J. Fluid Mech. (2006), *vol.* 559, *pp.* 215–254.
doi:10.1017/S0022112006000346 Printed in the United Kingdom

On proper orthogonal decomposition of randomly perturbed fields with applications to flow past a cylinder and natural convection over a horizontal plate

By DANIELE VENTURI

Department of Energy, Nuclear and Environmental Engineering, University of Bologna, Bologna, 40136 Italy
daniele.venturi@mail.ing.unibo.it

(Received 10 March 2005 and in revised form 19 December 2005)

The connections between the random elements of a discrete random flow field and the uncertainty in the hierarchical set of its spatio-temporal scales, obtained by the symmetric version of the proper orthogonal decomposition (POD) method, are investigated. It is shown that the relevant statistics for the energy levels, the temporal modes and the spatial modes can be expressed in an explicit form as power series of the flow field standard deviation. Such expansions characterize accurately interesting phenomena of mixing between different flow scales. The basis of the present work is the assumption that the randomness is characterized by a Gaussian uncorrelated random field. Two applications of the theory developed are proposed: to the incompressible flow past a cylinder at Reynolds number $Re = 100$ and to the natural convective flow over an isothermal horizontal plate at Rayleigh number $Ra = 4.75 \times 10^6$. The theoretical predictions are confirmed well by Monte Carlo simulations and interesting relations between the random flows and the relevant statistics of their POD spatio-temporal scales are determined and discussed.

1. Introduction

The proper orthogonal decomposition (POD) method was proposed by Lumley (1970) for detection of spatial coherent patterns in turbulent flows. He introduced it in the field of hydrodynamics when there was a need for a mathematical definition of coherent structures in turbulence. To analyse such temporally and spatially evolving flows Aubry, Guyonnet & Lima (1991) introduced the concept of biorthogonal decomposition which is a deterministic space–time version of the POD (Aubry 1991; Aubry, Guyonnet & Lima 1995; Aubry & Lima 1995). One of the most remarkable feature of such a decomposition is that it gives access to the complexity of the spatial and temporal dynamics simultaneously. The flow field is decomposed into a hierarchical set of spatial and temporal orthogonal modes which are coupled. This generalizes the notion of spatial and temporal structures which, for example, can be followed through the various instabilities that the flow undergoes as Reynolds number increases. Low-dimensional linear Galerkin (Noack *et al.* 2003; Noack, Papas & Monkewitz 2005; Ma *et al.* 2003; Deane *et al.* 1991; Cazemier, Verstappen & Veldman 1998), nonlinear Galerkin (Ma, Karamanos & Karniadakis 2000; Ma *et al.* 2000) and spectral viscosity (Sirisup & Karniadakis 2004) models of various flows have been

successfully developed by using the POD modes. Moreover, the method is a way to analyse and reconstruct (Venturi & Karniadakis 2004; Tan, Willcox & Damodaran 2003; Tan, Damodaran & Willcox 2004) space time information such as numerical data and experimental data measured simultaneously at various locations by means of recently developed experimental techniques such as digital particle image velocimetry (DPIV), digital particle image thermometry (DPIT) (Ma *et al.* 2002, 2003), laser scanning techniques, cross-stream rakes of X-wires (Delville *et al.* 2003; Gordeyev & Thomas 2000) or magnetic resonance imaging (MRI).

However, if the observed flow is noisy, perturbed or, more generally, it is considered as a superposition of *random* and *non-random* elements then its space–time structures should be considered random as well. The quantification of the relation between the stochastic flow and its random space–time structures extracted by the POD method is a challenging task as it involves random perturbations of the autocorrelation operator's spectral properties. The effects of infinitesimal deterministic perturbations in the ensemble of data have been investigated in the work by Rathinam & Petzold (2003) by introducing a POD 'sensitivity factor'. Other attempts to quantify directly the statistics for noisy correlation (or covariance) matrices have recently been made by several Authors. Everson & Roberts (2000) use a Bayesian inference method to obtain posterior densities for each random eigenvalue; Sengupta & Mitra (1999) propose a diagrammatic expansion and saddle point integration methods to quantify the empirical eigenvalue density; Hachem, Loubaton & Najim (2005, 2004) and Dozier & Silverstein (2004) give a characterization of the eigenvalue density in terms of its Stieltjes transform; Hoyle & Rattray (2004) use a statistical mechanics approach (variational mean-field theory) to give an analytical approximation to the eigenvalue spectral density.

In this paper we propose a new method based upon the perturbation theory for linear operators (Kato 1995). We show that the relevant statistics for the perturbed energy levels and the perturbed temporal modes can be expressed in the form of an explicit power series in the random-flow standard deviation. Therefore we provide a new point of view which, in turn, will give us a connection between the random elements of the flow and the uncertainty in the hierarchical set of its spatio-temporal scales. At the basis of the present work is the assumption that the randomness is characterized by a Gaussian uncorrelated random field although it is possible to apply the same approach to more general random fields. The key idea lies in a theoretical estimation of the overlap between perturbed energy levels through a resolution indicator function (§ 3.2). We emphasize that our method can also be used to represent Karhunen–Loève expansions of randomly perturbed covariance kernels. This is very important, for instance, when an experimental-based (noisy) stochastic quantity is used as random input in effective uncertainty propagation methods such as generalized polynomial chaos (Xiu & Karniadakis 2002, 2003; Xiu 2004) or multi-element generalized polynomial chaos (Wan & Karniadakis 2005*a*, *b*).

We propose two applications of the theory developed: to the incompressible flow past a cylinder at $Re = 100$ and to the natural convective flow above an isothermal, highly elongated, horizontal plate at Rayleigh number $Ra = 4.75 \times 10^6$. We have selected these flows as test problems due to the numerous experimental and theoretical investigations carried out during recent years as well as because of their differences. The flow past a cylinder at $Re = 100$ exhibits a periodic laminar wake (Zdravkovich 1997) which, in turn, gives a POD decomposition of the velocity field madeup of coupled degenerate periodic eigenmodes with a pronounced energy decay (Ma & Karniadakis 2002; Noack *et al.* 2003; Ma *et al.* 2003). On the other hand the

natural convective flow above an isothermal horizontal plate at $Ra = 4.75 \times 10^6$ is chaotic. Because of the high Rayleigh number, which activates many intermodal energy transfers between the various scales (Couplet, Sagaut & Basdevant 2003), the temperature field decomposition is characterized by simple eigenvalues with relatively small separation. Despite the differences between these two flows the perturbation of their deterministic states shows interesting analogies.

This paper is organized as follows. In § 2 we formulate the POD perturbation problem. In § 3 we describe the perturbation theory for the POD eigen-states and their Fourier analysis. In § 4 we analyse the link between the random dynamics in time and space and we stress a symmetry breaking between the resolvable part of the temporal and spatial dynamic. In § 5 we propose two applications of the theory developed: to the incompressible flow past a circular cylinder at $Re = 100$ and the natural convective flow above isothermal horizontal plates in confined domains at Rayleigh number $Ra = 4.75 \times 10^6$. The main findings and their implications are outlined and discussed in § 6.

2. POD theory and the perturbation of the deterministic state

Given a d-dimensional real vector flow field $\boldsymbol{u}(\boldsymbol{x}, t)$ in the space–time domain $\Omega \times T$ (which means that the ordered pair $(\boldsymbol{x}, t)$ is such that $\boldsymbol{x} \in \Omega$ (spatial domain) and $t \in T$ (temporal domain)) we can look for a biorthogonal representation (Aubry *et al.* (1991)) of $\boldsymbol{u}(\boldsymbol{x}, t)$ in the form

$$\boldsymbol{u}(\boldsymbol{x}, t) = \sum_{k=1}^{\infty} \sqrt{\mu_k} \boldsymbol{\Phi}_k(\boldsymbol{x}) \psi_k(t). \tag{2.1}$$

Let us consider the inner product

$$(\boldsymbol{u}, \boldsymbol{v}) := \int_{\Omega} \int_{T} \boldsymbol{u}(\boldsymbol{x}, t) \cdot \boldsymbol{v}(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}t \tag{2.2}$$

and the functional (energy norm)

$$\mathscr{L}[\boldsymbol{\Phi}_i, \psi_j] := \left(\boldsymbol{u}(\boldsymbol{x}, t) - \sum_{k=1}^{\infty} \sqrt{\mu_k} \boldsymbol{\Phi}_k(\boldsymbol{x}) \psi_k(t), \boldsymbol{u}(\boldsymbol{x}, t) - \sum_{k=1}^{\infty} \sqrt{\mu_k} \boldsymbol{\Phi}_k(\boldsymbol{x}) \psi_k(t) \right). \tag{2.3}$$

It is well known that the minimization of (2.3) with respect to an arbitrary variation of $\boldsymbol{\Phi}_k(\boldsymbol{x})$ and $\psi_k(t)$ leads to the Euler–Lagrange equations

$$\sqrt{\mu_k} \psi_k(t) = \int_{\Omega} \boldsymbol{u}(\boldsymbol{x}, t) \cdot \boldsymbol{\Phi}_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \tag{2.4a}$$

$$\sqrt{\mu_k} \boldsymbol{\Phi}_k(\boldsymbol{x}) = \int_{T} \boldsymbol{u}(\boldsymbol{x}, t) \psi_k(t) \, \mathrm{d}t. \tag{2.4b}$$

These equations are called dispersion relations and they provide the link between the spatial and the temporal evolution of the system. Following Aubry *et al.* (1991) it is useful to define the linear integral operators $\boldsymbol{U}$ and $\boldsymbol{U}^{\dagger}$:

$$(\boldsymbol{U}\boldsymbol{\Phi}_k)(t) := \int_{\Omega} \boldsymbol{u}(\boldsymbol{x}, t) \cdot \boldsymbol{\Phi}_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \sqrt{\mu_k} \psi_k(t), \tag{2.5a}$$

$$(\boldsymbol{U}^{\dagger} \psi_k)(\boldsymbol{x}) := \int_{T} \boldsymbol{u}(\boldsymbol{x}, t) \psi_k(t) \, \mathrm{d}t = \sqrt{\mu_k} \boldsymbol{\Phi}_k(\boldsymbol{x}). \tag{2.5b}$$

If we apply the operator $\boldsymbol{U}^\dagger$ to (2.5a) and the operator $\boldsymbol{U}$ to (2.5b) we obtain the eigenvalue problems:

$$(\boldsymbol{U}^\dagger \boldsymbol{U}\boldsymbol{\Phi}_k)(\boldsymbol{x}) = \int_\Omega \mathscr{S}(\boldsymbol{x}, \boldsymbol{x}')\boldsymbol{\Phi}_k(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}' = \mu_k \boldsymbol{\Phi}_k(\boldsymbol{x}), \tag{2.6a}$$

$$(\boldsymbol{U}\boldsymbol{U}^\dagger \psi_k)(t) = \int_T \mathscr{T}(t, t')\psi_k(t')\,\mathrm{d}t' = \mu_k \psi_k(t). \tag{2.6b}$$

In (2.6a) and (2.6b) μ_k are the *repeated* eigenvalues of $\boldsymbol{U}^\dagger\boldsymbol{U}$ and $\boldsymbol{U}\boldsymbol{U}^\dagger$, while we will denote by λ_k the eigenvalues counted with their multiplicity. The kernels of the integral operators $\boldsymbol{U}^\dagger\boldsymbol{U}$ and $\boldsymbol{U}\boldsymbol{U}^\dagger$ are symmetric and positive, representing the spatial and the temporal autocorrelation respectively:

$$\mathscr{S}(\boldsymbol{x}, \boldsymbol{x}') := \int_T \boldsymbol{u}(\boldsymbol{x}, t)\cdot\boldsymbol{u}(\boldsymbol{x}', t)\,\mathrm{d}t, \tag{2.7a}$$

$$\mathscr{T}(t, t') := \int_\Omega \boldsymbol{u}(\boldsymbol{x}, t)\cdot\boldsymbol{u}(\boldsymbol{x}, t')\,\mathrm{d}\boldsymbol{x}. \tag{2.7b}$$

The symmetric version (Aubry 1991) of the proper orthogonal decomposition (2.1) corresponds to the spectral analysis of the operator $\boldsymbol{U}$ and consists of square root of proper values (eigenvalues) and proper functions (eigenfunctions) of the correlation operators (2.6a) and (2.6b). Following the so-called 'snapshots method' (Sirovich 1987; Aubry (1991)) the decomposition (2.1) is computed in practice by solving the eigenvalue problem (2.6b) using a time scale roughly of the order of or greater than the correlation time and then the spatial mode relation (2.4b). We order the eigenvalues as

$$\mu_j \geqslant \mu_{j+1}, \quad \lambda_k > \lambda_{k+1}, \quad j, k \in \mathbb{N}. \tag{2.8}$$

Now we consider a random additive perturbation of the deterministic flow field $\boldsymbol{u}(\boldsymbol{x}, t)$ in the form

$$\widetilde{\boldsymbol{u}}(\boldsymbol{x}, t) = \boldsymbol{u}(\boldsymbol{x}, t) + \delta\boldsymbol{u}(\boldsymbol{x}, t). \tag{2.9}$$

Note that (2.9) can also be seen as a decomposition of a turbulent quantity $\widetilde{\boldsymbol{u}}$ into a mean flow $\boldsymbol{u}(\boldsymbol{x}, t) = \langle\widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)\rangle$ and a zero-mean random fluctuation $\delta\boldsymbol{u}(\boldsymbol{x}, t)$, where $\langle\cdot\rangle$ is a suitable averaging operation. The substitution of (2.9) into (2.7b) gives

$$\widetilde{\mathscr{T}}(t, t') = \mathscr{T}(t, t') + \mathscr{H}^{(1)}(t, t') + \mathscr{H}^{(2)}(t, t') \tag{2.10}$$

where

$$\mathscr{H}^{(1)}(t, t') := \int_\Omega (\boldsymbol{u}(\boldsymbol{x}, t)\cdot\delta\boldsymbol{u}(\boldsymbol{x}, t') + \boldsymbol{u}(\boldsymbol{x}, t')\cdot\delta\boldsymbol{u}(\boldsymbol{x}, t))\,\mathrm{d}\boldsymbol{x}, \tag{2.11a}$$

$$\mathscr{H}^{(2)}(t, t') := \int_\Omega \delta\boldsymbol{u}(\boldsymbol{x}, t)\cdot\delta\boldsymbol{u}(\boldsymbol{x}, t')\,\mathrm{d}\boldsymbol{x}. \tag{2.11b}$$

The perturbed temporal autocorrelation $\widetilde{\mathscr{T}}(t, t')$ appears to be a superimposition of three symmetric kernels: $\mathscr{T}(t, t')$ is the temporal autocorrelation of the unperturbed field (2.7b), $\mathscr{H}^{(1)}(t, t')$ is the temporal correlation between the unperturbed field and the perturbation and $\mathscr{H}^{(2)}(t, t')$ is the temporal autocorrelation of the perturbation. Equivalently we can say that the autocorrelation of $\widetilde{\boldsymbol{u}}$ consists of two 'free' autocorrelations $\mathscr{T}$ and $\mathscr{H}^{(2)}$ plus an 'interaction' term $\mathscr{H}^{(1)}$. Equation (2.10) implies the following expansion for the perturbed temporal correlation operator $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger$:

$$\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger = \boldsymbol{U}\boldsymbol{U}^\dagger + \boldsymbol{H}^{(1)} + \boldsymbol{H}^{(2)} \tag{2.12}$$

where $\boldsymbol{H}^{(1)}$ and $\boldsymbol{H}^{(2)}$ are integral operators with kernel (2.11*a*) and (2.11*b*) respectively. We define

$$\boldsymbol{H} := \boldsymbol{H}^{(1)} + \boldsymbol{H}^{(2)} \tag{2.13}$$

and write the perturbation problem (2.12) as

$$\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger} = \boldsymbol{U}\boldsymbol{U}^{\dagger} + \boldsymbol{H}. \tag{2.14}$$

Obviously, the operator $\boldsymbol{H}$ is an integral operator with kernel

$$\mathscr{H}(t, t') := \mathscr{H}^{(1)}(t, t') + \mathscr{H}^{(2)}(t, t'). \tag{2.15}$$

In this paper we will determine the relevant statistics for the eigenvalues and the eigenfunctions of $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}$, given the deterministic $\boldsymbol{U}\boldsymbol{U}^{\dagger}$, under the assumption that $\delta\boldsymbol{u}(\boldsymbol{x}, t)$ is a Gaussian uncorrelated (in time and space) random field with variance $\sigma^2(\boldsymbol{x}, t)$ which is homogeneous and stationary, i.e. not dependent on $\boldsymbol{x}$ and t. In a finite-dimensional representation (see Appendix A) this means that the joint probability density of $\delta\boldsymbol{u}$ has the form

$$\pi(\delta\boldsymbol{u}(\boldsymbol{x}_1, t_1), \ldots, \delta\boldsymbol{u}(\boldsymbol{x}_N, t_S); \sigma) = K \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{l=1}^{S}\delta\boldsymbol{u}(\boldsymbol{x}_i, t_l)\cdot\delta\boldsymbol{u}(\boldsymbol{x}_i, t_l)\right] \tag{2.16}$$

where N is the number of spatial points, S is the number of time instants and K is a suitable normalization constant. Note that the proposed method can also be applied to more general types of random perturbations $\delta\boldsymbol{u}$ (by a suitable recalculation of the expectation relations given in Appendix B).

3. Statistics for the perturbed energy levels and temporal modes

We consider the perturbed correlation operator $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}$ as a second-order analytic perturbation expansion in the random-flow standard deviation σ (perturbation parameter). To this end we write (2.12) as

$$\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}(\sigma) = \boldsymbol{U}\boldsymbol{U}^{\dagger} + \sigma\left(\frac{\boldsymbol{H}^{(1)}}{\sigma}\right) + \sigma^2\left(\frac{\boldsymbol{H}^{(2)}}{\sigma^2}\right). \tag{3.1}$$

We also define the following normalized perturbation operators

$$\widehat{\boldsymbol{H}}^{(1)} := \frac{\boldsymbol{H}^{(1)}}{\sigma}, \quad \widehat{\boldsymbol{H}}^{(2)} := \frac{\boldsymbol{H}^{(2)}}{\sigma^2}. \tag{3.2a,b}$$

The substitution of (3.2*a*) and (3.2*b*) into (3.1) gives

$$\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}(\sigma) = \boldsymbol{U}\boldsymbol{U}^{\dagger} + \sigma\widehat{\boldsymbol{H}}^{(1)} + \sigma^2\widehat{\boldsymbol{H}}^{(2)}, \tag{3.3}$$

which, by virtue of (2.12), is exact for perturbations of arbitrary magnitude. Note that the kernels of $\widehat{\boldsymbol{H}}^{(1)}$ and $\widehat{\boldsymbol{H}}^{(2)}$ contain a normalized random field $\delta\widehat{\boldsymbol{u}}$, i.e. a random field with unit variance. We shall also specify a relation between the standard deviation of $\delta\boldsymbol{u}$ (or, equivalently, $\widetilde{\boldsymbol{u}}$) and the unperturbed flow $\boldsymbol{u}$. To this end it is better to recast (in non-dimensional form) the perturbation problem (3.3) as

$$\frac{\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}}{u^2}(\sigma) = \frac{\boldsymbol{U}\boldsymbol{U}^{\dagger}}{u^2} + \left(\frac{\sigma}{u}\right)\frac{\widehat{\boldsymbol{H}}^{(1)}}{u} + \left(\frac{\sigma}{u}\right)^2\widehat{\boldsymbol{H}}^{(2)}, \tag{3.4}$$

where u is a given scaling factor for the field $\boldsymbol{u}(\boldsymbol{x}, t)$. In (3.4) both $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}/u^2$ and $\boldsymbol{U}\boldsymbol{U}^{\dagger}/u^2$ are non-dimensional while σ/u is the new dimensionless perturbation parameter,

which is widely used in engineering practice (given, for instance, as a percentage of the instrument's range). In the next subsections we will calculate the relevant statistics for the eigenvalues (energy levels) and the eigenfunctions (temporal modes) of $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}$ building upon the Kato's perturbation theory for linear operators (Kato 1995). The statistics for the aforementioned spectral properties will be given as power series in the powers of the random-flow standard deviation σ or, equivalently, σ/u. The random nature of $\widehat{\boldsymbol{H}}^{(1)}$ and $\widehat{\boldsymbol{H}}^{(2)}$ in (3.4) implies that such expansions converge with a certain probability which is closely related to the statistical separation of the perturbed energy levels discussed in § 3.2.

