_i) \right) \Delta t + \sigma \left( W_{h_{k+1}}^i - W_{h_k}^i \right). \tag{2.3.15}
\]

**Theorem 2.16.** Under Assumption 2.12, \(\forall 0 < \eta < 1, \forall f \in \mathcal{C}^{4+\eta}_b(\mathbb{R}), \exists N_0 \in \mathbb{N}^*, C > 0, \forall N \geq N_0:\)

\[
\mathbb{E} \left( \left| I(m_T), f \right| - \frac{1}{N} \sum_{i=1}^{N} h \left( \tilde{X}^{i,N}_{0}(\theta_i), \theta_i \right) f \left( \tilde{X}^{i,N}_{T}(\theta_i) \right) \right) \leq C \left( \Delta t + \sqrt{\log N} N^{-\frac{s}{2s + 1}} \right), \tag{2.3.16}
\]

where \(s\) is defined in Point 5 of Assumption 2.12.

Notice first that we have released some of the assumptions needed in [22]. In [22], the law \(v_G(da)\) is assumed to be absolutely continuous with respect to the Lebesgue measure on \(\mathbb{R}\), with a strictly positive Lipschitz continuous density supported by a compact interval.

The asymptotic convergence rate of Method 3, while still slower than the one of Method 1, is better than the one obtained for Method 2. If \(\Phi \circ G^{-1}\) is Lipschitz continuous \((s = 1)\), we have a rate slightly slower than \(N^{-1/3}\), and in the worst case, when \(s = 1/2\), we obtain a rate of order \(N^{-1/4}\), which is still more accurate than the preceding \(N^{-1/8}\).

The optimal choice of window \(h_N\) in Method 2 usually depends on the knowledge of the law of \(\tilde{X}_i(\theta)\), and is often hard to obtain. This is not the case for the resolution level \(I_1^N\) (2.3.13), which depends only on \(s\).

If we use the Mallat cascade algorithm, with a complexity of order \(O(N)\) (see [9], Chapter 12, p. 223, or Mallat [13] Chapter VII, sections 7.3 and 7.5), the complexity of Method 3 remains in \(O(N^2/\Delta t)\), which is comparable with Method 2.
3. Convergence rate of the particle approximation

In this section, we will prove Theorem 2.16. We work under Assumption 2.12. The Euler scheme associated with SDE (2.2.4) is given by: \( \forall k \in [0, K - 1], \)
\[
\bar{X}_{tk+1}(\theta) = \bar{X}_{tk}(\theta) + u_b(tk, \bar{X}_{tk}(\theta), \theta)\Delta t + \sigma(W_{tk+1} - W_{tk}), \quad \mathcal{L}(\bar{X}_0(\theta)) = p_0(x, \theta)dx.
\]
(3.0.17)

We consider as well i.i.d. copies of this Euler scheme coupled with the particles defined in (2.3.15) (same parameters \( (\theta_i)_{i \in [1, N]} \), same initial conditions \( (\bar{X}_{0i}^{N}(\theta_i))_{i \in [1, N]} \) and same Brownian motions \( (W^i)_{i \in [1, N]} \); \( \forall i \in [1, N], \forall k \in [0, K - 1], \)
\[
\bar{X}_{tk}^{i}(\theta_i) = \bar{X}_{tk}^{i}(\theta_i) + u_b(tk, \bar{X}_{tk}^{i}(\theta_i), \theta_i)\Delta t + \sigma(W_{tk+1}^i - W_{tk}^i), \quad \bar{X}_{0}^{i} = \bar{X}_{0}^{i,N}. \quad (3.0.18)
\]

We introduce the analogue of the coefficient estimator \( \hat{\beta}_I^{(k)}(x) \) (2.3.14) that is computed on the nonlinear independent particles (3.0.18): \( \forall k \in [0, K], \forall x \in \mathbb{R}, \forall I \in [-1, I_1^N] \times \mathbb{Z}, \)
\[
\hat{\tilde{\beta}}_I^{(k)}(x) = \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) h(\bar{X}_{0}^{(j)}(\theta(j)), \theta(j))b(x, \bar{X}_{tk}^{(j)}(\theta(j))), \quad (3.0.19)
\]
and the associated regression estimator: \( \forall a \in \Theta, \forall k \in [0, K], \forall x \in \mathbb{R}, \)
\[
\bar{u}_b(tk, x, a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \hat{\beta}_I^{(k)}(x) \psi_I(G_N(a)). \quad (3.0.20)
\]

In the sequel, we will also need the wavelet decomposition of the function \( a \in \Theta \mapsto \bar{u}_b(tk, x, a) \) := \( \mathbb{E}(h(\bar{X}_0(\theta), \theta)b(x, \bar{X}_tk(\theta))|\theta = a) \) on the warped wavelet basis \( (\psi_I \circ G)_I \):
\[
\forall a \in \Theta, \forall k \in [0, K], \forall x \in \mathbb{R}, \quad \bar{u}_b(tk, x, a) = \sum_{I_1=-1}^{+\infty} \sum_{I_2 \in \mathbb{Z}} \hat{\tilde{\beta}}_I^{(k)}(x) \psi_I(G(a)), \quad (3.0.21)
\]
with:
\[
\hat{\tilde{\beta}}_I^{(k)}(x) = \mathbb{E} \left( \psi_I(G(a))h(\bar{X}_0(\theta), \theta)b(x, \bar{X}_tk(\theta)) | \theta = a \right) v_G(da) \quad v_G(da) = \mathbb{E} \left( \psi_I(G(\theta))h(\bar{X}_0(\theta), \theta)b(x, \bar{X}_tk(\theta)) \right). \quad (3.0.22)
\]

**Proposition 3.1.** Let \( t \in [0, T], k \in [0, K] \) and \( x \in \mathbb{R} \). We have the following assertions, with constants independent from \( x \).

(i) The map \( a \in \Theta \mapsto u_b(t, x, a) \) (resp. \( a \in [0, 1] \mapsto u_b(t, x, G^{-1}(a)) \)) is Lipschitz-continuous (resp. \( s \)-Hölder continuous, \( s > 1/2 \) being defined in Point 5 of Assumption 2.12).

(ii) The map \( a \in \Theta \mapsto \bar{u}_b(tk, x, a) \) (resp. \( a \in [0, 1] \mapsto \bar{u}_b(tk, x, G^{-1}(a)) \)) is Lipschitz-continuous (resp. \( s \)-Hölder continuous).

**Proof.** We only study the Lipschitz continuities; the \( s \)-Hölder continuities are obtained in the same manner. We first prove Point (i). Let \( a_1, a_2 \in \Theta \), let \( (W_t)_{t \in [0, T]} \) be a Brownian motion, let \( X_0(a_1) \) and \( X_0(a_2) \) be two random variables of laws \( p_0(y, a_1)dy \) and \( p_0(y, a_2)dy \) and let...
(X_t(a_1))_{t \in [0,T]}, (X_t(a_2))_{t \in [0,T]} be the associated solutions of SDEs (2.2.6). From (2.3.2):

\[ |u_b(t, x, a_1) - u_b(t, x, a_2)| \leq A \|b\|_{Lip} \mathbb{E}(|X_t(a_1) - X_t(a_2)|) + \|b\|_\infty \mathbb{E}(|h(X_0(a_1), a_1) - h(X_0(a_2), a_2)|). \]  

(3.0.23)

For the second term on the right hand side of (3.0.23), we have from (2.2.1):

\[ \mathbb{E}(|h(X_0(a_1), a_1) - h(X_0(a_2), a_2)|) \leq A_1 + A_2 + A_3 + A_4, \]

with:

\[ A_1 = \mathbb{E}(1_{v_0(X(a_1), a_1) > 0} |v_0(., a_1)| \|v_0(., a_2)|), \]

\[ A_2 = \mathbb{E}(1_{v_0(X(a_1), a_1) > 0} - 1_{v_0(X(a_2), a_2) > 0} |v_0(., a_2)|). \]

We define A_3 and A_4 as A_1 and A_2 by replacing 1_{v_0(X(a_1), a_1) > 0} by 1_{v_0(X(a_1), a_1) \leq 0} for i \in \{1, 2\}. By Point 3 of Assumption 2.12, A_1 \leq L_0|a_1 - a_2|, and A_3 \leq L_0|a_1 - a_2|. By Point 4:

\[ A_2 = A_4 = |\mathbb{P}(v_0(X(a_1), a_1) > 0)\Delta v_0(X(a_2), a_2) > 0)|\|v_0(., a_2)| \leq L_1A|a_1 - a_2|. \]

This gives:

\[ \mathbb{E}(|h(X_0(a_1), a_1) - h(X_0(a_2), a_2)|) \leq 2(L_0 + L_1A)|a_1 - a_2|. \]  

(3.0.24)

Let us now consider the first term in the right hand side of (3.0.23).

\[ \mathbb{E}(|X_t(a_1) - X_t(a_2)|) \leq \mathbb{E}(|X_0(a_1) - X_0(a_2)|) + \int_0^t (B_1(s) + B_2(s) + B_3(s))ds \]

\[ B_1(s) = \mathbb{E}(\mathbb{E}(h(X_t(a_1), a_1)b(x_s, X_s(a_1)))|x = X_s(a_1) - h(X_t(a_1), a_1)b(x_s, X_s(a_1)))|x = X_s(a_2)) \]

\[ B_2(s) = \mathbb{E}(\mathbb{E}(h(X_t(a_1), a_1)b(x_s, X_s(a_1)))|x = X_s(a_2) - h(X_t(a_1), a_1)b(x_s, X_s(a_2)))|x = X_s(a_2)) \]

\[ B_3(s) = \mathbb{E}(\mathbb{E}(h(X_t(a_1), a_1)b(x_s, X_s(a_2)))|x = X_s(a_2) - h(X_t(a_2), a_2)b(x_s, X_s(a_2)))|x = X_s(a_2)). \]

From (A_1) and (A_3), B_1(s) and B_2(s) are upper bounded by A \|b\|_{Lip} \mathbb{E}(|X_s(a_1) - X_s(a_2)|). From (3.0.24), B_3(s) \leq 2\|b\|_\infty (L_0 + L_1A)|a_1 - a_2|. It follows by Gronwall’s Lemma that:

\[ \mathbb{E}(|X_t(a_1) - X_t(a_2)|) \leq (\mathbb{E}(|X_0(a_1) - X_0(a_2)|) + 2\|b\|_\infty (L_0 + L_1A)|a_1 - a_2|)e^{2A\|b\|_{Lip}T}. \]  

(3.0.25)

No assumption has been made on the joint law of (X_0(a_1), X_0(a_2)) yet, and we can take the infimum on every possible joint laws with marginals p_0(y, a_1)dy and p_0(y, a_2)dy, and thus replace \mathbb{E}(|X_0(a_1) - X_0(a_2)|) by \mathbb{V}_1(p_0(y, a_1)dy, p_0(y, a_2)dy) in (3.0.25). From (3.0.23)–(3.0.25) and Point 5 of Assumption 2.12:

\[ |u_b(t, x, a_1) - u_b(t, x, a_2)| \leq C(A, T, L_0, L_1, L_2)|a_1 - a_2|, \]

(3.0.26)

where the constant C(A, T, L_0, L_1, L_2) does not depend on x.

