J. Fluid Mech., page 1 of 27
doi:10.1017/S0022112011000322

Realizing turbulent statistics

JÉRÔME HŒPFFNER†, YOSHITSUGU NAKA AND KOJI FUKAGATA

Department of Mechanical Engineering, Keio University, Yokohama 223-8522, Japan

(Received 20 November 2009; revised 23 November 2010; accepted 16 January 2011)

How to design an artificial inflow condition in simulations of the Navier–Stokes equation that is already fully turbulent? This is the turbulent inflow problem. This first question is followed by: How much of the true turbulence must be reproduced at the inflow? We present a technique able to produce a random field with the exact two-point two-time covariance of a given reference turbulent flow. It is obtained as the output of a linear filter fed with white noise. The method is illustrated on the simulation of a turbulent free shear layer. The filter coefficients are obtained from the solution of the Yule–Walker equation, and the computation can be performed efficiently using a recursive solution procedure. The method should also be useful in the study of flow receptivity, when the processes of transition to turbulence are sensitive to the perturbation environment.

Key words: shear layer turbulence, turbulence simulation

1. Introduction

In numerical studies of fluid flow, one usually focuses on a domain in time and space, which contains the features whose mechanisms are to be understood: this is the system to be studied, delimited by its boundaries. Inside the computational domain, the flow dynamics is modelled by the Navier–Stokes equations, and boundary conditions are imposed at the boundaries. The boundary conditions are here to reflect the impact on the studied domain of outside processes: how the rest of the world acts on our system. Suppose that we are interested in the wake of a cylinder in a fully turbulent flow, as could be realized experimentally by placing a turbulence-generating grid in a wind tunnel, upstream of the cylinder. Certainly the wake of the cylinder will be affected by the particular properties of the incoming turbulent flow: spatial and temporal scales of the turbulence as compared to the cylinder radius and the frequency of the vortices in the cylinder wake. For this study, one would like to concentrate computational effort on a proper resolution of the flow about and behind the cylinder, and certainly not spend computational resources on a precise computation of the transition to turbulence throughout the grid. The purpose of the present paper is to build an inflow condition which is already fully turbulent. This inflow condition – which is the input to our simulation – is a model for the output of the complex processes of transition to turbulence.

This topic is reviewed in Keating *et al.* (2004): '*A priori* and *a posteriori* tests of inflow conditions for large-eddy simulation'. We can read in their abstract that

† Email address for correspondence: jerome.hoepffner@upmc.fr

'. . .consistent with previous investigations, it is found that the spectral content of the inflow velocity is important. Inflow conditions based on random noise, or small-scale eddies only, dissipate quickly.'

Turbulent flows are advantageously characterized by their statistics; it is on a statistical description that we would like to base our inflow condition. Two questions can now be raised: Which statistics must be accurately reproduced? And how can we generate a field evolving in time with precisely these statistics? The first question is that of modelling and the second is that of the realization. It is interesting to note that in Keating *et al.* (2004), a distinction is not drawn between the technical details of the generation of the inflow and a question of most fundamental nature: How much of the true turbulence must be reproduced? This confusion probably results from the lack of a technique – once the needs are clearly identified – to produce an artificial field that contains faithfully that much of the true turbulence.

Just reproducing the mean-flow profile is not enough: all information on the fluctuations is lost. The mean-flow profile plus some random noise is still not enough: uncorrelated noise does not have the structure that will sustain the dynamics. Description of the magnitude, spatial and temporal spectrum of the fluctuations is contained in the two-point two-time covariance

$$P_{ij}(r, t, r', t') = \overline{u_i(r, t)u_j(r', t')}, \tag{1.1}$$

where the overbar denotes the statistical average. These data describe how the velocity in the direction i at one position r and time t is related to the velocity in the direction j at another position r' and another time t'. We will focus on these second-order statistics. The question of the realization, on the other hand, will be the central part of this paper. We will show that it is possible to generate a random spatiotemporal field by filtering a field of uncorrelated noise; this artificial field will be tuned to have the exact two-point two-time covariance of the turbulent field. Our practical concern will be to choose the structure of the filter and compute efficiently its coefficients. We will use an auto-regressive filter, whose coefficients will be obtained from the numerical solution of the Yule–Walker equation.

The development of the filtering technique used in the present paper took place in the context of signal modelling. For an excellent textbook on this, see Hayes (1996): *Statistical Digital Signal Processing and Modeling*. The widespread use of digital signal filtering opened the door to sophisticated designs. Signal modelling is one application of filter design where one tries to reproduce a given signal by the output of a filter with, for instance, impulses at the input. This type of method is used, for instance, for data compression. Consider, for instance, an audio signal consisting of a pure sine wave. This signal can be sampled in time, each time sample carrying the information of the signal magnitude. On the other hand, a sine wave is exactly described by its phase, amplitude, frequency and duration. One can thus imagine an audio compression based on piecewise Fourier interpolation. The audio device should then have the capacity to translate back the wave information to a temporal signal. We can generalize the example of sine-wave interpolants by including as well an information of growth of amplitude in time. Pursuing further this process of generalization, we can use as interpolant the impulse responses of linear filters. Of course, growing or decaying sine waves are simple special cases of such signals. The parametrization of the signal is no longer amplitude and frequency, but it is the filter's numerical coefficients. Let us consider a commonplace family of filters, known as auto-regressive moving average

(ARMA) with input sequence v^i and output sequence x^i:

$$\underbrace{a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots}_{\text{auto-regressive terms}} = \underbrace{b_0v^n + b_1v^{n-1} + b_2v^{n-2} + b_3v^{n-3} + \cdots}_{\text{moving average terms}}. \tag{1.2}$$

The input sequence is $v^0, v^1, v^2, \ldots, v^n, \ldots$, where n is the index of a time-like scansion. For an impulse, we have $v_0 = 1$ and $v_n = 0, \forall n > 0$. The terms with coefficients a_i on the left-hand side correspond to the auto-regressive function: they specify how the output at time n depends on the output at previous times. The terms with coefficients b_i on the right-hand side correspond to the moving average function: an average of the input at previous times. The properties of such filters can be studied rigorously: When is it stable? What is the duration of its impulse response? What is the frequency content of its output? Its properties can be described in the time domain as impulse responses, or in the frequency domain as a transfer function. These kinds of filters are commonly used for audio or video compression. It is also used for usual filtering purpose, as for example in noise-reduction devices. For noise reduction, stochastic signal modelling is often used. The goal is no longer to fit a time sequence, but to fit a given statistics. The output is desired to have a given cross-covariance; it is the statistics that are the object of the interpolation. This type of method is instrumental for optimal filtering designs, such as methods based on Wiener filtering. For instance, a filter can be optimally designed so as to remove from a signal a noise whose statistics is known; for this, one must in a first step build a filter that interpolates the noise's statistical properties. Such filters equip, for instance, radar devices and mobile phones, whose signal is typically very noisy.

It is interesting to note that even though the Yule–Walker equation is a widely used and mature tool, it is seldom found in fluid mechanics applications. Perhaps one reason is that it is most practical for systems with few channels, whereas data obtained from the Navier–Stokes equation have a large number of degrees of freedom. We will see that the formal application of the Yule–Walker equations to data derived from systems of partial differential equations raises a few challenges, which we note in Appendix C.

We turn now to turbulent inflow generation. An overview of the literature can be found in Keating *et al.* (2004). Many approaches and contexts are analysed in this review. Also, the salient properties expected from an artificial inflow are discussed. We present below four specific papers that should serve as references to understand and evaluate our contribution.

We first consider Klein, Sadiki & Janicka (2003) and di Mare *et al.* (2006). The authors understood that the statistics of the fluctuations should be considered: correlation in space and correlation in time. They use a filtering technique to generate a correlated field out of random noise. The intuition of using a filter to generate a field whose properties mimic that of a turbulent field was successful; indeed, this is the necessary ingredient that introduces spatial and temporal scales for the fluctuations. For this task, they have chosen moving average filters: their output is a linear combination of past random inputs, i.e. $a_n = 0, \forall n > 0$ in (1.2). Unfortunately, this particular filter structure allows little freedom in the design of the output correlation: the procedure could not be pushed to the needed fidelity, that is, to obtain a field with statistics as close as desired to the reference statistics. To avoid this aspect, they have chosen to use reference statistics of exponential type, for which the filter coefficients can be derived so as to fit the measured turbulent properties.

A different approach was tested in Perret *et al.* (2008). Here an experimental turbulent flow is considered as the reference data. Using a rake of hot-wire probes, they extract coherent flow structures by a proper orthogonal decomposition technique. Knowledge of these structures is equivalent to the knowledge of the two-point covariance. Then, using a Galerkin projection, they build a dynamic system which is essentially a reduced-order model of the measured turbulent flow. After verifying that this system has the compulsory property of stability and that its attractor faithfully contains information on time scales of the reference flow, this model is marched in time and its state is used as an inflow condition. This method has the interesting feature of using a nonlinear model for generating the artificial turbulence. Nonlinear model reduction methods are receiving much attention but the methodology is still rather constraining for turbulence. In particular, the authors were restricted to using just a few representative coherent structures lest the reduced model be unstable. Here, again, we find a limitation to the need for fidelity.

A third approach was presented in Druault *et al.* (2004). After demonstrating that inflow fluctuations uncorrelated in space or time lead to an immediate decay of the turbulence, they devise an ingenious technique based on stochastic estimation. They assume that two aspects of the reference field are known: its two-point covariance, which can be measured once and for all using two mobile hot-wire probes, and exact time series of the velocity at three points in space. From the time series and the spatial correlation, it is possible to estimate the field away from these chosen points in space: it is obtained from considerations of conditional probability. Here, the ingredient of the spatial scales comes from the two-point covariance, whereas the form of the time evolution is contained in a few exact time series. Three time series are needed to account for the several processes happening simultaneously in the reference field.

The method developed in our work owes a direct filiation to Klein *et al.* (2003). Here, too, random noise is filtered to generate a correlated inflow. But we have chosen a different filter structure: auto-regressive as opposed to moving average, as will be described in the body of the paper. This choice was instrumental for a full flexibility of the output statistics, but it also gives a framework in which it is possible to impose explicitly correlations in space and time.

To get more insight into the properties needed from an inflow generation technique, let us consider three cases in which such a technique could be useful. One may first think about the comparison between an experiment and a numerical simulation. In the experimental set-up, it is typically possible to measure very long sample series of complex flows, but spatial resolution or time resolution is a limitation, as when using, for example, hot-wire probes and particle-image velocimetry. If the inflow could be reproduced numerically, one could concentrate the simulation on a domain where the effects of interest are confined. The simulation would typically provide a very detailed description of the flow and pressure field, but for short time samples. Thus, experiment and simulation can provide complementary viewpoints if their inputs are reproduced.

Considering now only simulation, another context is derived from a building-blocks view of turbulent flows in complex geometries. For instance, the transition from the laminar to turbulent regime is simulated in a channel. Flow statistics at the channel outlet are regenerated at the inlet of a diffuser, whose outlet statistics could be regenerated at the inflow of a leading edge simulation and so on. Each block represents the mutation of the turbulence regime from one equilibrium state to the other one, depending on the particularities of the flow: indeed, the statistics of the turbulence in the channel, the diffuser and the leading edge are very different.

