

WAVELET DENOISING FOR POSTPROCESSING OF A 2D
PARTICLE - IN - CELL CODE *, **, ***, ****

SALIMOU GASSAMA¹, ÉRIC SONNENDRÜCKER¹, KAI SCHNEIDER², MARIE FARGE³ AND
MARGARETE O. DOMINGUES⁴

Abstract. In this paper, we aim at improving the accuracy of a Vlasov-Poisson solver using a 2D Particle-In-Cell (PIC) scheme, by denoising its charge density field. To this end, we have used an improvement of Donoho and Johnstone's wavelet denoising technique. To some extent, our work is a continuation of that performed by Chehab *et al.* [6]. Indeed, they made such a study in the one dimensional case and validated their analysis by considering the simulation of the Landau damping phenomenon. They concluded on the efficiency of the method in reducing the number of particles. However, our approach is quite different, since we do not use wavelets to directly interpolate the charge density, but we smooth the density field calculated by the PIC code. This is carried out via an iterative wavelet denoising technique introduced by Azzalini *et al.* [2]. Our work consists in studying the application of the method as a post-processing tool, in view of a future embedding into the PIC code. The results are the following: first, we showed that the hypotheses underlying the application of this method are valid. Secondly, we can infer from this study that it is possible to significantly reduce the amount of data needed for a simulation.

Résumé. L'objectif de ces travaux est d'accroître la précision d'un solveur Vlasov-Poisson basé sur un schéma de type Particle-In-Cell (PIC). Pour ce faire, à l'instar de Chehab et coll. [6], nous avons choisi d'utiliser des techniques d'analyse multirésolution par ondelettes pour débruiter la densité de charge. Cependant, à la différence de ces auteurs, nous effectuons un post-traitement de la densité de charge en appliquant une technique itérative mise au point par Azzalini et coll. [2]. Cette étude est un préalable à l'implémentation d'un algorithme de débruitage dynamique du code PIC. Les expériences numériques menée sur le cas-test de l'amortissement Landau nous ont permis de valider les hypothèses sous-jacentes et portent à croire qu'il est possible de réduire de manière significative le nombre de particules nécessaires pour une précision donnée.

* CIRM

** CEMRACS 2005

*** CNRS

**** DFG (Graduiertenkolleg 366)

¹ Institut de Recherche Mathématique Avancée, UMR CNRS 7501 Université Louis Pasteur, Strasbourg, France

² MSNM-CNRS & CMI, Université de Provence, 39 rue Joliot-Curie, 13453 Marseille cedex 13, France

³ LMD-CNRS, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris cedex 05, France

⁴ Laboratório Associado de Computação e Matemática Aplicada (LAC), Instituto Nacional de Pesquisas Espaciais (INPE), Av. dos Astronautas, 1758, 12227-010 São José dos Campos, Brasil

INTRODUCTION

We shall first recall the physical background of the problem at hand, as well as its mathematical modeling. Given the main purpose of this report, we will only point out the aspects which are useful to seize the core of the matter discussed in this paper. For a thorough insight into it, one could refer to e.g.: [5, 16, 17, 20] and the bibliographies therein. Though notations will be introduced as we encounter new expressions, it is convenient to note that from now on, all magnitudes denoted with a subscript e , i , or M , will respectively refer to those relative to electronic, ionic and model magnitudes.

As far as we are concerned, we will consider the kinetic model of plasmas of charged particles in which the collisions between particles are neglected. Moreover, due to the fact that we consider a non relativistic plasma in a quasi-stationary regime we will neglect the magnetic field. Thus the Vlasov-Poisson equation for a single particle species of charge q and mass m reads as follows:

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - q \cdot (E \cdot \nabla_p) f = 0, \quad (x, v) \in \mathbb{R}^d \times \mathbb{R}^d, \quad (0.1)$$

where $f(x, p, t)$ is the particle the distribution function with respect to position x , momentum $p = mv$ and time t . E is the electrostatic field. This equation models the invariance of the distribution function f along the trajectories of the particles moving through the electrostatic field. The self-consistent field produced by the charge of particles is:

$$E_{self}(x, t) = -\nabla_x \phi_{self}(x, t) \quad \text{with} \quad \begin{cases} -\Delta_x \phi_{self} & = \rho(x, t) \\ \rho(x, t) & = q \int f(x, p, t) dp \end{cases} \quad (0.2)$$

In order to confine the particles in a bounded domain, we consider the classical case of an additional external potential $\phi_{ext}(x)$, which can be generated e.g. by a constant background of ions such that:

$$E(x, t) = E_{self} + E_{ext} \equiv -(\nabla_x \phi_{self} + \nabla_x \phi_{ext}) \quad (0.3)$$

To close the systems to be dealt with, we shall also fix an initial distribution function, via:

$$f(x, p, 0) = f_0(x, p) \quad (0.4)$$

Particle simulations [5, 17] try to mimic this behavior at best. In this study, we use Particle-In-Cell (PIC) methods. PIC methods consist in modeling the medium as a finite set of individual elements (particles) interacting with one another. However, if we also assume (see [5]) that the Debye length is invariant, that is: $\lambda_{D,M} = \lambda_D$, then the model and real plasma temperatures will be different: $T_{e,M} = \alpha T_e$, though the thermal velocities remain the same:

$$v_{T,M} = \left(\frac{T_{e,M}}{m} \right)^{1/2} = v_T \quad (0.5)$$

Consequently, the mean free path in a PIC plasma is much smaller than that of a real plasma. As a consequence, the level of shot noise is much higher in the model plasma. A means of reducing this effect could consist in increasing the number of particles per cell in the Eulerian mesh. However, it proves to be highly computationally expensive.

Thus in this paper, we try to control the noise intrinsically generated by the model. More precisely, we aim at enhancing accuracy, meanwhile significantly reducing the number of particles per cell. That is why we used the optimal compression abilities of wavelets [10–12, 21], in order to have such an adaptive code. To meet these requirements, we passively denoised the PIC code via an ad-hoc postprocessing tool. This is a preliminary step before a future embedding of the procedure into the PIC code.

In what follows, we will first explain the numerical techniques used to solve the Cauchy problem formed by equations 0.1 to 0.4. Then in section 2, we will present the methods used for denoising. Finally, sections 3

and 4 are respectively devoted to the presentation and discussion of the application of these techniques to the classical Landau damping.

1. PLASMA MODELING VIA PARTICLE-IN-CELL METHODS

Now, we focus on the numerical schemes used to model the aforementioned physical problem. For a comprehensive presentation, the reader may refer to [3].

First and foremost, the initial distribution function $f_N^0(x, p)$, of a species is approximated via a linear combination of N Dirac mass tensor products associated to macro-particles with initial states (x_k^0, p_k^0) in phase space:

$$f_N^0(x, p) = \sum_{k=1}^N w_k \delta(x - x_k^0) \delta(p - p_k^0) \tag{1.6}$$

Besides, we draw the reader’s attention to the fact that the choice of plasma particle initial distribution in phase space coordinates is of great importance in discrete models, since it affects the possibility of an adequate investigation of the physical processes considered. Here, we use a Monte Carlo method called "quietstart" (see e.g. [17] for more details on statistical PIC methods). Then, each time step of the simulation is split into two phases:

During the first stage , the interactions or the result of the joint effect of particles with fixed state are calculated. This is done by solving the Poisson equation 0.2. To solve the Poisson equation, we use the Fourier transform method. For the Vlasov equation, we solve it by considering its characteristic curves:

$$\begin{cases} \frac{dX(t)}{dt} = V(t) \\ \frac{dV(t)}{dt} = \frac{q}{m} E(X(t), t) \end{cases} \tag{1.7}$$

