C. R. Physique 12 (2011) 123–131

Contents lists available at ScienceDirect

Comptes Rendus Physique

www.sciencedirect.com

Propagation and plasmas: new challenges, new applications

Coherent vorticity extraction in resistive drift-wave turbulence: Comparison of orthogonal wavelets versus proper orthogonal decomposition

Filtrage en ondelettes orthogonales pour l'étude de la turbulence dans les plasmas : Comparaison entre la méthode CVE et la méthode POD

Shimpei Futatani [a], Wouter J.T. Bos [b], Diego del-Castillo-Negrete [c], Kai Schneider [d], Sadruddin Benkadda [a], Marie Farge [e,*]

[a] *International Institute for Fusion Science, CNRS - Université de Provence, case 321, 13397 Marseille cedex 20, France*
[b] *LMFA, CNRS UMR 5509, École Centrale de Lyon - Université de Lyon, 69134 Ecully cedex, France*
[c] *Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[d] *M2P2-CNRS and CMI, Université de Provence, 38 rue Frédéric Joliot-Curie, 13453 Marseille Cedex 13, France*
[e] *LMD-CNRS, École Normale Supérieure, 45, rue d'Ulm, 75230 Paris cedex 05, France*

ARTICLE INFO

Keywords:
Proper orthogonal decomposition
Wavelets
Plasma turbulence
Coherent structures

Mots-clés :
POD
Ondelettes
Turbulence plasma
Structures cohérentes

ABSTRACT

We assess two techniques for extracting coherent vortices out of turbulent flows: the wavelet based Coherent Vorticity Extraction (CVE) and the Proper Orthogonal Decomposition (POD). The former decomposes the flow field into an orthogonal wavelet representation and subsequent thresholding of the coefficients allows one to split the flow into organized coherent vortices with non-Gaussian statistics and an incoherent random part which is structureless. POD is based on the singular value decomposition and decomposes the flow into basis functions which are optimal with respect to the retained energy for the ensemble average. Both techniques are applied to direct numerical simulation data of two-dimensional drift-wave turbulence governed by Hasegawa–Wakatani equation, considering two limit cases: the quasi-hydrodynamic and the quasi-adiabatic regimes. The results are compared in terms of compression rate, retained energy, retained enstrophy and retained radial flux, together with the enstrophy spectrum and higher order statistics.

RÉSUMÉ

Cet article compare deux méthodes permettant d'extraire les contributions cohérentes dans les écoulements turbulents : la méthode CVE (Coherent Vorticity Extraction), basée sur la représentation en ondelettes, et la méthode POD (Proper Orthogonal Decomposition). La première méthode, CVE, décompose l'écoulement en base d'ondelettes orthogonales puis, grâce au filtrage des coefficients d'ondelettes, permet de séparer l'écoulement entre une contribution cohérente ayant un comportement statistique non Gaussien et un écoulement résiduel aléatoire ne présentant pas de structures. La seconde méthode, POD, est basée sur la décomposition en valeurs singulières. Elle décompose l'écoulement dans la base de fonctions qui retient le mieux l'énergie d'un ensemble de réalisations de l'écoulement. Ces deux méthodes sont ensuite utilisées pour analyser les résultats de simulations

* Corresponding author.
E-mail address: farge@lmd.ens.fr (M. Farge).

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2010.12.004

numériques directes calculant un écoulement turbulent bidimensionnel d'ondes de dérive, gouverné par l'équation d'Hasegawa–Wakatani, ceci pour deux cas limites : le régime quasi-hydrodynamique et le régime quasi-adiabatique. Les résultats sont comparés en fonction des taux de compression, des proportions d'enstrophie et de flux radial retenues, ainsi que du spectre d'enstrophie (statistique d'ordre deux) et de statistiques d'ordre plus élévé.

1. Introduction

Experimental and numerical investigations of plasma and fluid turbulence generate very large amounts of data. The use of state-of-the-art tools to analyze these data plays a key role in the understanding of the relevant physical phenomena, in particular to study the role of coherent structures (CS) which emerge within turbulent fluctuations and correspond to organized vortical motions of the flow (see, e.g., [1]). These structures are responsible for turbulent mixing, but also for turbulent transport of heat, mass and momentum. They are efficient in trapping particles for long times. The identification and extraction of coherent structures are nowadays one of the main goals of turbulence research and we will compare two methods to perform this.

Many mathematical techniques have been used to identify and describe coherent structures in fluid and plasma turbulence, either the Fourier transform, the proper orthogonal decomposition [2–4], the wavelet transform [5–8], or the wavelet packet transform [9]. Each of these techniques decomposes the flow in different ways and to do so considers different fields, e.g., velocity, pressure, or their gradients. The question then arises: which method is the best to detect coherent structures? This work is dedicated to explore this question. We report here on the advantages and disadvantages of the proper orthogonal decomposition (POD) and of wavelet methods to identify and extract coherent structures out of turbulent flows. We will also compare the data compression thus obtained.

Proper orthogonal decomposition, which is based on the Karhunen–Loève orthogonal decomposition, has been extensively used in fluid flows for the detection of coherent structures and the construction of reduced models since the pioneering work of Lumley [2,3,1]. Recently, it has also been applied to plasma physics. For example, Ref. [4] showed that POD can be used to unveil the spatio-temporal multiscale properties of plasma turbulence and transport. A key signature of the POD approach is that the spatio-temporal modes are empirical in the sense that they are data dependent, i.e., the modes are not given a priori, as is the case for Fourier and wavelet methods, but are computed from an ensemble of realizations. Such flexibility allows the construction of optimal representations of the data leading to an efficient multiscale analysis. Regarding data compression, it was shown in [10] that POD techniques provide very efficient methods to compress three-dimensional magnetic field data of interest to magnetically confined fusion plasmas. POD techniques have also been used to efficiently reduce the noise in particle based plasma calculations [11].

