# A Wavelet Particle Approximation for McKean-Vlasov and Navier-Stokes Spatial Statistical Solutions

Tran Viet Chi*

15th November 2004

**Abstract**

We are interested in the computation of spatial statistical solutions' intensities for the McKean-Vlasov and 2d-vortex equations. Inspired by the probabilistic approach of Talay and Vaillant [33], we introduce and study a new adaptive particle approximation which uses wavelet estimators with hard thresholding. Convergence rates are proved by adapting to our dependent particle model a wavelet regression approach inspired by Kerkyacharian and Picard [18] in random and independent design. The difficulties arise from the fact that our particles are no more in mean field interaction. Because of the special dependency, we can not use usual coupling methods like in Sznitman [31] or Méléard [25]. The new key is to prove spectral gap inequalities in order to deal directly with the interacting particles, as Malrieu and Talay did in a different context [22].

The vortex equation associated with the 2d-Navier-Stokes equation is carried as an illustration. Additional difficulties are introduced since the interaction kernel is not bounded and since the initial conditions may have a weight and a sign. □

In this paper, we are interested in partial differential equations (*PDEs* in the following) with random initial conditions. The flow of such PDE is therefore random, and its law is called *statistical solution* of the problem.

Statistical solutions are particularly interesting when modelling complex phenomena or when introducing the notion of uncertainty in the initial state. The 2d-vortex equation we study in Part 5 of this article has been one of the equations motivating these developments. It is obtained from the famous 2d-Navier-Stokes equation, which models the velocity of a viscous incompressible fluid in the plane. In the case of this equation, the theory of statistical solutions is for instance an attempt to take into account the turbulence arising with high velocities and low viscosities. Vishik and Fursikov [35] or Constantin and Wu [8] have studied such problems with analytical tools.

Talay and Vaillant [33, 34] generalized the probabilistic approach of McKean-Vlasov equations developed by Sznitman [31] and Méléard [25] to the case of a random initial condition. In particular, they proposed an original stochastic particle method with random weights to compute numerically the moments of the statistical solutions. They left however the case of the 2d-vortex equation open.

We consider the McKean-Vlasov and 2d-vortex equations with random initial conditions, and our aim is to provide a new numerical method for the computation of the intensities of the associated spatial statistical solutions, which are the time marginals of

*Université Paris X-Nanterre, Equipe Modal'X, batiment G, 200 avenue de la République, 92101 Nanterre Cedex, France. *viet_ chi.tran@u-paris10.fr*

the statistical solutions. We follow the probabilistic approach in Talay and Vaillant [33]. The new approximation we propose is based on wavelet regression estimators. This allows us to extend the method of Talay and Vaillant [33] to a larger class of initial conditions and to obtain better asymptotic convergence rates. Another advantage is that we obtain an adaptive numerical scheme. The latter does not require any *a priori* knowledge of regularities for the parameters which govern the randomness of the initial condition. Our main result is presented as Theorem 1.2, for McKean-Vlasov equations, and is used to obtain Theorems 5.1 and 5.3, for the 2d-vortex equation.

First, we work in the quite general frame of McKean-Vlasov equations. In section 1, we recall the probabilistic setting we will consider and define the statistical solutions and their intensities. We then introduce our new particle systems and explain the advantages of using the nonlinear wavelet regression estimators inspired by the work of Kerkyacharian and Picard [18]. Parts 2 to 4 are devoted to the proof of our main result, Theorem 1.2, concerning the rate of convergence of our approximation. Because of the nonlinearity of the estimators we use, we can not couple our interacting particles with independent nonlinear diffusions as it is usually done (see Sznitman [31], Méléard [25]). The new key is to establish the concentration inequalities, which we need to obtain the convergence rates, directly on the interacting particle system, with the use of spectral gap inequalities. The calculations are similar to the ones carried by Malrieu and Talay in [22] when they construct confidence intervals for Euler schemes. We show here how these ideas are well suited for the use of statistical tools on particle systems.

Finally, Section 5 is an adaptation of our numerical method to the 2d-vortex equation. Additional difficulties arise since the interaction kernel is not bounded and since the initial conditions may have a weight and a sign. We use a cut-off equation, and a trick due to Jourdain [15] to overcome these problems. The main results for the vortex equation are given in Theorems 5.1 and 5.3. We conclude with some simulations which seem to confirm the efficiency of the algorithm that we propose.

**Notation:**

We denote by $\mathcal{C}(E, F)$ the set of continuous maps from $E$ to $F$. The space $\mathcal{C}_b^k(E, F)$ is the set of functions of class $\mathcal{C}^k$, bounded, and whose successive partial derivatives are continuous and bounded to the order $k \in \mathbb{N}$. The space $\mathcal{C}_b^{k+\varepsilon}(E, F)$ is the set of functions of $\mathcal{C}_b^k(E, F)$ whose derivatives of order $k$ are $\varepsilon$-Hölder continuous. The space $\mathcal{B}_b(E, F)$ denotes the set of bounded measurable functions from $E$ to $F$. We denote by $L^p(E)$ the usual space of measurable functions $f$ from $E$ to $\mathbb{R}$ such that $\int |f|^p < \infty$. Finally, we use $\mathcal{D}(E)$ for the set of density functions on $E$.

For any random variable $X$, we write $\mathcal{L}(X)$ or $P^X$ for its law.

For a measurable space $E$, we write $\mathcal{P}(E)$ the set of probabilities on $E$. For $Q \in \mathcal{P}(E)$ and $f \in \mathcal{B}_b(E, \mathbb{R})$, we use the notation $\langle Q, f \rangle$ for $\int_E f dQ$.

Let $(\Omega, P)$ be a probability space which will characterize the randomness of the initial condition.

The stochastic processes studied here will be considered as random variables defined on $(\mathcal{C}([0, T], \mathbb{R}^n), \mathbb{P}_W)$, where $\mathbb{P}_W$ is the Wiener probability measure. The canonical process is thus a Brownian motion in $\mathbb{R}^n$, which we will write $(W_t)_{t \in [0,T]}$. $\mathbb{E}$ is the expectation operator under the Wiener law. $(\mathcal{F}_t)_{t \in [0,T]}$ denotes the natural filtration associated with $(W_t)_{t \in [0,T]}$.

We will be lead to introduce the space $\mathcal{C}([0, T], \mathbb{R}^n) \times \mathbb{R}$, on which we will define a reference measure $\mathbb{P}^\nu := \mathbb{P}_W \otimes \nu$, with $\nu$ a probability measure on $\mathbb{R}$. $\mathbb{E}^\nu$ is the expectation operator under this measure.

When we will be dealing with Euler schemes, we will write $\Delta t$ for the discretization step, which will be assumed constant and less than 1 for the sake of simplicity. On the interval $[0,T]$, with $T > 0$, we can choose for instance $\Delta t = \frac{T}{K}$, which defines $K+1$ discretization times $t_k = k\,\Delta t$ for $k = 0$ to $K$.

Finally, $C$ is a positive real constant that can change from line to line.

# 1 A Probabilistic Approach for the Computation of McKean-Vlasov Spatial Statistical Solutions' Intensities

## 1.1 Statistical Solutions for a McKean-Vlasov PDE and their Intensities

Let us consider a random function $\widehat{p}_0(x,\omega)$ depending on $x \in \mathbb{R}^n$ and on an alea $\omega \in \Omega$, such that $P(d\omega)$-almost surely ($P(d\omega)$-*a.s.* in the following) $\widehat{p}_0(.,\omega)$ is a density function. We call $\mu$ the law of probability of $\widehat{p}_0$, considered as a $L^1(\mathbb{R}^n)$-valued random variable with support in $\mathcal{D}(\mathbb{R}^n)$.We will suppose, in all the article, that the randomness of $\widehat{p}_0$ arises through its dependence on a scalar random variable $\theta \in \mathbb{R}$ of law $\nu$:

$$P(d\omega) - a.s., \widehat{p}_0(.,\omega) = p_0(.,\theta(\omega)),$$

where $p_0$ is a deterministic measurable function on $\mathbb{R}^n \times \mathbb{R}$. In this case, $\mu$ is the image measure of $\nu$ through the mapping:

$$\Phi \,:\, a \mapsto p_0(.,a).$$

We consider the following *McKean-Vlasov equation with random initial condition* $\widehat{p}_0$: $P(d\omega) - a.s.$, $\forall t \in [0,T]$, $\forall x \in \mathbb{R}^n$,

$$\left\{\begin{array}{rcl} \frac{\partial}{\partial t}p(t,x,\theta) & = & -\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(u_{b,i}(t,x,\theta)p(t,x,\theta)\right) \\ & + & \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}\left((u_\sigma(t,x,\theta)u_\sigma^*(t,x,\theta))_{ij}p(t,x,\theta)\right) \\ p(0,x,\theta) & = & p_0(x,\theta) \\ u_b(t,x,\theta) & = & \int_{\mathbb{R}^n} b(x,y)p(t,y,\theta)dy \\ u_\sigma(t,x,\theta) & = & \int_{\mathbb{R}^n} \sigma(x,y)p(t,y,\theta)dy, \end{array}\right. \tag{1}$$

where $b$ and $\sigma$ are functions respectively defined from $(\mathbb{R}^n)^2$ to $\mathbb{R}^n$ and from $(\mathbb{R}^n)^2$ to $\mathcal{M}_{n\times n}(\mathbb{R})$, the set of $n \times n$ real matrices.

The probabilistic approach which constitutes the frame of our study relies on the *weak form* of PDE (1).

**Definition 1.** *We say that the random variable in $\mathcal{C}([0,T],\mathcal{P}(\mathbb{R}^n))$, $(Q_t(dx,\theta))_{t\in[0,T]}$, is a weak measure-solution of the McKean-Vlasov equation if it satisfies:*

$$P(d\omega) - a.s., \forall \phi \in \mathcal{C}_b^2(\mathbb{R}^n,\mathbb{R}), \forall t \in [0,T],$$

$$\left\{\begin{array}{rcl} \int \phi(x)Q_t(dx,\theta) & = & \int \phi(x)p_0(x,\theta)dx + \int_0^t \int \Big(\sum_{i=1}^n u_{b,i}(s,x,\theta)\frac{\partial \phi}{\partial x_i}(x) \\ & + & \frac{1}{2}\sum_{i,j=1}^n (u_\sigma(s,x,\theta)u_\sigma^*(s,x,\theta))_{ij}\frac{\partial^2 \phi}{\partial x_i \partial x_j}(x)\Big)\, Q_s(dx,\theta)\, ds \\ u_b(t,x,\theta) & = & \int_{\mathbb{R}^n} b(x,y)Q_t(dy,\theta) \\ u_\sigma(t,x,\theta) & = & \int_{\mathbb{R}^n} \sigma(x,y)Q_t(dy,\theta) \\ Q_0(dx,\theta) & = & p_0(x,\theta)dx. \end{array}\right. \tag{2}$$

$\diamondsuit$

**Remark 1.** *With an abuse of notation, we will say that a random probability measure $Q(\theta)$ on $\mathcal{C}\left([0,T],\mathbb{R}^n\right)$ is a weak measure-solution of PDE (2) when its time marginals constitute a weak measure-solution in the sense of Definition 1.* $\square$

**Remark 2.** *When $P(d\omega) - a.s.$ and for all $t \in [0,T]$, the time marginals $Q_t(dx, \theta)$ of the weak measure-solution of PDE (2) admit densities $p(t,x,\theta)$ with respect to the Lebesgue measure $dx$ on $\mathbb{R}^n$, the family of these densities is called weak function-solution of the PDE. It can be viewed as a random function in $\mathcal{C}([0,T], L^1(\mathbb{R}^n))$, the continuity being a $L^1$-weak continuity.* □

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

**Definition 2.** *When the evolution problem (2) admits a unique weak measure-solution $(Q_t(dx,\theta))_{t\in[0,T]}$ (up to a null-set), then the following map is $\mu$-a.s. well defined:*

$$\begin{array}{rccl} S: & L^1(\mathbb{R}^n) & \rightarrow & \mathcal{C}\left([0,T], \mathcal{P}(\mathbb{R}^n)\right) \\ & p = p_0(.,a) & \mapsto & (Q_t(dx,a))_{t\in[0,T]}. \end{array}$$

*The law $m \in \mathcal{P}\left(\mathcal{C}\left([0,T], \mathcal{P}(\mathbb{R}^n)\right)\right)$ of the weak measure-solution of (2) can thus be written as the image measure of $\mu$ through the mapping $S$, and is called statistical solution of the McKean-Vlasov problem:*

$$m = \mu \circ S^{-1} = \nu \circ (S \circ \Phi)^{-1}. \tag{3}$$

*The marginal $Q_t(dx,\theta)$ at time $t$ of the weak measure-solution of (2) is a random probability measure on $\mathbb{R}^n$ whose law $m_t \in \mathcal{P}\left(\mathcal{P}(\mathbb{R}^n)\right)$ is the $t$-time marginal of $m$. The probability measure $m_t$ is called spatial statistical solution at time $t$ of the McKean-Vlasov equation.* ♢

The setting of the problem can be summed up in the following diagram:

$$\begin{array}{ccccccc} \Omega & \xrightarrow{\theta} & \mathbb{R} & \xrightarrow{\Phi} & L^1(\mathbb{R}^n) & \xrightarrow{S} & \mathcal{C}\left([0,\,T],\, \mathcal{P}\left(\mathbb{R}^n\right)\right) \\ \omega & \mapsto & a & \mapsto & p_0(.,\,a) & \mapsto & Sp_0(dx,\,a) = (Q_t(dx,\,a))_{t\in[0,T]} \\ P & & \nu & & \mu = \nu \circ \Phi^{-1} & & m \end{array}$$

When dealing with statistical solutions, only mean quantities can be computed. Since we consider random probabilities, interesting mean quantities are their intensities (see Sznitman [31]). The intensity $I(m_t) \in \mathcal{P}\left(\mathbb{R}^n\right)$ of the spatial statistical solution $m_t$ at time $t$ is defined by:

$$\begin{aligned} \forall f \in \mathcal{B}_b\left(\mathbb{R}^n, \mathbb{R}\right), \langle I(m_t), f\rangle &= \int_{\mathcal{P}(\mathbb{R}^n)} \langle Q, f\rangle m_t(dQ) \\ &= \int_{\mathbb{R}} \langle (S \circ \Phi(a))_t, f\rangle \, \nu(da). \end{aligned} \tag{4}$$

The aim of this article is to approximate this last expression (4) by using a wavelet stochastic particle method.

**Remark 3.** *When there exists a unique weak function-solution to PDE (2), and when the initial mean energy is finite ($\int_{\mathbb{R}} ||p_0(.,a)||_{L^2}\nu(da) < \infty$), the intensity $I(m_t)$ can be linked to the first moment of the statistical solution as defined in Talay and Vaillant [33]: $\forall t \in [0,T],\ \exists! M_1(t) \in L^2(\mathbb{R}^n),\ \forall f \in L^2(\mathbb{R}^n),\ \langle I(m_t), f\rangle = \langle M_1(t), f\rangle_{L^2(\mathbb{R}^n)}$.* □

## 1.2 Probabilistic Approach of the Problem

We now recall the probabilistic interpretation for the McKean-Vlasov equation given in Talay and Vaillant [33].

We deal here with two different sources of randomness. The first one comes from the initial condition and will be considered through the random variable $\theta$. The second one is related to the probabilistic approach, which consists in looking for Markovian processes whose time marginals satisfy the weak PDE (2). Here, the process of interest is given by the nonlinear stochastic differential equation (*SDE*) introduced in Theorem 1.1.

We state existence and uniqueness results for this SDE, then we enounce existence and uniqueness results for the weak PDE (2).

**Assumption 1.** *The following assumptions will be made for $b$, $\sigma$ and $p_0$:*

- $(A_1)$ *The functions $b$ and $\sigma$ are Lipschitz continuous functions:* $\exists L > 0$, $\forall x, y, z, u \in \mathbb{R}^n$, $|b(x,y) - b(z,u)| + ||\sigma(x,y) - \sigma(z,u)|| \leq L(|x-z| + |y-u|)$,
- $(A_2)$ *The functions $b$ and $\sigma$ are bounded,*
- $(A_3)$ $P(d\omega) - a.s.$, $p_0(.,\theta)$ *is a density function,*
- $(A_4)$ $P(d\omega) - a.s.$, $\int x^2 p_0(x,\theta)dx < \infty$.

**Theorem 1.1.** *Suppose that Assumptions $(A_1)$ to $(A_4)$ are satisfied. Let $(W_t)_{t\in[0,T]}$ be a Brownian motion on $\mathbb{R}^n$, let $\theta$ be a random variable of law $\nu$ and let $(X_0(a))_{a\in\mathbb{R}}$ be a family of random variables such that $a \mapsto X_0(a)$ is measurable and such that $P(d\omega) - a.s.$, $\mathcal{L}(X_0(\theta)) = p_0(x,\theta)dx$. 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 are available for the following SDE: $P(d\omega) - a.s.$, $\forall t \in [0,T]$,*

$$\begin{cases} dX_t(\theta) = u_b(t, X_t(\theta), \theta)dt + u_\sigma(t, X_t(\theta), \theta)dW_t \\ u_b(t,x,\theta) = \int_{\mathbb{R}^n} b(x, y)P^{X_t(\theta)}(dy) \\ u_\sigma(t,x,\theta) = \int_{\mathbb{R}^n} \sigma(x, y)P^{X_t(\theta)}(dy) \\ \mathcal{L}\left(X_0(\theta)\right) = p_0(x,\theta)dx. \end{cases} \tag{5}$$

□

*Proof.* Notice first that the law of $X_t(\theta)$ depends continuously on the initial condition $X_0(\theta)$ (see Kunita [19]), which is itself measurable in $\theta$. The conditional law of $X_t(\theta)$ knowing $\theta$ is therefore well defined. The idea of the proof is that existence and uniqueness results for the conditioned diffusion knowing $\theta$ are similar to their non-conditioned counterparts.

Since the random variable $\theta$ is independent of the Brownian motion $(W_t)_{t\in[0,T]}$ and of the family of random variables $(X_0(a))_{a\in\mathbb{R}}$, looking at the conditional version knowing $\theta$ of SDE (5) amounts to considering this equation for every realization $a$ of $\theta$. Hence, we are lead to study existence and pathwise properties for the following SDE:

$$\begin{cases} dX_t(a) = u_b(t, X_t(a), a)dt + u_\sigma(t, X_t(a), a)dW_t \\ \mathcal{L}\left(X_0(a)\right) = p_0(x,a)dx. \end{cases} \tag{6}$$

The solution $(X_t(\theta))_{t\in[0,T]}$ of SDE (5) is then linked to the solution $(X_t(a))_{t\in[0,T]}$ of SDE (6) by:

$$\mathcal{L}\left(X_t(\theta)\,|\,\theta = a\right) = \mathcal{L}\left(X_t(a)\right). \tag{7}$$

For a given realization $a$ of $\theta$, we can use the proofs in Sznitman [31] and Méléard [25]. Let $\mu = (\mu_t(dx))_{t\in[0,T]}$ be a given family of probabilities in $\mathcal{C}([0,T], \mathcal{P}(\mathbb{R}^n))$. Let us define:

$$u_b^\mu(t,x) = \int_{\mathbb{R}^n} b(x,y)\mu_t(dy), \qquad u_\sigma^\mu(t,x) = \int_{\mathbb{R}^n} \sigma(x,y)\mu_t(dy). \tag{8}$$

We can associate with the nonlinear SDE (6) a linear SDE by replacing the coefficients $u_b(t,x,a)$ and $u_\sigma(t,x,a)$ with $u_b^\mu(t,x)$ and $u_\sigma^\mu(t,x)$ respectively:

$$X_t^\mu(a) = X_0(a) + \int_0^t u_\sigma^\mu(s, X_s^\mu(a))dW_s + \int_0^t u_b^\mu(s, X_s^\mu(a))ds, \quad \text{for } t \in [0,T]. \tag{9}$$

Pathwise existence and uniqueness of the solution of the linear SDE (9) hold, thanks to Theorems 2.5 and 2.9 (pages 287 and 289) in Karatzas and Shreve [16].

To deduce the existence and uniqueness of a weak solution for SDE (6), we use a fixed point theorem. The weak solutions of SDE (6) and SDE (9) belong to the space:

$$\begin{aligned}\mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)) \;=\; & \big\{P \text{ probability law on } \mathcal{C}([0,T],\mathbb{R}^n) \;\mid\; \langle P, \sup_{t\leq T}|X_t|^2\rangle < \infty \\ & \text{ where } X_t \text{ is the canonical process on } \mathcal{C}([0,T],\mathbb{R}^n)\big\}.\end{aligned}$$

This space, endowed with the weak convergence, is metrisable with the Vaserstein complete metric $\rho_T$, defined for $m^1, m^2 \in \mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n))$ by (see Rachev [29]):

$$\begin{aligned}\rho_T^2(m^1,m^2) \;=\; & \inf\Big\{\int_{\mathcal{C}([0,T],\mathbb{R}^n)^2}\sup_{t\leq T}|x_t-y_t|^2 m(dx,dy) \;\mid\; m\in\mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)\times\mathcal{C}([0,T],\mathbb{R}^n)) \\ & \text{with marginals } m^1 \text{ and } m^2\Big\}.\end{aligned}$$

Existence and uniqueness of the weak solution of SDE (6) are equivalent to the existence and uniqueness of a fixed point for the mapping:

$$\begin{aligned}\zeta_a \;:\; \mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)) &\rightarrow \mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)) \\ \mu &\mapsto \mathcal{L}(X^\mu(a)),\end{aligned}$$

where $X^\mu(a)$ solves SDE (9). The latter result is obtained thanks to the completeness of $(\mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)), \rho_T)$ and to the following inequality (see Sznitman [31] or Méléard [25] for a complete proof):

$$\forall t\in[0,T],\ \forall m^1,m^2\in\mathcal{P}_2(\mathcal{C}([0,T],\mathbb{R}^n)), \rho_t^2(\zeta_a(m^1),\zeta_a(m^2)) \leq C_T\int_0^t \rho_u^2(m^1,m^2)du.$$

Finally, for a given initial condition $X_0(a)$ and a given Brownian motion $(W_t)_{t\in[0,T]}$, pathwise existence and uniqueness for the linear SDE (9) where $\mu$ has been set to the unique weak solution of SDE (6) imply pathwise existence and uniqueness for SDE (6). ∎

**Proposition 1.1.** *Under Assumptions $(A_1)$ to $(A_4)$, there exists a weak measure-solution to PDE (2).* □

*Proof.* This can be seen by using Itô's formula to compute $\mathbb{E}^\nu\phi(X_t(\theta))$, where $\phi\in\mathcal{C}_b^2(\mathbb{R}^n,\mathbb{R})$ and $(X_t(\theta))_{t\in[0,T]}$ is the solution of SDE (5). ∎

**Proposition 1.2.** *Suppose that Assumptions $(A_1)$ to $(A_4)$ are satisfied. Then if $(P_t^1(dx,\theta))_{t\in[0,T]}$ and $(P_t^2(dx,\theta))_{t\in[0,T]}$ are two weak measure-solutions of PDE (2):*

$$P(d\omega)-a.s., \forall t\in[0,T],\ P_t^1(dx,\theta) = P_t^2(dx,\theta).$$

□

*Proof.* As in the proof of Theorem 1.1, we can work conditionally to a realization $a$ of $\theta$ such that Assumptions $(A_1)$ to $(A_4)$ are satisfied.

Let $\mu = (\mu_t(dx))_{t\in[0,T]}$ be a given family of probabilities in $\mathcal{C}([0,T],\mathcal{P}(\mathbb{R}^n))$. We can associate with the nonlinear PDE (2) and with the nonlinear SDE (5) linear versions by replacing the coefficients $u_b(t,x,a)$ and $u_\sigma(t,x,a)$ by $u_b^\mu(t,x)$ and $u_\sigma^\mu(t,x)$ defined in (8).

Take $\mu_a^1 = (P_t^1(dx,a))_{t\in[0,T]}$ and $\mu_a^2 = (P_t^2(dx,a))_{t\in[0,T]}$ two measure-solutions of non-linear PDE (2).

Consider the solutions $X^1(a) = (X_t^1(a))_{t\in[0,T]}$ and $X^2(a) = (X_t^2(a))_{t\in[0,T]}$ of the linear SDE (9) with respectively $\mu_a^1$ and $\mu_a^2$. Using Itô's formula, we see that $\mathcal{L}(X^1(a))$ and $\mathcal{L}(X^2(a))$ are solutions of the linear PDE associated with (2) with respectively $\mu_a^1$ and $\mu_a^2$.

Since $\mu_a^1$ and $\mu_a^2$ solve the nonlinear PDE (2), they also solve the associated linear PDEs defined respectively with $\mu_a^1$ and $\mu_a^2$.

Since the generator of the linear PDE associated with (2) for a given $\mu$ is a diffusion generator with bounded coefficients, we can use Theorem 5.2 in Bhatt and Karandikar [3] to obtain uniqueness of the weak measure solution of this PDE. Thus: $\forall t \in [0,T]$, $\mathcal{L}(X_t^1(a)) = P_t^1(dx,a)$, and $\mathcal{L}(X_t^2(a)) = P_t^2(dx,a)$.

Consequently, $X^1(a)$ and $X^2(a)$ also solve the nonlinear SDE (6) (Notice that the nonlinearity in SDE (6) only plays through the solutions' marginal laws). Thanks to Theorem 1.1, uniqueness implies: $\mathcal{L}(X^1(a)) = \mathcal{L}(X^2(a))$.

Thus: $\forall t \in [0,T]$, $P_t^1(dx,a) = P_t^2(dx,a)$. ■

We can now reformulate the intensity of the spatial statistical solutions defined in (4) with the nonlinear diffusions (5) and (6):

$$\forall t \in [0,T], \forall f \in \mathcal{B}_b(\mathbb{R}^n,\mathbb{R}), \langle I(m_t), f\rangle = \int_{\mathbb{R}} \mathbb{E} f\left(X_t(a)\right) \nu(da) \tag{10}$$

$$= \mathbb{E}^\nu f(X_t(\theta)). \tag{11}$$

### 1.3 Particle Approximations

**Remark 4.** *Until section 4, we will assume, for the sake of simplicity, that $n = 1$.* □

We review here three approximations for the computation of the intensity $I(m_T)$ applied to a test function $f \in \mathcal{B}_b(\mathbb{R}^n,\mathbb{R})$. This allows us to explain our purpose. The first approximation relies on the formulation (10), while the second and third ones use the expression (11). The two first approximations have been studied by Talay and Vaillant [33]. The third one is the original approximation which makes the object of this article.