3.1. *Statistics for the random energy levels*

We consider a third-order perturbation expansion for the weighted mean of the perturbed energy levels $\widetilde{\lambda}_j(\sigma)$ (Kato 1995)†

$$\delta\lambda_j \equiv \widetilde{\lambda}_j(\sigma) - \lambda_j = \sigma\widehat{\lambda}_j^{(1)} + \sigma^2\widehat{\lambda}_j^{(2)} + \sigma^3\widehat{\lambda}_j^{(3)} + O\big(\sigma^4\widehat{\lambda}_j^{(4)}\big) \tag{3.5}$$

where

$$\widehat{\lambda}_j^{(1)} = \frac{1}{m_j}\mathrm{tr}\,\big(\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j\big), \tag{3.6a}$$

$$\widehat{\lambda}_j^{(2)} = \frac{1}{m_j}\mathrm{tr}\,\big(\widehat{\boldsymbol{H}}^{(2)}\boldsymbol{P}_j - \widehat{\boldsymbol{H}}^{(1)}\boldsymbol{S}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j\big), \tag{3.6b}$$

$$\begin{aligned}\widehat{\lambda}_j^{(3)} = \frac{1}{m_j}\mathrm{tr}\,\big(&-\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{S}_j\widehat{\boldsymbol{H}}^{(2)}\boldsymbol{P}_j - \widehat{\boldsymbol{H}}^{(2)}\boldsymbol{S}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j + \widehat{\boldsymbol{H}}^{(1)}\boldsymbol{S}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{S}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j \\ &- \widehat{\boldsymbol{H}}^{(1)}\boldsymbol{S}_j^2\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j\big).\end{aligned} \tag{3.6c}$$

In (3.6*a*)–(3.6*c*) m_j is the multiplicity of the unperturbed eigenvalue λ_j, $\mathrm{tr}\,(\cdot)$ denotes the trace operator, $\boldsymbol{P}_j$ the projection (A 15) onto the spectral subspace corresponding to λ_j and $\boldsymbol{S}_j$ the reduced resolvent (A 16) with respect to λ_j. We use the expectation relations (B 3*a*)–(B 3*g*) obtained in Appendix B to discard the odd powers of σ and write the expected value and the standard deviation of $\delta\lambda_j$ as

$$\langle\delta\lambda_j\rangle = \sigma^2\big\langle\widehat{\lambda}_j^{(2)}\big\rangle + O\big(\sigma^4\big\langle\widehat{\lambda}_j^{(4)}\big\rangle\big), \tag{3.7}$$

$$\sigma_{\delta\lambda_j} = \big[\sigma^2\big\langle\big(\widehat{\lambda}_j^{(1)}\big)^2\big\rangle + \sigma^4\big(\big\langle\big(\widehat{\lambda}_j^{(2)}\big)^2\big\rangle - \big\langle\widehat{\lambda}_j^{(2)}\big\rangle^2 + \big\langle\widehat{\lambda}_j^{(1)}\widehat{\lambda}_j^{(3)}\big\rangle\big) + \cdots\big]^{1/2} \tag{3.8}$$

where the averaging operator $\langle\cdot\rangle$ is defined in (B 2).

3.1.1. *The averaged energy levels*

By applying (B 3*b*) and (B 3*d*) to $\langle\widehat{\lambda}_j^{(2)}\rangle$ it is easy to show that (3.7) leads to

$$\begin{aligned}\langle\widetilde{\lambda}_j\rangle = \lambda_j + \sigma^2\bigg[\Delta t|\Omega|d - \frac{\Delta t^2}{m_j}\big(\mathrm{tr}\,(\boldsymbol{S}_j)\mathrm{tr}\,\big(\mathscr{T}^{(2)}\boldsymbol{P}_j\big) + \mathrm{tr}\,(\boldsymbol{P}_j)\mathrm{tr}\,\big(\mathscr{T}^{(2)}\boldsymbol{S}_j\big)\big)\bigg] \\ + O\big(\sigma^4\big\langle\widehat{\lambda}_j^{(4)}\big\rangle\big),\end{aligned} \tag{3.9}$$

† It is well known that if the unperturbed eigenvalue λ_j is degenerate then, after the perturbation, it will generally split into several eigenvalues $\widetilde{\mu}_{j_k}(\sigma)$ (the λ_j-group). It is very hard to keep track of the behaviour of each of these perturbed eigenvalues. In principle this can be done successfully using the reduction theory (Kato 1995). However here we are dealing with random operators so the first stage of the reduction process is the determination of the spectral properties for the random matrix $\boldsymbol{P}_j\widehat{\boldsymbol{H}}^{(1)}\boldsymbol{P}_j$ which is itself a very demanding problem.

where Δt is the time interval between two snapshots of the flow field, $|\Omega|$ is the measure of the spatial domain, d the dimension of the flow field and $\mathscr{T}^{(2)}$ is defined in (A 10). Equivalently, we define the antisymmetric matrix

$$\Lambda_{kj} := \frac{\mathrm{tr}\,(\boldsymbol{P}_k)\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{P}_j\right) + \mathrm{tr}\,(\boldsymbol{P}_j)\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{P}_k\right)}{\lambda_k - \lambda_j}, \quad k \neq j, \tag{3.10}$$

and write (3.9) as

$$\langle\widetilde{\lambda}_j\rangle = \lambda_j + \sigma^2\left[\Delta t|\Omega|d - \frac{\Delta t^2}{m_j}\sum_{k\neq j}\Lambda_{kj}\right] + O\left(\sigma^4\langle\widehat{\lambda}_j^{(4)}\rangle\right). \tag{3.11}$$

We note that $\langle\delta\lambda_j\rangle \geqslant 0\ \forall j$, which means that, in a third-order approximation, Gaussian uncorrelated random perturbations (2.16) always increase the average energy of the POD modes. Moreover, because of the skew-symmetry of Λ_{kj}, from (3.11) we have that the averaged total energy of the perturbation is

$$\sum_j\langle\delta\lambda_j\rangle = \sigma^2|\Omega||T|d. \tag{3.12}$$

This result is in agreement with the exact relation (3.21) obtained in § 3.3. The constraint (3.21) also implies that all the subsequent higher-order terms in (3.7) must be represented by antisymmetric matrices.

3.1.2. *The standard deviations*

After some computations of the terms in (3.8) the following expression for the standard deviation of $\widetilde{\lambda}_j$ (or, equivalently, $\delta\lambda_j$) is obtained:

$$\begin{aligned}\sigma_{\widetilde{\lambda}_j} = \Bigg[&4\frac{\sigma^2\Delta t^2}{m_j^2}\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{P}_j\right) + 2\frac{\sigma^4}{m_j^2}\big[\Delta t^2 d\Upsilon_2 m_j - 2\Delta t^3\big(2\mathrm{tr}\,(\boldsymbol{S}_j)\mathrm{tr}\left(\mathscr{T}^{(3)}\boldsymbol{P}_j\right)\\
&+ m_j\mathrm{tr}\left(\mathscr{T}^{(3)}\boldsymbol{S}_j\right)\big) + \Delta t^4\big(2\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{S}_j^2\mathscr{T}^{(2)}\boldsymbol{P}_j\right)(2-m_j)\\
&+ 2\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{S}_j^2\right)\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{P}_j\right) + m_j\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{S}_j\mathscr{T}^{(2)}\boldsymbol{S}_j\right)\\
&+ 4\mathrm{tr}\,(\boldsymbol{S}_j)\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{S}_j\mathscr{T}^{(2)}\boldsymbol{P}_j\right) - 3\mathrm{tr}\left(\boldsymbol{S}_j^2\right)\mathrm{tr}\left(\mathscr{T}^{(2)}\boldsymbol{P}_j\mathscr{T}^{(2)}\boldsymbol{P}_j\right)\big)\big] + \cdots\Bigg]^{1/2},\end{aligned} \tag{3.13}$$

where $\mathscr{T}^{(2)}$, $\mathscr{T}^{(3)}$ and Υ_2 are defined in (A 10) and (A 13). As we will see in § 5 the coefficient which multiplies σ^4 characterizes an interesting phenomenon of mixing between different scales of the flow. If only a first-order approximation is needed then from (3.9) and (3.13) we have

$$\langle\widetilde{\lambda}_j\rangle \simeq \lambda_j, \tag{3.14a}$$

$$\sigma_{\widetilde{\lambda}_j} \simeq 2\frac{\sigma\Delta t}{m_j}\sqrt{\mathrm{tr}\,\left(\mathscr{T}^{(2)}\boldsymbol{P}_j\right)}. \tag{3.14b}$$

This means that the perturbed weighted mean for the jth λ-group must lie in the neighbourhood of the unperturbed jth eigenvalue.

3.2. *Separation between perturbed energy levels*

We propose a simple criterion to identify the number of resolvable POD modes as a function of the random-flow standard deviation σ. Let ξ be a real positive parameter.

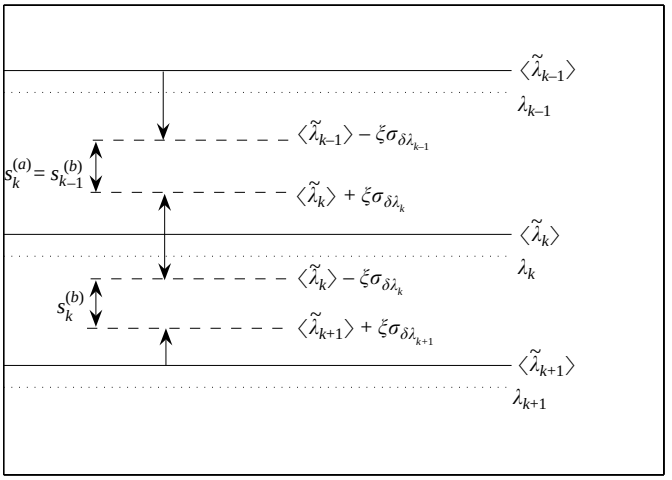

FIGURE 1. Separation between perturbed energy levels.

With probability related to the choice of ξ, the separation of the kth perturbed energy level from nearby (upper and lower) ones can be written as

$$\left.\begin{aligned} s_k^{(a)} &:= \langle\widetilde{\lambda}_{k-1}\rangle - \langle\widetilde{\lambda}_k\rangle - \xi\left(\sigma_{\widetilde{\lambda}_{k-1}} + \sigma_{\widetilde{\lambda}_k}\right) \quad \text{from above,} \\ s_k^{(b)} &:= \langle\widetilde{\lambda}_k\rangle - \langle\widetilde{\lambda}_{k+1}\rangle - \xi\left(\sigma_{\widetilde{\lambda}_k} + \sigma_{\widetilde{\lambda}_{k+1}}\right) \quad \text{from below.} \end{aligned}\right\} \tag{3.15}$$

A graphical representation of $s_k^{(a)}$ and $s_k^{(b)}$ is given in figure 1. Note that $s_k^{(a)} = s_{k-1}^{(b)}$. Therefore the condition

$$\min\left\{s_k^{(b)}, s_k^{(a)}\right\} = \min\left\{s_k^{(b)}, s_{k-1}^{(b)}\right\} \geqslant 0 \tag{3.16}$$

guarantees, with probability related to the choice of ξ, that the kth energy level remains separated from nearby ones. Equivalently, we define the function

$$\mathscr{D}_k(\sigma) := \frac{\sigma_{\widetilde{\lambda}_k} + \sigma_{\widetilde{\lambda}_{k+1}}}{\langle\widetilde{\lambda}_k\rangle - \langle\widetilde{\lambda}_{k+1}\rangle} \tag{3.17}$$

and write the condition (3.16) as

$$\mathscr{M}_k(\sigma) := \max\{\mathscr{D}_k(\sigma), \mathscr{D}_{k-1}(\sigma)\} \leqslant \frac{1}{\xi}. \tag{3.18}$$

If the inequality (3.18) is satisfied then the dimension of the kth perturbed eigenspace is constant with σ, which is the necessary condition for the convergence of all the perturbation series in § 3.1 and § 3.4 (see Kato 1995). In other words, if we solve the inequality (3.18) for σ (graphically or analytically), we obtain the maximum admissible random-flow standard deviation which ensures, with probability related to the choice of ξ, that the kth POD mode is resolvable. Therefore $\mathscr{M}_k(\sigma)$ will be called the *resolution indicator function* for the kth eigenspace and (3.18) the *resolution inequality*. We should mention that other authors have proposed different rules to identify the so-called 'knee' of the POD eigenspectrum (see Johnstone 2001; Sengupta & Mitra 1999 and the references therein).

3.3. *Statistics for the total energy of the perturbed field*

It is well known (Lumley 1970; Holmes, Lumley & Berkooz 1996) that the total energy of the flow field equals half of the trace of its autocorrelation operator. Given the unperturbed $\boldsymbol{U}\boldsymbol{U}^{\dagger}$ we have from (2.14) the trace identity

$$\operatorname{tr}(\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}) = \operatorname{tr}(\boldsymbol{U}\boldsymbol{U}^{\dagger}) + \operatorname{tr}(\boldsymbol{H}), \tag{3.19}$$

which is an energy balance. In fact equation (3.19) states that the total energy of the perturbed field $\widetilde{\boldsymbol{u}}$ equals the sum of the energies of the unperturbed field $\boldsymbol{u}$ and the perturbation. The average of $\operatorname{tr}(\boldsymbol{H})$ is

$$\langle \operatorname{tr}(\boldsymbol{H}) \rangle = \int_T \langle \mathscr{H}(t,t) \rangle \, \mathrm{d}t. \tag{3.20}$$

By substituting (2.15) into (3.20) and using (B 3*a*) and (B 3*d*) we see that (3.19) leads to

$$\left\langle \sum_{k=1}^{S} \widetilde{\mu}_k \right\rangle = \sum_{k=1}^{S} \mu_k + d\sigma^2 |\Omega||T|, \tag{3.21}$$

where S is the number of snapshots; $\widetilde{\mu}_k$ ($\widetilde{\mu}_k \geqslant 0$) are the (unknown) repeated eigenvalues of $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}$; μ_k ($\mu_k \geqslant 0$) are the repeated eigenvalues of $\boldsymbol{U}\boldsymbol{U}^{\dagger}$; $|\Omega|$, $|T|$ are the measures of the spatial and the temporal domains respectively. Therefore Gaussian uncorrelated perturbations always increase the average energy of the field. Interestingly this increment does not depend on the number of snapshots. As is easily seen, we can reach the same conclusion (3.21) in a more general setting, where the random field $\delta\boldsymbol{u}(\boldsymbol{x}, t)$ has zero mean and each component is in a wide sense stationary (Papoulis 1991; Lumley 1970). Using (B 3*b*), (B 3*e*) and (B 3*f*) we find that the standard deviation of $\operatorname{tr}(\boldsymbol{H})$ is

$$\sigma_{\operatorname{tr}(\boldsymbol{H})} = \sigma \Delta t \sqrt{4\Theta_2 + 2d\sigma^2 S \Upsilon_2} \tag{3.22}$$

where Θ_2 and Υ_2 are defined in (A 11) and (A 13) respectively. Formulae (3.21) and (3.22) are exact for random perturbations (2.16) of arbitrary 'magnitude' σ.

3.4. *Statistics for the random temporal modes*

One of the most remarkable results of the analytic perturbation theory for symmetric operators is the existence of a unitary transformation $Z(\sigma)$ (depending smoothly on the perturbation parameter σ) which transforms orthonormal families of eigenfunctions of $\boldsymbol{U}\boldsymbol{U}^{\dagger}$ into orthonormal families of eigenfunctions of $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^{\dagger}(\sigma)$. For single eigenspaces, under the hypothesis of absence of eigenvalue splitting, such a transformation takes the form

$$\boldsymbol{W}_j(\sigma) = \boldsymbol{P}_j + \sum_{n=1}^{\infty} \sigma^n \boldsymbol{W}_j^{(n)}, \tag{3.23}$$

where

$$\boldsymbol{W}_j^{(1)} = -\boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j, \tag{3.24a}$$

$$\boldsymbol{W}_j^{(2)} = -\boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(2)} \boldsymbol{P}_j + \boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j - \boldsymbol{S}_j^2 \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j - \tfrac{1}{2} \boldsymbol{P}_j \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j^2 \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j, \tag{3.24b}$$

$$\vdots$$

(see Kato 1995). The transformation (3.23) was derived by Kato (1950) in connection with the adiabatic theorem in quantum mechanics. Given an orthonormal family of eigenfunctions $\{\psi_{j_k}(t)\}$ of the unperturbed correlation operator $\boldsymbol{U}\boldsymbol{U}^\dagger$ the corresponding (not normalized†) perturbed eigenfunctions of $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger$ can be obtained by the mapping

$$\widetilde{\psi}_{j_k}(t) = \boldsymbol{W}_j(\sigma)\psi_{j_k}(t), \qquad k = 1, \dots, m_j, \tag{3.25}$$

i.e.

$$\widetilde{\psi}_{j_k}(t) = \psi_{j_k}(t) + \sigma \boldsymbol{W}_j^{(1)}\psi_{j_k}(t) + \sigma^2 \boldsymbol{W}_j^{(2)}\psi_{j_k}(t) + \cdots. \tag{3.26}$$

3.4.1. *The averaged temporal modes*

Using (B 3*a*) we have $\langle \boldsymbol{W}_j^{(1)} \rangle = 0$ and therefore (3.26) gives

$$\langle \widetilde{\psi}_{j_k}(t) \rangle = \psi_{j_k}(t) + \sigma^2 \langle \boldsymbol{W}_j^{(2)} \rangle \psi_{j_k}(t) + O\big(\sigma^4 \langle \boldsymbol{W}_j^{(4)} \rangle \psi_{j_k}(t)\big). \tag{3.27}$$

The operator $\langle \boldsymbol{W}_j^{(2)} \rangle$ is obtained in Appendix C:

$$\langle \boldsymbol{W}_j^{(2)} \rangle = \Delta t^2 \big(\big[\operatorname{tr}(\boldsymbol{S}_j)\boldsymbol{S}_j - \operatorname{tr}(\boldsymbol{P}_j)\boldsymbol{S}_j^2 - \tfrac{1}{2}\operatorname{tr}\big(\boldsymbol{S}_j^2\big)\boldsymbol{P}_j \big] \mathscr{T}^{(2)} - \tfrac{1}{2}\operatorname{tr}\,\big(\mathscr{T}^{(2)}\boldsymbol{S}_j^2\big)\,\boldsymbol{I}\big). \tag{3.28}$$

Equations (3.27) and (3.28) provide an approximation for the expected perturbed temporal modes which is accurate to the third order in σ.

3.4.2. *The standard deviation of the temporal modes*

We approximate the variance of the random temporal modes to the second order in σ. After some computations it is easy to see that the standard deviation can be expressed as

$$\begin{aligned} \sigma_{\widetilde{\psi}_{j_k}} &= \big[\sigma^2 \big\langle \big(\boldsymbol{W}_j^{(1)}\psi_{j_k}\big)^2 \big\rangle + \cdots\big]^{1/2} \\ &= \big[\sigma^2 \Delta t \big(\operatorname{diag}\big(\boldsymbol{S}_j \mathscr{T}^{(2)} \boldsymbol{S}_j\big) + \operatorname{diag}\big(\boldsymbol{S}_j^2\big)\big(\mathscr{T}^{(2)}\psi_{j_k}, \psi_{j_k}\big)_T\big) + \cdots\big]^{1/2} \end{aligned} \tag{3.29}$$

where $(\mathscr{T}^{(2)}\psi_{j_k}, \psi_{j_k})_T$ denotes the inner product $\sum_{l=1}^{S} [\mathscr{T}^{(2)}\psi_{j_k}]_l \psi_{j_k}(t_l)\Delta t$ and $\operatorname{diag}(\boldsymbol{A})$ stands for a vector whose components are the diagonal entries of the matrix $\boldsymbol{A}$. Equations (3.27) and (3.29) show that the perturbed temporal modes must lie in the neighbourhood of the unperturbed ones.