In the POD procedure a large amount of flow field realizations serve to extract the most representative modes having a major influence from an energetic view point, both on the mean flow and on the turbulent fluctuations intensity fields, but only statistically. This means that the most representative modes in the POD sense are statistical observables which are not encountered in each flow realization. Moreover, if the flow is statistically homogeneous, the POD most representative modes are spread all over space (e.g., Fourier modes) while the coherent structures observed in each flow realization are well-localized and translate in space (i.e., their positions change from one realization to the other which results in the statistical homogeneity).

In contrast to POD, the wavelet based coherent vortex extraction (CVE) method does not require an ensemble of realizations since wavelet modes are given a priori, as it is the case for Fourier modes. It then separates each turbulent flow realization into two orthogonal components: the coherent structures and a noise-like background flow. The procedure is computationally very efficient since the fast wavelet transform used for this has an operation count which scales as the resolution, which is thus asymptotically faster than the FFT. It has been applied to several fluid flows in two dimensions [6] and in three dimensions [7,8]. Ref. [12] compares POD and CVE methods for extracting coherent structures out of three-dimensional homogeneous isotropic fluid turbulent flows and shows that, for the same compression rate, the former separates the flow into large scale and small scale components while the later separates the flow into coherent and incoherent components which are spread all over scales but exhibit different probability distributions and power-law spectra.

One important issue in fusion research is the understanding and control of the turbulent radial flux of particles and heat in magnetized plasmas, in order to improve the confinement properties of fusion devices. Indeed, the coherent structures which are observed in such flows have a large influence on those radial fluxes. The CVE method has been applied to experimental signals of ion density measured in the scrape-off layer (SOL) of the tokamak Tore-Supra in Cadarache to separate the coherent bursts from an incoherent background noise [13]. It has also been applied to numerical experiments of drift-wave turbulence [14]. Another application of wavelets to plasma physics is the denoising of particle based computations which has been proposed in [15] and recently compared to POD denoising technique in [16].

The goal in this article is to compare the POD and CVE methods for extracting coherent structures out of fluid or plasma turbulent flows, and to assess their differences and advantages. Both techniques have already been applied to drift-wave

turbulence [17,4,18,14], but without comparing the two methods on the same test cases. The paper is organized as follows. Section 2 briefly describes the resistive drift-wave turbulence model in the Hasegawa–Wakatani approximation which is used to generate the flows which will be taken as test cases. Section 3 presents a summary of the proper orthogonal decomposition and of the wavelet technique used to analyze the flow datasets. Section 4 describes the results obtained with the two methods and compares them.

2. Model equations – resistive drift-wave turbulence

In this section we present a brief discussion of the resistive drift-wave model used for modeling electrostatic turbulence in the plasma edge. The model of our direct numerical simulations (DNS) of resistive drift-wave turbulence is the Hasegawa–Wakatani (HW) system in a slab geometry.

The standard form of the two-dimensional HW model is given by [19]

$$\left(\frac{\partial}{\partial t} - D\nabla^2\right)n + \frac{\partial \phi}{\partial y} + c(n-\phi) = [n,\phi] \quad \text{and} \quad \left(\frac{\partial}{\partial t} - \nu\nabla^2\right)\nabla^2\phi + c(n-\phi) = \left[\nabla^2\phi, \phi\right] \tag{1}$$

where n is the fluctuating component of the plasma density, ϕ is the electrostatic potential, D is the diffusion coefficient and ν is the kinematic viscosity. The Poisson brackets are represented as $[f,g] = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}$ and c is the adiabaticity parameter which is related to parallel electron conductivity. It is defined as $c = (T_e k_z^2)/(e^2 n_0 \eta \omega_{ci})$ with the electron temperature T_e, the electron resistivity η and the ion cyclotron frequency ω_{ci}. The above equations assume that the equilibrium magnetic field is constant in the z-direction, $\boldsymbol{B} = B_0\hat{\boldsymbol{e}}_z$. The x- and y-directions correspond to the radial and poloidal coordinates. The ions are considered as cold and temperature gradient effects are neglected. The equilibrium plasma density n_0 is assumed to be homogeneous in the poloidal direction and to decrease in the radial direction. The equilibrium density length scale $L_n \equiv n_0/|\mathrm{d}n_0/\mathrm{d}x|$ is constant. In writing the system (1) we use dimensionless variables defined as

$$\frac{e\phi}{T_e} \to \phi, \qquad \frac{n_1}{n_0} \to n, \qquad \frac{c_s}{L_n}t \to t, \qquad \frac{x}{\rho_s} \to x \tag{2}$$

where ρ_s is the hybrid Larmor radius and c_s is the sound speed. The HW system is linearly unstable and exhibits a turbulent saturated regime in its nonlinear phase which corresponds to fully developed turbulence. Let us note that for $c \gg 1$, we get the adiabatic regime where the Hasegawa–Wakatani system reduces to the Hasegawa–Mima equation [20]. In this case, there is no phase shift between n and ϕ fields, therefore there is no transport. The case $c \approx 1$ represents the quasi-adiabatic regime which is more realistic as it allows a phase shift between n and ϕ so that transport can occur. For $c \ll 1$ the HW system reduces to a two-dimensional Navier–Stokes equation for the $E \times B$ flow vortices and passive advection of the density fluctuations. The latter case is called the hydrodynamic regime.