This is the most difficult task, since we have to deal with a nonlinear coupling. The electric field known at the nodes of the Eulerian grid are interpolated to the positions (in phase space) of the particles. To interpolate this magnitude, we use shape functions $\{S_i\}$ associated to the Eulerian nodes $\{x_i\}$. If we denote by V_i the volume of the cell associated to node x_i , then we shall impose

$\sum V_i S_i(x) = 1 \quad \forall x \in \mathbb{R}^2$ to ensure charge conservation. Concerning the spline functions, we recall that a spline of order m can be obtained from the 0-order spline $1|_{[-1/2, 1/2]}$ via m convolutions of the latter by itself. The electric field is interpolated as being:

$$E(x_k(t)) = V_k \sum_i E(x_i, t) S_i(x_k(t)) \tag{1.8}$$

During the second stage, by using a leapfrog scheme, we can advance particles in two steps according to the Newton law for momentum:

$$\begin{cases} \frac{v_k^{n+1/2} - v_k^{n-1/2}}{\Delta t} = \frac{q}{m} E_k^n \\ \frac{x_k^{n+1} - x_k^n}{\Delta t} = v_k^{n+1/2} \end{cases} \tag{1.9}$$

Therefore, we have to calculate the new charge density $\{\rho_h(x_i, t)\}$ at the nodes of the Eulerian grid. This is the second interpolation and is carried out by using the same shape functions as above, in order to avoid the apparition of self-forces [18]:

$$\begin{aligned} \rho_N(x, t) &= q \int_{\mathbb{R}^2} f_N(x, p, t) dp \\ &= q \sum_{k=1}^N w_k \delta(x - x_k(t)) \\ \rho_h(x_i, t) &= \int_{\mathbb{R}^2} \rho_N(x, t) S_i(x) dx \\ &= q \sum_{k=1}^N w_k S_i(x_k(t)) \end{aligned} \tag{1.10}$$

Finally, knowing the charge density, we can compute the electric field for the next time step. Let us recall the fact that all the difficulty lies in the interpolation between Lagrangian and Eulerian meshes. Indeed, it may lead to charge non conservation. To solve this problem, several solutions have been suggested, such as constraints via Lagrange multipliers (see [3]). However, for the general Vlasov-Maxwell equations, methods calculating the current such that charge conservation is satisfied are noisier. To reduce numerical noise, one can use higher order shape forms. In the next section, we will introduce another technique to reduce that noise: wavelet filtering.

2. WAVELET DENOISING OF A 2D PIC CODE: TECHNICAL ASPECTS

2.1. Basics on wavelet theory

In what follows, we will recall some basic assumptions on wavelets, so that the paper is self-contained. Besides, we will only consider the case of discrete wavelet transforms. For further readings, we refer to [7,9,23].

In a nutshell, the wavelet transform decomposes the original data into a coarse approximation and a sequence of finer and finer details, keeping the size of the data constant. In this respect the wavelet transform has adaptive time/space frequency resolution. Moreover, it is also faster¹ than the Fast Fourier Transform. The framework of this approach is laid on a special construction of orthonormal bases of $L^p(\mathbb{R})$ with $1 < p < \infty$ generally taken to be equal to 2. Such a basis is obtained thanks to a couple (φ, ψ) of functions respectively called the *scaling* and the *wavelet* function. These bases are defined by an iterative algorithm of dilatation and translation of the functions φ and ψ . Such a procedure has led to the concept called multiresolution analysis (MRA). MRA gives rise to the multiscale local analysis of the signal and fast numerical algorithms. To give more details about this procedure, let us first recall the following definitions:

Definition 2.1. *Riesz² basis* of a Hilbert space: Let $(u_k)_{k \in \mathbb{Z}}$ be a sequence of a Hilbert space. It is then called a Riesz basis, if there exists two strictly positive constants A and B such that, for any finite sequence $(x_k)_{k \in I \subset \mathbb{Z}}$ we have:

$$A \|x\|_2 \leq \|(u_k x_k)_{k \in \mathbb{Z}}\|_2 \leq B \|x\|_2 \quad (2.11)$$

Definition 2.2. *Multiresolution analysis* of \mathbb{R} : it consists of a function $\varphi \in L^2(\mathbb{R})$ of unit norm and an increasing sequence $\{V_j\}_{j \in \mathbb{Z}}$ of closed linear subspaces of $L^2(\mathbb{R})$ defined by: $V_j = \text{span}(\{\varphi_{j,k} \equiv 2^{j/2} \varphi(2^j \cdot -k)\}_{k \in \mathbb{Z}})$, with the following properties:

- (1) The family $(\varphi_{0,k})_{k \in \mathbb{Z}}$ is a Riesz basis
- (2) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$
- (3) $\bigcup_{j \in \mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$
- (4) $(\forall j \in \mathbb{Z})(V_j \subset V_{j+1})$

We will also consider the following orthogonality hypothesis:

- (5) $(\forall j, k, l \in \mathbb{Z})(\langle \varphi_{j,k}, \varphi_{j,l} \rangle = \delta_{k,l})$

Remark 2.3. From the definition of a Riesz basis, we may infer from the first and the fourth statements of this definition, that there exists a sequence – called the *low-pass filter* – $(h_k)_{k \in I \subset \mathbb{Z}}$ of finite³ support, such that:

$$\varphi(\cdot) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \varphi(2 \cdot -k) \quad (2.12)$$

This equation is called *refinement equation*.

To construct orthogonal wavelets, let W_0 be the orthogonal complement of V_0 in V_1 , that is: $V_1 = V_0 \oplus W_0$. Then there exists a function, called the *mother of wavelets* ψ such that $(\psi_{0,k} \equiv \psi(\cdot - k))_{k \in \mathbb{Z}}$ is an orthonormal

¹ $O(N \log N)$ versus $O(N)$, N being a characteristic size of the system

²So far, the most general context for wavelet design lies on the notion of *frames*, but for the sake of concision and because it is sufficient, we will only refer to Riesz bases.

³In the case of a compactly supported φ function

basis of W_0 . Accordingly, $(\psi_{j,k} \equiv 2^{j/2}\psi(2^j \cdot -k))_{j,k \in \mathbb{Z}}$ is an orthonormal basis of W_j , where $V_{j+1} = V_j \oplus W_j$. Similarly to the case of φ , from $\psi \in W_0 \subset V_1$, we have another filter – the *high-pass filter* – $(g_k)_{k \in \mathbb{Z}}$ such that:

$$\psi(\cdot) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \varphi(2 \cdot -k) \tag{2.13}$$

If the low-pass filter h has D non-vanishing coefficients, then a permissible choice for g is : $(\forall k \in \mathbb{Z})(g_k = (-1)^k h_{D-k-1})$. Given this setting, we can decompose any function of $L^2(\mathbb{R})$ via its projections on these spaces. Let P_j and Q_j be respectively the orthogonal projections on spaces V_j and W_j . Then we have:

$$(\forall f \in L^2(\mathbb{R}))(P_{j+1}f = P_j f + Q_j f) \tag{2.14}$$

$$\text{With: } \begin{cases} P_j f &= \sum_{k \in \mathbb{Z}} \langle f, \varphi_{j,k} \rangle \varphi_{j,k} \\ Q_j f &= \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k} \end{cases} \tag{2.15}$$

Consequently, we have these two decompositions of $L^2(\mathbb{R})$:

$$L^2(\mathbb{R}) = \oplus_j W_j = V_{j_0} \oplus \oplus_{j \geq j_0} W_j \tag{2.16}$$

