Annales de l'ISUP
2015, Volume 59 Issue 1-2 Pages 57-85
http://archimer.ifremer.fr/doc/00320/43080/

Reduced flow models from a stochastic Navier-Stokes representation

Resseguier Valentin [1, 2, *], Memin Etienne [1], Chapron Bertrand [2]

[1] Inria Rennes, Bretagne Atlantique, Fluminance team, 35042 Rennes, France
[2] IFREMER, LOS, 29280 Plouzané, France

* Valentin Resseguier, email address : valentin.resseguier@inria.fr

Abstract :

In large-scale Fluids Dynamics systems, the velocity lives in a broad range of scales. To be able to simulate its large-scale component, the flow can be de- composed into a finite variation process, which represents a smooth large-scale velocity component, and a martingale part, associated to the highly oscillating small-scale velocities. Within this general framework, a stochastic representation of the Navier-Stokes equations can be derived, based on physical conservation laws. In this equation, a diffusive sub-grid tensor appears naturally and gener- alizes classical sub-grid tensors. Here, a dimensionally reduced large-scale simulation is performed. A Galerkin projection of our Navier-Stokes equation is done on a Proper Orthogonal De- composition basis. In our approach of the POD, the resolved temporal modes are differentiable with respect to time, whereas the unresolved temporal modes are assumed to be decorrelated in time. The corresponding reduced stochastic model enables to simulate, at low computational cost, the resolved temporal modes. It allows taking into account the possibly time-dependent, inhomoge- neous and anisotropic covariance of the small scale velocity. We proposed two ways of estimating such contributions in the context of POD-Galerkin. This method has proved successful to reconstruct energetic Chronos for a wake flow at Reynolds 3900, even with a large time step, whereas standard POD- Galerkin diverged systematically. This paper describes the principles of our stochastic Navier-Stokes equation, together with the estimation approaches, elaborated for the model reduction strategy.

Keywords : Stochastic calculus, uid dynamics, large eddy simulation, Proper Orthogonal Decomposition, reduced order model, uncertainty quantification

1. Introduction

Modeling accurately and understanding geophysical uid dynamics is a main issue in current researches. Indeed, beyond economic applications linked to weather forecasting, the need for accurate climate projections is becoming more and more important. Studying such systems using physics is challenging, especially in regard to the non-linearity of the Navier-Stokes equations. Since these equations make large-scale velocities interact with small-scale velocity uctuations, the main velocity tendency cannot be simulated alone ([1] and [2]). The e_ects of the unresolved small-scale, so-called turbulent, uctuations have to be taken into account.

A first way of modeling 3D turbulence is to only consider the direct energy cascade, which means that energy goes from the large scales to smaller scales, until it is dissipated by molecular viscosity ([3]). This can be done by assuming that the large-scale velocity follows the Navier-Stokes equation with an additive diffusive term, parametrized by a so-called eddy viscosity ([4]). Such types of additive terms are called sub-grid tensors, since they represent the effect of velocity living at smaller scale than the simulation grid. However, the diffusion matrix or diffusion coefficient and its temporal and spatial dependence have to be determined. In many cases, it is done empirically and/or using scaling assumptions. This is the case for LES (Large Eddies Simulation) where the large-scale flow is defined by a spatially low-pass-filtered velocity ([5],[6]) and RANS (Reynolds Average Numerical Simulation) where the large-scale component is defined by a statistical average of the velocity. The same type of models are used for large-scale modeling of tracers evolution, using eddy diffusivity instead of eddy viscosity.
Another drawback of this approach is the assumption of a permanent direct energy cascade. In real systems, there are intermittent back-scattering of energy from smaller scales toward the larger one. Therefore, some authors proposed to include terms that artificially bring energy to the system. Sometimes, it is done by a locally negative eddy viscosity ([7]). Another solution consists in setting up a system forced by a Gaussian process decorrelated in time ([8]). The spatial covariance of this forcing is a parameter that has to be determined. A stationary assumption greatly simplifies the problem and the associated model.
Considering a random velocity is now widely used ([1], [2] and [9]). In addition to theoretical physics constraints, it enables uncertainty quantification, and the use of ensemble based methods such as filtering ([10], [11] and [12]). The stochastic model, described in the previous paragraph, adds a random force to the equation without deep theoretical justification. This additive noise can be interpreted as an explicit error of the model. However, within this prospect, why would an additive noise be more adequate than a multiplicative noise or any other model of noise? According to [13], in reduced models for geophysical fluid dynamics applications, the coupling of an additive and a multiplicative noise is a good choice, leading naturally to heavy tails processes. But as far as we know, no theoretical justifications of this choice have been provided.
The study [14] and this paper follow another approach, introduced by [15] and [16]. The aim is to bring up naturally a physically based uncertainty quantification and a sub-grid-tensor model without strong assumptions. The velocity is assumed to be random and partially decorrelated in time. From the Lagrangian point of view, it defines a general semimartingale flow. Using stochastic calculus and classical fluid dynamics principles, one can prove a stochastic representation of the so-called Reynolds transport theorem. It describes the time-space evolution of a scalar transported by this semimartingale flow. Using energy and mass conservation, it leads to time-space evolution of respectively the temperature and the density. Then, the transport theorem applied to momentum and the second Newton Law lead to a stochastic version