3. POD versus wavelet decomposition methods

3.1. Proper orthogonal decomposition

Many multivariate signal processing techniques have been developed in various domains of research. In the following we shall consider one particular technique, which is called proper orthogonal decomposition (POD). Proper orthogonal decomposition is known under several other names depending on the community: Karhunen–Loève expansion in statistical physics, Principal Component Analysis (PCA) or Eigen Orthogonal Functions (EOF) in meteorology, and Singular Value Decomposition (SVD) in numerical linear algebra. In comparison with the Fourier transform, POD gives by construction the optimal modal representation in terms of the energy content. Moreover, Fourier analysis has some disadvantages. The fields are represented as a linear combination of plane waves whose shape is given a priori. However, information related to both position and time in physical space is completely hidden among the phases of all Fourier modes. This is an advantage when dealing with waves, but it is no more so when dealing with localized spatial structures like convective coherent vortices. When the dependence is harmonic, Fourier methods work well. However, in the presence of intermittent transport, the turbulent flux exhibits strong intermittency (e.g., in the edge of tokamak plasmas), the time evolution presents bursts of activity that are ill resolved by Fourier analysis. Advection of impurities by a turbulent plasma flow exhibits a complex spatio-temporal dynamics and we then need information about transfers between different scales to model the flow and to get further physical insight. This calls for a representation that decomposes fields into contributions of different scales as well as different locations. POD allows us to decompose the fields into a basis that optimally represents a flow in the energy norm. It is a powerful technique to define basis functions that represent the ensemble averaged structures of turbulent flows.

In the POD procedure the modes are defined a posteriori from all realizations. This is in contrast to wavelet based CVE, which performs the separation a priori and selects the retained modes from a given set of basis functions, the orthogonal wavelets, as those having the strongest coefficients. For this it uses a threshold which depends only on the total enstrophy and on the sampling of each flow realization. Hence the selection procedure is nonlinear, as the retained basis functions depend on the modulus of the wavelet coefficients of each flow realization. From a statistical point of view, the CVE method is based on a Bayesian approach while POD is based on a non-Bayesian (also called frequentist) approach.

Let $A(x, y)$ be a scalar field representing one of the components of Eqs. (1) and (2), where x- and y-components are radial and poloidal direction. It is convenient to define a two-dimensional rectangular grid (x_i, y_j) for $i = 1, N_I$ and $j = 1, N_J$. The discretized field data specified at these grid points are denoted by the $N_I \times N_J$ matrix $A_{ij} = A(x_i, y_j)$. The maximum number of terms in the POD of A_{ij} is $N_{POD} = \min(N_I, N_J)$. The fundamental SVD theorem states that any two-dimensional matrix A_{ij} may be expanded as follows [21]:

$$A_{ij} = \sum_{k=1}^{N_{POD}} \sigma_k u_k(x_i) v_k(y_j) \tag{3}$$

The weights of the singular values σ_k are either positive or equal to zero and conventionally they are sorted in descending order that $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_k$. Coherent structures that are highly correlated in space appear in heavily weighted components because they have higher energy. On the other hand, random fluctuations appear in modes that carry small weights.

The orthonormality of POD modes reads:

$$\sum_{i=1}^{N_{POD}} u_k(x_i) u_l(x_i) = \sum_{j=1}^{N_{POD}} v_k(y_j) v_l(y_j) = \delta_{kl} \tag{4}$$

The POD basis functions are optimal to study turbulence from an energy viewpoint since the basis functions are eigenfunctions of the two-point correlation tensor of an ensemble of realizations, which is not the case of Fourier or wavelet bases. This method is particularly appropriate when analyzing complex physical systems, where different dynamical regimes coexist. POD allows us to identify these regimes and to characterize their energetics and their mean spatial structure. The key to the interpretation of a proper orthogonal decomposition is an examination of the weight distribution of the singular values. This allows us not only to assess the degree of compressibility of the data, but also provides valuable information on properties such as mean spatio-temporal symmetries. Several parameters are used to quantify the weight distribution. The total energy is defined as the sum of the squared weights of the singular values,

$$E_\sigma = \sum_{k=1}^{N_{POD}} \sigma_k^2 \tag{5}$$

The dimensionless energy is $p_k = \sigma_k^2 / E_\sigma$ which measures the relative amount of energy and is a useful quantity for comparing different data sets. It has the properties of a probability distribution

$$0 \leqslant p_k \leqslant 1 \quad \text{and} \quad \sum_{k=1}^{N_{POD}} p_k = 1 \tag{6}$$

In this context we also introduce the normalized entropy of the signal, defined as the entropy normalized by its maximum possible value

$$H = -\frac{1}{\log N_{POD}} \sum_{k=1}^{N_{POD}} p_k \log p_k \tag{7}$$

This parameter measures the level of complexity of the data. It is equal to zero, if and only if the signal is separable. The maximum value $H = 1$ is reached when all the components have equal weights for the singular values. The POD algorithm can be summarized in the following three-step procedure:

- *Decomposition*: compute the POD weights σ_k using the proper orthogonal decomposition for each time step;
- *Truncation*: discarding the coefficients with absolute values smaller than a certain threshold;
- *Reconstruction*: reconstruct the coherent vorticity field from the retained POD modes.

The computational cost of POD scales with the cube of the number of data points N, which is due to the SVD decomposition. In practice for the two-dimensional datasets considered in this paper the computational cost remains negligible. The memory necessary to store one POD mode is equal to the number of data points N, thus the memory requirement of the retained coherent modes is N times their number.

3.2. Wavelet decomposition

Definitions and details about the orthogonal wavelet transform and its extension to higher dimensions can be found, e.g., in Refs. [5] and [22]. In the following, we fix the notation for the orthogonal wavelet decomposition of a two-dimensional scalar valued field taken at a fixed time. The wavelet transform unfolds the field into scales, positions, and directions using a set of dilated, translated, and rotated basis functions, called wavelets. Each wavelet is well-localized in space, oscillating

(i.e., it has at least a vanishing mean, or better its first m moments vanish), and smooth (i.e., its Fourier transform exhibits fast decay for wavenumbers tending to infinity). We apply the coherent vorticity extraction algorithm [7,6] here using orthogonal wavelets. In dimension two, orthogonal wavelets span three directions (horizontal, vertical, and diagonal), due to their tensor product construction. To go from one scale to the next, wavelets are dilated by a factor 2 and the translation step doubles at each scale accordingly. Wavelet coefficients are thus represented on a dyadic grid [5].