Let us now turn to the proof of (ii). We consider the two Euler schemes (\tilde{X}_k(a_1))_{k \in [0,K]} and (\tilde{X}_k(a_2))_{k \in [0,K]}, defined as in (3.0.17) with the same Brownian motion and the same
initial conditions as \((X_t(a_1))_{t\in[0,T]}\) and \((X_t(a_2))_{t\in[0,T]}\) introduced in the first part of the proof. Proceeding as in (3.0.23) and using (3.0.24), we are lead to:

\[
|\bar{u}_b(t_k, x, a_1) - \bar{u}_b(t_k, x, a_2)| \leq A\|b\|_{\text{Lip}} \mathbb{E}\left(|\bar{X}_{t_k}(a_1) - \bar{X}_{t_k}(a_2)|\right) \\
+ 2\|b\|_{\infty}(L_0 + L_1A)|a_1 - a_2|.
\] (3.0.27)

For \(k \in [0, K - 1]\):

\[
\mathbb{E}\left(|\bar{X}_{t_{k+1}}(a_1) - \bar{X}_{t_{k+1}}(a_2)|\right) \leq \mathbb{E}\left(|\bar{X}_{t_k}(a_1) - \bar{X}_{t_k}(a_2)|\right) \\
+ \mathbb{E}\left(|u_b(t_k, \bar{X}_{t_k}(a_1), a_1) - u_b(t_k, \bar{X}_{t_k}(a_1), a_2)|\right) \Delta t \\
+ \mathbb{E}\left(|u_b(t_k, \bar{X}_{t_k}(a_1), a_2) - u_b(t_k, \bar{X}_{t_k}(a_2), a_2)|\right) \Delta t \\
\leq (1 + \Delta tA\|b\|_{\text{Lip}}) \mathbb{E}\left(|\bar{X}_{t_k}(a_1) - \bar{X}_{t_k}(a_2)|\right) \\
+ \Delta tC(A, T, L_0, L_1, L_2)|a_1 - a_2|,
\]

where \(C(A, T, L_0, L_1, L_2)\) is the constant of (3.0.26). Since: \(\forall C > 0,\)

\[
(1 + C\Delta t)^{T/\Delta t} = \exp\left(\frac{T}{\Delta t} \log(1 + C\Delta t)\right) \leq \exp\left(\frac{T}{\Delta t}C\Delta t\right) = e^{CT},
\] (3.0.28)

we obtain by induction, and with our choice of initial conditions \((X_0(a_1), X_0(a_2))\):

\[
\mathbb{E}\left(|\bar{X}_{t_k}(a_1) - \bar{X}_{t_k}(a_2)|\right) \leq C(T, A)|a_1 - a_2|.
\] (3.0.29)

The Lipschitz continuity of \(a \mapsto \bar{u}_b(t_k, x, a)\) follows from (3.0.27). ■

**Remark 3.2.** From now on, we will write \(\bar{X}_{t_k}^{i,N}, \bar{X}_{t_k}^{i,N}, \bar{X}_{t_k}^{i,\bar{N}}\), and \(\bar{X}_{t_k}^{i,N}\) instead of \(\bar{X}_{t_k}(\theta_1), \bar{X}_{t_k}(\theta_2), \bar{X}_{t_k}^{i,\bar{N}}(\theta_1)\), \(\bar{X}_{t_k}(\theta_2)\), and \(\bar{X}_{t_k}^{i,N}(\theta_1)\).

**Proof of Theorem 2.16.** We can decompose the approximation error at time \(T\) into three sources:

\[
(I(m_T), f) - \frac{1}{N} \sum_{i=1}^{N} h(\bar{X}_0^{i,N}, \theta_i) f\left(\bar{X}_T^{i,N}\right) = T_1 + T_2 + T_3,
\] (3.0.30)

\[
T_1 = (I(m_T), f) - \mathbb{E}\left(h(\bar{X}_0(\theta), \theta) f(\tilde{X}_T(\theta))\right)
\]

\[
T_2 = \mathbb{E}\left(h(\bar{X}_0(\theta), \theta) f(\tilde{X}_T(\theta))\right) - \frac{1}{N} \sum_{i=1}^{N} h(\bar{X}_0^{i}, \theta_i) f\left(\bar{X}_T^{i}\right)
\]

\[
T_3 = \frac{1}{N} \sum_{i=1}^{N} h(\bar{X}_0^{i}, \theta_i) f\left(\bar{X}_T^{i}\right) - \frac{1}{N} \sum_{i=1}^{N} h(\bar{X}_0^{i,N}, \theta_i) f\left(\bar{X}_T^{i,N}\right).
\]

From (2.3.1), and choosing \(\bar{X}_0(\theta) = X_0(\theta)\), it appears that the discretization error \(T_1\) can be upper bounded by a result due to Talay and Vaillant ([22], Proposition 5.1), which generalizes a result from Talay and Tubaro [21]: \(\forall f \in C_b^{4+\varepsilon}(\mathbb{R}, \mathbb{R}), \exists C = C(T, f, b, \sigma) > 0,\)

\[
|\mathbb{E}\left(f(\bar{X}_T(\theta))\right) - \mathbb{E}\left(f(\tilde{X}_T(\theta))\right)| \leq C A \Delta t.
\] (3.0.31)

The statistical error \(T_2\) can be upper bounded thanks to the Central Limit Theorem: \(\exists C > 0,\)

\[
\mathbb{E}\left(|\mathbb{E}\left(h(\bar{X}_0(\theta), \theta) f(\tilde{X}_T(\theta))\right) - \frac{1}{N} \sum_{i=1}^{N} h(\bar{X}_0^{i}, \theta_i) f\left(\bar{X}_T^{i}\right)|\right) \leq \frac{CA\|f\|_{\infty}}{\sqrt{N}}.
\] (3.0.32)
Now let us focus on the Term $T_3$. As the particles $\tilde{X}^{i,N}$ are not in mean-field interactions, the classical propagation of chaos is not available. Since $\forall i \in [1, N]$, $\tilde{X}^{i,N}_0 = \bar{X}^{i}_0$:

$$|T_3| \leq \frac{A\|f\|_{\operatorname{Lip}}}{N} \sum_{i=1}^{N} \left| \bar{X}^{i}_T - \bar{X}^{i,N}_T \right|.$$ 

From (2.3.15) and (3.0.18), we have: $\forall k \in [0, K - 1]$,

$$S_N(t_{k+1}) := \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( |\bar{X}^{i}_k - \bar{X}^{i,N}_k| \right)$$

$$\leq S_N(t_k) + \frac{\Delta t}{N} \sum_{i=1}^{N} \mathbb{E} \left( |u_b(t_k, \bar{X}^{i}_k, \theta_i) - \bar{u}_b(t_k, \bar{X}^{i,N}_k, \theta_i)| \right)$$

$$\leq S_N(t_k) + \Delta t \left( B_1(k, N) + B_2(k, N) + B_3(k, N) \right), \quad (3.0.33)$$

$$B_1(k, N) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( |u_b(t_k, \bar{X}^{i}_k, \theta_i) - \bar{u}_b(t_k, \bar{X}^{i}_k, \theta_i)| \right)$$

$$B_2(k, N) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( |\bar{u}_b(t_k, \bar{X}^{i}_k, \theta_i) - \bar{u}_b(t_k, \bar{X}^{i,N}_k, \theta_i)| \right)$$

$$B_3(k, N) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( |\bar{u}_b(t_k, \bar{X}^{i}_k, \theta_i) - \bar{u}_b(t_k, \bar{X}^{i,N}_k, \theta_i)| \right).$$

The term $B_1(k, N)$ can be handled in a way similar to (3.0.31), and $B_1(k, N) \leq C A \Delta t$.

The term $B_2(k, N)$ looks like a regression error, except that the function being regressed, $a \mapsto \bar{u}_b(t_k, \bar{X}^{i}_k, a)$ depends on the random parameter $\bar{X}^{i}_k$.

**Lemma 3.3. Under Assumption 2.12, \( \exists \gamma > 0, \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall K > 0, \forall k \in [0, K], \)**

$$\mathbb{E} \left( |\bar{u}_b(t_k, \bar{X}^{i}_k, \theta_i) - \bar{u}_b(t_k, \bar{X}^{i,N}_k, \theta_i)| \right) \leq C \sqrt{\log N} N^{-\gamma + \frac{\kappa}{2}}. \quad (3.0.34)$$

It is in this part of the error that the regularity $s$ of $a \mapsto \bar{u}_b(t_k, x, G^{-1}(a))$ appears. This was expected, since this regularity can be translated in terms of properties of the wavelet coefficients. Recall that under Point 7 of Assumption 2.12, the wavelet coefficients $\beta_{I_1 I_2}$ of any $s$-Hölder continuous function satisfy (Theorem 9.6 p. 121 in [9]):

$$\exists C > 0, \forall I_1 \geq -1, \sup_{I_2 \in \mathbb{Z}} |\beta_{I_1 I_2}| \leq C 2^{-I_1 \left( s + \frac{1}{2} \right)}. \quad (3.0.35)$$

The proof of Lemma 3.3 (in Section 3.3) is based on Lemma 3.4 (proved in Section 3.2).

**Lemma 3.4. Under Assumption 2.12, the coefficients $\tilde{\beta}^{(k)}_I(x)$ and $\bar{\hat{\beta}}^{(k)}_I(x)$, defined in (3.0.19) and (3.0.22), satisfy: $\forall \gamma > 0, \exists \kappa = \kappa(\gamma, c_0, L_0, b, \sigma, \phi, \psi), \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall K > 0, \forall k \in [0, K], \)**

$$\mathbb{P} \left( |\tilde{\beta}^{(k)}_I(\bar{X}^{i}_k) | - |\bar{\hat{\beta}}^{(k)}_I(\bar{X}^{i}_k)| \geq \kappa \frac{\sqrt{\log N}}{N} \right) \leq \frac{C}{N^\gamma}. \quad (3.0.36)$$
In the proof of Lemma 3.3, we will use the result of Lemma 3.4 with \( \gamma > 3/2 \). The constant \( \kappa \) that should be chosen in Lemma 3.3 is given by \( \kappa(3/2, c_0, L_0, b, \sigma, \phi, \psi) \) of Lemma 3.4 (see Eq. (3.2.22) in the proof for the precise condition).

An upper bound for \( B_3(k, N) \) is given by the following lemma, proved in Section 3.1:

**Lemma 3.5.** Under Assumption 2.12, \( \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall K > 0, \forall k \in [0, K], \)

\[
\frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left| \frac{1}{N} \sum_{l_1=-1}^{I_1} \sum_{l_2 \in \mathbb{Z}} \left( \tilde{\beta}^{(k)}(\tilde{X}_{l_1}^i) - \tilde{\beta}^{(k)}(\tilde{X}_{l_1}^{i,N}) \right) \psi_I \left( \frac{i}{N} \right) \right| \right) \leq \frac{C}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left| \tilde{X}_{l_1}^i - \tilde{X}_{l_1}^{i,N} \right| \right).
\]

**End of the proof of Theorem 2.16.** We conclude the computation of an upper bound on \( T_3 \):

\[
S_N(t_{k+1}) \leq (1 + C \Delta t) S_N(t_k) + C \Delta t \left( \Delta t + \sqrt{\log N} N^{-\frac{\gamma}{1+\gamma}} \right)
\]

by induction, by using (3.0.28), and since \( S_N(0) = 0 \). Thus:

\[
\sup_{k \in [1, T/\Delta t]} \mathbb{E} \left( \left| \frac{1}{N} \sum_{i=1}^{N} h \left( \tilde{X}_{l_0}^i, \theta_i \right) \left( f \left( \tilde{X}_{l_1}^i \right) - f \left( \tilde{X}_{l_1}^{i,N} \right) \right) \right| \right) \leq C(T, A) \| f \|_{\text{Lip}} \left( \Delta t + \sqrt{\log N} N^{-\frac{\gamma}{1+\gamma}} \right).
\]

Gathering (3.0.31), (3.0.32) and (3.0.37) gives the result announced in Theorem 2.16. ■

The Sections 3.1–3.3 are devoted to the proofs of Lemmas 3.5, 3.4 and 3.3 respectively.