Then, the means to apply a discrete wavelet transform will only consist in performing a change of basis transformation between the bases of V_{j+1} and $V_j \oplus W_j$. By using respectively 2.12 and 2.13, in the expressions of $\varphi_{j,k}$ and $\psi_{j,k}$, we are led to:

$$\begin{cases} \varphi_{j,k} &= \sum_{l=2k}^{D+2k-1} h_{l-2k} \varphi_{j+1,l} \\ \psi_{j,k} &= \sum_{l=2k}^{D+2k-1} g_{l-2k} \psi_{j+1,l} \end{cases} \tag{2.17}$$

From which we deduce the following decomposition algorithm connecting coefficients of successive approximations:

$$(\forall f \in L^2(\mathbb{R})) \begin{cases} c_{j,k} \equiv \langle f, \varphi_{j,k} \rangle &= \sum_{l=2k}^{D+2k-1} h_{l-2k} c_{j+1,l} \\ d_{j,k} \equiv \langle f, \psi_{j,k} \rangle &= \sum_{l=2k}^{D+2k-1} g_{l-2k} d_{j+1,l} \end{cases} \tag{2.18}$$

Finally, thanks to 2.17, 2.14 and the orthogonality property, we obtain the reconstruction formula:

$$c_{j+1,k} = \sum_{l=0}^{D-1} h_{k-2l} c_{j,l} + \sum_{l=0}^{D-1} g_{k-2l} d_{j,l} \tag{2.19}$$

Last but not least, for practical reasons, one has to consider a finite subset of integers for the scale index j . So, let $J_c < J_f$ be the chosen coarsest and finest resolution levels, the applying 2.18 recursively yields:

$$V_{J_f} = V_{J_c} \oplus_{j=J_c}^{J_f-1} W_j \tag{2.20}$$

Thus, we have:

$$f \approx \sum_{k=0}^{D-1} \langle f, \varphi_{J_f,k} \rangle \varphi_{J_f,k} = \sum_{k=0}^{D-1} \langle f, \varphi_{J_c,k} \rangle \varphi_{J_c,k} + \sum_{j=J_c}^{J_f-1} \sum_{k=0}^{D-1} \langle f, \psi_{j,k} \rangle \psi_{j,k} \tag{2.21}$$

Remark 2.4. A means to make a multidimensional wavelet analysis of data consists in building tensor product spaces of the spaces $\{V_j\}_{j \in \mathbb{Z}}$. In dimension two, we have:

$$\forall j \in \mathbb{Z}, \quad V_{j+1} = V_{j+1} \otimes V_{j+1} \tag{2.22}$$

$$= (V_j \otimes V_j) \oplus (W_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes W_j) \tag{2.23}$$

As a consequence, we need one scaling function basis $(\varphi_{j,k_1} \otimes \varphi_{j,k_2})_{j,k_1,k_2 \in \mathbb{Z}}$ and three different wavelet functions $(\psi_{j,k_1} \otimes \varphi_{j,k_2})_{j,k_1,k_2 \in \mathbb{Z}}$, $(\varphi_{j,k_1} \otimes \psi_{j,k_2})_{j,k_1,k_2 \in \mathbb{Z}}$ and $(\psi_{j,k_1} \otimes \psi_{j,k_2})_{j,k_1,k_2 \in \mathbb{Z}}$ to span each resolution set V_j . From

a practical point of view, if we have to handle a square matrix, then we perform a 1D wavelet transform on each column and line for a given resolution level.

At this stage, it is necessary to recall that nowadays, there is a whole range of wavelet families available. Hence the necessity of knowing how to choose those which would best meet the requirements of the application at hand. These requirements may concern e.g. time/space localization, frequency localization, the speed of the transform, or the compression/reconstruction abilities. In our case, the latter aspect is of first importance. However, all these properties are interrelated. To obtain such and such property, one has to analyse the following characteristics:

- The size of the support of the wavelet function.
- The symmetry of the wavelet function.
- The number of vanishing moments⁴.
- The regularity of the wavelet and scaling functions.
- The orthogonality of the basis.

For the problem at hand, we should bear in mind that compression is all the more better than the wavelets have a large number of vanishing moments. However, due to the errors introduced by the thresholding process, we will need wavelets with high regularity in order to enhance the quality of the reconstructed data. Moreover, since we assume the additive noise to be well modeled by independent and identically distributed Gaussian variables, the transform should preserve the independence property. Yet, this can be granted by the use of orthonormal wavelets (see [23]).

Among the wavelets that meet the above stated requirements, are the classical Daubechies orthonormal wavelets with compact support. Actually, it can be proved that a Daubechies wavelet with r vanishing moments has μr ($\mu \approx 0.2$) continuous derivatives. However, we should be aware of the fact that as regularity increases, so does the support. Now, a small support is needed for a faster transform, while it implies relatively few vanishing moments and low regularity. This means that the requirements above cannot be satisfied equally well. Indeed, the size of the support, the symmetry of the wavelet function affect time/space localization, while the number of vanishing moments and the regularity affect the frequency localization. Good time/space localization requires a small support and high symmetry, while good frequency localization requires many vanishing moments and high regularity. A small support is also needed for a faster transform. However, a small support implies relatively few vanishing moments and low regularity. In addition, orthogonality implies asymmetry, except for the simplest wavelet (the Haar wavelets). (Bi-)orthogonality⁵ weakens the coupling between the properties of the decomposition and reconstruction wavelets, and allows perfect symmetry. In order to fulfill the requirements of the problem at hand we decided to choose a wavelet family called 'Coiflets', which has the advantage that both the scaling and the wavelet function have vanishing moments.

2.2. Wavelet denoising techniques

In this section, we shall detail the nonlinear wavelet denoising techniques that we used to denoise the charge density field. Our purpose is to significantly reduce the number of particles needed to perform the simulation described in section 2. Indeed, due to the numerical noise, so far good accuracy can be obtained only with a large number of particles. To denoise our 2D PIC code, we use the most important abilities of wavelets: multiresolution and compression. Actually, the wavelet transform cuts up the signal into different frequency components and then enables us to study each component with a resolution matched to its scale. The first application to denoising was developed by D. Donoho and I. Johnstone. For further readings, we refer to [10–12, 21] for wavelet denoising theory, and [13–15, 24, 25, 27] for its application to the study of some physical processes.

⁴A wavelet ψ has N vanishing moments when $\int_{-\infty}^{\infty} x^k \psi(x) dx = 0$, for $0 \leq k \leq N - 1$. Where x denotes time or space. In particular, all "normal" wavelets have zero mean (i.e. $N \geq 1$), since under rather general assumptions, this is related to the admissibility condition for the existence of the inverse transform.

⁵Biorthogonal wavelets use two wavelet bases both built according to the above mentioned MRA construction, except that the fourth hypothesis (orthogonality between translates of a same scale level) is not required.

