J. Fluid Mech. (2008), *vol.* 606, *pp.* 339–367.
doi:10.1017/S0022112008001821 Printed in the United Kingdom

Stochastic low-dimensional modelling of a random laminar wake past a circular cylinder

DANIELE VENTURI[1], XIAOLIANG WAN[2] AND GEORGE EM KARNIADAKIS[2]†

[1]Department of Energy, Nuclear and Environmental Engineering, University of Bologna, Italy

[2]Division of Applied Mathematics, Brown University, RI 02912, USA

(Received 18 September 2007 and in revised form 21 March 2008)

We present a new compact expansion of a random flow field into stochastic spatial modes, hence extending the proper orthogonal decomposition (POD) to noisy (non-coherent) flows. As a prototype problem, we consider unsteady laminar flow past a circular cylinder subject to random inflow characterized as a stationary Gaussian process. We first obtain random snapshots from full stochastic simulations (based on polynomial chaos representations), and subsequently extract a small number of deterministic modes and corresponding stochastic modes by solving a temporal eigenvalue problem. Finally, we determine optimal sets of random projections for the stochastic Navier–Stokes equations, and construct reduced-order stochastic Galerkin models. We show that the number of stochastic modes required in the reconstruction does not directly depend on the dimensionality of the flow system. The framework we propose is general and it may also be useful in analysing turbulent flows, e.g. in quantifying the statistics of energy exchange between coherent modes.

1. Introduction

The proper orthogonal decomposition (POD) approach was originally proposed by Lumley (1970) for detecting spatially coherent structures in turbulent flows. An important extension that led to significant computational reductions was proposed by Sirovich (1987), and since then POD has been used in many applications in analysing both laminar and turbulent flows using experimental or numerical simulation data, e.g. Delville *et al.* (2003), Gordeyev & Thomas (2000), Deane *et al.* (1991), Ma *et al.* (2003), Noack *et al.* (2003), see also Rempfer (2003), Holmes, Lumley & Berkooz (1996), and references therein. A particularly useful approach to POD was developed by Aubry (1991) (see also Aubry, Guyonnet & Lima 1995) who introduced the biorthogonal decomposition, a deterministic space–time symmetric version of the POD. 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.

However, if the observed flow is noisy, perturbed or, more generally, it is considered as a superimposition of *random* and *deterministic* elements, then its space–time structures should be considered random as well. A compact mathematical representation of the relationship between a noisy flow and its random space–time structures is

† Author to whom correspondence should be addressed: gk@dam.brown.edu.

lacking. Following the standard POD framework is challenging as it involves random perturbations of the autocorrelation operator's spectral properties; a number of recent efforts have focused on this, e.g. Everson & Roberts (2000), Sengupta & Mitra (1999), Hachem, Loubaton & Najim (2006), Dozier & Silverstein (2007), Hoyle & Rattray (2004). An alternative approach, developed recently by Venturi (2006), provides an analytical characterization of the fluctuations of eigenvalues and eigenfunctions of the response to random perturbations by employing Kato's perturbation theory for linear operators (Kato 1995). The statistics for the perturbed energy levels and the perturbed modes are expressed explicitly in terms of a power series of the random flow standard deviation. As pointed out in the conclusions of Venturi (2006), however, the issue of *stochastic* low-dimensional modelling and simulation is still an open question. In particular, a critical point is how 'randomness' propagates in Galerkin models or in modal energy flow analyses (Noack, Papas & Monkiewitz (2005)). Similarly, Burkardt & Webster (2007) pointed out that "...little work has been done on extending the POD method to the field of stochastic PDEs and yet this is a field of study which seems ideal for this approach, given the need to consider a vast ensemble of solutions".

In terms of representing random processes spectrally, the polynomial chaos approach (see Ghanem & Spanos 1998) and its generalization (Xiu & Karniadakis 2002), allows accurate representations. While mathematically elegant, however, standard polynomial chaos stochastic Galerkin schemes suffer form the 'curse of dimensionality': the number of basis functions increases exponentially with the size of random vector characterizing the system. In order to mitigate such computational cost, several authors have attempted to extend the POD low-dimensional modelling ideas to the stochastic framework. For example, Acharjee & Zabaras (2006) proposed expanding each polynomial chaos spatio-temporal mode according to the classical deterministic space–time biorthogonal decomposition. Also, Doostan, Ghanem & Red-Horse (2007) proposed a multiscale procedure based on two steps: a *coarse*-scale analysis to identify the low-dimensional manifold corresponding to the stochastic solution (assuming it exists!), and a subsequent *fine*-scale analysis based on the results of the coarse scale. However, many of these attempts are often straightforward applications of the standard POD theory in the deterministic (Acharjee & Zabaras 2006) or in the statistical (Doostan *et al.* 2006) sense and do not provide a new theoretical framework for analysing noisy flow systems.

In this paper we propose a new framework based on an expansion that employs random modes to represent the noisy flow field. We assume that we have available several random snapshots of the flow and we attempt to minimize the approximation error in different norms, appropriately defined both in random and deterministic spaces. The advantage of our method with respect to the deterministic POD (which is included in our theory as a subcase) is that if one is able to compute the random flow in any representation (e.g. Monte Carlo, polynomial chaos, etc.), then it is possible to represent the entire ensemble of the random flow behaviour in a compact form for purposes of both analysis and low-dimensional modelling. To this end, we develop a theory for random projections, based on which we can construct, for example, a single model for all flows past a circular cylinder with the Reynolds number as a stochastic input in the range $[Re_l, Re_u]$. This, in turn, allows us to study the instabilities the flow undergoes as a function of the Reynolds number, which is now a parameter that explicitly appears in the system and which we can directly control. It is also possible to analyse energy cascades and correlation between random energy levels in the statistical sense, which can be very useful in analysing turbulent flows. For example, using the standard POD approach to first relate the (standard) POD modes

to the coherent structures of the flow, we can then use the theory we present here to quantify how these structures exchange energy. This, in turn, means that we can analyze both the energy cascade and determine which structures have high correlation in a statistical sense.

As an illustrative example, we consider the flow past a circular cylinder, and we study the response of this flow to a random inflow perturbation, which is assumed to be a Gaussian homogeneous stationary process with Gaussian covariance. Random flow past a circular cylinder has been recently investigated by various stochastic methods (see e.g. Wan & Karniadakis 2006*a*; Kamiński & Carey 2005; Xiu & Karniadakis 2003). We obtain the initial condition from a fully developed deterministic (time-periodic) state, and we study the transient state corresponding to the random inflow condition. In Venturi (2006), randomness was included in the system through a Gaussian uncorrelated random perturbation superimposed on a deterministic state. This kind of noise was chosen empirically to model an experimental-like type of uncertainty. The spectral properties of the random autocorrelation function were obtained by perturbation theory, using well-known properties of Gaussian processes. In the current work, randomness comes from the response of the Navier–Stokes equations to a random inflow boundary condition, and therefore this randomness is related to the intrinsic behaviour of the fluid flow.

The paper is organized as follows. In § 2 we develop a representation of the random flow in terms of a superimposition of random, weakly orthogonal, spatial structures evolving in time. In § 3 we apply the new methodology to the random laminar wake past a circular cylinder corresponding to a random inflow condition whose representation is described in detail in § 3.2. In § 3.4 we obtain a reduced-order representation of the random wake as a superimposition of random spatial structures, and in § 3.7 we study the stochastic low-dimensional model arising from the random projection. The main findings and their implications are summarized in § 4. We also include three brief Appendices for the interested reader to be able to reproduce our results.

2. Flow decomposition into a superimposition of random modes

It is often convenient to represent complex flows as a superimposition of modes. Clearly the method for decomposition and the criterion for representation depend on the particular problem. For instance, Fourier transforms are standard tools to determine energy cascades in turbulent flows while proper orthogonal decompositions have been extensively used to detect spatially coherent structures in various applications (see e.g. Gordeyev & Thomas 2000; Delville *et al.* 2003).

The first step in obtaining useful information about the flow field through a particular type of decomposition is to determine what kind of information we are trying to extract or we are mainly interested in. This results in the identification of a criterion for decomposition (see table 1). In the presence of random fields, this operation is much more challenging than in the corresponding deterministic case. In fact, the complexity and the high dimensionality of these fields naturally yields a wide range of possible choices for decomposing them. In § 2.1, we develop tools for a decomposition theory based on three criteria:

optimal representation of the mean flow;
optimal representation of the second-order moment;
optimal representation of the standard deviation.

Random field $\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$ (any representation) obtained from stochastic numerical simulations or from experiments through estimation methods		
↓		
Choice of the criterion for decomposition	→	• Optimal representation in the mean sense • Optimal representation in second-order moment sense • Optimal representation in the standard deviation sense } § 2.1 • Decomposition theories based on nonlinear field dependent correlations to satisfy other criteria (§ 2.2 and § 3.6)
↓		
Computation of random modes and associated random energies (§ 3.4)		
↓		
Galerkin projection of Navier–Stokes equations onto random modes, reduced order modelling and simulation (§ 3.7)		

TABLE 1. Orthogonal decomposition and Galerkin modelling: organization chart.

The resulting three types of expansions are, in some sense, special in view of the many possibilities arising from using nonlinear kernel methods (Scholköpf, Smola & Müller 1998). These new theories and their relation to our formulation are discussed in some detail in § 2.2. Once a criterion for decomposition has been identified, it is possible to proceed with the computation of the orthogonal representation and perform a modal analysis of the random flow as illustrated in § 3.4. Finally, the weak orthogonality of the random spatial modes is used to develop low-dimensional Galerkin representations of Navier–Stokes equations in the presence of uncertainty (§ 3.7). The main steps to perform the analysis are summarized in table 1.

2.1. *Formulation of the decomposition theory*

We consider a random field $\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$ in a space–time domain $\Omega \times T$. This means that the ordered pair $(\boldsymbol{x}, t)$ is such that $\boldsymbol{x} \in \Omega$ (spatial domain) and $t \in T$ (temporal domain). The argument $\boldsymbol{\xi}$ labels an outcome of the random field. We look for biorthogonal representations in the form

$$\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi}) = \sum_{i=1}^{\infty} \sqrt{\mu_i^{(h)}} \Psi_i^{(h)}(t)\, \boldsymbol{\Phi}_i^{(h)}(\boldsymbol{x}; \boldsymbol{\xi}) \,, \tag{2.1}$$

where the superscript (h) denotes that we will construct these expansions according to different variational principles.

We assume that the temporal modes $\psi_i^{(h)}$ are strongly orthogonal in time while the spatial modes $\boldsymbol{\Phi}_i^{(h)}$ are weakly orthogonal in space with respect to appropriate inner products defined below. We denote by $(\,,\,)_T$ the inner product in the temporal domain and by $\{\,,\,\}_h$ $(h = 0, 1, 2)$ different types of inner products in the spatial domain. In this paper we will consider

$$(\psi_i, \psi_j)_T := \int_T \psi_i(t)\psi_j(t)\,\mathrm{d}t. \tag{2.2}$$

For every pair of square-integrable random fields $\boldsymbol{b}(\boldsymbol{x}; \boldsymbol{\xi})$ and $\boldsymbol{c}(\boldsymbol{x}; \boldsymbol{\xi})$ in the spatial domain we define the following three types of inner products:

$$\{\boldsymbol{b}, \boldsymbol{c}\}_0 := \int_\Omega \langle \boldsymbol{b} \rangle \cdot \langle \boldsymbol{c} \rangle \,\mathrm{d}\boldsymbol{x} \,, \tag{2.3}$$

$$\{\boldsymbol{b}, \boldsymbol{c}\}_1 := \int_{\Omega} \langle \boldsymbol{b} \cdot \boldsymbol{c} \rangle \, \mathrm{d}\boldsymbol{x}\,, \tag{2.4}$$

$$\{\boldsymbol{b}, \boldsymbol{c}\}_2 := \int_{\Omega} \big(\langle \boldsymbol{b} \cdot \boldsymbol{c} \rangle - \langle \boldsymbol{b} \rangle \cdot \langle \boldsymbol{c} \rangle \big) \, \mathrm{d}\boldsymbol{x}\,, \tag{2.5}$$

where the averaging operation $\langle \cdot \rangle$ is defined as

$$\langle f \rangle := \int f(\boldsymbol{\xi})\, W(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}\,, \tag{2.6}$$

and $W(\boldsymbol{\xi})$ is the (non-degenerate) multivariate joint probability density of $\boldsymbol{\xi}$. In the context of stochastic methods for the solution of stochastic partial differential equations (Ghanem & Spanos 1998; Xiu & Karniadakis 2003; Webster 2007) $\boldsymbol{\xi}$ corresponds to the vector of random inputs forcing the system. The orthogonality requirements for $\psi_i^{(h)}$ and $\boldsymbol{\Phi}_i^{(h)}$ are

$$\big(\psi_i^{(h)}, \psi_j^{(h)}\big)_T = \delta_{ij}, \tag{2.7a}$$

$$\big\{\boldsymbol{\Phi}_i^{(h)}, \boldsymbol{\Phi}_j^{(h)}\big\}_h = \delta_{ij}. \tag{2.7b}$$

Now we consider the positive-definite functionals

$$\|\boldsymbol{u}\|_h^2 := \int_T \{\boldsymbol{u}, \boldsymbol{u}\}_h \, \mathrm{d}t\,, \qquad h = 0, 1, 2\,. \tag{2.8}$$

By elementary arguments of the calculus of variations we minimize the 'distance' (in the generic norm $\|\cdot\|_h$, $h = 0, 1, 2$) between the expansion (2.1) and the random field $\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$. Physically, this corresponds to looking for expansions which minimize the error in the mean sense (case $h = 0$), in the second-order moment sense (case $h = 1$), and in the standard deviation sense (case $h = 2$).