In the above, we have focused on turbulent-flow regimes, but description of flow fluctuations could as well be instrumental in the study of the processes of transition to turbulence. The classical view of transition to turbulence consists of the emergence of an unstable eigenmode, growing out of a featureless background noise. In this view, the selection of the appearing wavelengths and flow structures depends solely on the dynamics of the system itself. Nevertheless, the study of flow receptivity shows that the flow response could be dramatically affected by the nature of the incoming perturbations: roughness of solid surfaces, free-stream turbulence outside a boundary layer, vorticity stretching at a leading edge ... (see Schmid & Henningson 2001). A distinction can be drawn between flows that are convectively unstable and flows that are absolutely unstable. A flow that is convectively unstable typically behaves as an amplifier for incoming perturbations; see Huerre & Monkevitz (1990) and Chomaz (2005). Also, the processes of selective sensibility of flow systems have been described in the context of transient growth. A flow system is said to be selectively sensitive whenever a large response is observed only for a very specific type of perturbations. A celebrated example is the generation of streamwise streaks in boundary layers, as the response to streamwise vortices known as the lift-up mechanism; see Butler & Farrell (1992) and Reddy & Henningson (1993). For flows whose dynamic response consists of a selective sensitivity, a precise description of the dynamics must be done hand in hand with a faithful description of their natural perturbative environment. An example where a linear system describing the stability of a flow is subjected to stochastic perturbations can be found in Farrell & Ioannou (1993). The method that we describe in the present paper can be useful in describing the natural noise environment. A similar context in which it is useful to properly describe perturbations is for flow control. Indeed, the better the incoming perturbations are known, the better the controller can be tuned to act efficiently. In control-optimization procedures, the description of the perturbation is a key element; see Hoepffner *et al.* (2005) and Chevalier *et al.* (2006).

We will illustrate our method in the case of a free mixing layer. The testing procedure will be as follows. We simulate the spatially developing mixing layer, with a hyperbolic tangent base-flow profile as inflow. We inject a large number of vortices, randomly located and with random circulation and radius. This is an artificial inflow, whose aim is to excite a rich response in the mixing layer. Some distance downstream, the flow has evolved into a natural turbulent mixing layer. This simulation will be called the 'reference case' here. We choose one downstream location at which the two-point two-time statistics (1.1) and the mean-flow profile are measured. We subsequently run new simulations, using for the inlet, the mean field and perturbations statistics regenerated from the reference case. We will use three degrees of quality of regeneration, in order to assess the performance of our method.

2. The flow and its statistics

In this section, we describe the flow that we consider, with its inlet condition, and show the statistics that we will, in the second step, seek to regenerate using the filtering technique.

The present numerical simulations are done by using the direct numerical simulation code of Fukagata, Kasagi & Koumoutsakos (2006), adapted to the mixing layer configuration. The computational domain has dimension $(L_x, L_y) = (80, 80)$ with the grid resolution $(N_x, N_y) = (1024, 192)$. The grid stretching in y is set similar to that of Druzhinin & Elghobashi (2001). The top and bottom of the computational domain

have a symmetry boundary condition, and at the outflow we impose a convective boundary condition for the velocity field and a vanishing pressure gradient.

For all simulations, the interior of the domain is initialized with an unperturbed mixing layer. At the inlet of the simulation, a hyperbolic tangent inlet profile is imposed as a base-flow profile. This inflow profile with two streams of different velocity will then diffuse slowly downstream due to viscous diffusion in addition to the complex vortex motions. Here, we will consider a Reynolds number of 400, large enough so that viscous diffusion remains minute. This Reynolds number is defined using as length scale the inflow vorticity thickness $(U_{top} - U_{bottom})/\max(U')$, where U' is the y derivative of the base flow, and using $U_{top} - U_{bottom}$ as reference velocity.

The characteristic of this flow is that it is unstable to the Kelvin–Helmholtz instability. The region of shear in the domain centreline has an inflectional profile. The instability will lead to the generation of vortices with a radius of the order of magnitude of the thickness of the shear region. These vortices are advected downstream by the mean flow. When these have grown and saturated, they start interacting with each other and pairing occurs: two or more vortices that are brought up in close vicinity will induce in each other a rotating motion and form a larger rotating structure; see Ho & Huerre (1984) for a review of the properties of mixing layers.

Statistics are obtained by sampling the velocity field every time unit during 10 000 time units. This time span corresponds to 125 through time on the lower velocity side of the inflow and 250 through time on the higher velocity side. The time step is 0.01 time unit.

The first step of our procedure consists in simulating a mixing layer which we will use as a reference case in the following sections. It is from this flow that we will extract the statistics that will be regenerated using filters later on. We could choose to obtain this reference case by simulating a mixing layer with the steady-inflow profile and let the instability play their role in generating the flow structures of interest out of the small numerical noise always present in simulations. Doing so would require a long spatial simulation domain for the transition to turbulence to occur. To help this process, we could introduce at the inlet, in addition to the mean profile, a small amount of noise, whose statistics would, for instance, be similar to that of the natural eigenmodes of the instability of the mixing layer (see e.g. Druzhinin & Elghobashi 2001).

Here, we adopt a slightly different approach; we choose to introduce a large number of disturbances at the inlet. This choice is motivated by the wish to stimulate in the mixing layer a wide range of flow structures, spatial and temporal scales. Note that this inflow is not constructed to mimic a natural configuration; it is intended to lead to a great variety of excitations. For this, we inject vortices at the inlet, whose properties, such as radius, circulation, position and time of appearance, are set randomly.

Each vortex has the following structure with azimuthal velocity:

$$V_{yt} = \frac{\Gamma}{2\pi r}\left[1 - \exp\left(-\frac{r^2}{r_c^2}\right)\right]. \tag{2.1}$$

We build an inflow field in y and time, as the sum of the velocity fields of N such vortices:

$$\sum_{i=1}^{N} V(y_i, t_i, \Gamma_i, r_{ci}). \tag{2.2}$$

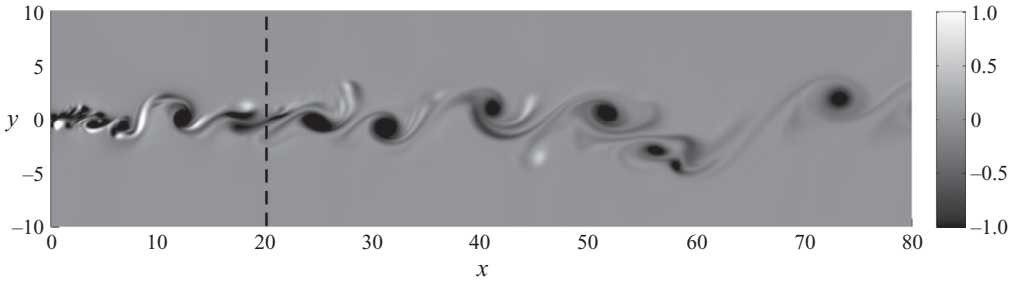

FIGURE 1. Snapshot of the vorticity of the reference simulation. The flow is from left to right, and the fastest stream is on the top. We will later consider in detail the flow statistics at $x = 20$, highlighted here by a dashed line.

The properties of each vortex are random variables with the following statistical properties. The vortex location y_i is Gaussian-distributed with zero mean and standard deviation 0.5, the time of vortex advection through the inlet t_i is uniformly distributed in the total time interval, with on average six vortices per unit time. The vortex circulation Γ_i is Gaussian-distributed with zero mean and standard deviation 0.5, and the vortex radius r_{ci} is log-normal distributed with standard deviation 0.25. In order to avoid the interaction of the vortices with the top and bottom boundary conditions of the computational domain, we introduce a smooth cutoff of the vortex-induced velocity beyond 5 vortex radii r_{ci}. The lesser domain height which we can thus use helps to concentrate the computational resources on long simulations for accurate statistics. The particular choice of this inflow structure is arbitrary, but it is a simple parametrization that serves the purpose of our reference inflow.

A snapshot of the vorticity field of a simulation excited in this fashion is represented in figure 1. One can clearly see the vortices appearing at the inlet, being advected by the mean flow, evolving and interacting. We can see the initiation of a pairing event at about $x = 50$. This simulation will now be referred to as the reference simulation. We consider that its statistics are the 'true' statistics that we will regenerate later.

We proceed now to describe in detail the flow field. Indeed, in a later section, we will start a simulation at one downstream location, using artificially regenerated inflow conditions, and compare the subsequent flow evolution with that of the reference case. In figure 2, we show a spatiotemporal evolution of the streamwise u and vertical v velocity components of the flow at $x = 20$. This x-location corresponds to the vertical dashed line in figure 1. We can clearly observe the signature of vortices being advected through this downstream location. In figure 3, we give a statistical description on the flow at this location. In figure 3(a), we can see the mean-flow profile. It is superimposed on the hyperbolic tangent profile imposed at the inlet for comparison. We see that the occurrence of vortices soon increases the height of the shearing region due to their mixing action. We also show the statistics of the fluctuations of the flow about this mean profile. It is the variation in y of the components of the Reynolds stress tensor

$$\overline{u(y,t)u(y,t)}, \quad \overline{v(y,t)v(y,t)}, \quad \overline{u(y,t)v(y,t)}, \tag{2.3}$$

where the overbar denotes the mean value. We now look at temporal statistics, in figure 3(b). It is the one-point two-time correlation for u and v:

$$\overline{u(y,t)u(y,t+\Delta t)}, \quad \overline{v(y,t)v(y,t+\Delta t)}, \quad \overline{u(y,t)v(y,t+\Delta t)}. \tag{2.4}$$

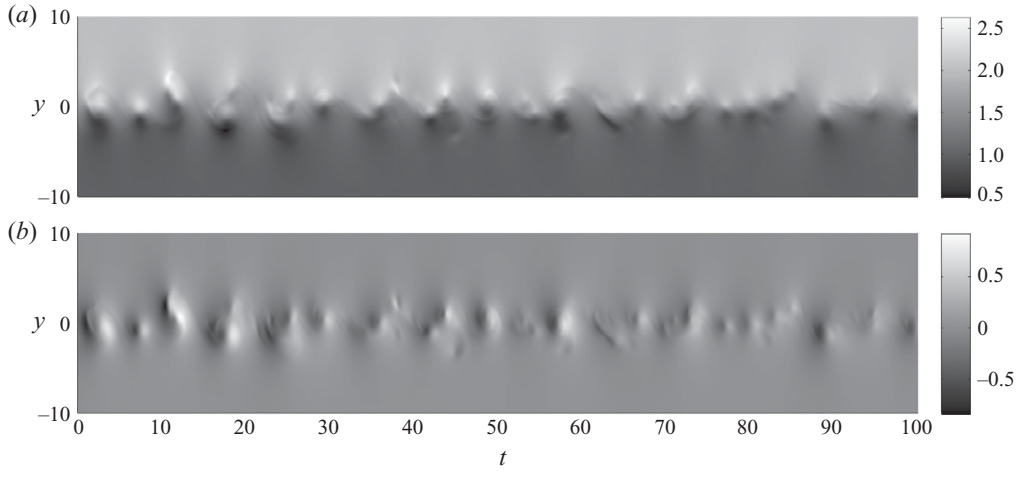

FIGURE 2. Temporal evolution of (*a*) u and (*b*) v at $x = 20$.