Let us assume that we wish to estimate the signal s , which is corrupted with additive Gaussian white noise, i.e. we observe:

$$f = s + w \quad (2.24)$$

Where $w \in \mathcal{N}(0, \sigma)$ is a Gaussian white noise with variance σ . Then, if we denote by \check{s} an estimator of s , we define a *risk* associated to this estimation by the following:

Definition 2.5. Let κ be a convex even function, called the *cost*. Then the the *risk* R associated to this cost is defined by:

$$R(\check{s}, s) = \mathbb{E}_s [\kappa(\check{s}(f) - s)] \quad (2.25)$$

For a large class of cost functions, simply choosing the estimator \check{s} equal to the value f leads to minimizing the risk operator:

$$R(f, s) = \inf_{\check{s}} \sup_s R(\check{s}, s) \quad (2.26)$$

Whence the calling *minimax*, since we have the best estimation in the worse case. If we have to handle an evolution problem, that is s varies with the time variable, then the problem at hands becomes trickier. However, in this case, the knowledge of some a priori information about the regularity of s may enable us to find its best estimator. Let E be a known regularity space to which s belongs. Then the estimator should be the solution of:

$$\inf_{\check{s}} \sup_{\|s\|_E \leq C} R(\check{s}, s) \quad (2.27)$$

Despite estimators have been found for several functional spaces, such an approach is limited at least due to two obstacles. First, from a computational point of view, the convergence rate of those algorithms depends on the functional space. Secondly, the unknown function may simultaneously belong to several functional spaces, or worse: we often do not have any clue about its regularity (e.g.: a noisy image). It is in this context that Donoho and Johnstone introduced new methods based on wavelet decomposition. Indeed, thanks to the *unconditionality*⁶ property of wavelet bases, such methods allow us to have convergence rates of the same order of magnitude on a whole set of functional spaces. The payback is then a not exactly – though near to – minimax risk. The first method – called *Wavelet shrinkage* – introduced by these authors in [11] and consists in replacing the wavelet coefficients $\{d_{j,k}\}_{j,k \in \mathbb{Z}}$ of the noisy function by:

$$\text{sign}(d_{j,k}) \inf(|d_{j,k}| - \epsilon, 0) \quad (2.28)$$

Where the threshold ϵ is fixed with respect to the known or estimated variance of the noise. In the original article of these authors, an optimal value of the threshold was proved to be equal to $\epsilon = \sqrt{2\sigma \log N}$, where N corresponds to the number of grid points or the size of the sample. This value was calculated under the assumption that the observations are independent identically distributed random variables with $\mathcal{N}(0, \sigma)$ density. This method is as efficient as the minimax estimator in many functional spaces and minimizes the maximum L^2 -risk – i.e.: $\kappa \equiv \|\cdot\|_2^2$ – in a whole class of finite energy signals including Hölder and Besov spaces. In addition, in the worse case, the risk estimation error is approximatively $\log N$.

In this paper, we used an improvement of this method suggested by [2], in which the estimation of the variance is made via an iterative algorithm performed on the threshold residuals. More precisely, the following algorithm is applied:

(1) *Initialisation*

- Apply a fast wavelet transform to the initial data $\{s_i\}_{1 \leq i \leq N}$ to obtain the wavelet coefficients $\{\check{s}_i\}_{1 \leq i \leq N}$.
- Compute the variance σ_0 of these coefficients and set the initial threshold to be equal to $t_0 = \sqrt{2\sigma_0 \log N}$.

⁶An unconditional basis is such that the convergence of the partial sums obtained by decomposing a function on them do not depend on the order of the terms of the series

- Set the noise counter variable ν_{Noise} to be equal to N .
- (2) *Main loop*
At each step i , repeat:
 - Set $\nu'_{\text{Noise}} = \nu_{\text{Noise}}$ and set ν_{Noise} to be the number of coefficients such that their modulus is smaller than t_i .
 - Compute the variance σ_{i+1} from the wavelet coefficients inferior to t_i and set the new threshold t_{i+1} to the value $\sqrt{2\sigma_{i+1} \log N}$.
 - Increment the counter i .
- Until $\nu'_{\text{Noise}} = \nu_{\text{Noise}}$.
- (3) *Final step*
 - Use the inverse fast wavelet transform to compute the denoised data from the coefficients such that their absolute value is greater than the last threshold value.

3. APPLICATION TO PASSIVE DENOISING OF A PIC CODE: NUMERICAL RESULTS

3.1. The Landau damping case test

Here, we shall discuss the results obtained from these numerical experiments and sketch some perspectives for future extensions of the method. The Landau damping test is known for its efficiency to assess the quality of Vlasov–Poisson or Vlasov–Maxwell codes. For the sake of clarity, let us briefly recall the physical background related to it.

In stable plasmas electrostatic waves are damped due to wave–particle interactions, and not to collisions. This is the meaning of the Landau damping. Indeed, let ω denote frequency and k the wave vector. Then let us consider particles with velocities just larger than the wave phase velocity $u \gtrsim \omega/k$. They can gain or lose energy depending on the relative phase of the wave; but if they gain energy, their velocity increases and they go out of the resonance: they can not exchange energy any more. If they lose energy, they slow down and stay longer in the resonance. So overall, they transfer energy to the wave. The opposite holds for particles with velocities just below the phase velocity $u \lesssim \omega/k$. Those that gained energy from the wave remain in the resonance longer, and the net effect is that they gain energy from the wave. The total energy balance is therefore given by the ratio between how many particles gain energy from the wave (with $u \lesssim \omega/k$ and how many give energy to the wave ($u \gtrsim \omega/k$). This balance can be deduced from the slope of the distribution function $f_0(u)$ around the resonance $u \approx \omega/k$. This is the meaning of the formula:

$$\gamma \propto \left. \frac{df_0}{du} \right|_{u=\omega/k} \quad (3.29)$$

i.e. the damping is proportional to the slope of the equilibrium distribution function at the wave phase velocity (the wave–particle resonance).

For our numerical experiments, we consider a configuration with periodic boundary conditions and the initial distribution function is chosen so as to be a perturbed Maxwellian:

$$f_0(x, v) = \frac{n_0}{\sqrt{2\pi}v_T} \exp\left(-\frac{v_x^2 + v_y^2}{2v_T^2}\right) (1 + \alpha \cos(k_x x)) \quad (3.30)$$

The phase space domain is denoted by $\Omega \equiv [0, L_x] \times [0, L_y] \times \mathbb{R}^2$. Besides, perturbation amplitude α , the thermal velocity v_T , as well as the x–component k_x of the wave vector are constant. Consequently, this perturbation of period $L_x = \frac{2\pi}{k_x}$ creates a perturbed electric field such that the norm of its k_x -th mode decays with an exponential rate according to the law:

$$E_0 \exp(\gamma t) \sin(k_x - \omega t), \quad (3.31)$$

where under the condition $k_x^2 \lambda_D^2 \ll 1$, we have:

$$\gamma \approx -\sqrt{\frac{\pi}{8}} \frac{\omega_e}{(k_x \lambda_D)^3} \exp\left(-\frac{1}{2(k \lambda_D)^2} - \frac{3}{2}\right) \tag{3.32}$$

$$\omega \approx \omega_e \sqrt{1 + 3(k_x \lambda_D)^2} \tag{3.33}$$

In order to diminish numerical roundoff errors caused by huge or tiny magnitudes, the PIC code used here was implemented by considering dimensionless Vlasov–Poisson and Maxwell equations. To this end, we make the following changes of variable in equations 0.1 to 0.4:

$$\begin{cases} t' = \bar{t}t, & x' = \bar{x}x, & v' = \bar{v}v, & E' = \bar{E}E, \\ \rho' = \bar{\rho}\rho, & q' = q_e q, & m' = m_e m, \end{cases} \tag{3.34}$$

where $m_e = 9.109410^{-31}kg$ is the electron mass, and variables with a prime symbol are those expressed in the SI unit system. Those with a bar represent the new units. This leads us to the following dimensionless Vlasov equation:

$$\frac{\partial f}{\partial t} + \frac{\bar{v}\bar{t}}{\bar{x}}v \cdot \nabla_x f + \frac{\bar{x}\bar{t}q\omega_e^2}{\bar{v}m}E \cdot \nabla_v f = 0 \tag{3.35}$$

In order to fix the units of the variables, we state that:

$$\begin{cases} \bar{t} = \omega_e^{-1}, & \bar{v} = v_T = c/8, & \bar{x} = \frac{\bar{v}\bar{t}}{\epsilon_0} \\ \bar{w} = \bar{f}\bar{x}^2\bar{v}^2 = n_0\bar{x}^2, & \bar{\rho} = q_e n_0, & \bar{E} = \frac{\bar{x}\bar{\rho}}{\epsilon_0} \end{cases} \tag{3.36}$$

Where c is the speed of light and \bar{w} is the new mass unit of macroparticles. In this new unit system, we have:

$$\begin{cases} \bar{v} = 3.747410^7 m/s, & \bar{t} = 1.772710^{-2} n_0^{-1/2} s, & \bar{x} = 6.643010^4 n_0^{-1/2} m \\ \bar{E} = 1.202010^{-2} n_0^{1/2} V/m, & \bar{\rho} = 1.602110^{-19} n_0 C/m^2, & \bar{w} = 4.412910^{11} n_0 \end{cases}$$

We carry out the subsequent numerical experiments with the following values for the constants:

$$\alpha = 0.1, \quad k_x = 0.5, \quad \omega_e = 1, \quad \lambda_D = 1$$

In these conditions, the theoretical exponential decay rate of the the electrostatic field is predicted to be $\gamma \approx -0.151$ with an oscillation frequency of $\omega \approx 1.323$. As for the mesh grid, we used a uniform grid of 128×128 in (x, y) -coordinate plane. Finally, we consider this configuration with different particle densities: from 50 to 2000 particles per cell.