### 3.1. Proof of Lemma 3.5

We first provide a technical lemma. The difficulty is that the suprema \( 2^{l/2} \| \psi \|_{\infty} \) of the functions \( \psi_I \) cannot be uniformly upper bounded when the resolution level \( I_1 \) increases.

**Lemma 3.6.** Assume that the father and mother wavelets \( \phi \) and \( \psi \) are Lipschitz continuous functions with compact support. Then, for a choice of \( I_1^N \) as in (2.3.13):

\[
\exists C > 0, \exists N_0 \in \mathbb{N}^*, \forall \alpha \in [0, 1], \forall N \geq N_0, \quad \frac{1}{N} \sum_{j=1}^{N} \left| \sum_{l_1=-1}^{I_1^N} \sum_{l_2 \in \mathbb{Z}} \psi_I(\alpha) \psi_I \left( \frac{j}{N} \right) \right| \leq C.
\]

**Proof.** Let \( \phi_{I_1^N I_2}(x) = 2^{l_1^N/2} \phi(2^{l_1^N} x - l_2) \), for \( I_2 \in \mathbb{Z} \), be the descendants of the wavelet father of Point 7 of Assumption 2.12. We can write \( \frac{1}{N} \sum_{j=1}^{N} \left| \sum_{l_1=-1}^{I_1^N} \sum_{l_2 \in \mathbb{Z}} \psi_I(\alpha) \psi_I \left( \frac{j}{N} \right) \right| \leq T_1 + T_2,
\]

\[
T_1 = \int_{0}^{1} \left| \sum_{l_1=-1}^{I_1^N} \sum_{l_2 \in \mathbb{Z}} \psi_I(\alpha) \psi_I \left( \frac{j}{N} \right) \right| dy = \int_{0}^{1} \left| \sum_{l_2 \in \mathbb{Z}} \phi_{I_1^N I_2}(\alpha) \phi_{I_1^N I_2} \left( \frac{j}{N} \right) \right| dy
\]

\[
\leq \int_{0}^{2^{l_1^N}} \sum_{l_2 \in \mathbb{Z}} |\phi(2^{l_1^N} x - l_2)\phi(x - l_2)| dx \leq C \| \phi \|_{\infty} \| \phi \|_{1},
\]

where \( C \) is a constant.
by using the fact that the orthogonal projection kernels associated with \((\psi_I)_{I_1 \in [-1, I_1^N]}, I_2 \in \mathbb{Z}\) and \((\phi_T^N, I_2)_{I_2 \in \mathbb{Z}}\) coincide (see [9], p. 117–118), and that the chosen wavelets have compact support.

\[
T_2 = \sum_{j=1}^{N} \int_{(j-1)/N}^{j/N} \left| \sum_{I_1=1}^{I_1^N} \sum_{I_2} \psi_I(\alpha) \left[ \psi_I(y) - \psi_I\left(\frac{j}{N}\right) \right] \right| dy
\]

\[
\leq \sum_{j=1}^{N} \sum_{I_1=1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} |\psi_{I_1, I_2}(\alpha)| \cdot 2^{3l/2} L_{\psi} \int_{(j-1)/N}^{j/N} \left( \frac{j}{N} - y \right) dy
\]

\[
+ \sum_{j=1}^{N} \sum_{I_2 \in \mathbb{Z}} |\phi(\alpha - I_2)| L_{\phi} \int_{(j-1)/N}^{j/N} \left( \frac{j}{N} - y \right) dy
\]

\[
\leq \sum_{j=1}^{N} \left( \sum_{I_1=0}^{I_1^N} \frac{C \|\psi\|_\infty L_{\psi} 2^{l_1}}{2N^2} \right) + \sum_{j=1}^{N} \frac{C \|\phi\|_\infty L_{\phi}}{2N^2}
\]

\[
\leq \frac{C \|\psi\|_\infty L_{\psi} 2^{l_1^N}}{2N} + \frac{C \|\phi\|_\infty L_{\phi}}{2N}.
\]

From the choice of \(I_1^N (2.3.13)\), we have \(2^{l_1^N} / N \to 0\), and hence \(T_2 \to 0\) when \(N \to \infty\).  

Proof of Lemma 3.5. Since \(\forall j \in [1, N], \bar{X}_j = \bar{X}_j^N\):

\[
\frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left| \tilde{u}_b(t_k, \bar{X}_k^i, \theta), \tilde{u}_b(t_k, \bar{X}_k^i, \theta) \right| \right)
\]

\[
\leq \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left( \sum_{I_1=1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \frac{1}{N} \sum_{j=1}^{N} \psi_I\left(\frac{j}{N}\right) h\left(\bar{X}_0^i, \theta(j)\right) \right) \left[ b\left(\bar{X}_k^i, \bar{X}_k^i\right) - b\left(\bar{X}_k^i, \bar{X}_k^i, \bar{X}_k^i, \theta(j)\right) \psi_I\left(\frac{i}{N}\right) \right] \right)
\]

\[
\leq \mathbb{E} \left( \frac{A \|b\|_{\text{Lip}}}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{I} \psi_I\left(\frac{i}{N}\right) \psi_I\left(\frac{j}{N}\right) \left( \left| \bar{X}_k^i - \bar{X}_k^i, \bar{X}_k^i \right| + \left| \bar{X}_k^i - \bar{X}_k^i, \bar{X}_k^i \right| \right) \right)
\]

\[
\leq \frac{2A \|b\|_{\text{Lip}} C}{N} \sum_{i=1}^{N} \mathbb{E} \left( \left| \bar{X}_k^i - \bar{X}_k^i \right| \right)
\]

and the Lemma is proved.  

3.2. Proof of Lemma 3.4

Let \(N, K \in \mathbb{N}^\ast\). We have, for all \(k \in [0, K]\), \(i \in [1, N]\), and \(I \in [-1, I_1^N] \times \mathbb{Z}\):

\[
\mathbb{P} \left( |\tilde{\beta}_I^{(k)}(X_k^i) - \tilde{\beta}_I^{(k)}(X_k^i)| > \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right) \leq C_1(I, k, i, \kappa, N) + C_2(I, k, i, \kappa, N)
\]

(3.2.1)
\[ C_1(I, k, i, \kappa, N) = \mathbb{P} \left( \left| \tilde{\beta}_I^{(k)}(\tilde{X}_t^i) - \mathbb{E} \left( \tilde{\beta}_I^{(k)}(x) \right) \right|_{x=\tilde{X}_t^i} > \frac{\kappa}{4} \frac{\log N}{N} \right) \]

\[ C_2(I, k, i, \kappa, N) = \mathbb{P} \left( \left| \mathbb{E} \left( \tilde{\beta}_I^{(k)}(x) \right) \right|_{x=\tilde{X}_t^i} - \tilde{\beta}_I^{(k)}(\tilde{X}_t^i) > \frac{\kappa}{4} \frac{\log N}{N} \right). \]

A difficulty in upper bounding \( C_1(I, k, i, \kappa, N) \) is that we deal with the wavelet coefficients of a function \( \alpha \in [0, 1] \mapsto \tilde{u}_p(t_k, \tilde{X}_t^i, G^{-1}(\alpha)) \) that depends on a random parameter \( x = \tilde{X}_t^i \).

We first fix \( x \in \mathbb{R} \) and prove Lemma 3.7 (Section 3.2.1). Then, we use a localization argument to obtain the upper bound of \( C_1(I, k, i, \kappa, N) \) announced in Corollary 3.8 (Section 3.2.2).

**Lemma 3.7.** Under Assumption 2.12 \( \exists N_0, C > 0, \forall N \geq N_0, \forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \forall x \in \mathbb{R}, \forall r > 0, \)

\[ \mathbb{P} \left( \left| \tilde{\beta}_I^{(k)}(x) - \mathbb{E} \left( \tilde{\beta}_I^{(k)}(x) \right) \right| > r \right) \leq 2 \exp \left( -\frac{r^2}{8} \left( \frac{4A^2\|b\|_G^2C^2}{N} + \frac{A\|b\|_G\|\psi\|_2^2|2^{i/2}r|}{3N} \right) \right) \]

\[ + Ce^{-\frac{Nr^2}{8\|b\|_G^2C^2|c|_0}} \quad (3.2.2) \]

where \( C' \) is a positive constant such that

\[ \forall N \geq N_0, \quad \sqrt{\frac{1}{N} \sum_{j=1}^{N} \psi_j^2 \left( \frac{j}{N} \right)} < C' \quad (3.2.3) \]

(This constant exists, since \( \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} \psi_j^2 \left( \frac{j}{N} \right) = \int_0^1 \psi_0^2(x)dx \leq 1 \).)

**Corollary 3.8.** Under Assumption 2.12, \( \forall \gamma > 0, \exists \kappa = \kappa(\gamma, b, \sigma, c_0, L_0, \phi, \psi) > 0, \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \)

\[ C_1(I, k, i, \kappa, N) \leq C/N^\gamma. \quad (3.2.4) \]

To upper bound \( C_2(I, k, i, \kappa, N), \) we study the difference:

\[ \left| \mathbb{E} \left( \tilde{\beta}_I^{(k)}(x) \right) \right|_{x=\tilde{X}_t^i} - \tilde{\beta}_I^{(k)}(\tilde{X}_t^i) \]

\[ = \left| \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \left( \psi_j \left( G_N(\theta_j) \right) - \psi_j \left( G(\theta_j) \right) \right) h(\tilde{X}_o^j, \theta_j)b(x, \tilde{X}_t^j) \right) \right|_{x=\tilde{X}_t^i}. \quad (3.2.5) \]

The intuitive idea of using the Lipschitz continuity of \( \psi \) does not lead to the result, since the Lipschitz constant \( 2^{3i/2}L_\psi \) is not counterbalanced by \( \mathbb{E}\|G_N - G\|_\infty \). We follow in Section 3.2.3 some ideas in Kerkyacharian and Picard [11] and use more deeply the structure of wavelets:

**Lemma 3.9.** Under Assumption 2.12: \( \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \)

\[ \left| \mathbb{E} \left( \tilde{\beta}_I^{(k)}(x) \right) \right|_{x=\tilde{X}_t^i} - \tilde{\beta}_I^{(k)}(\tilde{X}_t^i) \leq C/\sqrt{N} \quad \text{implying} \quad C_2(I, k, i, \kappa, N) = 0, \quad (3.2.6) \]

with \( \kappa \) as in Corollary 3.8.
The result announced in Lemma 3.4 is obtained from (3.2.1), Corollary 3.8 and Lemma 3.9.

3.2.1. Proof of Lemma 3.7

A difficulty comes from the fact that the terms in the sum defining $\tilde{\beta}_I^{(k)}(x)$ (3.0.19) are not independent. These terms are, however, independent conditionally to $(\theta_1, \ldots, \theta_N)$.

$$\Pr(|\tilde{\beta}_I^{(k)}(x) - \mathbb{E}(\tilde{\beta}_I^{(k)}(x))| > r) \leq D_1(I, k, x, r) + D_2(I, k, x, r),$$

(3.2.7)

$$D_1(I, k, x, r) = \Pr\left(\left|\tilde{\beta}_I^{(k)}(x) - \mathbb{E}(\tilde{\beta}_I^{(k)}(x) \mid \theta_1, \ldots, \theta_N)\right| \geq \frac{r}{2}\right)$$

(3.2.8)

$$D_2(I, k, x, r) = \mathbb{E}\left[\left|\mathbb{E}(\tilde{\beta}_I^{(k)}(x) \mid \theta_1, \ldots, \theta_N) - \mathbb{E}(\tilde{\beta}_I^{(k)}(x))\right| \geq \frac{r}{2}\right].$$

Upper bounds for $D_1(I, k, x)$ and $D_2(I, k, x)$ are given by Lemmas 3.10 and 3.11.