The minimization of the error functional

$$\mathscr{E}_h\big[\psi_1^{(h)}, \ldots, \psi_M^{(h)}\big] := \left\| \boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi}) - \sum_{i=1}^{M} \sqrt{\mu_i^{(h)}}\, \psi_i^{(h)}(t)\, \boldsymbol{\Phi}_i^{(h)}(\boldsymbol{x}; \boldsymbol{\xi}) \right\|_h^2$$

with respect to an arbitrary variation of $\psi_k^{(h)}$ leads to the Euler–Lagrange equations

$$\psi_k^{(h)}(t) = \frac{1}{\sqrt{\mu_k^{(h)}}} \big\{\boldsymbol{u}, \boldsymbol{\Phi}_k^{(h)}\big\}_h. \tag{2.9}$$

From (2.1) and the orthogonality requirements (2.7) we obtain

$$\boldsymbol{\Phi}_k^{(h)}(\boldsymbol{x}; \boldsymbol{\xi}) = \frac{1}{\sqrt{\mu_k^{(h)}}} \int_T \boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})\, \psi_k^{(h)}(t) \, \mathrm{d}t. \tag{2.10}$$

Substitution of (2.10) into (2.9) gives the temporal eigenvalue problem

$$\mu_k^{(h)} \psi_k^{(h)}(t) = \int_T \mathscr{T}^{(h)}(t, t') \psi_k^{(h)}(t') \, \mathrm{d}t', \tag{2.11}$$

where the autocorrelation is

$$\mathscr{T}^{(h)}(t, t') := \{\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi}), \boldsymbol{u}(\boldsymbol{x}, t'; \boldsymbol{\xi})\}_h. \tag{2.12}$$

Depending on the choice of the inner product $\{,\}_h$, we have different types of kernels $\mathscr{T}^{(h)}(t,t')$. Also, it follows from (2.3), (2.4) and (2.5) that

$$\mathscr{T}^{(1)}(t,t') = \mathscr{T}^{(0)}(t,t') + \mathscr{T}^{(2)}(t,t'). \tag{2.13}$$

The spectral decomposition of these kernels leads to constructions of different expansions (2.1) according to different 'sections' of the random field $\boldsymbol{u}$: through $\mathscr{T}^{(0)}$ we take optimal sections for the representation of the mean field; through $\mathscr{T}^{(1)}$ we take optimal sections for the representation of the second-order moment and through $\mathscr{T}^{(2)}$ we take optimal sections for the representation of the standard deviation. There are connections between these different orthogonal representations; in particular, there exist appropriate rotation operators which transform one representation into another (see Appendix A for details).

It is often convenient to re-write the expansion (2.1) as

$$\boldsymbol{u}(\boldsymbol{x},t;\boldsymbol{\xi}) = \sum_{j=1}^{\infty} \psi_j^{(h)}(t)\,\boldsymbol{a}_j^{(h)}(\boldsymbol{x};\boldsymbol{\xi}), \tag{2.14}$$

where by definition

$$\boldsymbol{a}_j^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) := \sqrt{\mu_j^{(h)}}\boldsymbol{\Phi}_j^{(h)}(\boldsymbol{x};\boldsymbol{\xi}). \tag{2.15}$$

The random energy of the kth random mode is

$$\theta_k^{(h)}(\boldsymbol{\xi}) := \int_{\Omega} \boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) \cdot \boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{x}. \tag{2.16}$$

It is clear that this represents the random energy level of the kth mode if we consider (2.14) and the fact that the temporal modes $\psi_j^{(h)}$ are orthonormal. In this sense the random energy level $\theta_k^{(h)}(\boldsymbol{\xi})$ is essentially the squared L_2 spatial norm of the (non-normalized) random spatial modes $\boldsymbol{a}_k^{(h)}$. Another important quantity we will consider is the correlation coefficient between different random energy levels

$$C_{ij}^{(h)} = \frac{\langle \theta_i^{(h)}\theta_j^{(h)} \rangle - \langle \theta_i^{(h)} \rangle \langle \theta_j^{(h)} \rangle}{\sigma_{\theta_i^{(h)}}\sigma_{\theta_j^{(h)}}}, \tag{2.17}$$

where $\sigma_{\theta_i^{(h)}}$ denotes the standard deviation of the ith energy level. This coefficient reveals whether there exists a correlation between different random structures, and, therefore, it allows study of the energy exchange between different random modes.

2.2. *Decomposition theories based on nonlinear field-dependent correlations*

In the previous section we have considered three different types of kernels, namely $\mathscr{T}^{(h)}$ $(h = 0, 1, 2)$, for optimal representation of a random field in the mean, the second-order moment and the standard deviation sense. These kernels were induced by a suitable choice of the inner product spaces and minimum variational principles. More precisely the inner product itself induces both the kernel and the minimization principle.

However, there are many other possibilities which make use of nonlinear methods to represent optimally other features of the random field $\boldsymbol{u}$. These methods are today known as kernel methods or nonlinear component analyses (Scholköpf *et al.* 1998; Tenenbaum, de Silva & Langford 2000 Ham *et al.* 2003). The main difference with respect to standard methods is that the field $\boldsymbol{u}$ is nonlinearly mapped onto another 'feature' space before decomposition is performed. Let us briefly discuss what are the

$\boldsymbol{G}(\boldsymbol{u})$	Induced correlation kernel for eigenvalue problem
$\langle \boldsymbol{u} \rangle$	$\mathscr{T}^{(0)}$
$\boldsymbol{u}$	$\mathscr{T}^{(1)}$
$\boldsymbol{u} - \langle \boldsymbol{u} \rangle$	$\mathscr{T}^{(2)}$
$\mathrm{e}^{-\boldsymbol{u}}$	$\int_{\Omega} \langle \mathrm{e}^{-\boldsymbol{u}} \cdot \mathrm{e}^{-\boldsymbol{u}'} \rangle \, \mathrm{d}\boldsymbol{x}$

TABLE 2. Special choices of $\boldsymbol{G}$ and corresponding correlation functions.

consequences of this approach. Given an arbitrary stochastic field $\boldsymbol{u}$, we first construct a nonlinear transformation

$$\boldsymbol{u} \rightarrow \boldsymbol{G}(\boldsymbol{u}) \tag{2.18}$$

and subsequently a decomposition theory for $\boldsymbol{G}(\boldsymbol{u})$ based on the arguments of § 2.1. The temporal correlation we obtain is

$$\mathscr{G}(t, t') := \int_{\Omega} \langle \boldsymbol{G}(\boldsymbol{u}) \cdot \boldsymbol{G}(\boldsymbol{u})' \rangle \, \mathrm{d}\boldsymbol{x}. \tag{2.19}$$

First we observe that the functions $\mathscr{T}^{(h)}$ considered in § 2.1 are particular subcases of this general theory as shown in table 2. There is a biunivocal correspondence between the nonlinear transformation $\boldsymbol{G}$ and the induced correlation $\mathscr{G}$. Given $\boldsymbol{G}$ it is straightforward to obtain the correlation $\mathscr{G}$ through (2.19). Moreover it can be proved that given a positive $\mathscr{G}$ there exist a mapping $\boldsymbol{G}$ such that (2.19) holds. As an example of the correspondence

$$\boldsymbol{G}(\boldsymbol{u}) = \mathrm{e}^{-\boldsymbol{u}} \tag{2.20}$$

obviously induces the following type of correlation

$$\mathscr{G}(t, t') = \int_{\Omega} \langle \mathrm{e}^{-\boldsymbol{u}} \cdot \mathrm{e}^{-\boldsymbol{u}'} \rangle \, \mathrm{d}\boldsymbol{x}. \tag{2.21}$$

Within this framework it is seen that very general types of decoposition theories can be formulated in order to satisfy general criteria. For instance Jenssen *et al.* (2007) and Paiva, Xu & Principe (2006) have recently constructed orthogonal expansions based on maximum Rényi (1961) second-order entropy principles. Statistical properties of kernel principal components have been recently investigated by Blanchard, Bousquet & Zwald (2007). In a field-theoretic view the correlation function $\mathscr{G}$ can also be generated though functional methods such as the path integral representation (Zinn-Justin 2002; Amit & Martín-Mayor 2005). This means that we can induce a nonlinear mapping $\boldsymbol{G}$ implicitly through the partition function theory and study complex interactions between the generalized modes by, e.g., perturbation expansions or diagrammatic representations (Hoyle & Rattray 2004).

We leave general decomposition approaches using nonlinear field-dependent kernels for future studies. In the next section we will consider representations of velocity fields according to the theory presented in § 2.1 and apply the method to the random wake

past a cylinder. In § 3.6 we describe applications of nonlinear kernel methods to nonlinear dimensionality reduction in random space.

3. Application to the random laminar wake past a cylinder

In this section we study the random laminar wake past a circular cylinder corresponding to random inflow. The purpose of this application is twofold. First, we want to use the decomposition developed in the previous section to analyse the random structures of this flow and to construct a Galerkin low-dimensional stochastic model for the evolution of the system. Second, we want to investigate how the representation (2.1) behaves when the dimensionality of the problem in random space increases, e.g. when we consider stochastic input processes having a small correlation length, or, more generally, when the system is subject to many simultaneous types of uncertainties. We use functional polynomial chaos representations to compute both the random flow and its decomposition into random modes (see Appendix B).

3.1. *Wiener–Hermite functional representation of the Navier–Stokes equations*

We consider the dimensionless form of the Navier–Stokes equations for an incompressible fluid

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} = -\nabla p + \frac{1}{Re}\nabla^2\boldsymbol{v}\,, \tag{3.1}$$

$$\nabla\cdot\boldsymbol{v} = 0\,, \tag{3.2}$$

where $\boldsymbol{v}(\boldsymbol{x},t;\boldsymbol{\xi})$ and $p(\boldsymbol{x},t;\boldsymbol{\xi})$ are dimensionless random velocity and pressure fields respectively. The Reynolds number is based on the velocity $\mathscr{U}$ and on the diameter D of the cylinder, i.e. $Re = \mathscr{U}D/\nu$, where $\mathscr{U}$ is the space–time average of the (statistical) mean inflow streamwise velocity component.

We represent the velocity and the pressure on a polynomial chaos basis $\Gamma_j(\boldsymbol{\xi})$

$$\boldsymbol{v}(\boldsymbol{x},t;\boldsymbol{\xi}) = \sum_{j=0}^{P} \widehat{\boldsymbol{v}}_j(\boldsymbol{x},t)\,\Gamma_j(\boldsymbol{\xi})\,, \tag{3.3}$$

$$p(\boldsymbol{x},t;\boldsymbol{\xi}) = \sum_{j=0}^{P} \widehat{p}_j(\boldsymbol{x},t)\,\Gamma_j(\boldsymbol{\xi}), \tag{3.4}$$

where $\{\Gamma_j(\boldsymbol{\xi})\}$ are polynomial functionals, which are orthogonal with respect to the joint probability density of the random input vector $\boldsymbol{\xi}$. In § 3.2 we will characterize this random input vector in terms of a set of normalized and independent Gaussian random variables. Correspondingly, the polynomial functionals $\{\Gamma_j(\boldsymbol{\xi})\}$ are the standard Wiener–Hermite functionals (Wiener 1966), which have been employed in turbulence modelling (Lee, Meecham & Hogge 1982; Meecham & Siegel 1964; Meecham & Jeng 1968; Bodner 1969).

A substitution of (3.3) and (3.4) into the Navier–Stokes equations and subsequent projection onto $\{\Gamma_j\}$ leads to the following system of deterministic coupled partial differential equations ($k = 0, \ldots, P$)

$$\frac{\partial \widehat{\boldsymbol{v}}_k}{\partial t} + \sum_{i,j=0}^{P} \frac{\langle \Gamma_i\Gamma_j\Gamma_k\rangle}{\langle \Gamma_k^2\rangle}(\widehat{\boldsymbol{v}}_i\cdot\nabla)\widehat{\boldsymbol{v}}_j = -\nabla\widehat{p}_k + \frac{1}{Re}\nabla^2\widehat{\boldsymbol{v}}_k\,, \tag{3.5}$$

$$\nabla\cdot\widehat{\boldsymbol{v}}_k = 0\,, \tag{3.6}$$

to be solved for the unknown 'chaos modes' $\widehat{\boldsymbol{v}}_j(\boldsymbol{x},t)$.

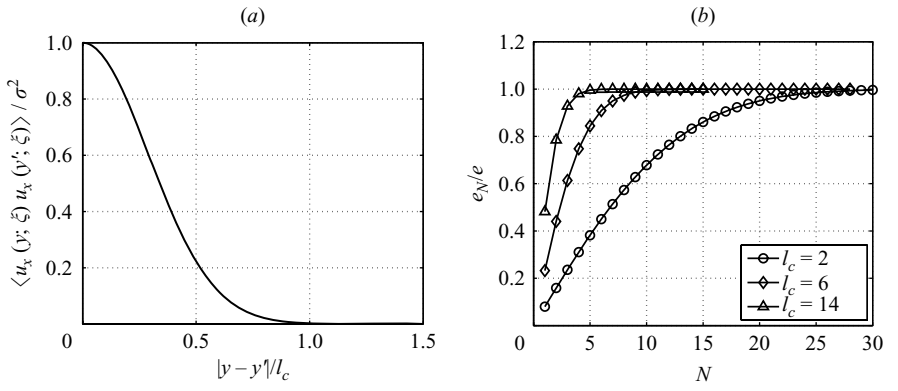

FIGURE 1. (*a*) Correlation function of streamwise random velocity fluctuation. (*b*) Relative energy of truncated Karhunen–Loève representations as a function of the number of modes.