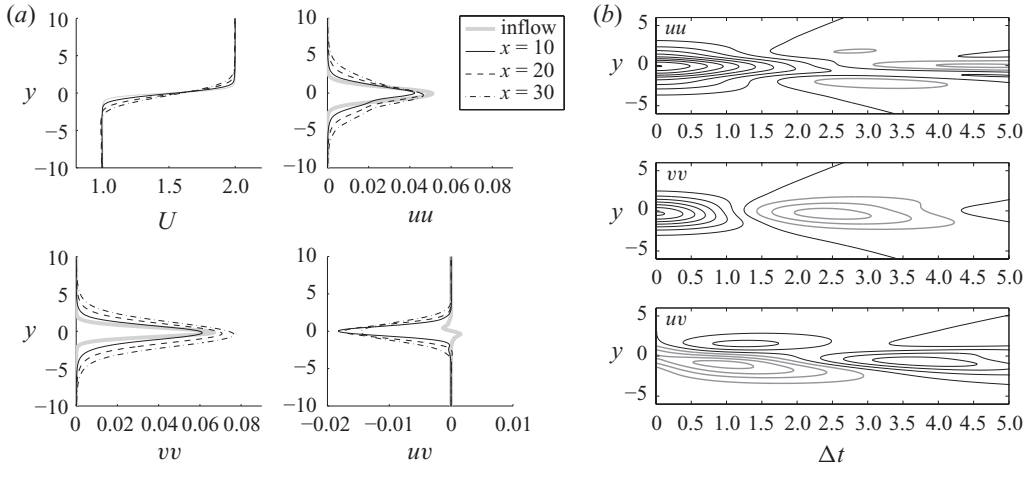

FIGURE 3. (*a*) Mean profiles U at $x = 10, 20, 30$ compared to the mean profile at inflow and the comparison of the Reynolds stresses $\overline{uu}$, $\overline{vv}$, $\overline{uv}$ from (2.3). (*b*) One-point two-time correlation from (2.4) at $x = 20$.

These data show how each velocity component is correlated with itself in time and give information on the typical time scales present at that x-location.

We will now continue the description of the flow at that location but from a slightly different point of view. We would like to extract spatial structures that are *coherent*. We will now briefly describe what can be meant by coherent spatial structures and how we can extract this information from the spatial statistics that we have shown in figure 3. The usefulness of these coherent structure as an alternative basis for the representation of the flow statistics will then clearly appear. Note that this change of basis is often referred to as 'proper orthogonal decomposition'; we will discuss its details and see how the spatial and temporal statistics can be accounted for. For a review on this decomposition technique in fluid mechanics, see Berkooz, Holmes & Lumley (1993).

We consider the spatial covariance of the velocity at this location, that is, the two-point, one-time covariance:

$$P_{uu} = \overline{u(y,t)u(y',t)}, \quad P_{vv} = \overline{v(y,t)v(y',t)}, \quad P_{uv} = \overline{u(y,t)v(y',t)}. \tag{2.5}$$

Naturally, the two velocity components u and v behave in a coordinated manner, this coordination being the consequence of the dynamical processes at play. We can thus consider that the flow field is described by a composite vector, formed by the concatenation of the two velocity components, say $\boldsymbol{q}$:

$$\boldsymbol{q}(y,t) = \begin{pmatrix} u(y,t) \\ v(y,t) \end{pmatrix}. \tag{2.6}$$

With this representation, the covariance information is carried by the following composite array:

$$P_{qq} = \overline{\boldsymbol{q}(y,t)\boldsymbol{q}(y',t)} = \begin{pmatrix} P_{uu} & P_{uv} \\ P_{vu} & P_{vv} \end{pmatrix}. \tag{2.7}$$

The long vector $\boldsymbol{q}$ has as elements the values at a given time of the velocity components at the mesh nodes in y. The top elements of this vector correspond to the velocity u and the bottom elements correspond to the velocity v. We can think that behind this representation, there is an implicit choice of a basis; this is a physical space basis. This choice of basis is convenient for the visualization of the data in physical space, that is, as long as it is the spatial coherence of the data that we are interested in. On the other hand, it is perfectly possible for another choice of basis to be of interest in view of our particular purpose.

As an alternative representation, we will now consider a very special basis, the basis in which the covariance matrix $\boldsymbol{P}_{qq}$ is diagonal. Of course, this basis is built out of the eigenvectors $\boldsymbol{V}_i(y)$ of $\boldsymbol{P}_{qq}$. As stated, the matrix is diagonal in this basis, and the diagonal elements are the eigenvalues λ_i associated with the eigenvectors. Let us now consider the meaning of this property of being diagonal.

The matrix $\boldsymbol{P}_{qq}$ was a covariance matrix in the physical space basis and remains a covariance matrix in the new basis; indeed, in the new basis, $\boldsymbol{q}$ is described not by the values of the velocity at the mesh point, but by the expansion coefficients c_i, the coordinates of $\boldsymbol{q}$ on the eigenvectors

$$\boldsymbol{q}(y,t) = \sum_i c_i(t)\boldsymbol{V}_i(y), \tag{2.8}$$

where the $\boldsymbol{V}_i(y)$ are the eigenvectors – evolving in space – and $c_i(t)$ are the coordinates in the new basis, evolving in time. $\boldsymbol{P}_{qq}$ expressed in the basis of the eigenvectors is thus the covariance matrix of the coordinates c_i, i.e.

$$\hat{\boldsymbol{P}}_{qq} = \overline{cc^T} = \begin{pmatrix} \overline{c_1c_1} & \overline{c_1c_2} & \dots & \overline{c_1c_m} \\ \overline{c_2c_1} & \overline{c_2c_2} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \overline{c_mc_1} & \dots & \dots & \overline{c_mc_m} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_N \end{pmatrix}, \tag{2.9}$$

where the superscript T denotes the transpose. The physical meaning of this covariance matrix being diagonal is thus clear: the spatial structures described by the eigenvectors have zero correlation. Since, at a given time, the knowledge of the expansion

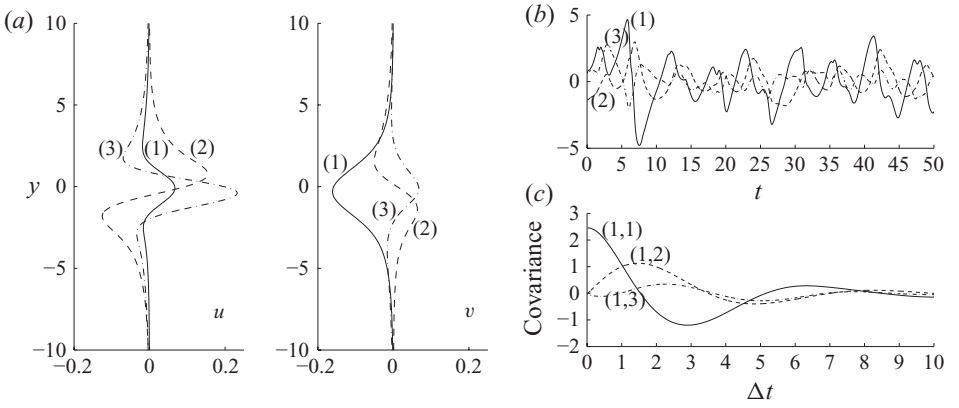

FIGURE 4. Coherent structures of the reference simulation at $x = 20$. (a) The three structures with largest variance, (b) a sequence of the time evolution of the first three coordinates c_1, c_2 and c_3, and (c) covariance of the first coordinate with itself, denoted (1, 1), and the two next ones (1, 2) and (1, 3), while varying the time lag.

coefficients c_i gives strictly no information about the expansion coefficient c_j for another eigenvector, the two eigenvectors are coherent structures.

This analysis corresponds to the plots in figure 4. In figure 4(a), we show the three coherent structures with largest variance. Note that each eigenvector has of course a u component and a v component. In this representation, it is not straightforward to identify each structure with a physical process since these are one-dimensional structures and time evolution is not accounted for. In figure 4(b), we represent the evolution in time of the coordinates c_i for these three structures. We can note that these three most energetic structures have similar amplitude and similar time scales. Typically, structures with lesser variance have shorter time scales.

These uncorrelated structures pertain to the two-point one-time information. Thus, they carry no information about the time behaviour of the structures, that is, the time evolution of the c_i must be studied *a posteriori*. For instance, we have seen that, by construction, we have the property

$$\overline{c_i(t)c_j(t)} = 0, \quad i \neq j, \tag{2.10}$$

but there is no reason why they should be uncorrelated at two times t and $t + \Delta t$. If we now consider the two-coordinate two-time covariance

$$\overline{c_i(t)c_j(t + \Delta t)}, \tag{2.11}$$

we recover the full statistical information of interest. This information is represented in figure 4(c) for the covariance of the first coordinate with itself, denoted as (1, 1), and the two next coordinates (1, 2) and (1, 3), while varying the time lag Δt. We observe that the first structure has a negative peak of correlation about $\Delta t = 3$, indication of a typical period of oscillation of twice that of the duration. As expected from the choice of basis, the first structure has zero correlation with the other ones for time lag $\Delta t = 0$, but we observe a peak of correlation between the first and second structures at about $\Delta t = 1.5$, the sign of the phase quadrature typical of structures combined to represent advected fluctuations.

3. Filtering noise

We would like to generate a field evolving in time that has the same two-point two-time statistics as that of the reference simulation at the chosen x-location. This field must have steady statistics, and the method should allow us to reproduce as accurately as desired the target statistics. The most simple random field is uncorrelated Gaussian noise. Its statistics are well defined, and it is convenient theoretically. It is also convenient from a computational point of view, since efficient algorithms are available to realize large arrays of Gaussian white noise. The approach that we will use to generate the inflow will be to *filter* white noise. The filter will introduce correlation in space and in time as the footprints of its structure and its coefficients. The realization problem, 'How can we generate a field with the desired statistics?', is thus transformed into the computation of the filter coefficients such that the output of the filter has the desired statistics.

3.1. *Scalar case*

We start by illustrating the method on a simple scalar case. For a very good textbook on filtering, please see Hayes (1996). A convenient structure is that of the auto-regressive filter. Consider the order one filter

$$x^n + a x^{n-1} = v^n. \tag{3.1}$$

Here, n denotes the index of time sequences; v^n is the input sequence, which we will assume has a variance Q and zero correlation from one time to another, i.e.

$$\overline{v^n v^n} = Q, \quad \overline{v^n v^{n+k}} = 0. \tag{3.2}$$

In (3.1), a is the single filter coefficient and x is the output sequence. We would now like to know how the statistics of x depend on a and Q. We denote by P the statistics of x. For instance, the variance of x is $P_0 = \overline{x^n x^n}$. We are also interested in the covariance of x at different times, so we denote

$$P_k = \overline{x^n x^{n+k}}. \tag{3.3}$$

From (3.1), we are able to derive a system of equations to determine P_k as a function of the parameters a and Q. For this, we multiply it first with x^n and take the mean value; we get

$$\underbrace{\overline{x^n x^n}}_{P_0} + a \underbrace{\overline{x^n x^{n-1}}}_{P_1} = \underbrace{\overline{x^n v^n}}_{?}. \tag{3.4}$$

To determine the term labelled with the question mark, we multiply (3.1) by v^n and take once again the mean value, to obtain

$$\overline{x^n v^n} + a \underbrace{\overline{x^{n-1} v^n}}_{0} = \underbrace{\overline{v^n v^n}}_{Q}, \tag{3.5}$$

where $\overline{x^{n-1} v^n}$ is zero because the output x^{n-1} cannot depend on future input due to causality in (3.1). We thus replace $\overline{x^n v^n}$ with Q to obtain the first equation for P:

$$P_0 + a P_1 = Q. \tag{3.6}$$

We need a second equation; for this we multiply (3.1) by x^{n-1} and average to get

$$P_{-1} + a P_0 = \underbrace{\overline{v^n x^{n-1}}}_{0}, \tag{3.7}$$

where, once more, x does not depend on future input. We then have the system of two equations:

$$\left.\begin{aligned} P_0 + aP_1 &= Q, \\ P_{-1} + aP_0 &= 0. \end{aligned}\right\} \tag{3.8}$$

This system is known as the Yule–Walker equation. From its definition, P_k must be even with respect to k: $P_k = P_{-k}$. The unique solution of this system is thus

$$P_0 = Q/(1-a^2), \quad P_1 = -aP_0. \tag{3.9}$$

Here, P_0 is the variance of x and must be positive, thus $|a| \geqslant 1$ would lead to a contradiction. In fact, the system is unstable whenever $|a| > 1$, in which case the mean of its output is not defined. We also see that P_1 will be smaller than P_0 in amplitude. From these two values of P_0 and P_1, we can also get P_2, P_3 and so on, by multiplying (3.1) by $x^{n-2}, x^{n-3}, \ldots$, which leads to the recursion

$$P_{k+1} = -aP_k. \tag{3.10}$$

In these derivations, we have assumed that the filter parameters a and Q were known, and that we were interested in obtaining the statistics of x, but we could do the other problem where one wants the output to have the statistics P_k, and would like to know the corresponding filter coefficients a and Q. This associated problem – the signal modelling problem – is formulated and solved in the same way. Here, the solution is given by (3.9),

$$a = -P_1/P_0, \quad Q = (1-a^2)P_0. \tag{3.11}$$

This is the problem that we are interested in for the present paper: the statistics of the inflow are known (P_k), and we want to build the associated filter.