Lemma 3.10. Under Assumption 2.12: $\exists N_0$, $C > 0$, $\forall N \geq N_0$, $\forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall K > 0$, $\forall k \in [0, K]$, $\forall x \in \mathbb{R}$, $\forall r > 0$,

$$D_1(I, k, x, r) \leq 2 \exp\left(-\frac{r^2}{8\left(\frac{4A^2C^2\|b\|_\infty^2}{N} + A\|b\|_\infty\|\psi\|_\infty 2^{1/2}r\right)}\right).$$

(3.2.9)

**Proof.** Since:

$$\mathbb{E}(\tilde{\beta}_I^{(k)}(x) \mid \theta_1, \ldots, \theta_N) = \frac{1}{N} \sum_{j=1}^{N} \psi_I\left(\frac{j}{N}\right)$$

$$\times \mathbb{E}\left(h(\tilde{X}_0^{(j)}, \theta(j))b(x, \tilde{X}_k^{(j)}) \mid \theta_1, \ldots, \theta_N\right)$$

$$= \frac{1}{N} \sum_{j=1}^{N} \psi_I\left(\frac{j}{N}\right) \mathbb{E}\left(h(\tilde{X}_0^{(j)}, \theta(j))b(x, \tilde{X}_k^{(j)}(\theta(j))) \mid \theta(j)\right)$$

$$= \frac{1}{N} \sum_{j=1}^{N} \psi_I\left(\frac{j}{N}\right) \mathbb{E}\left(h(\tilde{X}_0(\theta), \theta)b(x, \tilde{X}_k(\theta)) \mid \theta = \theta(j)\right),$$

we have:

$$\tilde{\beta}_I^{(k)}(x) - \mathbb{E}(\tilde{\beta}_I^{(k)}(x) \mid \theta_1, \ldots, \theta_N) = \frac{1}{N} \sum_{j=1}^{N} \psi_I\left(\frac{j}{N}\right)$$

$$\times \left(h(\tilde{X}_0^{(j)}, \theta(j))b(x, \tilde{X}_k^{(j)}(\theta(j))) - \mathbb{E}\left(h(\tilde{X}_0, \theta)b(x, \tilde{X}_k(\theta)) \mid \theta = \theta(j)\right)\right),$$

(3.2.10)

which, conditionally to $(\theta_1, \ldots, \theta_N)$, is a sum of independent centered variables which are upper bounded by $2^{1/2}\|\psi\|_\infty \times 2A\|b\|_\infty / N$, and such that for sufficiently large $N$:

$$\sum_{j=1}^{N} \mathbb{E}\left[\left|\frac{1}{N} \psi_I\left(\frac{j}{N}\right) \left(h(\tilde{X}_0^{(j)}, \theta(j))b(x, \tilde{X}_k^{(j)}(\theta(j))) - \mathbb{E}\left(h(\tilde{X}_0(\theta), \theta)b(x, \tilde{X}_k(\theta)) \mid \theta = \theta(j)\right)\right)\right|^2 \mid \theta_1, \ldots, \theta_N\right]$$
Proposition 3.1. Let \( \psi \) be a normalized function and \( b \) be a fixed parameter. Then, the common law \( v_m \) satisfies an Inequality formulated as

\[
\int |x|^2 v_m(dx) \leq \frac{4A^2}{N} \sum_{j=1}^{N} \left( \frac{j}{N} \right)^2 \| b \|^2_{\infty} \leq \frac{4A^2 C^2 \| b \|^2_{\infty}}{N},
\]

where \( A^2 := \sum_{j=1}^{N} \left( \frac{j}{N} \right)^2 \). (C’ is defined in (3.2.3)). Bernstein’s inequality (see for instance [9] p. 241) yields (3.2.9). \( \bbox{\hspace{1em}} \)

**Lemma 3.11.** Under Assumption 2.12: \( \exists N_0, C > 0, \forall N \geq N_0, \forall I = (I_1, I_2) \in [-1, 1]^N \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \forall x \in \mathbb{R}, \forall r > 0,

\[
D_2(I, k, x, r) \leq C \exp \left( -\frac{N r^2}{8 \| \bar{u}_b \|^2_{\text{Lip}} C^2 c_0} \right).
\]

**Proof.** Let \( I \in [-1, 1]^N \times \mathbb{Z}, k \in [0, K] \) and \( x \in \mathbb{R} \). Recall that:

\[
\mathbb{E} \left( \beta_I^{(k)}(x) \mid \theta_1, \ldots, \theta_N \right) - \mathbb{E} \left( \bar{\beta}_I^{(k)}(x) \right) = \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) \left( \bar{u}_b(t_k, x, \theta(j)) - \mathbb{E} \left( \bar{u}_b(t_k, x, \theta(j)) \right) \right).
\]

Let us define the following function:

\[
F_{I, k, x} : \mathbb{R}^N \to \mathbb{R} \quad (a_1, \ldots, a_N) \mapsto \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) \bar{u}_b(t_k, x, a(j)),
\]

where \( (a(j))_{j \in [1, N]} \) are the terms \( (a_i)_{i \in [1, N]} \) ranked in increasing order. By Proposition 3.1, the map \( a \mapsto \bar{u}_b(t_k, x, a) \) is Lipschitz continuous with a constant that does not depend on \( k \) nor on \( x \). Thus, \( \forall (a_1, \ldots, a_N), (a'_1, \ldots, a'_N) \in \mathbb{R}^N:

\[
|F_{I, k, x}(a_1, \ldots, a_N) - F_{I, k, x}(a'_1, \ldots, a'_N)| \leq \frac{\| \bar{u}_b \|_{\text{Lip}}}{N} \sum_{j=1}^{N} \left| \psi_I \left( \frac{j}{N} \right) \right| |a(j) - a'_j| \leq \frac{\| \bar{u}_b \|_{\text{Lip}} C'}{\sqrt{N}} \sqrt{\sum_{j=1}^{N} |a_j - a'_j|^2},
\]

by the Cauchy–Schwarz inequality, with \( C' \) defined in (3.2.3).

By Point 2 of Assumption 2.12, the common law \( v_G(da) \) of the random variables \( (\theta_i)_{i \in [1, N]} \) satisfies a Logarithmic–Sobolev inequality with constant \( c_0 \). By independent tensorization (see Ané et al. [1], Theorem 3.2.3 p. 31), the law \( v^{\otimes N}_G \) of \( (\theta_i)_{i \in [1, N]} \) also satisfies a Logarithmic-Sobolev inequality with the same constant \( c_0 \). Applying the inequality of concentration of the measure ([1], Theorem 6.4.1 p. 74) to \( v^{\otimes N}_G \) and to the function \( F_{I, k, x} \) gives (3.2.11). \( \bbox{\hspace{1em}} \)

### 3.2.2. Proof of Corollary 3.8

Our purpose is to prove (3.2.4) by replacing in Inequality (3.2.2) the fixed parameter \( x \) with the random position \( X_k^i \), and by choosing \( r = \kappa/4 \sqrt{\log N/N} \) with a proper choice of \( \kappa \). We will use a localization argument, and thus need some tail bounds for the law of \( X_k^i \).
Lemma 3.12. Under Point 6 of Assumption 2.12, \( \exists \varsigma_0 > 0, \forall \varsigma \geq \varsigma_0, \exists C_1, C_2 > 0, \forall N \in \mathbb{N}^*, \forall i \in [1, N], \forall k \in \mathbb{N}^*, \forall k \in [0, K], \)

\[
P \left( |\tilde{X}_{t_k}^i| > \varsigma \right) \leq C_1 e^{-C_2 \varsigma^2/2} + \exp \left( -\frac{\varsigma^2/2 - A\|b\|_{\infty} T}{2\sigma^2 T} \right). \tag{3.2.13}
\]

**Proof.** We have, by Point 6 of Assumption 2.12:

\[
P \left( |\tilde{X}_{t_k}^i| > \varsigma \right) \leq P \left( |\tilde{X}_{0}^i| > \frac{\varsigma}{2} \right) + P \left( |\tilde{X}_{0}^i| \leq \frac{\varsigma}{2}, |\tilde{X}_{t_k}^i| > \varsigma \right) \leq C_1 e^{-C_2 \varsigma^2/2} + \exp \left( -\frac{\varsigma^2/2 - A\|b\|_{\infty} T}{2\sigma^2 T} \right).
\]

Let \( \lambda \in \mathbb{R} \). Since \( u_b \) is bounded by \( A\|b\|_{\infty} \), we obtain by induction:

\[
E(e^{\lambda (\tilde{X}_{t_k}^i - \tilde{X}_{0}^i)}) = E[e^{\lambda (\tilde{X}_{t_k-1} - \tilde{X}_{0}^i)}] 
\times E(e^{\lambda (u_b(\theta_{t_k-1} - \tilde{X}_{t_k-1}^i, \theta_{t_k})) \Delta t + \sigma \sqrt{\Delta t} (W_{t_k}^i - W_{t_k}^{i-1}))}) | \tilde{X}_{t_k-1}^i, \tilde{X}_{0}^i]) \leq E(e^{\lambda A\|b\|_{\infty} \Delta t + \frac{\varsigma^2 \Delta t}{2}}) \leq e^{\lambda A\|b\|_{\infty} T + \frac{\varsigma^2 T}{2}}.
\]

Using Markov’s Inequality and optimizing the bound in \( \lambda \) gives the result for \( \varsigma > 2A\|b\|_{\infty} T \). \( \blacksquare \)

**Proof of Corollary 3.8.** We now use the announced localization argument. Let us define:

\[
C_3(x) = \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( G_N(\theta_j) \right) h \left( \tilde{X}_{0}^j, \theta_j \right) b \left( x, \tilde{X}_{t_k}^j \right) \right| - E \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( G_N(\theta_j) \right) h \left( \tilde{X}_{0}^j, \theta_j \right) b \left( x, \tilde{X}_{t_k}^j \right) \right).
\]

Let \( r > 0 \) and \( K = [-\varsigma, \varsigma] \) with \( \varsigma > 2A\|b\|_{\infty} T \). Since \( K \) is a compact interval, it can be covered with a finite number of balls \( |x_l - \varrho, x_j + \varrho|, \) with \( (x_l)_{l \in [1, \tilde{\ell}]} \) a finite sequence of \( K \) and \( \varrho > 0 \). It is possible to choose \( \tilde{\ell} = \lfloor \varsigma/\varrho \rfloor + 1 \), where \( \lfloor . \rfloor \) stands for the integer part. We have:

\[
P \left( C_3(\tilde{X}_{t_k}^i) > r \right) \leq \sum_{l=1}^{\tilde{\ell}} P \left( C_3(\tilde{X}_{t_k}^i) > r \text{ and } |\tilde{X}_{t_k}^i - x_l| \leq \varrho \right) + P \left( \tilde{X}_{t_k}^i \notin K \right). \tag{3.2.14}
\]