We are interested here in the case where the law $\nu$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$. We refer to Talay and Vaillant [33] for the discrete case.

#### 1.3.1 Existing Particle Approximations and Problematics

**Method 1:** The first "naive" idea is to return to the standard McKean-Vlasov case, using (10). This first method has been mentioned by Talay and Vaillant [33].

We set $\theta$ to $a$ and approximate the expectation $\mathbb{E}f(X_T(a))$ under the integral in (10) by computing a mean over interacting particles. The latter are simulated by replacing the unknown law $\mathcal{L}(X_t(a))$ appearing in the coefficients $u_b(t,x,a)$ and $u_\sigma(t,x,a)$ by the empirical law of the particle system. The convergence of the empirical law to $\mathcal{L}(X_t(a))$ is known as *propagation of chaos* and has been described in Méléard [25] or Sznitman [31]. Once we have done this, we evaluate in turn the integral in $a$ with a Monte-Carlo approximation.

The algorithm is thus the following:

1. We simulate $N_1$ random variables $(\theta_l)_{l\in[1,N_1]}$ i.i.d. of law $\nu$.
2. For each $\theta_l$, we simulate $N_2$ particles, $\left(\bar{Y}^{i,l,N_2}(\theta_l)\right)_{i\in[1,N_2]}$. To this purpose, we simulate $N_2$ random initial conditions $\left(\bar{Y}_0^{i,l,N_2}(\theta_l)\right)_{i\in[1,N_2]}$ i.i.d. of law $p_0(x,\theta_l)dx$. Then, we define the paths of the particle system with an Euler scheme of discretization step $\Delta t$ and of discretization times $t_k = k\Delta t$: $\forall l \in [1,N_1]$, $\forall i \in [1,N_2]$, $\forall k \in [0,K]$,

$$\begin{aligned}\bar{Y}_{t_{k+1}}^{i,l,N_2}(\theta_l) &= \tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} b\left(\bar{Y}_{t_k}^{i,l,N_2}(\theta_l), \bar{Y}_{t_k}^{j,l,N_2}(\theta_l)\right)\Delta t \\ &+ \tfrac{1}{N_2}\textstyle\sum_{j=1}^{N_2} \sigma\left(\bar{Y}_{t_k}^{i,l,N_2}(\theta_l), \bar{Y}_{t_k}^{j,l,N_2}(\theta_l)\right)(W_{t_{k+1}}^{i,l} - W_{t_k}^{i,l}),\end{aligned}$$

where $(W^{i,l})_{i\in[1,N_1],l\in[1,N_2]}$ is a $N_1 \times N_2$-dimensional Brownian motion, independent from the random variables $(\theta_l)_{l\in[1,N_1]}$ and from the initial conditions $\left(\bar{Y}_0^{i,l,N_2}(\theta_l)\right)_{i\in[1,N_2],l\in[1,N_1]}$.

3. An approximation of $\langle I(m_T), f\rangle$ is given by:

$$\langle I(m_T), f\rangle \simeq \frac{1}{N_1}\sum_{l=1}^{N_1}\frac{1}{N_2}\sum_{i=1}^{N_2} f\left(\bar{Y}_T^{i,l,N_2}(\theta_l)\right).$$

The corresponding approximation error, though not given by Talay and Vaillant [33], can be fairly obtained:

**Proposition 1.3.** *If* $b$, $\sigma \in \mathcal{C}_b^{4+\varepsilon}(\mathbb{R}^2, \mathbb{R})$*, with* $\varepsilon > 0$*, and if the initial condition* $p_0$ *satisfies* $(A_3)$ *and* $(A_4)$*, then:* $\forall f \in \mathcal{C}_b^{4+\varepsilon}(\mathbb{R}, \mathbb{R})$, $\exists C = C(T, f, b, \sigma) > 0$,

$$\mathbb{E}|\langle I(m_T),\, f\rangle - \frac{1}{N_1}\sum_{l=1}^{N_1}\frac{1}{N_2}\sum_{i=1}^{N_2} f(\bar{Y}_T^{i,l,N_2}(\theta_l))| \quad\leq\quad C\left(\frac{1}{\sqrt{N_1}} + \frac{1}{\sqrt{N_2}} + \Delta t\right).$$

□

The asymptotic rate of convergence here is the one of the Central Limit Theorem ($CLT$), which is the best we can hope with stochastic methods.

However, as pointed out by Talay and Vaillant, this method is numerically very expensive, since it uses a two-steps imbricated simulation procedure. The number of simulations at each stage has to be high. The total number of simulations, $N_1 \times N_2$, with $N_1$ and $N_2$ high, is therefore very large: Vaillant ([34], section 3.7 page 79) showed that the complexity of this method is of order $O\left(N_1 N_2^2/\Delta t\right)$.

Notice also, that contrarily to the two other methods, Method 1 does not require any assumption on the dimension of the space where $\theta$ takes its values.

**Methods 2 and 3:** The aim is to find numerically more efficient methods. In order to avoid imbricated simulations, we follow Talay and Vaillant [33] and consider particle approximations based on Expression (11) and SDE (5).

We approximate directly the expectation $\mathbb{E}^\nu f(X_T(\theta))$ in (11) by computing a mean over interacting particles $(\bar{X}_T^{i,N}(\theta_i))_{i\in[1,N]}$ whose laws are expected to be close to $\mathcal{L}(X_T(\theta))$.

To simulate these particles, we have to replace the unknown coefficients $u_b(t, x, \theta)$ and $u_\sigma(t, x, \theta)$ by quantities depending on the empirical measure of the system. Since SDE (5) can be rewritten as: $P(d\omega) - a.s.$, $\forall t \in [0, T]$,

$$\left\{\begin{array}{l} dX_t(\theta) = \mathbb{E}^\nu\left((b(x, X_t(\theta))|\theta\right)|_{x=X_t(\theta)}\, dt + \mathbb{E}^\nu\left((\sigma(x, X_t(\theta))|\theta\right)|_{x=X_t(\theta)}\, dW_t \\ \mathcal{L}\left(X_0(\theta), \theta\right) = p_0(x, a)dx\, \nu(da) \\ W \text{ is a Brownian motion independent of the random variable } \theta \text{ and of } X_0(\theta), \end{array}\right. \tag{12}$$

(we will choose for $(t, x) \mapsto \mathbb{E}^\nu\left(b(x, X_t(\theta))|\theta\right)$ and $(t, x) \mapsto \mathbb{E}^\nu\left(\sigma(x, X_t(\theta))|\theta\right)$ continuous modifications of the conditional expectation processes), this amounts to approximating the expectations in the law of $X_t(\theta)$ conditionally to $\theta$. To this purpose, we will use regression estimators.

The particle approximations which are considered here are based on the following algorithm:

1. Simulate $N$ random variables $(\theta_i)_{i\in[1,N]}$ of law $\nu$,

2. For all $i \in [1,N]$, associate a (single) particle, defined by its initial condition $\bar{X}_0^{i,N}(\theta_i)$ of law $p_0(x,\theta_i)dx$ and by its trajectory described by the following Euler scheme: $\forall i \in [1,N]$, $\forall k \in [0,K]$,

$$\bar{X}_{t_{k+1}}^{i,N}(\theta_i) = \bar{X}_{t_k}^{i,N}(\theta_i) + \widehat{u_b}\left(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i\right)\Delta t + \widehat{u_\sigma}\left(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i\right)\left(W_{t_{k+1}}^i - W_{t_k}^i\right), \quad (13)$$

where $\widehat{u_b}(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i)$ and $\widehat{u_\sigma}(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i)$ are some regression estimators constructed on our particle system: we compute the regression of the observations $\left(b(\bar{X}_{t_k}^{i,N}(\theta_i), \bar{X}_{t_k}^{j,N}(\theta_j))\right)_{j\in[1,N]}$ on $(\theta_j)_{j\in[1,N]}$ to approximate $a \mapsto u_b(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), a) = \mathbb{E}^\nu\left(b(x, X_{t_k}(\theta))|\theta = a\right)|_{x=\bar{X}_{t_k}^{i,N}(\theta_i)}$. The value of the regression function at point $\theta_i$ gives us $\widehat{u_b}(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i)$. The same goes for $\widehat{u}_\sigma$.

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

$$\langle I(m_T), f\rangle \simeq \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}_T^{i,N}(\theta_i)\right). \quad (14)$$

The two following stochastic particle methods use this approach, but with different regression estimators.

**Method 2: The Particle Method with Random Weights**.

This method has been introduced by Talay and Vaillant [33]. It is based on the use of Nadaraya-Watson regression estimators to compute $\widehat{u_b}$ and $\widehat{u_\sigma}$: $\forall i \in [1,N]$, $\forall k \in [0,K]$,

$$\begin{aligned}
\widehat{u_b}\left(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i\right) &= \sum_{j=1}^N \frac{H_N(\theta_j - \theta_i)}{\sum_{l=1}^N H_N(\theta_j - \theta_l)} b\left(\bar{X}_{t_k}^{i,N}(\theta_i), \bar{X}_{t_k}^{j,N}(\theta_j)\right) \\
\widehat{u_\sigma}\left(t_k, \bar{X}_{t_k}^{i,N}(\theta_i), \theta_i\right) &= \sum_{j=1}^N \frac{H_N(\theta_j - \theta_i)}{\sum_{l=1}^N H_N(\theta_j - \theta_l)} \sigma\left(\bar{X}_{t_k}^{i,N}(\theta_i), \bar{X}_{t_k}^{j,N}(\theta_j)\right), \quad (15)
\end{aligned}$$

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 [4] for more information. A possible choice for $H$ is the Gaussian density function for instance).

Talay and Vaillant [33] and Vaillant [34] computed the accuracy of this second method:

**Proposition 1.4.** *(Talay and Vaillant [33]) Under the following assumptions:*

1. *The functions $b$ and $\sigma$ are in $\mathcal{C}_b^{4+\varepsilon}(\mathbb{R},\mathbb{R})$,*
2. *The law $\nu$ of $\theta$ is absolutely continuous w.r.t. the Lebesgue measure and has a strictly positive Lipschitz continuous density $g$ supported by a compact interval $\Theta \subset \mathbb{R}$,*
3. *The application $\Phi : a \mapsto p_0(.,a)$ is Lipschitz continuous for the norm in $L^1(\mathbb{R})$,*
4. *$\forall a \in \Theta$, $p_0(.,a)$ is a density function,*
5. *$\sup_{a\in\Theta}\int_{\mathbb{R}} x^4 p_0(x,a)dx < \infty$,*
6. *$\lim_{N\to\infty} h_N \to 0$ and $\lim_{N\to\infty} \frac{\log N}{N h_N^2} = 0$,*
7. *The kernel $H$ used in the weights is a Parzen-Rosenblatt kernel.*

*Consider the particle system (13) constructed with the estimators defined in (15), then:*

$$\forall 0 < \varepsilon < 1,\ \forall f \in \mathcal{C}_b^{4+\varepsilon}(\mathbb{R},\mathbb{R}),\ \exists C = C(T,f,b,\sigma) > 0,\ \exists N_0,\ \forall N \geq N_0,$$

$$\mathbb{E}^\nu|\langle I(m_T), f\rangle - \frac{1}{N}\sum_{i=1}^N f(\bar{X}_T^{i,N}(\theta_i))| \leq C\left(\Delta t + \frac{1}{N^{1/4}h_N^{1/4}} + h_N^{1/4}\right). \quad (16)$$

□

The asymptotic rate is not as accurate as the one obtained with Method 1 because convergence rates of non-parametric regression estimators are slower than the ones of empirical means.

An optimisation of the right hand side in $h_N$ tells us that a window $h_N \sim 1/\sqrt{N}$ gives a convergence rate in $N^{-1/8}$. This rate is however never attained, since the condition on the window constrains us to choose $h_N$ such that $\log N/(Nh_N^2) \to 0$.

However, simulations are faster than in Method 1 and less demanding in memory. The random weights $H_N(\theta_j - \theta_i)/\left(\sum_{l=1}^N H_N(\theta_j - \theta_l)\right)$ depend only on the realizations of $(\theta_i)_{i\in[1,N]}$ and neither on the function being regressed, nor on $t_k$. These weights are thus computed once and for all at the beginning of the algorithm, when the $(\theta_i)_{i\in[1,N]}$ are simulated, and are then used for the computation of $\widehat{u_b}$ and $\widehat{u_\sigma}$ at each date $t_k$.

From the computations of Vaillant ([34], section 3.7 page 79), we deduce that the complexity of this method is of order $O\left(N^2/\Delta t\right)$.

### 1.3.2 Method 3: The Wavelet Particle Approximation

We now propose a third method. It is based on the algorithm (13) with the choice of truncated warped wavelet regression estimators for the computation of $\widehat{u_b}$ and $\widehat{u_\sigma}$. These estimators are inspired by the works of Kerkyacharian and Picard [18]. Since the two regression estimators are similar, we deal only with $\widehat{u_b}$. The results are the same for $\widehat{u_\sigma}$. A short review on the wavelet theory and on wavelet regression in random design is given in Appendices A and B.

Let us consider a Multi-Resolution Analysis (*MRA*) generated by the father wavelet $\phi \in L^2(\mathbb{R})$ and the mother wavelet $\psi \in L^2(\mathbb{R})$, satisfying the following concentration and moment properties:

**Assumption 2.** *1. Concentration Property* $(H)$*:* $\forall \alpha \in \mathbb{R}$, $\sum_{I_2\in\mathbb{Z}} |\phi(\alpha + I_2)| \leq C < \infty$.

*2. M Null-Moments Property* $(M)$*:*

$$\begin{array}{lcl} \exists M \geq 1, & \int \left(1+|\alpha|^M\right)|\phi(\alpha)|d\alpha & < & \infty \\ & \int \left(1+|\alpha|^M\right)|\psi(\alpha)|d\alpha & < & \infty, \\ \forall m \in [0, M-1], & \int \alpha^m \psi(\alpha)d\alpha & = & 0. \end{array}$$

We define the descendants of $\psi$ by $\psi_I(\alpha) = 2^{I_1/2}\psi(2^{I_1}\alpha - I_2)$, where $I = (I_1, I_2)$ is a double index with $I_1 \geq 0$ and $I_2 \in \mathbb{Z}$. To simplify notations, $\phi$ is often written $\psi_{-1,0}$, and we define for any $I_2 \in \mathbb{Z}$, $\psi_{-1 I_2}(\alpha) = \phi(\alpha - I_2)$.

A wavelet decomposition on the MRA $(\psi_I)_{I_1\geq -1, I_2\in\mathbb{Z}}$ is available for any function of $L^2$ (see Appendix A).

Recall that our aim is to approximate the conditional expectation:

$$u_b(t_k, x, a) = \mathbb{E}^\nu\left(b(x, X_{t_k}(\theta))|\theta = a\right).$$

In the following, we denote by $G$ the distribution function of $\theta$. We will assume that the law $\nu$ of $\theta$ admits a density $g$ w.r.t. the Lebesgue measure on $\mathbb{R}$ and that this density has a connected support $\Theta \subset \mathbb{R}$, possibly $\mathbb{R}$ itself. The distribution function $G$ hence defines a bijection from the interior of $\Theta$ into $]0, 1[$. We also define $G_N$ the empirical equivalent of $G$:

$$\forall a \in \mathbb{R},\ G_N(a) = \frac{1}{N}\sum_{i=1}^N 1_{\theta_i \leq a}. \tag{17}$$

Since $\alpha \mapsto u_b(t_k, x, G^{-1}(\alpha))$ is a bounded application with support in $[0,1]$, it belongs to $L^2([0,1])$ and a wavelet expansion on the MRA $(\psi_I)_I$ is thus available for this function:

$$\forall \alpha \in [0,1],\ \forall k \in [0,K],\ \forall x \in \mathbb{R},\ u_b(t_k, x, G^{-1}(\alpha)) = \sum_I \beta_I^{(b,t_k)}(x)\psi_I(\alpha).$$

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

$$\forall a \in \Theta,\ \forall k \in [0,K],\ \forall x \in \mathbb{R},\ u_b(t_k, x, a) = \sum_I \beta_I^{(b,t_k)}(x)\psi_I(G(a)). \tag{18}$$

The coefficients $\beta_I^{(b,t_k)}(x)$ can be expressed with respect to these two equivalent expansions:

$$\beta_I^{(b,t_k)}(x) = \int_{[0,1]} \psi_I(\alpha) u_b(t_k, x, G^{-1}(\alpha)) d\alpha = \int_{\mathbb{R}} \psi_I(G(a)) u_b(t_k, x, a)\nu(da). \tag{19}$$

Using (18) and (19), we propose the following regression estimator for $\widehat{u_b}$ inspired by Kerkyacharian and Picard [18] (To see how this estimator has been built, refer to Appendix B): $\forall a \in \Theta$, $\forall k \in [0,K]$, $\forall x \in \mathbb{R}$,

$$\widehat{u_b}(t_k, x, a) \quad = \quad \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(x)\psi_I(G_N(a)), \tag{20}$$

with:

$$\widehat{\beta}^{(b,t_k)}(x) = \sum_{j=1}^{N} \frac{1}{N}\psi_I(G_N(\theta_j))\, b\left(x, \bar{X}_{t_k}^{j,N}(\theta_j)\right), \text{ and } \widetilde{\beta}_I^{(b,t_k)}(x) = thr\left(\widehat{\beta}^{(b,t_k)}(x), t_N\right), \tag{21}$$

where $thr(x,\lambda) = x 1_{|x|>\lambda}$ is the thresholding function, $t_N$, the threshold level and $I_1^N$, the resolution level.

The expression of estimators (21) is simpler when we rank the couples $\left(\theta_j, \bar{X}_{t_k}^{j,N}(\theta_j)\right)_{j\in[1,N]}$ in increasing order of $(\theta_j)_{j\in[1,N]}$. The ranked sequence of $(\theta_j)_{j\in[1,N]}$ is denoted by $(\theta_{(j)})_{j\in[1,N]}$. The ranked couples are noted $(\theta_{(j)}, \bar{X}_{t_k}^{(j),N}(\theta_{(j)}))_{j\in[1,N]}$ with the index $j$ between parenthesis. For any $x \in \mathbb{R}$, estimators (21) can then be rewritten as:

$$\widehat{\beta}_I^{(b,t_k)}(x) = \sum_{j=1}^{N} \frac{1}{N}\psi_I\left(\frac{j}{N}\right) b\left(x, \bar{X}_{t_k}^{(j),N}(\theta_{(j)})\right). \tag{22}$$

When we plug the regression estimators (20) in the particle system defined in (13), with $\bar{X}_{t_k}^{i,N}(\theta_i)$ as $x$ and $\theta_i$ as $a$, for the regressions defining the path of $\bar{X}_{t_k}^{i,N}(\theta_i)$, we obtain the particle system on which Method 3 is based:

**Definition 3.** *Let $(\theta_i)_{i\in[1,N]}$ be i.i.d. variables of law $\nu$, let $\left(\bar{X}_0^{i,N}(\theta_i)\right)_{i\in[1,N]}$ be independent random variables of initial laws $p_0(x,\theta_i)dx$, and let $W = (W^1, \cdots, W^N)$ be a $N$-dimensional Brownian motion independent of the variables $(\theta_i)_{i\in[1,N]}$ and of $(\bar{X}_0^{i,N}(\theta_i))_{i\in[1,N]}$. Introduce the coefficients $\widetilde{\beta}_I^{(b,t_k)}(x)$ and $\widetilde{\beta}_I^{(\sigma,t_k)}(x)$ as in (21). We can define the following particle system: $\forall i \in [1,N]$, $\forall k \in [0,K]$,*

$$\begin{aligned} \bar{X}_{t_{k+1}}^{(i),N}(\theta_{(i)}) \quad &= \quad \bar{X}_{t_k}^{(i),N}(\theta_{(i)}) + \textstyle\sum_{I_1=-1}^{I_1^N} \sum_{I_2\in\mathbb{Z}} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}(\theta_{(i)}))\psi_I\left(\frac{i}{N}\right)\Delta t \\ &+ \quad \textstyle\sum_{I_1=-1}^{I_1^N} \sum_{I_2\in\mathbb{Z}} \widetilde{\beta}_I^{(\sigma,t_k)}(\bar{X}_{t_k}^{(i),N}(\theta_{(i)}))\psi_I\left(\frac{i}{N}\right)\left(W_{t_{k+1}}^i - W_{t_k}^i\right). \end{aligned} \tag{23}$$

◊

In the sequel, it will be useful to notice that the sequence of particles $((\bar{X}^{(i),N}_{t_k}(\theta_{(i)}))_{i\in[1,N]})_{k\in[1,K]}$ for the discretization times $t_k$ can be considered as a $\mathbb{R}^N$-Markov chain with transition kernel $S$ defined by:

$$\forall x \in \mathbb{R}^N,\ \forall f \in \mathcal{C}_b(\mathbb{R}^N, \mathbb{R}),$$

$$\begin{aligned} Sf(x) \;=\; & \mathbb{E}\Bigg[ f\Bigg( x + \left( \frac{1}{N}\sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right) 1_{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N}) b(x_i,x_j)|>t_N} b(x_i,x_j)\Delta t \right)_{1\leq i\leq N} \\ & + \left( \frac{1}{N}\sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right) 1_{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N}) \sigma(x_i,x_j)|>t_N} \sigma(x_i,x_j)\sqrt{\Delta t} B_i \right)_{1\leq i\leq N} \Bigg)\Bigg] \end{aligned} \tag{24}$$

where the $(B_i)_{i\in[1,N]}$ are independent centered reduced Gaussian variables.

Our main result deals with the convergence rate of this method:

**Theorem 1.2.** *Consider the following assumptions:*

1. *The functions* $b,\ \sigma \in \mathcal{C}_b^{4+\varepsilon}(\mathbb{R}^2, \mathbb{R})$,
2. *The law* $\nu$ *of* $\theta$ *admits a density* $g$ *on* $\mathbb{R}$ *w.r.t. the Lebesgue measure. The support* $\Theta$ *of* $g$ *is assumed to be connected, and can be* $\mathbb{R}$ *itself,*
3. $P(d\omega)-a.s.$, $p_0(.,\theta)$ *is a density function with moments of order 2,*
4. *The marginal law in* $x$ *of* $p_0(x,a)g(a)da\,dx$, *satisfies a Poincaré inequality with the positive deterministic constant* $c_0$:

$$\forall f \in \mathcal{C}^1(\mathbb{R},\mathbb{R}),\ \int_{\mathbb{R}} f^2(x)\left(\int_{\mathbb{R}} p_0(x,a)g(a)da\right)dx - \left(\int_{\mathbb{R}} f(x)\left(\int_{\mathbb{R}} p_0(x,a)g(a)da\right)dx\right)^2$$
$$\leq c_0 \int_{\mathbb{R}} |\nabla f(x)|^2 \left(\int_{\mathbb{R}} p_0(x,a)g(a)da\right)dx,$$

5. *The application* $\Phi\circ G^{-1}\ :\ \alpha\in[0,1]\mapsto p_0(.,G^{-1}(\alpha))$ *is s-Hölder continuous for the norm in* $L^1(\mathbb{R})$, *with* $s>1/2$, *and* $G$ *the distribution function of* $\theta$,
6. *The father and mother wavelets* $\phi$ *and* $\psi$ *used to construct the estimators* $\widehat{u_b}$ *and* $\widehat{u_\sigma}$ *are compactly supported, Lipschitz continuous, and satisfy Assumptions (H) and (M),*
7. *We threshold the estimators with* $t_N=\kappa(\log N)/\sqrt{N}$ *where* $\kappa$ *is a positive constant that depends only on* $b$, $\sigma$, $c_0$, $\phi$, $\psi$, $T$ *and* $\Delta t$ *(see (55) in the proof for more information on the choice of* $\kappa$*),*
8. *The resolution level* $I_1^N$ *satisfies* $2^{I_1^N}\sim\frac{\sqrt{N}}{\log N}$,

*Then the particle system (23) is well defined and satisfies:*

$$\forall 0<\varepsilon<1,\ \forall f\in\mathcal{C}_b^{4+\varepsilon},\ \exists C=C(T,f,b,\sigma)>0,\ \exists N_0,\ \forall N\geq N_0,$$

$$\mathbb{E}^{\nu}|\langle I(m_T),f\rangle - \frac{1}{N}\sum_{i=1}^N f(\bar{X}_T^{i,N}(\theta_i))| \;\leq\; C\left(\Delta t + \log N\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{1+2s}}\right).$$

□

The asymptotic convergence rate of this method, still slower than the one of Method 1, is better than the one of Method 2, given by Proposition 1.4. The Nadaraya-Watson estimators depend indeed linearly on the empirical measure of the particle system and do not always attain minimax rates (see Härdle *et al.* [14] section 10.4, Donoho *et al.*

[11, 12] or Kerkyacharian and Picard [17]). Since they are nonlinear, the Kerkyacharian-Picard estimators can achieve better convergence rates. If $\Phi \circ G^{-1}$ is Lipschitz continuous, we have a rate slightly slower than $N^{-1/3}$ which is more accurate than the preceding $N^{-1/8}$.

The assumptions of Theorem 1.2 are also weaker than those of Proposition 1.4.

First, Theorem 1.2 allows weaker regularities for $G$ and $\Phi$, which characterize the alea of the initial condition. For instance, when the law $\nu$ of $\theta$ admits a positive Lipschitz density $g$ with respect to the Lebesgue measure on a compact set $\Theta$, as in Proposition 1.4, it suffices that $\Phi$ be $s$-Hölder continuous to achieve the assumption of Theorem 1.2, which is weaker than $\Phi$ Lipschitz in Proposition 1.4.

Secondly, in Theorem 1.2, the support of the law $\nu$ does not need to be a compact interval of $\mathbb{R}$ any more, as it was required in Proposition 1.4. A much larger class of probability laws, including Gaussian laws and Gaussian mixing in particular, is then available to parameterize the randomness of the initial condition $\widehat{p_0}$.