The system in (3.1) is an order one system: the present value of the output depends on the previous step of the output and on the present input. Due to this choice, there is only one filter coefficient a, thus little freedom in shaping the output statistics. We can nevertheless imagine filters of a higher order:

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_p x^{n-p} = v^n, \tag{3.12}$$

with a scalar filter of order p, input sequence v and output sequence x. For this filter, we obtain a Yule–Walker system as for (3.8), but with $p+1$ equations.

In Appendix A, we have included a comparison of the ARMA filter with the state space filter formulation and indicated the similarities between the Yule–Walker equation and the Lyapunov equation.

3.2. *Filter for the inflow*

The system (3.1) that we have used as an example had scalar input and output sequences v and x. For the inflow of our mixing layer, the sequence that we need to generate is the value of the u and v velocities at the mesh nodes in y at one downstream x-location; it is thus a vector sequence. We can equivalently consider a formulation in the basis of the coherent structures, in which case the filter output must be the time sequence of the coordinate vector of the inflow in this basis. We thus need a filter with a vector input sequence and a vector output sequence; this filter will thus have matrix coefficients. In the signal-filtering literature, such a system is called a multi-channel filter.

We still denote the output at step n as $\boldsymbol{x}^n$, but now $\boldsymbol{x}^n$ is a vector of length m, and the filter coefficients will be denoted $\boldsymbol{A}_0, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$, thus p matrices, where p is the filter order. The input sequence v is still uncorrelated in time, but we need to allow

a correlation between its components in space. This is described by the one-time covariance matrix $\boldsymbol{Q}$. Thus, (3.1) becomes

$$\boldsymbol{x}^n + \boldsymbol{A}_1\boldsymbol{x}^{n-1} + \boldsymbol{A}_2\boldsymbol{x}^{n-2} + \cdots + \boldsymbol{A}_p\boldsymbol{x}^{n-p} = \boldsymbol{v}^n. \tag{3.13}$$

The two-time statistics of x will now be denoted by P and defined by

$$\boldsymbol{P}_k = \overline{\boldsymbol{x}^n \boldsymbol{x}^{(n-k)T}}, \tag{3.14}$$

where T denotes the transpose. From this definition, we see that $\boldsymbol{P}_k$ has the property $\boldsymbol{P}_{-k} = \boldsymbol{P}_k^T$. Now $\boldsymbol{P}_k$ is an $N \times N$ matrix of the covariance coefficients of the m-elements vector $\boldsymbol{x}^n$. Performing the same manipulations as for (3.1), we obtain

$$\left.\begin{array}{llll} \boldsymbol{P}_0 & +\boldsymbol{A}_1\boldsymbol{P}_1 & \ldots & +\boldsymbol{A}_p\boldsymbol{P}_p = \boldsymbol{Q}, \\ \boldsymbol{P}_{-1} & +\boldsymbol{A}_1\boldsymbol{P}_0 & \ldots & +\boldsymbol{A}_p\boldsymbol{P}_{p-1} = \boldsymbol{0}, \\ \vdots & \vdots & & \vdots \quad\quad \vdots \\ \boldsymbol{P}_{-p} & +\boldsymbol{A}_1\boldsymbol{P}_{-p+1} & \ldots & +\boldsymbol{A}_p\boldsymbol{P}_0 = \boldsymbol{0}, \end{array}\right\} \tag{3.15}$$

which we can rewrite conveniently as a matrix product:

$$\underbrace{(\boldsymbol{I} \mid \boldsymbol{A}_1 \ \ldots \ \boldsymbol{A}_p)}_{\mathscr{A}} \underbrace{\left(\begin{array}{c|ccc} \boldsymbol{P}_0 & \boldsymbol{P}_{-1} & \ldots & \boldsymbol{P}_{-p} \\ \hline \boldsymbol{P}_1 & \boldsymbol{P}_0 & \ldots & \boldsymbol{P}_{1-p} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{P}_p & \boldsymbol{P}_{p-1} & \ldots & \boldsymbol{P}_0 \end{array}\right)}_{\left(\begin{array}{c|c} \mathscr{P}_{11} & \mathscr{P}_{12} \\ \hline \mathscr{P}_{21} & \mathscr{P}_{22} \end{array}\right)} = \left(\boldsymbol{Q} \mid \boldsymbol{0} \quad \ldots \quad \boldsymbol{0}\right). \tag{3.16}$$

If we now partition $\mathscr{P}$ as $\mathscr{P}_{11}, \mathscr{P}_{12}, \mathscr{P}_{21}$ and $\mathscr{P}_{22}$, we have

$$\left.\begin{array}{l} \mathscr{A}\mathscr{P}_{22} = -\mathscr{P}_{12}, \\ \boldsymbol{Q} = \mathscr{P}_{11} + \mathscr{A}\mathscr{P}_{21}. \end{array}\right\} \tag{3.17}$$

We see that the system can be solved in two steps, first for the filter coefficients $\mathscr{A}$ and second for the covariance $\boldsymbol{Q}$ of the input sequence. The computation of $\mathscr{A}$ can be performed directly by any linear system solver as

$$\mathscr{A} = -\mathscr{P}_{12}\mathscr{P}_{22}^{-1}, \tag{3.18}$$

when the number of elements in the output vector and the filter order are not too large. When, on the other hand, the size of the matrices in this system precludes a direct solution, we can use a very efficient recursive solution procedure, as described in Appendix B. In Appendix C, we discuss cases in which $\mathscr{P}_{22}$ might not be invertible.

Now that we have a filter, the last technical step is to generate the random field with the desired statistics: the realization. As stated, the input of the filter must be white noise, that is, a vector sequence $\boldsymbol{v}^n$ uncorrelated in time, whose element-to-element statistics are described by $\boldsymbol{Q}$, that is, $\boldsymbol{Q}$ is the one-time covariance matrix of v:

$$\boldsymbol{Q} = \overline{\boldsymbol{v}^n \boldsymbol{v}^{nH}}. \tag{3.19}$$

It is easy to generate large fields of uncorrelated white noise, as for instance using the function `randn` in Matlab; we will now see how to give it a covariance $\boldsymbol{Q}$. This

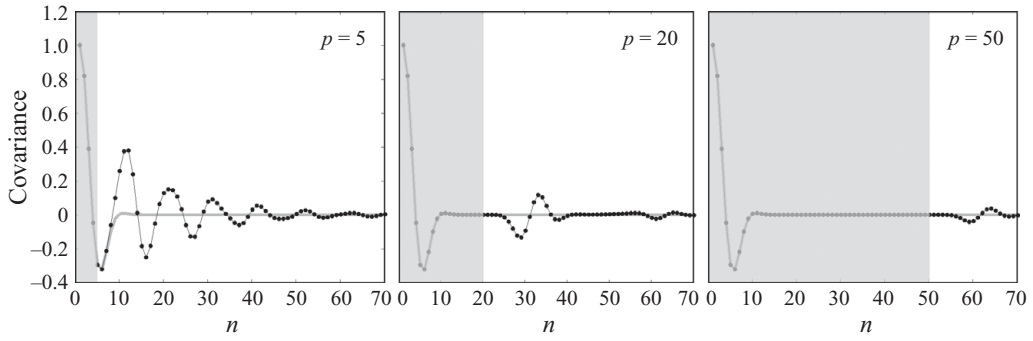

FIGURE 5. Comparison of the target time correlation (thick line) to the statistics of the output of the filter for three filter orders. Time lags where the fit is exact are highlighted by the shaded area.

is in fact a filtering problem too, a much simpler filtering problem. For this, we start by deriving the covariance of $v = \boldsymbol{F}\boldsymbol{w}$, where $\boldsymbol{w}$ is an uncorrelated random vector and $\boldsymbol{F}$ is an arbitrary matrix. By construction, the covariance matrix of $\boldsymbol{w}$ is the identity matrix $\boldsymbol{I}$ and the covariance of $\boldsymbol{F}\boldsymbol{w}$ is

$$\boldsymbol{Q} = \overline{(\boldsymbol{F}\boldsymbol{w})(\boldsymbol{F}\boldsymbol{w})^T} = \boldsymbol{F}\underbrace{\overline{\boldsymbol{w}\boldsymbol{w}^{\mathrm{H}}}}_{\boldsymbol{I}}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{F}^T. \tag{3.20}$$

We can thus construct the desired random input v by building $\boldsymbol{F}$ via the Cholesky factorization $\boldsymbol{F}\boldsymbol{F}^T = \boldsymbol{Q}$. The sequence $\boldsymbol{v}^n$ thus obtained is to be fed to the filter (3.13). To start the sequence, we need p initial steps of the output, for which we can use $\boldsymbol{x}^i = 0, i = 1, \ldots, p$ at the expense of an initial transient. During this transient the statistics of the output will be different from the steady-state statistics, which are the useful ones for our purpose. This transient has time duration of about that of the filter impulse response.

Let us now quickly summarize the steps followed to generate a field with given statistics. First, the two-point two-time statistics of the reference field is measured. From the two-point covariance, we obtain the coherent structures $\boldsymbol{V}_i(y)$. By transforming the field in the basis of these coherent structures we obtain the time evolution of the coordinates $c_i(t)$. From these data, we can build the two-coordinate two-time covariance $\boldsymbol{P}$. We can now build the filter. The filter matrices are obtained by solving (3.17) for $\mathscr{A}$ in a first step, and the input one-time covariance $\boldsymbol{Q}$ is obtained from $\mathscr{A}$ in a second step. In order to generate the input sequence v from a completely random field $\boldsymbol{w}$, we need to factorize $\boldsymbol{Q}$ as $\boldsymbol{F}\boldsymbol{F}^T$ to have $\boldsymbol{v} = \boldsymbol{F}\boldsymbol{w}$. We are now in position to generate the desired field by marching (3.13), with the initial condition $\boldsymbol{x}^0 = \cdots = \boldsymbol{x}^p = 0$.