of the Navier-Stokes equations. The corresponding scalar and velocity evolution laws involve an inhomogeneous time-dependent anisotropic diffusive sub-grid-tensor and additive and multiplicative noise.
For some industrial applications, the resolution of a system of partial differential equations may be too time consuming. A solution consists in deriving a model of reduced dimension, like in the case of the Proper Orthogonal Decomposition (POD) [17]. Within this model, the velocity at a fixed time, t, is assumed to live in a small dimensional subspace of functions of space. The basis of this subspace, so-called spatial modes, is determined by a Principal Component Analysis (PCA) on a sequence of velocity snapshots. The coefficient of the velocity in the reduced basis are called temporal modes. Then, the partial differential equation of interest is projected on the function of this basis. It leads to a finite set of coupled ordinary differential equations which describe the time evolution of the temporal modes. One problem for non-linear models such as Navier-Stokes equations is that keeping only a small number of modes destabilizes the system. In order to overcome this, some authors empirically add a diffusive term, parametrized by an eddy viscosity model to the reduced model. Several modeling of this eddy viscosity have been proposed. For instance, [18] proposed a constant coefficient, whereas [19] introduced the modal model with one eddy viscosity per mode. Recently, [20] and [7] proposed an eddy viscosity model that depends on the instantaneous energy of the temporal modes. Other authors ([21]) perform non-linear Galerkin methods, with the same spatial modes. It leads to another form of the reduced model, that will not be investigated in this paper.
In our approach, the unresolved temporal modes are assumed to be random and decorrelated in time whereas the resolved ones are deterministic. Thus, according to our stochastic Navier-Stokes model, an explicit sub-grid tensor appears both in the PDE and in the associate reduced model. The parameters of this sub-grid tensor can then be easily estimated on the residual velocity, through a statistical estimator. By residual velocity, we mean the part of the velocity snapshots which is not represented by the PCA. As will be demonstrated here, this sub-grid tensor successfully stabilizes the reduced system.
The paper is organized as follow. The first section presents the stochastic fluid dynamics model, on which we rely. The second section is a reminder of the classical POD approach. The third one presents our POD based reduced model under uncertainty. The fourth section presents some numerical results and comparisons. Finally, the last section concludes and provides perspectives.

2. The proposed stochastic model

In this work, an Eulerian stochastic description of the velocity and tracer evolution is used, as proposed in [14]. Unlike classical stochastic methods, a random part, encoding an uncertainty on the velocity expression, is added to the Lagrangian velocity before any model derivation. Thanks to this decomposition, a stochastic represen-

tation of the so-called Reynolds transport theorem, cornerstone of the deterministic fluid dynamic theory, can be derived. Thus, assuming a dynamical balance similarly to the second Newton law, a stochastic Navier-Stokes expression can be derived. It should be noticed that the equations, described below, are derived from fundamental physical laws only.
The time differentiation of a trajectory X_t of a particle is noted:

$$dX_t = w(X_t,t)dt + \sigma(X_t,t)dB_t, \tag{2.1}$$

where $\sigma(.,t)$ is an Hilbert-Schmidt operator on $(L^2(\mathbb{R}^d))^d$ defined by its kernel $\breve{\sigma}(.,.,t)$: $\forall f \in (L^2(\mathbb{R}^d))^d$, $\sigma(.,t)f \stackrel{\triangle}{=} \int_\Omega \breve{\sigma}(.,y,t)f(y)dy$ and $t \mapsto B(t)$ is a cylindrical I_d-Wiener process (see [22] and [23] for more information on infinite dimensional Wiener process and cylindrical I_d-Wiener process). Then, $(x,t) \mapsto \sigma(x,t)dB_t$ is a centered Gaussian process with the following covariance:

$$\forall x,y \in \mathbb{R}^d, \mathbb{E}\left((\sigma(x,t)dB_t)\left(\sigma(y,t')dB_{t'}\right)^T\right) \stackrel{\triangle}{=} a(x,y)\delta(t-t')dt,$$

where:

$$\begin{aligned} a(x,y)dt &= \int_\Omega \breve{\sigma}(x,z)\breve{\sigma}^T(y,z)dzdt \stackrel{\triangle}{=} \sigma(x)\sigma(y)^T dt \\ &= d\left\langle \int_0^t \sigma(x,t')dB_{t'}, \left(\int_0^t \sigma(y,t'')dB_{t''}\right)^T \right\rangle. \end{aligned}$$

The notation $< f,g >$ is the quadratic cross-variation of f and g, used in stochastic calculus, and its expression is recall in Appendix A. The term $\int_0^t wdt'$ represents the large-scale part of the flow whereas $\int_0^t \sigma dB_{t'}$ represents the small-scale part. The real physical small-scale flows are differentiable w.r.t. (with respect to) time. But, the time sampling used for large-scale modeling or observation is often larger than the smaller physical time scale of the real velocity. Thus, at this large scale time sampling, the smallest scales of the flow are non differentiable almost everywhere w.r.t. time.
The semimartingale Lagrangian formulation (2.1) together with stochastic calculus theory allows us differentiating and integrating random physical quantities. Some basic notions of stochastic calculus, concerning finite variation processes, martingales and semimartingales, are provided in Appendix A.

2.1. Stochastic representation of the Reynolds-transport theorem

Thanks to the previous decomposition, it is possible to derive a stochastic representation of the so-called Reynolds transport theorem. Unlike [14], we will not assume that, for each x,y and t, the matrix $\breve{\sigma}(x,y,t)$ is symmetric. Furthermore, the time differentiable part of the flow, $w(x,.)$, will not be assumed to be deterministic anymore, but rather, to be a continuous semimartingale. Nevertheless, if we exactly follow the same procedure, the very same stochastic transport theorem can be derived.

Theorem 2.1. *Stochastic Reynolds transport theorem*

Noting ϕ the stochastic flow defined by:

$$\forall x \in \Omega, t \in \mathbb{R}^+, \phi(x,t) = x + \int_0^t w(\phi(x,t'),t')dt' + \int_0^t \sigma(\phi(x,t'),t')dB_{t'},$$

and denoting $V(t) = \phi(V(0),t)$ a material volume transported by the stochastic flow, we have:

(2.2)
$$d\int_{V(t)} q(x,t)dx = \int_{V(t)} \left(d_t q + \nabla \cdot \left(q dX_t + q\sigma(\nabla\cdot\sigma)^T dt - \frac{1}{2}\nabla\cdot(aq)^T dt \right) \right) dx.$$

The mathematical equivalence between formulation (2.2) and the stochastic Reynolds transport theorem of [14] is proven in Appendix B. If q is a passive tracer, transported by the stochastic flow, $d\int_{V(t)} q(x,t)dx = 0$, and:

(2.3)
$$q(x,t) - q(x,0) = -\int_0^t \nabla\cdot\left(qw + q\sigma(\nabla\cdot\sigma)^T - \frac{1}{2}\nabla\cdot(aq)^T \right) dt' - \int_0^t \nabla\cdot(q\sigma)\, dB_{t'}.$$

This equation is the unique decomposition of the continuous semimartingale q, into a finite variation process (the integral in dt) and a local martingale (the integral in dB_t) [24]. Physically, the finite variation process varies slowly and is responsible of the large time-scale variation of q, whereas $\nabla\cdot(q\sigma)\, dB_{t'}$ is decorrelated in time and null in average. From this point of view, the two components live in two different spaces, and hence the semimartingale decomposition is unique. If we make the hypothesis of a constant density ρ, the last equation applied to $q = \rho$ and the uniqueness of the decomposition leads to:

$$0 = \nabla\cdot\sigma, \tag{2.4}$$

$$0 = \nabla\cdot\left(w + \sigma(\nabla\cdot\sigma)^T - \frac{1}{2}(\nabla\cdot a)^T \right) = \nabla\cdot\left(w - \frac{1}{2}(\nabla\cdot a)^T \right). \tag{2.5}$$