3.2. Numerical results of the passive denoising

In this part of the numerical experiment, we simulated the aforementioned case test with different numbers of particles per cell (nppc): from one to 2000. We use the Coiflet filters⁷, which yield wavelets (and scaling functions) of length 12, which yield wavelets with 4 vanishing moments. Then, we used a postprocessing program based on the algorithm described in section 3, to denoise the resulting fields.

To assess the time evolution of the noise, we studied the variation of some image quality parameters with respect to the number of particles per cell. Namely, we observed the time trend of the compression rate (in %), the signal to noise ratio SNR (in dB) and the variance ratio of the coherent and noisy solutions (in %). Let us recall that in what is following, coherent part and denoised signal are taken for synonyms. The same applies to incoherent part and noise component. A quantitative summary of the evolution of these features is provided by table 1.

⁷Let us recall that we perform our wavelet transform until the resolution level 1.

nppc	Compression rate: $100 \frac{NWC_{coh}}{NWC_{tot}}$ (%)	SNR: $10 \log(\frac{RMS_{inc}}{RMS_{tot}})$ (dB)	$100 \frac{RMS_{coh}}{RMS_{tot}}$ (%)
50	0.06	2.5	44.4
500	0.02	11.0	92.0
1000	0.15	14.8	96.7
1500	0.15	16.7	97.9
2000	0.18	18.1	98.4

TABLE 1. Summary of the compression rate and the SNR as a function of the nppc.

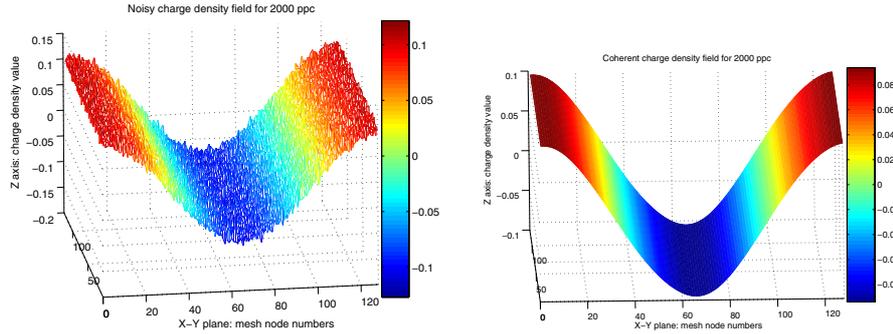


FIGURE 1. Surface plots of the charge density fields for 2000 ppc. Noisy PIC solution (left), denoised PIC solution (right)

Let us sum up the main deductions that can be drawn from the study of these data. We denote by NWC the number of coefficients of the wavelet decomposition, and by RMS the standard deviation of these coefficients. Subscripts 'coh', 'inc', and 'tot' respectively refer to the coherent, the incoherent and the whole signal. As it was announced in the previous sections, we observe that the incoherent part of the signal becomes all the more negligible that the $nppc$ increases (see also figures 3 and 4). As a consequence, a simulation with a high enough $nppc$ may be taken as a reference solution for the study of accuracy features. Another information given by this table, is that we need extremely few wavelet coefficients to reconstruct the solution (less than 0.2 %). It indicates that the solution is highly regular. Finally, we see that the signal is highly blurred for low $nppc$, since less than half of the total energy is considered as being coherent for 50 ppc.

Images on figure 1 are obtained with 2000 ppc and give us an insight of the exact solution. According to table 1 and the comments we made upon it, it seems reasonable to take the case of 2000 ppc as a reference to study the influence of the number of particles on the accuracy of the reconstruction. Thanks to the symmetry of the problem (whence of the solution), we could use cuts made at the middle mesh node number 64. In a first stage, we compared our reference solution with coherent solutions obtained with different $nppc$ (see the curves of figure 2). In the second stage, we compared all coherent solutions with their respective noisy counterparts. The results are furnished by the curves of figures 3 and 4.

As we can see through figure 2, even for the lowest $nppc$ (i.e. 1), the solution obtained is close to the reference solution. On this figure, we can distinguish three groups of quasi superimposed curves: those with 1 to 10 ppc, those with 25 to 100 ppc, and those with 250 to 2000 ppc. If we look at the number of coefficients kept to reconstruct the signal, we effectively ascertain the fact that it may be responsible for this disjunction into characteristic groups. Actually, the coherent part is reconstructed with 5 to 8 coefficients for the first group, 8 to 11 for the second and 24 to 29 for the last. An other criterion justifying this splitting is the SNR level, which goes from 0.07 to 0.55 for the first group, from 1.3 to 4.1 for the second and from 7.9 to 18.1 for the last. The first group corresponds to the very low $nppc$ ($SNR < 1$ and $\sigma \geq 0.09$), while the two others show less scattered ($\sigma < 0.02$) point sets. In addition, we can notice that the solution rapidly converges to the reference

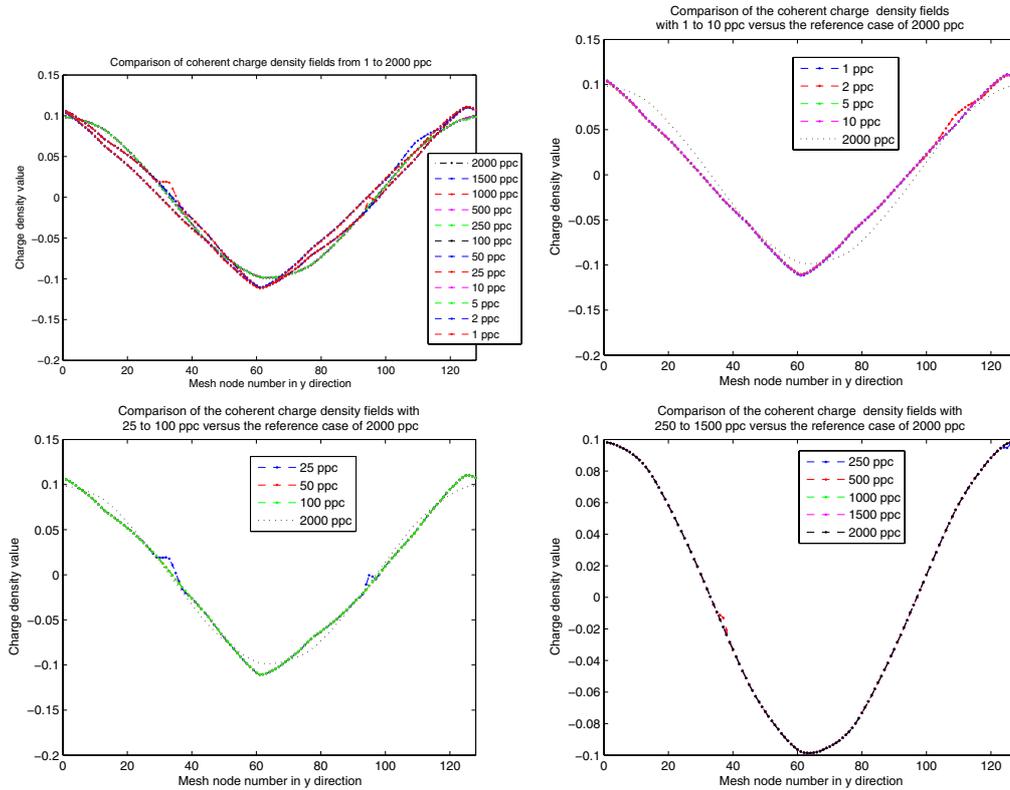


FIGURE 2. Comparison of the reference coherent PIC solution with those having less particles per cell. Cuts of the charge density field for different pppc