3.3. *A scalar illustration*

We apply the solution of the Yule–Walker equation to a reference scalar sequence with time covariance

$$P_k = \cos(0.6\pi t_k)\exp\left(-t_k^2/2\right), \tag{3.21}$$

where the time $t \in [0, 20]$ is discretized at 70 linearly spaced sample times. It is fitted using filters of three different orders: $p = 5, 20$ and 50. The temporal correlation of the output of the three filters is shown in figure 5. By construction, the fit is exact to machine accuracy for a time sample from 0 to the order of the filter. For time lags longer than the filter order, the filter output statistics is not constrained. This may

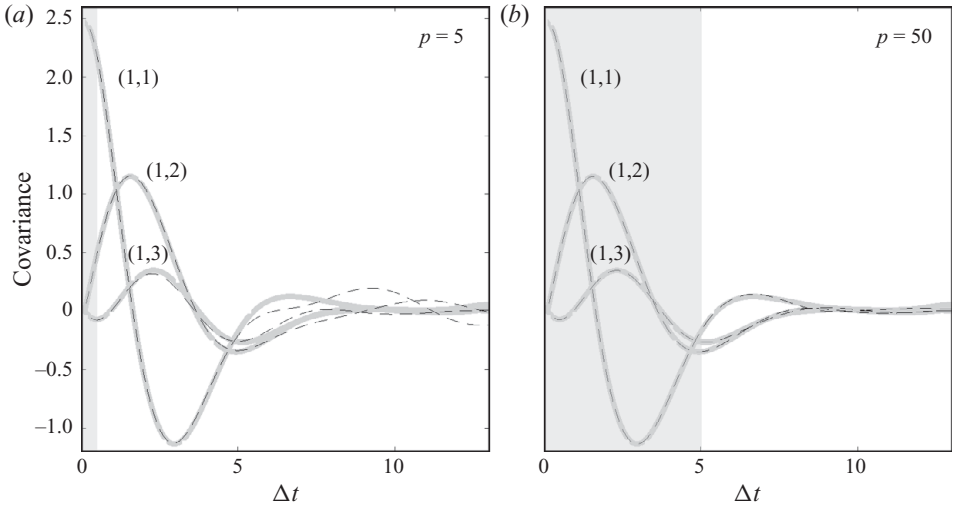

FIGURE 6. Correlation in time of the first coordinate c_1 with itself and the two following ones, for the reference simulation (plain grey) and for two choices of filter order, (*a*) $p = 5$ and (*b*) $p = 50$ (dashed). The time sample period is here 0.1.

result in spurious correlations for time lags larger than the filter order. This effect is analogous to Gibbs' phenomenon for polynomial interpolation and is visible for the present example.

Indeed, in (3.15), we have imposed that the cross-correlation be exactly that of the given statistics for all time lags up to a filter order p. For a larger time lag, nothing is imposed on the cross-correlation of the filter output. In cases where this effect is strong, an alternative formulation can be used where more equations are written than the filter order. This overdetermined system is then solved by the least-squares method. With this choice, the correlation is fitted approximately, but up to larger time lags. For the purpose of the present paper, thanks to the recursive solution method of the Yule–Walker equation described in Appendix B, we are not restricted to the filter order. Also, we will not observe spurious oscillations when fitting the turbulent inflow in § 4; we have thus kept to this first formulation. Another motivation for this choice is that we are not aware of a recursive solution algorithm adapted to the least-squares approach.

3.4. *Realizing the inflow statistics*

In § 2, we described the flow field at one downstream location: space/time diagrams and different representations of its statistics. Then, in § 3.2 we introduced a filtering technique to generate an artificial field evolving in time with desired statistics. In the present section, we will construct these filters and discuss their properties for the particular application of the inflow generation. Using the filter output for the inflow of new simulations of the Navier–Stokes equation will be done in the next section.

We have discussed in § 2 the possibility of choosing different bases to represent the data. We saw that the eigenmodes of the two-point covariance matrix were particularly interesting because they have zero correlation for zero time lag. This will prove a useful property, as discussed in detail in Appendix C. The filter will be built to realize the two-coordinate two-time statistics of the inflow $\overline{c_i(t)c_j(t+\Delta t)}$.

In figure 6, we compare the variance of the first coordinate c_1 and its covariance with the two following coordinates c_2 and c_3. We have built two filters, with order

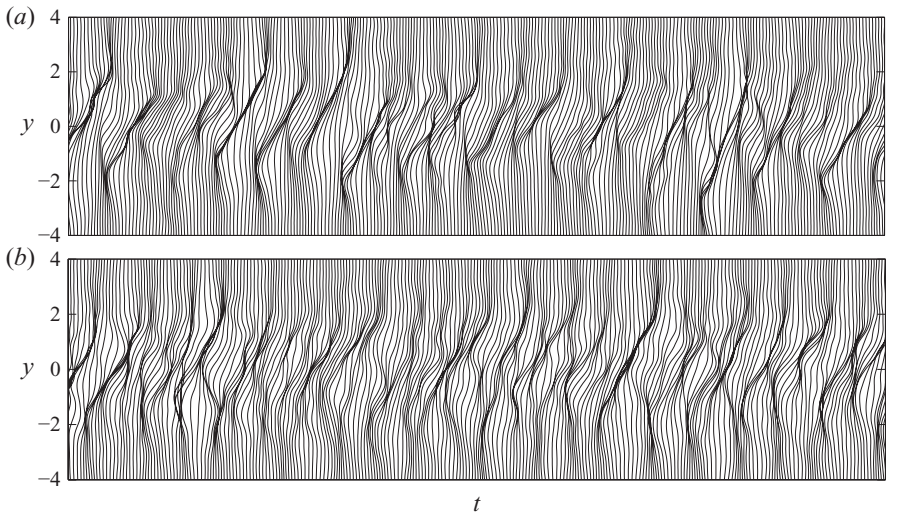

FIGURE 7. The streamwise velocity evolution in time for (a) the reference field at $x = 20$ and (b) a realization with the same two-point two-time covariance. Here the mean-flow profile is added to the fluctuations. We use here a filter of order $p = 200$.

respectively $p = 5$ and $p = 50$. By construction (see (3.15)), the fit is exact for time lags up to filter order, and this is of course respected here; for larger time lags, nothing is imposed. Nevertheless, we can see here that the order 5 filter shows a good fit for much larger time lags. It means that the statistics of the coordinates is itself in fact not much more 'complicated' than an order 5 process. In the figure, we also show the fit for the order 50 filter, which is able to reproduce in detail the low-energy oscillations of the statistics' tail.

Once they are generated, we can transform the field of coordinates varying in time to physical space by back-projection using the eigenmodes, as in (2.8). In figure 7, we show a sequence of the time evolution of the streamwise velocity at the considered x-location and visually compare it to the output of the filter. Here, we have added the mean-flow profile. These two fields originate from very different processes: one is a subset of the velocity field of a numerical solution of the Navier–Stokes equation – that is the state evolving in time of a nonlinear partial differential equation – and the other is the output of a linear filter whose input is white noise. By construction, these two fields have the same two-point two-time statistics. The visual agreement is fair; we can observe similar time scales at different heights in the flow domain and similar amplitude of the fluctuations.

The physical space comparison of the reference statistics and the realized statistics is done in figure 8. There we first show the variance of the Reynolds stresses, varying in y, and then a choice of two-point two-time statistics. For these results, we have used a filter order $p = 200$. Here, the statistics of the filter output are computed using a sample average, which might allow for a slight discrepancy at time lags lower than filter order.

In the filter setting, there are two parameters which define the accuracy to which the reference statistics is fitted. These two parameters are m, the number of retained coherent structures, and p, the order of the filter. The number of modes decides how many degrees of freedom are kept in the flow description: each of these degrees of freedom is the coordinate onto the associated coherent structure, and will correspond

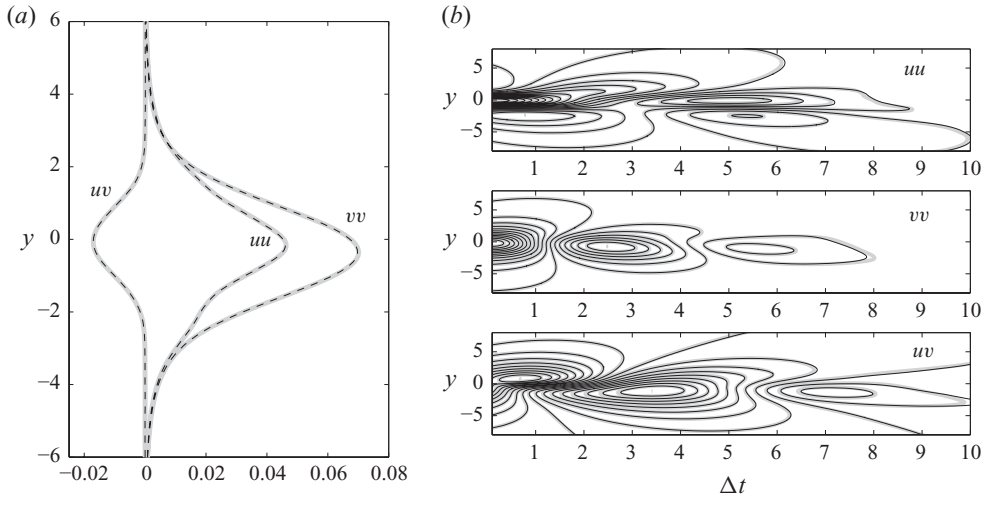

FIGURE 8. Comparison in physical space of the statistics of the reference flow at $x = 20$ and the filter output. (*a*) Variation of the Reynolds stress in y. (*b*) Contours of the two-point two-time covariance $\overline{u(0,0)u(y,t)}$ $\overline{v(0,0)v(y,t)}$ and $\overline{u(0,0)v(y,t)}$; thick grey line, reference; thin line, filter output.

to one channel of the filter. The filter order p sets the extent in time (the time lag) up to which the fit is exact.

To test the method of inflow generation, we have chosen to overperform the accuracy in terms of the channel number and filter order. We are testing the idea of regenerating the two-point two-time correlation and not the specific technical details of the regeneration. We did so also because the efficient recursive solution technique allows it. We see this efficiency as a strong aspect of the method which we would like to emphasize.

We found that 50 coherent structures carry more than 99.9 % of the turbulent kinetic energy. We will thus use an $m = 50$ channels filter. For the time extent, we see in figure 6 that a time lag of 20 time units is well over the correlation time; we thus generate the inflow condition with filters of order $p = 200$.

We now touch upon a subtle aspect of the inflow generation. Since the inflow is a plane instead of a volume, it is not strictly speaking concerned with fluid incompressibility. On the other hand, it is reasonable to assume that inlet fluctuations are for short distances simply advected at the mean-flow velocity, thus leading to a convective spatial derivative in the streamwise direction. If the continuity equation evaluated with this approximation was violated, a rapid unrealistic evolution would be forced through the pressure. The reference field satisfies incompressibility, and the continuity equation is linear; thus, any field reproducing the two-point two-time covariance also satisfies this property. For the same reason, each single coherent structure also satisfies this property.

We have shown the regenerated streamwise velocity in figure 7. It is more informative on the other hand to represent the vorticity since it has more structured features, tracing the advected vortices. We use the convective derivative in the streamwise direction to derive the vorticity at the fixed $x = 10$, corresponding to the location of the regenerated inlet condition. The vorticity field in y and time is shown in figure 9. We see on this graph that despite generating the correct two-point two-time covariance of the reference field, the large-scale vortical structures are absent

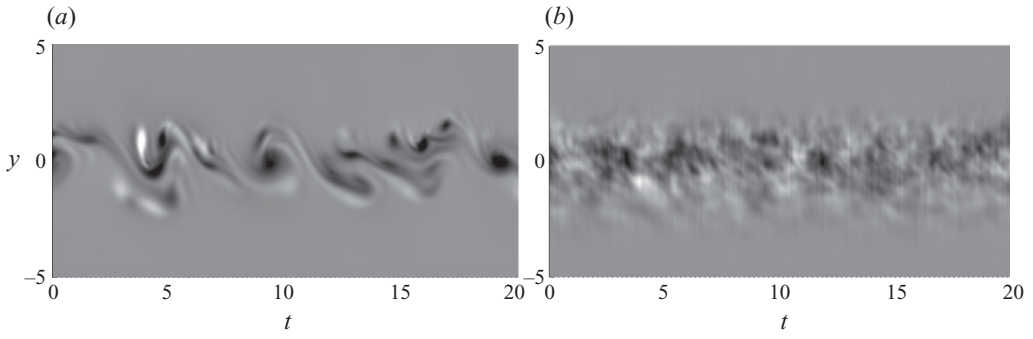

FIGURE 9. Flow vorticity for fixed x varying in time for (a) the reference field at $x = 10$ and (b) the regenerated inflow; the vorticity is computed using the convective derivative in the streamwise direction. We see that the regenerated inflow lacks the large-scale features of the passing vortices.

from the artificial inlet. Therefore, it takes a recovering length downstream of the inlet before the artificial flow can recover the proper vortical elements. This recovering length is related to the mean vortex size at the inlet location.