3.5. *Fourier analysis*

In view of the application presented in § 5.1 we need some tools to perform an analysis of the random temporal modes which is phase independent. It is well known (Lumley 1970; Therrien 1988; Oppenheim & Shafer 1975; Papoulis 1991) that the Fourier transform of the autocorrelation function

$$\widetilde{\mathscr{R}}(\alpha) = \int_{-\infty}^{\infty} \langle \widetilde{\psi}(t' + \alpha)\widetilde{\psi}(t') \rangle \, \mathrm{d}t' \tag{3.30}$$

equals the averaged power density of the process $\widetilde{\psi}(t')$, i.e. the expected Fourier power spectrum

$$\langle |\widetilde{\Psi}|^2 \rangle(w) = \int_{-\infty}^{\infty} \widetilde{\mathscr{R}}(\alpha) \mathrm{e}^{-\mathrm{j}w\alpha} \, \mathrm{d}\alpha. \tag{3.31}$$

This in turn implies that the autocorrelation function (3.30) ignores the phase information of $\widetilde{\psi}$.

† It is also possible to derive the expression for the transformation $\boldsymbol{Z}(\sigma)$ that preserves the norm (see Kato 1995). Here we only discuss the simpler form (3.23) that preserves orthogonality.

3.5.1. *The autocorrelation function*

We use the following estimator for (3.30) (see Therrien 1988; Oppenheim & Shafer 1975):

$$\widetilde{\mathscr{R}}_{|l|} = \Delta t \sum_{m=1}^{S-|l|} \langle \widetilde{\psi}_{m+|l|} \widetilde{\psi}_m \rangle, \qquad |l| \leqslant S. \tag{3.32}$$

Note that $\widetilde{\mathscr{R}}_l = \widetilde{\mathscr{R}}_{-l}$ and that $\widetilde{\mathscr{R}}_l$ is the sum of the elements in the lth diagonal of the autocorrelation matrix (F 1) given in Appendix F. The symmetry of the latter ensures that the autocorrelation function is an even function. Thus its Fourier transform is real and equals the averaged power density of the stochastic process $\widetilde{\psi}$. We remark that (3.32), calculated with (F 1), is valid under the hypothesis of absence of eigenvalue splitting.

4. The link between the random dynamics in time and space

In the fluid mechanics framework suitable combinations of POD spatial modes have been recognized to be coherent structures (Holmes *et al.* 1996; Liu 1988; Lumley 1970; Gad-El-Hak 2000). We derive a perturbation expansion for such modes in a rather general framework, i.e. for a random field $\widetilde{\boldsymbol{u}} = \boldsymbol{u} + \delta \boldsymbol{u}$ under only the hypothesis of absence of eigenvalue splitting. This can be successfully done using the dispersion relation (2.4*b*):

$$\boldsymbol{\Phi}_{j_k}(\boldsymbol{x}) = \frac{1}{\sqrt{\mu_{j_k}}} \int_T \boldsymbol{u}(\boldsymbol{x}, t) \psi_{j_k}(t)\, \mathrm{d}t. \tag{4.1}$$

We recall that the μ_{j_k} are the repeated eigenvalues of the temporal autocorrelation while the $\psi_{j_k}(t)$ are the unperturbed normalized temporal modes. If we assume that there is no splitting of eigenvalues, i.e. $\widetilde{\mu}_{j_k}(\sigma) = \widetilde{\lambda}_j(\sigma)$ for $k = 1, \ldots, m_j$, then the perturbed spatial modes can be calculated by the following equation:

$$\widetilde{\boldsymbol{\Phi}}_{j_k}(\boldsymbol{x}, \sigma) = \frac{1}{\sqrt{\widetilde{\lambda}_j(\sigma)}} \int_T (\boldsymbol{u}(\boldsymbol{x}, t) + \sigma \delta \widehat{\boldsymbol{u}}(\boldsymbol{x}, t)) \boldsymbol{W}_j(\sigma) \psi_{j_k}(t), \qquad k = 1, \ldots, m_j, \tag{4.2}$$

where $\delta \widehat{\boldsymbol{u}}(\boldsymbol{x}, t)$ is a random perturbation with unit variance and $\boldsymbol{W}_j(\sigma)$ is given in (3.23). As shown in Appendix D we can expand $\widetilde{\boldsymbol{\Phi}}_{j_k}(\boldsymbol{x}, \sigma)$ as follows:

$$\widetilde{\boldsymbol{\Phi}}_{j_k}(\boldsymbol{x}, \sigma) = \boldsymbol{\Phi}_{j_k}(\boldsymbol{x}) + \sigma \boldsymbol{N}_j^{(1)} \boldsymbol{\Phi}_{j_k} + \cdots. \tag{4.3}$$

where the operator $\boldsymbol{N}_j^{(1)}$, which represents the first-order correction for the unperturbed structure $\boldsymbol{\Phi}_{j_k}$, is

$$\boldsymbol{N}_j^{(1)} := \frac{1}{\lambda_j} \left(\boldsymbol{U}^\dagger \boldsymbol{W}_j^{(1)} \boldsymbol{U} + \delta \boldsymbol{U}^\dagger \boldsymbol{U} - \tfrac{1}{2} \widehat{\lambda}_j^{(1)} \boldsymbol{I} \right) \tag{4.4}$$

and

$$\delta \boldsymbol{U}^\dagger \phi := \int_T \delta \widehat{\boldsymbol{u}}\,(\boldsymbol{x}, t)\, \phi(t)\, \mathrm{d}t. \tag{4.5}$$

Note that the presence of $\delta \boldsymbol{U}^\dagger \boldsymbol{U}$ in (4.4) implies that it is possible to have a perturbation in the kth spatial mode even though, for a certain determination of $\delta \boldsymbol{u}(\boldsymbol{x}, t)$, the kth temporal mode is perfectly resolved (i.e. $\boldsymbol{W}_j^{(k)} = 0$, $\widehat{\lambda}_j^{(k)} = 0$). Also, the perturbation is in inverse proportionality to λ_k (which is the energy of the unperturbed kth mode). The operator $\boldsymbol{N}_j^{(1)}$ breaks the symmetry between the resolvable part of the temporal

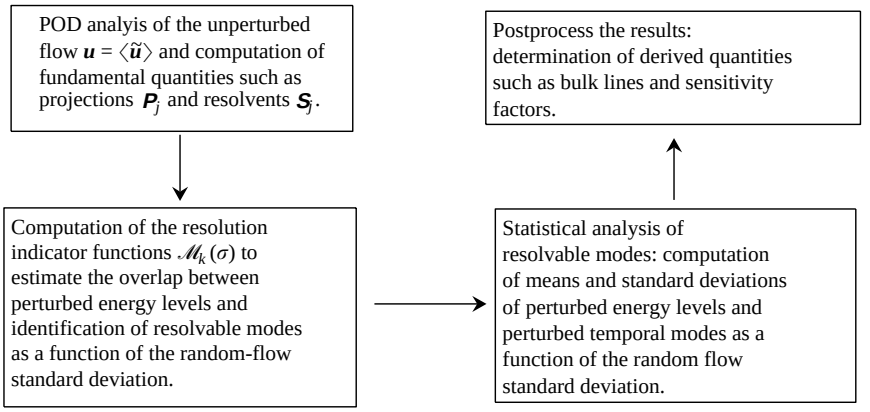

FIGURE 2. Principle sketch for perturbation analysis.

and the spatial dynamics. In other words one could not expect that the presence of perturbations in the ensemble of data are equally 'shared' between the space and the time dynamic. A similar analysis was performed qualitatively in Venturi & Karniadakis (2004). Moreover we observe that first-order perturbation in the structure $\boldsymbol{\Phi}_{j_k}(\boldsymbol{x})$ is due to the following three contributions:

(i) the first-order temporal perturbation $\boldsymbol{W}_j^{(1)} \psi_{j_k}(t)$;
(ii) the first-order perturbation in energy $\widehat{\lambda}_j^{(1)}$;
(iii) the perturbation in the data through the operator (4.5).

5. Applications

In this section we address two applications of the POD perturbation theory developed in § 2 and § 3 to the incompressible flow past a cylinder at $Re = 100$ and to the natural convective flow above an isothermal, highly elongated, horizontal plate at Rayleigh number $Ra = 4.75 \times 10^6$. The perturbation analysis of these flows will be performed following the principles shown in figure 2.

The first step is the POD analysis of the unperturbed (or 'mean') flow $\langle \widetilde{\boldsymbol{u}} \rangle = \boldsymbol{u}$ and the computation of quantities defined in Appendix A such as projections $\boldsymbol{P}_j$, reduced resolvents $\boldsymbol{S}_j$, matrices $\mathscr{T}^{(z)}$ and traces Θ_z. Having the POD modes of $\langle \widetilde{\boldsymbol{u}} \rangle$ and the statistical characterization of the perturbation $\delta \boldsymbol{u}$ (through the probability density (2.16) and relations derived in Appendix B) we compute the statistics for the POD modes of $\widetilde{\boldsymbol{u}} = \langle \boldsymbol{u} \rangle + \delta \boldsymbol{u}$ as function of the standard deviation of $\delta \boldsymbol{u}$ (or, equivalently, $\widetilde{\boldsymbol{u}}$) using the equations derived in § 3. As discussed in § 3.2 these statistics make sense for the so-called resolvable modes which are identifiable using the resolution indicator functions (3.18).

5.1. *Incompressible flow past a cylinder*

We consider two-dimensional incompressible flow past a circular cylinder (Zdravkovich 1997) at Reynolds number $Re = 100$, based on the cylinder diameter D and on the inflow velocity. We have simulated the time-dependent Navier–Stokes flow using the spectral/hp finite element method described in Karniadakis & Sherwin (1999) on a mesh consisting of 412 spectral triangular elements (figure 3) of order $p = 8$. After we had established the fully periodic state, we extracted from the DNS 40 equidistant snapshots in a period. The POD analysis and the low-dimensional

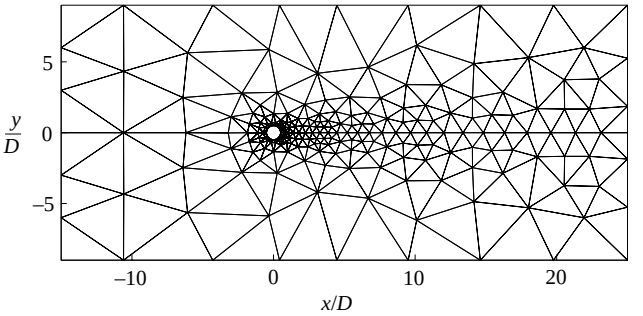

FIGURE 3. Geometry and computational mesh. The domain is composed of 412 spectral elements each with 9×9 collocation points.

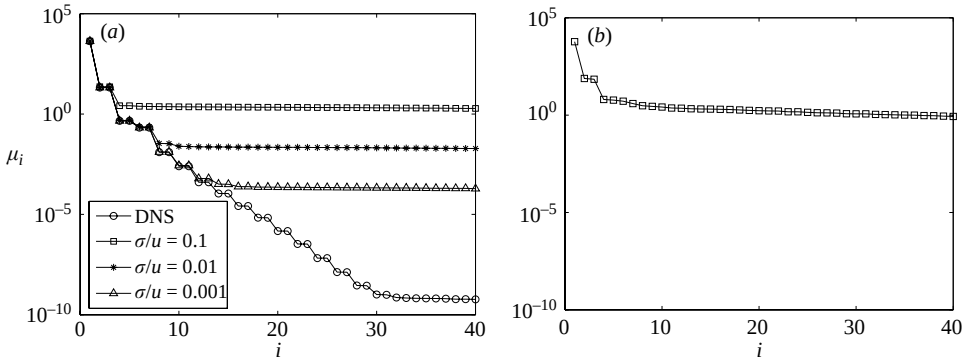

FIGURE 4. (*a*) DNS-based POD eigenspectrum (unperturbed flow) and spectra of perturbed flows corresponding to different σ/u. (*b*) Experimental POD eigenspectrum from 2D-PIV measurements of flow past a cylinder at $Re \simeq 100$. (Courtesy of B. Pulvirenti, University of Bologna.)

modelling of such unperturbed flows have been studied extensively during the past few years (Ma & Karniadakis 2002; Ma *et al.* 2003; Sirisup & Karniadakis 2004; Ma 2001; Noack *et al.* 2003; Deane *et al.* 1991; see also Venturi & Karniadakis 2004).

In order to have consistent results for an entire class of flows we use the non-dimensional version of the perturbation theory (3.4) where the dimensionless perturbation parameter σ/u is scaled by the unperturbed inflow streamwise velocity u. We have simulated random space–time additive perturbations $\delta\boldsymbol{u}(\boldsymbol{x}_i, t_l)$ according to the probability density (2.16) at different levels σ/u and constructed an ensemble of 40 000 independent possible realizations of the perturbed flow $\widetilde{\boldsymbol{u}}$ for each level. The effects of such perturbations on the POD eigenspectrum are shown in figure 4 where we also plot an experimental eigenspectrum from 2D-PIV measurement of this flow (courtesy of B. Pulvirenti, University of Bologna). The similarity shown in figure 4 between the perturbed spectrum corresponding to $\sigma/u = 0.1$ (*a*) and the experimental PIV-based spectrum (*b*) suggests a guideline for designing the experiment in order to increase the number of POD-resolved modes. In fact, assuming that the standard deviation of the PIV system is approximately constant within a certain range of displacements (Raffel, Willert & Kompenhans 1998; Westerweel 2000) and that enough spatial and temporal resolution is available, the increase of the inflow streamwise velocity u will lower the ratio σ/u and improve the POD-resolution (see

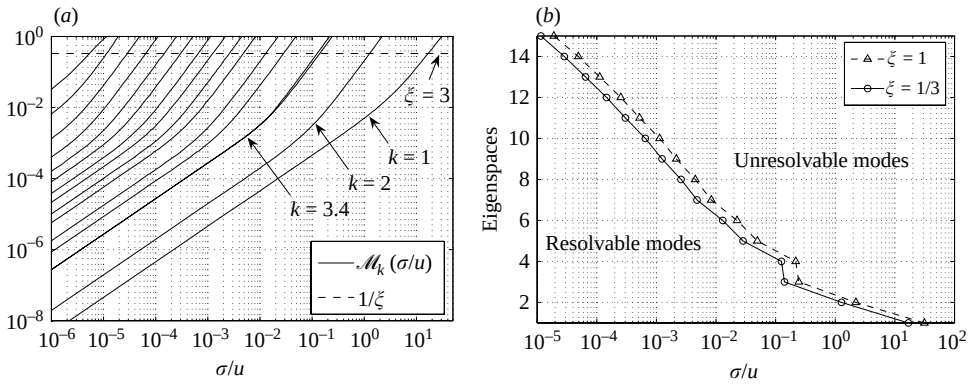

FIGURE 5. Resolvable eigenmodes as a function of the dimensionless random-flow standard deviation σ/u.

figure 4). Obviously the cylinder's diameter needs to be reduced in order to keep Re constant. We conclude that we can obtain an enhancement of the POD-resolution by going towards smaller length scales.

5.1.1. *Statistics for the perturbed energy levels*

As shown in figure 4 at $Re = 100$ all the unperturbed POD eigen-states (except the first, which approximately represents the mean field) are degenerate with degeneracy 2.† Therefore equations (3.9) and (3.13) characterize the relevant statistics for the weighted means of the perturbed energy levels. In other words if $\mu_{2_1} = \mu_{2_2}$ are two repeated unperturbed eigenvalues and $\lambda_2 = \mu_{2_1} = \mu_{2_2}$ is their weighted mean then it is possible to compute the statistics for $\widetilde{\lambda}_2 = (\widetilde{\mu}_{2_1} + \widetilde{\mu}_{2_2})/2$. In the absence of splitting we also have $\widetilde{\lambda}_2 = \widetilde{\mu}_{2_1} = \widetilde{\mu}_{2_2}$.

5.1.2. *The separation of the energy levels and the 'bulk' spectrum*

In figure 5(*a*) we plot the resolution indicator functions (3.18) for all the eigenspaces. As pointed out in § 3.2, the energy level of the kth perturbed temporal eigenspace is distinguishable from the other levels if the inequality (3.18) is satsified. We explain the meaning of the figure 5(*a*) by a simple example. Let us assume for simplicity that the probability density of all the perturbed energy levels is Gaussian. If the random-flow standard deviation is 1 % of the inflow streamwise velocity, i.e. $\sigma/u = 0.01$, then from figure 5 we see that it is possible to resolve 6 eigenspaces with probability 100 % ($\xi = 4.59$); 7 eigenspaces with probability 43.81 % ($\xi = 0.58$). Equivalently, the intersection of the resolution indicator function $\mathscr{M}_k(\sigma/u)$ with the line $1/\xi = 1/3$ identifies the maximum random-flow dimensionless standard deviation for which, with probability 99.73 %, the kth eigenspace is resolvable. Therefore all the perturbed eigenspaces represented by the continuous lines below the dashed line 1/3 are resolvable with very high probability. A comparison between figure 5 and figure 4 supports this conclusion. However, the fact that the weighted means of different perturbed energy levels are distinguishable does not guarantee that single perturbed energy levels belonging to different $\widetilde{\lambda}$-groups are not allowed to interfere. In figure 6 we plot the weighted means of the perturbed energy levels for the sixth and seventh

† As is well known the stated degeneracy of the wake POD is exact in the case of exact streamwise periodicity. In most cases we should refer to an approximate degeneracy.

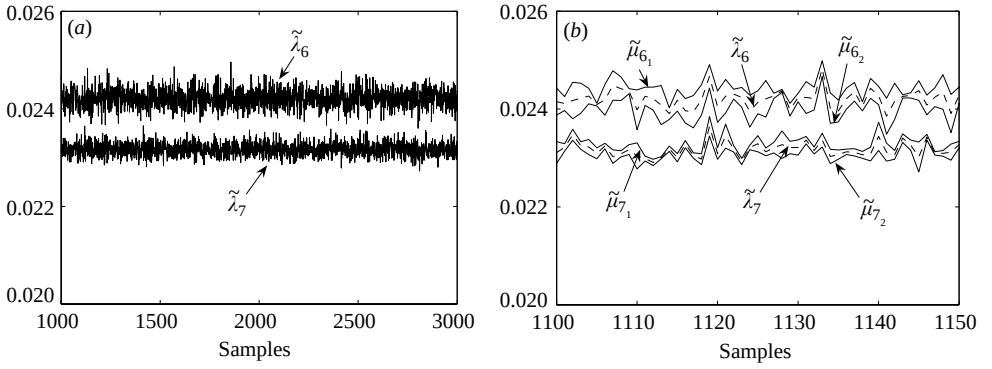

FIGURE 6. $\sigma/u = 0.01$. (*a*) Weighted means of perturbed eigenvalues and (*b*) single perturbed eigenvalues versus the number of samples.

eigenspaces as well as single energy fluctuations versus the sample number. We see that even though the weighted means of these eigenvalues remain distinguishable, single perturbed split eigenvalues may interfere.

We conclude that with degenerate unperturbed eigenvalues the meaning of figure 5(*a*) is semi-quantitative in the sense that there is a region close the dashed line that should be treated with caution for inferences about single eigenvalues (it is rigorous for their weighted means). If all the unperturbed eigenvalues are simple (as they are in § 5.2) or if they do not split at all, then the predictions of figure 5 are rigorous.

5.1.3. *Average energies and standard deviations*

In figure 7 we plot the expected values (3.9) and the standard deviations (3.13) for all the weighted means of the perturbed eigenvalues $\widetilde{\lambda}_k$ as a function of the dimensionless perturbation parameter σ/u. The comparison with the ensemble means and the ensemble standard deviations obtained by the Monte Carlo method described in § 5.1 supports the accuracy of the theoretical predictions.

An interesting result is represented by the asymptotic trend of all the standard deviations to the same line which we shall call *bulk line*. This definition arises from the fact that the mixing between different random energy levels occurs when they have the same standard standard deviation (the *bulk* level) which is a function of the flow uncertainty σ/u. It is easy to see from the structure of the power series expansion (3.13)

$$\sigma_{\widetilde{\lambda}_j} = \sqrt{b_j \left(\frac{\sigma}{u}\right)^2 + c_j \left(\frac{\sigma}{u}\right)^4} \tag{5.1}$$

that the bulk line has equation

$$\mathscr{B} := \sqrt{c}\left(\frac{\sigma}{u}\right)^2 \tag{5.2}$$

where the coefficient $c \simeq 4.333$ depends only on the unperturbed dimensionless flow and turns out to be a constant for all the perturbed eigenmodes, that is $c = c_j \ \forall j$. This coefficient is therefore an interesting invariant for this flow and it is approximately equal to $10^{-3}\lambda_1$. As we will see in § 5.2 the standard deviation for the perturbed energy levels of the convective flow exhibits the same behaviour.