Notice moreover that the optimal choice of the window $h_N$ in Method 2 (see Proposition 1.4) usually depends on the regularities of $\Phi$ and $g$, which may be unknown (see Bosq and Lecoutre [4]). The wavelet regression estimator introduced by Kerkyacharian and Picard [18] is adaptive: it adjusts itself automatically to the unknown regularity of the regression function we estimate. Here, the parameters $t_N$ and $I_N^1$ only depend on the number of observations and on basic bounds for $c_0$, $b$ and $\sigma$ (and not on the regularity $s$ of $\Phi \circ G^{-1}$).

Numerically, this method is more demanding in computations than Method 2, since the regression estimators have to be recomputed entirely at each date, because of the thresholding procedure. Still, if we use for instance the Mallat cascade algorithm, which complexity is of order $O(N)$ (see Härdle *et al.* [14], Chapter 12, page 223, or Mallat [21] Chapter VII, sections 7.3 and 7.5), the complexity of Method 3 remains in $O\left(N^2/\Delta t\right)$, which is better than Method 1.

## 2 Convergence Rate of the Wavelet Particle Approximation

In this section, we will prove Theorem 1.2, which gives the asymptotic convergence rate of the wavelet approximation (14) constructed on the particle system (23). We thus work under the assumptions of this theorem.

Let us first introduce some objects we will use for the proof. The Euler scheme associated with SDE (5) is given by: $\forall k \in [0, K]$,

$$\begin{cases} \bar{X}_{t_{k+1}}(\theta) = \bar{X}_{t_k}(\theta) + u_b(t_k, \bar{X}_{t_k}(\theta), \theta)\Delta t + u_\sigma(t_k, \bar{X}_{t_k}(\theta), \theta)\left(W_{t_{k+1}} - W_{t_k}\right) \\ \bar{X}_0 = X_0. \end{cases} \tag{25}$$

Consider as well i.i.d. copies of this Euler scheme coupled with the particles defined in (23) (same parameters $(\theta_{(i)})_{i\in[1,N]}$, same initial conditions $(\bar{X}_0^{i,N}(\theta_{(i)}))_{i\in[1,N]}$ and same Brownian motions $(W^i)_{i\in[1,N]}$), and which are also ranked in increasing order of $(\theta_i)_{i\in[1,N]}$: $\forall i \in [1, N]$, $\forall k \in [0, K]$,

$$\begin{cases} \bar{X}^{(i)}_{t_{k+1}}(\theta_{(i)}) = \bar{X}^{(i)}_{t_k}(\theta_{(i)}) + u_b(t_k, \bar{X}^{(i)}_{t_k}(\theta_{(i)}), \theta_{(i)})\Delta t + u_\sigma(t_k, \bar{X}^{(i)}_{t_k}(\theta_{(i)}), \theta_{(i)})\left(W^i_{t_{k+1}} - W^i_{t_k}\right) \\ \bar{X}^{(i)}_0 = \bar{X}^{(i),N}_0. \end{cases} \tag{26}$$

Since the (unordered) particles $\left(\theta_i, \bar{X}^{i,N}_{t_k}(\theta_i)\right)_{i\in[1,N]}$ defined with the regression estimators (21) are exchangeable, they have a common law. We will call $\mathcal{L}(\bar{X}^{\cdot,N}_{t_k}(\theta))$ the common first marginal law, and denote by $\mathbb{E}^\nu\left(b(x, \bar{X}^{\cdot,N}_{t_k}(\theta))|\theta\right)$ the expectation of $b(x,.)$ under this law conditioned by the random variable $\theta$. Notice that the conditional laws $\mathcal{L}\left(\bar{X}^{i,N}_{t_k}(\theta_i) \mid \theta_i\right)$ may be all different. This makes the difficulty.

In the developments, we will also need the wavelet decomposition of the function $a \in \Theta \mapsto \bar{u_b}(t_k, x, a) := \mathbb{E}^\nu\left(b(x, \bar{X}^{\cdot,N}_{t_k}(\theta))|\theta = a\right)$ on the warped wavelet basis $(\psi_I \circ G)_I$:

$$\forall a \in \Theta,\, \bar{u_b}(t_k, x, a) = \sum_I \bar{\beta}^{(b,t_k)}_I(x)\psi_I(G(a)), \tag{27}$$

with:

$$\bar{\beta}^{(b,t_k)}_I(x) = \int \psi_I(G(a))\mathbb{E}^\nu\left(b(x, \bar{X}^{\cdot,N}_{t_k}(\theta)) \mid \theta = a\right)\nu(da). \tag{28}$$

**Remark 5.** *From now on, we will write with a notational abuse* $\bar{X}^{(i)}_{t_k}$ *and* $\bar{X}^{(i),N}_{t_k}$ *for the particles ranked in increasing order of* $(\theta_i)_{i\in[1,N]}$*, instead of* $\bar{X}^{(i)}_{t_k}(\theta_{(i)})$ *and* $\bar{X}^{(i),N}_{t_k}(\theta_{(i)})$ *respectively.* □

*Proof of Theorem 1.2.* We can decompose the approximation error at time $T$ in three sources:

$$\langle I(m_T), f\rangle - \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i),N}_T\right) = (1) + (2) + (3), \tag{29}$$

where:

$$\begin{aligned}
(1) &= \langle I(m_T), f\rangle - \mathbb{E}^\nu\left(f(\bar{X}_T(\theta))\right) \\
(2) &= \mathbb{E}^\nu\left(f(\bar{X}_T(\theta))\right) - \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i)}_T\right) \\
(3) &= \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i)}_T\right) - \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i),N}_T\right).
\end{aligned}$$

The discretization error (1) can be upper bounded thanks to the following result due to Talay and Vaillant ([33], Proposition 5.1), which generalizes a result from Talay and Tubaro [32]:

**Theorem 2.1.** *(Talay and Vaillant [33]) For $b$ and $\sigma$ in $\mathcal{C}^{4+\varepsilon}_b(\mathbb{R}^2, \mathbb{R})$:*

$$\forall f \in \mathcal{C}^{4+\varepsilon}_b(\mathbb{R}, \mathbb{R}),\, \exists C = C(T, f, b, \sigma) > 0,\, \forall a \in \Theta,\, |\mathbb{E}f(X_T(a)) - \mathbb{E}f(\bar{X}_T(a))| \le C\,\Delta t. \tag{30}$$

*The constant $C = C(T, f, b, \sigma)$ can be upper bounded uniformly in $a$ by a sum of terms of type $||\partial^{(i)}_x b||_\infty ||\partial^{(j)}_x \sigma||_\infty ||f^{(k)}||_\infty$ for $0 \le i, j, k \le 4$. Integrating over $a$, we obtain:*

$$\forall f \in \mathcal{C}^{4+\varepsilon}_b(\mathbb{R}, \mathbb{R}),\, \exists C = C\,(T, f, b, \sigma) > 0,\, |\mathbb{E}^\nu\left(f(X_T(\theta))\right) - \mathbb{E}^\nu\left(f(\bar{X}_T(\theta))\right)| \le C\,\Delta t. \tag{31}$$

□

The $L^1$-norm of the statistical error (2) can be upper bounded thanks to the CLT:

$$\exists C > 0,\, \mathbb{E}^\nu\left|\mathbb{E}^\nu\left(f(\bar{X}_T(\theta))\right) - \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i)}_T\right)\right| \le \frac{C||f||_\infty}{\sqrt{N}}. \tag{32}$$

Now let us focus on Term (3). This term appears because we use approximated coefficients to overcome nonlinearity. The difficulty here is that the classical propagation of chaos is no longer available since the particles $\bar{X}^{(i),N}$ are no more in mean-field interactions.

From the definition, for any $i \in [1,N]$ and any $k \in [0,K]$:

$$\begin{aligned}\bar{X}^{(i)}_{t_{k+1}} - \bar{X}^{(i),N}_{t_{k+1}} &= \bar{X}^{(i)}_{t_k} - \bar{X}^{(i),N}_{t_k} + \left(u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\Delta t \\ &+ \left(u_\sigma(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\left(W^i_{t_{k+1}} - W^i_{t_k}\right).\end{aligned}$$

Taking the square, then the expectation:

$$\begin{aligned}&\mathbb{E}^\nu\left(\mid \bar{X}^{(i)}_{t_{k+1}} - \bar{X}^{(i),N}_{t_{k+1}} \mid^2\right) \\ =\ &\mathbb{E}^\nu\left(\mid \bar{X}^{(i)}_{t_k} - \bar{X}^{(i),N}_{t_k} \mid^2\right) + \mathbb{E}^\nu\left(\mid u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) \mid^2\right)\Delta t^2 \\ +\ &\mathbb{E}^\nu\left(\mid u_\sigma(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) \mid^2\right)\Delta t \\ +\ &2\Delta t\,\mathbb{E}^\nu\left(\left(\bar{X}^{(i)}_{t_k} - \bar{X}^{(i),N}_{t_k}\right)\left(u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\right) \\ +\ &2\mathbb{E}^\nu\left(\left(\bar{X}^{(i)}_{t_k} - \bar{X}^{(i),N}_{t_k}\right)\left(u_\sigma(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\left(W^i_{t_{k+1}} - W^i_{t_k}\right)\right) \\ +\ &2\Delta t\mathbb{E}^\nu\left(\left(u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\left(u_\sigma(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right)\right. \\ &\left.\left(W^i_{t_{k+1}} - W^i_{t_k}\right)\right).\end{aligned} \tag{33}$$

The last two terms are zero. This can be seen when conditioning by $\mathcal{F}_{t_k} \vee \sigma(\theta_i, i \in [1,N])$. Given that $2ab \leq a^2 + b^2$, an upper bound of the fourth term in the right hand side of (33) is:

$$\Delta t\left[\mathbb{E}^\nu\left(|\bar{X}^{(i)}_{t_k} - \bar{X}^{(i),N}_{t_k}|^2\right) + \mathbb{E}^\nu\left(\mid u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) \mid^2\right)\right].$$

Moreover:

$$\begin{aligned}&\mathbb{E}^\nu\left(\left|u_b(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right|^2\right) \\ \leq\ &2\left[\mathbb{E}^\nu\left(\left|\mathbb{E}^\nu\left(b(x,X_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}}\right|^2\right)\right. \\ +\ &\left.\mathbb{E}^\nu\left(\left|\mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right|^2\right)\right].\end{aligned}$$

The expectation $\mathbb{E}^\nu\left(\left|u_\sigma(t_k,\bar{X}^{(i)}_{t_k},\theta_{(i)}) - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right|^2\right)$ can be treated similarly.

Let us introduce the notation:

$$S_N(t_{k+1}) = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}^\nu|\bar{X}^{(i)}_{t_{k+1}} - \bar{X}^{(i),N}_{t_{k+1}}|^2.$$

By summation, and for $\Delta t \leq 1$, we deduce from (33):

$$\forall k \in [0,K],\ S_N(t_{k+1}) \leq (1+\Delta t)S_N(t_k) + C\,\Delta t\left(\frac{1}{N}\sum_{i=1}^{N}A_1(i,t_k) + \frac{1}{N}\sum_{i=1}^{N}A_2(i,t_k)\right), \tag{34}$$

where $A_1$ and $A_2$ are given by: $\forall i \in [1,N]$, $\forall k \in [0,K]$,

$$\begin{aligned}A_1(i,t_k) &= \mathbb{E}^\nu \mid \mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \widehat{u_b}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) \mid^2 \\ &+ \mathbb{E}^\nu \mid \mathbb{E}^\nu\left(\sigma(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \widehat{u_\sigma}(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) \mid^2 \qquad (35) \\ A_2(i,t_k) &= \mathbb{E}^\nu \mid \mathbb{E}^\nu\left(b(x,X_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} \mid^2 \\ &+ \mathbb{E}^\nu \mid \mathbb{E}^\nu\left(\sigma(x,X_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} - \mathbb{E}^\nu\left(\sigma(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}} \mid^2 \qquad (36)\end{aligned}$$

To complete the computation of an upper bound for the $L^1$-norm of Term (3) in (29), we need to consider the terms $A_1$ and $A_2$ defined in (35) and (36). Since the terms in $b$ and in $\sigma$ are similar, we do the computations only for the terms with $b$.

Term $A_2$ results from the discretization step and can be handled with arguments similar to those used to obtain (30). We replace $T$ with $t_k$, $f$ with $b\left(\bar{X}^{(i)}_{t_k}(\theta_{(i)}),.\right)$, and choose $\theta_{(i)}$ as $a$. Summing over the particles gives: $\frac{1}{N}\sum_{i=1}^{N} A_2(i,t_k) \leq C\,(\Delta t)^2$.

Now, turn to Term $A_1$. Because of the nonlinearity of the wavelet regression estimator, we do not use coupling methods as in Sznitman [31] or Méléard [25]. We choose to work at once on the part of the error linked to the estimation:

$$\mathbb{E}^{\nu}\left(\left|\widehat{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) - \mathbb{E}^{\nu}\left(b\left(x,\bar{X}_{t_k}(\theta)\right)|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}}\right|^2\right) \leq C\left(\mathbb{E}^{\nu}(A)^2+\mathbb{E}^{\nu}(B)^2+\mathbb{E}^{\nu}(C)^2\right), \quad (37)$$

where:

$$\begin{aligned}
(A) &= \widehat{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) - \bar{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\\
(B) &= \bar{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)}) - \mathbb{E}^{\nu}\left(b\left(x,\bar{X}_{t_k}(\theta)\right)|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i),N}_{t_k}}\\
(C) &= \mathbb{E}^{\nu}(b(x,\bar{X}_{t_k}(\theta))|\theta=\theta_{(i)})\,|_{x=\bar{X}^{(i),N}_{t_k}} - \mathbb{E}^{\nu}\left(b\left(x,\bar{X}_{t_k}(\theta)\right)|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i)}_{t_k}}.
\end{aligned}$$

Since Term $(B)$ can be rewritten as:

$$(B) = \mathbb{E}^{\nu}\left(b\left(x,\bar{X}^{\cdot,N}_{t_k}(\theta)\right) - b\left(x,\bar{X}_{t_k}(\theta)\right)|\theta=\theta_{(i)}\right)|_{x=\bar{X}^{(i),N}_{t_k}},$$

we have:

$$\mathbb{E}^{\nu}(B)^2 \leq L\,\mathbb{E}^{\nu}(|\bar{X}^{(i),N}_{t_k} - \bar{X}^{(i)}_{t_k}|^2).$$

To deal with $(C)$, we use the Lipschitz property of $x \mapsto \mathbb{E}^{\nu}\left(b\left(x,\bar{X}_{t_k}(\theta)\right)|\theta=\theta_{(i)}\right)$. Thus:

$$\mathbb{E}^{\nu}(C)^2 \leq L\,\mathbb{E}^{\nu}(|\bar{X}^{(i),N}_{t_k} - \bar{X}^{(i)}_{t_k}|^2).$$

To upper bound $\mathbb{E}^{\nu}(A)^2$, we use the two following lemmas:

**Lemma 2.1.** *Consider the particle system defined in (23). Under the assumptions of Theorem 1.2, the coefficients* $\widehat{\beta}^{(b,t_k)}_I(x)$ *and* $\bar{\beta}^{(b,t_k)}_I(x)$*, defined in (22) and (28) respectively, satisfy:*$\exists N_0$, $C>0$, $\forall\gamma>0$, $\forall N\geq N_0$, $\forall i\in[1,N]$, $\forall I=(I_1,I_2)\in[-1,I_1^N]\times\mathbb{Z}$, $\forall k\in[0,K]$, $\exists\kappa=\kappa(\gamma,T,\Delta t,b,\sigma,\phi,\psi)$,

$$\mathbb{P}^{\nu}\left(|\widehat{\beta}^{(b,t_k)}_I(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}^{(b,t_k)}_I(\bar{X}^{(i),N}_{t_k})| \geq \frac{\kappa\log N}{2\sqrt{N}}\right) \quad \leq \quad \frac{C}{N^{\gamma}}. \quad (38)$$

□

This lemma is proved in Section 3. It can be noticed that in the sequel, we will choose $\gamma > 7/2$ (see Lemma 4.1). Then, the above property determines the constant $\kappa = \kappa(7/2,T,\Delta t,b,\sigma,\phi,\psi)$ that we should use for the wavelet thresholding procedure (see Equation (55) in the proof for the accurate expression of $\kappa$).

The second lemma we will need gives us the regularity $s$ of the function $\alpha\mapsto\bar{u}_b(t_k,x,G^{-1}(\alpha))$ that will eventually appear in the convergence rate of our method.

**Lemma 2.2.** *Under Assumptions* 5) *of Theorem 1.2, the application* $\alpha \in [0,1] \mapsto \bar{u}_b(t_k, x, G^{-1}(\alpha))$ *is s-Hölder continuous.* □

*Proof.* A sufficient condition is that $\Phi \circ G^{-1}$ is $s$-Hölder continuous, where $\Phi : a \mapsto p_0(.,a)$ and $G^{-1}$ is the quantile function for $\theta$.

To see this, notice that for $-1 \leq a_1 \leq a_2 \leq \cdots \leq a_N \leq 1$, the law of the particle system $(\bar{X}^{(i),N}_{t_k}(a_i))_{i\in[1,N]}$ is: $(p_0(.,a_1),\cdots,p_0(.,a_N))\, S^k$, where $S$ has been defined in (24). For any $j \in [1,N]$ and for the measurable bounded function $f_j : (y_1,\cdots,y_N) \in \mathbb{R}^N \mapsto b(x,y_j) \in \mathbb{R}$, we have: $\forall \alpha_1, \alpha_2 \in [0,1]$,

$$\begin{aligned} & \left|\mathbb{E}^\nu\left(b\left(x, \bar{X}^{(j),N}_{t_k}(G^{-1}(\alpha_1))\right)\right) - \mathbb{E}^\nu\left(b\left(x, \bar{X}^{(j),N}_{t_k}(G^{-1}(\alpha_2))\right)\right)\right| \\ \leq\ & ||p_0(.,G^{-1}(\alpha_1)) - p_0(.,G^{-1}(\alpha_2))||_{L^1}||S^k f_j||_\infty \\ \leq\ & C\,|\alpha_1 - \alpha_2|^s, \end{aligned}$$

thanks to Assumption 5) in Theorem 1.2.

Since this is true for every $j \in [1,N]$, we deduce that $\alpha \mapsto \mathbb{E}^\nu\left(b\left(x, \bar{X}^{\cdot,N}_{t_k}(G^{-1}(\alpha))\right)\right)$ is also $s$-Hölder continuous. ■

The following corollary then concludes the computation of an upper bound for $A_1$. It is proved in Section 4:

**Corollary 2.1.** *Let us consider the particle system defined in (23). Under the assumptions of Theorem 1.2, and with the choice of* $\kappa$ *for the threshold* $t_N$ *as in Lemma 2.1, the estimator* $\widehat{u}_b(t_k, \bar{X}^{(i),N}_{t_k}, \theta_{(i)})$ *defined in (20) satisfies:* $\exists N_0, C > 0, \forall N \geq N_0, \forall i \in [1,N], \forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}, \forall k \in [0,K]$,

$$\mathbb{E}^\nu\left(\left|\widehat{u}_b(t_k, \bar{X}^{(i),N}_{t_k}, \theta_{(i)}) - \bar{u}_b(t_k, \bar{X}^{(i),N}_{t_k}, \theta_{(i)})\right|^2\right) \leq C(\log N)^2\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}},$$

*where* $\bar{u}_b(t_k,x,a)$ *has been defined in (27) and is s-Hölder continuous thanks to Lemma 2.2.* □

Let us come back to the proof of Theorem 1.2. Gathering the upper bounds of $A_1$ and $A_2$ in (34), we deduce:

$$\begin{aligned} S_N(t_{k+1}) &= (1 + C\,\Delta t)S_N(t_k) + C\,\Delta t\left((\Delta t)^2 + (\log N)^2\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}\right) \\ &\leq (1 + C\,\Delta t)^{k+1}S_N(0) + \frac{(1 + C\,\Delta t)^k - 1}{C\,\Delta t}C\,\Delta t\left((\Delta t)^2 + (\log N)^2\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}\right). \end{aligned}$$

As $k \leq T/\Delta t$ and $\log(1+x) \leq x$ for any $x > -1$:

$$(1 + C\Delta t)^k = \exp\left(k\log(1 + C\Delta t)\right) \leq \exp(Ck\Delta t) \leq \exp(CT), \tag{39}$$

Thus, if we notice in addition that $S_N(0) = 0$:

$$S_N(t_{k+1}) \leq C(T)\left((\Delta t)^2 + (\log N)^2\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}\right).$$

We can now conclude the computation of an upper bound for the $L^1$-norm of Term (3) in (29). Using the fact that $f$ is Lipschitz continuous with constant $||f'||_\infty$ and the Cauchy-Schwarz inequality, we have:

$$\mathbb{E}^\nu\left|\frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i)}_T\right) - \frac{1}{N}\sum_{i=1}^N f\left(\bar{X}^{(i),N}_T\right)\right| \leq C(T)||f'||_\infty\left(\Delta t + \log N\left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{1+2s}}\right). \tag{40}$$

The result announced in Theorem 1.2 is finally obtained by gathering (31), (32) and (40). ■

# 3 Proof of Lemma 2.1

We have:

$$\begin{aligned}
&\mathbb{P}^{\nu}\left(\left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right| \geq \frac{\kappa \log N}{2\sqrt{N}}\right) \\
\leq\ &\mathbb{P}^{\nu}\left(\left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \mathbb{E}^{\nu}\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}}\right| \geq \frac{\kappa \log N}{4\sqrt{N}}\right) \\
+\ &\mathbb{P}^{\nu}\left(\left|\mathbb{E}^{\nu}\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}} - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right| \geq \frac{\kappa \log N}{4\sqrt{N}}\right). \qquad (41)
\end{aligned}$$

The first term in the right hand side of (41) is dealt with Lemma 3.1 which is proved in Paragraph 3.1. The second term vanishes for sufficiently large $N$, thanks to Lemma 3.2 which is proved in Paragraph 3.2. This concludes the proof of Lemma 2.1. ■

**Lemma 3.1.** *We consider the particle system (23) and work under the assumptions of Theorem 1.2. Then:* $\exists\, N_0, C > 0$, $\forall N > N_0$, $\forall i \in [1, N]$, $I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall k \in [0, K]$, $\forall \gamma > 0$, $\exists \kappa = \kappa(\gamma, T, \Delta t, b, \sigma, \phi, \psi)$ :

$$\mathbb{P}^{\nu}\left(\left|\widehat{\beta}_I^{(b,t_k)}\left(\bar{X}_{t_k}^{(i),N}\right) - \mathbb{E}^{\nu}\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}}\right| > \frac{\kappa \log N}{4\sqrt{N}}\right) \leq \frac{C}{N^{\gamma}}. \qquad (42)$$

□

**Lemma 3.2.** *We consider the particle system (23) and work under the assumptions of Theorem 1.2. In particular, we define* $t_N$ *and* $I_1^N$ *as in Theorem 1.2, with the constant* $\kappa$ *appearing in Lemma 3.1 for the thresholding. Then:* $\exists N_0$, $C > 0$, $\forall N > N_0$, $\forall i \in [1, N]$, $\forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall k \in [0, K]$ :

$$\mathbb{P}^{\nu} - a.s., \left|\mathbb{E}^{\nu}\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}} - \bar{\beta}_I^{(b,t_k)}\left(\bar{X}_{t_k}^{(i),N}\right)\right| \leq \frac{C}{\sqrt{N}}.$$

□

## 3.1 Proof of Lemma 3.1

The difficulty in establishing the inequality (42) of Lemma 3.1 lies in the fact that our data are dependent. This prevents us from using the same techniques as in Kerkyacharian and Picard [18]: Bernstein and Rosenthal inequalities (see Härdle *et al.* [14], page 239) do not hold any longer. We do not deal with a classical mixing case either since the addition of a supplementary particle changes the definition of the whole system $\left(\bar{X}_{t_k}^{(i),N}\right)_{i\in[1,N]}$. The idea is to prove a spectral gap inequality and to deduce a concentration of measure phenomenon.

Let us first recall some definitions (see Ane and al. [1], Bakry [2] or Ledoux [20] for further details).

**Definition 4.** ***Poincaré Inequality (or Spectral Gap Inequality)***
*A probability measure* $\mu \in \mathcal{P}(\mathbb{R}^N)$ *satisfies a Poincaré inequality with constant* $\rho > 0$ *if for every function* $f \in \mathcal{C}^1(\mathbb{R}^N, \mathbb{R})$*:*

$$Var_{\mu}(f) := \langle \mu, f^2 \rangle - (\langle \mu, f \rangle)^2 \leq \rho \langle \mu, |\nabla f|^2 \rangle.$$

□

**Remark 6.** *Let us give two examples:*

*The Gaussian measure on $\mathbb{R}^N$, $\mathcal{N}(m, \Sigma)$ satisfies a Poincaré inequality with constant $\rho > 0$ the greatest eigenvalue in absolute value of the covariance matrix $\Sigma$ (and thus also with all other real number greater than $\rho$)(see Ané and al. [1] page 10).*

*The Dirac measure $\delta_x$, $x \in \mathbb{R}^N$, satisfies a Poincaré inequality with any positive constant.* □

An interest in looking for a Poincaré inequality is that it implies the following Concentration Phenomenon:

**Corollary 3.1.** ***Concentration Phenomenon*** *(Ledoux, [20])*
*Let $\mu$ be a probability measure that satisfies a Poincaré inequality with constant $\rho > 0$. For every $r \geq 0$, for every Lipschitz application $f$ with Lipschitz constant $L_f$, we have:*

$$\exists C > 0,\ \mu\left(|f - \langle \mu, f \rangle| \geq r\right) \leq C \exp\left(-\frac{r}{2\, L_f \sqrt{\rho}}\right).$$

*The constant $C$ does not depend on $f$ nor on $\mu$.* □

Our purpose in the sequel is to show that the law of the particle system $\left(\bar{X}^{(i),N}_{t_k}\right)_{i \in [1,N]}$ satisfies at each discretization time $t_k$ ($k \in [0, K]$) a Poincaré inequality. Then, we will use the Concentration Phenomenon to obtain inequality (42).

### 3.1.1 A Technical Lemma

We first provide a technical lemma, which will be useful to prove that some of the quantities which will appear are well-defined. The difficulty lies in the fact that wavelets can not be uniformly upper bounded: their suprema $2^{I_1/2}||\psi||_\infty$ tend to $\infty$ when the resolution levels $I_1$ increase.