Usually, the evolution of an intensive property, q, can be computed from equation (2.3), through the knowledge of the small-scale velocity characteristic, σ, and of the large-scale drift, w. Indeed, the evolution of all intensive property statistical moments can be formalized through equation (2.3). For instance, the equation of the conditional expectation of the scalar, given the velocity w for all time, $\bar{q} \overset{\triangle}{=} \mathbb{E}(q|w)$, is:

$$\frac{\partial \bar{q}}{\partial t} + \nabla\cdot(\bar{q}w^*) = \nabla\cdot\left(\frac{1}{2}a\nabla\bar{q}\right) \text{ where } w^* = w + \sigma(\nabla\cdot\sigma)^T - \frac{1}{2}(\nabla\cdot a)^T. \tag{2.6}$$

This is a classical advection-diffusion equation. Indeed, since a is symmetric positive-semidefinite, $\nabla\cdot\left(\frac{1}{2}a\nabla\bar{q}\right)$ leads only to diffusion. The expectation, $\bar{q}$, is advected by

an effective drift, w^*, and undergoes a diffusion through the tensor $\frac{1}{2}a$. In the case of a constant density, w^* is naturally divergence-free (see equation (2.5)). For large-scale tracers, this advection-diffusion equation, derived from physical laws, has the same form as the widely used empirical advection-diffusion equation setup through an eddy diffusivity assumption ([3]). However, unlike most of these classical models, the sub-grid diffusion we got is time-dependent, anisotropic and inhomogeneous.

2.2. Stochastic Navier-Stokes model

Similarly to the Newton second law, a dynamical balance between the temporal differentiation of the stochastic momentum, ρdX_t, and general stochastic forces action is assumed. This leads, applying (2.2) to ρdX_t and ρ, to the following stochastic Navier-Stokes representation.

Theorem 2.2. *Stochastic Navier-Stokes representation*
If w is a finite variation process and f the integral of the pressure p along time can be decomposed as a general continuous semimartingale $\int_0^t (p'dt + d\hat{p})$, then

$$\rho\left(\frac{\partial w}{\partial t} + (w\cdot\nabla)w + f\times w\right) = \tau(w) + \rho g - \nabla p' + f_V(w), \tag{2.7}$$

$$\rho\left((\sigma dBt\cdot\nabla)w + f\times\sigma dBt\right) = -\nabla d\hat{p} + f_V(\sigma)dBt, \tag{2.8}$$

where

$$\begin{cases} f_V(h) &= \mu\left(\nabla^2 h + \frac{1}{3}\nabla\left(\nabla\cdot h\right)\right), \\ \forall k, \tau_k(w) &= \frac{1}{2}\left(\nabla\cdot\left(\nabla\cdot\left(\rho a w_k\right)\right)^T - \nabla\cdot\left(\nabla\cdot\left(\rho a\right)\right)^T w_k - 2*\rho\left((\nabla\cdot\sigma)\sigma^T\nabla\right)w_k\right). \end{cases}$$

As a consequence, if the large-scale component, w, is a finite variation process (i.e. if it is time differentiable) and if the density ρ is deterministic, then w is deterministic, knowing the initial conditions. It can be noted again that the kernel $\breve{\sigma}$ is not assumed pointwise symmetric. The equivalence between formulation (2.7) and the stochastic Navier-Stokes model of [14] is proven in Appendix B.
Expression (2.7) can be seen as a generalization of several classical turbulence models. For instance, if the small-scale infinitesimal displacement σdB_t is isotropic and divergence free, and if the density is constant, the sub-grid tensor simplifies to $\tau(w) = \rho\frac{a}{2}\triangle w$. We retrieve the simplest expression of the Boussinesq assumption, with a constant eddy viscosity given by $\frac{a}{2}$. Generally speaking, we may wonder whether the sub-grid tensor, τ, is dissipative, like in a theoretical 3D direct energy cascade ([3], [8]). If ρ is assumed to be constant, and if a or w and their derivatives are assumed to be null on the border of Ω, then τ is dissipative. The proof is provided in Appendix C.
The knowledge of small-scale physical flow realizations allows estimating σ and a:

$$adt = \mathbb{E}((\sigma dB_t)(\sigma dB_t)^T) \text{ and } (\nabla\cdot\sigma)\sigma^T dt = \mathbb{E}((\nabla\cdot\sigma dB_t)(\sigma dB_t)^T). \tag{2.9}$$