We conclude that there exists a limiting lower level of uncertainty in the energy of the perturbed modes for which the modes themselves are resolvable. As shown in

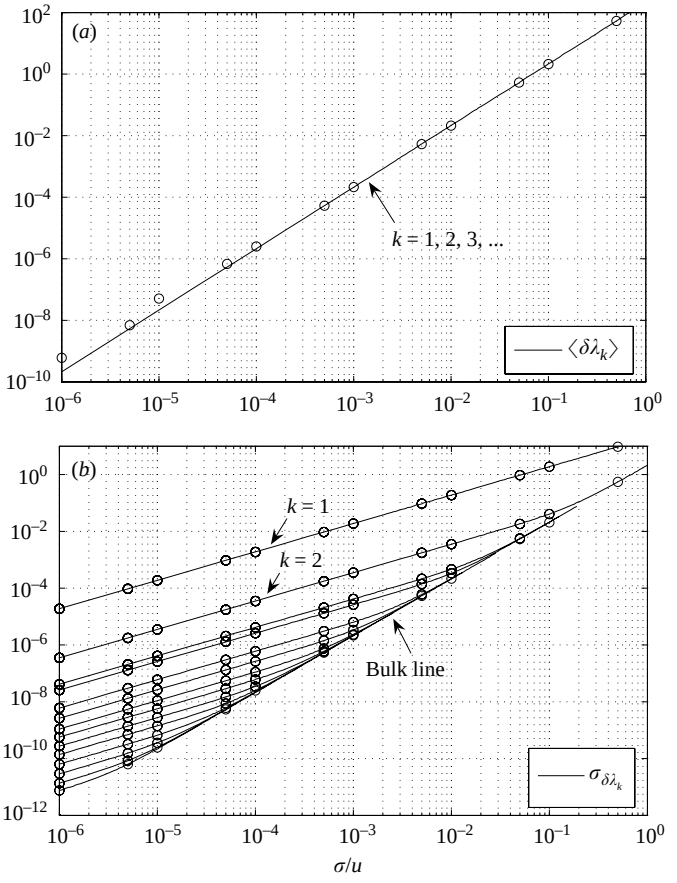

FIGURE 7. (*a*) Expected value and (*b*) standard deviation for all the perturbed energy levels as a function of the dimensionless random flow standard deviation σ/u. Comparison between the theoretical formulae (3.9), (3.13) (continuous lines) and the results (circles) from Monte Carlo simulations (40 000 samples).

figure 7(*b*) and figure 8 this limiting value can be much lower than the uncertainty in the flow field σ/u and it is given explicitly by equation (5.2). Equivalently we shall say that every inference about the energy levels of the perturbed temporal modes is meaningless if it is more accurate than $\sqrt{c}(\sigma/u)^2$.

All the averaged perturbed energy levels (3.9) are increased by approximately the same quantity $(\sigma/u)^2\Delta t|\Omega|d$ ($|\Omega|$, Δt are dimensionless) which is mode-independent. In fact we have found that the other term in (3.9) is negligible with respect $(\sigma/u)^2\Delta t|\Omega|d$ for all the indexes. The comparison with the ensemble means obtained by the Monte Carlo method shown in figure 7(*a*) supports this conclusion. This also implies that the ratio between the averaged perturbation energies $\langle\delta\lambda_k\rangle$ and the bulk line (5.2) is a constant which depends only on the dimensionless space–time domain and $\sqrt{c}$:

$$\frac{\langle\delta\lambda_k\rangle}{\mathscr{B}} = \frac{|\Omega|\Delta t d}{\sqrt{c}}. \tag{5.3}$$

The deviation of the ensemble means from the theoretical means observed in figure 7(*a*) for very small σ/u is due to the limited sample size. In fact we observe

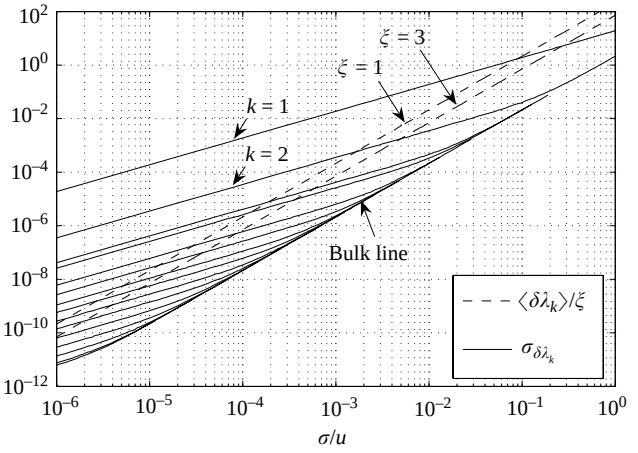

FIGURE 8. Statistics for the perturbed energy levels: $\sigma_{\widetilde{\lambda}_k}$ (continuous lines) and $\langle \delta\lambda_k \rangle / \xi$ (dashed lines).

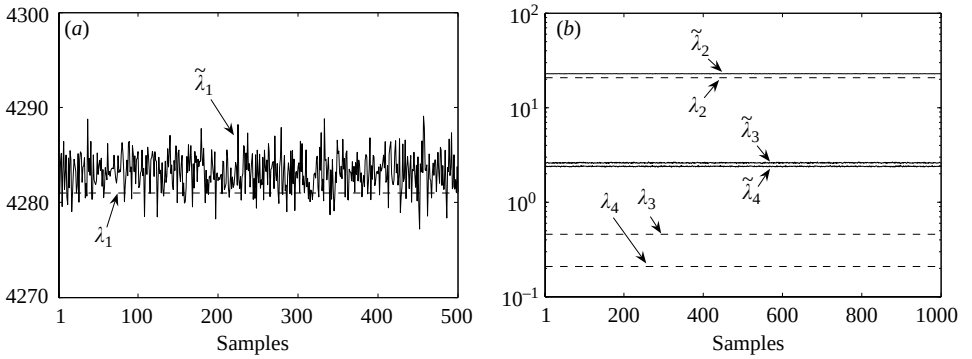

FIGURE 9. $\sigma/u = 0.1$. Weighted means λ_k and $\widetilde{\lambda}_k$, for (*a*) $k = 1$ and (*b*) $k = 2, 3, 4$ versus the number of samples.

that if certain perturbed eigenvalues have a standard deviation which is much greater than their mean then 40000 samples could not be enough to ensure the closeness of the ensemble mean to the 'true' analytical mean (see Appendix G).

We may ask when the perturbed energy levels are greater than the unperturbed ones. The condition

$$\sigma_{\widetilde{\lambda}_k} \leqslant \frac{1}{\xi} \langle \delta\lambda_k \rangle, \quad \xi > 0, \tag{5.4}$$

answers this question with statistical confidence related to the choice of ξ. The substitution of equations (3.9) and (3.13) into (5.4) gives an inequality which can be solved analytically for σ/u. Graphically we can draw both $\sigma_{\widetilde{\lambda}_k}$ and $\langle \delta\lambda_k \rangle / \xi$ and look for the resolved eigenspaces inside the strip between the 'bulk' line (5.2) and the lines $\langle \delta\lambda_k \rangle / \xi$ (see figure 8). For instance we see that if $\sigma/u = 0.1$ then four eigenspaces are resolvable: the second, the third and the fourth fall inside the aforementioned strip. Therefore there is a high probability that they have an energy which is greater than the corresponding unperturbed energy. To document this we plot in figure 9 the energy fluctuations versus the sample number for $\sigma/u = 0.1$. We note that the

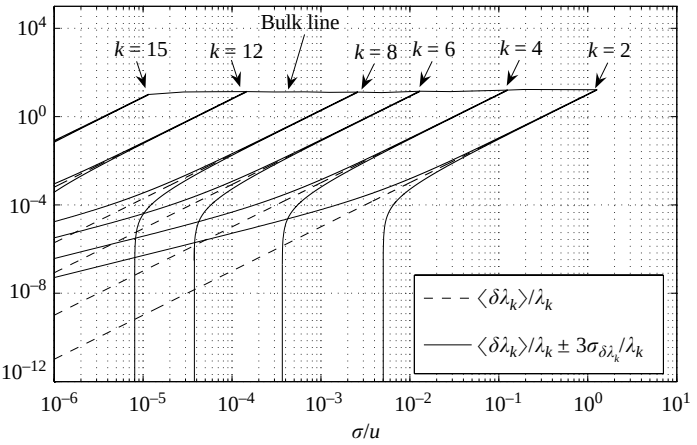

FIGURE 10. Sensitivity of the energy levels to perturbations: $\langle\delta\lambda_k\rangle/\lambda_k$ (dashed lines) and $\langle\delta\lambda_k\rangle/\lambda_k \pm 3\sigma_{\delta\lambda_k}/\lambda_k$ (continuous lines) for the second, fourth, sixth, eighth, twelfth and fifteenth eigenspaces as function of the dimensionless perturbation parameter σ/u.

predictions of figure 8 are confirmed, i.e. the energy of the perturbed mean field ($\widetilde{\lambda}_1$) can be greater or less than the unperturbed energy while the energy levels of the other perturbed eigenspaces are certainly greater. Note that the third and the fourth eigenspaces are practically mixed as they are very close to the bulk line for $\sigma/u = 0.1$ (see also figure 5).

5.1.4. *Disturbance-sensitive structures*

We would like to study the sensitivity of the energy levels with respect to perturbations in the flow. To this end in figure 10 we plot

$$\frac{\langle\delta\lambda_k\rangle}{\lambda_k} \quad k = 2, 4, \ldots \tag{5.5}$$

and the following confidence intervals

$$\frac{\langle\delta\lambda_k\rangle}{\lambda_k} \pm \xi\frac{\sigma_{\delta\lambda_k}}{\lambda_k}, \quad k = 2, 4, \ldots, \tag{5.6}$$

for $\xi = 3$. Figure 10 shows that for every mode the 'bulk' phenomenon takes place when the average perturbation energy $\langle\delta\lambda_k\rangle$ is approximately 10 times the corresponding unperturbed energy level λ_k. This ratio is determined by the intersections of the resolution indicator functions $\mathscr{M}_k(\sigma/u)$ shown in figure 5(a) with the line $1/\xi = 1/3$.

5.1.5. *Statistics for the total energy of the perturbed field and the norm of* $\boldsymbol{H}$

The relevant statistics for the total energy of the perturbation and the norm perturbation operator $\boldsymbol{H}$ are given in § 3.3 and Appendix E respectively. In figure 11 we compare such theoretical results to the ensemble means and ensemble standard deviations obtained by the Monte Carlo method. The difference between the ensemble means and the theoretical means observed in figure 11(a) for small σ/u is due to the limited sample size. For $\sigma/u \geqslant 7 \times 10^{-3}$ there is a high probability (level $\xi = 3$) that the total energy of the perturbed field is greater than the energy of the unperturbed one.

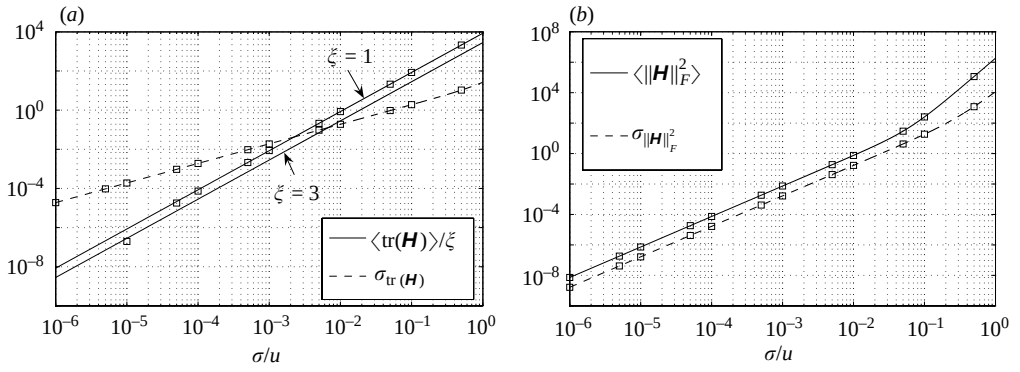

FIGURE 11. Statistics for (*a*) the total energy of the perturbation and (*b*) the Frobenius norm of the perturbation operator $\boldsymbol{H}$ as function of the dimensionless perturbation parameter σ/u. The squares are the results from Monte Carlo simulations (40 000 samples).

This limiting value can be expressed analytically by (3.21) and (3.22) with statistical confidence related to the choice of ξ:

$$\langle \mathrm{tr}\,(\boldsymbol{H})\rangle \geqslant \xi \sigma_{\mathrm{tr}\,(\boldsymbol{H})} \quad \Rightarrow \quad \frac{\sigma}{u} \geqslant 2\xi \sqrt{\frac{\Theta_2}{d^2|\Omega|^2 S^2 - 2dS\Upsilon_2}}, \tag{5.7}$$

where all the quantities on the right-hand side are calculated with the unperturbed dimensionless field and the dimensionless geometry.

5.1.6. *Random temporal modes*

As shown in figure 4, every unperturbed POD eigenspace (except the mean field) is two-dimensional and spanned by periodic functions which represent the time evolution of the Kármán–Bénard downstream wake. It is well known that any linear combination of eigenfunctions corresponding to the same degenerate eigenvalue is still an eigenfunction. This implies that the periodic POD eigenmodes are defined up to an arbitrary phase. Therefore to test our formulae we need to perform a Fourier analysis of the sampled temporal modes to remove this arbitrary phase. As pointed out in § 3.5 instead of the Fourier spectrum we shall compare its generating function, i.e. the autocorrelation function (3.30). This is done in figure 12 for the second, third, seventh and ninth modes. In figure 13 we plot the relevant statistics for the third and ninth random temporal modes. We note that even though these modes are near the bulk spectrum, for $\sigma/u = 0.1$ and $\sigma/u = 0.01$ respectively (see figure 4), the theoretical autocorrelations (3.32) obtained by a first-order transformation function (3.26) are very close to the sampled autocorrelations. This leads to the conclusion that the resolvable perturbed temporal modes are not very sensitive to perturbations.

5.1.7. *Statistics for the fluctuating components of the velocity field*

We consider the fluctuating components of the perturbed velocity field

$$\widetilde{\boldsymbol{u}}'(\boldsymbol{x}, t) := \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t) - \boldsymbol{U}(\boldsymbol{x}) \tag{5.8}$$

where the mean flow $\boldsymbol{U}(\boldsymbol{x})$ is defined as

$$\boldsymbol{U}(\boldsymbol{x}) := \frac{1}{|T|}\int_T \langle \widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)\rangle \,\mathrm{d}t = \frac{1}{|T|}\int_T \boldsymbol{u}(\boldsymbol{x}, t)\,\mathrm{d}t. \tag{5.9}$$

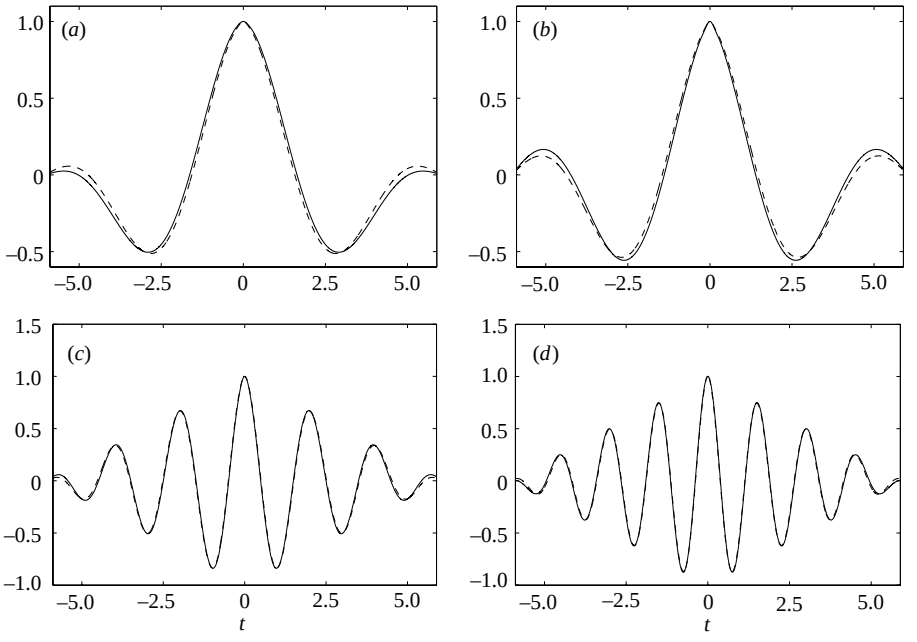

FIGURE 12. Comparison between the theoretical (—) and the Monte Carlo sampled (– –) autocorrelation functions for (*a*) the second, (*b*) the third, (*c*) the seventh and (*d*) the ninth random temporal modes. The perturbation magnitude is $\sigma/u = 0.1$ for (*a*) and (*b*); $\sigma/u = 0.01$ for (*c*) and (*d*).

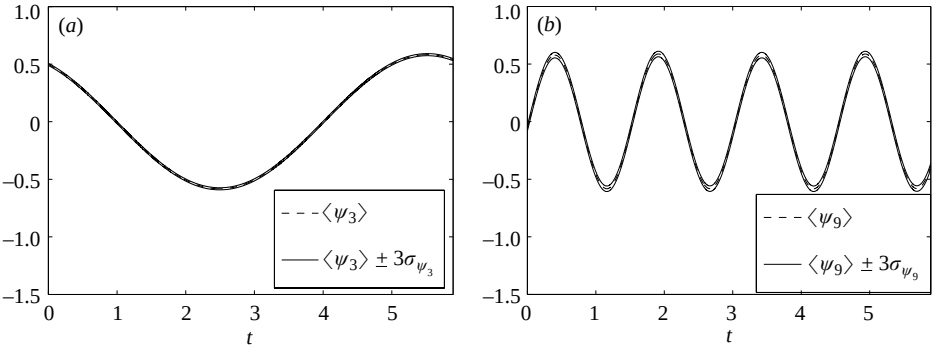

FIGURE 13. (*a*) Third and (*b*) ninth random temporal modes corresponding to $\sigma/u = 0.1$ and $\sigma/u = 0.01$ respectively.

The perturbation analysis is performed following the sketch of figure 2. The POD eigenspectrum of the unperturbed fluctuating velocity field

$$\boldsymbol{u}'(\boldsymbol{x}, t) := \boldsymbol{u}(\boldsymbol{x}, t) - \boldsymbol{U}(\boldsymbol{x}) \tag{5.10}$$

is shown in figure 14(*a*) together with some samples of perturbed spectra corresponding to different σ/u; we note that the unperturbed spectrum is in very good agreement with most POD studies of wakes past a cylinder (see for instance Noack *et al.* 2003; Ma & Karniadakis 2002). In figures 14(*b*), 15, 16 we show the results of the perturbation theory applied to $\widetilde{\boldsymbol{u}}'$. Here the perturbation parameter σ/u is still scaled on the unperturbed inflow streamwise velocity u and the number

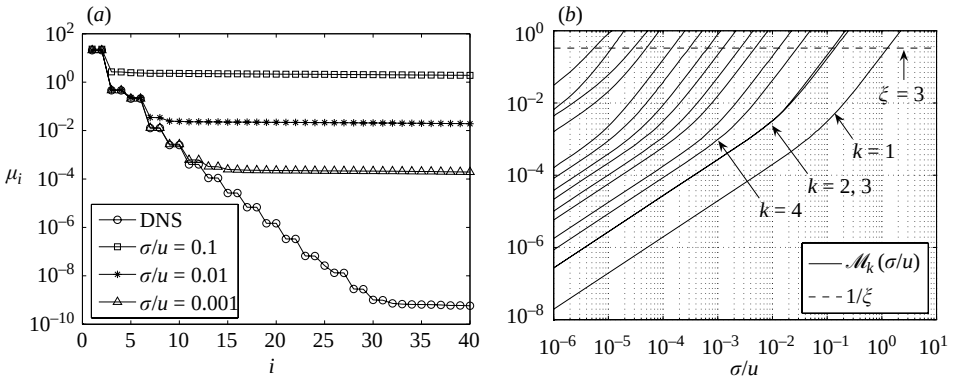

FIGURE 14. (*a*) POD eigenspectrum of the fluctuating components of the velocity field and (*b*) resolvable eigenmodes as a function of the dimensionless random-flow standard deviation σ/u.