**Lemma 3.3.** *Assume that the father and mother wavelets $\phi$ and $\psi$ are Lipschitz continuous functions with compact support. Then, for any functions $\eta_i : (x, y) \in \mathbb{R}^2 \mapsto \eta_i(x, y) \in \mathbb{R}$, $i \in \{1, 2\}$, such that $\eta_2$ is bounded by $M$, and for the choice of $t_N = \kappa(\log N)/\sqrt{N}$ and $I_1^N$ such that $2^{I_1^N} \sim t_N^{-1}$, terms:*

$$\frac{1}{N}\sum_{j=1}^{N}\left|\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I(\alpha)\,\psi_I\left(\frac{j}{N}\right)1_{\{|\frac{1}{N}\sum_{j=1}^{N}\psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}}\eta_2(z,x_j)\right|,$$

*are bounded by $C \times M$, where $C$ is a constant, uniformly in $N \in \mathbb{N}$, $\kappa \in \mathbb{R}$, $\alpha \in [0, 1]$, $z \in \mathbb{R}$ and $x = (x_1, \cdots, x_N)$.* □

*Proof.* Let us write:

$$\frac{1}{N}\sum_{j=1}^{N}\left|\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I(\alpha)\,\psi_I\left(\frac{j}{N}\right)1_{\{|\frac{1}{N}\sum_{j=1}^{N}\psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}}\eta_2(z,x_j)\right| \leq T_1 + T_2,$$

with:

$$\begin{aligned}
T_1 &= \left|\sum_{j=1}^{N}\int_{(j-1)/N}^{j/N} K_{I_1^N}(\alpha, y)\,dy\eta_2(z, x_j)\right|,\\
T_2 &= \sum_{j=1}^{N}\left|\frac{1}{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I(\alpha)\,\psi_I\left(\frac{j}{N}\right)1_{\{|\frac{1}{N}\sum_{j=1}^{N}\psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}}\eta_2(z,x_j)\right.\\
&\quad \left. - \int_{(j-1)/N}^{j/N} K_{I_1^N}(\alpha, y)\,dy\eta_2(z, x_j)\right|,
\end{aligned}$$

where: $K_{I_1^N}(x,y) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I(x)\psi_I(y) 1_{\{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}}$ looks like the projection kernel on the space $V_{I_1^N}$ generated by $(\psi_I)_{-1\le I_1 \le I_1^N,\, I_2\in\mathbb{Z}}$ (see Appendix B).

To upper bound $T_1$, let us write:

$$\begin{aligned} T_1 &\le \sum_{j=1}^{N} \int_{(j-1)/N}^{j/N} \left| K_{I_1^N}(\alpha, y)\right| dy \, |\eta_2(z,x_j)| \\ &\le M \int_0^1 \sum_{I_1=-1}^{I_1^N} \sum_{I_2} |\psi_I(\alpha)||\psi_I(y)| dy \le MC, \end{aligned}$$

by using the Concentration Assumption $(H)$ (introduced in Paragraph 1.3.2).

Let us now work on $T_2$:

$$\begin{aligned} T_2 &\le \sum_{j=1}^{N} \int_{(j-1)/N}^{j/N} \left| \sum_{I_1=-1}^{I_1^N} \sum_{I_2\in\mathbb{Z}} \psi_I(\alpha)\, \psi_I\left(\frac{j}{N}\right) 1_{\{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}} - K_{I_1^N}(\alpha,y) \right| |\eta_2(z,x_j)|\, dy \\ &\le M \sum_{j=1}^{N} \int_{(j-1)/N}^{j/N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} |\psi_I(\alpha)| \left| \psi_I\left(\frac{j}{N}\right) - \psi_I(y) \right| |1_{\{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N})\eta_1(z,x_j)|>t_N\}}| dy \end{aligned}$$

Since $\psi$ is Lipschitz continuous (we denote by $L_\psi$ its Lipschitz norm), $\psi_{I_1,I_2}$ is Lipschitz continuous with constant $2^{3I_1/2}L_\psi$. Moreover, using compact supported wavelets implies that the sum on $I_2$ corresponds to a finite number $C$ of non nul terms that does not depend on $I_1$. Resuming the above computation gives:

$$\begin{aligned} T_2 &\le M \sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2\in\mathbb{Z}} |\psi_{I_1,I_2}(\alpha)|\, 2^{3I_1/2} L_\psi \left| \int_{j-1/N}^{j/N} \left(\frac{j}{N} - y\right) dy \right| \\ &\le M \sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} C\, 2^{I_1/2} ||\psi||_\infty 2^{3I_1/2} L_\psi \frac{1}{2N^2} \\ &\le C\, M\, \frac{2^{2I_1^N} ||\psi||_\infty L_\psi}{2N}. \end{aligned}$$

From the choice of $I_1^N$ such that $2^{I_1^N} \sim \frac{\sqrt{N}}{\kappa \log N}$, we have $\frac{2^{2I_1^N}}{N} \to 0$. ■

### 3.1.2 Spectral Gap Inequality for the Transition Kernel associated with a discretization step

**Lemma 3.4.** *For any $x \in \mathbb{R}^N$, for the transition kernel $S$ defined in (24), and for $\rho_1$ an upper bound of:*

$$\sup_{x\in\mathbb{R}^N} \max_{i\in[1,N]} \left( \left( \frac{1}{N} \sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I\left(\frac{i}{N}\right) \psi_I\left(\frac{j}{N}\right) 1_{|\frac{1}{N}\sum_{j=1}^N \psi_I(\frac{j}{N})\sigma(x_i,x_j)|>t_N} \sigma(x_i,x_j) \right)^2 \right),$$

*$\delta_x S$ satisfies a Poincaré inequality with constant $\rho_1 \Delta t$:*

$$\forall f \in \mathcal{C}_b^1(\mathbb{R}^N,\mathbb{R}),\; Var_{\delta_x S}(f) \le \rho_1\, \Delta t\, S\left(|\nabla f|^2\right)(x). \tag{43}$$

□

*Proof.* From (24), $\delta_x S$ is the law of a Gaussian variable with expectation:

$$x + \left( \frac{1}{N} \sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I \left( \frac{i}{N} \right) \psi_I \left( \frac{j}{N} \right) 1_{\left| \frac{1}{N} \sum_{j=1}^N \psi_I \left(\frac{j}{N}\right) b(x_i,x_j) \right| > t_N} b(x_i, x_j) \Delta t \right)_{1 \le i \le N},$$

and with covariance matrix the diagonal matrix:

$$diag \left[ \left( \frac{1}{N} \sum_{j=1}^{N} \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I \left( \frac{i}{N} \right) \psi_I \left( \frac{j}{N} \right) 1_{\left| \frac{1}{N} \sum_{j=1}^N \psi_I \left(\frac{j}{N}\right) \sigma(x_i,x_j) \right| > t_N} \sigma(x_i, x_j) \right)^2 \Delta t \right]_{1 \le i \le N}.$$

(Lemma 3.3 tells us that the terms in the expectation and of the covariance matrix are well-defined.)

The conclusion is a consequence of Remark 6. ■

### 3.1.3 Permutation Formula

We now want to obtain a Poincaré inequality for the transition kernels of our particle system (23) between 0 and any discretization time $t_k$. The reiteration we will consider in Paragraph 3.1.4 will let terms like $\delta_x |\nabla S^k f|^2$ appear. To deal with these terms and prove a Poincaré inequality for $\delta_x S^k$, an interversion formula of $\nabla$ and $S$ is established.

**Lemma 3.5.** *For the transition kernel $S$ defined in (24) and for any probability law $\mu(dx)$ on $\mathbb{R}^N$, we have for almost every real $\kappa$ (that appears in the definition $t_N = \kappa \frac{\log N}{\sqrt{N}}$):*

$$\mu(dx) - a.s., \exists \rho_2 > 0, \forall f \in \mathcal{C}_b^1(\mathbb{R}^N, \mathbb{R}), |\nabla S f(x)|^2 < \rho_2 \, S \, |\nabla f|^2 (x). \tag{44}$$

□

*Proof.* Let us compute $\nabla S f(x)$. Looking at the definition (24), we would like to take the derivatives under the expectation. The term under the expectation is however not differentiable everywhere since discontinuities originate from the thresholding.

However, the set $\left\{ \kappa > 0 \;\mid\; \mu \left( \left| \frac{1}{N} \sum_{j=1}^N \psi_I \left( \frac{j}{N} \right) b(x_i, x_j) \right| = \kappa \frac{\log N}{\sqrt{N}} \right) > 0 \right\}$ is at most countable. Let us choose $\kappa > 0$, such that $\mu \left( \left| \frac{1}{N} \sum_{j=1}^N \psi_I \left( \frac{j}{N} \right) b(x_i, x_j) \right| = \kappa \frac{\log N}{\sqrt{N}} \right) = 0$. Then, for $f$ bounded and smooth enough, we can $\mu(dx)$-almost surely take the derivatives under the expectation:

$$\begin{aligned} & |\nabla S f(x)| \\ \le \; & \mathbb{E} \left| A(x) \nabla f \left( x + \left( \frac{1}{N} \sum_{j=1}^N \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I \left(\frac{i}{N}\right) \psi_I \left(\frac{j}{N}\right) 1_{\left| \frac{1}{N} \sum_{j=1}^N \psi_I \left(\frac{j}{N}\right) b(x_i,x_j) \right| > t_N} b(x_i,x_j) \Delta t \right)_{1 \le i \le N} \right. \right. \\ + \; & \left. \left. \left( \frac{1}{N} \sum_{j=1}^N \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \psi_I \left(\frac{i}{N}\right) \psi_I \left(\frac{j}{N}\right) 1_{\left| \frac{1}{N} \sum_{j=1}^N \psi_I \left(\frac{j}{N}\right) \sigma(x_i,x_j) \right| > t_N} \sigma(x_i,x_j) \sqrt{\Delta t} B_i \right)_{1 \le i \le N} \right) \right|, \end{aligned} \tag{45}$$

where the $(B_i)_{i \in [1,N]}$ are centered and reduced Gaussian variables, and where $A(x)$ is a random matrix that can be decomposed into: $A(x) = Id + M(x) + D(x)$, with: $\forall (i,j) \in [1,N]$

$$
\begin{aligned}
M_{ij}(x) &= \frac{1}{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)b(x_i,x_j)|>t_N}\nabla_y b(x_i,x_j)\Delta t \\
&+ \frac{1}{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)\sigma(x_i,x_j)|>t_N}\nabla_y \sigma(x_i,x_j)\sqrt{\Delta t}B_i,
\end{aligned}
$$

and:

$$
\begin{aligned}
D(x) &= diag\Bigg[\frac{1}{N}\sum_{j=1}^{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)b(x_i,x_j)|>t_N}\nabla_x b(x_i,x_j)\Delta t \\
&+ \frac{1}{N}\sum_{j=1}^{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)\sigma(x_i,x_j)|>t_N}\nabla_x \sigma(x_i,x_j)\sqrt{\Delta t}B_i\Bigg]_{1\le i\le N}.
\end{aligned}
$$

Recall that: $\forall R \in \mathcal{M}_{N\times N}(\mathbb{R}), \forall v \in \mathbb{R}^N$, $|Rv|^2 \le \rho(RR^*)^2|v|^2$, where $\rho(RR^*)$ is the greatest eigenvalue of $RR^*$. Let us hence look for an upper bound for $\rho(A(x)A^*(x))$.

Using Lemma 3.3 and the fact that $b$ and $\sigma$ have bounded derivatives, we deduce the following upper bound:

$$
\sup_{x\in\mathbb{R}^N}\sup_{i\in[1,N]}|D_{ii}(x)| \le C\sup_{i\in[1,N]}(||\nabla_x b||_\infty\Delta t + ||\nabla_x\sigma||_\infty\sqrt{\Delta t}\,|B_i|).
$$

On the other hand, we are also able to control the eigenvalues of $M(x)$, thanks to Hadamard's Theorem (see [30]). Take $\lambda(x)$ the greatest eigenvalue of $M(x)$ in absolute value:

$$
\begin{aligned}
|\lambda(x)| &\le \sup_{x\in\mathbb{R}^N}\max_{i\in[1,N]}\sum_{j=1}^N|M_{ij}(x)| \\
&\le \sup_{x\in\mathbb{R}^N}\max_{i\in[1,N]}\left(\frac{1}{N}\sum_{j=1}^N\left|\sum_{I_1=-1}^{I_1^N}\sum_{I_2\in\mathbb{Z}}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)\nabla_y b(x_i,x_j)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)b(x_i,x_j)|>t_N}\Delta t\right|\right. \\
&+ \left.\frac{1}{N}\sum_{j=1}^N\left|\sum_{I_1=-1}^{I_1^N}\sum_{I_2\in\mathbb{Z}}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)\nabla\sigma_y(x_i,x_j)1_{|\frac{1}{N}\sum_{j=1}^N\psi_I\left(\frac{j}{N}\right)\sigma(x_i,x_j)|>t_N}\sqrt{\Delta t}B_i\right|\right). \quad (46)
\end{aligned}
$$

Lemma 3.3 again assures us that the spectre of $M(x)$ is uniformly controlled in $x$ by:

$$
\sup_{i\in[1,N]} C\left(||\nabla_y b||_\infty\Delta t + ||\nabla_y\sigma||_\infty\sqrt{\Delta t}\,|B_i|\right).
$$

The spectre of $A(x)A^*(x)$ is therefore controlled by:

$$
\sup_{i\in[1,N]}\left(1 + C\left((||\nabla_x b||_\infty + ||\nabla_y b||_\infty)\Delta t + (||\nabla_x\sigma||_\infty + ||\nabla_y\sigma||_\infty)\sqrt{\Delta t}\,|B_i|\right)\right)^2,
$$

which is almost surely finite. Moreover:

$$
\begin{aligned}
&\mathbb{E}\left(|\rho(A(x)A^*(x))|^2\right) \\
\le\ &\sup_{i\in[1,N]}\mathbb{E}\left(1 + C\left((||\nabla_x b||_\infty + ||\nabla_y b||_\infty)\Delta t + (||\nabla_x\sigma||_\infty + ||\nabla_y\sigma||_\infty)\sqrt{\Delta t}\,|B_i|\right)\right)^4 \\
\le\ &\rho_2 < \infty. \quad (47)
\end{aligned}
$$

Resuming the preceding computation (45) gives that $\mu(dx)-a.s.$:

$$
\begin{aligned}
&|\nabla Sf(x)| \\
\leq\ & \mathbb{E}^{\nu}\left[|\rho(A(x)A^*(x))|\left|\nabla f\left(x+\left(\frac{1}{N}\sum_{j=1}^{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(x_i,x_j\right)|>t_N}b(x_i,x_j)\Delta t\right)_{1\leq i\leq N}\right.\right.\right. \\
+\ & \left.\left.\left.\left(\frac{1}{N}\sum_{j=1}^{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right)1_{|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)\sigma\left(x_i,x_j\right)|>t_N}\sigma(x_i,x_j)\sqrt{\Delta t}B_i\right)_{1\leq i\leq N}\right)\right|\right] \\
\leq\ & \sqrt{\mathbb{E}^{\nu}\left(\rho^2(A(x)A^*(x))\right)}\sqrt{S\,|\nabla f|^2}\ \leq\ \sqrt{\rho_2}\sqrt{S\,|\nabla f|^2}.
\end{aligned}
$$

■

### 3.1.4 Spectral Gap Inequality for the Law of the Particle System at Discretization Time $t_k$

Now, we are ready to prove that, at each discretization time $t_k$, the law of the particle system (23) conditioned on its initial position $x\in\mathbb{R}^N$ satisfies a Poincaré inequality:

**Proposition 3.1.** *Consider the transition kernel $S$ defined in (24). Then:* $\exists\rho_1,\rho_2>0$, *for almost every* $\kappa>0$, $\forall x\in\mathbb{R}^N$, $\forall k\in[0,K]$, $\forall f\in\mathcal{C}_b^1(\mathbb{R}^N,\mathbb{R})$,

$$Var_{\delta_x S^k}(f)\leq\rho_1\,\Delta t\,\frac{1-\rho_2^k}{1-\rho_2}S^k\left(|\nabla f|^2\right)(x).$$

□

*Proof.* We follow here the computations in Talay and Malrieu [22].

$$
\begin{aligned}
Var_{\delta_x S^k}(f) &= S^k\left(f^2\right)(x)-\left(S^kf(x)\right)^2=\sum_{i=1}^{k}\left[S^i\left(\left(S^{k-i}f\right)^2\right)(x)-S^{i-1}\left(\left(S^{k-i+1}f\right)^2\right)(x)\right] \\
&= \sum_{i=1}^{k}S^{i-1}\left[S\left(\left(S^{k-i}f\right)^2\right)(x)-\left(S\left(S^{k-i}f\right)(x)\right)^2\right] \\
&= \sum_{i=1}^{k}S^{i-1}Var_{\delta_x S}\left(S^{k-i}f\right) \\
&\leq \sum_{i=1}^{k}S^{i-1}\rho_1\,\Delta t\,S\left(|\nabla S^{k-i}f|^2\right)(x) \qquad (48)\\
&\leq \sum_{i=1}^{k}\rho_1\,\Delta t\,\rho_2^{k-i}S^k\left(|\nabla f|^2\right)(x) \qquad (49)\\
&\leq \rho_1\,\Delta t\,\frac{1-\rho_2^k}{1-\rho_2}S^k\left(|\nabla f|^2\right)(x).
\end{aligned}
$$

Inequality (48) is obtained with Lemma 3.4 and Inequality (49) comes from the use of Lemma 3.5 with the measures $\mu=\delta_xS^i$, $i\in[1,k]$. ■

### 3.1.5 Generalization to Non-Dirac Initial Measures

Now, we generalize the result of Proposition 3.1 to the case where the initial condition at time $t=0$ is not a Dirac in $x\in\mathbb{R}^N$ but any probability measure $u_0\in\mathcal{P}(\mathbb{R}^N)$ (we have in mind the law of $\left(\bar{X}_0^{(i),N}\right)_{i\in[1,N]}$).

**Theorem 3.1.** *Let $u_0$ be a probability measure that satisfy a Poincaré inequality with constant $c_0$: $Var_{u_0} f \leq c_0 \int |\nabla f|^2 du_0$. Consider a transition kernel $\widetilde{S}$ on $\mathbb{R}^N$ satisfying a Poincaré inequality with constant $\rho > 0$ for every initial condition $x \in \mathbb{R}^N$: $Var_{\delta_x \widetilde{S}} f \leq \rho \widetilde{S}\left(|\nabla f|^2\right)(x)$, and for which the following permutation inequality is available: $u_0(dx) - a.s.$, $|\nabla \widetilde{S} f(x)|^2 \leq C S\left(|\nabla f|^2\right)(x)$.*
*Then, the measure $u_0 \widetilde{S}$ satisfies a Poincaré inequality with constant: $\rho + c_0 C$.* □

*Proof.* Take $f \in \mathcal{C}_b^1(\mathbb{R}^N, \mathbb{R})$. We have:

$$\begin{aligned} Var_{u_0 \widetilde{S}} f &= \langle u_0, \left(Var_{\delta_x \widetilde{S}} f\right)\rangle + Var_{u_0}\left(\widetilde{S} f\right) \\ &\leq \rho u_0 \widetilde{S}\left(|\nabla f|^2\right) + c_0 u_0 \left(|\nabla \widetilde{S} f|^2\right) \\ &\leq (\rho + c_0 C)\, u_0 \widetilde{S}\left(|\nabla f|^2\right). \end{aligned}$$

We used the spectral gap inequality for $u_0$ to obtain the first inequality and the permutation formula to obtain the second one. ■

Notice that in our numerical procedure, the particles at time $t = 0$ are independent. Hence, the following tensorization lemma and Assumption *4)* in Theorem 1.2 ensure that the law of $\left(\bar{X}_0^{(i),N}\right)_{i \in [1,N]}$ satisfies a spectral gap inequality with constant $c_0$:

**Lemma 3.6.** ***Tensorization****(see Ledoux [20]) Assume that the measures $\mu_1, \cdots, \mu_N$ satisfy Poincaré inequalities with constants $C_1, \cdots, C_N$ respectively. Then $\otimes_{i=1}^N \mu_i$ satisfies a Poincaré inequality with constant $C \leq \max_{i \in [1,N]} C_i$.* □

We can finally deduce the spectral gap inequality we were looking for:

**Corollary 3.2.** *Under the assumptions of Theorem 1.2, the law at discretization time $t_k$ of the particle system defined in (23) satisfies a spectral gap inequality with constant: $D_k = \rho_1 \Delta t \frac{1-\rho_2^k}{1-\rho_2} + c_0 \rho_2^k$.* □

### 3.1.6 End of Lemma 3.1's Proof

We use Corollary 3.2 and the Concentration Phenomenon (Corollary 3.1) to obtain a first inequality, that we will exploit to obtain Inequality (42) in Lemma 3.1.

**Lemma 3.7.** *We consider the particle system (23) and assume that the law of $(\bar{X}_0^{(i),N})_{i \in [1,N]}$ satisfies a Poincaré inequality with constant upper bounded by $c_0$. Then, we have: $\exists C > 0$, $\exists C' > 0$, $\forall k \in [0, K]$, $\forall N \in \mathbb{N}$, $\forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall x \in \mathbb{R}$, $\forall r > 0$,*

$$\mathbb{P}^\nu\left(\left|\widehat{\beta}_I^{(b,t_k)}(x) - \mathbb{E}^\nu\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\right| > r\right) \leq C \exp\left(-\frac{\sqrt{N} r}{2C' L \sqrt{D_k}}\right), \tag{50}$$

*where $D_k = \rho_1 \Delta t \frac{1-\rho_2^k}{1-\rho_2} + c_0\, \rho_2^k$ is a positive constant independent from $N$. (Recall that $L$ is the Lipschitz constant for $b$).* □

*Proof.* If $f : \mathbb{R} \mapsto \mathbb{R}$ is Lipschitz continuous with constant $L$, then the application: $F(x) = 1/N \sum_{j=1}^N \psi_I(j/N) f(x_j)$ from $\mathbb{R}^N$ to $\mathbb{R}$ is Lipschitz continuous with constant $C' L/\sqrt{N}$, where $C'$ is an upper bound for the terms $\sqrt{1/N \sum_{j=1}^N \psi_I^2(j/N)}$ which converge deterministically to $\sqrt{\int_{[0,1]} \psi_I^2(x)\, dx}$ upper bounded by 1, when $N$ grows to $\infty$.

We have indeed, using $f$ Lipschitz and Cauchy-Schwarz inequality:

$$\begin{aligned}
\left|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)f\left(x_j\right)-\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)f\left(y_j\right)\right| &\leq \frac{L}{\sqrt{N}}\sqrt{\frac{1}{N}\sum_{j=1}^{N}\psi_I^2\left(\frac{j}{N}\right)}\sqrt{\sum_{j=1}^{N}|x_j-y_j|^2}\\
&\leq \frac{C'L}{\sqrt{N}}|x-x'|.
\end{aligned}$$

The result then follows from the use of the Concentration Phenomenon (Corollary 3.1) with the Lipschitz function $F$ since Corollary 3.2 tells us that the law at time $t_k$ of the particle system satisfies a Poincaré inequality with constant $D_k$. ■

Now let us prove Lemma 3.1. Our purpose is to extend inequality (50) by replacing the parameter $x$ with the random position of particle $(i)$, $\bar{X}_{t_k}^{(i),N}$. To achieve this, we will show the tail bound somehow holds uniformly in $x$:

*Proof of Lemme 3.1.* Let $r>0$, and let us define the interval $\mathcal{K}=[\mathbb{E}^\nu\left(\bar{X}_{t_k}^{(i),N}\right)-\varsigma,\mathbb{E}^\nu\left(\bar{X}_{t_k}^{(i),N}\right)+\varsigma]$ with $\varsigma>0$. From Corollaries 3.1 and 3.2, we deduce that:

$$\mathbb{P}^\nu\left(\left|\bar{X}_{t_k}^{(i),N}-\mathbb{E}^\nu\left(\bar{X}_{t_k}^{(i),N}\right)\right|>\varsigma\right)\leq Ce^{-\frac{\varsigma}{2\sqrt{D_k}}}. \tag{51}$$

Since $\mathcal{K}$ is a compact interval, it can be covered by a finite number of balls $]x_l-\varrho,x_l+\varrho[$, with $(x_l)_{l\in[1,\ell]}$ a finite sequence of $\mathcal{K}$ and $\varrho>0$. It is possible to choose $\ell=[\varsigma/\varrho]+1$, where $[.]$ stands for the integer part. Then:

$$\begin{aligned}
&\mathbb{P}^\nu\left(\left|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(\bar{X}_{t_k}^{(i),N},\bar{X}_{t_k}^{(j),N}\right)-\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(x,\bar{X}_{t_k}^{(j),N}\right)\right)\Bigg|_{x=\bar{X}_{t_k}^{(i),N}}\right|>r\right)\\
\leq\ &\mathbb{P}^\nu\left(\left|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(\bar{X}_{t_k}^{(i),N},\bar{X}_{t_k}^{(j),N}\right)-\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(x,\bar{X}_{t_k}^{(j),N}\right)\right)\Bigg|_{x=\bar{X}_{t_k}^{(i),N}}\right|>r\right.\\
&\left.\text{and}\quad \bar{X}_{t_k}^{(i),N}\in\mathcal{K}\right)+\mathbb{P}^\nu\left(\bar{X}_{t_k}^{(i),N}\notin\mathcal{K}\right)\\
\leq\ &\sum_{l=1}^{\ell}\mathbb{P}^\nu\left(\left|\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(\bar{X}_{t_k}^{(i),N},\bar{X}_{t_k}^{(j),N}\right)-\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^{N}\psi_I\left(\frac{j}{N}\right)b\left(x,\bar{X}_{t_k}^{(j),N}\right)\right)\Bigg|_{x=\bar{X}_{t_k}^{(i),N}}\right|>r\right.\\
&\left.\text{and}\quad \left|\bar{X}_{t_k}^{(i),N}-x_l\right|\leq\varrho\right)+\mathbb{P}^\nu\left(\bar{X}_{t_k}^{(i),N}\notin\mathcal{K}\right).
\end{aligned}\tag{52}$$