4. Using the realized field as inlet condition

We have confirmed in the previous section that the filtering technique is able to generate inflow conditions with the desired statistics. We will now run simulations using these regenerated fields and associated mean profiles as inflow conditions.

For this, we will use three qualities of regenerated inflows based on the measured statistics at $x = 10$. In the first one, we set to zero the correlation in time. This is simply implemented in our framework by building a filter with order $p = 0$: the filter is thus simply $x^n = v^n$. The input sequence v and thus x has the spatial correlation Q. The inflow thus generated has the exact two-point one-time correlation of the reference flow, but zero correlation in time. We will refer to this method as 'no time correlation'.

The second quality of inflow generation is as follows. We keep the exact time correlation, but we discard the correlation between the different coherent structures. To generate such an inflow, we apply the complete filter procedure with $m = 50$ and $p = 200$, but prior to this, we artificially set to zero the off-diagonal elements of the $\boldsymbol{P}_i$ matrices in (3.15). This amounts, for instance, to having the curves labelled (1, 2) and (1, 3) set to zero in figure 6. This leads to an inflow smooth in time and in space, but which lacks the fine details on the interplay of the different structures. We will refer to this inflow generation as 'no cross-correlation'.

The third quality of inflow is that it reproduces the exact two-point two-time correlation. This is the 'full correlation' inflow. For the three methods we will regenerate the reference statistics at $x = 10$, a position where the vortex inflow of the reference simulation has had time to settle.

The set-up is illustrated in figure 10: a snapshot of vorticity of the reference simulation is compared to snapshots of the simulations using as inlet the filter outputs based on the three methods.

Of course, the instantaneous fields of the reference simulation and those of the regenerated inflow are not supposed to be identical; on the other hand, their statistics should be the same if our method is successful. It is thus the statistics of the reference field and those of the regenerated simulations that should be compared. A first

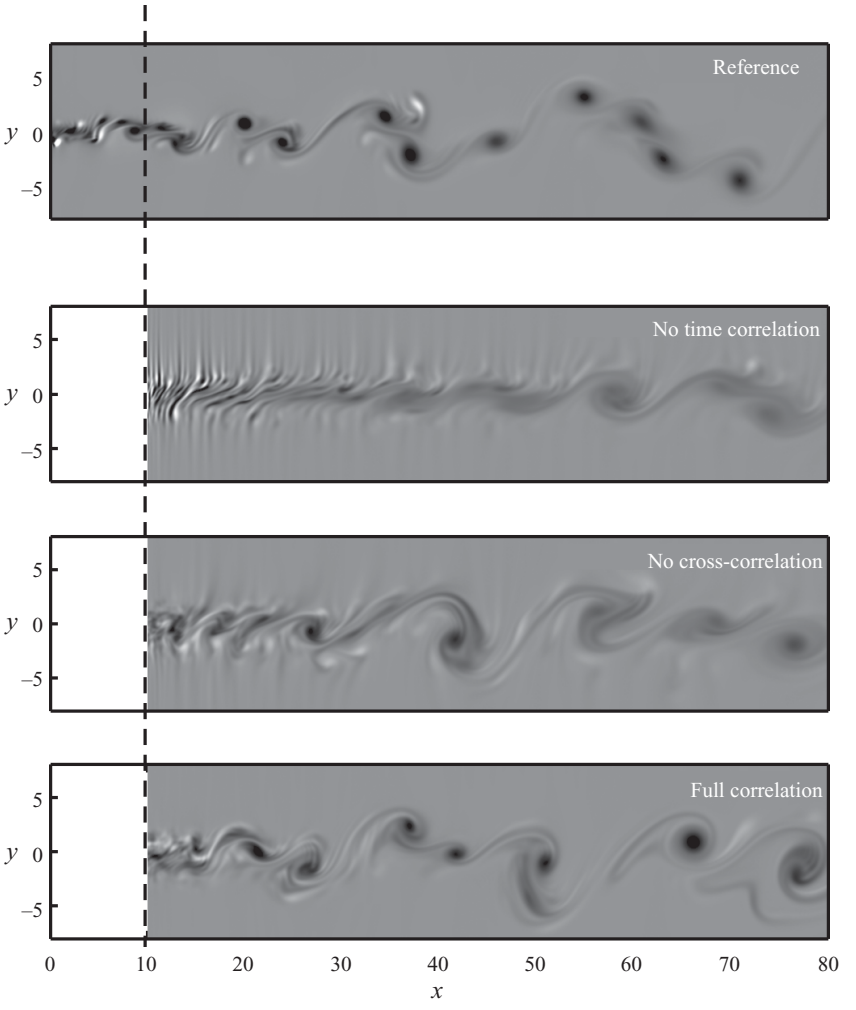

FIGURE 10. Snapshots of the vorticity field of the reference simulation compared to that of the three inflow generation methods starting at $x = 10$. We see at the inlet fluctuations reminding us of the lack of phase relationship observed in figure 9. Nevertheless, we confirm here that the flow field quickly adjusts, giving confidence that the decisive ingredients are preserved.

comparison is presented in figure 11, where we show the downstream evolution of uu, vv and uv along the centreline $y = 0$.

We can see that the vortex inflow of the reference simulation has a quick adjustment where the amplitude of the fluctuation decays and then grows again slowly once relaxed to fluctuations that the mixing layer can naturally sustain. We first concentrate on the regenerated flow with no time correlation. At $x = 10$, the position of the inflow, all components of the Reynolds stress have exactly the value of the reference; indeed, the correct statistics are imposed at the inflow location, but we can see that these fluctuations cannot be sustained by the Navier–Stokes equations and decay rapidly. The amplitude of the fluctuations then starts growing again, thanks to the instabilities in the mixing layer, but can never reach the fluctuation amplitude of the reference flow. Turning now to the inflow generated with no cross-correlation, we can see that the fluctuation amplitude follows a similar trend to the no correlation case; we observe

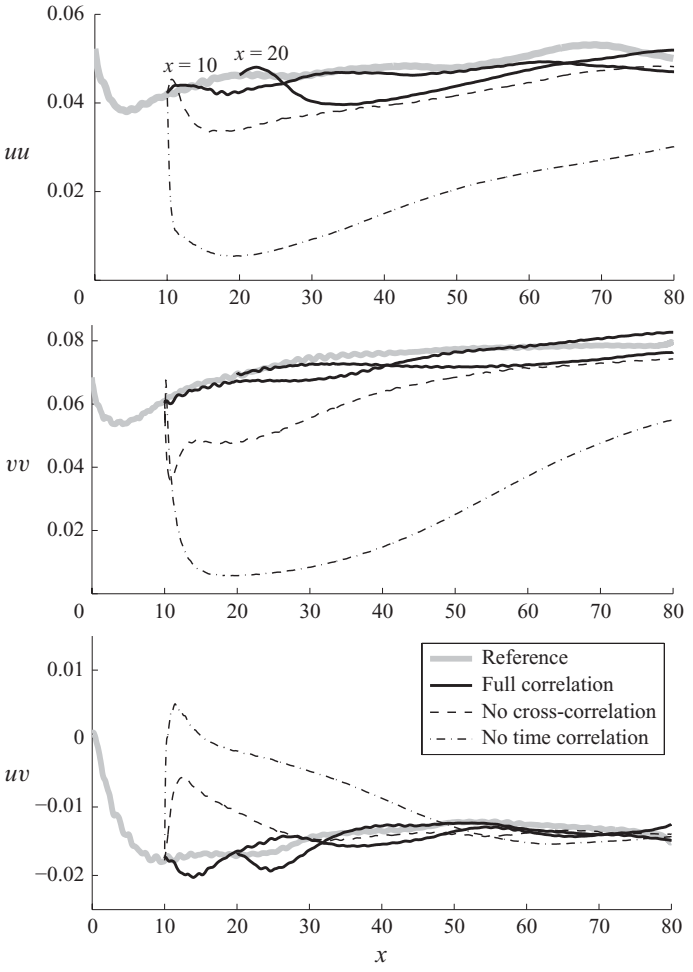

FIGURE 11. Downstream evolution of the Reynolds stress components along the flow centreline $y = 0$ for the reference simulation and the three inflow regenerations for the inlet location at $x = 10$. We have also drawn the centreline evolution of the full correlation case with inlet starting at a position $x = 20$ farther downstream.

a quick decay in amplitude, though of lesser extent. Once again after this decay, the remaining fluctuations can feed the natural instability of the mixing layer and we observe growth again. These two behaviours can now be compared with that of the full correlation method. For uu and vv, we cannot distinguish a phase of decaying downstream of the inlet: the fluctuations have the structure necessary to be sustained by the natural flow dynamics. The fluctuation amplitude grows slowly, following in that the amplitude of the reference simulation.

We have also represented the centreline Reynolds stresses for the full correlation case with the inlet condition at $x = 20$, further downstream. We see a trend similar to the case of $x = 10$, but degraded with a first overshoot of uu then decay of the fluctuation amplitude below that of the reference case. The reason for the degradation of the inlet condition for $x = 20$ can probably be understood by considering figure 9, where we saw that the artificial inlet field does not carry the clear vortical structure. This is in opposition to the reference flow at $x = 20$, since at that position the

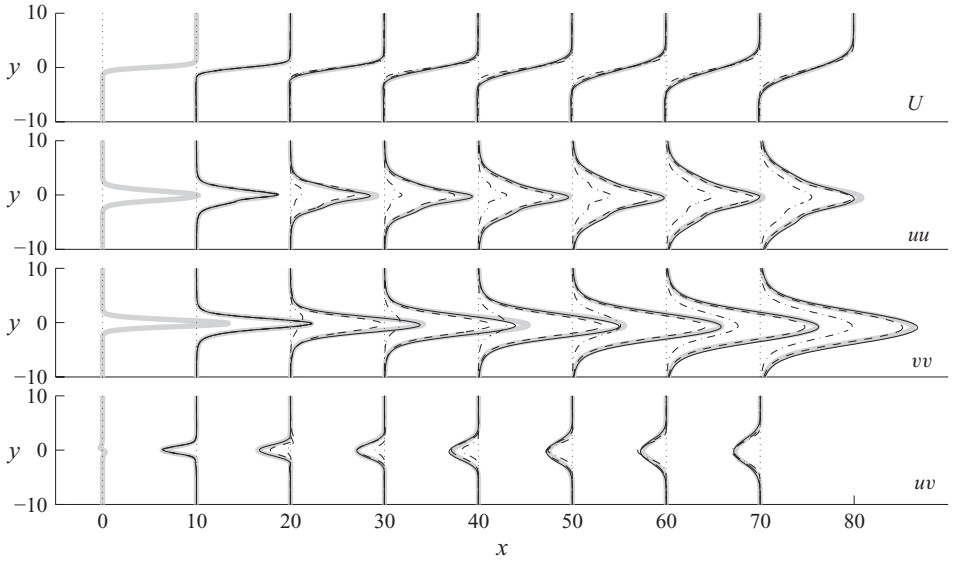

FIGURE 12. Profiles of the mean flow and Reynolds stresses for the reference simulation compared to that of the simulation with regenerated inflows (line styles similar to figure 11).