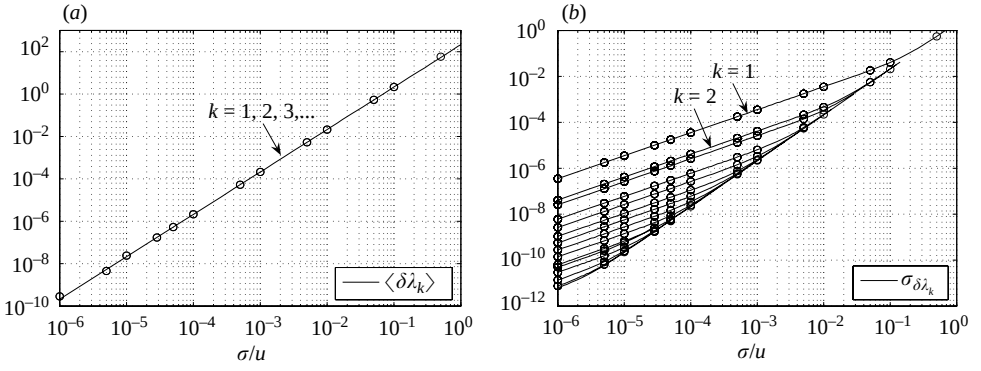

FIGURE 15. (*a*) Mean and (*b*) standard deviation of perturbed energy levels for the fluctuating components of the velocity field as a function of the dimensionless random-flow standard deviation σ/u. Comparison between the theoretical formulae (continuous lines) and the Monte Carlo simulations (circles).

of independent samples generated by the Monte Carlo method is still 40000. As is easily seen, the conclusions that can be drawn from the perturbation analysis are of exactly the same type as in the previous sections. However here we note the interesting behaviour of the 13th eigenspace whose unperturbed eigenvalue is simple (number 25 in figure 14*a*). For this mode the bulk phenomenon takes place at a standard deviation which is greater than the bulk line level $\mathscr{B}$. Moreover the hierarchy of magnitudes of the standard deviations is broken (see figure 16). This result, confirmed by the Monte Carlo simulations, can be due to the different multiplicity of the 13th eigenvalue ($\lambda_{13} = \mu_{25}$).

5.2. *Natural convective flow over an isothermal horizontal plate*

We consider the heat transfer and natural convective two-dimensional flow over an isothermal upward-facing plate. Natural convective flows induced over horizontal and inclined heated plates have been the subject of numerous investigations (Goldstein & Kei-Shun 1983; Fujii & Imura 1972; Kitamura & Kimura 1995; Chambers & Lee 1997; Yousef, Tarasuk & McKenn 1982; Sparrow & Carlson 1986; Al-Arabi & El-Riedy 1976; Hassan & Mohamed 1970) in recent decades. We have performed a

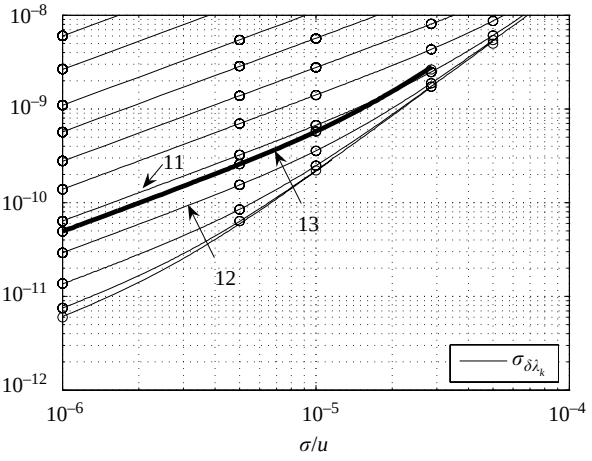

FIGURE 16. Behaviour of the 13th eigenspace for the fluctuating components of the velocity field (bold line): the bulk phenomenon takes place at a standard deviation which is greater than the bulk line level and the hierarchy of magnitudes of the standard deviations is broken. The Monte Carlo simulations (circles) confirm this result.

direct numerical simulation of the time-dependent Boussinesq equations

$$\left.\begin{gathered}\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = \frac{Gr}{Re^2}\theta\,\boldsymbol{j} - \nabla p + \frac{1}{Re}\nabla^2\boldsymbol{u}, \\ \frac{\partial \theta}{\partial t} + \boldsymbol{u}\cdot\nabla\theta = \frac{1}{RePr}\nabla^2\theta, \\ \nabla\cdot\boldsymbol{u} = 0,\end{gathered}\right\} \tag{5.11}$$

at Rayleigh number $Ra = 4.75\times10^6$ ($Pr = 0.72$, $Re = 1000$), using the high-order spectral/hp finite element method described in Karniadakis & Sherwin (1999). In (5.11) $\boldsymbol{j}$ is the upward unit vector while $\boldsymbol{u}$ and θ are the dimensionless velocity and temperature fields respectively. The dimensionless temperature is defined by

$$\theta = \frac{T - T_\infty}{T_w - T_\infty} \tag{5.12}$$

where T_w is the uniform temperature at the plate surface and T_∞ is the reference temperature. Also,

$$Gr = \frac{g\beta\,(T_w - T_\infty)\,L^3}{\nu^2}, \quad Pr = \frac{\alpha}{\nu}, \quad Ra = GrPr. \tag{5.13}$$

Figure 17 shows the computational domain and the boundary conditions, which are very similar to those used by Martorell, Herrero & Grau (2003). Kimura *et al.* (2002), Kitamura & Kimura (1995) and Martorell *et al.* (2003) performed experiments with horizontal and inclined heated plates with imposed heat flux in similar domains. The flow starts from rest.

A Fourier analysis of the time series for the velocity (see figure 18) and the temperature fields reveals that the highest dimensionless time frequency is everywhere less than $f_{sup} = 6$. We extracted $S = 65$ regular spaced snapshots from the DNS using a dimensionless sampling frequency $f = 16 > 2f_{sup}$ in order to satisfy Nyquist's criterion. The dimensionless frequency 0.246 characterizes the most energetic dimensionless vortex shedding close to the heated plate (see figure 18). In figure 19 and figure 20

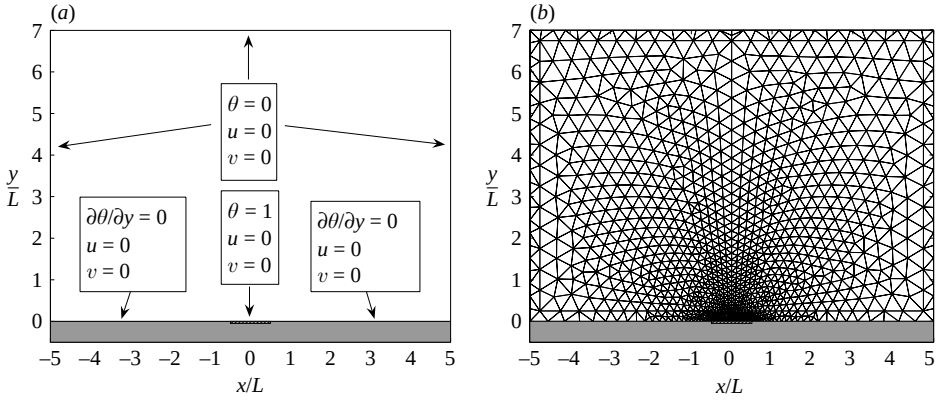

FIGURE 17. (*a*) Schematic of the geometry, boundary conditions and (*b*) computational mesh. The domain is composed of 2188 spectral elements each with 9×9 collocation points.

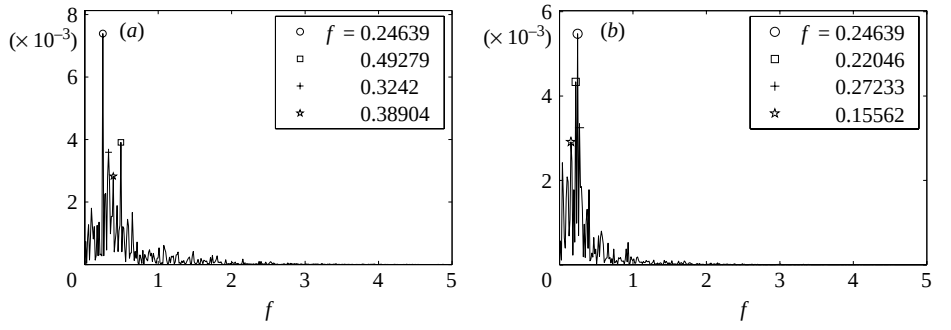

FIGURE 18. Fourier spectrum for the time series of (*a*) the y and (*b*) x velocity components at the point $x^* = -0.1$, $y^* = 0.23$.

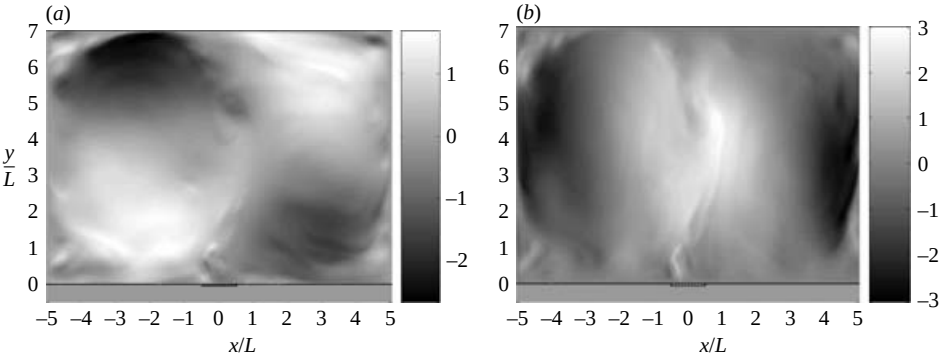

FIGURE 19. (*a*) Horizontal and (*b*) vertical components of the dimensionless velocity field.

we show one snapshot of the dimensionless flow and temperature fields. In order to compare the heat transfer features of plates of various shapes Goldstein, Sparrow & Jones (1973) proposed a length scale L defined as the active surface divided by its perimeter. For our infinite horizontal strip we have

$$L = \lim_{b\to\infty} \frac{ab}{2(a+b)} = \frac{a}{2}, \tag{5.14}$$

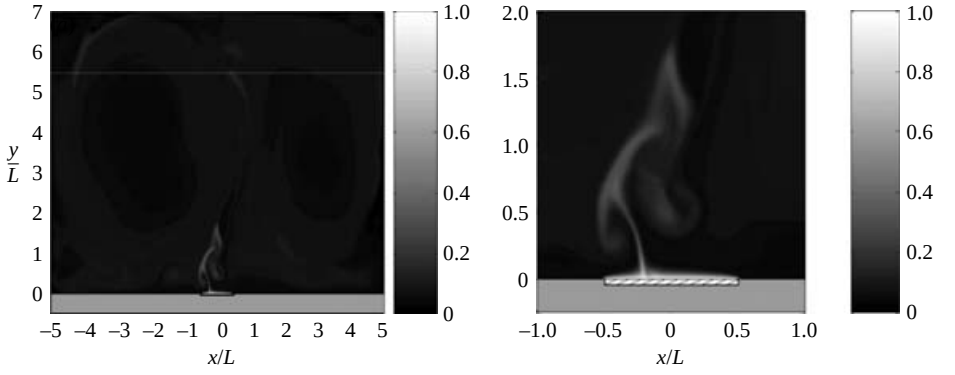

FIGURE 20. Dimensionless temperature field. Snapshot on the right is a close-up.

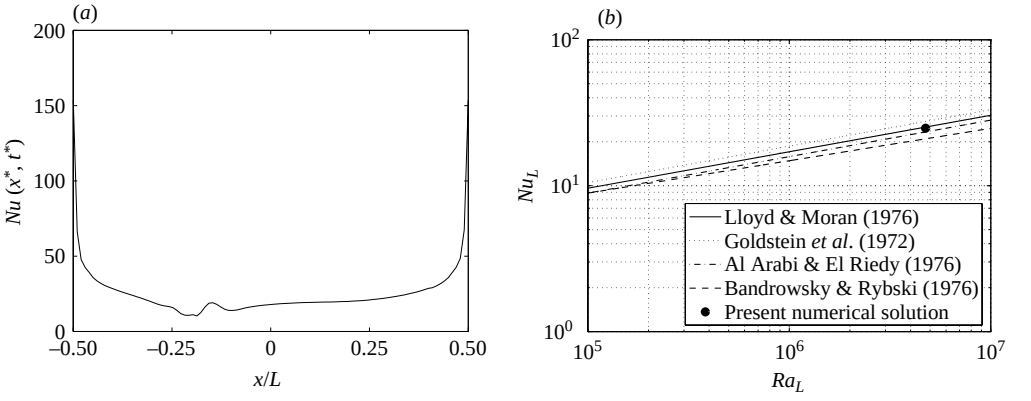

FIGURE 21. (*a*) Local Nusselt number corresponding to the snapshot shown in figure 20 and (*b*) published heat and mass transfer correlations for natural convection over horizontal plates. The area over perimeter ratio is used as characteristic length.

where a and b are the width and the length of the plate respectively. Many experimental $Nu - Ra$ correlations such as those given by Al-Arabi & El-Riedy (1976) and Bandrowski & Rybski (1976), originally based upon the width of the plate a, have been recalculated by Goldstein & Kei-Shun (1983) as a function of the area to perimeter length scale L. In figure 21(*b*) we plot the correlations proposed by Al-Arabi & El-Riedy (1976), Lloyd & Moran (1974), Bandrowski & Rybski (1976) and Goldstein *et al.* (1973) for the dependence of the averaged Nusselt number on the Rayleigh number when these dimensionless groups are based on the area to perimeter ratio (see Goldstein & Kei-Shun 1983). The correlations of Bandrowski & Rybski (1976) and Goldstein *et al.* (1973) are extrapolated from experiments at $Ra_L \simeq 2.5 \times 10^5$ and $Ra_L \simeq 10^4$ respectively. Summarizing the main findings of other studies (Hassan & Mohamed 1970; Fujii & Imura 1972; Yousef *et al.* 1982; Sparrow & Carlson 1986) we can say that the averaged Nusselt number corresponding to our Rayleigh number should be between 20 and 30 (see also Martorell *et al.* 2003). This is what we have found as demonstrated in figure 21(*a*) where we plot the local Nusselt number on the plate surface

$$Nu(x^*, t^*) = \left(\frac{\partial \theta}{\partial y^*}\right)_{y=0}, \quad x^* \in [-0.5, 0.5], \tag{5.15}$$

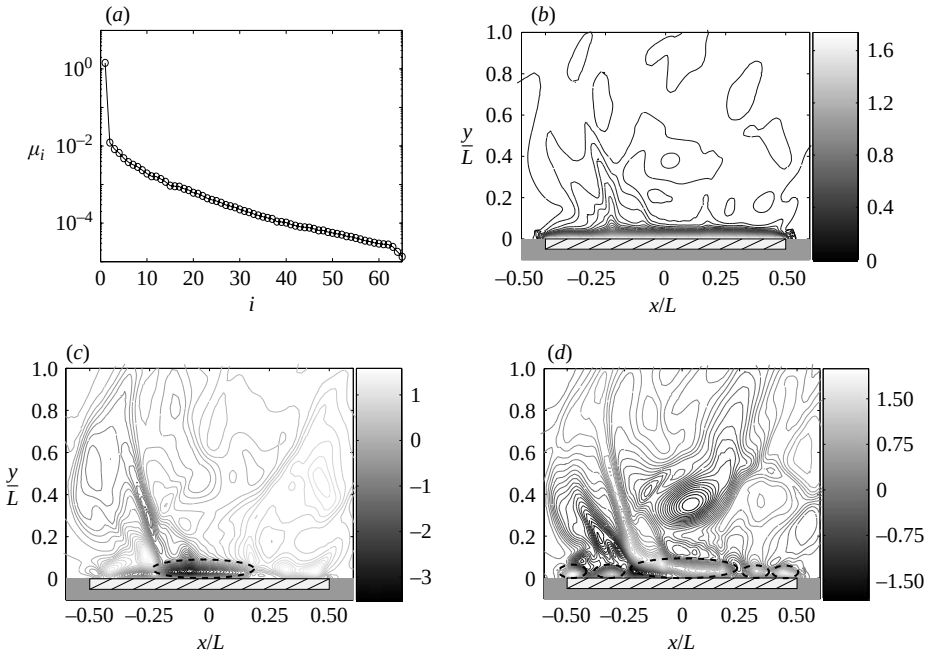

FIGURE 22. (*a*) Eigenspectrum and (*b–d*) isocontours of the first three temperature POD eigenmodes Φ_1 (*b*), Φ_2 (*c*) and Φ_3 (*d*).

versus the dimensionless $x^* = x/L$ for the snapshot shown in figure 20. The averaged Nusselt number

$$Nu_L = \frac{1}{|\tau|}\int_{\tau}\int_{-0.5}^{0.5} Nu(x^*, t^*)\,\mathrm{d}x^*\,\mathrm{d}t^* \tag{5.16}$$

where $|\tau|$ is the period of observation, is found to be $Nu_L = 26.36$, in very good agreement with the experiment-based correlations.

The POD analysis of the unperturbed temperature field is shown in figure 22. The first temperature mode carries the 96 % of the total energy. From the structure of the second and the third temperature POD modes shown in figure 22 (spatial modes) and figure 29 below in § 5.2.3 (temporal modes), we can draw conclusions on the main mechanism responsible for the heat transfer. The symmetric-like structure of these spatial modes and the periodic-like behaviour of the corresponding temporal modes (figure 29) implies that the heat flux at the plate surface is waveform-like in time and symmetric in space, 0.246 being the main dimensionless time frequency for both the aforementioned temperature modes. This frequency is in very good agreement with the main vortex shedding time frequency. This supports the conclusion that this temperature coherent structure is determined by advection. Therefore the most energetic temporal evolution of the heat transfer is determined by source zones highlightened with dashed lines in figure 22 moving on the plate surface symmetrically and periodically from the centre to the edges.

5.2.1. *Statistics for the perturbed energy levels*

At $Ra = 4.75 \times 10^6$ all the unperturbed temperature POD eigenstates are simple. Therefore there is no chance of having an eigenvalue splitting phenomenon and the theoretical formulae in § 3.1 characterize the statistics of each perturbed energy

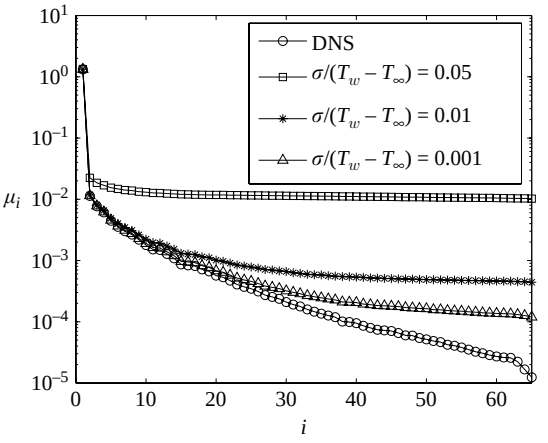

FIGURE 23. DNS-based temperature eigenspectrum and spectra of perturbed temperature fields corresponding to different $\sigma/(T_w - T_\infty)$.