For a given \( l \in [1, \tilde{\ell}] \):

\[
C_3(\tilde{X}_{t_k}^i) \leq \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( G_N(\theta_j) \right) h \left( \tilde{X}_{0}^j, \theta_j \right) \left( b \left( \tilde{X}_{t_k}^j, \tilde{X}_{t_k}^i \right) - b \left( x_l, \tilde{X}_{t_k}^j \right) \right) \right| + C_3(x_l) + \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( G_N(\theta_j) \right) h \left( \tilde{X}_{0}^j, \theta_j \right) \left( b \left( x_l, \tilde{X}_{t_k}^j \right) - b \left( x, \tilde{X}_{t_k}^j \right) \right) \right|_{x=\tilde{X}_{t_k}^j} \leq 2AC'\|b\|_{\text{Lip}} |x_l - \tilde{X}_{t_k}^i| + C_3(x_l), \tag{3.2.15}
\]

where $C'$ is defined in (3.2.3). Thus:

$$
\Pr \left( C_3(\bar{x}_t^i) > r \text{ and } |\bar{x}_t^i - x_t| \leq \varrho \right) \leq \Pr \left( C_3(x_t) > r - 2AC'\|b\|_{\text{Lip}} \varrho \right)
$$

$$
\leq 2 \exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8(4A^2\|b\|_\infty^2 C' + (r - 2AC'\|b\|_{\text{Lip}} \varrho)A\|b\|_\infty \|\psi\|_\infty 2^{l_1/2}/3) \right)
$$

$$
+ C \exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8 C'\|\bar{u}_b\|_{\text{Lip}}^2 c_0} \right),
$$

by Lemma 3.7. Thus, from (3.2.14):

$$
\Pr \left( C_3(\bar{x}_t^i) > r \right) \leq \left( \left\lceil \frac{\varsigma}{\varrho} \right\rceil + 1 \right) \times 2 \exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8(4A^2\|b\|_\infty^2 C' + (r - 2AC'\|b\|_{\text{Lip}} \varrho)A\|b\|_\infty \|\psi\|_\infty 2^{l_1/2}/3) \right)
$$

$$
+ C \exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8 C'\|\bar{u}_b\|_{\text{Lip}}^2 c_0} \right) + C_1 \exp (-C_2\varsigma/2) + \exp \left( -\frac{(\varsigma/2 - A\|b\|_\infty T)^2}{2\sigma^2 T} \right). \tag{3.2.16}
$$

We now choose $\varrho = r/(4AC'\|b\|_{\text{Lip}})$, $r = \kappa/4\sqrt{\log N/N}$, and upper bound the different terms in the right hand side of (3.2.16).

- For the first parentheses, we can find a constant $C_3 > 0$ such that:

$$
\left\lceil \frac{\varsigma}{\varrho} \right\rceil + 1 \leq \frac{16\varsigma C' A\|b\|_{\text{Lip}} \kappa}{\log N} \leq C_3 \varsigma \sqrt{\frac{N}{\log N}}. \tag{3.2.17}
$$

- By choice of $l_1^N$, $r 2^{l_1/2} \leq 2r 2^{l_1/2} \leq \kappa \sqrt{\log N N^{-1/4}/4} \leq \kappa/4$ for sufficiently large $N$, and:

$$
\exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8(4A^2\|b\|_\infty^2 C' + (r - 2AC'\|b\|_{\text{Lip}} \varrho)A\|b\|_\infty \|\psi\|_\infty 2^{l_1/2}/3) \right)
$$

$$
\leq \exp \left( -\frac{N(\kappa/8 \times \sqrt{\log N/N})^2}{8(4A^2\|b\|_\infty^2 C' + \kappa A\|b\|_\infty \|\psi\|_\infty /24) \right) \leq \exp \left( -\frac{\kappa^2 \log N}{512(4A^2\|b\|_\infty^2 C' + \kappa A\|b\|_\infty \|\psi\|_\infty /24) \right). \tag{3.2.18}
$$

- For the second exponential term,

$$
\exp \left( -\frac{N(r - 2AC'\|b\|_{\text{Lip}} \varrho)^2}{8 C'\|\bar{u}_b\|_{\text{Lip}}^2 c_0} \right) \leq \exp \left( -\frac{\kappa^2 \log N}{512 C'\|\bar{u}_b\|_{\text{Lip}}^2 c_0} \right). \tag{3.2.19}
$$

- When $\varsigma > 2(A\|b\|_\infty T + 1)$, there exists $C_4, C_5 > 0$ such that:

$$
C_1 \exp (-C_2\varsigma/2) + \exp \left( -\frac{(\varsigma/2 - A\|b\|_\infty T)^2}{2\sigma^2 T} \right)
$$
\[ \leq C_1 \exp\left(-C_2 \xi/2\right) + \exp\left(-\frac{\xi/2 - A\|b\|_{\infty}T}{2\sigma^2 T}\right) \]
\[ \leq \left(C_1 + \exp\left(\frac{A\|b\|_{\infty}T}{2\sigma^2 T}\right)\right) \exp\left(-\frac{\xi}{2} \min\left(C_2, \frac{1}{2\sigma^2 T}\right)\right) \]
\[ \leq C_4 \exp\left(-C_5 \xi/2\right). \]  

(3.2.20)

From (3.2.16)–(3.2.20):
\[ \mathbb{P}\left(C_3(\tilde{X}_i^j) > r\right) \leq C_6 \xi N^{-\tilde{\gamma}(\kappa)} + C_4 \exp\left(-C_5 \xi/2\right). \]  

(3.2.21)

with
\[ \tilde{\gamma}(\kappa) = \left(\frac{\kappa^2}{512(4A^2\|b\|_{\infty}^2 C^2 + \kappa A\|b\|_{\infty} \|\psi_{\infty}/24\|^2)} \right)^\wedge \left(\frac{\kappa^2}{512 C^2 \|\tilde{b}\|_{\text{Lip}}^4 c_0}\right) - \frac{1}{2}. \]

Let us show that the right hand side of (3.2.21) can be upper bounded by a term of order \( N^{-\gamma} \). We introduce \( f(\xi) = C_6 \xi N^{-\tilde{\gamma}(\kappa)} + C_4 \exp\left(-C_5 \xi/2\right) \). Its derivative in \( \xi \) vanishes for \( \xi_0 = (2/C_5) \ln\left(\left(\frac{C_4 C_5 N^{\tilde{\gamma}(\kappa)}}{2 C_6}\right)\right) \), which is equivalent to \((2\tilde{\gamma}(\kappa)/C_5) \log N \) when \( N \to +\infty \), and is thus greater than \( 2(4\|b\|_{\infty} T + 1) \) for sufficiently large \( N \).

For \( \xi_0 \):
\[ f(\xi) \leq \left(\frac{2C_6}{C_5} \ln\left(\frac{C_4 C_5}{2 C_6}\right) + \frac{2C_6}{C_5} \tilde{\gamma}(\kappa) \right) \ln(N) + \frac{2C_6}{C_5} \right) N^{-\tilde{\gamma}(\kappa)}. \]

If we choose \( \kappa \) such that:
\[ \left(\frac{\kappa^2}{512(4A^2\|b\|_{\infty}^2 C^2 + \kappa A\|b\|_{\infty} \|\psi_{\infty}/24\|^2)} \right) > \gamma + \frac{1}{2} \]  

(3.2.22)

we have \( \tilde{\gamma}(\kappa) > \gamma \) and \( \exists C > 0, \exists N_0 \in \mathbb{N}^*, \forall N \geq N_0, f(\xi) \leq C_7/N^\gamma \), which proves the lemma. A sufficient condition is that \( \kappa \) fulfills the following conditions:
\[ \kappa > C' \|\tilde{b}\|_{\text{Lip}} \sqrt{512c_0 \left(\gamma + \frac{1}{2}\right)} \]
\[ \kappa > \left[\left(\gamma + \frac{1}{2}\right) 512 \left(4A^2\|b\|_{\infty}^2 C^2 + \frac{A\|b\|_{\infty} \|\psi_{\infty}\|}{24}\right)\right]. \]

3.2.3. Proof of Lemma 3.9

To prove Lemma 3.9, we cannot use the Lipschitz continuity of the function \( \psi_I \). We rather take advantage of the regularity of \( \alpha \mapsto \tilde{u}_b(I, x, G^{-1}(\alpha)) \) by an integration by part-like formula. To this purpose, let us recall that if \( \psi \) is a wavelet with compact support that satisfies the moment properties of Point 8 of Assumption 2.12, then there exists a compactly supported Lipschitz continuous function \( \Psi \) such that: \( \psi = \Delta_{-h} (\Psi) = \Psi(-h) - \Psi(.) \), with \( h = 2^{-1} \) (see [11]). Thus:
\[ \psi_I = \Delta_{-h_I} (\Psi_I), \quad \text{with } h_I = 2^{-I_1-1} \text{ and } \Psi_I(y) = 2^{I_1/2} \psi(2^{I_1}y - I_2). \]

Let us introduce the following notation: for \( \alpha \in [0, 1], \)
\[ U_N(\alpha) = \frac{1}{N} \sum_{i=1}^{N} 1_{[-\infty, \alpha]}(G(\theta_i))(= G_N(G^{-1}(\alpha))). \]
Lemma 3.9 can then be rewritten as
\[ U_N^{(-j)}(\alpha) = \frac{1}{N} \sum_{i,j=1}^{N} 1_{[-\infty, \alpha]}(G(\theta_i)) = U_N(\alpha) - \frac{1_{\{G(\theta_j)\leq \alpha < 1\}}}{N}. \]

Recall the Dvoretzky–Kiefer–Wolfowitz inequality (see [306]):
\[ \exists C > 0, \forall r > 0, \quad \mathbb{P} \left( \sup_{\alpha \in [0,1]} |U_N(\alpha) - \alpha| \geq r \right) \leq C e^{-2Nr^2}. \]  
(3.2.23)

Integrating this inequality in \( r \) gives:
\[ \exists C > 0, \quad \mathbb{E} \left( \sup_{\alpha \in [0,1]} |U_N(\alpha) - \alpha| \right) \leq C \sqrt{\frac{\pi}{2N}}. \]  
(3.2.24)

In the sequel, we will use the following set:
\[ B_N(\varsigma) = \left\{ \sup_{\alpha \in [0,1]} |U_N(\alpha) - \alpha| \geq \varsigma \sqrt{\frac{\log N}{N}} \right\}, \]  
(3.2.25)

where the constant \( \varsigma \) has to be chosen (our choice will be \( \varsigma > \sqrt{3/8} \)). Using (3.2.23) yields:
\[ \mathbb{P}(B_N(\varsigma)) \leq C/N^2\varsigma^2. \]  
(3.2.26)

We are now ready to prove Lemma 3.9. From (3.2.5), we have:
\[ E := \tilde{\beta}_l^{(k)}(\tilde{X}_t^i) - \mathbb{E}(\tilde{\beta}_l^{(k)}(x))|_{x=\tilde{X}_t^i} = \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \left[ \Delta_{-h_{\bar{I}}} (\Psi_I(U_N(G(\theta_j))) - \Delta_{-h_{\bar{I}}} (\Psi_I(G(\theta_j))) \right] \right. \]
\[ \times h \left( \tilde{X}_0^j, \theta_j \right) b(x, \tilde{X}_t^j(\theta_j)) \left|_{x=\tilde{X}_t^i} \right. \].

We add and subtract the terms \( \Psi_I(U_N^{(-j)}(G(\theta_j))) \) and:
\[ \Psi_I(U_N^{(-j)}(G(\theta_j) - h_{\bar{I}})1_{h_{\bar{I}} \leq G(\theta_j) \leq 1} + \Psi_I(G(\theta_j) - h_{\bar{I}})1_{0 \leq G(\theta_j) \leq h_{\bar{I}}} \]  
(3.2.27)

to \( \Delta_{-h_{\bar{I}}} (\Psi_I(U_N(G(\theta_j))) - \Delta_{-h_{\bar{I}}} (\Psi_I(G(\theta_j))) \). By convention, let us define
\[ U_N^{(-j)}(\alpha) = \alpha \quad \text{for} \quad \alpha < 0. \]  
(3.2.28)

The term (3.2.27) can then be rewritten as \( \Psi_I(U_N^{(-j)}(G(\theta_j) - h_{\bar{I}})) \), and:
\[ \Delta_{-h_{\bar{I}}} (\Psi_I(U_N(G(\theta_j))) - \Delta_{-h_{\bar{I}}} (\Psi_I(G(\theta_j))) \]
\[ = \left( \Psi_I(U_N^{(-j)}(G(\theta_j) - h_{\bar{I}})) - \Psi_I(G(\theta_j) - h_{\bar{I}}) \right) \]
\[ - \left( \Psi_I(U_N^{(-j)}(G(\theta_j))) - \Psi_I(G(\theta_j)) \right) \]
\[ + \Psi_I \left( U_N^{(-j)}(G(\theta_j)) + \frac{1}{N} - h_{\bar{I}} \right) - \Psi_I(U_N^{(-j)}(G(\theta_j) - h_{\bar{I}})) \]
+ \Psi_I(U_N^{(-j)}(G(\theta_j))) - \Psi_I\left(U_N^{(-j)}(G(\theta_j)) + \frac{1}{N}\right).