For a given $l \in [1, \ell]$, we have:

$$
\begin{aligned}
& \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( \bar{X}_{t_k}^{(i),N}, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x, \bar{X}_{t_k}^{(j),N} \right) \right) \Bigg|_{x=\bar{X}_{t_k}^{(i),N}} \right| \\
\leq & \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) \left( b\left( \bar{X}_{t_k}^{(i),N}, \bar{X}_{t_k}^{(j),N} \right) - b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) \right) \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E}^{\nu} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) \right) \right| \\
+ & \left| \mathbb{E}^{\nu} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) \right) - \mathbb{E}^{\nu} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x, \bar{X}_{t_k}^{(j),N} \right) \right) \Bigg|_{x=\bar{X}_{t_k}^{(i),N}} \right| \\
\leq & \; 2C'L \left| x_l - \bar{X}_{t_k}^{(i),N} \right| \\
+ & \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E}^{\nu} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) \right) \right|, \qquad (53)
\end{aligned}
$$

where $C'$ is an upper bound for $\left(\sqrt{1/N \sum_{j=1}^{N} \psi_I^2(j/N)}\right)_{N \in \mathbb{N}}$. Thus, using Lemma 3.1:

$$
\begin{aligned}
& \mathbb{P}^{\nu} \Bigg( \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( \bar{X}_{t_k}^{(i),N}, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x, \bar{X}_{t_k}^{(j),N} \right) \right) \Bigg|_{x=\bar{X}_{t_k}^{(i),N}} \right| > r \\
& \quad \text{and} \quad \left| \bar{X}_{t_k}^{(i),N} - x_l \right| \leq \varrho \Bigg) \\
\leq & \; \mathbb{P}^{\nu} \left( \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x_l, \bar{X}_{t_k}^{(j),N} \right) \right) \right| > r - 2C'L\varrho \right) \\
= & \; \mathbb{P}^{\nu} \left( \left| \widehat{\beta}_I^{(b,t_k)}(x_l) - \mathbb{E}\left( \widehat{\beta}_I^{(b,t_k)}(x_l) \right) \right| > r - 2C'L\varrho \right) \\
\leq & \; C\, e^{-\frac{\sqrt{N}(r - 2C'L\varrho)}{2C'L\sqrt{D_k}}}.
\end{aligned}
$$

Thus:

$$
\begin{aligned}
& \mathbb{P}^{\nu} \left( \left| \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( \bar{X}_{t_k}^{(i),N}, \bar{X}_{t_k}^{(j),N} \right) - \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^{N} \psi_I \left( \frac{j}{N} \right) b\left( x, \bar{X}_{t_k}^{(j),N} \right) \right) \Bigg|_{x=\bar{X}_{t_k}^{(i),N}} \right| > r \right) \\
\leq & \; C \left( \left[ \frac{\varsigma}{\varrho} \right] + 1 \right) e^{-\frac{\sqrt{N}(r - 2C'L\varrho)}{2C'L\sqrt{D_k}}} + e^{-\frac{\varsigma}{2\sqrt{D_k}}}. \qquad (54)
\end{aligned}
$$

If we choose $\varrho = \frac{r}{4C'L}$, we can upper bound the right hand side of (54) by:

$$
f(\varsigma) = 4C'CL\frac{\varsigma}{r} e^{-\frac{\sqrt{N}r}{4C'L\sqrt{D_k}}} + e^{-\frac{\varsigma}{2\sqrt{D_k}}}.
$$

The derivative in $\varsigma$ of this bound vanishes for:

$$
\varsigma_0 = 2\sqrt{D_k} \log\left( \frac{r}{8C'CL\sqrt{D_k}} \right) + \frac{\sqrt{N}r}{2C'L},
$$

and:

$$
f(\varsigma_0) \leq C \left( \frac{8C'L\sqrt{D_k}}{r} \log\left( \frac{r}{8C'CL\sqrt{D_k}} \right) + 2\sqrt{N} + \frac{8C'CL\sqrt{D_k}}{r} \right) e^{-\frac{\sqrt{N}r}{4C'L\sqrt{D_k}}}.
$$

Let us finally choose $r = \kappa \log N/(4\sqrt{N})$, then:

$$\begin{aligned} f(\varsigma_0) &\leq \frac{C}{N^{\frac{\kappa}{16C'L\sqrt{D_k}}}}\left(\frac{24C'L\sqrt{D_k}\sqrt{N}}{\kappa\log N}\log\left(\frac{\kappa\log N}{24C'CL\sqrt{D_k}\sqrt{N}}\right)+2\sqrt{N}+\frac{24C'CL\sqrt{D_k}\sqrt{N}}{\kappa\log N}\right) \\ &\leq \frac{C}{N^{\frac{\kappa}{16C'L\sqrt{D_k}}-\frac{1}{2}}}. \end{aligned}$$

Choosing:

$$\kappa > 16C'L\sqrt{D_k}\left(\gamma+\frac{1}{2}\right), \tag{55}$$

allows us to obtain the announced result. ■

## 3.2 Proof of Lemma 3.2

Recall that $\bar{\beta}_I^{(b,t_k)}(x)$ and $\widehat{\beta}_I^{(b,t_k)}(x)$ are defined in (28) and (22) respectively. We can rewrite them as:

$$\begin{aligned} \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) &= \mathbb{E}^\nu\left(\frac{1}{N}\sum_{j=1}^N \psi_I\left(G(\theta_j)\right)\mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}^{\cdot;N}(\theta))\ \mid\ \theta=\theta_j\right)\right)\Bigg|_{x=\bar{X}_{t_k}^{(i),N}} \\ \mathbb{E}^\nu\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}} &= \mathbb{E}^\nu\left(\frac{1}{N}\sum_{j=1}^N \psi_I\left(G_N(\theta_j)\right)\mathbb{E}^\nu\left(b(x,\bar{X}_{t_k}^{\cdot;N}(\theta))\ \mid\ \theta=\theta_j\right)\right)\Bigg|_{x=\bar{X}_{t_k}^{(i),N}}. \end{aligned}$$

Then:

$$\begin{aligned} A &= \mathbb{E}^\nu\left(\widehat{\beta}_I^{(b,t_k)}(x)\right)\Big|_{x=\bar{X}_{t_k}^{(i),N}} - \bar{\beta}_I^{(b,t_k)}\left(\bar{X}_{t_k}^{(i),N}\right) \\ &= \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left(\left[\psi_I\left(G_N(\theta_j)\right)-\psi_I\left(G(\theta_j)\right)\right]\bar{u}_b(t_k,x,\theta_j)\right)\big|_{x=\bar{X}_{t_k}^{i,N}} \end{aligned}$$

The immediate use of the Lipschitz continuity of $\psi_I$ does not lead to the result we are looking for, because the Lipschitz constant $2^{3I_1/2}||\psi||_\infty$ is not counterbalanced by $\mathbb{E}^\nu||G_N - G||_\infty \leq \frac{C}{\sqrt{N}}$. We follow here some ideas in Kerkyacharian and Picard [18] and use more deeply the structure of wavelets. The idea is to take advantage of the regularity of $\alpha \mapsto \bar{u}_b(t_k, x, G^{-1}(\alpha))$ by an integration by parts-like formula.

To this purpose, let us recall that if $\psi$ is a wavelet with compact support, then, there exists a compact supported Lipschitz continuous function $\Psi$ such that: $\psi = \Delta_{-h}(\Psi) = \Psi(. - h) - \Psi(.)$, with $h = 2^{-1}$. Thus:

$$\psi_I = \Delta_{-h_I}(\Psi_I),\ \text{ with } h_I = 2^{-I_1-1}\ \text{ and } \Psi_I(y) = 2^{I_1/2}\Psi(2^{I_1}y - I_2).$$

Let us also introduce the following notations:

$$\begin{aligned} U_N(\alpha) &= \frac{1}{N}\sum_{i=1}^N 1_{]-\infty,\alpha]}(G(\theta_i))(= G_N(G^{-1}(\alpha)).) \\ U_N^{(-j)}(\alpha) &= \frac{1}{N}\sum_{i=1,i\neq j}^N 1_{]-\infty,\alpha]}(G(\theta_i)) = U_N(\alpha) - \frac{1_{\{\alpha\geq G(\theta_j)\}}}{N}. \end{aligned}$$

Recall the Dvoretsky-Kiefer-Wolfowitz inequality $(DKW)$ (see [13]):

$$\exists K > 0,\ \forall\lambda > 0,\ \mathbb{P}^\nu\left(\sup_{\alpha\in[0,1]}|U_N(\alpha)-\alpha| \geq \lambda\right) \leq Ke^{-2N\lambda^2}. \tag{56}$$

Integrating this inequality in $\lambda$ gives:

$$\exists K > 0,\ \mathbb{E}^\nu\left(\sup_{\alpha\in[0,1]} |U_N(\alpha)-\alpha|\right) \leq K\sqrt{\frac{\pi}{2N}}. \tag{57}$$

We are now ready to prove Lemma 3.2. We first use the independence of the $(\theta_j)_{j\in[1,N]}$ and separate the term $A$ into two parts. The first one ($A_1$ in the sequel) allows us to carry out the integration by parts mentioned previously. The second one ($A_2$) is a residual term.

$$\begin{aligned} A &= \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left([\Delta_{-h_I}(\Psi_I)(U_N(G(\theta_j))) - \Delta_{-h_I}(\Psi_I)(G(\theta_j))]\,\mathbb{E}^\nu\left(b(x,\bar{X}^{\cdot,N}(\theta)) \mid \theta=\theta_j\right)\right)\Big|_{x=\bar{X}^{i,N}_{t_k}} \\ &= A_1 + A_2, \end{aligned}$$

with:

$$A_1 = \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left(\int_0^1 \Delta_{-h_I}\left(\Psi_I\left(U_N^{(-j)}(.)\right) - \Psi_I(.)\right)(\alpha)\bar{u}_b(t_k,x,G^{-1}(\alpha))\right)\Big|_{x=\bar{X}^{i,N}_{t_k}} d\alpha \tag{58}$$

$$\begin{aligned} A_2 &= \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left(\int_0^1 \left(\Psi_I(U_N^{(-j)}(\alpha)-h_I) - \Psi_I(U_N^{(-j)}(\alpha-h_I))\right)\bar{u}_b(t_k,x,G^{-1}(\alpha))d\alpha\right)\Big|_{x=\bar{X}^{i,N}_{t_k}} \\ &+ \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I\left(U_N^{(-j)}(\alpha)+\frac{1}{N}\right)\right)\bar{u}_b(t_k,x,G^{-1}(\alpha))d\alpha\right)\Big|_{x=\bar{X}^{i,N}_{t_k}} \end{aligned} \tag{59}$$

We first upper bound $A_1$ defined in (58). Using an integration by parts formula as well as the properties of MRAs on an interval (see Cohen *et al.* [7]), we have:

$$A_1 = \frac{1}{N}\sum_{j=1}^N \mathbb{E}^\nu\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right)\Delta_{h_I}\bar{u}_b(t_k,x,G^{-1}(.))(\alpha)d\alpha\right)\Big|_{x=\bar{X}^{i,N}_{t_k}},$$

Thanks to Lemma 2.2, the application $\alpha \mapsto \bar{u}_b(t_k,x,G^{-1}(\alpha))$ is $s$-Hölder continuous with $s > 1/2$ for every $x\in\mathbb{R}$. We thus have:

$$|A_1| \leq \frac{C\sqrt{h_I}}{N}\sum_{j=1}^N \left|\mathbb{E}^\nu\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right)d\alpha\right)\right|. \tag{60}$$

Let us introduce the following set:

$$B_N(\varsigma) = \left\{\sup_{\alpha\in[0,1]} |U_N(\alpha)-\alpha| \geq \varsigma\sqrt{\frac{\log N}{N}}\right\}, \tag{61}$$

where the constant $\varsigma$ has to be properly chosen (in the sequel, we will be lead to choose $\varsigma > \sqrt{3/8}$). Using DKW (56) yields:

$$\mathbb{P}^\nu\left(B_N(\varsigma)\right) \leq \frac{K}{N^{2\varsigma^2}}. \tag{62}$$

Then:

$$\left|\mathbb{E}^\nu\left(\int_0^1 \left(\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right)d\alpha\right)\right| \leq A_{11}^{(j)} + A_{12}^{(j)}, \tag{63}$$

where:

$$\begin{aligned} A_{11}^{(j)} &= \mathbb{E}^\nu\left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right| d\alpha\, 1_{B_N(\varsigma)}\right) \\ A_{12}^{(j)} &= \mathbb{E}^\nu\left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right| d\alpha\, 1_{B_N^c(\varsigma)}\right). \end{aligned}$$

Using (62), we obtain the following bound for $A_{11}^{(j)}$:

$$A_{11}^{(j)} \leq 2 \times 2^{I_1/2} ||\Psi||_\infty \mathbb{P}^\nu \left(B_N(\varsigma)\right) \leq \frac{2 \times 2^{I_1/2} ||\Psi||_\infty K}{N^{2\varsigma^2}}. \tag{64}$$

Now consider the term $A_{12}^{(j)}$. On the set $B_N(\varsigma)$, we have:

$$\sup_{\alpha \in [0,1]} |U_N^{(-j)}(\alpha) - \alpha| \leq \sup_{\alpha \in [0,1]} |U_N(\alpha) - \alpha| + \frac{1}{N} \leq \varsigma \sqrt{\frac{\log N}{N}} + \frac{1}{N}.$$

Thus, the support of $\alpha \mapsto \Psi\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)$ is included in a compact interval of length $C(\varsigma)$ depending only on $\varsigma$, $\Psi$ and $N$. For any double index $I \in [-1, I_1^N] \times \mathbb{Z}$, there exists an interval $\mathcal{I}_I$ of length $C(\varsigma)2^{-I_1}$ such that:

$$\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right| d\alpha \, 1_{B_N^c(\varsigma)} = \int_{\mathcal{I}_I} \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I(\alpha)\right| d\alpha \, 1_{B_N^c(\varsigma)}. \tag{65}$$

Thus:

$$\begin{aligned} A_{12}^{(j)} &\leq 2^{3I_1/2} L_\Psi \mathbb{E}^\nu \left(\int_{\mathcal{I}_I} \left|U_N^{(-j)}(\alpha) - \alpha\right| d\alpha \, 1_{B_N^c(\varsigma)}\right) \\ &\leq 2^{3I_1/2} L_\Psi \mathbb{E}^\nu \left(\sup_\alpha \left|U_N^{(-j)}(\alpha) - \alpha\right|\right) C(\varsigma) 2^{-I_1} \\ &\leq 2^{I_1/2} L_\Psi C(\varsigma) K \sqrt{\frac{\pi}{2N}}, \end{aligned} \tag{66}$$

where the third inequality has been obtained thanks to (57).

Therefore, from (60), (63), (64) and (66) we deduce:

$$|A_1| \leq C\left(\frac{1}{N^{2\varsigma^2}} + \frac{1}{\sqrt{N}}\right) \leq \frac{C}{\sqrt{N}}, \tag{67}$$

if we choose $\varsigma$ such that $\varsigma > 1/2$.

Now turn to the residual term $A_2$ defined in (59). We consider as before the set $B_N(\varsigma)$ defined in (61):

$$\begin{aligned} |A_2| &\leq \frac{||b||_\infty}{N} \sum_{j=1}^N \mathbb{E}^\nu \left(\int_0^1 \left|\Psi_I(U_N^{(-j)}(\alpha) - h_I) - \Psi_I(U_N^{(-j)}(\alpha - h_I))\right| d\alpha\right) \\ &+ \frac{||b||_\infty}{N} \sum_{j=1}^N \mathbb{E}^\nu \left(\int_0^1 \left|\Psi_I\left(U_N^{(-j)}(\alpha)\right) - \Psi_I\left(U_N^{(-j)}(\alpha) + \frac{1}{N}\right)\right| d\alpha\right) \\ &\leq \frac{||b||_\infty}{N} \sum_{j=1}^N \left[A_{21}^{(j)} + A_{22}^{(j)}\right], \end{aligned} \tag{68}$$

where:

$$
\begin{aligned}
A_{21}^{(j)} &= \mathbb{E}^{\nu}\left(\int_0^1 \left(\left|\Psi_I(U_N^{(-j)}(\alpha)-h_I)-\Psi_I(U_N^{(-j)}(\alpha-h_I))\right|\right.\right.\\
&+ \left.\left.\left|\Psi_I\left(U_N^{(-j)}(\alpha)\right)-\Psi_I\left(U_N^{(-j)}(\alpha)+\frac{1}{N}\right)\right|\right) d\alpha 1_{B_N(\varsigma)}\right)\\
A_{22}^{(j)} &= \mathbb{E}^{\nu}\left(\int_0^1 \left(\left|\Psi_I(U_N^{(-j)}(\alpha)-h_I)-\Psi_I(U_N^{(-j)}(\alpha-h_I))\right|\right.\right.\\
&+ \left.\left.\left|\Psi_I\left(U_N^{(-j)}(\alpha)\right)-\Psi_I\left(U_N^{(-j)}(\alpha)+\frac{1}{N}\right)\right|\right) d\alpha 1_{B_N^c(\varsigma)}\right).
\end{aligned}
$$

We have for $A_{21}^{(j)}$:

$$
A_{21}^{(j)} \leq 4\times 2^{I_1/2}||\Psi||_\infty \mathbb{P}^{\nu}\left(B_N(\varsigma)\right) \leq \frac{4\times 2^{I_1/2}||\Psi||_\infty K}{N^{2\varsigma^2}}. \tag{69}
$$

Now let us study the term $A_{22}^{(j)}$. Similarly to (65), we have:

$$
\begin{aligned}
A_{22}^{(j)} &= \mathbb{E}^{\nu}\left(\int_{\mathcal{I}_I} \left(\left|\Psi_I(U_N^{(-j)}(\alpha)-h_I)-\Psi_I(U_N^{(-j)}(\alpha-h_I))\right|\right.\right.\\
&+ \left.\left.\left|\Psi_I\left(U_N^{(-j)}(\alpha)\right)-\Psi_I\left(U_N^{(-j)}(\alpha)+\frac{1}{N}\right)\right|\right) d\alpha 1_{B_N^c(\varsigma)}\right)\\
&\leq 2^{3I_1/2}L_\Psi C(\varsigma)2^{-I_1}\sup_{\alpha\in[0,1]}\left(\mathbb{E}^{\nu}\left|U_N^{(-j)}(\alpha)-U_N^{(-j)}(\alpha-h_I)-\frac{N-1}{N}h_I\right|+\frac{h_I+1}{N}\right).
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}^{\nu}\left|U_N^{(-j)}(\alpha)-U_N^{(-j)}(\alpha-h_I)-\frac{N-1}{N}h_I\right| &\leq \sqrt{\mathbb{E}^{\nu}\left[\left(\frac{1}{N}\sum_{i=1,i\neq j}^{N}\left(1_{]\alpha-h_I,\alpha]}G(\theta_i)-h_I\right)\right)^2\right]}\\
&\leq \sqrt{\frac{1}{N}Var\left(1_{]\alpha-h_I,\alpha]}G(\theta)\right)}\\
&\leq \sqrt{\frac{h_I(1-h_I)}{N}}.
\end{aligned}
$$

Then, we finally have:

$$
\begin{aligned}
A_{22}^{(j)} &\leq 2^{I_1/2}L_\Psi C(\varsigma)\left(\sqrt{\frac{h_I(1-h_I)}{N}}+\frac{h_I+1}{N}\right)\\
&\leq C\left(\frac{1}{\sqrt{N}}+\frac{2^{I_1/2}}{N}\right).
\end{aligned} \tag{70}
$$

From (68), (69), (70), and as $I_1\in[1,I_1^N]$:

$$
|A_2| \leq C\left(\frac{1}{\sqrt{\log N}N^{2\varsigma^2-1/4}}+\frac{1}{\sqrt{N}}+\frac{1}{\sqrt{\log N}\,N^{3/4}}\right)\leq \frac{C}{\sqrt{N}}, \tag{71}
$$

if we choose $\varsigma$ such that $\varsigma>\sqrt{3/8}$.

From (67), (71), and if we choose $\varsigma>\sqrt{3/8}$, we conclude that:

$$
|A| \leq \frac{C}{\sqrt{N}},
$$

and Lemma 3.2 is proved. ■

# 4 Proof of Corollary 2.1

Let us now prove Corollary 2.1. Since:

$$
\begin{aligned}
& \mathbb{E}^\nu\left(\left|\widehat{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})-\bar{u}_b(t_k,\bar{X}^{(i),N}_{t_k},\theta_{(i)})\right|^2\right)\\
\leq\ & \mathbb{E}^\nu\left(\sup_{\alpha\in[0,1]}\left|\widehat{u}_b(t_k,\bar{X}^{(i),N}_{t_k},G^{-1}(\alpha))-\bar{u}_b(t_k,\bar{X}^{(i),N}_{t_k},G^{-1}(\alpha))\right|^2\right),\\
\leq\ & 2\mathbb{E}^\nu\left(\sup_{\alpha\in[0,1]}\left|\widehat{u}_b(t_k,\bar{X}^{(i),N}_{t_k},G^{-1}(\alpha))-\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\psi_I(G(G^{-1}(\alpha)))\right|^2\right)\\
+\ & 2\mathbb{E}^\nu\left(\sup_{\alpha\in[0,1]}\left|\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\psi_I(\alpha)-\bar{u}_b(t_k,\bar{X}^{(i),N}_{t_k},G^{-1}(\alpha))\right|^2\right), \qquad (72)
\end{aligned}
$$

it appears that we can conclude with a proof similar to the classical computation of uniform convergence rates for wavelet regression estimators (see Donoho *et al.*). We first give a lemma allowing to control some moments, then we upper bound the first term in the right hand side of (72) thanks to Lemma 4.2 and the second term thanks to Lemma 4.3.

**Lemma 4.1.** *Under the assumptions of Lemma 2.1, we have:* $\forall\gamma>0$, $\exists N_0, C>0$, $\forall N\geq N_0$, $\forall i\in[1,N]$, $\forall I=(I_1,I_2)\in[-1,I_1^N]\times\mathbb{Z}$, $\forall k\in[0,K]$, $\exists\kappa=\kappa(\gamma,T,\Delta t,b,\sigma,\phi,\psi)$ *(as in Lemma 2.1),*

$$
\mathbb{P}^\nu\left(\sup_{I_2}|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})-\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq\frac{\kappa\log N}{2\sqrt{N}}\right)\ \leq\ \frac{C}{\log N N^{\gamma-1/2}}. \qquad (73)
$$

*Moreover, if we choose* $\gamma>7/2$*, we can deduce from the preceding inequality (73) that:* $\forall p\in[1,6]$,

$$
\mathbb{E}^\nu\left(\sup_{I_2}|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})-\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|^p\right)\ \leq\ \frac{C(\log N)^p}{N^{p/2}}. \qquad (74)
$$

□

*Proof.* The tail upper bound of Lemma 2.1 implies that:

$$
\begin{aligned}
& \mathbb{P}^\nu\left(\sup_{I_2}|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})-\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq\frac{\kappa\log N}{2\sqrt{N}}\right)\\
\leq\ & \sum_{I_2}\mathbb{P}^\nu\left(|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})-\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq\frac{\kappa\log N}{2\sqrt{N}}\right)\\
\leq\ & C2^{I_1}\mathbb{P}^\nu\left(|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})-\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq\frac{\kappa\log N}{2\sqrt{N}}\right)\\
\leq\ & \frac{C}{\log N N^{\gamma-1/2}},
\end{aligned}
$$

since the number of non nul coefficients for a given level $I_1$ is of order $2^{I_1}$ when we consider MRAs generated by compact supported wavelets. From the preceding inequality, we have:

$\forall p \geq 1$,

$$\begin{aligned}
& \mathbb{E}^\nu \left( \sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^p \right) \\
\leq \; & \frac{t_N^p}{2^p} P^\nu \left( \sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| < \frac{t_N}{2} \right) \\
+ \; & (2C')^p ||b||_\infty^p \mathbb{P}^\nu \left( \sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \geq \frac{t_N}{2} \right) \\
\leq \; & \frac{\kappa^p (\log N)^p}{2^p N^{p/2}} + \frac{(2C')^p ||b||_\infty^p}{\log N N^{\gamma - 1/2}},
\end{aligned}$$

since $\sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|$ is bounded by $2C'||b||_\infty$, where $C'$ is an upper bound for $\sqrt{1/N \sum_{j=1}^N \psi_I^2(j/N)}$ that converges to 1. The choice of $\gamma > 7/2$ gives the desired conclusion. ■

**Lemma 4.2.** *Under the assumptions of Theorem 1.2, and with the choice of $\kappa$ for the threshold $t_N$ as in Lemma 2.1:* $\exists N_0, C > 0$, $\forall N \geq N_0$, $\forall i \in [1, N]$, $\forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall k \in [0, K]$,

$$\mathbb{E}^\nu \left( \sup_{\alpha \in [0,1]} \left| \widehat{u}_b(t_k, \bar{X}_{t_k}^{(i),N}, G^{-1}(\alpha)) - \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) \psi_I(G(G^{-1}(\alpha))) \right|^2 \right) \leq C \left( \frac{\log N}{\sqrt{N}} \right)^{\frac{4s}{1+2s}}.$$

□

**Lemma 4.3.** *Under the assumptions of Theorem 1.2, and with the choice of $\kappa$ for the threshold $t_N$ as in Lemma 2.1:* $\exists N_0, C > 0$, $\forall N \geq N_0$, $\forall i \in [1, N]$, $\forall I = (I_1, I_2) \in [-1, I_1^N] \times \mathbb{Z}$, $\forall k \in [0, K]$,

$$\mathbb{E}^\nu \left( \sup_{\alpha \in [0,1]} \left| \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) \psi_I(\alpha) - \bar{u}_b(t_k, \bar{X}_{t_k}^{(i),N}, G^{-1}(\alpha)) \right|^2 \right) \leq C (\log N)^2 \left( \frac{\log N}{\sqrt{N}} \right)^{\frac{4s}{1+2s}}.$$

□

Let us notice that the convergence rates depend on the regularity $s$ of $\alpha \mapsto \bar{u}_b(t_k, x, G^{-1}(\alpha))$ (see Lemma 2.2). This was to be expected, since this regularity can be traduced in term of properties of the wavelet coefficients $\bar{\beta}_I^{(b,t_k)}$. Recall indeed that when $\phi$ and $\psi$ satisfy Assumptions $(H)$ and $(M)$ (see Paragraph 1.3.2), the wavelet coefficients $\beta_I$ of any $s$-Hölder continuous function $f$ satisfy:

$$\exists C > 0, \forall I_1 \geq -1, \sup_{I_2} |\beta_{I_1 I_2}| \leq C \, 2^{-I_1 \left(s + \frac{1}{2}\right)}. \tag{75}$$

## 4.1 Proof of Lemma 4.2

Let us introduce some notations. We write $\Delta_{I,N}(a)$ for $\psi_I(G(a)) - \psi_I(G_N(a))$. As we deal with compact supported wavelets, for any given $I_1$, the number of index $I_2$ such that $\Delta_{I,N}(a)$ does not vanish is finite and does not depend on $I_1$. Since $\psi_I$ is Lipschitz continuous with constant $L_\psi 2^{3I_1/2}$, we notice that:

$$\forall p \geq 1, \, \mathbb{E}^\nu ||\Delta_{I,N}||_\infty^p \leq \frac{C \, 2^{3I_1 p/2}}{N^{p/2}}. \tag{76}$$

We are now ready to start the proof of Lemma 4.2:

$$\mathbb{E}^\nu\left( \|\widehat{u}_b(t_k, \bar{X}^{(i),N}_{t_k}, G^{-1}(\alpha)) - \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\psi_I(G(G^{-1}(\alpha)))\|^2_\infty \right)$$
$$\leq \; I_1^N \sum_{I_1=-1}^{I_1^N} \mathbb{E}\left( \| \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})(\psi_I \circ G_N - \psi_I \circ G) \|^2_\infty \right).$$

The number $C$ of non nul terms in the sum $\sum_{I_2}$ is finite and depends only on $\psi$ and $\phi$. Thus, we have:

$$|| \sum_{I_2} \widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\Delta_{I,N}(a)||^2_\infty \leq C \sup_{I_2} |\widetilde{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|^2 ||\Delta_{I,N}||^2_\infty \leq (1) + (2) + (3),$$

where:

$$(1) \;=\; C \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq t_N\}} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\leq t_N/2\}} ||\Delta_{I,N}||^2_\infty$$
$$(2) \;=\; C \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq t_N\}} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|> t_N/2\}} ||\Delta_{I,N}||^2_\infty$$
$$(3) \;\leq\; C \sup_{I_2} \left|\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq t_N\}} ||\Delta_{I,N}||^2_\infty.$$

Let us work on Term (1):

$$\mathbb{E}^\nu (1) \;=\; \mathbb{E}^\nu \left( \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq t_N\}} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\leq t_N/2\}} \|\Delta_{I,N}\|^2_\infty \right)$$
$$\leq \left( \mathbb{E}^\nu \left( \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^6 \right)\right)^{1/3} \left( P\left( \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right| \geq \frac{t_N}{2} \right)\right)^{1/3}$$
$$\times \; \left(\mathbb{E}^\nu \left(\|\Delta_{I,N}\|^6_\infty\right)\right)^{1/3}.$$

Using Lemma 4.1 and (76), we obtain that:

$$\mathbb{E}^\nu (1) \leq C \frac{(\log N)^2}{N} \times \frac{1}{(\log N)^{1/3} N^{\gamma/3 - 1/6}} \times \frac{2^{3I_1}}{N} \leq C \frac{2^{3I_1} (\log N)^{5/3}}{N^{11/6 + \gamma/3}}.$$

Hence:

$$I_1^N \sum_{I_1=-1}^{I_1^N} \mathbb{E}^\nu (1) \leq \frac{C}{(\log N)^{1/3} N^{1/3 + \gamma/3}} \leq C \left( \frac{\log N}{\sqrt{N}} \right)^{\frac{4s}{1+2s}}, \tag{77}$$

certainly if we choose $\gamma$ such that $1/3 + \gamma/3 > 2s/(1+2s)$ *i.e.* $\gamma > (4s-1)/(1+2s)$. This is the case if we choose $\gamma > 7/2$ as in Lemma 4.1, since $\forall s > 1/2$, $(4s-1)/(1+2s) < 2$.