We apply the CVE algorithm to the vorticity fields $\omega = \nabla^2\phi$ of both the quasi-hydrodynamic and the quasi-adiabatic cases. The extraction is performed from the vorticity since enstrophy is an inviscid invariant in the hydrodynamic limit. Moreover, vorticity is Galilean invariant in contrast to velocity and stream function. We consider the quasi-stationary state of the simulations, i.e., when a saturated regime is reached. We then project the vorticity field, given at resolution $N = 2^{2J}$, onto a wavelet basis,

$$\omega(x, y) = \sum_{\lambda\in\Lambda} \omega_\lambda \psi_\lambda(x, y) \tag{8}$$

where the multi-index $\lambda = (j, i_x, i_y, d)$ denotes the scale j the position $\boldsymbol{i} = (i_x, i_y)$ and the three directions $d = 1, 2, 3$, corresponding to horizontal, vertical, and diagonal wavelets, respectively. The corresponding index set Λ is $\Lambda = \{\lambda = (j, i_x, i_y, d), j = 0, \ldots, J-1;\ i_x, i_y = 0 \ldots 2^j - 1;\ d = 1, 2, 3\}$. Due to orthogonality, the wavelet coefficients are given by $\omega_\lambda = \langle\omega, \psi_\lambda\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the L^2-inner product defined as $\langle f, g\rangle = \int f(x, y)g(x, y)\,\mathrm{d}x\,\mathrm{d}y$. The wavelet coefficients measure fluctuations of ω at scale 2^{-j} around the position $\boldsymbol{i}$, in one of the three directions d. Here a Coifman 30 wavelet is used, which is orthogonal and has 10 vanishing moments [22], $\int x^n\psi(x)\,\mathrm{d}x = 0$ for $n = 0, \ldots, 9$. As the two-dimensional wavelets are constructed using tensor products of one-dimensional wavelets and one-dimensional scaling functions, the resulting two-dimensional wavelets only have vanishing moments in the direction where the wavelet is present.

The CVE algorithm can be summarized in the following three-step procedure:

- *Decomposition*: compute the wavelet coefficients ω_λ using the fast wavelet transform [22];
- *Thresholding*: apply the thresholding function ρ_ϵ to the wavelet coefficients ω_λ, thus discarding the coefficients with absolute values smaller than the threshold ϵ;
- *Reconstruction*: reconstruct the coherent vorticity field ω_C from the thresholded wavelet coefficients using the fast inverse wavelet transform.

The incoherent vorticity field is obtained by simple subtraction, i.e., $\omega_I = \omega - \omega_C$. The thresholding function is given by

$$\rho_\epsilon(a) = \begin{cases} a & \text{if } |a| > \epsilon \\ 0 & \text{if } |a| \leqslant \epsilon \end{cases}$$

where ϵ denotes the threshold $\epsilon = \sqrt{4Z\ln N}$, with the enstrophy $Z = \frac{1}{2}\langle\omega, \omega\rangle$, which corresponds to half of the variance of the vorticity fluctuations and N is the resolution. This threshold value allows for optimal denoising in a min–max sense, assuming the noise to be additive, Gaussian, and white [7].

In summary, this decomposition yields $\omega = \omega_C + \omega_I$. Due to orthogonality, we have $\langle\omega_C, \omega_I\rangle = 0$ and hence it follows that enstrophy is conserved, i.e., $Z = Z_C + Z_I$.

The computational cost of the fast wavelet transform (FWT) scales linearly with the number of data points, i.e., like $O(N)$. The memory requirements of the extracted coherent vortices are equal to the number of retained coefficients, plus the number of filter coefficients (here 30) defining the wavelet.

4. Results and comparison

Direct numerical simulations of the Hasegawa–Wakatani equations are performed with a resolution of 512^2 grid points. For the details about the numerical method used we refer to [4]. The diffusion and kinematic viscosity coefficients are set to 0.01. Our study is concentrated on two cases: the quasi-adiabatic case, where the adiabaticity parameter is $c = 0.7$, and the quasi-hydrodynamic case, where it is much weaker with the value $c = 0.01$. The computations are performed up to time $t = 612$. At $t \approx 100$ the kinetic energy saturates and a statistically steady state is reached, independent of the random initial conditions. Fig. 1 shows the vorticity field $\omega = \nabla^2\phi$ in a turbulent saturated state for $c = 0.01$ and for $c = 0.7$. The quasi-hydrodynamic case (Fig. 1, left) has fewer and larger vortices compared to those of the quasi-adiabatic case (Fig. 1, right). We then apply the POD and CVE techniques upon these two turbulent flows and compare the results thus obtained. The total number of modes thus obtained is $N_{POD} = 512$ for POD and $N_{CVE} = 512^2$ for CVE because each POD mode corresponds to 512 wavelet modes in term of complexity.

4.1. Energy compression

Fig. 2 shows the energy $E = \langle\phi, \nabla^2\phi\rangle$ and enstrophy $Z = \langle\omega, \omega\rangle$ of the singular values of the POD modes in function number of retained modes for both the quasi-hydrodynamic and the quasi-adiabatic cases.

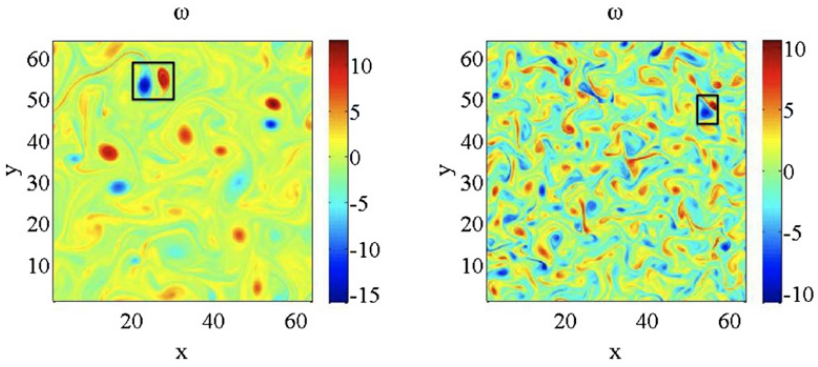

Fig. 1. Vorticity fields $\omega = \nabla^2\phi$ of quasi-hydrodynamic (left) and quasi-adiabatic case (right). The black squares show the dipole structures selected for the scatter plot in Fig. 6.