3.2. *Representation of the random inflow boundary condition*

The dimensionless inflow streamwise velocity component v_x is assumed to be a time-independent wide-sense stationary Gaussian stochastic process in the form

$$v_x(y;\boldsymbol{\xi}) = 1 + u_x(y;\boldsymbol{\xi})\,. \tag{3.7}$$

(Wide sense stationarity means that the mean velocity $\langle v_x\rangle$ is not y-dependent while the covariance function depends only on $|y-y'|$.)

The zero mean streamwise random fluctuation $u_x(y;\boldsymbol{\xi})$ follows a Gaussian correlation function

$$\langle u_x(y;\boldsymbol{\xi})u_x(y';\boldsymbol{\xi})\rangle = \sigma^2 \exp\left[-A\frac{(y-y')^2}{l_c^2}\right], \tag{3.8}$$

where A is a normalization constant. The parameter σ defines in some sense the 'magnitude' of the perturbation u_x. The correlation function (3.8), suitably normalized, is shown in figure 1(*a*). We choose here $A = 6$ in order for the correlation length to have a direct physical meaning, i.e. the correlation becomes approximately zero at $|y-y'| = l_c$.

The spectral analysis of (3.8) leads to a Karhunen–Loève representation of the inflow streamwise velocity component in the form

$$v_x(y;\xi_1,\ldots,\xi_N) = 1 + \sigma\sum_{i=1}^{N}\sqrt{\lambda_i}\,g_i(y)\xi_i\,, \tag{3.9}$$

where $\{\xi_i\}$ is a set of zero-mean, uncorrelated and normalized Gaussian random variables†; λ_i, $g_i(y)$ are, respectively, eigenvalues and eigenfunctions of an integral operator having kernel $\langle u_x(y;\boldsymbol{\xi})u_x(y';\boldsymbol{\xi})\rangle/\sigma^2$. The random input multivariate

† We remark that zero-mean, uncorrelated Gaussian random variables are also necessarily independent. This is an important remark since, as is well known, the Wiener–Hermite functional polynomial representation (3.3) and (3.4) is constructed using a very special independent increment (Wiener) process, i.e. it assumes stochastic independence of all the components of the random input vector $\boldsymbol{\xi}$.

probability density $W(\boldsymbol{\xi})$, which appears in the definition (2.6), has the form

$$W(\boldsymbol{\xi}) = K \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\xi_i^2\right], \tag{3.10}$$

where K is a suitable normalization constant. Consequently, the functional polynomial chaos $\{\Gamma_k(\boldsymbol{\xi})\}$ is the classical Wiener–Hermite chaos. The parameter σ represents the standard deviation of the random perturbation u_x. This can be seen if we consider

$$\langle v_x(y;\boldsymbol{\xi})\rangle = 1, \tag{3.11a}$$

$$\langle v_x(y;\boldsymbol{\xi})^2\rangle = 1 + \sigma^2\sum_{i=1}^{\infty}\lambda_i g_i(y)^2 \tag{3.11b}$$

and note that the latter summation equals 1 because of the spectral decomposition of (3.8):

$$\langle u_x(y;\boldsymbol{\xi})u_x(y';\boldsymbol{\xi})\rangle = \sigma^2\sum_{i=1}^{\infty}\lambda_i g_i(y)\, g_i(y') \quad \Rightarrow \quad \sum_{i=1}^{\infty}\lambda_i g_i(y)^2 = 1. \tag{3.12}$$

This means that if we consider infinite terms in the representation (3.9) the standard deviation of the velocity streamwise component at the inflow theoretically does not depend on y. However, when we consider a finite number of terms in the expansion the standard deviation at the inflow may be a function of y. The truncation process has to be performed with some care since we must check that the approximated correlation function is still at least approximately Gaussian, which means, in particular, that the energy of the neglected modes has to be negligible. In figure 1(b) we plot the relative energy of the truncated expansion with respect to the energy of the 'full' expansion, where e_N and e are defined as

$$e_N := \sum_{i=1}^{N}\lambda_i, \quad e := \sum_{i=1}^{\infty}\lambda_i. \tag{3.13}$$

Quantitative results are shown in table 3. We note that when the correlation length l_c becomes smaller we need many more terms to properly represent the stochastic process since it is approaching an independent increment process (Segall & Kailath 1976).

It is important to note that the correlation length l_c of the velocity field at the inflow cannot be chosen arbitrarily because it is connected to the Reynolds number of the flow. It is physically rather unlikely that a realization of a low-Reynolds-number flow exists for which the velocity field has a small correlation length at the inflow. Small correlation length would make the velocity profile very irregular and noisy, which is not observed in experiments at low Reynolds number (Zdravkovich 1997). In § 3.3 we consider the random laminar wake at $Re = 100$ and, therefore, we choose $l_c = 14$, which means that the random inflow velocity is correlated up to 28 diameters of the cylinder. Some realizations of these profiles are shown in figure 2(a). In figure 2(b) we show that a truncation up to order $N = 6$ of the Karhunen–Loève expansion (3.9) gives an approximated correlation function (3.12) essentially indistinguishable from the exact Gaussian correlation (3.8) for $l_c = 14$ (see also table 3).

3.3. *Stochastic direct numerical simulation (DNS) of the random wake*

We have simulated the time-dependent random flow at $Re = 100$ by discretizing the system of equations (3.5) and (3.6) using the spectral/hp element method described

		N				
		2	4	6	8	10
l_c	2	15.18%	31.05%	44.98%	57.31%	67.82%
	6	44.06%	74.77%	90.92%	97.39%	99.40%
	14	78.51%	98.05%	99.91%	100.00%	100.00%

TABLE 3. Relative energy e_N/e of Karhunen–Loève representation as a function of the correlation length l_c and number N of Gaussian independent random variables.

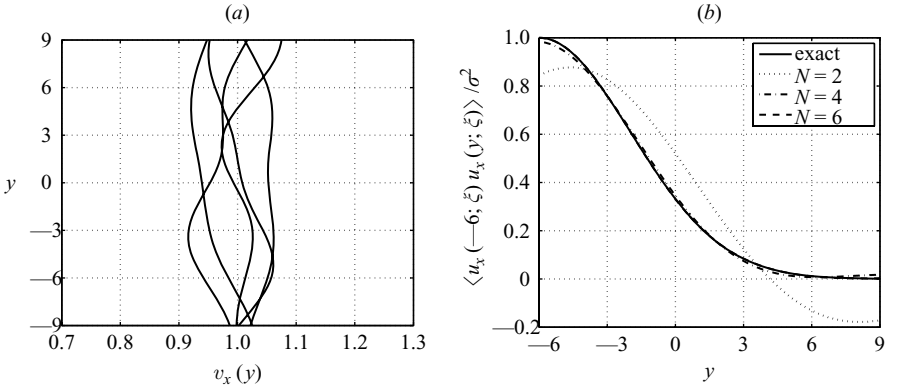

FIGURE 2. (*a*) Some samples of the dimensionless random inflow boundary condition for correlation length $l_c = 14$ and standard deviation $\sigma = 0.05$. (*b*) Convergence to the correct correlation function for truncated Karhunen–Loève representations of the random inflow process ($l_c = 14$).

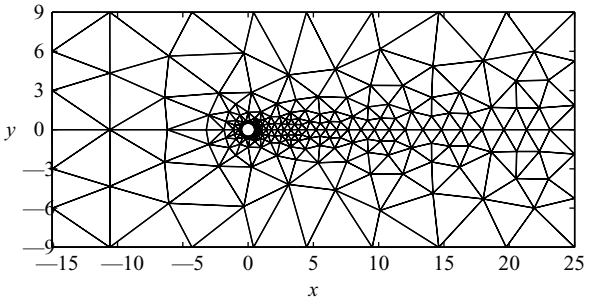

FIGURE 3. Geometry and computational mesh composed of 412 spectral elements of polynomial order 11.

in Karniadakis & Sherwin (2005). The computational domain is shown in figure 3 and it is composed of 412 spectral elements of polynomial order 11 per direction; see computational details for this problem in Ma, Karamanos & Karniadakis (2000). All the results presented hereafter are based on a representation of the random input process having $l_c = 14$ and $\sigma = 0.05$ in terms of $N = 6$ random variables. We use a third-order Wiener–Hermite representation for velocity and pressure fields. This means that the number of coupled partial differential equations (3.5) we have solved simultaneously in the space–time domain is $P = 83$.

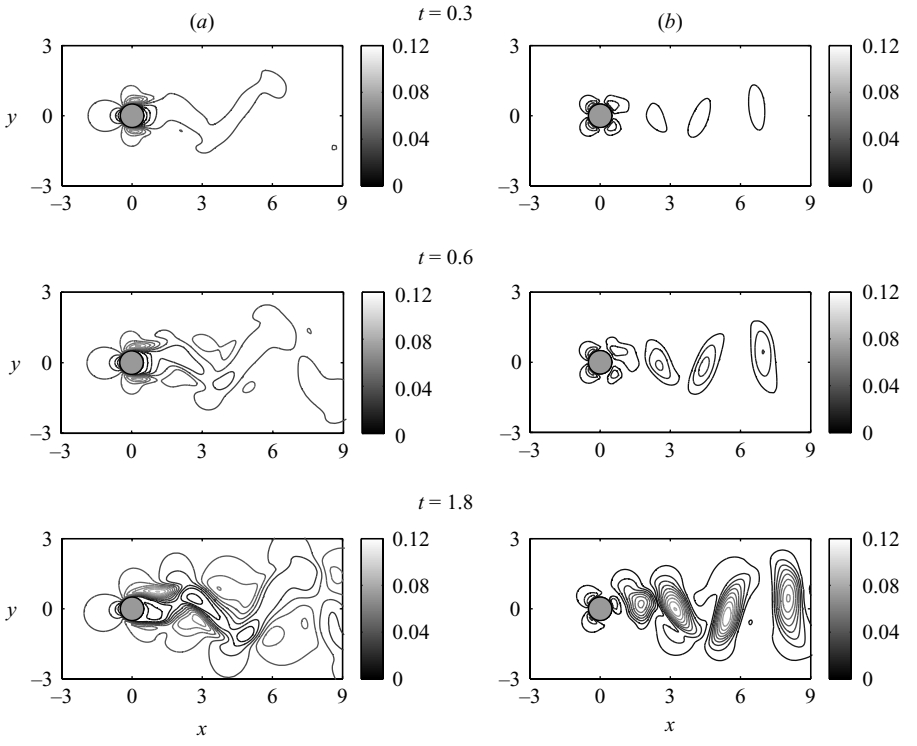

FIGURE 4. Standard deviation of streamwise (a) and crossflow (b) velocity components at different time instants. The shedding period is about $|T| = 6$ convective time units.

The initial condition for the system of equations (3.5) is a fully developed deterministic wake at $Re = 100$. The random inflow boundary condition generates a transient state for the stochastic velocity as shown in figure 4. The standard deviation of the velocity field increases in time due to the randomness of the streamwise velocity profile at the inflow and after long-term integration it will saturate, approaching a constant field having $y = 0$ symmetry. The mean flow approximately still follows the deterministic evolution in one shedding period, which is approximately $|T| = 6$ convective time units. Since the period of integration is quite small, the Wiener–Hermite functional representation captures accurately the statistics of the flow field (Wan & Karniadakis 2006*a*). This claim will also be verified in § 3.7, where we compare the prediction of the low-dimensional Galerkin system with the DNS results.

3.4. *Stochastic eigen-decomposition and modal analysis*

In order to perform the decomposition (2.1) we have extracted 40 equidistant stochastic snapshots of the random flow field in one shedding period of the mean flow. This means that each snapshot contains a complete statistical description of the flow. In other words the snapshots we are considering are in the form $\boldsymbol{v}\left(\boldsymbol{x}, t_j; \boldsymbol{\xi}\right)$ $(j = 1, \ldots, 40)$. First, we decompose the random velocity field $\boldsymbol{v}(\boldsymbol{x}, t; \boldsymbol{\xi})$ into a mean flow $\boldsymbol{U}$ with a superimposed random fluctuation $\boldsymbol{u}$

$$\boldsymbol{v}(\boldsymbol{x}, t; \boldsymbol{\xi}) = \boldsymbol{U}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi}) , \tag{3.14}$$

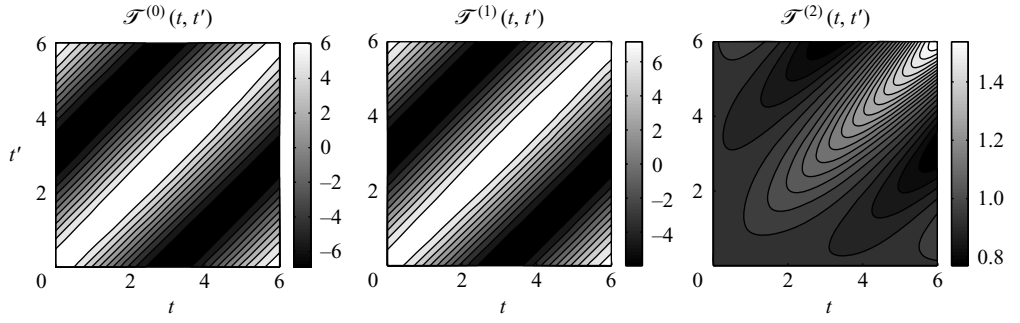

FIGURE 5. Correlations functions $\mathscr{T}^{(0)}(t,t')$, $\mathscr{T}^{(1)}(t,t')$ and $\mathscr{T}^{(2)}(t,t')$.

where by definition

$$\boldsymbol{U}(\boldsymbol{x}) := \frac{1}{|T|}\int_T \langle \boldsymbol{v}(\boldsymbol{x},t;\boldsymbol{\xi})\rangle \,\mathrm{d}t \tag{3.15}$$

and $|T|$ denotes the measure of the temporal domain, i.e. the period of integration. Next, we expand the random fluctuation $\boldsymbol{u}$ according to (2.1), or equivalently (2.14).