Thus: $E = E_1(I, k, i) + E_2(I, k, i)$, with:

$$E_1(I, k, i) = \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}(\mathbb{E}(\Delta_{-h_1}\left(\Psi_I(U_N^{(-j)}(\cdot)) - \Psi_I(\cdot))(G(\theta_j))
\times \bar{u}_b(t_k, x, \theta_j)|\theta_i, i \neq j)|_{x=\bar{\chi}^i_k}
= \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\left(\int_0^1 \Delta_{-h_1}\left((\Psi_I(U_N^{(-j)}(\cdot)) - \Psi_I(\cdot))(\alpha)\bar{u}_b(t_k, x, G^{-1}(\alpha))\right)d\alpha\right)|_{x=\bar{\chi}^i_k}
$$

by independence of the $((\theta_j))_{j \in [1, N]}$, and:

$$E_2(I, k, i) = \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha) + \frac{1}{N} - h_1\right) - \Psi_I(U_N^{(-j)}(\alpha - h_1))\right)\bar{u}_b(t_k, x, G^{-1}(\alpha))d\alpha\right)|_{x=\bar{\chi}^i_k}
+ \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I\left(U_N^{(-j)}(\alpha) + \frac{1}{N}\right)\right)\bar{u}_b(t_k, x, G^{-1}(\alpha))d\alpha\right)|_{x=\bar{\chi}^i_k}.
(3.2.29)
$$

The term $E_1(I, k, i)$ allows us to carry out the integration by parts mentioned previously. From Proposition 3.1 and $s > 1/2$, we obtain:

$$|E_1(I, k, i)|
= \left| \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right)\Delta_{h_1}\bar{u}_b(t_k, x, G^{-1}(\cdot)(\alpha))d\alpha\right)\right|_{x=\bar{\chi}^i_k}
\leq C\frac{\sqrt{h_1}}{N} \sum_{j=1}^{N} \mathbb{E}\left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right|d\alpha\right)
\leq C\frac{\sqrt{h_1}}{N} \sum_{j=1}^{N} (E_{11}^{(j)} + E_{12}^{(j)}),
(3.2.30)
$$

$$E_{11}^{(j)} = \mathbb{E}\left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right|d\alpha 1_{B_N(\varsigma)}\right) \leq 2 \times 2^{h_1/2}\|\Psi\|_{\infty}\mathbb{P}(B_N(\varsigma))
\leq \frac{2 \times 2^{h_1/2}\|\Psi\|_{\infty}C}{N^2\varsigma^2},
(3.2.31)
$$

$$E_{12}^{(j)} = \mathbb{E}\left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right|d\alpha 1_{B_N^c(\varsigma)}\right).
$$
On the set $B_N^c(\zeta)$, we have
\[
\sup_{\alpha \in [0,1]} |U_N^{(-j)}(\alpha) - \alpha| \leq \sup_{\alpha \in [0,1]} |U_N(\alpha) - \alpha| + \frac{1}{N} \leq \zeta \sqrt{\log N} + \frac{1}{N}.
\]
In this case, the support of $\alpha \mapsto \Psi(U_N^{(-j)}(\alpha)) - \Psi(\alpha)$ is included in an interval of length $C(\zeta, \Psi)$ depending only on $\zeta$ and $\Psi$. For any double index $I \in [-1, I_1^N] \times \mathbb{Z}$, there hence exists an interval $I(I)$ of length $C(\zeta, \Psi)2^{-l_1}$ such that:
\[
\int_0^1 \left| \Psi_I \left( U_N^{(-j)}(\alpha) \right) - \Psi_I(\alpha) \right| d\alpha 1_{B_N^c(\zeta)} = \int_{I(I)} \left| \Psi_I \left( U_N^{(-j)}(\alpha) \right) - \Psi_I(\alpha) \right| d\alpha 1_{B_N^c(\zeta)}.
\]
Thus:
\[
E_{12}^{(j)} \leq 2^{l_1/2} L \Psi \mathbb{E} \left( \sup_{\alpha} \left| U_N^{(-j)}(\alpha) - \alpha \right| \right) C(\zeta, \Psi)2^{-l_1} \leq 2^{l_1/2} L \Psi C(\zeta, \Psi) \sqrt{2N}.
\]
From (3.2.30), (3.2.31) and (3.2.33), we deduce, with $\zeta > 1/2$, that:
\[
\exists C > 0, |E_1(I, k, i)| \leq C \left( \frac{1}{N^{2\zeta^2}} + \frac{1}{N} \right) \leq \frac{C}{N^{\zeta}}.
\]
Now we turn to the residual term $E_2(I, k, i)$ defined in (3.2.29).
\[
|E_2(I, k)| \leq \frac{A\|b\|_{\infty}}{N} \sum_{j=1}^N \left[ E_{21}^{(j)} + E_{22}^{(j)} \right],
\]
\[
E_{21}^{(j)} = \mathbb{E} \left( \int_0^1 \left| \Psi_I \left( U_N^{(-j)}(\alpha) + \frac{1}{N} - h_1 \right) - \Psi_I \left( U_N^{(-j)}(\alpha - h_1) \right) \right| d\alpha 1_{B_N(\zeta)} \right) \leq 4 \times 2^{l_1/2} \mathbb{E}(B_N(\zeta)) \leq \frac{4 \times 2^{l_1/2} \|\Psi\|_{\infty} C}{N^{2\zeta^2}}.
\]
\[
E_{22}^{(j)} = \mathbb{E} \left( \int_0^1 \left| \Psi_I \left( U_N^{(-j)}(\alpha) + \frac{1}{N} - h_1 \right) - \Psi_I \left( U_N^{(-j)}(\alpha - h_1) \right) \right| d\alpha 1_{B_N^c(\zeta)} \right) \leq \frac{4 \times 2^{l_1/2} \|\Psi\|_{\infty} C}{N^{2\zeta^2}}.
\]
Since on $B_N^c(\zeta)$:
\[
\sup_{\alpha \in [0,1]} \left| U_N^{(-j)}(\alpha) + \frac{1}{N} - h_1 - U_N^{(-j)}(\alpha - h_1) \right| \leq \sup_{\alpha \in [0,1]} \left| U_N^{(-j)}(\alpha) - \alpha \right| + \sup_{\alpha \in [0,1]} \left| \alpha - h_1 - U_N^{(-j)}(\alpha - h_1) \right| + \frac{1}{N}.
\]
there exists, similarly to (3.2.32), a positive constant $C'(\varsigma, \Psi)$ such that:

\[
E_{22}^{(j)} \leq 2^{1/2}L_{\Psi}C'(\varsigma, \Psi)2^{-I_1}
\times \sup_{\alpha \in [0,1]} \left( \mathbb{E}\left| U_N^{(-j)}(\alpha) - U_N^{(-j)}(\alpha - h_1) - \frac{N - 1}{N}h_1 \right| + \frac{h_1 + 2}{N} \right).
\]

Using the Cauchy–Schwarz inequality, then, the independence of $(\theta_i)_{i \in [1, N]}$:

\[
\mathbb{E}\left| U_N^{(-j)}(\alpha) - U_N^{(-j)}(\alpha - h_1) - \frac{N - 1}{N}h_1 \right| \\
\leq \left\{ \mathbb{E} \left[ \left( \frac{1}{N} \sum_{i=1, i \neq j}^N (1_{[\alpha - h_1, \alpha]}G(\theta_i) - h_1) \right)^2 \right] \right\}^{1/2}
\leq \sqrt{\frac{1}{N} \text{Var}(1_{[\alpha - h_1, \alpha]}G(\theta))} \leq \sqrt{\frac{h_1(1 - h_1)}{N}}.
\]

Thus:

\[
E_{22}^{(j)} \leq 2^{1/2}L_{\Psi}C'(\varsigma, \Psi) \left( \sqrt{\frac{h_1(1 - h_1)}{N}} + \frac{h_1 + 2}{N} \right) \leq C \left( \frac{1}{\sqrt{N}} + \frac{2^{i_1/2}}{N} \right). \tag{3.2.38}
\]

From (3.2.35), (3.2.36) and (3.2.38), for $I_1 \in [1, I_1^N]$ with $I_1^N$ as in (2.3.13) and $\varsigma > \sqrt{3}/8$:

\[
|E_2(I, k, i)| \leq C \left( \frac{2^{I_1^N/2}}{N^{2\varsigma^2}} + \frac{1}{\sqrt{N}} + \frac{2^{I_1^N/2}}{N} \right)
\leq CN^{-\min(2\varsigma^2 - \frac{1}{2(\varsigma+1)}, 1/2 - \frac{1}{2(\varsigma+1)})} \leq \frac{C}{\sqrt{N}}. \tag{3.2.39}
\]

From (3.2.34) and (3.2.39), we then have $|E| \leq C/\sqrt{N}$, and Lemma 3.9 is proved. This concludes the proof of Lemma 3.4. We have now all the tools to prove Lemma 3.3.

### 3.3. Proof of Lemma 3.3

Recall the definitions of $\bar{u}_b$ and $\bar{u}_b$ given in (3.0.20) and (3.0.21). We have:

\[
\mathbb{E}\left( \left| \bar{u}_b(t_k, \bar{X}_k^i, \theta_i) - \bar{u}_b(t_k, \bar{X}_k^i, \theta_i) \right| \right)
\leq \mathbb{E}\left( \sup_{\alpha \in [0,1]} \left| \bar{u}_b(t_k, \bar{X}_k^i, G^{-1}(\alpha)) - \bar{u}_b(t_k, \bar{X}_k^i, G^{-1}(\alpha)) \right| \right).
\]

We first give a lemma allowing us to control some moments.
Lemma 3.13. Under Assumption 2.12: \( \forall \gamma > 0, \exists \kappa > 0 \) as in Lemma 3.4, \( \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall I \subseteq [-1, I^N_1] \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \)

\[
\mathbb{P}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \geq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right) \leq \frac{C 2^{l_1}}{N^\gamma} \leq \frac{C}{N^{\gamma-1/(2s+1)}}. \tag{3.3.1}
\]

As a consequence, \( \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall I \subseteq [-1, I^N_1] \times \mathbb{Z}, \forall K > 0, \forall k \in [0, K], \)

\[
\forall p \in \{1, 2\}, \mathbb{E}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)|^p \right) \leq \frac{C (\log N)^{p/2}}{N^{p/2}}. \tag{3.3.2}
\]