solution as soon as we have 250 ppc. This gives us incentives for reducing the number of particles per cell while enhancing accuracy.

Meanwhile, let us observe the curves of figures 3 and 4. They clearly give us two information. First, with the help of the curves related to the convergence feature (see figure 2), these results demonstrate the high quality of the reconstruction pattern chosen. Indeed, we tried such experiments with other (e.g. median filtering) techniques and they did not prove such efficiency.

Finally, in order to assess the validity of the normality assumption made on the noise distribution, we used Lilliefors’s non-parametric adjustment test (see [22]). This test is an adaption of the Kolmogorov–Smirnov test to the case where the variance and the mean of the data are unknown. It is based on the comparison between the empirical and theoretical distribution functions of the sample. It can be proved that this test is statistically more powerful than Pearson’s χ^2 -test and that it is more reliable than a simple kurtosis and skewness evaluation. Let us briefly recall its main principles:

- We standardize the data $(s_i)_{1 \leq i \leq N}$ according to the transformation: $s_i \mapsto z_i = \frac{s_i - m}{s}$. Where $m = \frac{1}{N} \sum_i s_i$ and $s = \frac{1}{N-1} \left[\sum_i s_i^2 - \frac{1}{N} (\sum_i s_i)^2 \right]$ are the empirical mean and unbiased RMS of the sample.
- We compute the cumulative distribution function C of the sample, via:

$$C(z_i) = \frac{\text{Number of values} \leq z_i \text{ in the sample}}{N}.$$

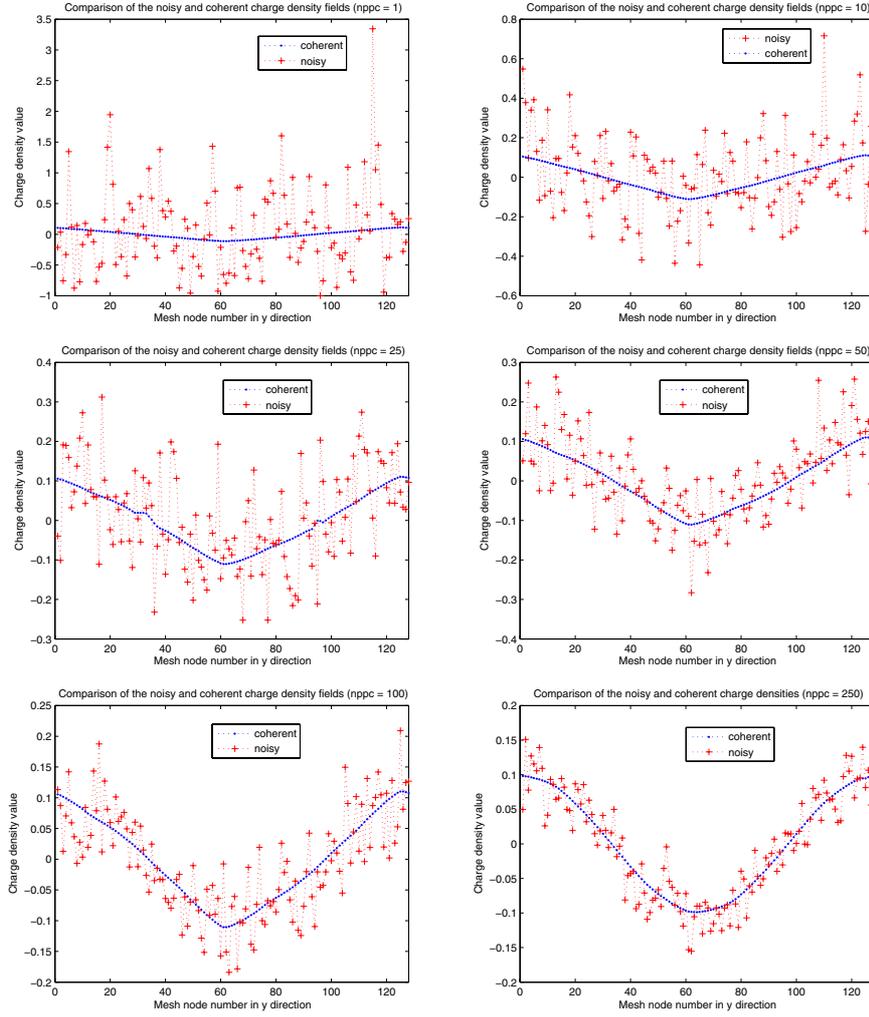


FIGURE 3. Comparison of the coherent PIC solutions with their noisy counterparts (from 1 to 250 ppc). Cuts of the charge density fields.

- In view of a bilateral test, we form the sequences $(\phi(z_i) - C(z_i))$ and $(\phi(z_i) - C(z_{i-1}))$ of differences between the empirical (C) and the theoretical (ϕ) cumulative distribution functions. The normal distribution with the estimated mean m and estimated variance σ is defined by the mapping $\phi : z = \frac{s-m}{\sigma} \mapsto \frac{1}{2} \left(1 + \operatorname{erf} \frac{z}{\sqrt{2}} \right)$.
- The Lilliefors parameter is equal to the maximum discrepancy in absolute value.

According to [22], the thresholds for rejecting the null hypothesis H_0 : 'the distribution is normal' with significance levels of 1% and 5% are respectively equal to $\frac{0.886}{\sqrt{N}}$ and $\frac{1.031}{\sqrt{N}}$. Namely, above these critical values, we can reject the null hypothesis with the corresponding significance level. Let us recall that the significance level of a test is the maximum probability of accidentally rejecting a true null hypothesis (a decision known as a Type I error). Table 2 gives the results of the test. Since N is always equal to 16384, the respective critical values are equal to 0.0081 and 0.0069. In addition to the Lilliefors test, we assessed the skewness $\gamma_1 = \frac{m_3}{m_2^{3/2}}$ and kurtosis