We compute different expansions corresponding to $h = 0, 1, 2$ (see (2.3), (2.4) and (2.5)). As previously mentioned, these different representations will give us optimal convergence in the mean, in the second-order moment and in the standard deviation for the representation of the random fluctuation $\boldsymbol{u}(\boldsymbol{x},t;\boldsymbol{\xi})$. We note that each random mode $\boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi})$ (or equivalently $\boldsymbol{\Phi}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi})$) is strongly divergence-free since (3.14), (2.7) and (3.2) imply

$$\nabla\cdot\boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) = \int_T \nabla\cdot\boldsymbol{u}(\boldsymbol{x},t;\boldsymbol{\xi})\,\psi_k^{(h)}(t)\,\mathrm{d}t = 0\,. \tag{3.16}$$

In figure 5 we show the correlation functions (2.12) obtained for $h = 0, 1, 2$. Note that the covariance $\mathscr{T}^{(2)}(t,t')$ increases along the diagonal $t = t'$ as it represents the (spatial) average of the flow standard deviation, which is obviously increasing in time (in the transient state only) due to the random inflow boundary condition. $\mathscr{T}^{(0)}(t,t')$ exhibits the classical pattern of a periodic phenomenon. The summation of $\mathscr{T}^{(0)}(t,t')$ and $\mathscr{T}^{(2)}(t,t')$ gives $\mathscr{T}^{(1)}(t,t')$ which closely resembles $\mathscr{T}^{(0)}(t,t')$ due to the relative small contribution of $\mathscr{T}^{(2)}(t,t')$. The eigenvalues of the correlation functions (2.12) are shown in figure 6. In figure 7, we show the streamwise component of the mean and the standard deviation of some normalized random spatial modes obtained using the inner products $\{,\}_1$. The symmetry of the mean modes (figure 7*a*) suggests a periodic vortex street pattern of the mean flow. A similar pattern analysis can be performed for random modes obtained using other inner products.

Now we study the energy of the random modes and their interaction. In figure 8 we show the mean and the standard deviation of the random energy levels (2.16) (we kept the same ordering of eigenvalues as in figure 6). The noisy behaviour observed in figure 8(*a*) deserves some comments. When we decompose the random velocity by a principle of minimum distance in the mean sense (inner product $\{,\}_0$) the amount of statistical information we are actually using to construct the orthogonal expansion is very limited and in practice it includes only the mean flow data. In representing the random energy (2.16), we need to compute correlations between orthogonal modes which obviously fail to capture any statistics of order higher than one if they are

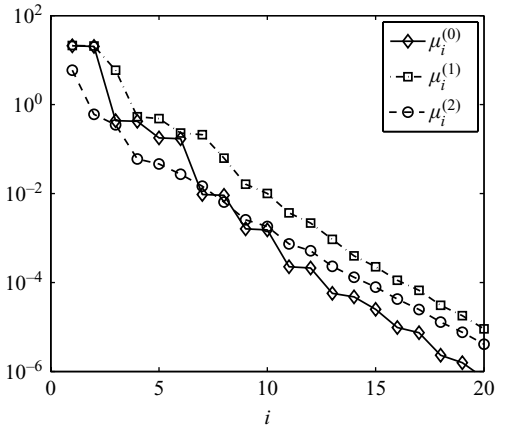

FIGURE 6. Spectra of the temporal correlations $\mathscr{T}^{(h)}$ for $h = 0, 1, 2$.

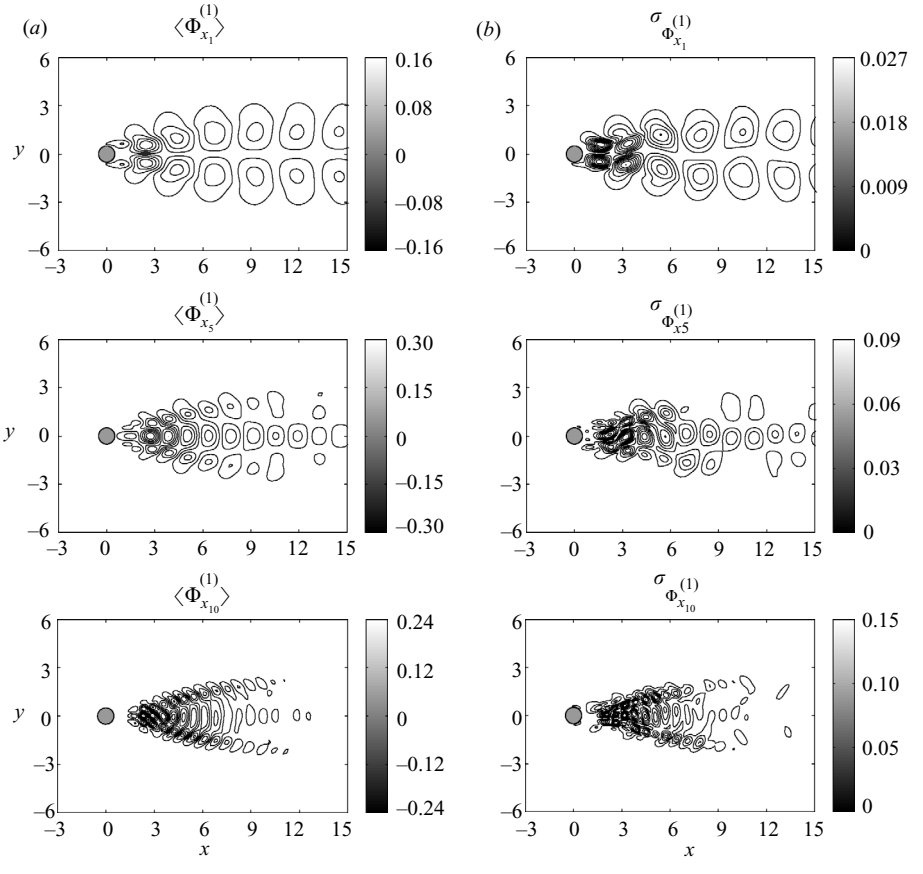

FIGURE 7. Streamwise component of normalized random spatial modes obtained using the second-order moment inner product $\{,\}_1$: (*a*) mean and (*b*) standard deviation.

obtained in the inner product $\{,\}_0$. In other words, the series expansion we construct in the inner product $\{,\}_0$, despite its optimality for the mean flow representation, is not complete in the statistical sense (we will discuss this important issue more in

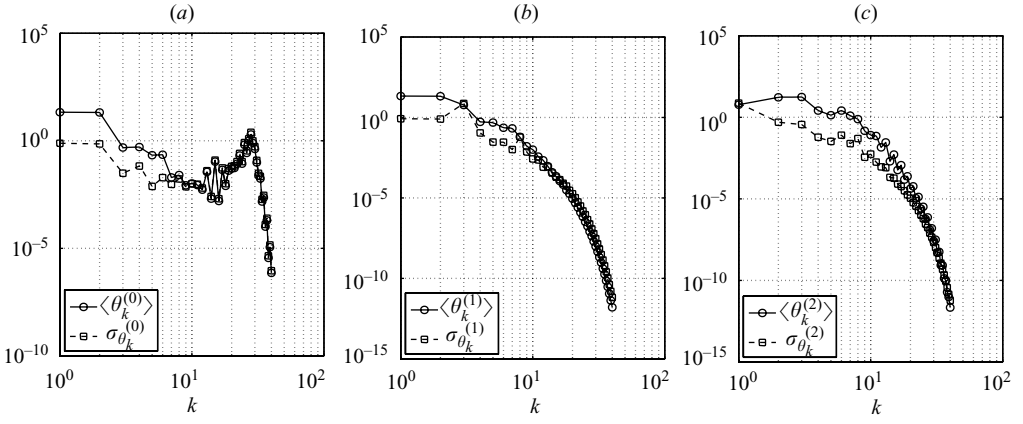

FIGURE 8. Random energy levels (2.16) obtained using different inner products: mean and standard deviation of $\theta_k^{(0)}(a)$, $\theta_k^{(1)}(b)$ and $\theta_k^{(2)}(c)$.

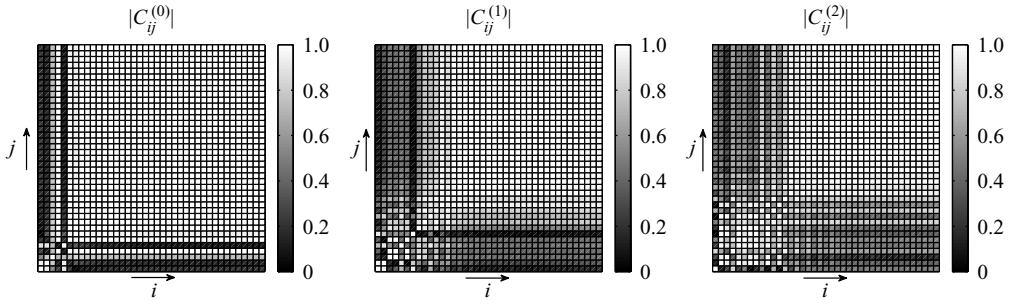

FIGURE 9. Absolute values of correlation coefficients between random energy levels.

detail in § 3.7.) and therefore the correlations (see (2.16)) required to compute $\theta_k^{(0)}$ are not accurate. The absolute value of the correlation coefficient (2.17) between different energy levels is shown in figure 9. The plot should be interpreted as follows: each cell is identified by two labels i and j (bottom left corner of each plot is $i, j = 1$). The grey intensity of each cell represents the value of the correlation coefficient between the energy levels i and j. It is interesting to note that higher-order modes are strongly correlated ($C_{ij}^{(h)} \simeq 1$) while the exchange of energy between lower-order modes is more irregular (this is influenced somewhat by the way we have ordered the levels $\theta_i^{(h)}$). In particular, we note that when using the representation $h = 2$ there are two zones of highly correlated energy levels (see figure 9):

$i, j \lesssim 13$ energy exchange between lower-order modes,

$i, j \gtrsim 14$ energy exchange between higher-order modes.

The correlation between the energy of higher-order and lower-order modes is weaker.

3.5. *Low-dimensionality of the random wake*

We study convergence in a space–time sense of the mean and the standard deviation of the random flow as a function of the number of stochastic POD modes M. To this

end, we define

$$\boldsymbol{u}_M(\boldsymbol{x},t;\boldsymbol{\xi}) := \sum_{j=1}^{M} \psi_j^{(h)}(t)\,\boldsymbol{a}_j^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) \,, \tag{3.17}$$

where M is equal to or less than the number of extracted temporal snapshots. The low-dimensional representation of the mean and the standard deviation of the velocity field is

$$\langle \boldsymbol{v}_M(\boldsymbol{x},t;\boldsymbol{\xi})\rangle = \boldsymbol{U}(\boldsymbol{x}) + \langle \boldsymbol{u}_M(\boldsymbol{x},t;\boldsymbol{\xi})\rangle \,, \tag{3.18}$$

$$\sigma_{\boldsymbol{v}_M}(\boldsymbol{x},t) = \sigma_{\boldsymbol{u}_M}(\boldsymbol{x},t) \,. \tag{3.19}$$

In figures 10, 11 and 12 we show convergence of the mean and the standard deviation of the streamwise and the crossflow component of the velocity at the downstream crossline $x = 3$.

We define the space–time errors in the $L_2(\Omega \times T)$ norm as follows

$$\|\langle \boldsymbol{v}\rangle - \langle \boldsymbol{v}_M\rangle\|^2_{\Omega\times T} := \int_T \int_\Omega |\langle \boldsymbol{v}\rangle - \langle \boldsymbol{v}_M\rangle|^2 \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}t, \tag{3.20}$$

$$\|\sigma_{\boldsymbol{v}} - \sigma_{\boldsymbol{v}_M}\|^2_{\Omega\times T} := \int_T \int_\Omega |\sigma_{\boldsymbol{v}} - \sigma_{\boldsymbol{v}_M}|^2 \mathrm{d}\boldsymbol{x}\,\mathrm{d}t. \tag{3.21}$$

In figure 13 we show the errors (3.20) and (3.21) with respect to DNS versus the number of modes M using different types of inner products. Not surprisingly, figure 13 shows that the expansion obtained using the mean inner product (case $h = 0$) cannot represent the standard deviation of the flow (see also figure 10). Similar behaviour is observed in figure 8(a) where it is shown that the random energy levels increase at high modes due to the inability of the modes $\boldsymbol{a}_j^{(0)}$ to properly represent the correlation functions appearing in (2.16). We remark that the error plots for the mean (figure 13(a) case $h = 0$) and for the standard deviation (figure 13(b) case $h = 2$) represent the minimum errors achievable, respectively, for the mean and the standard deviation given a certain number of modes M.

3.6. *Accuracy of orthogonal expansions as a function of the number of random inputs*

In this section we study the accuracy and sensitivity of the orthogonal representations (2.1) as a function of the number of random inputs forcing the system. The objective of this study is to prove the robustness of our approach in modelling high-dimensional random input vectors such as those arising from the representation of stochastic flow processes having small correlation length or in the design of fluid systems with a great number of uncertain parameters.