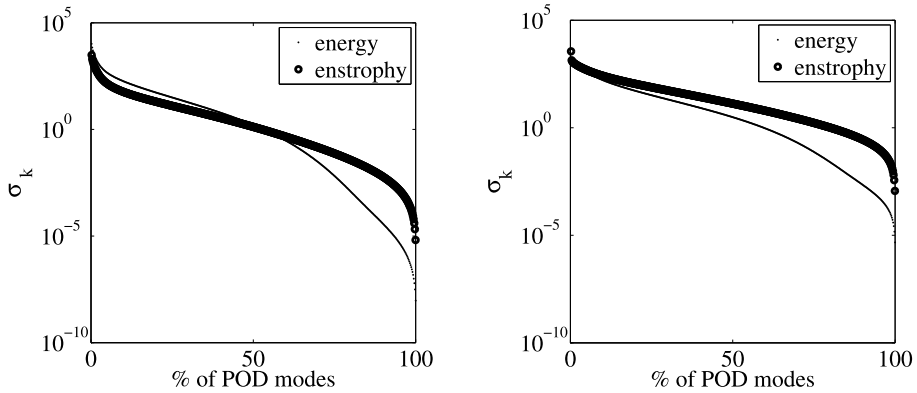

Fig. 2. Decay of singular values σ_k in quasi-hydrodynamic (left) and quasi-adiabatic cases (right).

Table 1
Compression rate related energy E, enstrophy Z and radial flux Γ_r, after applying the POD method and the CVE method to the vorticity field of the quasi-hydrodynamic ($c = 0.01$) and quasi-adiabatic ($c = 0.7$) two-dimensional drift-wave turbulence simulations. Note that the total number of POD modes is 512 and the total number of wavelet modes is 512^2. The CVE results are reproduced from [14].

	Compression (%)	Number of modes	E	Z	Γ_r
Quasi-hydrodynamic (POD)	14.4	74	99.9	99.9	99.8
Quasi-adiabatic (POD)	14.2	73	99.0	97.0	99.8
Quasi-hydrodynamic (CVE)	1.3	3408	99.9	97	99
Quasi-adiabatic (CVE)	1.8	4719	99.0	93	98

The results of the extraction are displayed in Table 1. An exponential decay is observed in the weight distribution of the singular values σ_k of energy E and of enstrophy Z. This indicates that in these quantities most of the energy is contained in the first POD modes. In particular, the first 14% of total modes contribute to about 99% of the total energy and enstrophy. Note that if the singular values decay fast, as is typically the case, most of the energy is contained in the first few modes, and the data can thus be well reproduced by using only few modes in Eq. (3). This is the principle behind the use of POD for data compression and is also directly related to identify coherent structures, which are supposed to carry most of the energy and enstrophy of the flow.

Fig. 3 shows a comparison between the compression curves obtained with the POD and the CVE methods for both the quasi-adiabatic and the quasi-hydrodynamic cases. The analyzed datasets correspond to those of Ref. [14]. The curves show the percentage of retained enstrophy versus the percentage of retained number of modes. For POD half of the retained enstrophy is obtained by summing up the square of the retained POD singular values σ_k, i.e., $\sum_k \sigma_k^2 / Z_\sigma$ renormalized by the total enstrophy Z_σ. For CVE we proceed as follows. The vorticity field is wavelet transformed and we compute the percentage of the wavelet coefficient variance, i.e., half their enstrophy, as a function of the number of retained wavelet coefficients. Fig. 3 shows that the curves increase more rapidly for CVE than for POD: only 1% of the wavelet coefficients is needed to reach 99% of the enstrophy, as was already noted in [14], while 14% of the POD modes are needed to retain the same level of enstrophy. This indicates that the wavelet method reaches a higher compression rate than the POD as also shown in Table 1. Both analyses, with CVE and POD, show that the compression rate of the quasi-adiabatic case is higher than the one of the quasi-hydrodynamic case. This could be related to the different amount of vortical structures in each type of flow. In Fig. 3 (right) we replot those curves in lin–log coordinates to focus on the behavior for high compression rates, which confirms that the CVE is more efficient than POD. We also observe that the compression rate curve of CVE starts close to zero while the curve for POD starts from 40 percent of retained enstrophy. This is due to the fact that the