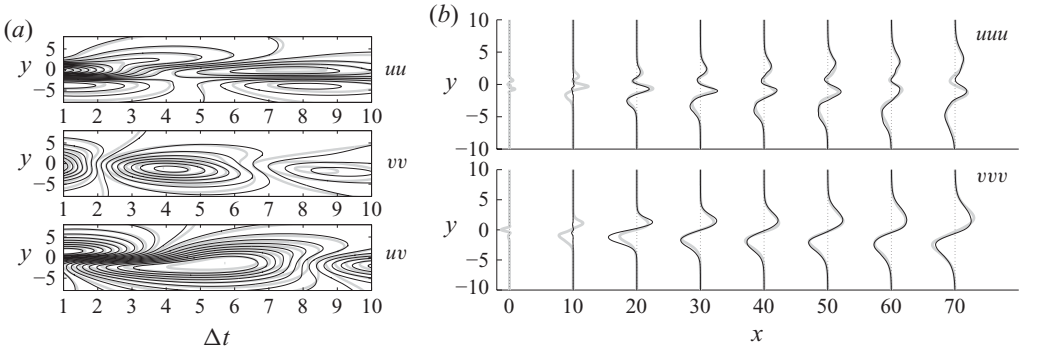

FIGURE 13. Comparison of the reference simulation and the regenerated simulations. (*a*) Comparison of the one-point two-time covariance at $x = 30$; thick grey line, reference; thin black line, simulation with inflow at $x = 10$. (*b*) Comparison of the third-order moments (line styles similar to figure 12).

fluctuating field has had time to organize itself into large vortices with slow time scales of evolution.

Figure 12 displays the profiles of the components of the Reynolds stress. We can see that the mean flow is not very sensitive to the inflow generation technique, but still, best fit is seen for the full correlation method. It is clear in this figure that all the methods reproduce exactly uu, vv and uv at the inlet, but that this is not sufficient for a proper flow behaviour, as can be seen by the profiles downstream of the inflow for 'no time correlation' and 'no cross-correlation'.

See figure 13(*a*) for a comparison of the one-point two-time data for the full correlation method at $x = 30$. Very similar structures are observed. We can now turn to third-order data. See figure 13(*b*) for uuu and vvv. We have seen that the original second-order moments are well recovered with artificial inflow – in all the domain

and not only at the inlet. It is interesting to consider third-order statistics, since these are not specified for the filter output. We can see in the figure that at the inlet, these third-order moments are zero, since their values are not specified through the filter design. We see that uuu is able to relax to a profile similar to that of the reference field in about 10 spatial units, but a longer relaxation is needed for the vvv component, which starts to resemble that of the reference after as much as 20 spatial units downstream of the inlet.

5. Discussion

The main results of this paper are presented in (3.17) for the Yule–Walker equation and in figure 10, where we compare the instantaneous vorticity fields of a reference computation of a turbulent free shear layer to simulations where the inflow is artificially regenerated. A quantitative check of the flow obtained in the domain is provided in figure 11. In that figure, the inflow with exact two-point two-time correlation is compared with two degraded versions of the method: discarding time correlation and discarding cross-correlation between the coherent structures. We see in Keating *et al.* (2004) that an inflow of low-quality induces a rapid decay of the fluctuation amplitude downstream. We found this effect in our cases too, for the 'no time correlation' and 'no cross-correlation' inflows. We may phrase this as whether the Navier–Stokes equation accepts or rejects the artificial inflow fluctuations. If the injected fluctuations cannot live on in the flow, then they are indeed rejected and the fluctuation amplitude rapidly decays. We can say that our correlated inflow is in this sense accepted by the Navier–Stokes equations.

When a method is proposed for the turbulent inflow problem, one should also have a series of tests to judge the quality of the method. In Keating *et al.* (2004), the chosen testbed is a plane channel flow. The inflow quality is tested by the question: How far downstream of the artificial inflow does the turbulence become natural again?' The longer it takes, the more costly the computations since a longer domain will be needed. Here, there is indeed a well-defined natural turbulent regime since this flow is homogeneous. The cases of non-homogeneous flows are more demanding, since simple inflow conditions based on recycling the outflow into the inflow are not directly applicable. If the non-homogeneous flow configuration has a self-similar development, then the statistics of its fluctuations may as well be self-similar, in which case an artificial inflow will relax to a natural evolution. This is the case in our evolving free shear layer for instance. If on the other hand there is no self-similarity, there is then no natural state of the turbulence: the statistics evolve with the flow from that of the source of excitations. Thus, in this case, the low fidelity of an artificial inflow cannot be compensated for by a longer computational domain.

The question 'How much of the true turbulence must be reproduced?' ultimately depends on the needed degree of accuracy. For the cylinder in a turbulent flow which we mentioned in our introduction, the quality of the artificial inflow should be such that the cylinder wake response is the same, that is, up to the sensitivity of the process itself. Unfortunately, if an inflow condition must be good for all possible purposes, it is probably all of the true turbulence that we will need to reproduce; this goal is beyond our reach.

We must thus seek to formulate the turbulent inflow problem in a relevant manner. This formulation should be independent of the precise case that we would like to use the inflow on and be useful for the development of new methods. The test procedure typically consists of a reference turbulent flow, and a simulation with an

inflow which reproduce the reference at one downstream location to a certain degree of accuracy. The flow evolution downstream of this inflow must then be compared to the reference. The useful question can then be: How much of the true turbulence must be reproduced at the inflow if that much of the true turbulence is to be reproduced downstream of the inflow?

We have been able to test the method on a two-dimensional flow case, and it is clear that its ability to reproduce the important features of a turbulent inflow cannot be safely assessed before a full three-dimensional flow is actually tested. The present paper was written with the scope of describing accurately the technicalities of the method and provides the conceptual context in which its testing is possible, while avoiding the heavy computations required by a full three-dimensional turbulent simulation. Fortunately, the filtering technique is formulated on the orthogonal basis of the coherent structures of the inflow. Once this transformation is done, only expansion coefficients are manipulated, which are indifferent to the flow dimensionality. Memory-efficient techniques are available to obtain the coherent structures of flows, using proper-orthogonal decomposition. This can be done, for instance, using the method of snapshots. It should thus be possible to get the statistical data for three-dimensional flows. Certainly, more coherent structures will then be necessary for a precise statistical description, but the filter computations will be tractable, thanks to the recursive resolution technique described in Appendix B.

The main technical limitation of the formulation pertains to the discrete nature of the filter structure we have used: discrete in time and discrete in space. Spatial and temporal manipulation of the input and output fields in the filter is not explicitly scaled as temporal and spatial derivatives. Therefore, the increase in the sample rates does not induce convergence of the filter coefficients towards an ideal continuous limit. We have pointed in Appendix C to references which offer alternative filter formulations that satisfy the desired convergence properties, at least for temporal sampling. A second shortcoming has a more subtle origin, pertaining to the linearity of the statistical data we desired to regenerate: the two-point two-time covariance. The information carried by the two-point two-time covariance is equivalent to the spatiotemporal spectrum and thus fails to describe the instantaneous phase relationship of the active structures of the turbulence. This loss of phase information is discussed in detail in Keating *et al.* (2004). We have shown in figure 9 how the vorticity field right at the artificial inlet did not carry the large-scale vortical structures that would be necessary for immediate adaptation from the inlet condition to the natural evolution downstream. We also used this argument to explain why in figure 11 the inflow regenerated at $x = 20$ was not as efficient as that for $x = 10$.

The use of this method relies on the knowledge of the mean-flow profile and on the two-point two-time covariance in that cross-flow plane which is to become the new inflow. These data can be measured in experiments. Carrying out these measurements is a necessary step whenever experiments and simulations are to be compared. In the general case, on the other hand, it would be very valuable if a database were to be made available where one could find the two-point two-time covariance along cross-flow planes for typical flow cases: channel flow, boundary layer, free jet, free shear layer, wall–jet, cylinder wake.

Appendix A. Relation with other filter structures

The works of Farrell & Ioannou (1993) and Bamieh & Dahleh (2001) have made popular the use of the Lyapunov equation to study flow response to stochastic

excitations. We describe here for convenience the manner in which Yule–Walker and Lyapunov are related. From the auto-regressive filter structure, one can build a state space representation of the filter. For example, let us consider a first-order auto-regressive filter

$$\boldsymbol{x}^n + A\boldsymbol{x}^{n-1} = B\boldsymbol{v}^n. \tag{A 1}$$

We can write the same relation between x and v as

$$\underbrace{\begin{pmatrix} \boldsymbol{x}^n \\ \boldsymbol{x}^{n-1} \end{pmatrix}}_{\hat{\boldsymbol{x}}^n} = \underbrace{\begin{pmatrix} -\boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix}}_{\hat{\boldsymbol{A}}} \underbrace{\begin{pmatrix} \boldsymbol{x}^{n-1} \\ \boldsymbol{x}^{n-2} \end{pmatrix}}_{\hat{\boldsymbol{x}}^{n-1}} + \underbrace{\begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{0} \end{pmatrix}}_{\hat{B}} \boldsymbol{v}^n, \tag{A 2}$$

where the state is X, the input is v and the output is

$$\boldsymbol{x}^n = (\boldsymbol{I} \quad \boldsymbol{0})\hat{\boldsymbol{x}}^n. \tag{A 3}$$

For the auto-regressive filter, the statistics of the output are obtained from the Yule–Walker equation, and for the state space system (A 3), it is obtained from the discrete-time Lyapunov equation:

$$\hat{\boldsymbol{A}}\hat{\boldsymbol{P}}\hat{\boldsymbol{A}}^T - \hat{\boldsymbol{P}} + \hat{\boldsymbol{B}}\hat{\boldsymbol{B}}^T = 0, \tag{A 4}$$

where the covariance matrix of the state is

$$\hat{\boldsymbol{P}} = \overline{\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^T} = \overline{\begin{pmatrix} \boldsymbol{x}^n\boldsymbol{x}^{nT} & \boldsymbol{x}^n\boldsymbol{x}^{(n-1)T} \\ \boldsymbol{x}^{n-1}\boldsymbol{x}^{nT} & \boldsymbol{x}^{n-1}\boldsymbol{x}^{(n-1)T} \end{pmatrix}} = \begin{pmatrix} \boldsymbol{P}_0 & \boldsymbol{P}_1 \\ \boldsymbol{P}_{-1} & \boldsymbol{P}_0 \end{pmatrix}. \tag{A 5}$$

The discrete time Lyapunov equation can be solved by standard software, for instance the function `dlyap.m` in Matlab; see Bartels & Stewart (1972).

Appendix B. Recursive solution of the Yule–Walker equation

The filter coefficients are obtained from the solution of the linear system of equations (3.17). The number of elements in the system matrix $\mathscr{P}_{22}$ is $(m \times p)^2$, where m is the number of degrees of freedom in the output sequence x and p is the filter order.

It is possible to use the Toeplitz structure of $\mathscr{P}_{22}$ to solve the system recursively. For instance, see Akaike (1973) for numerical algorithms for the solution of linear systems with the Toeplitz structure. For the particular case of the Yule–Walker equation, the combined structure of $\mathscr{P}_{22}$ and $\mathscr{P}_{12}$ allows for an even faster algorithm, the Levinson–Durbin recursion; see for instance Durbin (1960). It was first developed for mono-channel filters ($m = 1$) and later extended to multi-channel filters by Whittle (1963). One nice fact is that the system is solved recursively upon filter order p. The recursion is initiated with the trivial solution for $p = 0$, and from this solution, one can compute the filter coefficients of order $p = 1$ and so on until the desired accuracy is reached.