Thus, the value of $a(x,t)$ can be used in a large-scale simulation ruled by equation (2.7). The tensor a and functions of σ can also be estimated from a single realization thanks to stochastic calculus, as explained later, or by assuming local time or space ergodicity as in [25]. Another interesting way of using (2.9) is through Monte-Carlo small-scale simulations, such as particle filtering, where particles correspond to several probable values of the small-scale velocity. A third alternative to estimate a and σ, without using (2.9), consists in relying directly on known statistical properties of small-scale measurements. Some works on this subject are currently ongoing. In all these methods, the estimation of the tensor a corresponds to a solution of a closure problem. Knowing the value of a should lead to a simulation of the drift, through (2.7), or of a tracer transport, through the stochastic transport theorem ([14]), in which the small-scale actions are taken into account in a statistical way.
It is also possible to follow a dual strategy with a downscaling approach, like, for instance, mixing diagnostics, which is, at the moment, an important issue in Meteorology and Oceanography ([26] and [27]). The evolution equation of both the averaged stochastic transport theorem (2.6) and the average stochastic Navier-Stokes model (2.7), applied to observed large-scale geophysical data can give information on the tensors a and σ. Indeed, analysing how a tracer is advected and diffused may help computing these two tensors. Related information, such as the local small-scale energy, the anisotropy created by the matrix $a(x,x,t)$, or the local divergence, which is linked to $\nabla\cdot\sigma$, can be inferred. These information could teach us, for instance, the likelihood of locally strong velocity or tracer gradient and the principal directions of mixing created by the main variance directions.
Theorems 2.1 and 2.2 provide the foundations of a physically relevant stochastic Fluid Dynamics framework. In this paper, we will rely on them for a reduced model application.

3. Classical model reduction using POD

Dimensional reduction techniques are methods allowing simplification of Partial Differential Equations (PDE), using dedicated basis specified from observed data. The Proper Orthogonal Decomposition (POD) is one of these methods, and below are recalled its main principles.
Here, we consider an observed multivariate field such as a velocity $u(x,t)$ depending on space $x \in \Omega$ and time $t \in [0,T]$. The goal consists in looking for a subspace of reduced dimension where $u(.,t)$ is likely to live for all t. We thus seek a finite orthonormal set of function of space, which spans this subspace. These functions $(\phi_i(x))_{1\leqslant i\leqslant N}$ are called spatial modes or *Topos* and are computed from a Karunen Loeve decomposition on a series of available velocity snapshots. In other words, a spectral analysis is done on the space (or time) autocorrelation tensor of observed

data:

$$\bar{u}(x) \triangleq \frac{1}{T}\int_0^T u(x,t)dt, \tag{3.1}$$

$$cov(x_1,x_2) \triangleq \frac{1}{T}\int_0^T (u(x_1,t)-\bar{u}(x_1))(u(x_2,t)-\bar{u}(x_2))^T dt, \tag{3.2}$$

$$\int_\Omega cov(x_1,x_2)\phi_i(x_1)dx_1 = \lambda_i\phi_i(x_2) \text{ with } \int_\Omega \phi_i(x)\phi_j(x)dx = \delta_{i,j}. \tag{3.3}$$

The *Topos* are sorted such that $\lambda_1 > ... > \lambda_N$, where N is the number of observed snapshots (if the number of points of the spatial grid is larger than N). It leads to the decomposition:

$$\forall (x,t) \in \mathbb{R}^d \times \mathbb{R},\ u(x,t) \approx \bar{u}(x) + \sum_{i=1}^{N} b_i(t)\phi_i(x). \tag{3.4}$$

The values $(b_i(t))_{1\leqslant i\leqslant N}$ are called temporal modes or *Chronos* and satisfy:

$$\forall i,j, \frac{1}{T}\int_0^T b_i(t)b_j(t)dt = \lambda_i\delta_{i,j}. \tag{3.5}$$