In table 4 we report normalized errors with respect to DNS data of low-dimensional representations using M modes as a function of the number random inputs N. We need to normalize these errors to make an effective comparison because if we increase the number of random variables N in the representation of the random boundary condition (3.9) for fixed correlation length, the standard deviation of the random inflow will increase as well (see figure 1b and figure 2b). In table 4 we report results for 2, 4 and 6 random variables.

We note that the number of random spatial modes required to achieve a desired level of accuracy does not seem to depend on the number of random inputs. To explain this important result, we recall that each random spatial mode contains a complete representation, which includes all these random variables. Therefore, if the dimensionality of the random input vector increases then the computation of

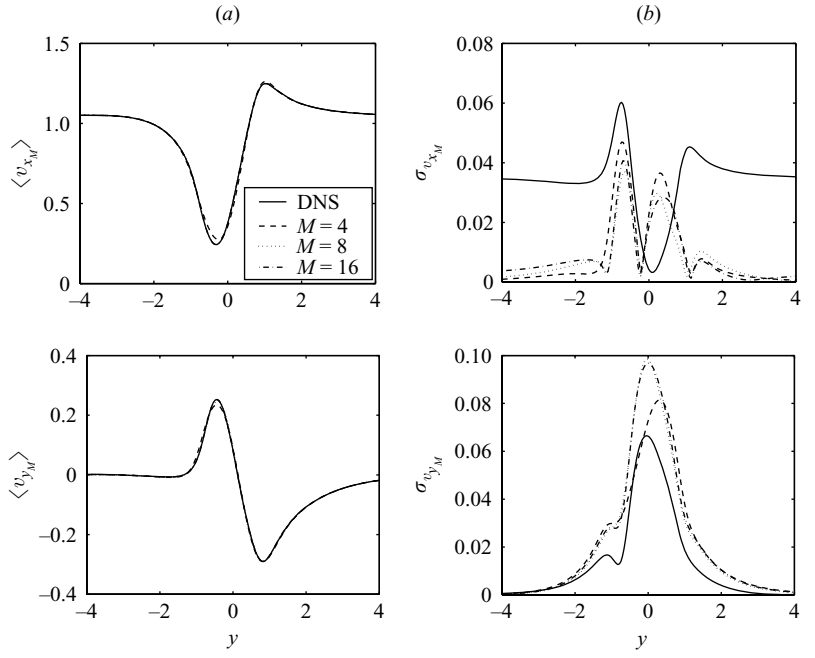

FIGURE 10. Mean (*a*) and standard deviation (*b*) of velocity components along the crossline $x = 3$ as a function of the number of modes M: convergence to stochastic DNS simulation. The orthogonal expansion is obtained in the mean inner product $\{,\}_0$.

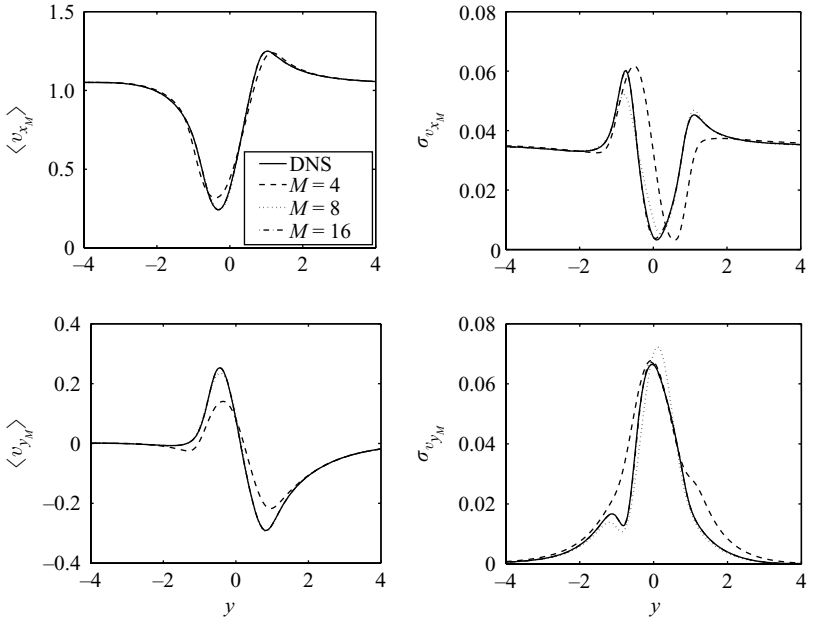

FIGURE 11. As figure 10 but the orthogonal expansion is obtained in the second-order moment inner product $\{,\}_1$.

each random mode requires more effort. (It can be proved that the number of terms is $P = (N + p)!/(N!p!)$, where N is the number of random inputs and p is the polynomial order of the functional chaos expansion representation). This leads to

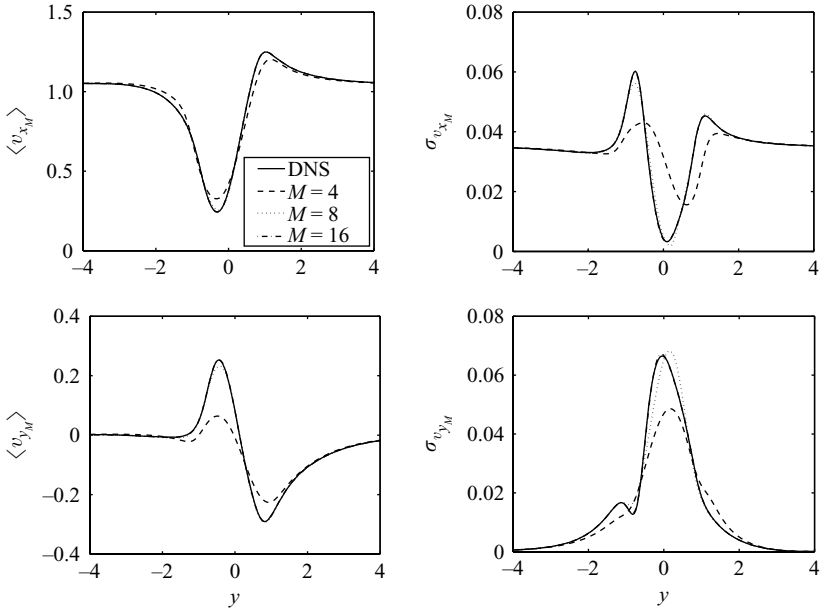

FIGURE 12. As figure 10 but the orthogonal expansion is obtained in the standard deviation inner product $\{,\}_2$.

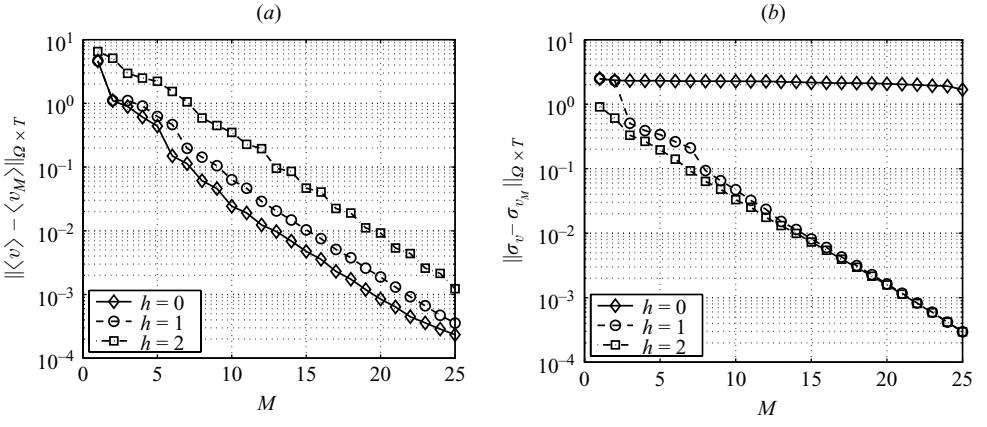

FIGURE 13. Error in the L^2 norm with respect to stochastic DNS of low-dimensional representation versus the number of modes: (*a*) error in mean; (*b*) error in the standard deviation. Shown are results obtained using different expansions.

substantial increases in computing; however, all the random modes can be determined offline in a pre-processing stage, so that they will be readily available for a real-time stochastic low-dimensional simulation. Moreover, one can resort to sparse tensor products or sparse quadrature grids in the multi-dimensional random space to reduce further the computational complexity for large N, e.g. see Gerstner & Griebel (1998), Novak & Ritter (1996, 1999). Also, new effective approaches for efficient nonlinear dimensionality reduction in random space have been proposed by different authors (Tenenbaum *et al.* 2000; Belkin & Niyogi 2003). As Ham *et al.* (2003) pointed out, many of these methods belong to the so-called kernel principal component analysis (Scholköpf *et al.* 1998), which makes use of a 'feature space' in order to construct a

		$\|\langle \boldsymbol{v} \rangle - \langle \boldsymbol{v}_M \rangle\|_{\Omega\times T}/\|\langle \boldsymbol{v} \rangle\|_{\Omega\times T}$					$\|\sigma_{\boldsymbol{v}} - \sigma_{\boldsymbol{v}_M}\|_{\Omega\times T}/\|\sigma_{\boldsymbol{v}}\|_{\Omega\times T}$				
					$h = 0$						
		M					M				
		5	10	15	20	25	5	10	15	20	25
N	1	0.67	0.037	0.0073	0.0013	0.00035	3.26	3.21	3.04	2.92	2.36
	2	0.66	0.037	0.0072	0.0013	0.00035	3.36	3.31	3.14	3.02	2.45
	4	0.66	0.037	0.0071	0.0013	0.00035	3.45	3.41	3.23	3.10	2.54
	6	0.66	0.037	0.0071	0.0013	0.00035	3.46	3.41	3.23	3.11	2.54
					$h = 1$						
		M					M				
		5	10	15	20	25	5	10	15	20	25
N	1	0.94	0.096	0.016	0.0029	0.00054	0.49	0.069	0.012	0.0024	0.00044
	2	0.94	0.096	0.016	0.0029	0.00054	0.50	0.070	0.012	0.0024	0.00044
	4	0.94	0.095	0.016	0.0028	0.00054	0.51	0.071	0.012	0.0025	0.00045
	6	0.94	0.095	0.016	0.0028	0.00054	0.51	0.071	0.012	0.0025	0.00045
					$h = 2$						
		M					M				
		5	10	15	20	25	5	10	15	20	25
N	1	3.88	0.81	0.18	0.029	0.0030	0.25	0.044	0.0096	0.0021	0.00042
	2	3.60	0.58	0.090	0.017	0.0022	0.28	0.047	0.011	0.0023	0.00044
	4	3.38	0.53	0.072	0.014	0.0019	0.29	0.050	0.011	0.0024	0.00045
	6	3.34	0.53	0.071	0.014	0.0019	0.29	0.050	0.011	0.0024	0.00045

TABLE 4. Percentage relative errors in the mean (left) and in the standard deviation (right) as a function of the number of random variables N forcing the system and the number of random modes M used for the low-dimensional representation. Shown are different results corresponding to $h = 0, 1, 2$.

kernel for capturing the nonlinear intrinsic properties of the ensembles of solutions. Before performing the standard eigen-decomposition of the covariance matrix which leads to dimensionality reduction, the procedure geometrically unfolds the manifold of the 'feature space', which can be also obtained by the so-called learning kernel methods (Weinberger, Sha & Saul 2004; Saul & Roweis 2003; Scholköpf & Smola 2002). In a polynomial chaos representation we already have an explicit analytic expression for such a curved manifold, which is given in polynomial form through the chaos basis $\Gamma_i(\boldsymbol{\xi})$. Therefore, the problem of nonlinear dimensionality reduction amounts to defining a minimum principle for the geodetic distance between points over this curved manifold, e.g. see Tenenbaum *et al.* (2000), Weinberger & Saul (2006). Geodetic distances over polynomial hypersurfaces can be expressed in analytical form because the metric of the surface can be computed explicitly. Therefore, it is possible to develop a general theory for nonlinear dimensionality reduction of polynomial chaos representations.

3.7. *Random projection of Navier–Stokes equations onto random spatial modes*

The deterministic nature of the temporal modes in the representation (2.1) opens the possibility of constructing stochastic low-dimensional models of random flows. In fact, the weak orthogonality property of random spatial modes can be used to define random projections in a quite straightforward manner. However, before proceeding with the computation and representation of the low-dimensional model for the wake,

we point out some important issues concerning the completeness of the projections. In the process of constructing the Galerkin system, one important consideration, which must be carefully checked, is that the approximated stochastic solution is consistent, i.e. it converges to the exact solution as the number of projections for the spectral representation increases to infinity. It is clear that the expansion obtained by decomposing the kernel $\mathscr{T}^{(0)}$ is not of this kind. In fact, it is essentially an expansion of the mean flow, which cannot obviously represent higher-order statistics if the system is nonlinear. In this particular application, this can be seen from the error plot reported in figure 13(*b*), in which it clearly appears that the case $h = 0$ cannot represent the standard deviation of the flow. Therefore, in considering the Navier–Stokes equations, here written in operatorial form as $\mathscr{N}(\boldsymbol{v}) = 0$, and constructing a low-dimensional model through an approximated solution $\boldsymbol{v}_M$, we must carefully check that

$$\|\mathscr{N}(\boldsymbol{v}) - \mathscr{N}(\boldsymbol{v}_M)\|^2 \to 0 \quad \text{as} \quad M \to \infty \tag{3.22}$$

in a proper norm; here, we require convergence in the $\|\cdot\|_h^2$ norms. If the representation of the stochastic solution $\boldsymbol{v}_M$ is not complete, then the condition (3.22) cannot be satisfied. This happens, for instance, when the representation obtained in the inner product $\{,\}_0$ is used as an approximated solution for the velocity field. In fact, as can be easily seen

$$\int_\Omega \langle \mathscr{N}(\boldsymbol{v}) - \mathscr{N}(\boldsymbol{v}_M)\rangle^2 \,\mathrm{d}\boldsymbol{x} \nrightarrow 0 \quad \text{as} \quad M \to \infty \tag{3.23}$$

because the operator $\mathscr{N}$ is nonlinear. The representation obtained in the inner product $\{,\}_1$ is complete and valid under the same hypothesis of the functional polynomial chaos theory.