The structure of the multi-channel algorithm introduces the use of an instrumental 'backward filter', whose coefficients will be denoted $\hat{A}_i$. This filter corresponds to a filter whose output depends on future input. At each step of the recursion, the coefficient matrices of both filters are computed to the current order. To emphasize the recursion, we denote by $\boldsymbol{A}_j^p$ and $\hat{\boldsymbol{A}}_j^p$, $j \in [0, p]$ the coefficient matrices of the filters at the step p of the recursion. We assume that $\boldsymbol{A}_0^p = \boldsymbol{I}$. The equations satisfied by the

two filter matrices are (3.17), and its equivalent for the backward filter,

$$\sum_{k=0}^{p} \boldsymbol{A}_k^p \boldsymbol{P}_{j-k} = 0, \quad \sum_{k=0}^{p} \hat{\boldsymbol{A}}_k^p \boldsymbol{P}_{k-j} = 0, \quad (j = 1, \ldots, p). \tag{B 1}$$

We define the matrices

$$\left.\begin{aligned} \boldsymbol{Q}^p &= \sum_{k=0}^{p} \boldsymbol{A}_k^p \boldsymbol{P}_{-k}, \qquad \boldsymbol{\Delta}^p = \sum_{k=0}^{p} \boldsymbol{A}_k^p \boldsymbol{P}_{p-k+1}, \\ \hat{\boldsymbol{Q}}^p &= \sum_{k=0}^{p} \hat{\boldsymbol{A}}_k^p \boldsymbol{P}_k, \qquad \hat{\boldsymbol{\Delta}}^p = \sum_{k=0}^{p} \hat{\boldsymbol{A}}_k^p \boldsymbol{P}_{-p+k-1}. \end{aligned}\right\} \tag{B 2}$$

Then the recursion is as follows: a new filter coefficient is added at the new step $p+1$, based on results of the step p,

$$\boldsymbol{A}_{p+1}^{p+1} = -\boldsymbol{\Delta}^p (\hat{\boldsymbol{Q}}^p)^{-1}, \quad \hat{\boldsymbol{A}}_{p+1}^{p+1} = -\hat{\boldsymbol{\Delta}}^p (\boldsymbol{Q}^p)^{-1}, \tag{B 3}$$

where we see that the elements from the forward and backward filters are combined. From these new coefficients, we can update the rest of the coefficient sequence from step p to step $p+1$ as follows:

$$\boldsymbol{A}_k^{p+1} = \boldsymbol{A}_k^p + \boldsymbol{A}_{p+1}^{p+1} \hat{\boldsymbol{A}}_{p-k+1}^p, \quad \hat{\boldsymbol{A}}_k^{p+1} = \hat{\boldsymbol{A}}_k^p + \hat{\boldsymbol{A}}_{p+1}^{p+1} \boldsymbol{A}_{p-k+1}^p, \tag{B 4}$$

for k from 1 to p. The recursion starts at $p=0$ from $\boldsymbol{A}_0^0 = \hat{\boldsymbol{A}}_0^0 = \boldsymbol{I}$, then from (B 2) $\boldsymbol{Q}_0 = \boldsymbol{P}_0$, $\boldsymbol{\Delta}_0 = \boldsymbol{P}_1$, $\hat{\boldsymbol{Q}}^0 = \boldsymbol{P}_0$, $\hat{\boldsymbol{\Delta}}^0 = \boldsymbol{P}_{-1}$.

This recursion requires at each step the solution of two linear system of equations (B 3), where the system matrices are respectively $\boldsymbol{Q}^p$ and $\hat{\boldsymbol{Q}}^p$, that is, the approximation at the current step of the covariance matrix of the input sequence $\boldsymbol{v}$. Thus, instead of solving at once m linear systems with the system matrix of size $p \times m$, we solve $2 \times p$ times a set of m linear systems with the matrix of size m. The main cost of the solution of the Yule–Walker equation comes mainly from large output size m, and increasing the filter order p does not result in much cost.

Appendix C. Two limitations of the discrete filter design

We discuss here two aspects of the discrete filter design which might lead to inaccuracy in the resolution of the Yule–Walker equation with finite-accuracy arithmetics. In essence, these originate from the discrete-in-time and discrete-in-space filter formulation, whereas our reference signal is continuous.

Consider as an example a simple scalar auto-regressive filter $x^n + a x^{n-1} = b v$, with x^n being the sample number n of a continuous signal. In the limit of short time interval Δt between samples, x^n must nearly equal x^{n-1}; thus, $a \to -1$ and $b \to 0$ whenever $\Delta t \to 0$. For multi-channel systems, the equivalent tendency is a clustering of the system eigenvalues in a small neighbourhood of 1 of the complex plane. This clustering will lead to numerical ill-conditioning of (3.17). This effect is overcome by deriving the filter structure from a continuous time model; see Vijayan *et al.* (1991). A possible structure of a continuous-time auto-regressive filter is

$$\frac{\partial^p x}{\partial t^p} + A_1 \frac{\partial^{p-1} x}{\partial t^{p-1}} + A_2 \frac{\partial^{p-2} x}{\partial t^{p-2}} + \cdots + A_p x = v, \tag{C 1}$$

where time derivatives appear up to filter order p. Proper scaling with respect to the time sample rate in the discretization of the time derivative leads to a well-behaved

continuous limit. We have not tested this method in our case. Essentially, x^n being very close to x^{n-1} can be interpreted as a sign that both need not be generated. The time step of our numerical simulations is 0.01, much shorter than the typical time scale of the inflow field, below which $x^n \approx x^{n-1}$, we can thus choose to generate the artificial inflow on a coarser sampling 0.1 and interpolate it afterwards to the needed time instants.

A second aspect of the Yule–Walker equation was analysed in Cybenko (1980), dealing with the cases of signals for which some channels are nearly exactly correlated, for instance if one channel is a replication of another channel. In this case, the matrix $\mathscr{P}_{22}$ of (3.17) has one nearly vanishing eigenvalue, preventing an accurate linear system solve for the filter coefficients. Several numerical techniques are proposed in Cybenko (1980) to circumvent this potentially harmful numerical conditioning.

For our inflow signal, which is obtained from the spatial sampling of a continuous field, channel correlation necessarily increases with the sample rate: in the limit of small spatial distance, velocity correlation is unity. We can therefore expect vanishing eigenvalues of our system matrix. This issue must however be judged immaterial: no need indeed to regenerate identical signals twice. Channel correlation is quantified by the number of energetical coherent structures. If all channels are exactly correlated for instance – that is, all channels are the replication of one single master signal – there will be one unique (self-)coherent structure with non-zero variance. Only the expansion coefficient of that structure needs to be dealt with. This sheds light on a very beneficial property for our change of basis. Increasing the spatial discretization accuracy leads to more coherent structures, but not to more energetical coherent structures: there is in this sense statistical convergence towards continuum. For our computations, we have used 50 coherent structures. This restriction is effectively imposed by the numerical conditioning of (3.17), but precisely this restriction provides us with the assurance that those 50 structures accurately describe the inflow statistics.

The authors are grateful to Dr Philippe Druault (Université Pierre et Marie Curie) for comments on the manuscript. This work was supported through grant-in-aid for JSPS Fellows (19-07821) as well as grant-in-aid for Scientic Research (A) by the Japan Society for the Promotion of Science (JSPS).

REFERENCES

Akaike, H. 1973 Block Toeplitz matrix inversion. *SIAM J. Appl. Math*. **24** (2), 234–241.

Bamieh, B. & Dahleh, M. 2001 Energy amplification in channel flows with stochastic excitation. *Phys. Fluids* **13** (11), 3258–3269.

Bartels, R. H. & Stewart, G. W. 1972 Solution of the matrix equation $AX + XB = C$: algorithm 432. *Commun. ACM* **15**, 820–826.

Berkooz, G., Holmes, P. & Lumley, J. L. 1993 The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.* **25**, 539–575.

Butler, K. M. & Farrell, B. F. 1992 Three-dimensional optimal perturbations in viscous shear flow. *Phys. Fluids* **4** (8), 1637–1650.

Chevalier, M., Hoepffner, J., Bewley, T. R. & Henningson, D. 2006 State estimation in wall-bounded flow systems. Part II. Turbulent flows. *J. Fluid Mech.* **552**, 167–187.

Chomaz, J.-M. 2005 Global instabilities in spatially developing flows: non-normality and nonlinearity. *Annu. Rev. Fluid Mech.* **37**, 357–392.

Cybenko, G. 1980 The numerical stability of the Levinson–Durbin algorithm for Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput*. **1** (3), 303–319.

Druault, P., Lardeau, S., Bonnet, J.-P., Coiffet, J., Delville, J., Lamballais, E., Largeau, J.-F. & Perret, L. 2004 Generation of three-dimensional turbulent inlet conditions for large-eddy simulation. *AIAA J.* **42** (3), 447–456.

Druzhinin, O. A. & Elghobashi, S. E. 2001 Direct numerical simulation of a three-dimensional spatially developing bubble-laden mixing layer with two-way coupling. *J. Fluid Mech.* **429**, 23–61.

Durbin, J. 1960 The fitting of time-series models. *Rev. Inst. Intl Stat.* **28**, 233–244.

Farrell, B. F. & Ioannou, P. J. 1993 Stochastic forcing of the linearized Navier–Stokes equations. *Phys. Fluids* **5** (11), 2600–2609.

Fukagata, K., Kasagi, N. & Koumoutsakos, P. 2006 A theoretical prediction of friction drag reduction in turbulent flow by superhydrophobic surfaces. *Phys. Fluids* **18** (051703).

Hayes, M., H. 1996 *Statistical Digital Signal Processing and Modeling*. Wiley.

Ho, C.-M. & Huerre, P. 1984 Perturbed free shear layers. *Annu. Rev. Fluid Mech.* **16**, 365–424.

Hoepffner, J., Chevalier, M., Bewley, T. R. & Henningson, D. 2005 State estimation in wall-bounded flow systems. Part I. Perturbed laminar flows. *J. Fluid Mech.* **534**, 263–294.

Huerre, P. & Monkevitz, P. A. 1990 Local and global instabilities in spatially developing flows. *Annu. Rev. Fluid Mech.* **22**, 473–537.

Keating, A., Piomelli, U., Balaras, E. & Kaltenbach, H.-J. 2004 *A priori* and *a posteriori* tests of inflow conditions for large-eddy simulation. *Phys. Fluids* **16** (12), 4696–4712.

Klein, M., Sadiki, A. & Janicka, J. 2003 A digital filter based generation of inflow data for spatially developing direct numerical or large eddy simulations. *J. Comput. Phys.* **186**, 652–665.

di Mare, L., Klein, M., Jones, W. P. & Janicka, J. 2006 Synthetic turbulence inflow conditions for large-eddy simulation. *Phys. Fluids* **18**, 025107.

Perret, L., Delville, J., Manceau, R. & Bonnet, J.-P. 2008 Turbulent inflow conditions for large-eddy simulation based on low-order empirical model. *Phys. Fluids* **20**, 075107.

Reddy, S. C. & Henningson, D. 1993 Energy growth in viscous channel flow. *J. Fluid Mech.* **252**, 209–238.

Schmid, P. & Henningson, D. 2001 *Stability and Transition in Shear Flows*. Springer.

Vijayan, R., Poor, V., Moore, J. B. & Goodwin, G. C. 1991 A Levinson-type algorithm for modeling fast-sampled data. *IEEE Trans. Autom. Control* **36** (3), 314–321.

Whittle, P. 1963 On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix. *Biometrika* **50** (1 and 2), 129–134.