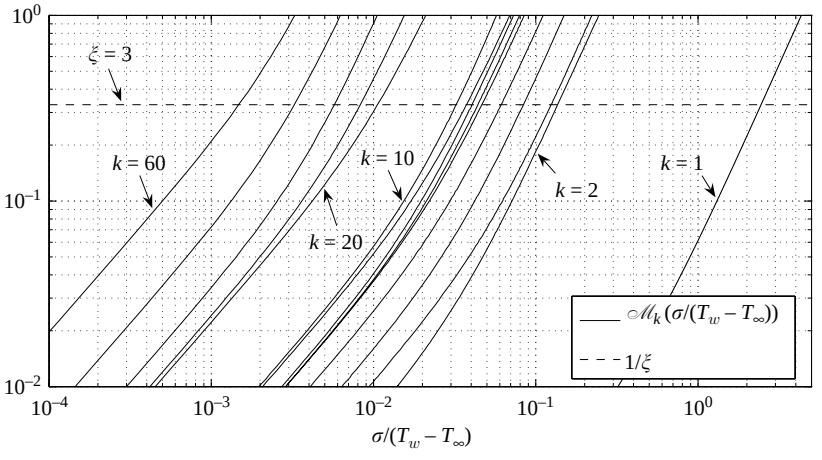

FIGURE 24. Resolvable eigenmodes as a function of the dimensionless temperature standard deviation $\sigma/(T_w - T_\infty)$.

level. As in § 5.1 we use the non-dimensional version of the perturbation theory (3.4) where the dimensionless perturbation parameter for the temperature field $\sigma/(T_w - T_\infty)$ is scaled by the reference temperature difference $T_w - T_\infty$. We have simulated random disturbances $\delta T(\boldsymbol{x}_i, t_l)$ according to the probability density (2.16) at different levels $\sigma/(T_w - T_\infty)$ and we constructed an ensemble of 40 000 independent possible realizations of the perturbed temperature field at each level. The effects of such perturbations on the unperturbed (DNS-based) POD eigenspectrum are shown in figure 23. In figure 24 we plot the resolution indicator functions (3.18). The small separation between the unperturbed energy levels makes it difficult to characterize clearly the number of resolvable modes for given $\sigma/(T_w - T_\infty)$ (the resolution indicator functions are clustered). For instance, if we assume for simplicity that the probability density for the perturbed energy levels is Gaussian then from figure 24 we see that correspondingly to $\sigma/(T_w - T_\infty) = 0.01$ it is possible to resolve 20 modes with probability 99.89 % ($\xi = 3.26$); 30 modes with probability 96.92 %

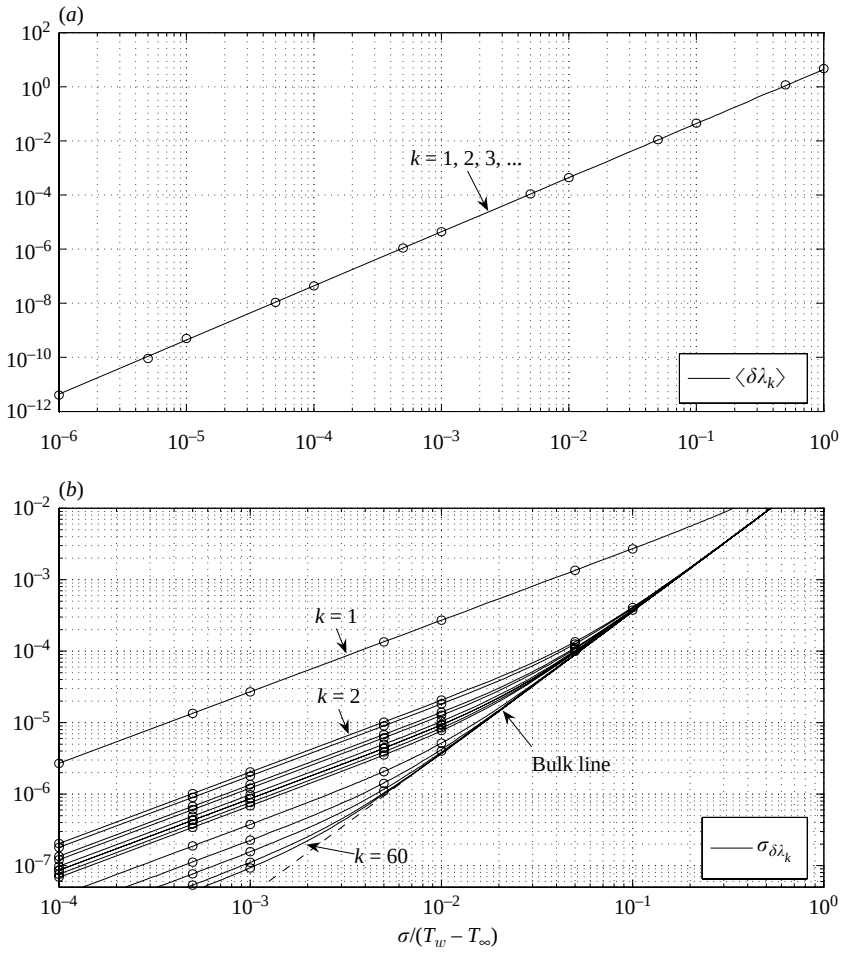

FIGURE 25. (*a*) Expected value and (*b*) standard deviation for all the perturbed energy levels as a function of the dimensionless temperature standard deviation $\sigma/(T_w - T_\infty)$. Comparison between the theoretical formulae (3.9), (3.13) (continuous lines) and the results (circles) from Monte Carlo simulations (40 000 samples).

($\xi = 2.16$); 40 modes with probability 71.99 % ($\xi = 1.08$). Note also that the most energetic time-dependent dynamic (ψ_2) will be almost completely over shadowed by the random fluctuations for $\sigma/(T_w - T_\infty) \geqslant 0.35$ (figure 24). Under the hypothesis of Gaussian density the resolution probability for the second mode corresponding to such a perturbation magnitude will be approximately 38.29 % ($\xi = 0.5$).

5.2.2. *Averaged energies and standard deviations*

In figure 25 we compare the predictions of the theoretical formulae (3.9) and (3.13) to the ensemble means and ensemble standard deviations obtained by the Monte Carlo method described in § 5.2.1. It is shown that the standard deviations of the perturbed energy levels tend asymptotically to the same line (bulk line) as the random temperature dimensionless standard deviation increases. This phenomenon is

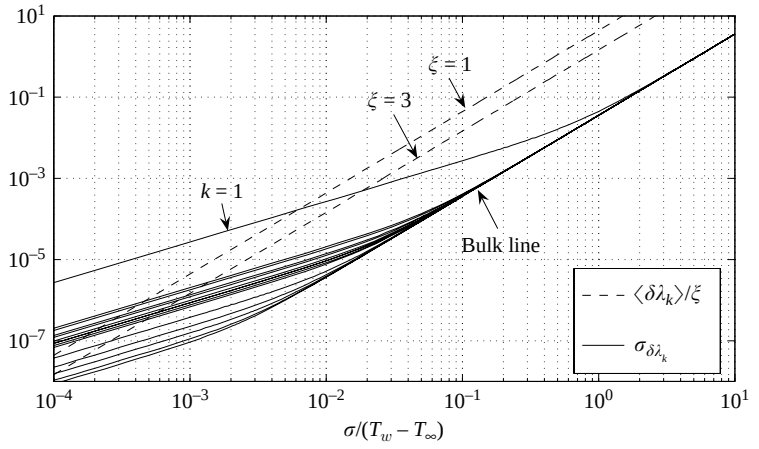

FIGURE 26. Expected value and standard deviation for all the perturbed energy levels.

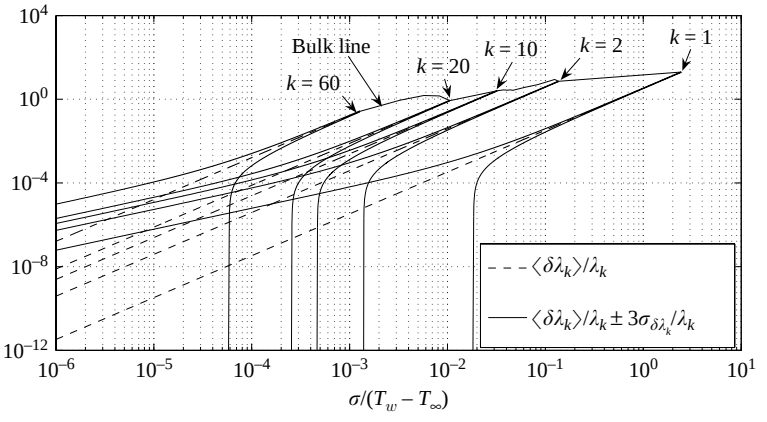

FIGURE 27. Sensitivity of energy levels to perturbations.

still represented by an equation similar to (5.2):

$$\mathscr{B} = \sqrt{c}\left(\frac{\sigma}{T_w - T_\infty}\right)^2, \quad c = 0.00127. \tag{5.17}$$

We note that $c \simeq 10^{-3}\lambda_1$. As for the cylinder flow discussed in § 5.1 all the expected perturbed energy levels turn out to be increased by approximately the same quantity $(\sigma/(T_w - T_\infty))^2 \Delta t |\Omega| d$. In figure 26 we summarize the relevant statistics for all the perturbed energy levels. With high probability the perturbed modes represented by the continuous lines inside the strip between the bulk line and the line $\langle \delta\lambda_k \rangle/3$ have an energy which is greater than the corresponding unperturbed energy. The sensitivity of the perturbed energy levels is studied in figure 27 where it is shown that the maximum perturbation supportable by each mode (at level $\xi = 3$ of resolution probability) depends on $\sigma/(T_w - T_\infty)$. The ratio $\langle \delta\lambda_k \rangle/\lambda_k$ ranges from 0.26 (60th mode) to 19.82 (1st mode). This means, for instance, that the energy level of the perturbed mean field will be indistinguishable from the energy of the second mode if the expected perturbation energy is approximately 20 times the unperturbed energy

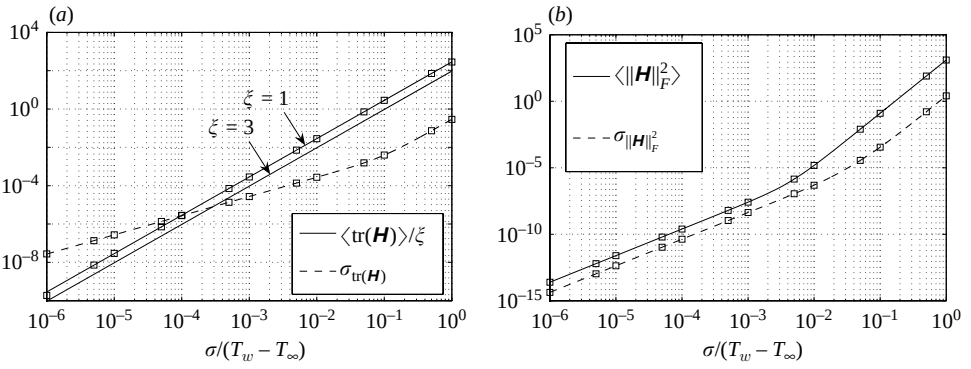

FIGURE 28. Statistics for (*a*) the perturbation energy and (*b*) the Frobenius norm, of the perturbation operator $\boldsymbol{H}$ as function of the dimensionless perturbation parameter $\sigma/(T_w - T_\infty)$. The squares are the results from Monte Carlo simulations (40 000 samples).

level. Figure 28 shows the relevant statistics for the total energy of the perturbation as well as the norm of the perturbation operator $\boldsymbol{H}$. With high probability ($\xi = 3$) the total energy of the perturbed field becomes greater than the unperturbed one for $\sigma/(T_w - T_\infty) > 3.34 \times 10^{-4}$.

5.2.3. *Random temporal modes*

All the unperturbed POD eigenstates are simple. Therefore the perturbation expansions in § 3.4 characterize exactly the mean and the standard deviation of each perturbed temporal mode. In figure 29 we compare the Monte-Carlo-based ensemble mean and ensemble standard deviation to the theoretical predictions of equations (3.27) and (3.29) for several random temporal modes. The comparison is made for $\sigma/(T_w - T_\infty) = 0.01$. A very good agreement is found up to the 30th mode which, under the hypothesis of Gaussian density, is resolvable with probability 96.92 %. Note that the normalization of the expected values and the standard deviations shown in figure 29 is needed for the comparison because the transformation (3.23) is not unitary. Finally in figure 30 we plot the expected value as well as a confidence interval for the 14th and 30th temporal modes.

6. Conclusions

We have developed a stochastic perturbation theory for the POD spatio-temporal scales of a discrete random flow under the assumption that the perturbation of the deterministic state is a Gaussian uncorrelated random field. We have shown that the relevant statistics for the perturbed energy levels and the perturbed temporal modes can be expressed in an explicit form as power series in the random-flow standard deviation (§ 3.1, § 3.4). Such expansions provide connections between the random elements of the flow and the uncertainty in its set of random spatio-temporal scales.

We applied the theory developed to the flow past a cylinder at $Re = 100$ (§ 5.1) and to the natural convective flow over an isothermal horizontal plate at $Ra = 4.75 \times 10^6$ (§ 5.2). Both of these studies show the existence of a limiting lower level of uncertainty in the energy of the random POD modes for which the modes themselves are resolvable. Interestingly this limiting value, characterized in terms of resolution indicator functions (3.18) and bulk lines (5.2), can be much lower than the uncertainty in the flow σ/u. The slope of the bulk line seems to be strongly correlated to the

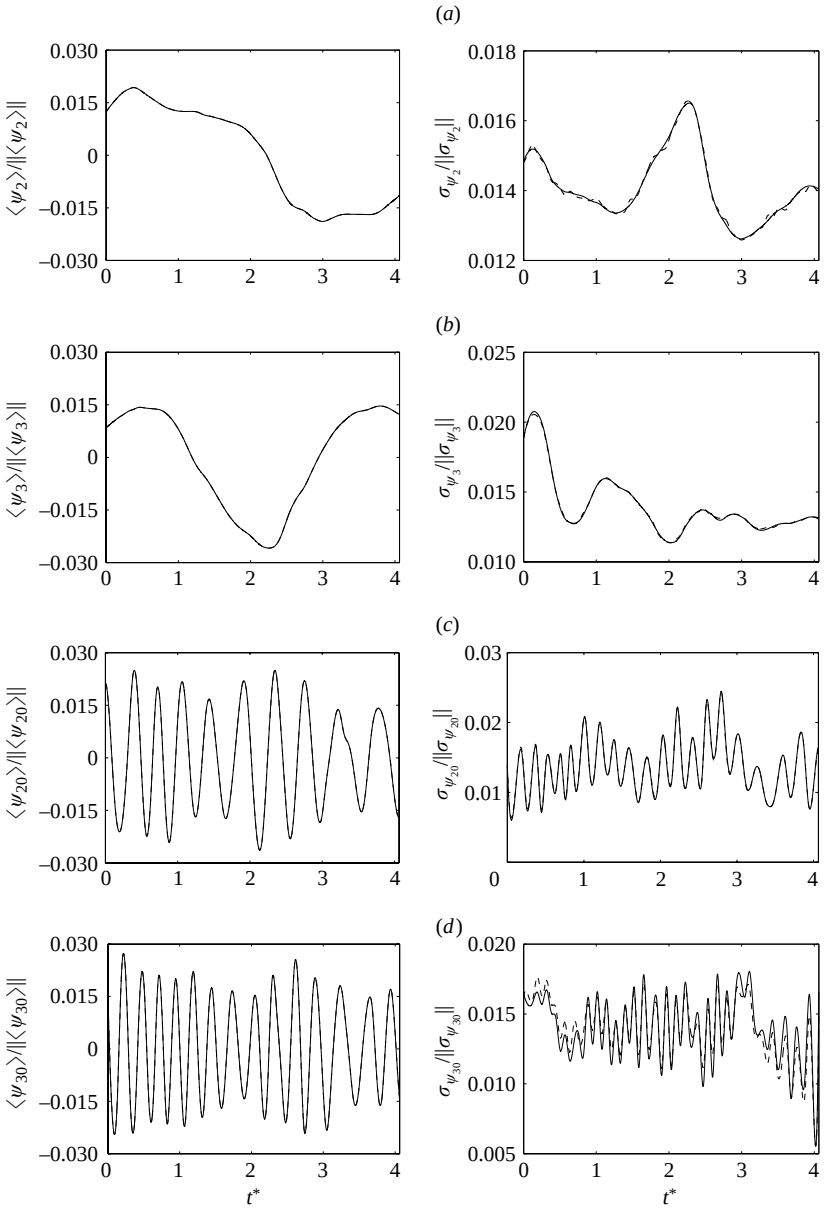

FIGURE 29. Random temporal modes corresponding to $\sigma/(T_w - T_\infty) = 0.01$. Comparison between the Monte Carlo simulations (——) and the theoretical predictions (—) for (*a*) the second, (*b*) the third, (*c*) the twentieth and (*d*) the thirtieth mode. Left: normalized expected value; right: normalized standard deviation.

largest unperturbed eigenvalue λ_1 ($c \simeq 10^{-3}\lambda_1$ in (5.2) and (5.17)) for both the flows discussed in this paper. The standard deviations of the random energy levels tend asymptotically to this line as the flow uncertainty increases. This means that mixing of different random flow scales occurs when the corresponding energy levels have the same standard deviations. However the perturbation analysis of the fluctuating velocity field for the cylinder wake (§ 5.1.7) shows an exception to this rule (figure 16).

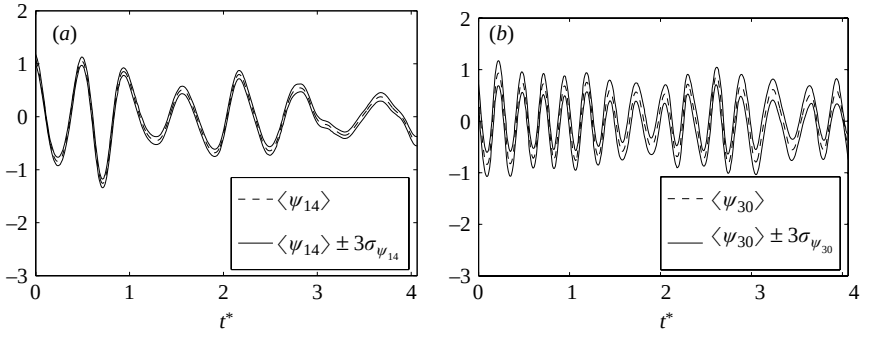

FIGURE 30. $\sigma/(T_w - T_\infty) = 0.01$. Relevant statistics for (*a*) the 14th and (*b*) the 30th random modes.

All the averaged perturbations in the energy levels are positive and exhibit a very weak dependence on the particular mode, increasing them by approximately the same quantity.

The expansions (3.27) and (3.29) for the statistics of random temporal modes turn out also to be accurate for the modes near the bulk spectrum as demonstrated in figures 12 and 29.

The specific new contributions of the present work are:

the development of a stochastic perturbation theory for the POD eigenmodes;

the characterization of the symmetry breaking between the resolvable part of the temporal and the spatial dynamic for randomly perturbed flows;

the characterization of the number of POD resolvable modes in terms of resolution indicator functions (3.18) and bulk lines (5.2);

the determination of the relevant statistics for the POD scales of the perturbed flow past a cylinder and the perturbed natural convective flow over an isothermal horizontal plate.

These results not only provide a better understanding of POD decompositions of numerical and experimental data but also, and more importantly, include the mechanism by which the uncertainty itself propagates from the flow field to its spatio-temporal modes. Our approach can also be applied to more general types of random perturbations $\delta\boldsymbol{u}$ (by a suitable recalculation of the expectation relations given in Appendix B).

The issue of stochastic POD low-dimensional modeling is still an open question. In particular a crucial point is how the randomness propagates in Galerkin models or modal energy flow analyses (Noack *et al.* 2005). In principle the expansion for the spatial modes (4.3) can be used to define random POD projectors. However this straightforward approach has the inconvenience that the corresponding temporal evolution will be random as well, thus not suitable for a system of ordinary differential equations. In addition a formula like (4.3) is valid under the hypothesis of the absence of eigenvalue splitting. A more promising approach for *stochastic* low-dimensional modelling lies within the possibility of obtaining a polynomial chaos expansion (Xiu & Karniadakis 2003; Wan & Karniadakis 2005*b*) for the POD spatial modes, given a chaos expansion for the random flow $\widetilde{\boldsymbol{u}}(\boldsymbol{x}, t)$. (By a suitable redefinition of the biorthogonal decomposition (2.1) and the inner product (2.2) it is possible to represent a random flow field in terms of random (weakly orthogonal) spatial modes and deterministic temporal modes.) A Galerkin projection onto these random modes defines, after a suitable averaging operation, a *deterministic* temporal evolution. The

reduced-order model of the random flow will then be some sort of deterministic temporal evolution of random spatial structures. This allows a rigorous quantification of the propagation of randomness in stochastic POD low-dimensional models of random flows.