In contrast, for the inner product $\{,\}_2$ we observe that we have convergence (in the L_2 sense) to the mean and to the standard deviation, but there might be some problems if the standard deviation of the flow is identically zero somewhere. The completeness of the representation obtained by decomposing the random field according to the inner product $\{,\}_2$ relies on its linear equivalence with the decomposition obtained in the inner product $\{,\}_1$.

In order to obtain a Galerkin system optimally convergent in the mean sense, we can use an oblique projection onto random spatial modes $\boldsymbol{a}_k^{(0)}$ in the form

$$\{\mathscr{N}(\boldsymbol{v}_M), \boldsymbol{a}_k^{(0)}\}_0 = 0. \tag{3.24}$$

We need to employ oblique projection because $\boldsymbol{v}_M$ cannot be represented using the expansion obtained in the inner product $\{,\}_0$. Therefore, we have to use another complete representation for $\boldsymbol{v}_M$ (for instance, the expansion obtained in the inner product $\{,\}_1$). Thus, the Galerkin system will include rotation operators in the form discussed in Appendix A (equation (A 3)).

Now, we consider the Navier–Stokes equations (3.1). A substitution of (3.14) into (3.1) yields

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{U}\cdot\nabla)\,\boldsymbol{U} + (\boldsymbol{U}\cdot\nabla)\,\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{U} + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u} = \nabla p + \frac{1}{Re}(\nabla^2\boldsymbol{U} + \nabla^2\boldsymbol{u}). \tag{3.25}$$

Next, we expand the fluctuating stochastic field $\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$ into our generalized series (3.17) and we perform a Galerkin projection onto the (non-normalized) strongly

divergence-free random spatial modes $\boldsymbol{a}_j^{(h)}$ according to the inner product $\{,\}_h$ ($h = 1, 2$). We obtain (hereafter, the summation convention on the pedices l and n is employed)

$$\left\{ \boldsymbol{a}_l^{(h)} \frac{\partial \psi_l^{(h)}}{\partial t} + (\boldsymbol{U}\cdot\nabla)\boldsymbol{U} + (\boldsymbol{U}\cdot\nabla)\boldsymbol{a}_l^{(h)}\psi_l^{(h)} + \left(\boldsymbol{a}_l^{(h)}\cdot\nabla\right)\boldsymbol{U} + \left(\boldsymbol{a}_l^{(h)}\cdot\nabla\right)\boldsymbol{a}_n^{(h)}\psi_l^{(h)}\psi_n^{(h)} + \nabla p - \frac{1}{Re}\left(\nabla^2\boldsymbol{U} + \psi_l\nabla^2\boldsymbol{a}_l^{(h)}\right),\ \boldsymbol{a}_m^{(h)} \right\}_h = 0\,. \quad (3.26)$$

The pressure term does not drop out because of the random inflow boundary condition. In general, it is well known that the pressure integral does not vanish and may not even be negligible. The need to model this term has already been anticipated by Holmes *et al.* (1996) from a Fourier-space representation. Also, Noack *et al.* (2005) pointed out the need for a pressure-term representation in empirical Galerkin models and modal energy flow analysis. In this paper, we do not represent the pressure integral in terms of the velocity field (see Noack *et al.* 2005 for details); instead, to simplify our analysis, we use the pressure field obtained from the stochastic DNS: From (3.26) the following system of ordinary differential equations is obtained:

$$\frac{\mathrm{d}\psi_m^{(h)}}{\mathrm{d}t} = \frac{1}{\mu_m^{(h)}}\left(-\mathscr{C}_m^{(h)} - \mathscr{L}_{ml}^{(h)}\psi_l^h - \mathscr{Q}_{lnm}^{(h)}\psi_l^h\psi_n^h\right), \quad (3.27)$$

where

$$\mathscr{C}_m^{(h)} := \left\{(\boldsymbol{U}\cdot\nabla)\boldsymbol{U},\ \boldsymbol{a}_m^{(h)}\right\}_h - \frac{1}{Re}\left\{\nabla^2\boldsymbol{U},\ \boldsymbol{a}_m^{(h)}\right\}_h + \left\{\nabla p,\ \boldsymbol{a}_m^{(h)}\right\}_h, \quad (3.28a)$$

$$\mathscr{L}_{ml}^{(h)} := \left\{(\boldsymbol{U}\cdot\nabla)\boldsymbol{a}_l^{(h)},\ \boldsymbol{a}_m^{(h)}\right\}_h + \left\{\left(\boldsymbol{a}_l^{(h)}\cdot\nabla\right)\boldsymbol{U},\ \boldsymbol{a}_m^{(h)}\right\}_h - \frac{1}{Re}\left\{\nabla^2\boldsymbol{a}_l^{(h)},\ \boldsymbol{a}_m^{(h)}\right\}_h, \quad (3.28b)$$

$$\mathscr{Q}_{lnm}^{(h)} := \left\{(\boldsymbol{a}_l^{(h)}\cdot\nabla)\boldsymbol{a}_n^{(h)},\ \boldsymbol{a}_m^{(h)}\right\}_h. \quad (3.28c)$$

The initial condition for $\psi_m^{(h)}$ is

$$\psi_m^{(h)}(0) = \frac{1}{\mu_m^{(h)}}\left\{\boldsymbol{u}(\boldsymbol{x}, 0; \boldsymbol{\xi}),\ \boldsymbol{a}_m^{(h)}(\boldsymbol{x}; \boldsymbol{\xi})\right\}_h. \quad (3.29)$$

Explicit expressions for the coefficients (3.28) corresponding to inner products $\{,\}_1$ and $\{,\}_2$ are given in Appendix C. Note that the projections $\{\cdot, \boldsymbol{a}_k^{(h)}\}_h$ ($h = 1, 2$) allow us to obtain two different specialized sets of equations which take different 'sections' of the solution domain. In particular, $\{\cdot, \boldsymbol{a}_k^{(2)}\}_2$ generates an optimal temporal dynamics for modelling of the standard deviation, while $\{\cdot, \boldsymbol{a}_k^{(1)}\}_1$ generates an optimal temporal dynamics for modelling of the second-order moment.

3.8. *Low-dimensional simulation of the random wake*

In figures 14 and 15 we compare the evolution of the temporal modes extracted from the stochastic DNS data to the evolution predicted by the system (3.27). As already mentioned, two different types of projections are considered: in figure 14 we report results using $\{\cdot, \boldsymbol{a}_k^{(1)}\}_1$ while in figure 15 we show results obtained using $\{\cdot, \boldsymbol{a}_k^{(2)}\}_2$. In both cases, we see that the accuracy is increasing as more projections are employed in the representation.

In figure 16 we show the errors in the L_2 spatial norm as a function of time for the mean and the standard deviation of the velocity fields obtained from the

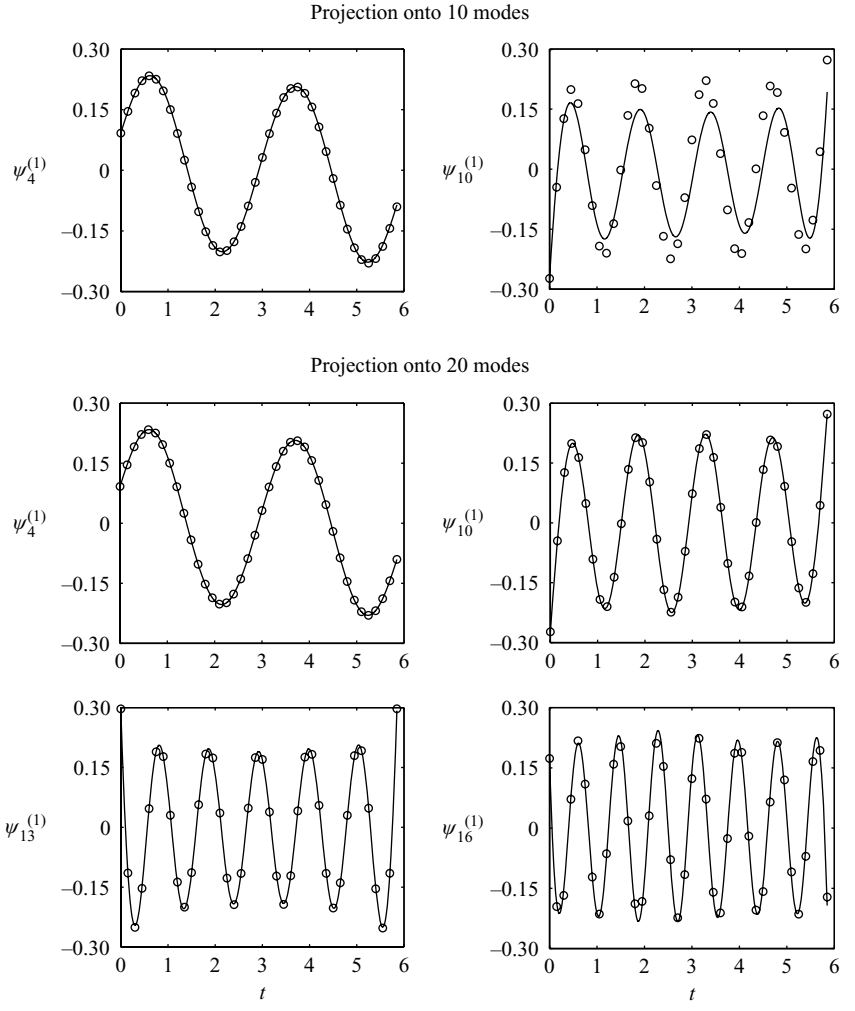

FIGURE 14. Case $h = 1$ (second-order-moment inner product). Comparison between temporal evolution predicted by the system (3.27) (——) and the DNS-based evolution (○).

low-dimensional model. These errors are defined as follows:

$$\|\langle \boldsymbol{v} \rangle - \langle \boldsymbol{v}_M \rangle\|_{\Omega}^2 := \int_{\Omega} |\langle \boldsymbol{v} \rangle - \langle \boldsymbol{v}_M \rangle|^2 \, \mathrm{d}\boldsymbol{x} \,, \tag{3.30}$$

$$\|\sigma_{\boldsymbol{v}} - \sigma_{\boldsymbol{v}_M}\|_{\Omega}^2 := \int_{\Omega} |\sigma_{\boldsymbol{v}} - \sigma_{\boldsymbol{v}_M}|^2 \, \mathrm{d}\boldsymbol{x} \,. \tag{3.31}$$

We report results for the projections $\{\cdot, \boldsymbol{a}_k^{(1)}\}_1$ and $\{\cdot, \boldsymbol{a}_k^{(2)}\}_2$ as a function of the number of modes M. Specifically, we have used 10, 20 and 30 modes for the simulation of the low-dimensional system and correspondingly 5, 10 and 20 modes for the computation of the errors. The Galerkin low-dimensional model allows computation of the relevant statistics (mean and standard deviation) of the system at some non-observed/off-design states, provided that the information for such a prediction is contained in the snapshot set (Burkardt & Webster 2007).

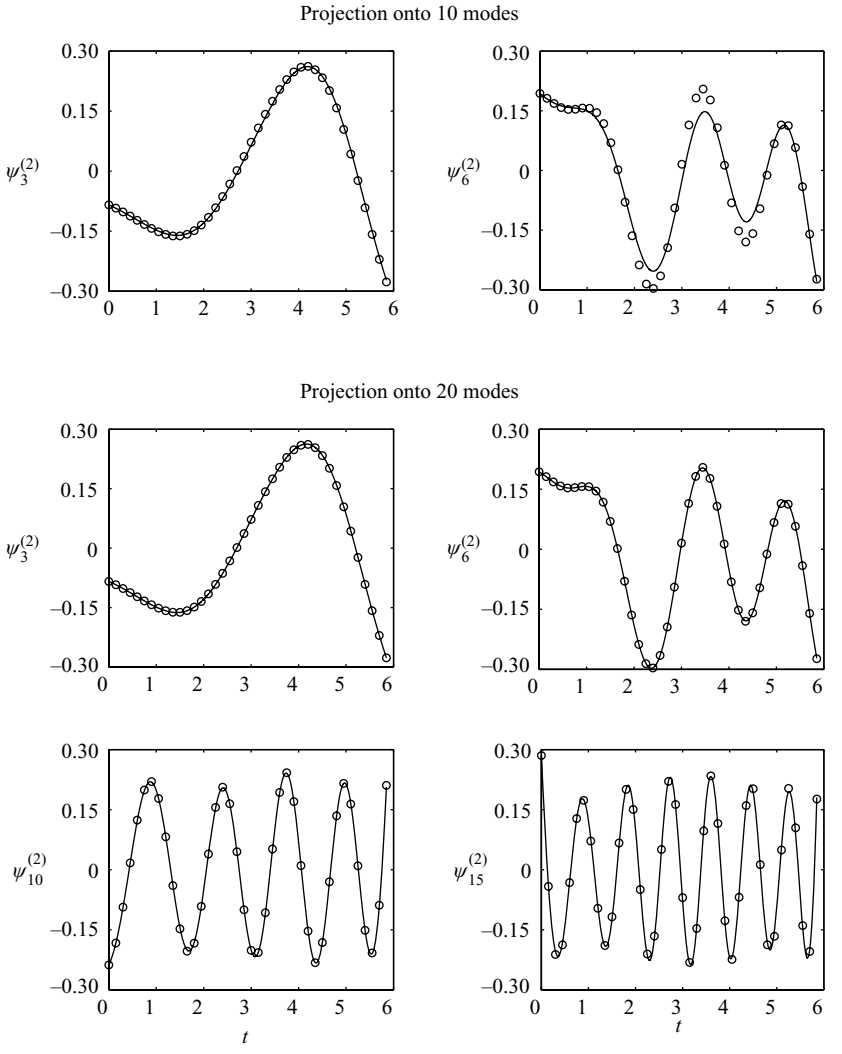

FIGURE 15. Case $h = 2$ (standard deviation inner product). Comparison between temporal evolution predicted by the system (3.27) (——) and the DNS-based evolution (○).