Now, consider Term (2):

$$\mathbb{E}^\nu \left( \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|\geq t_N\}} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})|> t_N/2\}} \|\Delta_{I,N}\|^2_\infty \right)$$
$$\leq \left( \mathbb{E}^\nu \left( \sup_{I_2} \left|\widehat{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})\right|^6 \right)\right)^{1/3} \left( \mathbb{P}^\nu \left( \sup_{I_2} |\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k})| > t_N/2 \right)\right)^{1/3}$$
$$\times \; \left(\mathbb{E}^\nu \left(\|\Delta_{I,N}\|^6_\infty\right)\right)^{1/3}$$
$$\leq \; C \, \frac{(\log N)^2}{N} \times \left( \mathbb{E}^\nu \left( \sup_{I_2} \frac{2^6 (\bar{\beta}_I^{(b,t_k)}(\bar{X}^{(i),N}_{t_k}))^6}{t_N^6} \right)\right)^{1/3} \times \frac{2^{3I_1}}{N}.$$

Since $\alpha \mapsto \bar{u}_b(t_k, x, G^{-1}(\alpha))$ is $s$-Hölder continuous for all $x \in \mathbb{R}$, we can use (75). Thus, if $1/2 < s < 1$:

$$I_1^N \sum_{I_1=-1}^{I_1^N} \mathbb{E}^\nu(2) \leq C \frac{\log N 2^{2I_1^N(1-s)}}{N} \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{2s} \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}},$$

since $2/(1+2s) < 1$. If $s \geq 1$, we have:

$$I_1^N \sum_{I_1=-1}^{I_1^N} \mathbb{E}^\nu(2) \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{2} \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}},$$

since $2s/(1+2s) < 1$.

Let us finally work on the expectation of term (3):

$$\begin{aligned}
\mathbb{E}^\nu\left(\sup_{I_2} \left|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right|^2 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \geq t_N\}} \|\Delta_{I,N}\|_\infty^2\right) &\leq C\sqrt{\mathbb{E}^\nu\left(\sup_{I_2}\left|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right|^4\right)} \frac{2^{3I_1}}{N} \\
&\leq \frac{2^{2I_1(1-s)}}{N}.
\end{aligned}$$

Thus, if $1/2 < s < 1$:

$$I_1^N \sum_{I_1^N} \mathbb{E}^\nu(3) \quad \leq \quad C \frac{\log N 2^{2I_1^N(1-s)}}{N} \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}.$$

And if $s \geq 1$:

$$I_1^N \sum_{I_1^N} \mathbb{E}^\nu(3) \quad \leq \quad C \left(\frac{\log N}{\sqrt{N}}\right)^{2} \leq C \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}.$$

■

## 4.2 Proof of Lemma 4.3

The proof of Lemma 4.3 is very similar to the usual computation of the uniform convergence rate for wavelet regression estimators (see Donoho *et al.* [11]). The main difference is that the term we upper bound is not really a convergence rate. It rather describes the difference between the regression approximation $\alpha \mapsto \widehat{u}_b(t_k, \bar{X}_{t_k}^{(i),N}, G^{-1}(\alpha))$ and the function $\alpha \mapsto \bar{u}_b(t_k, \bar{X}_{t_k}^{(i),N}, G^{-1}(\alpha))$, which is random (as the second argument $x$ is taken at the random position $\bar{X}_{t_k}^{(i),N}$) and depends on $N$. We have:

$$\begin{aligned}
&\mathbb{E}^\nu\left(\|\sum_{I_1=-1}^{I_1^N}\sum_{I_2} \widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) 1_{\{|\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \geq t_N\}} \psi_I(G(a)) - \bar{u}_b(t_k, \bar{X}_{t_k}^{(i),N}, a)\|_\infty^2\right) \\
&\leq \quad C\left[(1) + (2) + (3) + (4) + (5)\right],
\end{aligned}$$

where:

$$
\begin{aligned}
(1) &= \mathbb{E}^{\nu}\left( \| \sum_{I_1=-1}^{I_1^N} \sum_{I_2} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \le t_N/2,\, |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| < t_N\}} \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) \psi_I \|_\infty^2 \right) \\
(2) &= \mathbb{E}^{\nu}\left( \| \sum_{I_1=-1}^{I_1^N} \sum_{I_2} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| > t_N/2,\, |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| < t_N\}} \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) \psi_I \|_\infty^2 \right) \\
(3) &= \mathbb{E}^{\nu}\left( \| \sum_{I_1=-1}^{I_1^N} \sum_{I_2} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \le t_N/2,\, |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \ge t_N\}} \left(\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right) \psi_I \|_\infty^2 \right) \\
(4) &= \mathbb{E}^{\nu}\left( \| \sum_{I_1=-1}^{I_1^N} \sum_{I_2} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| > t_N/2,\, |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \ge t_N\}} \left(\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})\right) \psi_I \|_\infty^2 \right) \\
(5) &= \mathbb{E}^{\nu}\left( \| \sum_{I_1 \ge I_1^N} \sum_{I_2} \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) \psi_I \|_\infty^2 \right).
\end{aligned}
$$

Using (75) and $s + 1/2 \ge 1$, we deduce that:

$$
(5) \le C 2^{-2I_1^N(s+1/2)} 2^{I_1^N} \le C \left(\frac{\log N}{\sqrt{N}}\right)^{2s} \le C \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}. \tag{78}
$$

Now, let us consider Term (1).

$$
(1) \le C\, I_1^N \sum_{I_1=-1}^{I_1^N} 2^{I_1} \|\psi\|_\infty^2 \mathbb{E}^{\nu}\left( \sup_{I_2} 1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \le t_N/2\}} |\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^2 \right).
$$

For $0 < r < 2$, we have:

$$
\begin{aligned}
1_{\{|\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \le t_N/2\}} |\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^2 &\le \left(\frac{t_N}{2}\right)^{2-r} |\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^r \\
&\le C \left(\frac{\log N}{\sqrt{N}}\right)^{2-r} 2^{-I_1 r(s+1/2)}.
\end{aligned}
$$

We deduce, with $r = 2/(2s+1)$:

$$
\begin{aligned}
(1) &\le C\, I_1^N \sum_{I_1=-1}^{I_1^N} \left(\frac{\log N}{\sqrt{N}}\right)^{2-r} 2^{-I_1 r(s+1/2-1/r)} \\
&\le C(\log N)^2 \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{2s+1}}.
\end{aligned}
$$

Now turn to Term (2). Using (75) again:

$$
\begin{aligned}
(2) &\le C\, I_1^N \sum_{I_1=-1}^{I_1^N} \sqrt{\mathbb{P}^{\nu}\left( \sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \ge \frac{t_N}{2} \right)} \\
&\quad \times \sqrt{\mathbb{E}^{\nu}\left( \sup_{I_2} |\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^4 \right)} \|\psi_I\|_\infty^2 \\
&\le C\, I_1^N \frac{1}{\sqrt{\log N} N^{\gamma/2-1/4}} \sum_{I_1=-1}^{I_1^N} 2^{-2I_1 s} \\
&\le \frac{C(\log N)^{3/2}}{N^{\gamma/2-1/4}}.
\end{aligned}
$$

Choosing $\gamma$ such that $\gamma/2 - 1/4 > 2s/(1+2s)$, *i.e.* $\gamma > (5s+1/2)/(1+2s)$, and this is the case with $\gamma > 7/2$, implies that this term decreases faster than the rate announced in Lemma 4.3.

Let us consider Term (3):

$$\begin{aligned}
(3) &\leq C\, I_1^N \sum_{I_1=-1}^{I_1^N} \sqrt{\mathbb{E}^\nu\left(\sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^4\right)} \\
&\times \sqrt{\mathbb{P}^\nu\left(\sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \geq t_N/2\right)} ||\psi_I||_\infty^2 \\
&\leq C\, I_1^N \frac{(\log N)^2}{N} \frac{1}{\sqrt{\log N}\, N^{\gamma/2-1/4}} 2^{I_1^N} \\
&\leq C \frac{(\log N)^{3/2}}{N^{\gamma/2+3/4}}.
\end{aligned}$$

Again, choosing $\gamma > (s-3/2)/(1+2s)$ (this is the case when we choose $\gamma > 7/2$) allows us to neglect this term.

Let us finally consider Term (4):

$$\begin{aligned}
(4) &\leq C\, I_1^N \sum_{I_1=-1}^{I_1^N} \sqrt{\mathbb{P}^\nu\left(\sup_{I_2} |\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})| \geq t_N/2\right)} \\
&\times \sqrt{\mathbb{E}^\nu\left(\sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^4\right)} 2^{I_1} \|\psi\|_\infty^2.
\end{aligned}$$

Thanks to (75), we know that all the wavelet coefficient $\bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})$ vanish for $I_1$ such that $2^{-I_1(s+1/2)} < t_N/2$. Thus the sum in $I_1$ is in fact a sum from $-1$ to $I_1^{(4)}$ only, where $I_1^{(4)}$ is such that $2^{-I_1^{(4)}(s+1/2)} = t_N/2$. Thus:

$$\begin{aligned}
(4) &\leq C\, I_1^N \sum_{I_1=-1}^{I_1^{(4)}} \sqrt{\mathbb{E}^\nu\left(\sup_{I_2} |\widehat{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N}) - \bar{\beta}_I^{(b,t_k)}(\bar{X}_{t_k}^{(i),N})|^4\right)} 2^{I_1} \|\psi\|_\infty^2 \\
&\leq C \log N \frac{(\log N)^2}{N} 2^{I_1^{(4)}} \\
&\leq C \log N \frac{(\log N)^2}{N} \left(\frac{t_N}{2}\right)^{-\frac{2}{1+2s}} \\
&\leq C \log N \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{4s}{1+2s}}.
\end{aligned}$$

■

# 5 Application to the 2d Navier-Stokes Equation

## 5.1 2d Vortex Equation

The statistical two-dimensional Navier-Stokes equation models the velocity $\mathbf{v} = (v_1, v_2)$ of a viscous incompressible fluid in the plane: $P(d\omega) - a.s.$, $\forall x = (x_1, x_2) \in \mathbb{R}^2$, $\forall t \in [0, T]$,

$$\begin{cases} \frac{\partial \mathbf{v}}{\partial t}(t,x,\theta) + (\mathbf{v} \cdot \nabla)\mathbf{v}(t,x,\theta) = (\sigma^2/2)\Delta \mathbf{v}(t,x,\theta) - \nabla p \\ \nabla \mathbf{v}(t,x,\theta) = 0 \\ \mathbf{v}(0,x,\theta) = \mathbf{v}_0(x,\theta) \\ \mathcal{L}(\theta) = \nu. \end{cases} \tag{79}$$

In the above equation, $p$ is the pressure and $\sigma^2/2$ the viscosity ($\sigma{>}0$), which we assume to be constant. The random initial condition $\mathbf{v}_0$ is as before parameterized by a random variable $\theta$ defined on a probability space $(\Omega, P)$ and with values in $\Theta \subset \mathbb{R}$ (possibly $\mathbb{R}$ itself). Detailed studies of the probabilistic approach for the Navier-Stokes equation with deterministic initial condition have been carried by Marchioro and Pulvirenti [6] and by Méléard in [26, 27, 28].

The probabilistic approach relies on the vortex equation which can be deduced from the Navier-Stokes equation by considering the *curl* of the velocity: $w = curl(\mathbf{v})$. We consider the following evolution problem, where $G$ is the 2 dimensional Poisson kernel ($\forall r > 0$, $G(r) = -(\ln r)/(2\pi)$) and $K$ is defined by $K(y) = \nabla^{\perp} G(|y|) = 1/(2\pi|y|^2)(-y_2, y_1)$:

$$P(d\omega) - a.s., \forall x = (x_1, x_2) \in \mathbb{R}^2, \forall t \in [0, T],$$

$$\begin{cases} \frac{\partial w}{\partial t}(t,x,\theta) = -(K * w \cdot \nabla) w(t,x,\theta) + (\sigma^2/2)\, \Delta w(t,x,\theta), \\ w(0,x,\theta) = w_0(x,\theta) \\ \mathcal{L}(\theta) = \nu. \end{cases} \tag{80}$$

A difficulty arises since the kernel $K$ explodes at 0. To overcome this problem, we will consider $w_0(x,\theta) \in L^1 \cap L^{\infty}$, the space of integrable bounded functions (see Marchioro-Pulvirenti [6], Méléard [28]).

The velocity $\mathbf{v}$ can be recovered from the vortex $w$ by: $\mathbf{v} = K * w$. Then, the mean velocity vector field associated with the Navier-Stokes equation with random initial condition (79) at point $x \in \mathbb{R}^2$ and time $T > 0$ expresses as:

$$\begin{aligned} \forall i \in \{1,2\}, \int_{\mathbb{R}} v_i(T,x,a)\nu(da) &= \int_{\mathbb{R}} \langle w(T,dy,a), K^i(x-y) \rangle \nu(da) \\ &= \langle I(m_T), K^i(x - .) \rangle. \end{aligned} \tag{81}$$

Our purpose is now to adapt the results of Part 1 to the special case of the vortex equation (80), in order to compute the mean velocity vector field (81).

## 5.2 Regularized Equation

The vortex equation (80) is a McKean-Vlasov equation with: $\forall x, y \in \mathbb{R}^2$, $\forall a \in \Theta$,

$$\begin{aligned} b(x,y) &= K(x-y) = \frac{1}{2\pi|x-y|^2} \begin{pmatrix} y_2 - x_2 \\ x_1 - y_1 \end{pmatrix} \\ u_b(t,x,a) &= K * w(t,x,a) \\ u_\sigma(t,x,a) &= \sigma. \end{aligned}$$

Assumptions $(A_1)$ and $(A_2)$ are not satisfied, since $K$ explodes at 0. Following Marchioro and Pulvirenti [6] we regularize the equation to overcome this problem. Take $\varepsilon > 0$, and consider the cut-off equation:

$$\forall t \in [0,T],\ P(d\omega)-a.s., \begin{cases} \frac{\partial w^\varepsilon}{\partial t}(t,x,\theta) = -\left((K_\varepsilon * w^\varepsilon).\nabla\right) w^\varepsilon(t,x,\theta) + (\sigma^2/2)\,\Delta w^\varepsilon(t,x,\theta) \\ w^\varepsilon(0,x,\theta) = w_0(x,\theta) \\ \mathcal{L}(\theta) = \nu, \end{cases} \tag{82}$$

where the function $K_\varepsilon$ is defined by:

$$\begin{aligned} K_\varepsilon(x) &= \nabla^\perp G_\varepsilon(|x|) = \left(G'_\varepsilon(|x|)\frac{x_1}{|x|}, -G'_\varepsilon(|x|)\frac{x_2}{|x|}\right), \qquad (83) \\ &\text{with:} \\ G_\varepsilon(|x|) &= \begin{cases} G(|x|), \text{ if } |x| \geq \varepsilon, \\ \text{extended in a } \mathcal{C}_b^\infty \text{ way on } \mathcal{B}(0,\varepsilon). \end{cases} \end{aligned}$$

It is possible to choose $G_\varepsilon$ such that its derivatives vanish at the origin and satisfy the following inequalities:

$$\forall r \geq 0,\ |G_\varepsilon^{(k)}(r)| \leq \sup_{u\geq\varepsilon} |G^{(k)}(u)| \leq \frac{1}{2\pi\varepsilon^k}\ ,\ \ k \geq 1. \tag{84}$$

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

### 5.3 Weighted and Signed Initial Measures

Another difficulty lies in the fact that the vortex initial condition $w_0$ is not necessarily a probability density. The function $w_0$ may take negative values and have a $L^1$-norm that differs from 1. Hence, Assumption $(A_3)$ also fails. We use a trick due to Jourdain [15] to pass from a density function $p_0$ to any $w_0 \in L^1 \cap L^\infty$. We assume here that:

$$\exists A > 0,\ P(d\omega) - a.s. \left(||w_0(.,\theta)||_{L^1} + ||w_0(.,\theta)||_{L^\infty}\right) < A. \tag{85}$$

Let us introduce the bounded random function $h$ defined by: $P(d\omega) - a.s.$, $\forall x \in \mathbb{R}^2$,

$$\begin{aligned} h(x,\theta) &= \frac{w_0(x,\theta)||w_0(.,\theta)||_{L^1}}{|w_0(x,\theta)|} \\ &= \text{sign}(w_0(x,\theta)) \times ||w_0(.,\theta)||_{L^1} \text{ or } 0, \text{ with the convention } \frac{0}{0} = 0. \end{aligned} \tag{86}$$

Then $P(d\omega) - a.s.$, $\forall x \in \mathbb{R}^2$, $w_0(x,\theta) = h(x,\theta)\frac{|w_0(x,\theta)|}{||w_0(.,\theta)||_{L^1}}$ with $\frac{|w_0(x,\theta)|}{||w_0(.,\theta)||_{L^1}}$ a probability density.

Let us also define, for a probability transition measure $Q(dx,a)$ on $\mathcal{C}([0,T],\mathbb{R}^2)$ measurable in $a \in \mathbb{R}$, the family $(\widetilde{Q}_t(dx,a))_{t\geq 0}$ of weighted signed transition measures on $\mathbb{R}^2$ measurable in $a$ by: $P(d\omega) - a.s.$, $\forall B$ Borel subset of $\mathbb{R}^2$, $\forall t \in [0,T]$,

$$\widetilde{Q}_t(B,\theta) = \langle Q(dx,\theta),\ 1_B(x(t))h(x(0),\theta)\rangle. \tag{87}$$

The following result from Jourdain [15] can be extended to the case of statistical solutions by working conditionally to $\theta$ as in the proof of Theorem 1.1.

**Proposition 5.1.** *(Jourdain [15]) If $P(d\omega) - a.s.$ the marginals $Q_t(dx,\theta)$ at time $t$ of $Q(dx,\theta)$ are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^2$, then $\widetilde{Q}_t(dx,\theta)$ is also $P(d\omega) - a.s.$ absolutely continuous.* □

## 5.4 Existence, Uniqueness and SDE Associated with the Regularized Equation

Part 1's general presentation can be adapted to the case of the regularized vortex equation (82) and we can extend the results obtained for a deterministic initial condition (see Méléard [26, 27, 28]) to the case of statistical solutions by working conditionally to $\theta$ again.

**Proposition 5.2.** *(Méléard [26]) There exists a unique weak function-solution of (82) in the sense of Definition 1 and Remark 2, which we denote by* $(w^\varepsilon(t,.,\theta))_{t\in[0,T]}$. *This weak solution is the density process of* $\left(\widetilde{Q}_t^\varepsilon(dx,\theta)\right)_{t\in[0,T]}$, *associated as in (87) with the weak-solution* $(Q_t^\varepsilon(dx,\theta))_{t\in[0,T]}$ *of the following nonlinear SDE:* $P(d\omega)-a.s.$, $\forall t\in[0,T]$,

$$\begin{cases} X_t^\varepsilon(\theta) = X_0(\theta) + \sigma W_t + \int_0^t K_\varepsilon * \widetilde{Q}_s^\varepsilon(X_s^\varepsilon(\theta),\theta)ds \\ \mathcal{L}(\theta) = \nu \\ \mathcal{L}(X_0(\theta)) = \frac{|w_0(x,\theta)|dx}{||w_0(.\theta)||_{L^1}} \\ W \textit{ is a Brownian motion independent of } \theta \textit{ and of the initial condition } X_0(\theta) \\ Q_s^\varepsilon(dx,\theta) = \mathcal{L}(X_s^\varepsilon(\theta)), \\ \widetilde{Q}_s^\varepsilon(dx,\theta) \textit{ is associated with } Q_s^\varepsilon(dx,\theta) \textit{ as in (87).} \end{cases} \tag{88}$$

*Moreover, pathwise existence and uniqueness are available for SDE (88) (for a given Brownian motion $W$, a given random parameter $\theta$ and a given initial condition $X_0(\theta)$).* □

## 5.5 Existence and Uniqueness of the Statistical Solution

Existence and uniqueness of a weak function-solution $(w(t,.,\theta))_{t\in[0,T]}$ for the original equation (80) hold by adapting again the proofs with deterministic initial condition (see Méléard [26, 27, 28]).

**Proposition 5.3.** *(Méléard [26])*
*(i) For a random initial condition* $w_0(.,\theta)$ *on* $L^1(\mathbb{R}^2)\cap L^\infty(\mathbb{R}^2)$, *there exists a unique weak function-solution* $(w(t,.,\theta))_{t\in[0,T]}$ *of (80) in the sense of Definition 1 and Remark 2. This weak solution is the density process of* $(\widetilde{Q}_t(dx,\theta))_{t\in[0,T]}$, *associated as in (87) with the weak-solution of the following nonlinear SDE:* $P(d\omega)-a.s.$, $\forall t\in[0,T]$,

$$\begin{cases} X_t(\theta) = X_0(\theta) + \sigma W_t + \int_0^t K * \widetilde{Q}_s(X_s(\theta),\theta)ds \\ \mathcal{L}(\theta) = \nu \\ \mathcal{L}(X_0(\theta)) = \frac{|w_0(x,\theta)|dx}{||w_0(.\theta)||_{L^1}} \\ W \textit{ is a Brownian motion independent of } \theta \textit{ and of the initial condition } X_0(\theta) \\ Q_s(dx,\theta) = \mathcal{L}(X_s(\theta)), \\ \widetilde{Q}_s(dx,\theta) \textit{ is associated with } Q_s(dx,\theta) \textit{ as in (87).} \end{cases} \tag{89}$$

*(ii) The function-solution* $(w(t,.,\theta))_{t\in[0,T]}$ *can be viewed as a random variable parameterized by $\theta$ with values in the following space:*

$$\mathcal{H} = \left\{ (q_t(x))_{t\leq 0} \ \mid \ \forall t\in[0,T],\ q_t\in L^1\cap L^\infty,\ \sup_t ||q_t||_{L^1}\leq A,\ \sup_t ||q_t||_{L^\infty}\leq A,\ A \textit{ given by } (85) \right\},$$

*endowed with the complete norm:* $|||q||| = \sup_{t\in[0,T]} |||q_t|||$, *with* $|||q_t||| = ||q_t||_{L^1} + ||q_t||_{L^\infty}$. □

The following lemma, from Méléard [26], is also still available. It shows that we can approximate the weak function-solution $w$ of (80) with the weak function-solution $w^\varepsilon$ of (82):

**Lemma 5.1.** *(Méléard [26])Take $\varepsilon' > \varepsilon > 0$. We have:*

$$P(d\omega)-a.s.,\ |||w^\varepsilon(.,.,\theta) - w^{\varepsilon'}(.,.,\theta)||| \leq C\,\varepsilon' \exp(CT). \tag{90}$$

*Inequality (85) tells us that the constant $C$ depends only on $A$ and not on $\omega$.* □

Using completeness of $(\mathcal{H}, |||.|||)$ and Lemma 5.1, we can show by adapting the proofs in Méléard [26], that $P(d\omega) - a.s.$, $w(.,.,\theta)$ is the limit in $\mathcal{H}$ of $w^\varepsilon(.,.,\theta)$ when $\varepsilon \to 0$. Letting $\varepsilon \to 0$ in Lemma 5.1, we deduce that:

$$P(d\omega) - a.s., \; |||w^\varepsilon(.,.,\theta) - w(.,.,\theta)||| \leq C\,\varepsilon \exp(C\,T). \tag{91}$$

## 5.6 Particle Approximation for $\langle I(m_T), f\rangle$

Since $P(d\omega) - a.s.$, $K_\varepsilon * \widetilde{Q}^\varepsilon_s(X^\varepsilon_s(\theta),\theta) = \mathbb{E}\left(h\left(X_0(\theta),\theta\right) K_\varepsilon(x - X^\varepsilon_s(\theta))\right)|_{x=X^\varepsilon_s(\theta)}$, SDE (88) can be rewritten as: $P(d\omega) - a.s.$, $\forall t \in [0,T]$,

$$\begin{cases} X^\varepsilon_t(\theta) = X_0(\theta) + \sigma W_t + \int_0^t \mathbb{E}^\nu\left(h\left(X_0(\theta),\theta\right) K_\varepsilon(x - X^\varepsilon_s(\theta))\,|\,\theta\right)|_{x=X^\varepsilon_s(\theta)} ds \\ \mathcal{L}(\theta) = \nu \\ \mathcal{L}(X_0(\theta)) = \frac{|w_0(x,\theta)|dx}{||w_0(.,\theta)||_{L^1}}, \\ W \text{ is a Brownian motion independent of } \theta \text{ and of the initial condition } X_0(\theta). \end{cases} \tag{92}$$

As in Part 1, we can now build a particle system similar to (23) in order to approximate the law of the diffusion (92). To this purpose, we replace the conditional expectation by a regression estimator.