In the following, $\bar{u}$ will be denoted ϕ_0 and $b_0 \triangleq 1$. Then, since only the first temporal modes concentrate the most significant part of the energy, a second truncation approximation is usually performed:

$$\forall (x,t) \in \mathbb{R}^d \times \mathbb{R},\ u(x,t) \approx \sum_{i=0}^{n} b_i(t)\phi_i(x) \text{ with } n \ll N. \tag{3.6}$$

A Galerkin projection enables us to look for an approximate solution of a PDE. The approximate solution at time t, $u(.,t)$, defined in (3.6), is assumed to live in a finite-dimensional sub-space, spanned by $(\phi_0,...,\phi_n)$, instead of an infinite-dimensional one. The time-space evolution equation of u (a PDE) is then expressed as the time evolution equations (a finite set of coupled ODEs) of *Chronos*. In fluid dynamics, PDE system describing the velocity evolution, such as the Navier-Stokes equations, have the general following abstract form:

$$\frac{\partial u}{\partial t} = I + L(u) + C(u,u), \tag{3.7}$$

where L and C are respectively linear and bilinear differential operators. The first term, I, includes pressure and gravity. The second one, L, involves molecular viscosity and possibly Coriolis force. The last one, C, encodes the non-linear advection

term. Projecting this PDE on each *Topos* leads to:

$$\forall i \leqslant n,\ \frac{db_i}{dt} = \underbrace{\left(\int_\Omega \phi_i \cdot I\right)}_{\triangleq i_i} + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \phi_i \cdot L(\phi_p)\right)}_{\triangleq l_{p,i}} b_p + \sum_{p,q=0}^{n} \underbrace{\left(\int_\Omega \phi_i \cdot C(\phi_p, \phi_q)\right)}_{\triangleq c_{p,q,i}} b_p b_q.$$

Because of the non-linearity, the temporal modes strongly interact with each others. In particular, even though the original model (with $n = N$) is computationally stable for moderate Reynolds number, the reduced one is generally not so. This particularity of the Navier-Stokes equation is not restricted to the POD framework. Simulating a large-scale flow, considering only the largest Fourier modes, leads also to strong instabilities and numerical explosions. A rough truncation cannot be considered without introducing a dissipative term whose role is to drain the energy brought by the larger modes beyond the truncation and thus to avoid an energy accumulation. Eddy viscosity models, which consist in enforcing the fluid viscosity, are often used for that purpose. This principle, which dates back to Boussinesq ([4]), is often used in large-scale simulation as well as in the context of POD ([18], [19], [20], [7]). In practice, these methods introduce empirically an additional damping term to the Navier-Stokes equation. This leads to a modified linear term in (3.7). Unfortunately, since this term is built from an empirical thermodynamical analogy, its precise form is difficult to justify and its parametrization has to be tuned for each simulation to get optimal results. The method proposed in the next section allows us to tackle these drawbacks.

4. Stochastic POD

To overcome the difficulties developed previously, we suggest to use our stochastic Navier-Stokes model instead of the classical Navier-Stokes equations. Let us outline that both systems address the same physics. They both rely on mass and momentum conservation and differ only in how they are taking into account small-scale missing information.

4.1. Model

The reduced dynamic system we propose is based on the stochastic Navier-Stokes model developed in (2.2), assuming that the density ρ is constant and the smooth part of the flow, w, is of bounded variations. To tackle the problem of modes interactions, [14] proposed to decompose u as follow : $udt = wdt + \sigma dB_t$ with $w = \sum_{i=0}^{n} b_i \phi_i$

(projection on the truncated subspace) and $\sum_{i=n+1}^{N} b_i \phi_i dt$ a realization of σdB_t (projection on the complementary "small-scale" subspace). Since $\nabla \cdot u = 0$, for all i, $\nabla \cdot \phi_i = 0$ and, then, $\nabla \cdot w = 0$. The drift, w, follows the finite variation part of the stochastic Navier-Stokes equation (2.7) in the incompressible case. Projecting on the divergence-free functions space, we have:

$$\frac{\partial w}{\partial t} + \mathcal{P}\left((w \cdot \nabla)w\right) = \mathcal{P}\left(\frac{1}{\rho}\tau(w)\right) + g + \nu \triangle w, \tag{4.1}$$

where

$$\frac{1}{\rho}\tau_k(w) = \frac{1}{2}\left(\nabla \cdot (\nabla \cdot (a w_k))^T - \nabla \cdot (\nabla \cdot a)^T w_k\right), \tag{4.2}$$

$$\mathcal{P} \stackrel{\triangle}{=} \mathbb{I}_d - \triangle^{-1} \nabla \nabla^T. \tag{4.3}$$