Appendix A. Finite-dimensional representations

In this Appendix we represent the perturbation problem formulated in § 2 in a suitable finite-dimensional setting. We assume that the spatial domain Ω and the temporal domain T are partitioned into N spatial points and S time instants respectively. We approximate the integrals over the spatial domain by a weighted formula (Karniadakis & Sherwin 1999) in the form

$$\int_{\Omega} u^{\alpha}(\boldsymbol{x}, t_l)\,\mathrm{d}\boldsymbol{x} \simeq \sum_{j=1}^{N} u^{\alpha}(\boldsymbol{x}_j, t_l) w_j, \qquad \alpha = 1, \ldots, d, \tag{A 1}$$

where N is the total number of grid points and u^{α} denotes the αth component of the vector flow field. We assume that we have available S equidistant snapshots $\{\boldsymbol{u}(\boldsymbol{x}, t_l)\}_{l=1,\ldots,S}$ of the flow field $\boldsymbol{u}$, each separated by Δt. For the integration in time we use a first-order approximation

$$\int_{T} u^{\alpha}(\boldsymbol{x}_i, t)\,\mathrm{d}t \simeq \Delta t \sum_{l=1}^{S} u^{\alpha}(\boldsymbol{x}_i, t_l), \qquad \alpha = 1, \ldots, d. \tag{A 2}$$

Moreover we set

$$\widetilde{\boldsymbol{u}}_{jl} \equiv \widetilde{\boldsymbol{u}}(\boldsymbol{x}_j, t_l), \quad \boldsymbol{u}_{jl} \equiv \boldsymbol{u}(\boldsymbol{x}_j, t_l), \quad \delta\boldsymbol{u}_{jl} \equiv \delta\boldsymbol{u}(\boldsymbol{x}_j, t_l). \tag{A 3}$$

The discrete analogue of (2.9) is

$$\widetilde{\boldsymbol{u}}_{jl} = \boldsymbol{u}_{jl} + \delta\boldsymbol{u}_{jl}. \tag{A 4}$$

We assume that the random variables $\{\delta u^{\alpha}_{il}\}$ are zero-mean jointly Gaussian and uncorrelated, with the same variance σ^2, i.e. the joint probability density function of $\{\delta u^{\alpha}_{il}\}$ has the form

$$\pi\big(\delta u^1_{11}, \ldots, \delta u^d_{NS}; \sigma\big) = K \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{N} \sum_{l=1}^{S} \sum_{\alpha=1}^{d} \big(\delta u^{\alpha}_{il}\big)^2\right], \tag{A 5}$$

where K is a suitable normalization constant. Therefore $\delta\boldsymbol{u}_{jl}$ is a discrete Gaussian uncorrelated random field. It is well know that a set of zero mean Gaussian uncorrelated random variables is also necessarily independent. We consider here the case where all the variances σ^2 are the same, independently of the time, the location and the vector component. One can also generalize the theory to the case of different variances in each component of $\widetilde{\boldsymbol{u}}$ or, more generally, by defining a function $\sigma^2(\boldsymbol{x}_k, t_m)$ as Hachem *et al.* (2004) have recently done. However, their results are limited to the Stieltjes transform of the normalized eigenvalue counting measure when the dimension of the problem goes to infinity.

The finite-dimensional representation of the unperturbed temporal autocorrelation (2.7*b*) is

$$\mathscr{T}_{lm} \equiv \mathscr{T}(t_l, t_m) = \sum_{i=1}^{N} \sum_{\alpha=1}^{d} u^{\alpha}_{il} u^{\alpha}_{im} w_i. \tag{A 6}$$

Also, the finite-dimensional representations of the operators $\boldsymbol{U}\boldsymbol{U}^{\dagger}$, $\boldsymbol{H}^{(1)}$ and $\boldsymbol{H}^{(2)}$ (see (2.12)) are

$$[\boldsymbol{U}\boldsymbol{U}^{\dagger}]_{lm} = \Delta t \sum_{\alpha=1}^{d} \sum_{i=1}^{N} u_{il}^{\alpha} u_{im}^{\alpha} w_i, \tag{A 7}$$

$$H_{lm}^{(1)} = \Delta t \sum_{\alpha=1}^{d} \sum_{i=1}^{N} \left(u_{il}^{\alpha} \delta u_{im}^{\alpha} + u_{im}^{\alpha} \delta u_{il}^{\alpha} \right) w_i, \tag{A 8}$$

$$H_{lm}^{(2)} = \Delta t \sum_{\alpha=1}^{d} \sum_{i=1}^{N} \delta u_{il}^{\alpha} \delta u_{im}^{\alpha} w_i. \tag{A 9}$$

Everson & Roberts (2000), Zwald, Bousquet & Blanchard (2004) and Taylor *et al.* (2002) analysed the relation between the correlation operator's eigenspectrum and the eigenspectrum of its finite-dimensional representation. It is useful to define other matrices and related quantities. Let

$$\mathscr{T}_{lm}^{(z)} := \sum_{i=1}^{N} \sum_{\alpha=1}^{d} u_{il}^{\alpha} u_{im}^{\alpha} (w_i)^z, \quad z \in \mathbb{N}. \tag{A 10}$$

We note that $\mathscr{T}_{lm}^{(1)}$ equals the unperturbed temporal autocorrelation kernel (A 6). We also define

$$\Theta_z := \operatorname{tr}\left(\mathscr{T}^{(z)}\right) = \sum_{l=1}^{S} \sum_{i=1}^{N} \sum_{\alpha=1}^{d} \left(u_{il}^{\alpha}\right)^2 (w_i)^z, \tag{A 11}$$

$$\left\| \mathscr{T}^{(z)} \right\|_F^2 := \sum_{l=1}^{S} \sum_{m=1}^{S} \left(\sum_{i=1}^{N} \sum_{\alpha=1}^{d} u_{il}^{\alpha} u_{im}^{\alpha} (w_i)^z \right)^2, \tag{A 12}$$

$$\Upsilon_z := \sum_{i=1}^{N} (w_i)^z. \tag{A 13}$$

Clearly we have

$$\Theta_1 = \frac{1}{\Delta t} \sum_{k=1}^{S} \mu_k, \qquad \left\| \mathscr{T}^{(1)} \right\|_F^2 = \frac{1}{\Delta t^2} \sum_{k=1}^{S} \mu_k^2, \qquad \Upsilon_0 = N, \qquad \Upsilon_1 = |\Omega|, \tag{A 14}$$

where $|\Omega|$ is the measure of the spatial domain; similarly $|T|$ is the measure of the temporal domain. We denote by $\{\psi_{j_k}(t)\}_{k=1,..,m_j}$ the m_j eigenfunctions of $\boldsymbol{U}\boldsymbol{U}^{\dagger}$ that span its jth eigenspace of dimension m_j. The projection matrix $\boldsymbol{P}_j$ onto such an eigenspace is

$$[\boldsymbol{P}_j]_{lm} := \Delta t \sum_{k=1}^{m_j} \psi_{j_k}(t_l) \psi_{j_k}(t_m), \quad l, m = 1, \ldots, S. \tag{A 15}$$

The reduced resolvent (Kato 1995) with respect to the eigenvalue λ_j (counted with its multiplicity) will be represented by the matrix

$$\boldsymbol{S}_j := \sum_{k \neq j} \frac{\boldsymbol{P}_k}{\lambda_k - \lambda_j}. \tag{A 16}$$

Appendix B. Expectation relations for $\boldsymbol{H}^{(1)}$ and $\boldsymbol{H}^{(2)}$ products

In this section we derive some important expectation relations which are widely used throughout the paper. First we note that if $\{\delta u_{il}^{\alpha}\}$ is a zero-mean discrete Gaussian uncorrelated random field with joint probability density (A 5) then the δu_{il}^{α} are necessarily independent. Therefore we can write

$$\langle \delta u_{il}^{\alpha} \rangle = 0, \tag{B 1a}$$

$$\langle \delta u_{il}^{\alpha} \delta u_{jm}^{\beta} \rangle = \sigma^2 \delta_{ij} \delta_{lm} \delta_{\alpha\beta}, \tag{B 1b}$$

$$\langle \delta u_{il}^{\alpha} \delta u_{jm}^{\beta} \delta u_{hn}^{\gamma} \rangle = 0, \tag{B 1c}$$

$$\begin{aligned} \langle \delta u_{il}^{\alpha} \delta u_{jm}^{\beta} \delta u_{hn}^{\gamma} \delta u_{kp}^{\eta} \rangle = \sigma^4 (&\delta_{ij}\delta_{hk}\delta_{lm}\delta_{np}\delta_{\alpha\beta}\delta_{\gamma\eta} + \delta_{ih}\delta_{jk}\delta_{ln}\delta_{mp}\delta_{\alpha\gamma}\delta_{\beta\eta} \\ &+ \delta_{ik}\delta_{jh}\delta_{lp}\delta_{mn}\delta_{\alpha\eta}\delta_{\beta\gamma}), \end{aligned} \tag{B 1d}$$

$$\langle \delta u_{il}^{\alpha} \delta u_{jm}^{\beta} \delta u_{hn}^{\gamma} \delta u_{kp}^{\eta} \delta u_{vr}^{\lambda} \rangle = 0, \tag{B 1e}$$

$$\vdots$$

where δ_{jk} is the Kronecker delta and $\langle \cdot \rangle$ denotes the expectation operator

$$\langle g \rangle := K \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g\left(\delta u_{11}^{1}, \dots, \delta u_{NS}^{d}\right) \exp\left[-\frac{1}{2\sigma^2} \sum_{i,l,\alpha} \left(\delta u_{il}^{\alpha}\right)^2\right] \prod_{i,l,\alpha} \mathrm{d}\left(\delta u_{il}^{\alpha}\right). \tag{B 2}$$

In equation (B 2) K is the normalization constant for the multivariate probability density function. Using the equations (B 1*a*)–(B 1*e*) it is easy to prove the following important expectation relations involving products of operators (A 8) and (A 9):

$$\langle H_{ln}^{(1)} \rangle = 0, \tag{B 3a}$$

$$\frac{1}{\sigma^2} \langle H_{ln}^{(1)} H_{pq}^{(1)} \rangle = \Delta t^2 \left(\mathscr{T}_{lp}^{(2)} \delta_{nq} + \mathscr{T}_{lq}^{(2)} \delta_{np} + \mathscr{T}_{np}^{(2)} \delta_{lq} + \mathscr{T}_{nq}^{(2)} \delta_{lp}\right), \tag{B 3b}$$

$$\langle H_{ln}^{(1)} H_{pq}^{(1)} H_{mr}^{(1)} \rangle = 0, \tag{B 3c}$$

$$\frac{1}{\sigma^2} \langle H_{ln}^{(2)} \rangle = \Delta t |\Omega| d \delta_{ln}, \tag{B 3d}$$

$$\langle H_{ln}^{(2)} H_{pq}^{(1)} \rangle = 0, \tag{B 3e}$$

$$\frac{1}{\sigma^4} \langle H_{ln}^{(2)} H_{pq}^{(2)} \rangle = \Delta t^2 \left(|\Omega|^2 d^2 \delta_{ln} \delta_{pq} + \Upsilon_2 d \left(\delta_{lp}\delta_{nq} + \delta_{lq}\delta_{np}\right)\right), \tag{B 3f}$$

$$\begin{aligned} \frac{1}{\sigma^4} \langle H_{ln}^{(2)} H_{pq}^{(1)} H_{mr}^{(1)} \rangle = {} & \Delta t^3 |\Omega| d \delta_{ln} \left(\mathscr{T}_{mp}^{(2)} \delta_{rq} + \mathscr{T}_{rp}^{(2)} \delta_{mq} + \mathscr{T}_{mq}^{(2)} \delta_{rp} + \mathscr{T}_{rq}^{(2)} \delta_{mp}\right) \\ & + \Delta t^3 \mathscr{T}_{mp}^{(2)} \left(\delta_{ql}\delta_{rn} + \delta_{qn}\delta_{rl}\right) + \Delta t^3 \mathscr{T}_{mq}^{(2)} \left(\delta_{pl}\delta_{rn} + \delta_{pn}\delta_{rl}\right) \\ & + \Delta t^3 \mathscr{T}_{rp}^{(2)} \left(\delta_{ql}\delta_{mn} + \delta_{qn}\delta_{ml}\right) + \Delta t^3 \mathscr{T}_{rq}^{(2)} \left(\delta_{pl}\delta_{mn} + \delta_{pn}\delta_{ml}\right), \end{aligned}$$

$$\vdots \tag{B 3g}$$

where $\mathscr{T}_{lm}^{(2)}$ and Υ_2 are defined in (A 10) and (A 13) respectively. We observe that if the sum of the superscripts of the coefficients inside the expectations is odd then the expected value is zero. This result will be used very often for the computation of the relevant statistics of the perturbed energy levels and temporal modes.

Appendix C. The expected operator $\langle \boldsymbol{W}_j^{(2)} \rangle$

In this Appendix we show how to compute the expected value of the second-order perturbation operator defined in (3.24*b*):

$$\begin{aligned}\langle \boldsymbol{W}_j^{(2)} \rangle = &-\boldsymbol{S}_j \langle \widehat{\boldsymbol{H}}^{(2)} \rangle \boldsymbol{P}_j + \boldsymbol{S}_j \langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(1)} \rangle \boldsymbol{P}_j \\ &- \boldsymbol{S}_j^2 \langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j \widehat{\boldsymbol{H}}^{(1)} \rangle \boldsymbol{P}_j - \tfrac{1}{2} \boldsymbol{P}_j \langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j^2 \widehat{\boldsymbol{H}}^{(1)} \rangle \boldsymbol{P}_j . \end{aligned} \tag{C 1}$$

Using (B 3*a*)–(B 3*d*) we find

$$\langle \widehat{\boldsymbol{H}}^{(2)} \rangle = d \Delta t |\Omega| \boldsymbol{S}_j \boldsymbol{P}_j , \tag{C 2a}$$

$$\langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j \widehat{\boldsymbol{H}}^{(1)} \rangle = \Delta t^2 \big(\mathscr{T}^{(2)} \boldsymbol{S}_j + \mathrm{tr}\,(\boldsymbol{S}_j) \mathscr{T}^{(2)} + \mathrm{tr}\,\big(\mathscr{T}^{(2)} \boldsymbol{S}_j \big) \boldsymbol{I} + \boldsymbol{S}_j \mathscr{T}^{(2)} \big), \tag{C 2b}$$

$$\langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{P}_j \widehat{\boldsymbol{H}}^{(1)} \rangle = \Delta t^2 \big(\mathscr{T}^{(2)} \boldsymbol{P}_j + \mathrm{tr}\,(\boldsymbol{P}_j) \mathscr{T}^{(2)} + \mathrm{tr}\,\big(\mathscr{T}^{(2)} \boldsymbol{P}_j \big) \boldsymbol{I} + \boldsymbol{P}_j \mathscr{T}^{(2)} \big), \tag{C 2c}$$

$$\langle \widehat{\boldsymbol{H}}^{(1)} \boldsymbol{S}_j^2 \widehat{\boldsymbol{H}}^{(1)} \rangle = \Delta t^2 \big(\mathscr{T}^{(2)} \boldsymbol{S}_j^2 + \mathrm{tr}\,\big(\boldsymbol{S}_j^2 \big) \mathscr{T}^{(2)} + \mathrm{tr}\,\big(\mathscr{T}^{(2)} \boldsymbol{S}_j^2 \big) \boldsymbol{I} + \boldsymbol{S}_j^2 \mathscr{T}^{(2)} \big) . \tag{C 2d}$$

Finally, substituting (C 2*a*)–(C 2*d*) into (C 1) and taking into account the well-known identities $\boldsymbol{S}_j \boldsymbol{P}_j = \boldsymbol{P}_j \boldsymbol{S}_j = 0$, $\boldsymbol{P}_j^2 = \boldsymbol{P}_j$ we get

$$\langle \boldsymbol{W}_j^{(2)} \rangle = \Delta t^2 \big(\big[\mathrm{tr}\,(\boldsymbol{S}_j) \boldsymbol{S}_j - \mathrm{tr}\,\big(\boldsymbol{P}_j \big) \boldsymbol{S}_j^2 - \tfrac{1}{2} \mathrm{tr}\,\big(\boldsymbol{S}_j^2 \big) \boldsymbol{P}_j \big] \mathscr{T}^{(2)} - \tfrac{1}{2} \mathrm{tr}\,\big(\mathscr{T}^{(2)} \boldsymbol{S}_j^2 \big) \boldsymbol{I} \big). \tag{C 3}$$

Appendix D. Computation of the expansion for spatial modes

Let us consider equation (4.2). We expand $1/\sqrt{\widetilde{\lambda}_j(\sigma)}$ in a Laurent series near $\sigma = 0$,

$$\begin{aligned}\frac{1}{\sqrt{\widetilde{\lambda}_j(\sigma)}} &= \frac{1}{\sqrt{\lambda_j + \sigma \widehat{\lambda}_j^{(1)} + \sigma^2 \widehat{\lambda}_j^{(2)} \ldots}} \\ &= \frac{1}{\sqrt{\lambda_j}} - \frac{1}{2} \sigma \frac{\widehat{\lambda}_j^{(1)}}{\lambda_j \sqrt{\lambda_j}} + \sigma^2 \left(\frac{3}{4} \frac{\big(\widehat{\lambda}_j^{(1)} \big)^2}{\lambda_j^2 \sqrt{\lambda_j}} - \frac{\widehat{\lambda}_j^{(2)}}{\lambda_j \sqrt{\lambda_j}} \right) + \cdots . \end{aligned} \tag{D 1}$$

The substitution of equations (D 1) and (3.26) into (4.2) gives

$$\begin{aligned}\widetilde{\boldsymbol{\Phi}}_{j_k}(\boldsymbol{x}, \sigma) &= \left(\frac{1}{\sqrt{\lambda_j}} - \frac{1}{2} \sigma \frac{\widehat{\lambda}_j^{(1)}}{\lambda_j \sqrt{\lambda_j}} + \ldots \right) \big(\boldsymbol{U}^\dagger + \sigma \delta \boldsymbol{U}^\dagger \big) \big(\psi_{j_k} + \sigma \boldsymbol{W}_j^{(1)} \psi_{j_k} + \cdots \big) \\ &= \frac{1}{\sqrt{\lambda_j}} \boldsymbol{U}^\dagger \psi_{j_k} + \sigma \left(\frac{\boldsymbol{U}^\dagger \boldsymbol{W}_j^{(1)}}{\sqrt{\lambda_j}} + \frac{\delta \boldsymbol{U}^\dagger}{\sqrt{\lambda_j}} - \frac{1}{2} \frac{\widehat{\lambda}_j^{(1)}}{\lambda_j \sqrt{\lambda_j}} \boldsymbol{U}^\dagger \right) \psi_{j_k} + \cdots \end{aligned} \tag{D 2}$$

where we have defined

$$\delta \boldsymbol{U}^\dagger \phi := \int_T \delta \widehat{\boldsymbol{u}}(\boldsymbol{x}, t) \phi(t) \, \mathrm{d}t . \tag{D 3}$$

Finally, using the dispersion relation (2.5*b*) we obtain

$$\begin{aligned}\widetilde{\boldsymbol{\Phi}}_{j_k}(\boldsymbol{x}, \sigma) &= \boldsymbol{\Phi}_{j_k}(\boldsymbol{x}) + \sigma \left(\frac{\boldsymbol{U}^\dagger \boldsymbol{W}_j^{(1)} \boldsymbol{U}}{\lambda_j} + \frac{\delta \boldsymbol{U}^\dagger \boldsymbol{U}}{\lambda_j} - \frac{1}{2} \frac{\widehat{\lambda}_j^{(1)}}{\lambda_j^2} \boldsymbol{U}^\dagger \boldsymbol{U} \right) \boldsymbol{\Phi}_{j_k} + \cdots \\ &= \boldsymbol{\Phi}_{j_k}(\boldsymbol{x}) + \sigma \left(\frac{\boldsymbol{U}^\dagger \boldsymbol{W}_j^{(1)} \boldsymbol{U}}{\lambda_j} + \frac{\delta \boldsymbol{U}^\dagger \boldsymbol{U}}{\lambda_j} - \frac{1}{2} \frac{\widehat{\lambda}_j^{(1)}}{\lambda_j} \boldsymbol{I} \right) \boldsymbol{\Phi}_{j_k} + \cdots , \end{aligned} \tag{D 4}$$

The definition (4.4) follows from (D 4).