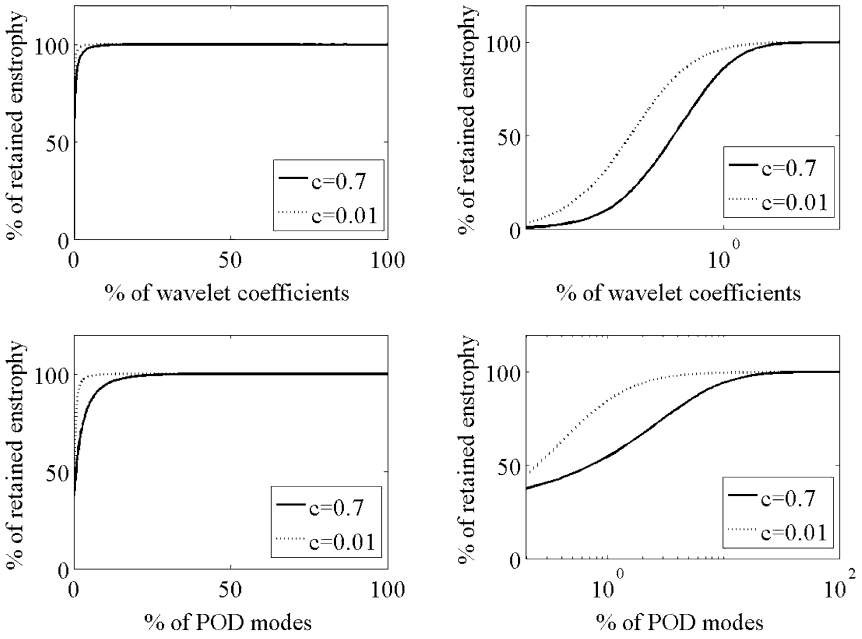

Fig. 3. Comparison of the retained enstrophy obtained by the CVE (top) and by the POD (bottom) methods. The curves show the percentage of retained enstrophy Z versus the percentage of the retained wavelets or POD modes for the quasi-adiabatic case ($c = 0.7$) and for the quasi-hydrodynamic case ($c = 0.01$). Left: lin–lin representation. Right: lin–log representation to enhance the high compressions where only few modes are retained. Note that the abscissas of the curves on the right are different. They range up to 2% compression for CVE and up to 100% for POD, to focus on the interesting part of the compression curves.

first POD modes already contain a significant amount of enstrophy, which is not the case for the wavelet modes. Note again that the total number of modes is very different for both methods: $N_{POD} = 512$ while $N_{CVE} = 512^2$.

In the following we split the vorticity field into coherent and incoherent contributions. For the wavelet based CVE method the criterion to decide what is coherent and what is incoherent is made automatically since it is based on the threshold whose value depends only on the total enstrophy and the total number of modes, both quantities being known a priori as discussed in Section 3.2. For the POD method no such automatic procedure is available to separate what is coherent from what is not. Therefore we decided to first apply the CVE method and find out the amount of retained energy in the coherent part, which gave 99.9%E for the quasi-hydrodynamic case and 99.0%E for the quasi-adiabatic case. For the POD method we then fixed those amounts of coherent energy and checked how many POD modes are necessary to obtain them. The definition of what we call coherent using POD is thus based on the amount of retained energy as decided by the CVE method.

Table 1 presents, for both quasi-adiabatic and quasi-hydrodynamic cases, the number of retained coherent modes and the corresponding compression rate, the percentage of coherent energy E, enstrophy Z by both methods, POD and CVE. We also show the percentage of the retained radial flux Γ_r, defined as $\Gamma_r = \langle n\partial\phi/\partial_y \rangle$, which is computed by decomposing $n\partial\phi/\partial_y$ at each time instant into a POD or wavelet basis, retaining only the given number of coherent modes, and then averaging over all time instants. Those quantitative results confirm that to reach 99.9% of the total energy, 14.4% of the total number of modes $N_{POD} = 512$ are needed for the POD technique, while for CVE only 1.3% of the total number of modes $N_{CVE} = 512^2$ are sufficient. We also show that POD needs much less modes than CVE. We found that 74 out of $N_{POD} = 512$ modes are retained for POD against 4719 out of $N_{CVE} = 512^2$ for CVE modes in the quasi-hydrodynamic case, and 73 against 4719 respectively in the quasi-adiabatic case. From a point of view of number of modes needed to reproduce the flow POD is therefore more efficient. However, in the present case, each POD mode corresponds to 512 wavelet modes in term of complexity so that in terms of numerical efficiency wavelets are here most efficient.

4.2. Wavenumber spectra and probability density functions

The probability distribution functions (PDFs) of vorticity yield information on the higher order statistics while the enstrophy spectra yield insight into the second order statistics. Both quantities have been averaged over 512 realizations during the time interval $100 < t \leqslant 612$. They are shown for the total, coherent, and incoherent vorticities computer either by POD (see Fig. 4) or by CVE (see Fig. 5). We find that the PDFs of the total and coherent vorticities almost coincide, exhibiting a stretched exponential shape in the quasi-hydrodynamic case and an almost Gaussian shape in the quasi-adiabatic case. The incoherent parts show for both POD and CVE a strongly reduced variance. In Fig. 4 we observe that the enstrophy spectra of the total and coherent enstrophies perfectly superpose at large scales and all along the inertial range, whereas they begin to depart in the dissipative range, for $k \geqslant 50$, the departure being slightly more pronounced for the POD than for the CVE. In contrast, the spectra of the incoherent enstrophy differ: POD concentrates it to the dissipative scales while it is spread all over inertial scales for CVE. Indeed, one observes that the incoherent enstrophy of POD (Fig. 4(c)–(d)) scales as $k^{11.5}$

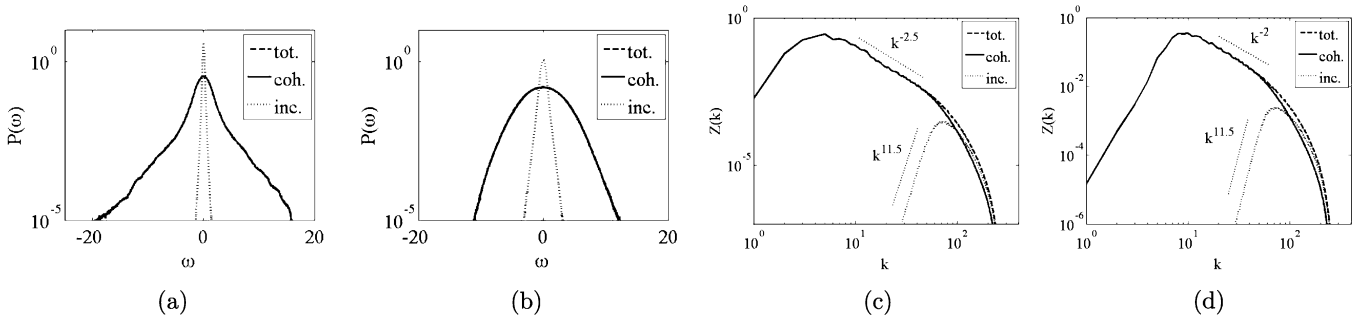

Fig. 4. POD results. (a, b): PDF of the vorticity. (c, d): wavenumber spectrum of enstrophy. (a, c): quasi-hydrodynamic case. (b, d): quasi-adiabatic case. Dashed line: total field, solid line: coherent part, dotted line: incoherent part. Note that the coherent contribution (solid) superposes the total field (dashed), which is thus hidden under the solid line in all four figures. The straight lines indicating power laws are plotted for reference.