**Proof.** From the tail upper bound of Lemma 3.4, we obtain:

\[
\mathbb{P}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \geq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right)
\]

\[
\leq \sum_{l_2 \in \mathbb{Z}} \mathbb{P}\left( |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \geq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right)
\]

\[
\leq C 2^{l_1} \mathbb{P}\left( |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \geq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right)
\]

\[
\leq \frac{C 2^{l_1}}{N^\gamma} \leq \frac{C 2^{l_1}}{N^\gamma} \leq \frac{C}{N^{\gamma-1/(2s+1)}}, \tag{3.3.3}
\]

since the number of non-zero coefficients for a given level \( l_1 \) is of order \( 2^{l_1} \) from Point (ii) of Proposition 2.13. From (3.3.3), we have for \( p \geq 1, \)

\[
\mathbb{E}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)|^p \right)
\]

\[
\leq \frac{\kappa^p (\log N)^{p/2}}{2^p N^{p/2}} \mathbb{P}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \leq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right)
\]

\[
\quad + (C' A \|b\|_{\infty})^p \mathbb{P}\left( \sup_{l_2} |\tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i) - \tilde{\beta}^{(k)}_I(\tilde{X}_{t_{k}}^i)| \geq \frac{\kappa}{2} \sqrt{\frac{\log N}{N}} \right)
\]

\[
\leq \frac{\kappa^p (\log N)^{p/2}}{2^p N^{p/2}} + (C' A \|b\|_{\infty})^p \frac{C}{N^{\gamma-1/(2s+1)}}. \tag{3.3.4}
\]

where \( C' \) has been defined in (3.2.3). If we choose \( \gamma > 3/2, \) then \( \gamma - 1/(2s+1) > \gamma - 1/2 > 1, \) since \( s > 1/2. \) For \( p \in \{1, 2\}, \) and with the appropriate choice of \( \kappa \) (see Lemma 3.4), the dominant term in (3.3.4) is then of order \( (\log N)^{p/2}/N^{p/2}. \) ■

Lemma 3.14. Under Assumption 2.12, \( \exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1, N], \forall K > 0, \forall k \in [0, K], \)

\[
\mathbb{E}\left( \sup_{\alpha \in [0, 1]} |\tilde{u}_b(t_k, \tilde{X}_{t_{k}}^i, G^{-1}(\alpha)) - \tilde{u}_b(t_k, \tilde{X}_{t_{k}}^i, G^{-1}(\alpha))| \right) \leq C \sqrt{\log N} N^{-\frac{3}{2(p+1)}}. \]
We can decompose:

\[
E \left( \sup_{a \in [0,1]} \left| \tilde{u}_b(t_k, \tilde{X}_{t_k}^i, G^{-1}(\alpha)) - \tilde{u}_b(t_k, \tilde{X}_{t_k}^i, G^{-1}(\alpha)) \right| \right) \leq C_1 + C_2 + C_3, \quad (3.3.5)
\]

\[
C_1 = E \left( \sup_{a \in [0,1]} \left| \sum_{l_1 > l_1^N} \sum_{l_2 \in \mathbb{Z}} (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i)) \psi(a) \right| \right)
\]

\[
C_2 = E \left( \sup_{a \in \Theta} \left| \sum_{l_1 = -1}^{l_1^N} \sum_{l_2 \in \mathbb{Z}} (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i)) \left( \psi(G(a)) - \psi(G(N(a))) \right) \right| \right)
\]

\[
C_3 = E \left( \sup_{a \in \Theta} \left| \sum_{l_1 = -1}^{l_1^N} \sum_{l_2 \in \mathbb{Z}} \left( (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i) - (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i))) \psi(G(a)) \right) \right| \right).
\]

Let us upper bound the term \( C_1 \).

\[
C_1 \leq C \sum_{l_1 > l_1^N} E \left( \sup_{l_2 \in \mathbb{Z}} \left| (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i)) \right| \right)^{2I_1/2} \| \psi \|_\infty \leq C \sum_{l_1 > l_1^N} 2^{I_1} \left( \frac{1}{2} - \frac{1}{2} \right) \leq CN^{-\frac{s}{2+1}},
\]

from Point (i) of Proposition 2.13 and (3.0.35). Let us now consider the term \( C_2 \). Let us define

\[
\Delta_{I,N}(a) = \psi(G(a)) - \psi(G(N(a))). \quad (3.3.6)
\]

From the law of the iterated logarithm (see Van der Vaart [7], Chapter 19, p. 268):

\[
\limsup_{N \to \infty} \sqrt{\frac{N}{2 \log \log N}} \| G - G_N \|_\infty \leq \frac{1}{2} \text{ almost surely.} \quad (3.3.7)
\]

For sufficiently large \( N \) and for every \( p \geq 1 \), we thus have:

\[
E \left( \| \Delta_{I,N} \|_\infty^p \right) \leq 2^{3I_1 p/2} L^p \left( \frac{\log \log N}{N} \right)^{p/2}. \quad (3.3.8)
\]

Since the chosen wavelets have compact support, we have:

\[
C_2 \leq E \left( C \sum_{l_1 = -1}^{l_1^N} \sup_{l_2 \in \mathbb{Z}} \left| (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i)) \right| \| \Delta_{I,N} \|_\infty \right) \leq D_1 + D_2,
\]

with:

\[
D_1 = E \left( C \sum_{l_1 = -1}^{l_1^N} \sup_{l_2 \in \mathbb{Z}} \left| (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i) - (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i))) \right| \| \Delta_{I,N} \|_\infty \right),
\]

\[
D_2 = E \left( C \sum_{l_1 = -1}^{l_1^N} \sup_{l_2 \in \mathbb{Z}} \left| (\tilde{\beta}_I^{(k)}(\tilde{X}_{t_k}^i)) \right| \| \Delta_{I,N} \|_\infty \right).
\]
Let us upper bound $D_1$. From (3.3.2) and (3.3.7):

$$D_1 \leq C \sum_{l_1=-1}^{l_1^n} \sqrt{\mathbb{E} \left( \sup_{l_2 \in \mathbb{Z}} \left| \tilde{\beta}^{(k)}_l (\tilde{X}_l^n) - \tilde{\beta}^{(k)}_l (\tilde{X}_l^n) \right|^2 \right)} \sqrt{\mathbb{E} \left( \| \Delta I, N \|_{\infty}^2 \right)}$$

$$\leq C \sqrt{\frac{\log N}{N}} \sum_{l_1=-1}^{l_1^n} 2^{3l_1/2} \sqrt{\frac{\log \log N}{N}}$$

$$\leq C \sqrt{\frac{\log N}{N}} \log \log N N^{-\frac{4s-1}{2s+1}} < CN^{-\frac{s}{2(s+1)}}, \quad (3.3.9)$$

since $(4s-1)/(2(1+2s)) > \frac{s}{2s+1}$ is equivalent to $s > 1/2$. Let us now upper bound $D_2$. Using (3.0.35) and (3.3.7):

$$D_2 \leq C \sum_{l_1=-1}^{l_1^n} 2^{-l_1(s+\frac{1}{2})} \mathbb{E} \left( \| \Delta I, N \|_{\infty} \right) \leq C \sum_{l_1=-1}^{l_1^n} 2^{l_1}\left( \frac{s-\frac{1}{2}}{2s+1} \right)^\frac{1}{2} \sqrt{\frac{\log \log N}{N}}.$$

If $1/2 < s < 1$ then $1 - s > 0$ and $D_2 \leq C \sqrt{\frac{\log \log N}{N}} N^{-\frac{4s-1}{2(s+1)}} < CN^{-\frac{s}{2(s+1)}}$, with the same argument as in (3.3.9). If $s \geq 1$ then $D_2 \leq C l_1^n \sqrt{\frac{\log \log N}{N}} < CN^{-\frac{s}{2(s+1)}}$, since $1/2 > s/(2s+1)$.

Let us finally consider the term $C_3$. From Point (i) of Proposition 2.13 and (3.3.2):

$$C_3 \leq C \sum_{l_1=-1}^{l_1^n} \mathbb{E} \left( \sup_{l_2 \in \mathbb{Z}} \left| \tilde{\beta}^{(k)}_l (\tilde{X}_l^n) - \tilde{\beta}^{(k)}_l (\tilde{X}_l^n) \right|^2 \right)^{l_1/2} \| \psi \|_{\infty}$$

$$\leq C \sqrt{\frac{\log N}{N}} 2^{l_1/2} = C \sqrt{\frac{\log N}{N}} \frac{2^{l_1/2}}{\frac{\log N}{N}^{\frac{s}{2(s+1)}}}.$$

With this last result, the proof of Theorem 2.16 is definitively achieved. ■

4. Application to the 2D-Navier–Stokes equations

We now consider the weak solutions of the 2D-Navier–Stokes equations with random initial condition (1.0.1), and the associated 2D-vortex equations (1.0.2). We regularize $K$ with a cut-off technique, and propose stochastic wavelet particle approximations to compute the intensity of their spatial statistical solutions thanks to the results of Section 2. This allows us to compute some physical quantities of interest like the mean velocity vector field.

4.1. Regularized equation

We regularize the kernel $K$ in order to apply the results of Section 2. Recall (see Marchioro and Pulvirenti [14] or Méléard [16]) that $K$ is defined by $\forall x, y \in \mathbb{R}^2, K(x - y) = \nabla \cdot \mathcal{G}(|x - y|)$, where $\mathcal{G}$ is the 2D-Poisson kernel defined by $\forall r > 0, \mathcal{G}(r) = - (\ln r)/(2\pi)$. The cut-off approximation proposed by [14] consists in regularizing the function $\mathcal{G}$ near the origin. Let $\varepsilon > 0$, and consider the cut-off equation: $\forall t \in [0, T], P(d\omega) - a.s., v^\varepsilon(0, x, \theta) = v_0(x, \theta), \mathcal{L}(\theta) = v_G(d\omega)$, and

$$\frac{\partial v^\varepsilon}{\partial t}(t, x, \theta) = - ((K_\varepsilon * v^\varepsilon)(x, \theta) \cdot \nabla) v^\varepsilon(t, x, \theta) + \nu \Delta v^\varepsilon(t, x, \theta), \quad (4.1.1)$$
where $K_\varepsilon(x) = \nabla^2 G_\varepsilon(|x|) = \left( -G'_\varepsilon(|x|) \frac{x}{|x|}, G'_\varepsilon(|x|) \frac{x}{|x|} \right)$, with $G_\varepsilon(|x|) = G(|x|)$ if $|x| \geq \varepsilon$, and extended in a $C^\infty$-way on $B(0, \varepsilon)$. It is possible to choose $G_\varepsilon$ such that its derivatives vanish at the origin and satisfy the following inequalities:

$$\forall r \geq 0, \quad |G_\varepsilon^{(k)}(r)| \leq \sup_{u \geq \varepsilon} |G^{(k)}(u)| \leq \frac{1}{2\pi \varepsilon^k}, \quad k \in \{1, 2\}. \quad (4.1.2)$$

In particular, $K_\varepsilon$ is bounded by $1/2\pi \varepsilon$ and Lipschitz continuous with constant $1/2\pi \varepsilon^2$.

### 4.2. Probabilistic approach for the 2D-Vortex equation

It is possible to generalize the work of Marchioro and Pulvirenti [14] and Méleard [15,16] with deterministic initial conditions to random initial conditions (see [23] for complete proofs). From the results of Section 2:

**Proposition 4.1.** Let $(W_t)_{t \in [0,T]}$, $\theta$ and $(X_0(a))_{a \in \mathbb{R}}$ be as in Theorem 2.7. Under (A3), (A4),

(i) The following SDE admits a unique pathwise solution: $P(\text{d}\omega)$-a.s., $\forall t \in [0, T]$,

$$
\begin{cases}
X_t^\varepsilon(\theta) = X_0(\theta) + \sqrt{2}W_t + \int_0^t K_\varepsilon \ast \tilde{Q}_s^\varepsilon(X_s^\varepsilon(\theta), \theta) \text{d}s \\
\mathcal{L}(\theta) = v_G(\text{d}a), \quad \mathcal{L}(X_0(\theta)) = \frac{|v_0(x, \theta)| \text{d}x}{\|v_0(\cdot, \theta)\|_1} = p_0(x, \theta) \text{d}x \\
\mathcal{Q}^\varepsilon(\text{d}x, \theta) = \mathcal{L}(X_0(\theta)), \quad \tilde{Q}^\varepsilon(\text{d}x, \theta) \text{ is associated with } \mathcal{Q}^\varepsilon(\text{d}x, \theta) \text{ as in (2.2.3)}.
\end{cases} \quad (4.2.1)
$$

(ii) There exists a unique weak function-solution of Eq. (4.1.1), denoted by $(v^\varepsilon(t, \cdot, \theta))_{t \in [0,T]}$.