We simulate $N$ realizations $(\theta_i)_{i\in[1,N]}$ of $\theta$, rank them, and associate with each of them a particle with initial condition $\bar{X}^{(i),N,\varepsilon}_0(\theta_{(i)})$ of law $|w_0(x,\theta_{(i)})|dx/||w_0(.,\theta_{(i)})||_{L^1}$. We compute the regressions component by component: $\forall i \in [1,N]$, $\forall k \in [0,K]$,

$$\begin{aligned} \bar{X}^{(i),N,\varepsilon}_{t_k} &= \bar{X}^{(i),N,\varepsilon}_{t_{k-1}} + \sigma(W^i_{t_k} - W^i_{t_{k-1}}) + \Bigg[\frac{1}{N}\sum_{j=1}^{N}\sum_{I_1=-1}^{I_1^N}\sum_{I_2}\psi_I\left(\frac{i}{N}\right)\psi_I\left(\frac{j}{N}\right) h\left(\bar{X}^{(j),N,\varepsilon}_0,\theta_{(j)}\right) \\ &\times \begin{pmatrix} 1_{|\frac{1}{N}\sum_{j=1}^N \psi_I\left(\frac{j}{N}\right) h\left(\bar{X}^{(j),N,\varepsilon}_0,\theta_{(j)}\right) K^1_\varepsilon\left(\bar{X}^{(i),N,\varepsilon}_{t_{k-1}} - \bar{X}^{(j),N,\varepsilon}_{t_{k-1}}\right)|>t_N} K^1_\varepsilon\left(\bar{X}^{(i),N,\varepsilon}_{t_{k-1}} - \bar{X}^{(j),N,\varepsilon}_{t_{k-1}}\right) \\ 1_{|\frac{1}{N}\sum_{j=1}^N \psi_I\left(\frac{j}{N}\right) h\left(\bar{X}^{(j),N,\varepsilon}_0,\theta_{(j)}\right) K^2_\varepsilon\left(\bar{X}^{(i),N,\varepsilon}_{t_{k-1}} - \bar{X}^{(j),N,\varepsilon}_{t_{k-1}}\right)|>t_N} K^2_\varepsilon\left(\bar{X}^{(i),N,\varepsilon}_{t_{k-1}} - \bar{X}^{(j),N,\varepsilon}_{t_{k-1}}\right) \end{pmatrix}\Bigg]\Delta t. \end{aligned} \tag{93}$$

where $\varepsilon > 0$ is fixed and $K^1_\varepsilon(x)$ and $K^2_\varepsilon(x)$ are the first and second components of $K_\varepsilon(x)$.

**Theorem 5.1.** *Under the following assumptions:*

1. *The law $\nu$ of $\theta$ admits a density $g$ with connected support on $\mathbb{R}$ w.r.t. the Lebesgue measure,*
2. *$P(d\omega) - a.s.$, $w_0(.,\theta) \in L^1 \cap L^\infty$,*
3. *$P(d\omega) - a.s.$, $\int x^2 |w_0(x,\theta)|dx < \infty$,*
4. *The marginal law in $x$ of $\frac{|w_0(x,a)|}{||w_0(.,a)||_{L^1}} g(a)da\,dx$, satisfies a Poincaré inequality with the positive deterministic constant $c_0$.*
5. *The application $\Phi \circ G^{-1} : \alpha \in [0,1] \mapsto \frac{|w_0(.,G^{-1}(\alpha))|}{||w_0(.,G^{-1}(\alpha))||_{L^1}}$ is s-Hölder continuous for the norm in $L^1(\mathbb{R})$, with $s > 1/2$, and $G$ the distribution function of $\theta$,*
6. *The father and mother wavelets $\phi$ and $\psi$ are compactly supported, Lipschitz continuous, and satisfy Assumptions (H) and (M),*
7. *We threshold the estimators with $t_N = \kappa\frac{\log N}{\sqrt{N}}$ where $\kappa = \kappa(\varepsilon)$ is a positive constant that has to be chosen (see (55) to see precisely how),*
8. *The resolution level $I_1^N$ satisfies $2^{I_1^N} \sim \frac{\sqrt{N}}{\log N}$,*

*Then,*

$$\forall \varepsilon > 0,\, \forall 0 < \eta < 1,\, \forall f \in \mathcal{C}_b^{4+\eta}(\mathbb{R}),\, \exists N_0,\, \forall N \geq N_0,$$
$$C_1 = C_1(f,h,\varepsilon), C_2 = C_2(f,h), C_3 = C_3(f,h) > 0,$$
$$\mathbb{E}^\nu |\langle I(m_T), f\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) f(\bar{X}_T^{(i),N,\varepsilon})| \leq C_1\,\Delta t + C_2 \log N \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{1+2s}} + C_3\varepsilon, \tag{94}$$

*where $h$ is defined as in (86). The constants $C_1(f,h,\varepsilon)$, $C_2(f,h)$ and $C_3(f,h)$ can be respectively upper bounded by $C||h||_\infty \left(\frac{\sum_{k=0}^4 ||f^{(k)}||_\infty}{\varepsilon^5} + \frac{1}{\varepsilon^{10}}\right)$, $C||h||_\infty\, ||f'||_\infty$ and $C||f||_\infty ||h||_\infty$.* □

*Proof.* We have:

$$\begin{aligned}
&\mathbb{E}^\nu |\langle I(m_T), f\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) f(\bar{X}_T^{(i),N,\varepsilon})| \\
\leq\ & \mathbb{E}^\nu |\langle I(m_T), f\rangle - \langle I^\varepsilon(m_T), f\rangle| + \mathbb{E}^\nu |\langle I^\varepsilon(m_T), f\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) f(\bar{X}_T^{(i),N,\varepsilon})| \quad (95)
\end{aligned}$$

Let us consider the first term in the right hand side of (95). If $X$ and $X^\varepsilon$ are solutions of the SDE (89) and (92) respectively, then:

$$\begin{aligned}
|\langle I(m_T), f\rangle - \langle I^\varepsilon(m_T), f\rangle| &= |\mathbb{E}^\nu \left(h(X_0(\theta),\theta) f(X_T(\theta))\right) - \mathbb{E}^\nu \left(h(X_0(\theta),\theta) f(X_T^\varepsilon(\theta))\right)| \\
&\leq \int_{\mathbb{R}} \int_{\mathbb{R}^2} |f(x)|\, |w(T,x,a) - w^\varepsilon(T,x,a)| dx\, \nu(da) \\
&\leq \int_{\mathbb{R}} ||f||_\infty\, |||w(T,.,a) - w^\varepsilon(T,.,a)|||\, \nu(da) \\
&\leq C(T,A)||f||_\infty \varepsilon,
\end{aligned}$$

thanks to (91).

The second term in the right hand side is dealt with a result similar to Theorem 1.2. As in the proof of Theorem 1.2, the approximation error can be decomposed in three terms as in (29): the error due to the use of an Euler scheme, the statistical error and the term linked to the propagation of chaos.

As in the proof of Theorem 1.2, the error linked to the use of an Euler scheme (31) is in $(C||h||_\infty \sum_{k=0}^4 ||f^{(k)}||_\infty)/\varepsilon^5$ and the statistical error (32) is in $C||h||_\infty ||f||_\infty/\sqrt{N}$.

The part of the error linked to the use of a particle system can be divided in two terms as in (34). The first term, which corresponds to the term $A_1$ in (34), can be upper bounded with Theorem 2.1. An upper bound is then $C_1\Delta t$, where the constant $C_1$ is upper bounded by $C||h||_\infty ||\partial_y K_\varepsilon^{(4)}||_\infty ||\partial_x K_\varepsilon^{(4)}||_\infty \leq C||h||_\infty/\varepsilon^{10}$, using (84).

The second term, which corresponds to the term $A_2$ in (34), can be upper bounded with spectral gap inequalities as it was done in paragraph 3.1. The proofs extend indeed to the case where the unknown drift:

$$\mathbb{E}^\nu \left(h\left(X_0(\theta),\theta\right) K_\varepsilon(x - X_t^\varepsilon(\theta))\,|\,\theta\right)|_{x=X_t^\varepsilon(\theta)},$$

also depends on $X_0(\theta)$. To see this, we have to show that the law of $(\bar{X}_0^{(i),N,\varepsilon}, \bar{X}_{t_k}^{(i),N,\varepsilon})_{i\in[1,N]}$ satisfies a spectral gap inequality similar to the one obtained in Corollary 3.1. This comes from the fact that $(\bar{X}_0^{(i),N,\varepsilon})_{i\in[1,N]}$ do not vary with time, and from the assumption that their laws satisfy spectral gap inequalities with constant $c_0$. An upper bound for the term $A_2$ is then $C||h||_\infty ||f'||_\infty \log N \left((\log N)/\sqrt{N}\right)^{\frac{2s}{1+2s}}$. Notice that the constant $C$ here does not depend on $\varepsilon$, but that the constant $\kappa$, which appears in the threshold $t_N = \kappa(\log N)/\sqrt{N}$ and which is set by considering this part of this error, does depend on $\varepsilon$. ■

Letting $N \to \infty$ in (94) shows that for any given $\varepsilon > 0$, the biais of our particle method is asymptotically of order:

$$\frac{\Delta t}{\varepsilon^{10}} + \varepsilon.$$

It seems however impossible to set $\varepsilon = \varepsilon_N$ and let $\varepsilon_N \to 0$ when $N \to \infty$ with our proof (Section 3). Changing the discretization step $\Delta t = o(\varepsilon_N^{11})$ allows the right hand side of (94) to converge to 0 with $N$ indeed, but we can not bound the threshold $\kappa = \kappa(\varepsilon_N)$ independently of $N$. As a matter of fact, $\kappa$ (defined in (55)) depends on the Poincaré constant $D_k$ (defined in Corollary 3.2) which we can control independently of $\Delta t$ thanks to arguments similar to (39) and on the Lipschitz constant of $K_\varepsilon$, which we can not control.

## 5.7 Mean Velocity Vector Field

### 5.7.1 Particle Approximation

Recall that the mean velocity vector field associated to the statistical solution of (79) at point $x \in \mathbb{R}^2$ and time $T > 0$ has been defined in (81). This computation corresponds to the following particular choice for $f$ in (81): $\forall y \in \mathbb{R}^2$, $f(y) = K(x - y)$.

Since $K(x-.)$ and its derivatives are not bounded at 0, we use a cut-off technique again and cut-off the test function with $\delta > 0$ as in (83). The approximation we will consider for $\langle I(m_T), K(x-.)\rangle$ is: $\frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) K_\delta(x - \bar{X}_T^{(i),N,\varepsilon})$.

**Theorem 5.2.** *Under the assumptions of Theorem 5.1:*

$$\forall \varepsilon, \delta > 0,\ \forall x \in \mathbb{R}^2,\ \exists N_0,\ \forall N \geq N_0,$$

$$\exists \kappa = \kappa(\varepsilon),\ C_1 = C_1(\delta, h, \varepsilon), C_2 = C_2(\delta, h), C_3 = C_3(\delta, h), C_4 > 0,$$

$$\begin{aligned} & \mathbb{E}^\nu |\langle I(m_T), K(x-.)\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) K_\delta(x - \bar{X}_T^{(i),N,\varepsilon})| \\ \leq \ & C_1\,\Delta t + C_2 \log N \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{1+2s}} + C_3\varepsilon + C_4\delta. \end{aligned}$$

*The constants $C_1$, $C_2$, $C_3$ can be respectively upper bounded by $C||h||_\infty \left(\frac{1}{\delta^5\varepsilon^5} + \frac{1}{\varepsilon^{10}}\right)$, $\frac{C||h||_\infty}{\delta^2}$ and $\frac{C||h||_\infty}{\delta}$.* □

*Proof.* We first decompose the error:

$$\begin{aligned} & \mathbb{E}^\nu |\langle I(m_T), K(x-.)\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) K_\delta(x - \bar{X}_T^{(i),N,\varepsilon})| \\ \leq \ & |\langle I(m_T), K(x-.)\rangle - \langle I(m_T), K_\delta(x-.)\rangle| \\ + \ & \mathbb{E}^\nu |\langle I(m_T), K_\delta(x-.)\rangle - \langle I^\varepsilon(m_T), K_\delta(x-.)\rangle| \\ + \ & \mathbb{E}^\nu |\langle I^\varepsilon(m_T), K_\delta(x-.)\rangle - \frac{1}{N}\sum_{i=1}^N h\left(\bar{X}_0^{(i),N,\varepsilon}, \theta_{(i)}\right) K_\delta(x - \bar{X}_T^{(i),N,\varepsilon})|. \end{aligned} \tag{96}$$

The two last terms in the right hand side of (96) can be dealt with Theorem 5.1. Let us now consider the first term:

$$\begin{aligned}
& |\langle I(m_T), K(x-.)\rangle - \langle I(m_T), K_\delta(x-.)\rangle| \\
\leq\ & |\mathbb{E}^\nu\left(h(X_0(\theta),\theta)K(x-X_T(\theta))\right) - \mathbb{E}^\nu\left(h(X_0(\theta),\theta)K_\delta(x-X_T(\theta))\right)| \\
\leq\ & \int_{\mathbb{R}}\int_{\mathbb{R}^2} |K(x-y)-K_\delta(x-y)|\,|w(T,y,a)|dy\,\nu(da) \\
\leq\ & 2\int_{\mathbb{R}}\int_{|x-y|\leq\delta} |K(x-y)||w(T,y,a)|dy\nu(da) \\
\leq\ & 2\int_{|x-y|\leq\delta} |K(x-y)|dy \int_{\mathbb{R}} |||w(T,.,a)|||\nu(da) \\
\leq\ & C(T,A)\delta.
\end{aligned}$$

■

The rate of convergence obtained in Theorem 5.3 does not depend on the point $x \in \mathbb{R}^2$ where we approximate the mean velocity vector field.

For any given $\varepsilon$ and $\delta$, the asymptotic biais when $N \to \infty$ is of order:

$$\frac{\Delta t}{\varepsilon^{10}} + \frac{\Delta t}{\delta^5\varepsilon^5} + \frac{\varepsilon}{\delta} + \delta. \qquad (97)$$

### 5.7.2 Numerical Experiments

We study the Navier-Stokes equation (80) with $\sigma = 10^{-4}$. The initial condition is given by: $\mathcal{L}(X_0(\theta)) = \mathcal{N}(\theta, 0.3)$, where the distribution of $\theta$ is a gaussian mixing:

$$\theta = 1_{U=1}\theta^{(1)} + 1_{U=0}\theta^{(2)},$$

with $U$ a binomial random variable that takes the value 1 with probability 0.3, and with $\theta^{(1)}$ and $\theta^{(2)}$ two gaussian random variables with respective expectations 1.3 and 0, and with standard deviation 0.2. The random variables $U$, $\theta^{(1)}$ and $\theta^{(2)}$ are independent.

We are interested in the mean velocity vector field and use formula (81), to simulate its time evolution. We apply our particle approximations to: $f_\delta = K_\delta(x-.)$, with $x = (x_1, x_2)$ varying on a grid of $0.5 \times 0.5$ on $[-3,3] \times [-3,3]$ and with $\delta = 10^{-1}$. In our particular example, we have $h \equiv 1$.

We compare the results provided by each of the three preceding methods for a given number of $2^8 = 256$ particles. The number of particles is a power of 2 because we use a Mallat algorithm to handle the wavelet expansions. The drift function is cut-off as in Paragraph 5.2, with $\varepsilon = 10^{-2}$. The diffusions are simulated by using Euler Schemes with $\Delta t = 0.5$ (the theoretical discretization step $\Delta t = 10^{-21}$ that we should use to be coherent with (97) requires very powerful computational abilities).

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 evolutions of the mean velocity vector fields on Figure 1 look similar in every case of our example. This was to be expected, since the example has good regularity properties. We can verify here that the three methods are concordant. Notice however that for this particular example the assumptions of Proposition 1.4 do not hold. Thus, the theoretical asymptotic convergence rate is not available for Method 2.

Notice also that the Markovian structure of the particles' diffusions as well as the travelling properties of the 2d-Brownian motion make the influence of the initial randomness decrease with time. Though, on our simulations, the center of the vortex evolves differently according to the chosen method.

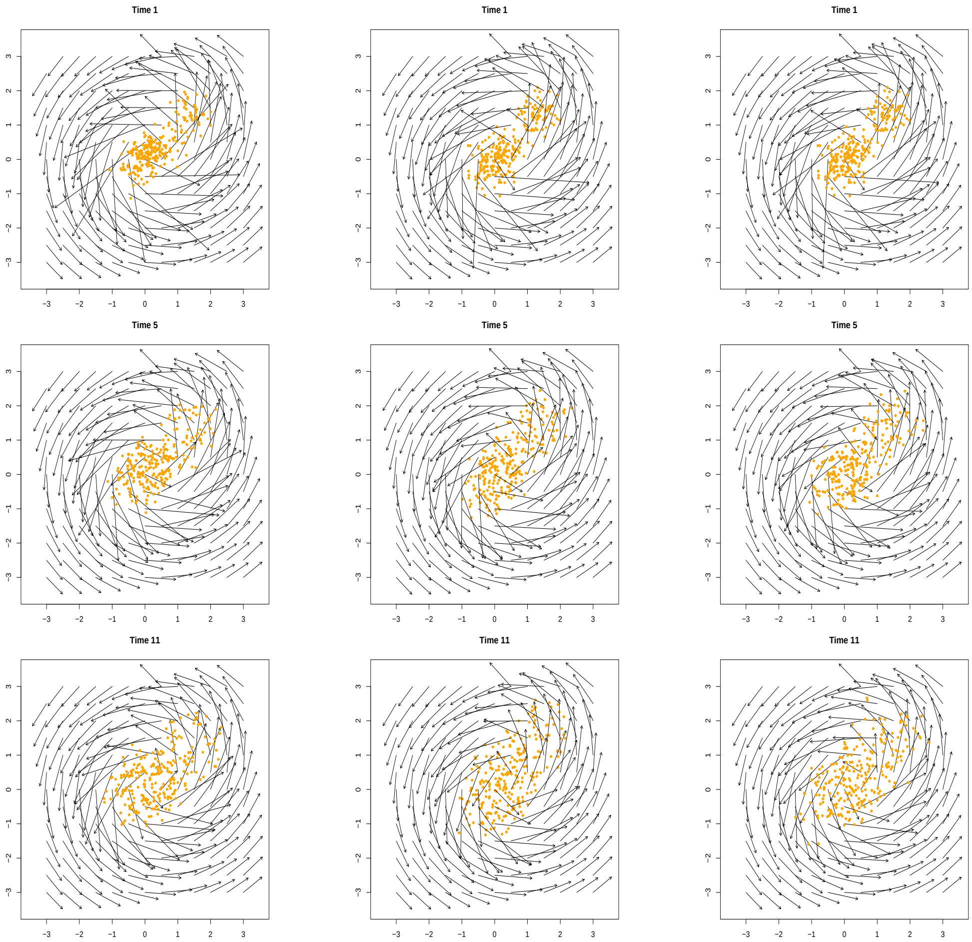


Figure 1: *Approximation of the mean velocity vector field given by the first particle method (left), the second particle method with Nadaraya-Watson estimators (center) and with Kerkyacharian-Picard estimators (right). The first line corresponds to the initialization, then the two other lines represent the approximation of the mean velocity vector field after respectively 5 and 11 discretization steps.*

Let us now have a look at the quality of the regression estimators. In each of the plots of Figure 2, we draw the points $\left(\theta_{(i)}, K^1_\varepsilon\left(\bar{X}^{(128),256,\varepsilon}_{t_1} - \bar{X}^{(i),256,\varepsilon}_{t_1}\right)\right)_{i\in[1,N]}$ with bullets and the first component of the regression estimator $a \mapsto \widehat{u}_{K_\varepsilon}(t_1, \bar{X}^{(128),256,\varepsilon}_{t_1}, a)$ in continuous line. We choose the particle number 128 as it corresponds to the median realization of $\theta$. Results are only presented for the first components of $K$ and $\widehat{u}_{K_\varepsilon}(t_1, \bar{X}^{(128),256,\varepsilon}_{t_1}, .)$.

It seems that the wavelet estimator fits better to the data. In particular, in the gap between the two peaks of our gaussian mixing, where fewer observations are available, the Nadaraya-Watson estimator is not robust at all. In contrary, in the "crowded" regions, where many realizations of $\theta$ can be found, the Nadaraya-Watson estimator seems to average more than the Kerkyacharian-Picard estimator and thus, it misses some aggregation features. This can give an intuition to understand why the method based on wavelet regression estimators is more accurate than the method based on Nadaraya-Watson estimators, and why it extends to less regular initial conditions.

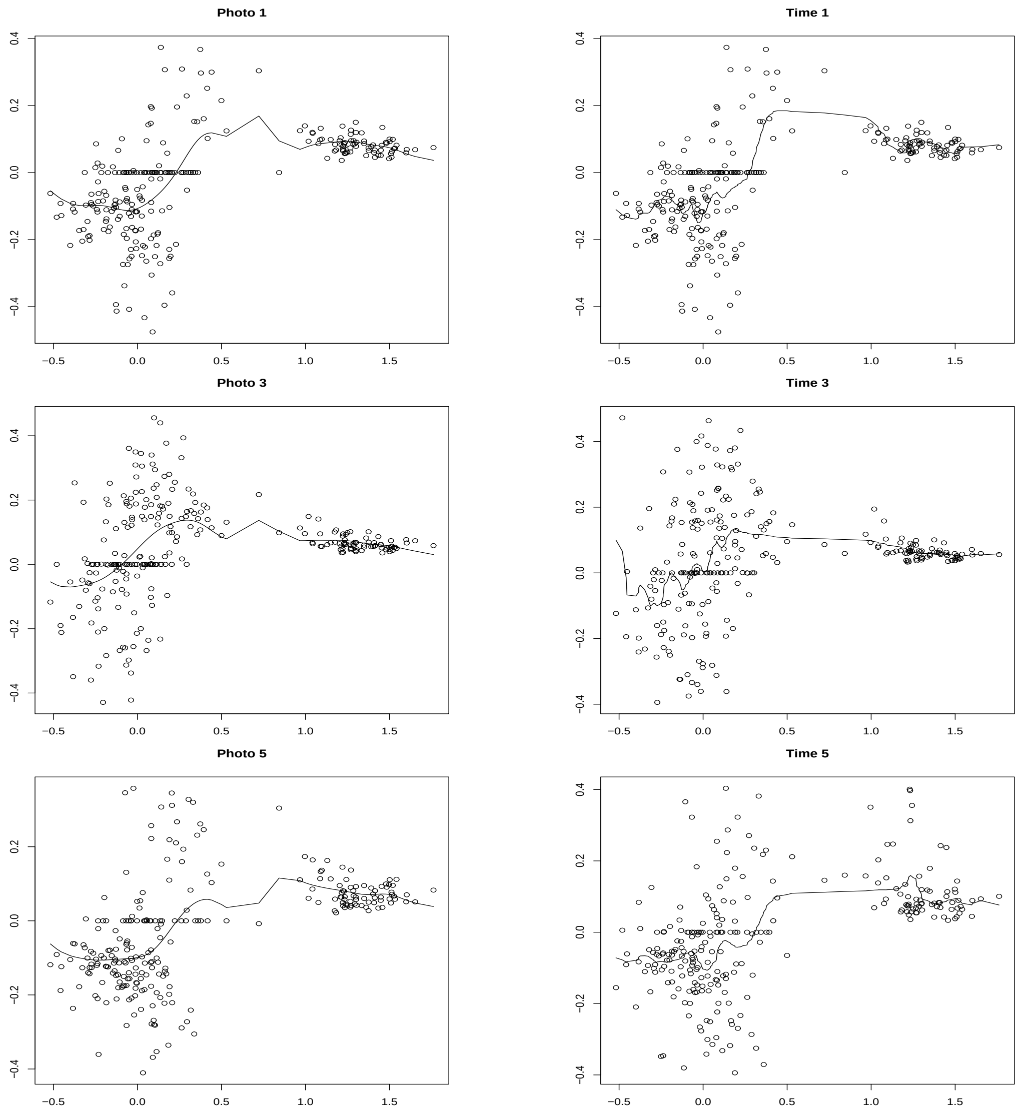


Figure 2: *On the x-axis are the values of* $(\theta_i)_{i\in[1,N]}$*. The circles correspond to the points* $\left(\theta_{(i)}, K^1_\varepsilon\left(\bar{X}^{(128),256,\varepsilon}_t - \bar{X}^{(i),256,\varepsilon}_t\right)\right)_{i\in[1,N]}$ *and in continuous line is the first component of the regression estimator* $a \mapsto \widehat{u}_{K_\varepsilon}(t, \bar{X}^{(128),256,\varepsilon}_t, a)$*. Left is the Nadaraya-Watson estimator with band-with 0.1, whereas right is the Kerkyacharian-Picard estimator with threshold 0.1 and level 4 resolution. The results are presented for the first, third and fifth discretization time.*

## 5.8 Test Case and Comparison of the Three Methods of Section 1

Finally, we end this paper with another numerical test in order to conclude the comparison between the three particle methods introduced in Section 1.