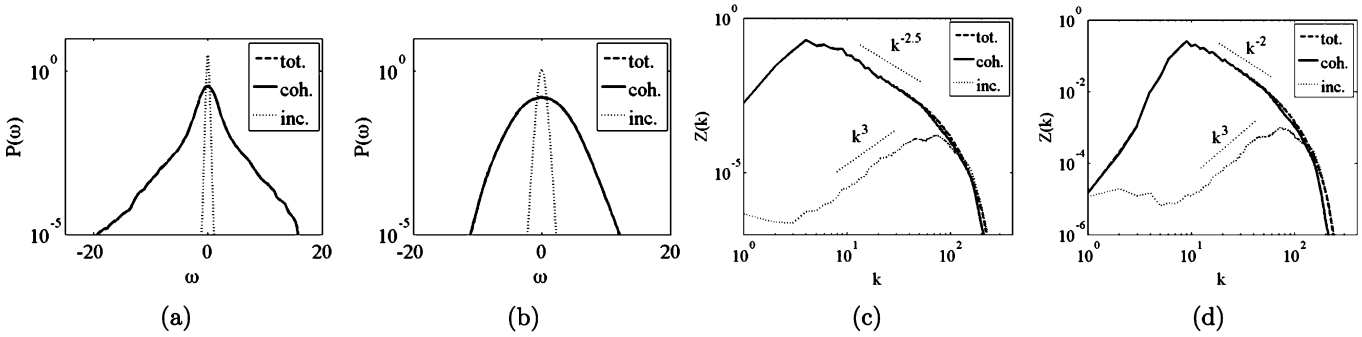

Fig. 5. CVE results. (a, b): PDF of the vorticity. (c, d): wavenumber spectrum of the enstrophy. (a, c): quasi-hydrodynamic case. (b, d): quasi-adiabatic case. Dashed line: total field, solid line: coherent part, dotted line: incoherent part. Note that the coherent contribution (solid) superposes the total field (dashed), which is thus hidden under the solid line in all four figures. The straight lines indicating power laws are plotted for reference.

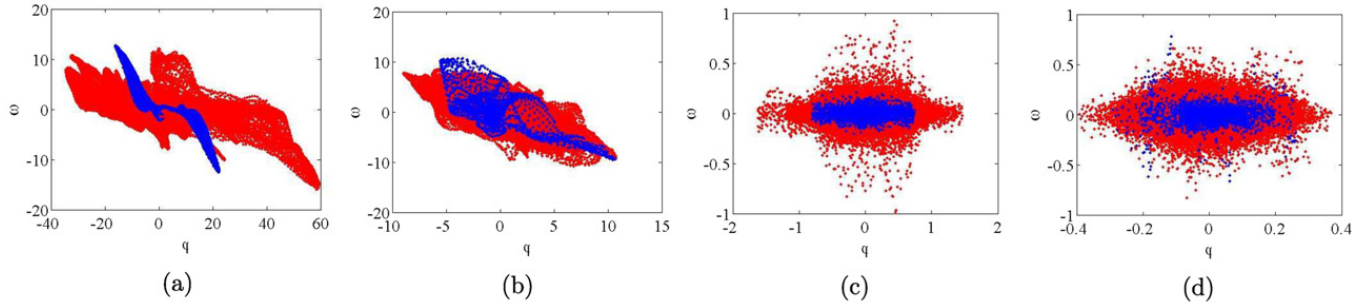

Fig. 6. POD results. Scatter plot of the vorticity against the streamfunction for the coherent and incoherent parts of the field: quasi-hydrodynamic case ((a) coherent and (c) incoherent); quasi-adiabatic case ((b) coherent and (d) incoherent). The red dots correspond to the total field, the blue dots correspond to the dipoles marked in Fig. 1.

while for CVE (Fig. 5(c)–(d)) it scales as k^3, and they then decay when reaching the dissipative scale $k = 50$. The k^3 scaling of the incoherent enstrophy found for CVE gives a k^1 scaling of the incoherent energy, which corresponds to an energy equipartition. Note that the slope $+1$ comes from the integrand when we integrate over all angles in k-space to compute the one-dimensional spectrum which is plotted here. This gives evidence for some kind of thermalization of the incoherent energy in the case of CVE.

4.3. Scatter plots

Joint PDFs between the electrostatic potential (which can be interpreted as a streamfunction) and the vorticity allow to detect functional relationships between both quantities. Such functional relationship implies that the Poisson bracket between the two quantities vanishes and thus leads to a depletion of nonlinearity in the governing equation for the vorticity. We show in Fig. 6 scatter plots of the streamfunction as a function of the vorticity for the fields in Fig. 1. Both the coherent part and the incoherent are shown for POD. The total part is visually almost identical to the coherent part and is thus omitted. Also shown, superposed in the same figures and visualized in blue is the scatter plot corresponding to one of the dipoles observed in Fig. 1. In the freely decaying hydrodynamic case, Joyce and Montgomery [23] showed that the functional relation $\phi(\omega) = \alpha \sinh(\beta\omega)$, α and β being constants, can be expected, corresponding to a final state of decay

depleted from nonlinearity. The dipoles in the hydrodynamic case are coherent structures depleted of nonlinearity and we thus observe such a sinh functional in the scatter plot. In the quasi-adiabatic case the scatter plots do not yield such a clear functional relation, but some relation between ϕ and ω may be discerned. In the incoherent parts no functional relation can be distinguished which gives evidence that the incoherent part does not contain coherent structures. Similar observation was found in [14] for CVE.