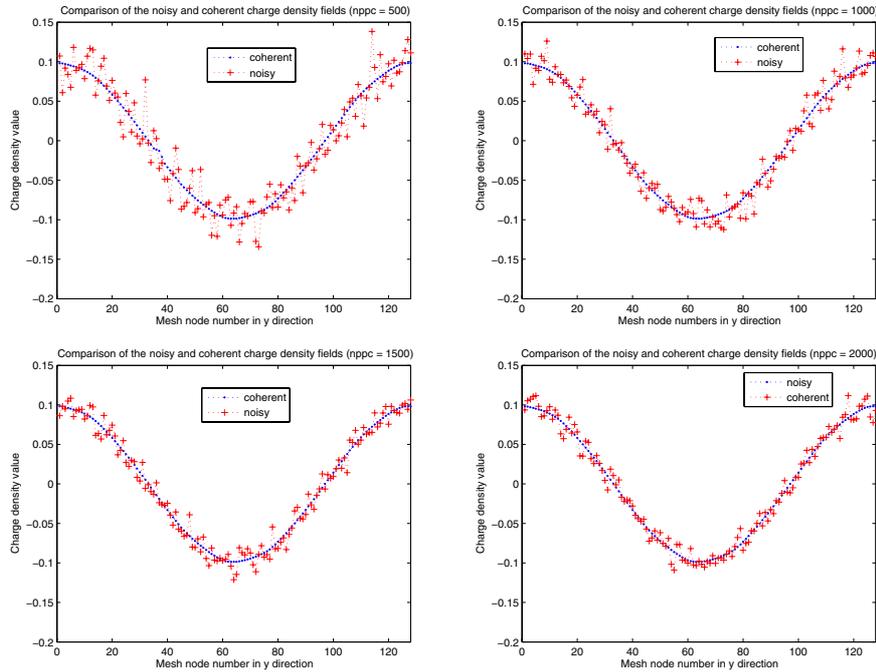


FIGURE 4. Comparison of the coherent PIC solutions with their noisy counterparts (from 500 to 2000 ppc. Cuts of the charge density fields.

$\gamma_2 = \frac{m_4}{m_2^2} - 3$ coefficients (with $m_k = \sum_{i=1}^N \frac{(s_i - m)^k}{N}$). For a normal distribution, we should have on the average: $\gamma_1 = 0$ and $\gamma_2 = 0$. According to the results recorded in table 2, the extracted noise can be considered as being Gaussian for $nppc > 50$. In spite of the (slight) failure of the Lilliefors test for the 2000 ppc case, we can ascertain the normality hypothesis. Indeed, figure 5 and 6 enable us to visually confirm the Gaussian trend of the distribution. For denoising methods related to non necessarily Gaussian noise, the reader may refer to e.g. to [1] and the bibliography therein.

nppc	Lilliefors' parameter	Skewness coefficient	Kurtosis coefficient
1	0.0592	0.8328	0.7759
2	0.0391	0.5292	0.2537
5	0.0265	0.3201	0.0492
10	0.0167	0.2397	0.1008
25	0.0093	0.1398	0.0613
50	0.0074	0.1076	0.0294
100	0.0072	0.0680	0.0027
250	0.0047	0.0240	-0.0061
500	0.0056	0.0322	-0.0329
1000	0.0045	0.0362	-0.0372
1500	0.0046	0.0320	-0.0548
2000	0.0086	0.0363	-0.1165

TABLE 2. Results of the normality tests: evolution of Lilliefors' parameter, along with the skewness and kurtosis coefficients for different nppc.

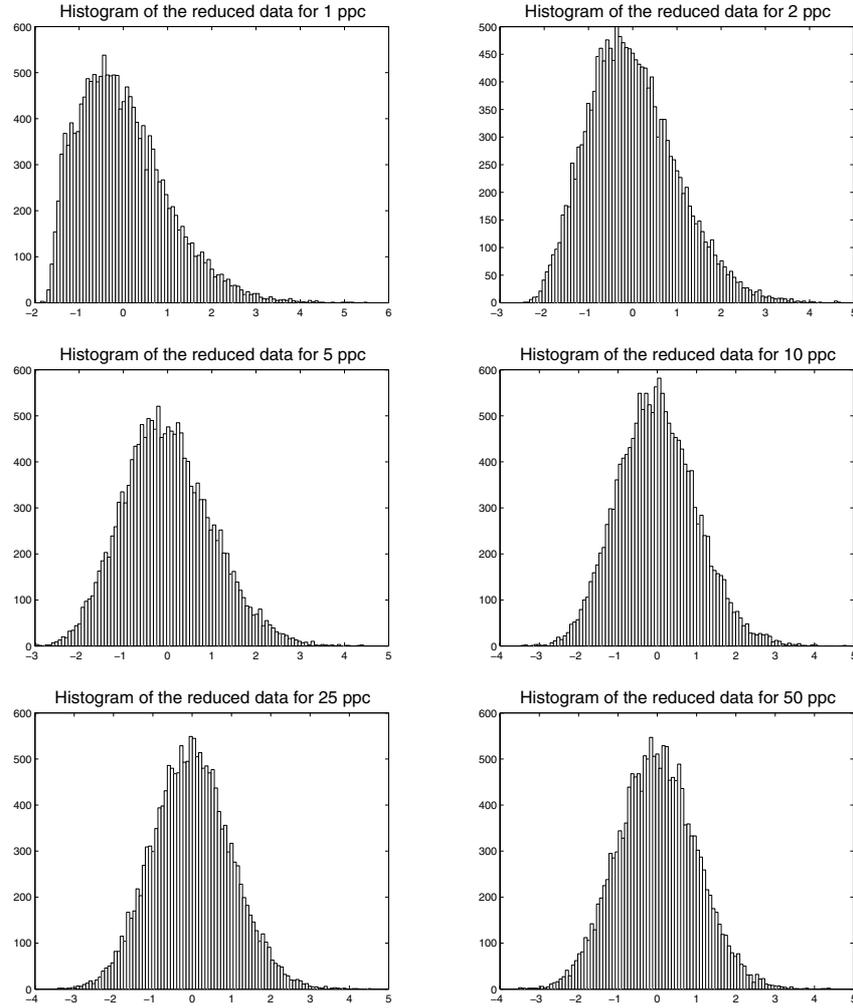


FIGURE 5. Noise distribution for different nppc using histograms with 100 bins. Note that the distributions have been centered and normalized by their variance, i.e. $z_i = \frac{s_i - m}{\sigma}$

4. CONCLUSIONS AND PERSPECTIVES

To conclude, this approach seems to be a promising step towards a significant reduction of the number of particles per cell. Moreover, it should allow us to enhance the accuracy of the scheme for a given number of particles per cell. Concerning ongoing and future work, several points should be pondered over. Actually, as for the embedding of this procedure into the PIC code, there still remains an important step before it could be achieved. Indeed, the denoising process may lead to large overshooting oscillations in the vicinity of some points. This naturally reminds us of the classical Gibbs phenomenon (see [19]) encountered in the case of Fourier series. Such problems occur when the convergence is not uniform over the interval considered. In the case of wavelets, a similar problem may happen, and it can be proved (cf. [26]) that such a phenomenon occurs for all known wavelet constructions, except for the Haar wavelets. Actually, due to the thresholding procedure, the reconstructed density has local discontinuities. This may be a consequence of the fact that the threshold is no longer adapted for such nppc. Namely, if we pay attention to the evolution of the threshold value with