### 5.8.1 Test Problem

Following Milinazzo and Saffman [24] and Bossy [5], we consider the vortex equation (80) and focus on the following quantity:

$$\forall t \in [0,T],\ \forall a \in \Theta,\ L(t,a) = \frac{\int_{\mathbb{R}^2} |x|^2 w(t,x,a)dx}{\int_{\mathbb{R}^2} w(t,x,a)dx},$$

which satisfies the following equation:

$$\forall t \in [0,T],\ \forall a \in \Theta,\ L(t,a) = L(0,a) + 2t\sigma^2.$$

In case $P(d\omega) - a.s.$, $w_0(.,\theta)$ is a density function, we have:

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

The idea is then to compute the approximations of $t \mapsto \langle I(m_t)(dx), |x|^2\rangle$ given by the three particle methods introduced in Section 1, and to see how they fit the theoretical line $t \mapsto \int_{\mathbb{R}} L(0,a)\nu(da) + 2t\sigma^2$. Since the map $x \in \mathbb{R}^2 \mapsto |x|^2$ is not bounded, we replace it with $x \in \mathbb{R}^2 \mapsto |x|^2 1_{\{|x|\leq M\}}$.

Let us mention the convergence theorem for the wavelet particle method:

**Theorem 5.3.** *Under the assumptions of Theorem 5.1, and with the following additional assumptions:*

1. $P(d\omega) - a.s.$, $w_0(.,\theta)$ *is a density function,*
2. $P(d\omega) - a.s.$, $\int_{\mathbb{R}^2} |x|^4 w_0(x,\theta)dx < \infty$,

*then:* $\forall \varepsilon > 0$, $\exists N_0$, $\forall N \geq N_0$, $\exists \kappa = \kappa(\varepsilon)$, $C_1 = C_1(\varepsilon, M), C_2, C_3 = C_3(M), C_4 > 0$, $\forall t \in [0,T]$,

$$\mathbb{E}^\nu \left| \langle I(m_t)(dx), |x|^2\rangle - \frac{1}{N}\sum_{i=1}^N \left|\bar{X}_t^{(i),N,\varepsilon}\right|^2 1_{\{|\bar{X}_t^{(i),N,\varepsilon}|\leq M\}} \right| \leq C_1\,\Delta t + C_2 \log N \left(\frac{\log N}{\sqrt{N}}\right)^{\frac{2s}{1+2s}} + C_3\varepsilon + \frac{C_4}{M}.$$

*The constants $C_1(\varepsilon, M)$ and $C_3(M)$ can be respectively upper bounded by $C\left(\frac{M^2+1}{\varepsilon^5} + \frac{1}{\varepsilon^{10}}\right)$ and $C\,M$.* □

*Proof.* As previously, we decompose the error:

$$\begin{aligned}\mathbb{E}^\nu \left| \langle I(m_t)(dx), |x|^2\rangle - \frac{1}{N}\sum_{i=1}^N \left|\bar{X}_t^{(i),N,\varepsilon}\right|^2 1_{\{|\bar{X}_t^{(i),N,\varepsilon}|\leq M\}} \right| &\leq |\langle I(m_t)(dx), |x|^2\rangle - \langle I(m_t)(dx), |x|^2 1_{\{|x|\leq M\}}\rangle| \\ &+\mathbb{E}^\nu|\langle I(m_t)(dx), |x|^2 1_{\{|x|\leq M\}}\rangle - \langle I^\varepsilon(m_t)(dx), |x|^2 1_{\{|x|\leq M\}}\rangle| \\ &+\mathbb{E}^\nu\left|\langle I^\varepsilon(m_t), |x|^2 1_{\{|x|\leq M\}}\rangle - \frac{1}{N}\sum_{i=1}^N \left|\bar{X}_t^{(i),N,\varepsilon}\right|^2 1_{\{|\bar{X}_t^{(i),N,\varepsilon}|\leq M\}}\right|. \quad (98)\end{aligned}$$

The two last terms in the right hand side of (98) can be dealt with Theorem 5.1. Let us now consider the first term:

$$\begin{aligned}|\langle I(m_t)(dx), |x|^2\rangle - \langle I(m_t)(dx), |x|^2 1_{\{|x|\leq M\}}\rangle| &= |\langle I(m_t)(dx), |x|^2 1_{\{|x|>M\}}\rangle| \\ &= \left|\mathbb{E}^\nu\left(|X_t(\theta)|^2 1_{\{|X_t(\theta)|>M\}}\right)\right| \\ &\leq \sqrt{\mathbb{E}^\nu\left(|X_t(\theta)|^4\right)}\sqrt{\mathbb{P}^\nu\left(|X_t(\theta)| > M\right)}.\end{aligned}$$

Our assumptions imply that: $\sup_{t\in[0,T]} \mathbb{E}^\nu\left(|X_t(\theta)|^4\right) < \infty$ and $\sup_{t\in[0,T]} \mathbb{P}^\nu\left(|X_t(\theta)| > M\right) \leq \sup_{t\in[0,T]} \mathbb{E}^\nu\left(|X_t(\theta)|^2\right)/M^2$, which concludes the proof. ■

### 5.8.2 Numerical Results and Conclusions

We consider the vortex equation (80) with $\sigma = 10^{-3}$ and choose as initial condition $\mathcal{L}(X_0(\theta)) = \mathcal{N}(\theta, 0.6)$ with a gaussian mixing of means -1 and 1.3 and of standard deviations 0.2 for the distribution of $\theta$. We apply our particle approximations with the choice of the test function: $f(x) = |x|^2$ and compare the results provided by each of the three preceding methods for a given number of $2^8 = 256$ particles. The drift function is cut-off as in Paragraph 5.2, with $\varepsilon = 5.10^{-4}$. The diffusions are simulated by using Euler Schemes with $\Delta t = 0.05$ and $T = 5$.

A criterium to compare how well the particle approximations of $t \mapsto \langle I(m_t)(dx), |x|^2\rangle$ fit to the theoretical line $t \mapsto \int_{\mathbb{R}} L(0,a)\nu(da) + 2t\sigma^2$ is given by the relative error. If we denote by $A(t)$ the particle approximation of $\int_{\mathbb{R}} L(t,a)\nu(da)$, the relative error can be defined by:

$$e(t) = \frac{\left|A(t) - A(0) - 2t\sigma^2\right|}{|A(t)|}.$$

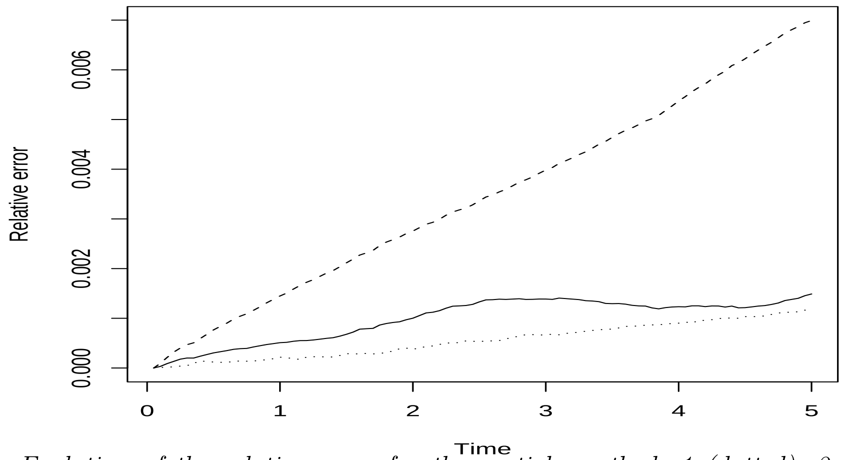


Figure 3: *Evolution of the relative error for the particle methods 1 (dotted), 2 (dashed) and 3 (plain).*

In term of relative error, it clearly appears on this set of simulations that Method 1 (dotted line) gives the best performance. Our particle method with wavelets (Method 3, in plain line) is a little less accurate, but remarkably better than Method 2 (dashed line). In term of simulation time, Method 3 is longer than Method 2 but avoids the imbricated simulation procedure of Method 1.

## Thanks

I am extremely grateful to my PhD advisor, Mrs. 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 M. Gérard Kerkyacharian and M. Patrick Cattiaux for very useful help.

## Appendix

We recall here some facts about wavelet decomposition and about the wavelet regression in random design. For further details, there is an abundant literature on the subject (see for instance Daubechies [9], Härdle and al. [14] or Meyer [23] for the wavelet theory, Kerkyacharian and Picard [17], Donoho and Johnstone [10] or Donoho and al. [11] for wavelet regression).

# A Wavelet Decomposition

**Notation:** In this paragraph, we denote by $\widehat{f}$ the Fourier-transform of $f \in L^2$ endowed with its canonical euclidian structure: $\widehat{f}(\omega) = \int e^{-it\omega} f(t)\, dt$.

Take $\phi \in L^2$. One defines the descendants of this application by $\phi_{I_1 I_2} = 2^{I_1/2}\phi(2^{I_1}a - I_2)$, for $I_1 \in \mathbb{N}$, $I_2 \in \mathbb{Z}$.

**Assumption 3.** *$\phi$ is said to be a scale function if it satisfies the following properties $(HO_1)$:*

1. *$\sum_{I_2=-\infty}^{\infty} |\widehat{\phi}(\omega + 2I_2\pi)|^2 = 1$ almost everywhere. This condition is equivalent to the orthogonality in $L^2$ of the translated functions $(\phi(. - I_2))_{I_2 \in \mathbb{Z}}$.*
2. *There exists a $2\pi$-periodic function $m_0 \in L^2[0, 2\pi]$ such that: $\widehat{\phi}(\omega) = \widehat{m_0}\left(\frac{\omega}{2}\right)\widehat{\phi}\left(\frac{\omega}{2}\right)$ which is equivalent to the increase in the inclusion sense of the spaces $V_{I_1}$ defined by:*

$$\begin{array}{rcl} V_0 & = & \overline{sp\{\phi_{0I_2},\, I_2 \in \mathbb{Z}\}} = \left\{f = \sum_{I_2} c_{I_2}\phi_{0I_2} \mid \sum |c_{I_2}|^2 < \infty\right\} \\ V_{I_1} & = & \left\{h(a) = f(2^{I_1}a) = \sum_{I_2} c_{I_2}\phi_{I_1I_2}(a) \mid f \in V_0, \sum |c_{I_2}|^2 < \infty\right\}, \quad I_1 \geq 1. \end{array}$$

**Assumption 4.** *We also assume that there exists a map $\psi \in L^2$ which satisfies the following properties $(HO_2)$, where the applications $(\psi_{I_1 I_2})$ are defined by the same way as the $(\phi_{I_1 I_2})$ were:*

1. *$\sum_{I_2=-\infty}^{\infty} |\widehat{\psi}(\omega + 2I_2\pi)|^2 = 1$ a.e.*
2. *There exists a $2\pi$-periodic function $m_1 \in L^2[0, 2\pi]$ verifying: $\widehat{\psi}(\omega) = \widehat{m_1}\left(\frac{\omega}{2}\right)\widehat{\phi}\left(\frac{\omega}{2}\right)$ which means that $\psi \in V_1$.*
3. *$\sum_{I_2} \widehat{\phi}(\omega + 2I_2\pi)\overline{\widehat{\psi}(\omega + 2I_2\pi)} = 0$, which tells us that $V_0$ and $W_0$ are orthogonal, where we have defined:*

$$\begin{array}{rcl} W_0 & = & \left\{f = \sum_{I_2} c_{I_2}\psi_{0,I_2} \mid \sum |c_{I_2}|^2 < \infty\right\} \\ W_{I_1} & = & \left\{h(a) = f(2^{I_1}a) = \sum_{I_2} c_{I_2}\psi_{I_1I_2}(a) \mid f \in W_0, \sum |c_{I_2}|^2 < \infty\right\}, \quad I_1 \geq 1. \end{array}$$

4. *$m_1(\omega) = e^{-i\omega}\overline{m_0(\omega + \pi)}\mu(\omega)$, where $\mu$ is $2\pi$-periodic with $|\mu(\omega)| = 1$ a.e. This implies that $V_1 = V_0 \oplus W_0$.*

*These conditions are sufficient to prove that, more generally, for a given $I_1 \geq 0$, the functions $(\psi_{I_1 I_2})_{I_2}$ form an orthonormal basis of $W_{I_1}$, and that $V_{I_1+1} = V_{I_1} \overset{\perp}{\oplus} W_{I_1}$.*

The function $\phi$ is called father wavelet, whereas $\psi$ is the mother wavelet.

**Assumption 5.** *Hypothesis $(HO_3)$ : $\phi \in L^1 \cap L^2$. This assumption implies the continuity of the Fourier transform and the fact that $\overline{\left(\bigcup_{I_1 \geq 0} V_{I_1}\right)}^{L^2} = L^2$. Notice that this assumption is implied by the concentration property (H) enounced in section 1.3.2.*

When all the previous properties are satisfied, the sequence of spaces $(V_{I_1})_{I_1 \geq 0}$ generated by $\phi$ is called MultiResolution Analysis (MRA) of $L^2$. The sets $(W_{I_1})_{I_1 \geq 0}$ are then the resolution levels of the MRA.

Then, a wavelet decomposition is available for every $f \in L^2$:

$$\forall f \in L^2,\, \forall a \in \mathbb{R},\, f(a) = \sum_{I_2 \in \mathbb{Z}} \alpha_{I_2} \phi_{0 I_2}(a) + \sum_{I_1=0}^{+\infty} \sum_{I_2 \in \mathbb{Z}} \beta_{I_1 I_2} \psi_{I_1 I_2}(a) = \sum_{I_1=-1}^{+\infty} \sum_{I_2 \in \mathbb{Z}} \beta_{I_1 I_2} \psi_{I_1 I_2}(a), \quad (99)$$

where we have written $\phi_{0I_2} := \psi_{-1I_2}$ and $\beta_{-1I_2} = \alpha_{I_2}$ in order to simplify the formulas.

The coefficients of the wavelet expansion are:

$$\beta_{-1I_2} = \alpha_{I_2} = \int f(a) \phi_{0I_2}(a) da \text{ and } \beta_{I_1 I_2} = \int f(a) \psi_{I_1 I_2}(a) da. \quad (100)$$

Heuristically, we may consider $I_1$ as a degree of liberty in the frequency scale. The coefficients $\alpha$ thus sum up general shape of $f$ whereas the coefficients $\beta$ correspond to "details". The index $I_2$ corresponds to a translation parameter and can be viewed as a degree of liberty in the time scale.

Using Expansion (99), we can approximate any $f \in L^2$, by its sequence of projections on $V_{I_1^N}$:

$$P_{V_{I_1^N}} f(a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \beta_I \psi_I(a).$$

We can show that the projection operator $P_{V_{I_1^N}}$ on $V_{I_1^N}$, for $I_1^N \geq 0$ is associated with the kernel $K_{I_1^N}(x,y) = \sum_{I_1=-1}^{I_1^N - 1} \sum_{I_2 \in \mathbb{Z}} \psi_I(x) \psi_I(y)$: $\forall f \in L^2$, $P_{V_{I_1^N}} f(x) = \int K_{I_1^N}(x,y) f(y)\, dy$. It can also be proved that under Assumption $(H)$, $\int |K_{I_1^N}(x,y)|\, dy$ is bounded uniformly in $N$ by the positive constant $C$ appearing in $(H)$.

Finally, let us notice that when $\phi$ and $\psi$ satisfy Assumptions $(H)$ and $(M)$, the wavelet coefficients of any function $f \in B_{s,\infty,\infty}$, the Besov space of $s$-Hölder continuous functions, satisfy:

$$\exists C > 0,\, \sup_{I_2} |\beta_{I_1 I_2}| \leq C\, 2^{-I_1 \left(s + \frac{1}{2}\right)} \quad (101)$$

In fact, this inequality characterizes the functions of the Besov space $B_{s,\infty,\infty}$.

# B Presentation of the Warped Wavelet Regression Estimator

We now present the warped wavelet regression estimator inspired by the work of Kerkyacharian and Picard [18]. Then, we give a theorem dealing with its $L^2$-error.

Suppose we are trying to estimate the regression function $f$ in:

$$Y_i = f(\theta_i) + \varepsilon_i, \;\; i \in [1, N], \quad (102)$$

with independent identically distributed observations $(Y_i, \theta_i)_{i \in [1,N]} \in (\mathbb{R} \times \mathbb{R})^N$, with the same law as $(Y, \theta)$, and unobserved $(\varepsilon_i)_{i \in [1,N]}$, which are assumed to be centered and orthogonal to the sigma-field $\sigma(\theta)$ generated by the $(\theta_i)_{i \in [1,N]}$.

The theoretical solution, which minimizes the variance of the residuals $\varepsilon_i$, is the conditional expectation: $\mathbb{E}(Y|\theta = a) = \arg\min_f \mathbb{E}\,(Y - f(\theta))^2$.

In the following, we also suppose that the law $\mathcal{L}(\theta)$ of $\theta_i$ has an unknown density which we will call $g$. We denote by $G$ its distribution function, by $G^{-1}$ the usual generalized inverse of $G$ and by $G_N$ the empirical equivalent of $G$ as we did in (17).

The supports of $f$ and $g$ are also assumed to be imbedded in a compact interval $\mathcal{I}$.

We look for estimators $\widehat{f}$ for $f$ of the form:

$$\widehat{f}(a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2} \widehat{\beta}_I \psi_I(a) 1_{|\widehat{\beta}_I| \geq t_N}, \tag{103}$$

where $\widehat{\beta}_I$ are estimators of $\beta_I$ appearing in (99) and where $I_1^N$ corresponds to the truncation of the serie (99). Small terms corresponding to estimated coefficients $\widehat{\beta}_I$ with absolute values under the threshold $t_N$ are eliminated.

When dealing with deterministic and equi-spaced data $\theta_i$ ($\theta_i = i/N$), one classically uses the following estimator for $\beta_I$ (see Donoho and Johnstone [10] or Donoho and al. [11]):

$$\widehat{\beta}_I^{\boxtimes} = \frac{1}{N} \sum_{i=1}^{N} \psi_I \left( \frac{i}{N} \right) Y_i. \tag{104}$$

When the explicatives $\theta_i$ are random, the estimator (104) is no more valid.

Instead of looking for a complicated estimator, whose choice would be conditioned on the chosen wavelet base, Kerkyacharian and Picard [18] tried to stay close to the statistical data and therefore used warped wavelets to construct natural estimators. One can indeed reduce the problem to the case of uniformly distributed data on [0, 1] since $G(\theta_i)$ are uniformly distributed on [0, 1]. A natural extension of (104) is then:

$$\widehat{\beta}_I^* = \frac{1}{N} \sum_{i=1}^{N} \psi_I(G(\theta_i)) Y_i. \tag{105}$$

This is an unbiased estimator of the coefficient $\beta_I$ appearing in the following expansion of $f \circ G^{-1}$ on the wavelet base $(\psi_I)$ or equivalently of the expansion of $f$ on warped wavelets $(\psi_I \circ G)$:

$$f(a) = \sum_{I_1 \geq -1} \sum_{I_2} \beta_I \psi_I(G(a)).$$

A natural estimator of $f$ would then be:

$$\widehat{f}^*(a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \widehat{\beta}_I^* \psi_I(G(a)) 1_{|\widehat{\beta}_I^*| \geq t_N}, \tag{106}$$

with $t_N$ and $I_1^N$ appropriately chosen.

Since this transformation relies on $G$, which is in fact unknown, we replace it by $G_N(a)$. The new estimators of the coefficients are then:

$$\widehat{\beta}_I' = \frac{1}{N} \sum_{i=1}^{N} \psi_I(G_N(\theta_i))\, Y_i = \frac{1}{N} \sum_{i=1}^{N} \psi_I \left( \frac{i}{N} \right) Y_{(i)}, \tag{107}$$

where the $Y_{(i)}$ correspond to the observations ranked by increasing values of $\theta_i$. The ranking is made here to simplify notation: we have just re-indexed the sum.

A generalization of the estimator (106) is therefore:

$$\widehat{f}'(a) = \sum_{I_1=-1}^{I_1^N} \sum_{I_2 \in \mathbb{Z}} \widehat{\beta}'_I \psi_I(G_N(a)) 1_{|\widehat{\beta}'_I| \geq t_N}, \tag{108}$$

with $t_N$ and $I_1^N$ appropriately chosen.

Notice that the expression (107) is very similar to (104), which is nice, since the computation of (104) is often already implemented in statistical softwares.

# References


[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses, vol. 10, Société Mathématique de France, Paris, 2000.

[2] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes*, Ecole d'Ete de Probabilités de Saint-Flour XXII (Berlin), Lecture Notes in Mathematics, vol. 1581, Springer, 1992, pp. 1–114.

[3] A.G. Bhatt and R.L. Karandikar, *Invariant measures and evolution equations for Markov processes characterized via martingale problems*, The Annals of Probability **21** (1993), no. 4, 2246–2268.

[4] D. Bosq and J.-P. Lecoutre, *Théorie de l'estimation fonctionnelle*, Ecole Nationale de la Statistique et de l'Administration Economique et Centre d'Etudes des Programmes Economiques, Economica, 1987.

[5] M. Bossy, Ph.D. thesis, 1995.

[6] M. Pulvirenti C. Marchioro, *Hydrodynamics in two dimensions and vortex theory*, Commun. Math. Phys. **84** (1982), 483–503.

[7] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, Applied and Computational Harmonic Analysis **1** (1993), 54–81.

[8] P. Constantin and J. Wu, *Statistical solutions of the Navier-Stokes equations on the phase space of vorticity and the inviscid limit*, Journal of Mathematical Physics **38** (1997), no. 6, 3031–3045.

[9] I. Daubechies, *Ten lectures on wavelets*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, Pennsylvania, 1992.

[10] D.L. Donoho and I.M. Johnstone, *Ideal spatial adaptation via wavelet shrinkage*, Biometrika **81** (1994), 425–455.

[11] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian, and D. Picard, *Wavelet shrinkage: Asymptopia ?*, Journal of the Royal Statistical Society B **57** (1995), 301–369, with discussion.

[12] ———, *Density estimation by wavelet thresholding*, The Annals of Statistics **24** (1996), no. 2, 508–539.

[13] A. Dvoretsky, J. Kiefer, and J. Wolfowitz, *Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator*, Annals of Mathematical Statistics **27** (1956), 642–669.

[14] W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov, *Wavelets, approximation, and statistical applications*, Lecture Notes in Statistics, vol. 129, Springer, New York, 1987.

[15] B. Jourdain, *Diffusion processes associated with nonlinear evolution equations for signed initial measures*, Methodol. Comput. Appl. Probab. **2** (2000), no. 1, 69–91.

[16] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, 2 ed., Springer-Verlag, New York, 1991.

[17] G. Kerkyacharian and D. Picard, *Thresholding algorithms, maxisets and well-concentrated bases*, Test - Sociedad de Estadistica e Investigacion Operativa **9** (2000), no. 2, 283–344.

[18] ______, *Regression in random design and warped wavelets*, Preprint Laboratoire MODAL'X, 01 2003.

[19] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, Ecole d'Eté de probabilités de Saint-Flour XII (Springer, ed.), Lectures Notes in Math., vol. 1097, 1982, pp. 143–303.

[20] M. Ledoux, *Concentration of measure and logarithmic Sobolev inequalities*, Séminaire de Probabilités XXXIII (Berlin) (Springer, ed.), Lectures Notes in Math., vol. 1709, 1999, pp. 120–216.

[21] S. Mallat, *A wavelet tour of signal processing*, 2 ed., Academic Press, 1999.

[22] F. Malrieu and D. Talay, *Concentration inequalities for Euler schemes*, Preprint, 2003.

[23] Y. Meyer, *Ondelettes*, vol. 1 to 3, Hermann, 1990.

[24] F. Milinazzo and P.G. Saffman, *The calculation of large reynolds number two dimensional flow using discrete vortices with random walks*, Journal of Computational Physics **23** (1977), 380–392.

[25] S. Méléard, *Asymptotic behaviour of some interacting particle systems, McKean-Vlasov and Boltzmann models*, CIME Lectures, Lecture Notes in Mathematics, vol. 1627, Springer, 1996, pp. 45–95.

[26] ______, *A trajectorial proof of the vortex method for the two-dimensional Navier-Stokes equation*, Annals of Applied Probability **10** (2000), no. 4, 1197–1211.

[27] ______, *Monte-Carlo approximations of the solution of two-dimensional Navier-Stokes equations with finite measure initial data*, P.T.R.F. **121** (2001), no. 3, 367–388.

[28] ______, *Stochastic particle approximations for two-dimensional Navier-Stokes equations*, Preprint Modal'X, II Workshop on Dynamics and Randomness, 12 2002.

[29] S.T. Rachev, *Probability metrics and the stability of stochastic models*, (1991).

[30] D. Serre, *Matrices, theory and applications*, Graduate Texts in Mathematics, vol. 216, Springer, 2000.

[31] A.S. Sznitman, *Topics in propagation of chaos*, Ecole d'Ete de Probabilités de Saint-Flour XIX (Berlin), Lecture Notes in Mathematics, vol. 1464, Springer, 1991, pp. 165–251.

[32] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Analysis and Applications **8** (1990), no. 4, 94–120.

[33] D. Talay and O. Vaillant, *A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations*, The Annals of Applied Probability **13** (2003), no. 1, 140–180.

[34] O. Vaillant, *Une méthode particulaire stochastique à poids aléatoires pour l'approximation de solutions statistiques d'équations de McKean-Vlasov-Fokker-Planck*, Ph.D. thesis, Université de Provence, Marseille, France, 2000.

[35] M.J. Vishik and A.V. Fursikov, *Mathematical problems of statistical hydromechanics*, Mathematics and its Applications, Kluwer Academic Publishers, 1980.