Under (A3) and (A4), existence and uniqueness results for the vortex Eq. (1.0.2) follow by a straightforward adaptation of [15,16] to the case of random initial conditions.

**Theorem 4.2.** Let $v_0(\cdot, \theta)$ be a random initial condition satisfying (A3) and (A4).

(i) The weak function-solutions $\left((v^\varepsilon(t, \cdot, \theta))_{t \in [0,T]}\right)_{\varepsilon > 0}$ of (4.1.1) is $P(\text{d}\omega) - a.s.$ a Cauchy sequence in the complete space $\mathcal{H}$ defined in (2.2.7) when $\varepsilon$ decreases to zero.

(ii) As a consequence, there exists a unique weak function solution $(v(t, \cdot, \theta))_{t \in [0,T]}$ to Eq. (1.0.2) in $\mathcal{H}$, which satisfies moreover: $P(\text{d}\omega)$-a.s., $\forall t \in [0, T], \forall \varepsilon > 0$,

$$\max \left(\|v^\varepsilon(t, \cdot, \theta) - v(t, \cdot, \theta)\|_1, \|v^\varepsilon(t, \cdot, \theta) - v(t, \cdot, \theta)\|_\infty\right) \leq C(T, A) \varepsilon. \quad (4.2.2)$$

### 4.3. Stochastic interacting particle system

Let $m_T$ be the spatial statistical solution at time $T$ of PDE (1.0.2). We are looking for particle approximations of $(I(m_T), f)$. We proceed as in Section 2.3, except that we compute the regressions component by component. We denote by $K_\varepsilon^1(x)$ and $K_\varepsilon^2(x)$ the first and second components of $K_\varepsilon(x)$. The particle system that we consider is: $\forall i \in [1, N]$, $\forall k \in [0, K - 1]$,

$$\tilde{X}_{i,k+1} = \tilde{X}_{i,k} + \tilde{u}_\varepsilon(t_k, \tilde{X}_{i,k}^{1:N, \varepsilon}, \theta_i) \Delta t + \sqrt{2}v(W_{i,k}^{1} - W_{i,k-1}^{1}). \quad (4.3.1)$$
where:

\[
\hat{h}_b^\varepsilon(t_k, \tilde{X}_{t_k}^{i,N}, \theta_i) = \left[ \sum_{l=1}^{I_1^N} \sum_{l_2 \in \mathbb{Z}} \sum_{j=1}^{N} \frac{1}{N} \psi_j \left( G_N(\theta_j) \right) \psi_j \left( G_N(\theta_j) \right) h \left( \tilde{X}_{0}^{j,N}, \theta_j \right) \times \left( K^1_b \left( \tilde{X}_{t_k}^{j,N} - \tilde{X}_{t_k}^{j,N} \right) \right) \right],
\]

with \( \varepsilon > 0 \) and with a resolution level \( I_1^N \) as in (2.3.13).

**Theorem 4.3.** Under Assumption 2.12, \( \forall \varepsilon > 0, \forall 0 < \eta < 1, \forall f \in C_b^{\varepsilon+\eta}(\mathbb{R}, \mathbb{R}), \exists N_0, C > 0, \forall N \geq N_0, \)

\[
\mathbb{E} \left( \langle I (m_T), f \rangle - \frac{1}{N} \sum_{i=1}^{N} h \left( \tilde{X}_{0}^{i,N}, \theta_i \right) f \left( \tilde{X}_{T}^{i,N} \right) \right) \leq C \left( \Delta t + \sqrt{\log NN}^{-\frac{\varepsilon}{2\Delta t + \varepsilon}} \right).
\]

**Sketch of proof.** Let \( m^\varepsilon \) be the statistical solution of PDE (4.1.1).

\[
| \langle I (m_T), f \rangle - \langle I^\varepsilon (m_T), f \rangle | \leq \int_{\mathbb{R}} \int_{\mathbb{R}^2} |f(x)| |v(T, x, a) - v^\varepsilon (T, x, a)| dx v_G(da) \leq C(T, A) \| f \|_{\infty, \varepsilon},
\]

by (4.2.2), which is sufficient to obtain the announced result.

Since it is possible to recover the velocity \( u = (u_1, u_2) \) from the vortex \( v \) by \( u = K * v \), the mean velocity vector field associated with the 2D-Navier–Stokes equation with random initial condition (1.0.1) at point \( x \in \mathbb{R}^2 \) and time \( T > 0 \) expresses as: \( \forall i \in \{1, 2\}, \)

\[
\int_{\mathbb{R}} u_i(T, x, a) v_G(da) = \int_{\mathbb{R}} \langle v(T, dy, a), K^i(x - y) \rangle v_G(da) = \langle I (m_T), K^i(x \cdot .) \rangle.
\]

The result of Theorem 4.3 can be applied to the computation of the mean velocity vector field (4.3.4), which corresponds to the particular choice of \( f : y \in \mathbb{R}^2 \mapsto K(x - y) \) (unbounded, with unbounded derivatives near zero, see [23] for the theoretical convergence rate).

5. **Numerical experiments**

5.1. **Test case**

To compare numerically the three particle methods introduced in Section 1 on the vortex equations (1.0.2), we use the following test problem (see Milinazzo and Saffman [18] and Bossy [4]), and we consider the following quantity: \( \forall t \in [0, T], \forall a \in \Theta, \)

\[
L(t, a) = \frac{\int_{\mathbb{R}^2} |x|^2 v(t, x, a) dx}{\int_{\mathbb{R}^2} v(t, x, a) dx}, \quad \text{which satisfies } L(t, a) = L(0, a) + 4tv.
\]
Fig. 1. Left: Evolution of the relative error for the particle methods 1 (with 1024 = 32 × 32 particles in dash–dots, with 32768 = 1024 × 32 particles in dots), 2 (dashed, 1024 particles) and 3 (plain, 1024 particles). \( \nu = 10^{-6}, \varepsilon = 5 \times 10^{-4}, \Delta t = 0.05 \). Right: Evolution of the relative error for the particle methods 1 (dotted), 2 (dashed) and 3 (plain). 1024 particles. \( \nu = 10^{-8}, \varepsilon = 10^{-2}, \Delta t = 0.5 \).

Fig. 2. Approximation of the mean velocity vector field given by our particle method at time \( t_0 = 0, t_5 \) and \( t_{11} \). \( \nu = 10^{-8}, \varepsilon = 10^{-2}, \Delta t = 0.5 \).

In case \( P(\omega) \)-a.s., \( v_0(\cdot, \theta) \) is a density function, we have:

\[
\forall t \in [0, T], \int_{\mathbb{R}} L(t, a) v_G(da) = \int_{\mathbb{R}} \int_{\mathbb{R}^2} |x|^2 v(t, x, a) dx \, v_G(da) = \langle I(m_t)(dx), |x|^2 \rangle.
\]

We compute the approximations \( A(t) \) of \( t \mapsto \langle I(m_t)(dx), |x|^2 \rangle \) with the three particle methods, and measure how they fit the theoretical line (5.1.1) thanks to the relative error:

\[
e(t) = |A(t) - A(0) - 4tv| / |A(t)|.
\]

(the map \( x \in \mathbb{R}^2 \mapsto |x|^2 \) is not bounded, see [23] for the theoretical convergence rate). All the simulations presented here are performed with the \texttt{R} open source software\(^1\) and the \texttt{wavethresh} package developed by Nason et al. [19].

We have chosen Daubechies orthonormal compactly supported wavelets \( N = 8 \) (see [6]), and use a thresholded version of the wavelet regression estimators (see the details and theoretical convergence rates in [23]). We use a Gaussian kernel for the Method 2, with a window \( h_N \) that was suggested by the software using a plug-in methodology. The choice of a Gaussian kernel is natural since we deal with Gaussian laws for \( \theta \) and \( X_0(a) \).

\(^1\)http://www.r-project.org/.
The random numbers generator has been seeded such that in each of the three methods, the same simulations are used for $\theta$ and for the Brownian motions underlying the particles’ diffusions.

The following examples have been chosen in order to emphasize the advantages of the wavelet regression estimator compared with Methods 1 and 2. The latter are known to lack robustness in cases with extreme values or empty regions. The numerical results are robust to the choice of kernels and wavelets, and we present the better-looking simulations. We refer to [23] for further details.

5.2. Example 1: An uncertainty parameter $\theta$ with large variance

The initial condition is given by: $\mathcal{L}(X_0(\theta)) = \mathcal{N}(U(\theta), 0.6)$, where $\mathcal{L}(\theta) = \mathcal{N}(0, 5)$, and where $U$ is a 1-Lipschitz function bounded up and below.

In Fig. 1 (left), we can see that Method 1 (dash–dots) works very badly, while Method 2 and 3 are equivalent. The poor performance of Method 1 is due to the fact that there are not enough simulations of $\theta$ given the large variance that we have chosen for $\mathcal{L}(\theta)$. If we want to obtain a relative error comparable with Methods 2 and 3, we have to increase the number $N_1$ of simulations for $\theta$, which implies that we do more simulations than with direct methods. This is achieved if we simulate 1024 realizations of $\theta$ and 32 particles for each of these realizations (i.e. 32 768 particles, in dots on Fig. 1 (left)).
5.3. Example 2: An uncertainty parameter $\theta$ with a mixture law

Let us now give an example of a case where the wavelet particle method (Method 3) proves more robust than the random weights particle method (Method 2).

The initial condition $L(X_0(\theta)) = N(U(\theta), 0.3)$, where $\theta = 1_{Y=1}\theta^{(1)} + 1_{Y=0}\theta^{(2)}$, with $Y$, $\theta^{(1)}$, and $\theta^{(2)}$ three independent random variables, such that $Y$ is a Bernoulli random variable with $P(Y = 1) = 0.3$, and $\theta^{(1)}$ and $\theta^{(2)}$ are two Gaussian random variables with respective expectations 1.3 and 0, and standard deviation 0.2.

We are interested in the mean velocity vector field and use formula (4.3.4) to simulate its time evolution (Fig. 2). A numerical comparison (see [23]) shows that the vortex obtained in the three methods have the same profile. Let us have a look at the quality of the regression (Fig. 3) to understand the differences that can however be observed. It seems that the wavelet estimators fit the data better. In particular, compared with the wavelet estimators, the Nadaraya–Watson estimators are not robust in the gap between the two peaks of our Gaussian mixing, where fewer observations are available, and average too much and miss some aggregation feature in the “crowded” regions.

Finally, we use the test case of Section 5.1 to compare the performances of the three methods (Fig. 1 (right)). In terms of relative error, it clearly appears on this set of simulations that Method 1 gives the best performance, but also has the highest complexity. Our particle method with wavelets is a little less accurate, but remarkably better than Method 2.

Acknowledgments

I am extremely grateful to my Ph.D. advisor, Professor Sylvie Méléard, for proposing this subject to me and for her permanent encouragements, her precise corrections and her pertinent advice. I would also like to thank Professor Gérard Kerkyacharian.

References