4. Summary

We have studied the random laminar wake past a circular cylinder corresponding to a random inflow boundary condition, which is represented by a wide-sense stationary Gaussian stochastic process. We simulated the random flow using an accurate stochastic DNS based on the Wiener–Hermite functional representation. Subsequently, we decomposed the random velocity field according to a new expansion developed in § 2.1, and in § 3.7 we constructed a low-dimensional stochastic model of the random wake though a Galerkin projection onto random spatial modes. These new tools allow us to analyse the flow from a new perspective, exploiting the relation between the random dynamics in space–time and its hierarchical set of random structures. They also give us the possibility of constructing low-dimensional stochastic models, which can be useful for controlling flow systems operating under uncertain conditions.

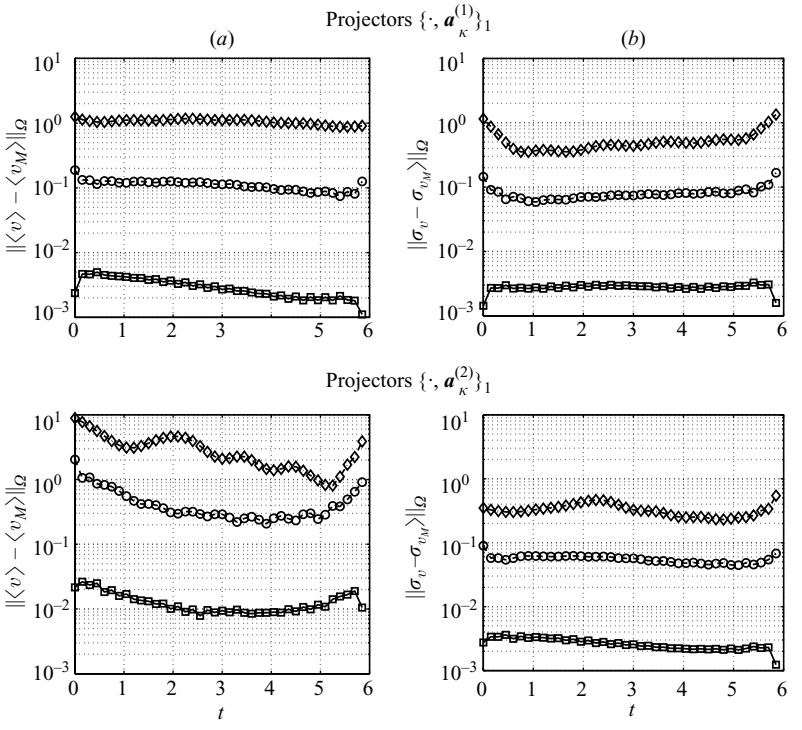

FIGURE 16. Errors for the low-dimensional simulation as a function of time. (*a*): Error in the mean (3.30); (*b*): error in the standard deviation (3.31). Shown are results obtained using different types of projections and a different number of modes in the simulation: 5 modes (⋄); 10 modes (○); 20 modes (□).

We have shown that the accuracy of the orthogonal expansion does not depend on the number of random inputs forcing the system although the computational complexity of each mode depends heavily on them ('curse of dimensionality'). However, all random modes can be computed offline so that they will be readily available for a real-time stochastic low-dimensional simulation. In addition, new methods that employ sparse multi-level quadrature formulae for multi-dimensional problems can be employed to expedite the offline computations while new learning algorithms can be exploited for further nonlinear dimensionality reduction.

Partial support for his work was provided by ONR, NSF and AFOSR. The second and third authors would like to acknowledge additional support by the ONR ERSD Consortium. The computations were performed on NSF's supercomputing centers (NCSA, PSC and SDSC).

Appendix A. Connections between different expansions through oblique projections

The decomposition of the kernels $\mathscr{T}^{(h)}(t, t')$ for $h = 0, 1, 2$ allows us to identify three different types of random modes and associated orthogonal projections in random space, which give rise to three expansions optimally convergent in the mean sense ($h = 0$), in the second-order moment sense ($h = 1$), and in the standard deviation sense ($h = 2$). Given a complete orthogonal representation of the stochastic field

$\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$, it is easy to transform it into another one by a suitable rotation of the functional space. To this end, we let

$$\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi}) = \sum_{i=1}^{M} \psi_i^{(1)}(t)\, \boldsymbol{a}_i^{(1)}(\boldsymbol{x}; \boldsymbol{\xi}) \,. \tag{A 1}$$

According to (2.9) we can transform the temporal modes from this representation into another one (for instance the $h = 2$) as

$$\psi_k^{(2)}(t) = \sum_{m=1}^{M} [\boldsymbol{R}_{1\to 2}]_{km}\, \psi_m^{(1)}(t) \,. \tag{A 2}$$

The operator

$$[\boldsymbol{R}_{1\to 2}]_{km} := \frac{\{\boldsymbol{a}_m^{(1)}, \boldsymbol{a}_k^{(2)}\}_2}{\mu_k^{(2)}} \tag{A 3}$$

is orthogonal (i.e. it is a rotation) since both $\psi_k^{(1)}$, $\psi_k^{(2)}$ are orthonormal with respect to the same inner product $(\,,\,)_T$. This rotation transforms the representation in such a way that the new decomposition is optimally convergent in the standard deviation sense. We also have the rotation from the representation $h = 1$ to the representation $h = 0$

$$[\boldsymbol{R}_{1\to 0}]_{km} := \frac{\{\boldsymbol{a}_m^{(1)}, \boldsymbol{a}_k^{(0)}\}_0}{\mu_k^{(0)}} \,. \tag{A 4}$$

Appendix B. Polynomial chaos representations

We assume that we have available a functional polynomial chaos expansion up to order P for the random field $\boldsymbol{u}(\boldsymbol{x}, t; \boldsymbol{\xi})$

$$\boldsymbol{u}(\boldsymbol{x}, t; \xi) = \sum_{l=0}^{P} \widehat{\boldsymbol{u}}_l(\boldsymbol{x}, t)\, \Gamma_l(\boldsymbol{\xi}) \,, \tag{B 1}$$

where $\{\Gamma_l(\boldsymbol{\xi})\}$ is the generalized chaos basis (Xiu & Karniadakis 2002; Wan & Karniadakis 2006*b*), which is orthogonal in the L_2 sense with respect to the joint probability density of the random input vector $\boldsymbol{\xi}$. We have denoted such a probability density by $W(\boldsymbol{\xi})$. The temporal autocorrelations $\mathscr{T}^{(h)}(t, t')$ defined in (2.12) have the following representations

$$\mathscr{T}^{(0)}(t, t') = \int_{\Omega} \widehat{\boldsymbol{u}}_0(\boldsymbol{x}, t) \cdot \widehat{\boldsymbol{u}}_0(\boldsymbol{x}, t')\, \mathrm{d}\boldsymbol{x} \,, \tag{B 2}$$

$$\mathscr{T}^{(1)}(t, t') = \sum_{l=0}^{P} \langle \Gamma_l^2 \rangle \int_{\Omega} \widehat{\boldsymbol{u}}_l(\boldsymbol{x}, t) \cdot \widehat{\boldsymbol{u}}_l(\boldsymbol{x}, t')\, \mathrm{d}\boldsymbol{x}, \tag{B 3}$$

$$\mathscr{T}^{(2)}(t, t') = \sum_{l=1}^{P} \langle \Gamma_l^2 \rangle \int_{\Omega} \widehat{\boldsymbol{u}}_l(\boldsymbol{x}, t) \cdot \widehat{\boldsymbol{u}}_l(\boldsymbol{x}, t')\, \mathrm{d}\boldsymbol{x}. \tag{B 4}$$

Having these expressions we can solve the eigenvalue problem (2.11) to obtain the eigenvalues $\mu_k^{(h)}$ and the temporal modes $\psi_k^{(h)}$. Subsequently, we consider the following

chaos representation for the normalized spatial modes:

$$\boldsymbol{\Phi}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) = \sum_{l=0}^{P} \widehat{\boldsymbol{\Phi}}_{kl}^{(h)}(\boldsymbol{x})\, \Gamma_l(\boldsymbol{\xi}) \,, \tag{B 5}$$

where

$$\widehat{\boldsymbol{\Phi}}_{kj}^{(h)}(\boldsymbol{x}) = \frac{1}{\sqrt{\mu_k^{(h)}}} \int_T \widehat{\boldsymbol{u}}_j(\boldsymbol{x},t)\, \psi_k^{(h)}(t)\, \mathrm{d}t \,. \tag{B 6}$$

The latter equality follows from (2.10), (B 1) and (B 5). Equivalently, we obtain the following expression for the non-normalized random spatial modes $\boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi})$:

$$\boldsymbol{a}_k^{(h)}(\boldsymbol{x};\boldsymbol{\xi}) = \sum_{l=0}^{P} \widehat{\boldsymbol{a}}_{kl}^{(h)}(\boldsymbol{x})\, \Gamma_l(\boldsymbol{\xi}) \,, \qquad \widehat{\boldsymbol{a}}_{kj}^{(h)}(\boldsymbol{x}) = \frac{1}{\mu_k^{(h)}} \int_T \widehat{\boldsymbol{u}}_j(\boldsymbol{x},t)\, \psi_k^{(h)}(t)\, \mathrm{d}t \,. \tag{B 7}$$

Associated with these modes, we define a 'random energy' (2.16), whose representation in the polynomial chaos basis is

$$\theta_k^{(h)}(\boldsymbol{\xi}) = \sum_{l,m=0}^{P} \Gamma_l(\boldsymbol{\xi})\, \Gamma_m(\boldsymbol{\xi}) \int_\Omega \widehat{\boldsymbol{a}}_{kl}^{(h)}(\boldsymbol{x}) \cdot \widehat{\boldsymbol{a}}_{km}^{(h)}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \,. \tag{B 8}$$

The mean and the second-order moment of these quantities are readily obtained as

$$\left\langle \theta_k^{(h)} \right\rangle = \sum_{l=0}^{P} \left\langle \Gamma_k^2 \right\rangle \int_\Omega \widehat{\boldsymbol{a}}_{kl}^{(h)} \cdot \widehat{\boldsymbol{a}}_{kl}^{(h)}\, \mathrm{d}\boldsymbol{x} \,, \tag{B 9a}$$

$$\left\langle \theta_k^{2(h)} \right\rangle = \sum_{l,m,n,p=0}^{P} \left\langle \Gamma_l \Gamma_m \Gamma_n \Gamma_p \right\rangle \int_\Omega \widehat{\boldsymbol{a}}_{kl}^{(h)} \cdot \widehat{\boldsymbol{a}}_{km}^{(h)}\, \mathrm{d}\boldsymbol{x} \int_\Omega \widehat{\boldsymbol{a}}_{kn}^{(h)} \cdot \widehat{\boldsymbol{a}}_{kp}^{(h)}\, \mathrm{d}\boldsymbol{x} \,, \tag{B 9b}$$

where the fields $\widehat{\boldsymbol{a}}_{kl}$ are defined in (B 7). The correlation coefficient (2.17) between different energy levels has a similar representation.

Appendix C. Coefficients arising from the Galerkin projection of the Navier–Stokes equations onto random spatial modes

Here, we obtain explicit expressions for the coefficients arising from the Galerkin projection using different projectors $\{\cdot, \boldsymbol{a}_j^{(h)}\}_h$ $(h = 1, 2)$. The pressure term does not drop out because of the random inflow boundary condition. We denote by $\partial\Omega$ the set of boundaries of the computational domain; it includes inflow, outflow, cylinder surface and the two (upper and lower) sides where periodicity is imposed. We employ integration by parts and the Gauss theorem to transform the pressure term into an integral along these boundaries. Using the fact that the random spatial modes are strongly divergence-free (see (3.16)) we obtain

$$\begin{aligned} \int_\Omega \boldsymbol{a}_m^{(h)} \cdot \nabla p \mathrm{d}\boldsymbol{x} &= -\int_\Omega \nabla \cdot \boldsymbol{a}_m^{(h)} p\, \mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} p \left(\boldsymbol{a}_m^{(h)} \cdot \boldsymbol{n}\right) \mathrm{d}\boldsymbol{x} \\ &= 0 + \int_{\partial\Omega} p \left(\boldsymbol{a}_m^{(h)} \cdot \boldsymbol{n}\right) \mathrm{d}\boldsymbol{x} \,, \end{aligned} \tag{C 1}$$

where $\boldsymbol{n}$ denotes the outward normal unit vector at the boundaries $\partial\Omega$. Depending on the choice of the projection $\{\cdot, \boldsymbol{a}_j^{(h)}\}_h$ $(h = 1, 2)$, the last integral has different expressions as shown in the following subsections § C.1 and § C.2.