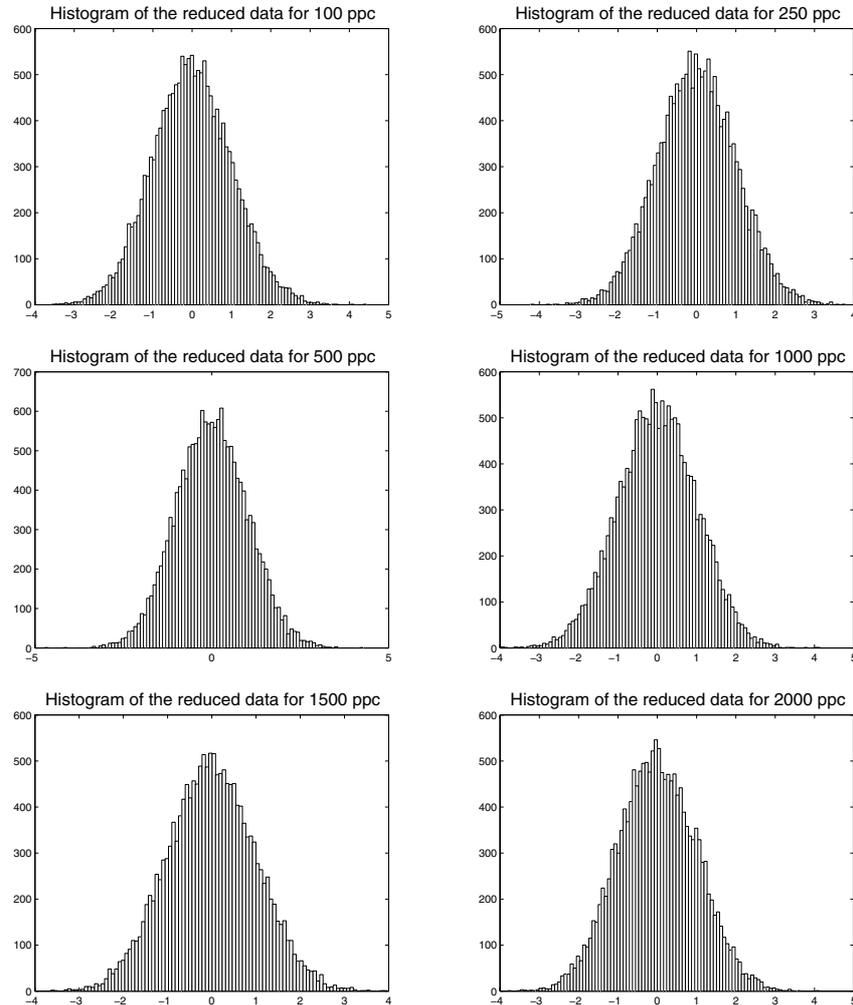


FIGURE 6. Noise distribution for different nppc using histograms with 100 bins. Note that the distributions have been centered and normalized by their variance, i.e. $z_i = \frac{s_i - m}{\sigma}$

respect to the nppc, we observe that it is all the more large, that the nppc is low. As a consequence, we observe strong oscillations in the vicinity of these loci. By incautiously injecting such non physical values in the iterative process, we are led to false solutions. A means to suppress this shortcoming could consist in the use of the so-called translation-invariant wavelets (see [4, 8, 26]). This approach consists in averaging the transform in the vicinity of these regions. After using this method, we observed that the oscillations diminish in amplitude. However, a thorougher study of this stage remains to be achieved both on theoretical and numerical implementation aspects. Last but not least, rigorous comparisons should be made with other types of solvers, in order to validate the quality of such improvements.

The first author is grateful to R. Barh elme who designed the initial 2D PIC code and allowed us to use it as a ground material for the project, and all the participants of the CEMRACS 2005 for their advice, help, and fruitful discussion.

Last but not least, since the project was also supported by TU Karlsruhe and ULP Strasbourg, we address our greatest thanks to these institutions.

REFERENCES

- [1] R. Averkamp and Houdré, *Wavelet thresholding for non (necessarily) Gaussian noise: a preliminary report*, CRM Proceedings and lecture notes, vol. 18, Centre de recherches mathématiques, American mathematical society, 1999, pp. 347–354.
- [2] A. Azzalini, M. Farge, and K. Schneider, *Nonlinear wavelet thresholding: A recursive method to determine the optimal denoising threshold.*, Appl. Comput. Harm. Anal. **18** (2005), no. 2, 177–185.
- [3] R. Barthelmé, *Le problème de conservation de la charge dans le couplage des équations de Vlasov et de Maxwell*, Ph.D. thesis, Université Louis Pasteur, Strasbourg, 2005, (in french).
- [4] K. Berkner and R.O. Wells ,Jr., *Smoothness estimates for soft-threshold denoising via translation-invariant wavelet transforms*, Applied and computational harmonic analysis **12** (2002), 1–24.
- [5] C.K. Birdsall and A.B. Langdon, *Plasma physics via computer simulation*, Series in plasma physics, Adam Hilger, Bristol, 1991.
- [6] J.-P. Chehab, A. Cohen, J. Roche, D. Jennequin, J.J. Nieto, and C. Roland, *Solution of Vlasov-Poisson equations using adaptive multiresolution methods*, CEMRACS 2003, IRMA Lectures in Mathematics and Theoretical Physics, EMS, pp. 29–42.
- [7] A. Cohen, *Numerical analysis of wavelet methods*, Elsevier, North-Holland, Amsterdam, 2003.
- [8] R.R. Coifman and D. Donoho, *Translation-invariant de-noising*, Lect. Notes Stat. **103** (1995), 125–150.
- [9] I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.
- [10] D. Donoho, *De-noising by soft-thresholding*, IEEE transactions on information theory **41** (1995), no. 3, 613–627.
- [11] D. Donoho and I. Johnstone, *Ideal spatial adaptation via wavelet shrinkage*, Biometrika **81** (1994), no. 3, 425–455.
- [12] D.L. Donoho and I.M. Johnstone, *Adapting to unknown smoothness via wavelet shrinkage*, Journal of the american statistical association **90** (1995), no. 432, 1200–1224.
- [13] I.M. Dremin, O.V. Ivanov, and V.A. Nechitailo, *Wavelets and their use*, Physics-Uspekhi **44** (2001), no. 5, 447–478.
- [14] M. Farge, *Wavelet transforms and their applications to turbulence*, Ann. Rev. Fluid Mech. **24** (1992), 395.
- [15] M. Farge, G. Pellegrino, and K. Schneider, *Coherent vortex extraction in 3D turbulent flows using orthogonal wavelets*, Phys. Rev. Lett. **87** (2001), no. 5, 45011–45014.
- [16] R.J. Goldston and P.H. Rutherford, *Introduction to plasma physics*, Institute of physics, Bristol, 1995.
- [17] Y.N. Grigoryev, V.A. Vshivkov, and M.P. Fedoruk, *Numerical “Particle-in-Cell“ Methods*, VSP BV, Utrecht, Boston, 2002.
- [18] J.W. Hockney and J.W. Eastwood, *Computer simulations using particles*, Institute of physics, Bristol, 1988.
- [19] A.J. Jerri, *Reducing the Gibbs phenomenon in a Fourier-Bessel series, Hankel and Fourier transform*, Contemporary mathematics **190** (1995), 179–194.
- [20] N.A. Krall and A.W. Trivelpiece, *Principles of plasma physics*, San Francisco Press, San Francisco, 1986.
- [21] H. Krim, D. Tucker, S. Mallat, and D. Donoho, *On denoising and best signal representation*, IEEE transactions on information theory **45** (1999), no. 7, 2225–2238.
- [22] H.W. Lilliefors, *On the Kolmogorov-Smirnov test for normality with mean and variance unknown*, J. Am. Stat. Assoc. **62** (1967), 399–402.
- [23] S. Mallat, *A wavelet tour of signal processing*, Academic Press, London, 1999.
- [24] O. Roussel, K. Schneider, and M. Farge, *Coherent vortex extraction in 3D homogeneous turbulence: comparison between orthogonal and biorthogonal wavelet decompositions*, J. of Turbulence **6** (2005), no. 1.
- [25] K. Schneider, M. Farge, G. Pellegrino, and M. Rogers, *Coherent vortex simulation of 3D turbulent mixing layers using orthogonal wavelets*, J. Fluid Mech. **534** (2005), no. 10, 39–66.
- [26] H.-T. Shim, *On the Gibbs phenomenon for wavelet expansions*, Journal of approximation theory **84** (1996), 74–95.
- [27] C. Torrence and G.P. Compo, *A practical guide to wavelet analysis*, Bulletin of american meteorological society **79** (1998), no. 1, 61–78.