Appendix E. Statistics for the norm of the perturbation operator $\boldsymbol{H}$

The integral operators defined in (2.6*a*), (2.6*b*), (2.13) and (2.14) are in the Hilbert–Schmidt class (Kato 1995; Riesz & Sz-Nagy 1953). Therefore they can be considered as vectors in the space that has with the inner product

$$(\boldsymbol{A}, \boldsymbol{B})_F = \sum_j (\boldsymbol{A}\phi_j, \boldsymbol{B}\phi_j) = \int_T \int_T \mathscr{A}(t,t')\mathscr{B}(t,t')\,\mathrm{d}t\,\mathrm{d}t'. \tag{E 1}$$

The Hilbert–Schmidt (or Frobenius) norm $\|\boldsymbol{H}\|_F := \sqrt{(\boldsymbol{H},\boldsymbol{H})_F}$ of the perturbation operator $\boldsymbol{H}$ is a unitary-invariant norm which takes into account a sort of mixture effect of the perturbation. If we consider the spectral representations (Kato 1995) of the correlation operators (2.14) and (2.6*b*)

$$\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger = \sum_k \widetilde{\lambda}_k \widetilde{\boldsymbol{P}}_k, \quad \boldsymbol{U}\boldsymbol{U}^\dagger = \sum_k \lambda_k \boldsymbol{P}_k, \tag{E 2a,b}$$

where $\boldsymbol{P}_k$, $\widetilde{\boldsymbol{P}}_k$ are spectral projectors of $\boldsymbol{U}\boldsymbol{U}^\dagger$ and $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger$ respectively then we have

$$\left\| \sum_k \widetilde{\lambda}_k \widetilde{\boldsymbol{P}}_k - \sum_j \lambda_j \boldsymbol{P}_j \right\|_F^2 = \|\boldsymbol{H}\|_F^2. \tag{E 3}$$

Other information can be obtained by considering the operators $\widetilde{\boldsymbol{U}}\widetilde{\boldsymbol{U}}^\dagger$ and $\boldsymbol{U}\boldsymbol{U}^\dagger$ (and their spectral projectors) as vectors in the space that has with the inner product (E 1). The squared norm of the symmetric perturbation $\boldsymbol{H}$ is

$$\|\boldsymbol{H}\|_F^2 = \mathrm{tr}\,(\boldsymbol{H}^2) = \int_T \int_T \mathscr{H}(t,t')^2\,\mathrm{d}t\,\mathrm{d}t'. \tag{E 4}$$

From (E 4) we see that the expected value of $\|\boldsymbol{H}\|_F^2$ is

$$\left\langle \left\|\boldsymbol{H}\right\|_F^2 \right\rangle = \int_T \int_T \langle \mathscr{H}(t,t')^2 \rangle\,\mathrm{d}t\,\mathrm{d}t'. \tag{E 5}$$

Using (2.15), (B 3*b*), (B 3*e*) and (B 3*f*) we find

$$\left\langle \left\|\boldsymbol{H}\right\|_F^2 \right\rangle = \sigma^2 \Delta t^2 ((S+1)(2\Theta_2 + d\sigma^2 S \Upsilon_2) + d^2\sigma^2 S |\Omega|^2). \tag{E 6}$$

The second-order moment of $\|\boldsymbol{H}\|_F^2$ is defined by

$$\left\langle \left\|\boldsymbol{H}\right\|_F^4 \right\rangle = \int_T \int_T \int_T \int_T \langle \mathscr{H}(t,t')^2 \mathscr{H}(s,s')^2 \rangle\,\mathrm{d}t\,\mathrm{d}t'\mathrm{d}s\,\mathrm{d}s'. \tag{E 7}$$

After some computations the following expression for the standard deviation of $\|\boldsymbol{H}\|_F^2$ is obtained:

$$\begin{aligned}\sigma_{\|\boldsymbol{H}\|_F^2} = \big[&8\sigma^4\Delta t^4\big(\Theta_2^2 + (S+2)\|\mathscr{T}^{(2)}\|_F^2\big) + 4\sigma^6\Delta t^4(1+S)d\Upsilon_2(4\Theta_2 + Sd\sigma^2\Upsilon_2) \\ &+ 8\sigma^6\Delta t^4 d^2|\Omega|^2(2\Theta_2 + Sd\sigma^2\Upsilon_2) + 16\sigma^6\Delta t^4|\Omega|d(1+S)(3\Theta_3 + Sd\sigma^2\Upsilon_3) \\ &+ 4\sigma^6\Delta t^4(5+5S+2S^2)(4\Theta_4 + Sd\sigma^2\Upsilon_4)\big]^{1/2}.\end{aligned} \tag{E 8}$$

Equations (E 6) and (E 8) are exact for random perturbations (2.16) of arbitrary 'magnitude' σ.

Appendix F. The autocorrelation matrix for the random temporal modes

Let ψ be a temporal mode belonging to the jth eigenspace of $\boldsymbol{U}\boldsymbol{U}^{\dagger}$. The autocorrelation matrix for $\widetilde{\psi}$ can be computed using Kato's transformation (3.23) as follows:

$$\begin{aligned}\langle\widetilde{\psi}(t_n)\widetilde{\psi}(t_m)\rangle &= \langle\big(\psi_n+\sigma\big(\boldsymbol{W}_j^{(1)}\psi\big)_n+\dots\big)\big(\psi_m+\sigma\big(\boldsymbol{W}_j^{(1)}\psi\big)_m+\dots\big)\rangle\\ &= \psi_n\psi_m+\sigma^2\langle\big(\boldsymbol{W}_j^{(1)}\psi\big)_n\big(\boldsymbol{W}_j^{(1)}\psi\big)_m\rangle+\sigma^2\psi_n\big(\langle\boldsymbol{W}_j^{(2)}\rangle\psi\big)_m\\ &\quad+\sigma^2\psi_m\big(\langle\boldsymbol{W}_j^{(2)}\rangle\psi\big)_n+\dots,\end{aligned}\tag{F 1}$$

where $\psi_n\equiv\psi(t_n)$, $\langle\boldsymbol{W}_j^{(2)}\rangle$ is given in (C 3) and

$$\langle\big(\boldsymbol{W}_j^{(1)}\psi\big)_n\big(\boldsymbol{W}_j^{(1)}\psi\big)_m\rangle=\Delta t\big[\big[\boldsymbol{S}_j\mathscr{T}^{(2)}\boldsymbol{S}_j\big]_{nm}+\big[\boldsymbol{S}_j^2\big]_{nm}\big(\mathscr{T}^{(2)}\psi,\psi\big)_T\big].\tag{F 2}$$

The symmetry with respect n and m implies that $\langle\widetilde{\psi}_n\widetilde{\psi}_m\rangle$ is represented by a symmetric matrix.

Appendix G. A simple estimation of the number of samples required to resolve the means of the perturbed eigenvalues

In this Appendix we use a basic result in probability theory to estimate the number of independent realizations of the pertubed eigenvalues required to resolve their statistical means with a certain confidence. We assume that all the possible determinations of $\delta\lambda_j$ are bounded certainly by $\pm3\sigma_{\widetilde{\lambda}_j}$ and we denote by $\mathscr{S}_{\widetilde{\lambda}_j}$ the number of independent samples of $\widetilde{\lambda}_j$. Using the Hoeffding inequality (Hoeffding 1963)

$$Pr\left\{\left|\frac{1}{\mathscr{S}_{\widetilde{\lambda}_j}}\sum_{k=1}^{\mathscr{S}_{\widetilde{\lambda}_j}}\delta\lambda_j^{(k)}-\langle\delta\lambda_j\rangle\right|\geqslant\epsilon\right\}\leqslant\exp\big(-\mathscr{S}_{\widetilde{\lambda}_j}\epsilon^2/\big(18\sigma_{\widetilde{\lambda}_j}\big)\big)\tag{G 1}$$

we conclude that in order to have the sample mean close to the decimal ($\epsilon=\langle\delta\lambda_j\rangle/10$) of the true mean with probability at least 90 %, we would need

$$\exp\big(-\mathscr{S}_{\widetilde{\lambda}_j}\epsilon^2/\big(18\sigma_{\widetilde{\lambda}_j}\big)\big)\leqslant0.1\quad\Rightarrow\quad\mathscr{S}_{\widetilde{\lambda}_j}\geqslant1800\ln(10)\frac{\sigma_{\widetilde{\lambda}_j}}{\langle\delta\lambda_j\rangle^2}\tag{G 2}$$

independent samples. For instance, if $\sigma/u=10^{-4}$ then (G 2) applied to the largest eigenvalue ($k=1$) of the wake POD (figure 8) gives $\mathscr{S}_{\widetilde{\lambda}_1}\gtrsim9\times10^{11}$. We remark that this simple estimation is not optimal and more refined bounds have been developed (see Lugosi 2004; Papoulis 1991; Bousquet 2002 and the references therein).

REFERENCES

Al-Arabi, M. & El-Riedy, M. 1976 Natural convection heat transfer from isothermal horizontal plates of different shapes. *Intl J. Heat Mass Transfer* **19**, 1399–1404.

Aubry, N. 1991 On the hidden beauty of the proper orthogonal decomposition. *Theoret. Comput. Fluid Dyn.* **2**, 339–352.

Aubry, N., Guyonnet, R. & Lima, R. 1991 Spatio-temporal analysis of complex signals: theory and applications. *J. Statist. Phys.* **64**, 683–739.

Aubry, N., Guyonnet, R. & Lima, R. 1995 Spatio-temporal symmetries and bifurcations via bi-orthogonal decomposition. *J. Nonlinear Sci.* **2**, 183–215.

Aubry, N. & Lima, R. 1995 Spatio-temporal and statistical symmetries. *J. Statist. Phys.* **81** (3/4), 793–828.

Bandrowski, J. & Rybski, W. 1976 Free convection mass transfer from horizontal plates. *Intl J. Heat Transfer* **19**, 827.

Bousquet, O. 2002 Concentration inequalities and empirical processes theory applied to the analysis of learning algorithms. PhD thesis, Ecole Polytechnique.

Cazemier, W., Verstappen, R. W. C. P. & Veldman, A. E. P. 1998 Proper orthogonal decomposition and low dimensional models for turbulent flows. *Phys. Fluids* **10**, 1685–1699.

Chambers, B. & Lee, T. Y. T. 1997 A numerical study of local and average natural convection Nusselt numbers for simultaneous convection above and below a uniformly heated horizontal thin plate. *J. Heat Transfer* **119**, 102–108.

Couplet, M., Sagaut, P. & Basdevant, C. 2003 Intermodal energy transfers in a proper orthogonal decomposition-Galerkin representation of a turbulent separated flow. *J. Fluid Mech.* **491**, 275–284.

Deane, A. E., Kevrekidis, I. G., Karniadakis, G. E. & Orszag, S. A. 1991 Low-dimensional models for complex geometry flows: application to grooved channels and circular cylinders. *Phys. Fluids* **3**, 2337–2354.

Delville, J., Ukeiley, L., Cordier, L., Bonnet, J. P. & Glauser, M. 2003 Examination of large-scale structures in a turbulent plane mixing layer. Part 1. Proper orthogonal decomposition. *J. Fluid Mech.* **497**, 335–363.

Dozier, R. B. & Silverstein, J. W. 2004 On the empirical distribution of eigenvalues of large dimensional information plus-noise type matrices. Manuscript, Mathematics Department, North Carolina State University.

Everson, R. M. & Roberts, S. J. 2000 Inferring the eigenvalues of covariance matrices from limited, noisy data. *IEEE Trans. Sig. Proc.* **48** (7), 2083–2091.

Fujii, T. & Imura, H. 1972 Natural convection heat transfer from a plate with arbitrary inclination. *Intl J. Heat Mass Transfer* **15**, 755–767.

Gad-El-Hak, M. 2000 *Flow Control*. Cambridge University Press.

Goldstein, R. J. & Kei-Shun, L. 1983 Laminar natural convection from a horizontal plate and the influence of plate edge extensions. *J. Fluid Mech.* **129**, 55–75.

Goldstein, R. J., Sparrow, E. M. & Jones, D. C. 1973 Natural convection mass transfer adjacent to horizontal plates. *Intl J. Heat Mass Transfer* **16**, 1025–1035.

Gordeyev, S. V. & Thomas, F. O. 2000 Coherent structure in the turbulent planar jet. Part 1. Extraction of proper orthogonal decomposition eigenmodes and their self-similarity. *J. Fluid Mech.* **414**, 145–194.

Hachem, W., Loubaton, P. & Najim, J. 2004 On the empirical distribution of eigenvalues of a Gram matrix with a given variance profile. *Ann. I'Institut Henry Poincavé, Probabilites et Statistiqués* (accepted).

Hachem, W., Loubaton, P. & Najim, J. 2005 The empirical eigenvalue distribution of a Gram matrix: from independence to stationarity. *Markov Processes and Related Fields* **11** (4).

Hassan, K. E. & Mohamed, S. 1970 Natural convection from isothermal flat surfaces. *Intl J. Heat Mass Transfer* **13**, 1873–1886.

Hoeffding, W. 1963 Probability inequalities for sums of bounded random variables. *J. Ame. Statist. Ass.* **58** (301), 13–30.

Holmes, P., Lumley, J. L. & Berkooz, G. 1996 *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press.

Hoyle, D. C. & Rattray, M. 2004 A statistical mechanics analysis of Gram matrix eigenvalue spectra. In *Learning Theory, 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004, Proce.* (ed. J. Shawe-Taylor & Y. Singer), pp. 579–593. Springer.

Johnstone, I. M. 2001 On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29**, 295–327.

Karniadakis, G. E. & Sherwin, S. 1999 *Spectral/hp Element Methods for CFD*. Oxford University Press.

Kato, T. 1950 On the adiabatic theorem of quantum mechanics. *J. Phys. Soc. Japan* **5**, 435–439.

Kato, T. 1995 *Perturbation Theory for Linear Operators*, 4th Edn. Springer.

Kimura, F., Yoshioka, T., Kitamura, K., Yamaguchi, M. & Asami, T. 2002 Fluid flow and heat transfer of natural convection at slightly inclined, upward-facing, heated plate. *Heat Transfer, Asian Research* **35** (5), 362–375.

Kitamura, K. & Kimura, F. 1995 Heat transfer and fluid flow of natural convection adjacent to upward-facing horizontal plates. *Intl J. Heat Mass Transfer* **38**, 3149–3159.

Liu, J. T. C. 1988 Contributions to the understanding of large-scale coherent structures in developing free turbulent shear flows. In *Advances in Applied Mathematics* (ed. M. F. Barnsley and S. G. Demko), vol. 26, pp. 183–309. Academic.

Lloyd, J. R. & Moran, W. R. 1974 Natural convection adjacent to horizontal surface of various platforms. *Intl J. Heat Mass Transfer* **96**, 443–447.

Lugosi, G. 2004 Concentration of measure inequalities. *Tech. Rep.* Australian National University.

Lumley, J. L. 1970 *Stochastic Tools in Turbulence*. Academic.

Ma, X. 2001 Hierarchical Galerkin and non-Galerkin models for laminar and turbulent wakes. PhD thesis, Brown University.

Ma, X., Karamanos, G. S. & Karniadakis, G. E. 2000 Dynamics and low-dimensionality of turbulent near-wake. *J. Fluid Mech.* **410**, 29–65.

Ma, X. & Karniadakis, G. E. 2002 A low dimensional model for simulating three dimensional cylinder flow. *J. Fluid Mech.* **458**, 181–190.

Ma, X., Karniadakis, G. E., Park, H. & Gharib, M. 2002 DPIV/T-driven convective heat transfer simulation. *Intl J. Heat & Mass Transfer* **45**, 3517–352.

Ma, X., Karniadakis, G. E., Park, H. & Gharib, M. 2003 DPIV-driven simulation: a new computational paradigm. *Proc. R. Soc. Lond.* A. **459**, 547–565.

Martorell, I., Herrero, J. & Grau, F. 2003 Natural convection from narrow horizontal plates at moderate Rayleigh numbers. *Intl J. Heat Mass Transfer* **46**, 2389–2402.

Noack, B. R., Afanasiev, K., Morzyński, M., Tadmor, G. & Thiele, F. 2003 A hierarchy of low dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.* **497**, 335–363.

Noack, B. R., Papas, P. & Monkewitz, P. A. 2005 The need of a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.* **523**, 339–365.

Oppenheim, A. V. & Shafer, R. W. 1975 *Digital Signal Processing*. Prentice-Hall.

Papoulis, A. 1991 *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill.

Raffel, M., Willert, C. E. & Kompenhans, J. 1998 *Particle Image Velocimetry: A Practical Guide*. Springer.

Rathinam, M. & Petzold, L. 2003 A new look at the proper orthogonal decomposition. *SIAM J. Numer. Anal.* **41**, 1893–1925.

Riesz, F. & Sz-Nagy, B. 1953 *Functional Analysis*. Dover.

Sengupta, M. & Mitra, P. P. 1999 Distribution of singular values of some random matrices. *Phys. Rev.* E **60**, 3389–3392.

Sirisup, S. & Karniadakis, G. E. 2004 A spectral viscosity method for correcting the long-term behavior of POD models. *J. Comput. Phys.* **194**, 92–116.

Sirovich, L. 1987 Turbulence and the dynamics of coherent structures: 1, 2, 3. *Q. Appl. Math.* **45**, 561–590.

Sparrow, E. M. & Carlson, C. K. 1986 Local and average natural convection Nusselt numbers for a uniformly heated, shrouded or unshrouded horizontal plate. *Intl J. Heat Mass Transfer* **29**, 369–379.

Tan, B. T., Damodaran, M. & Willcox, K. 2003 Proper orthogonal decomposition extensions for parametric applications in transonic aerodynamics. *AIAA Paper* 2003–4213.

Tan, B. T., Damodaran, M. & Willcox, K. 2004 Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition. *AIAA J* **42** (8), 1505–1516.

Taylor, J. S., Williams, C., Crisantini, N. & Kandola, J. 2002 Eigenspectrum of the Gram matrix and its relationship to the operator eigenspectrum. In *Algorithmic Learning Theory: 13th International Conference. ALT 2002* (ed. N. Cesa-Bianchi, M. Numao & R. Reischuk), pp. 23–40. Springer.

Therrien, C. 1988 *Discrete Random Signals and Statistical Signal Processing*. Prentice-Hall.

Venturi, D. & Karniadakis, G. E. 2004 Gappy data and reconstruction procedures for flow past a cylinder. *J. Fluid Mech.* **519**, 315–336.

Wan, X. & Karniadakis, G. E. 2005*a* An adaptive multi–element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.* **209**, 617–642.

Wan, X. & Karniadakis, G. E. 2005*b* Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput*. (submitted).

Westerweel, J. 2000 Theoretical analysis of the measurement precision in particle image velocimetry. *Exps. Fluids (Suppl.)* pp. 3–12.

Xiu, D. 2004 Generalized (Wiener-Askey) polynomial chaos. PhD thesis, Brown University.

Xiu, D. & Karniadakis, G. E. 2002 The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput*. **24**, 619–644.

Xiu, D. & Karniadakis, G. E. 2003 Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.* **187**, 137–167.

Yousef, W. W., Tarasuk, J. D. & McKenn, W. J. 1982 Free convection heat transfer from upward-facing isothermal horizontal surfaces. *J. Heat Transfer* **104**, 493–500.

Zdravkovich, M. M. 1997 *Flow around Circular Cylinders*. Oxford University Press.

Zwald, L., Bousquet, O. & Blanchard, G. 2004 Statistical properties of kernel principal component analysis. In *Learning Theory, 17th Annual Conference on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004, Proc*. (ed. J. Shawe-Taylor & Y. Singer), pp. 594–608. Springer.