C.1. *Representation of the Galerkin system using the projectors* $\{\cdot, \boldsymbol{a}_j^{(1)}\}_1$

In the second-order moment inner product (2.4) we obtain

$$\mathscr{C}_m^{(1)} = \int_\Omega (\boldsymbol{U}\cdot\nabla)\boldsymbol{U}\cdot\langle \boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x} - \frac{1}{Re}\int_\Omega \nabla^2\boldsymbol{U}\cdot\langle \boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x} + \int_{\partial\Omega}\langle p\boldsymbol{a}_m^{(1)}\rangle\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{x}\,, \quad (C\,2)$$

$$\mathscr{L}_{ml}^{(1)} = \int_\Omega \langle (\boldsymbol{U}\cdot\nabla)\boldsymbol{a}_l^{(1)}\cdot\boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x} + \int_\Omega \langle (\boldsymbol{a}_l^{(1)}\cdot\nabla)\boldsymbol{U}\cdot\boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x} - \frac{1}{Re}\int_\Omega \langle \nabla^2\boldsymbol{a}_l^{(1)}\cdot\boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x}\,, \quad (C\,3)$$

$$\mathscr{Q}_{lnm}^{(1)} = \int_\Omega \langle (\boldsymbol{a}_l^{(1)}\cdot\nabla)\boldsymbol{a}_n^{(1)}\cdot\boldsymbol{a}_m^{(1)}\rangle \,\mathrm{d}\boldsymbol{x}\,. \quad (C\,4)$$

The functional polynomial chaos representation of these coefficients takes the form (Einstein's summation convention on repeated indexes from 0 to P is assumed hereafter)

$$\mathscr{C}_m^{(1)} = \int_\Omega (\boldsymbol{U}\cdot\nabla)\boldsymbol{U}\cdot\widehat{\boldsymbol{a}}_{m0}^{(1)}\,\mathrm{d}\boldsymbol{x} + \frac{1}{Re}\int_\Omega \nabla\boldsymbol{U}\cdot\nabla\widehat{\boldsymbol{a}}_{m0}^{(1)}\,\mathrm{d}\boldsymbol{x}, + \langle \Gamma_n^2\rangle\int_{\partial\Omega}\widehat{p}_n\widehat{\boldsymbol{a}}_{mn}^{(1)}\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{x}\,, \quad (C\,5)$$

$$\mathscr{L}_{ml}^{(1)} = \langle \Gamma_n^2\rangle\int_\Omega (\boldsymbol{U}\cdot\nabla)\widehat{\boldsymbol{a}}_{ln}^{(1)}\cdot\widehat{\boldsymbol{a}}_{mn}^{(1)}\,\mathrm{d}\boldsymbol{x} + \langle \Gamma_n^2\rangle\int_\Omega (\widehat{\boldsymbol{a}}_{ln}^{(1)}\cdot\nabla)\boldsymbol{U}\cdot\widehat{\boldsymbol{a}}_{mn}^{(1)}\,\mathrm{d}\boldsymbol{x} + \langle \Gamma_n^2\rangle\frac{1}{Re}\int_\Omega \nabla\widehat{\boldsymbol{a}}_{ln}^{(1)}\cdot\nabla\widehat{\boldsymbol{a}}_{mn}^{(1)}\,\mathrm{d}\boldsymbol{x}, \quad (C\,6)$$

$$\mathscr{Q}_{lnm}^{(1)} = \langle \Gamma_p\Gamma_q\Gamma_r\rangle\int_\Omega (\widehat{\boldsymbol{a}}_{lp}^{(1)}\cdot\nabla)\widehat{\boldsymbol{a}}_{nq}^{(1)}\cdot\widehat{\boldsymbol{a}}_{mr}^{(1)}\,\mathrm{d}\boldsymbol{x}\,. \quad (C\,7)$$

C.2. *Representation of the Galerkin system using the projectors* $\{\cdot, \boldsymbol{a}_j^{(2)}\}_2$

From the inner product definitions (2.3), (2.4) and (2.5), for every pair of random fields $\boldsymbol{b}(\boldsymbol{x};\boldsymbol{\xi})$, $\boldsymbol{c}(\boldsymbol{x};\boldsymbol{\xi})$ in the spatial domain we have

$$\{\boldsymbol{b},\boldsymbol{c}\}_2 = \{\boldsymbol{b},\boldsymbol{c}\}_1 - \{\boldsymbol{b},\boldsymbol{c}\}_0\,. \quad (C\,8)$$

Therefore, the coefficients arising from Galerkin projection using the 'standard deviation' inner product (2.5) are

$$\mathscr{C}_m^{(2)} = \int_{\partial\Omega}\left(\langle p\boldsymbol{a}_m^{(2)}\rangle - \langle p\rangle\langle \boldsymbol{a}_m^{(2)}\rangle\right)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{x}\,, \quad (C\,9)$$

$$\mathscr{L}_{ml}^{(2)} = \int_\Omega \left(\langle (\boldsymbol{U}\cdot\nabla)\boldsymbol{a}_l^{(2)}\cdot\boldsymbol{a}_m^{(2)}\rangle - \langle (\boldsymbol{U}\cdot\nabla)\boldsymbol{a}_l^{(2)}\rangle\cdot\langle \boldsymbol{a}_m^{(2)}\rangle\right)\mathrm{d}\boldsymbol{x} + \int_\Omega \left(\langle (\boldsymbol{a}_l^{(2)}\cdot\nabla)\boldsymbol{U}\cdot\boldsymbol{a}_m^{(2)}\rangle - \langle (\boldsymbol{a}_l^{(2)}\cdot\nabla)\boldsymbol{U}\rangle\cdot\langle \boldsymbol{a}_m^{(2)}\rangle\right)\mathrm{d}\boldsymbol{x} \quad (C\,10)$$

$$- \frac{1}{Re}\int_\Omega \left(\langle \nabla^2\boldsymbol{a}_l^{(2)}\cdot\boldsymbol{a}_m^{(2)}\rangle - \langle \nabla^2\boldsymbol{a}_l^{(2)}\rangle\cdot\langle \boldsymbol{a}_m^{(2)}\rangle\right)\mathrm{d}\boldsymbol{x}\,, \quad (C\,11)$$

$$\mathscr{Q}_{lnm}^{(2)} = \int_\Omega \left(\langle (\boldsymbol{a}_l^{(2)}\cdot\nabla)\boldsymbol{a}_n^{(2)}\cdot\boldsymbol{a}_m^{(2)}\rangle - \langle (\boldsymbol{a}_l^{(2)}\cdot\nabla)\boldsymbol{a}_n^{(2)}\rangle\cdot\langle \boldsymbol{a}_m^{(2)}\rangle\right)\mathrm{d}\boldsymbol{x}\,. \quad (C\,12)$$

The functional chaos expansion representation of these coefficients can be obtained by substitution of (B 7) into the above expressions.

REFERENCES

ACHARJEE, S. & ZABARAS, N. 2006 A concurrent model reduction approach on spatial and random domains for the solution of stochastic PDEs. *Intl J. Numer. Anal. Model.* **66** (12), 1934–1954.

Amit, D. J. & Martín-Mayor, V. 2005 *Field Theory, The Renormalization Group and Critical Phenomena*, 3rd Edn. World Scientific.

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

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

Belkin, M. & Niyogi, P. 2003 Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**, 1373–1396.

Blanchard, G., Bousquet, O. & Zwald, L. 2007 Statistical properties of kernel principal component analysis. *Mach. Learn.* **66**, 259–294.

Bodner, S. E. 1969 Turbulence theory with a time-varying Wiener–Hermite basis. *Phys. Fluids* **12** (1), 33–38.

Burkardt, J. & Webster, C. 2007 Reduced order modelling of some nonlinear stochastic partial differential equations. *Intl J. Numer. Anal. Model.* **4** (3–4), 368–391.

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** (10), 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.

Doostan, A., Ghanem, R. & Red-Horse, J. 2007 Stochastic model reduction for chaos representation. *Comput. Meth. Appl. Mech. Engng* **196**, 3951–3966.

Dozier, R. B. & Silverstein, J. W. 2007 On the empirical distribution of eigenvalues of large dimensional information plus-noise type matrices. *J. Multivariate Anal.* **98** (4), 678–694.

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

Gerstner, T. & Griebel, M. 1998 Numerical integration using sparse grids. *Numer. Algorithms* **18** (3–4), 209–232.

Ghanem, R. G. & Spanos, P. D. 1998 *Stochastic Finite Elements: A Spectral Approach*. Springer.

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. 2006 On the empirical distribution of eigenvalues of a Gram matrix with a given variance profile. *Ann. l'Inst. Henri Poincaré (B), Probability Statist.* **42** (6), 649–670.

Ham, J., Lee, D. D., Mika, S. & Scholköpf, B. 2003 A kernel view of the dimensionality reduction of manifolds. *Tech. Rep.* TR-110. Max Plank Institute for Biological Cybernetics.

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 Conf. on Learning Theory, COLT 2004, Banff, Canada, July 1–4, 2004, Proc.* (ed. J. Shawe-Taylor & Y. Singer), pp. 579–593. Springer.

Jenssen, R., Eltoft, T., Girolami, M. & Erdogmus, D. 2007 Kernel maximum entropy data transformation and an enhanced spectral clustering algorithm. In *Advances in Neural Information Processing Systems (NIPS) 19*, pp. 633–640. MIT Press.

Kamiński, M. & Carey, G. F. 2005 Stochastic perturbation-based finite element approach to fluid flow problems. *Int. J. Numer. Meth. Heat Fluid FLow* **15** (7), 671–697.

Karniadakis, G. E. & Sherwin, S. 2005 *Spectral/hp Element Methods for CFD*, 2nd Edn. Oxford University Press.

Kato, T. 1995 *Perturbation Ttheory for Linear Operators*, 4th Edn. Springe.

Lee, C. P., Meecham, W. C. & Hogge, H. D. 1982 Application of the Wiener–Hermite expansion to turbulence of moderate Reynolds number. *Phys. Fluids* **25** (8), 1322–1327.

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

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

Meecham, W. C. & Jeng, D. T. 1968 Use of the Wiener–Hermite expansion for nearly normal turbulence. *J. Fluid Mech.* **32** (2), 225–249.

Meecham, W. C. & Siegel, A. 1964 Wiener–Hermite expansion in model turbulence at large Reynolds numbers. *Phys. Fluids* **7** (8), 1178–1190.

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

Novak, E. & Ritter, K. 1996 High-dimensional integration of smooth functions over cubes. *Numer. Math.* **75**, 79–97.

Novak, E. & Ritter, K. 1999 Simple cubature formulas with high polynomial exactness. *Constructive Approximation* **15**, 499–522.

Paiva, A. R. C., Xu, J. & Principe, J. C. 2006 Kernel principal components are maximum entropy projections. In *Proc. 6th Intl Conf. on Independent Component Analysis and Blind Signal Separation (ICA), Charleston, SC, USA*, pp. 846–853.

Rempfer, D. 2003 Low-dimensional modelling and numerical simulation of transition in simple shear flows. *Annu. Rev. Fluid. Mech.* **35**, 229–265.

Rényi, A. 1961 On measures of information and entropy. In *Proc. 4th Berkeley Symposium on Mathematics, Statistics and Probability*.

Saul, K. L. & Roweis, S. T. 2003 Think globally, fit locally: unsupervised learning of low dimensional manifolds. *J. Machine Learning Res.* **4**, 119–155.

Scholköpf, B. & Smola, A. J. 2002 *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press.

Scholköpf, B., Smola, A. J. & Müller, K. R. 1998 Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation* **10**, 1299–1319.

Segall, A. & Kailath, T. 1976 Orthogonal functionals of independent-increment processes. *IEEE Trans. Inf. Theory* **22** (3), 287–298.

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

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

Tenenbaum, J. B., de Silva, V. & Langford, J. C. 2000 A global geometric framework for nonlinear dimensionality reduction. *Science* **290**, 2319–2323.

Venturi, D. 2006 On proper orthogonal decomposition of randomly perturbed fields with applications to flow past a cylinder and natural convection over a horizontal plate. *J. Fluid Mech.* **559**, 215–254.

Wan, X. & Karniadakis, G. E. 2006*a* Long term behavior of polynomial chaos in stochastic flow simulations. *Comput. Meth. Appl. Mech. Engng* **195**, 5582–5596.

Wan, X. & Karniadakis, G. E. 2006*b* Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.* **28** (3), 901–928.

Webster, C. 2007 Sparse collocation techniques for the numerical solution of stochastic partial differential equations. PhD thesis, The Florida State University.

Weinberger, K. Q. & Saul, L. K. 2006 Unsupervised learning of image manifolds by semidefinite programming. *Intl J. Computer Vision* **70** (1), 77–90.

Weinberger, K. Q., Sha, F. & Saul, L. K. 2004 Learning a kernel matrix for nonlinear dimensionality reduction. In *Proc. 21st Int Conf. on Machine Learning, Banff, Alberta, Canada*, p. 106.

Wiener, N. 1966 *Nonliner Problems in Random Theory*. MIT Press.

Xiu, D. & Karniadakis, G. E. 2002 The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.* **24** (2), 619–644.

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

Zdravkovich, M. M. 1997 *Flow around Circular Cylinders*, vol. 1,2. Oxford Univ. Press.

Zinn-Justin, J. 2002 *Quantum Field Theory and Critical Phenomena*, 4th Edn. Oxford University Press.