5. Conclusion

The wavelet based CVE method was compared with the POD method to extract coherent structures out of resistive drift-wave turbulent flows computed by direct numerical simulation. There is for CVE an objective criterion, based on denoising by thresholding the wavelet coefficients, which allows to split each flow realization into coherent structures and an incoherent noise-like background flow. In order to compare both methods, we have chosen for POD an ad hoc energy based criterion which corresponds to the coherent energy level found by the CVE. We assessed the compression properties of both methods for the quasi-adiabatic, i.e., the plasma relevant regime, and the quasi-hydrodynamic regime and showed that CVE required less modes in terms of percentage of the total number of modes, while POD requires less modes in terms of their total number. Nevertheless, we found that both methods allow to extract coherent vortices out of turbulent flows. The statistical properties of the total, coherent, and incoherent flow contributions obtained by POD and by CVE were investigated, in terms of probability distribution function of vorticity and in terms of enstrophy spectrum. We found that the coherent and total flows almost exhibit the same statistics, which is not surprising as most of the energy and enstrophy is retained. The incoherent flow obtained by CVE exhibits an energy equipartition which corresponds to the spatial decorrelation of the incoherent noise-like contribution. From a computational point of view the complexity of POD is prohibitive for large datasets, as the computation of the POD basis scales with the cube of the number of data points N. In contrast CVE is optimal since its operation count only increases linearly with N. The memory requirements of POD are determined by the number of retained modes, which is typically very low, but each POD mode is represented by N grid points which have to be stored. For CVE the retained wavelet coefficients have to be stored, which are typically in total much larger than for POD, however no wavelet mode has to be stored as wavelets are defined by a quadratic mirror filter [5], characterizing the wavelet basis one has a priori chosen, and its length is relatively short, i.e., 30 values to be stored for the wavelet chosen here. Future work could deal with modeling aspects of the Hasegawa–Wakatani equations using either POD or wavelet projections. This will allow to perform numerical simulations using a reduced number of degrees of freedom while preserving the nonlinear flow dynamics.

Acknowledgements

The authors acknowledge financial support from the Agence Nationale de la Recherche under contract ‘M2TFP’. D.dC.N. is thankful for the support from the Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725. D.dC.N. also gratefully acknowledges the hospitality and financial support of the École Centrale de Marseille during the elaboration of this work. K.S. and M.F. are thankful to the Association CEA-EURATOM and the French Research Federation for Fusion Studies for supporting their work within the framework of the European Fusion Development Agreement under contract V.3258.001. The views and opinions expressed herein do not necessarily reflect those of the European Commission. M.F. thanks the Wissenschaftskolleg zu Berlin for its kind hospitality while writing this paper.

References

[1] P. Holmes, J.L. Lumley, G. Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, 1996.
[2] J.L. Lumley, in: R.E. Meyer (Ed.), Transition and Turbulence, Academic Press, New York, 1981, pp. 215–242.
[3] N. Aubry, R. Lima, M. Rahibe, Chaos 13 (2003) 541.
[4] S. Futatani, S. Benkadda, D. del-Castillo-Negrete, Phys. Plasmas 16 (2009) 042506.
[5] M. Farge, Annu. Rev. Fluid Mech. 24 (1992) 395.
[6] M. Farge, K. Schneider, N. Kevlahan, Phys. Fluids 11 (1999) 2187.
[7] M. Farge, G. Pellegrino, K. Schneider, Phys. Rev. Lett. 87 (2001) 054501.
[8] N. Okamoto, K. Yoshimatsu, K. Schneider, M. Farge, Y. Kaneda, Phys. Fluids 19 (11) (2007) 11519, 1–13.
[9] M. Farge, E. Goirand, Y. Meyer, F. Pascal, M.V. Wickerhauser, Fluid Dyn. Res. 10 (1992) 229–250.
[10] D. del-Castillo-Negrete, S.P. Hirshman, D.A. Sponga, E.F. D'Azevedo, J. Comput. Phys. 222 (2007) 265.
[11] D. del-Castillo-Negrete, D.A. Sponga, S.P. Hirshman, Phys. Plasmas 15 (2008) 092308.
[12] M. Farge, K. Schneider, G. Pellegrino, A.A. Wray, R.S. Rogallo, Phys. Fluids 15 (10) (2003) 2886–2896.
[13] M. Farge, K. Schneider, P. Devynck, Phys. Plasmas 13 (2006) 042304.
[14] W.J.T. Bos, S. Futatani, S. Benkadda, M. Farge, K. Schneider, Phys. Plasmas 15 (2008) 072305.
[15] S. Gassama, E. Sonnendrücker, K. Schneider, M. Farge, M. Domingues, ESAIM Proc. 16 (2006) 195–210.
[16] R. Nguyen van yen, D. del-Castillo-Negrete, K. Schneider, M. Farge, G. Chen, J. Comput. Phys. 229 (2010) 2821.
[17] S. Benkadda, T. Dudok de Wit, A. Verga, A. Sen, ASDEX team, X. Garbet, Phys. Rev. Lett. 73 (1994) 3403.
[18] A.J. Gámez, C.S. Zhou, A. Timmermann, J. Kurthst, Nonlinear Process. Geophys. 11 (2004) 393.
[19] A. Hasegawa, M. Wakatani, Phys. Rev. Lett. 50 (1983) 682.
[20] A. Hasegawa, K. Mima, Phys. Rev. Lett. 39 (1977) 205.
[21] G.H. Golub, D.F. Van Loan, Matrix Computations, third ed., The John Hopkins University Press, London, 1996.
[22] S. Mallat, A Wavelet Tour of Signal Processing, Academic, New York, 1998.
[23] G. Joyce, D. Montgomery, J. Plasma Phys. 19 (1973) 107.