In Fourier space, the projector on divergence-free functions space reads $\hat{\mathcal{P}} = \mathbb{I}_d - \frac{k k^T}{\|k\|_2^2}$. Equation (4.1) can be rewritten as:

$$\frac{\partial w}{\partial t} = I + L(w) + C(w, w) + F(a, w), \tag{4.4}$$

where F is a bilinear differential operator. Projecting this equation along ϕ_i for each $i \in [\![1, n]\!]$, gives the evolution equation of $b \stackrel{\triangle}{=} (b_i)_{1 \leqslant i \leqslant n}$.

$$\forall i \in [\![1, n]\!], \frac{db_i}{dt} = i_i + \left(l_{\cdot i} + \breve{f}(a)_{\cdot i}\right)^T b + b^t c_{\cdot\cdot i} b, \tag{4.5}$$

$$\text{with } \breve{f}(a)_{j,i} \stackrel{\triangle}{=} \int_\Omega \phi_i \cdot F(a, \phi_j), \tag{4.6}$$

where the coefficients $(i_i)_{1 \leqslant i \leqslant n}, (\breve{f}(a)_{j,i})_{1 \leqslant i,j \leqslant n}, (l_{j,i})_{1 \leqslant i,j \leqslant n}$ and $(c_{k,j,i})_{1 \leqslant i,j,k \leqslant n}$ are computed through the integration over the whole space of the terms of (4.1). Those dynamical coefficients depend on both the resolved *Topos* and the unresolved velocity variance tensor, a. This system includes a natural small-scale dissipation mechanism, through the tensor τ. To fully specify this system, we need to estimate the quadratic variance tensor a. This important issue is developed in subsection 4.3. But first we will elaborate further on the choice of a characteristic time step related to the truncation operated.

4.2. Choice of the time step

For several applications, the simulation of the most energetic large-scale component of the solution is sufficient. However, this simulation needs to be fast, implying a low complexity evolution model and a large time step. The structure of our stochastic model enables to reach both goals.
Indeed, as long as the resolved modes, which represent w, are differentiable w.r.t.

time, our stochastic reduced model is valid. Thus, the time step can be chosen as large as desired, as long as these modes remain smooth. The Shannon-Nyquist sampling theorem provides a natural upper bound to fix this time step. This theorem states that a function can be sampled, without loss of information, if the sampling frequency is twice as large as the largest frequency of the original function. Otherwise, the sampled function undergoes an aliasing artifact characterized by a back folding of the Fourier spectrum. If the resolved POD modes and their evolution equations are not affected by aliasing phenomena, the required smoothness is assumed to be reached. Since the evolution equations are quadratic, a sufficient condition for the necessary smoothness is:

$$\frac{1}{\Delta t} \geqslant 4\ max_{i \leqslant n} \left(f_{max} \left(b_i \right) \right), \tag{4.7}$$

where $f_{max}\left(b_i\right)$ is the maximum frequency of the i-th temporal mode.
Of course, aliasing will occur in the unresolved temporal modes, associated to smaller time scales. However, our stochastic model is derived from a decorrelation assumption of the small-scale unresolved part of the velocity. A strong subsampling of these components strengthens the decorrelation property of these modes.

4.3. Estimation of the uncertainty variance tensor

After having estimated the *Topos* and fixed the time step, we need to estimate the uncertainty variance tensor a. This estimation will enable us to get a full expression of the dynamical coefficients of the *Chronos* evolution equations (4.5). To that end, additional modeling assumptions must be imposed. The first natural hypothesis consists in assuming an uncertainty field that is stationnary in time – and spatially non homogenous. In this stationary case, the uncertainty variance tensor is constant in time.

4.3.1. The uncertainty variance tensor is constant in time

This case corresponds to the assumption used in [14]. To understand the consequence of this hypothesis, we recall that $\frac{a}{\Delta t}$ is the variance of the residual velocity $u-w$. The process is decorrelated in time and Gaussian. The snapshots are hence independent. Therefore, the $(u-w)(u-w)^T(t_i)$ are independent and identically distributed. So, the expectation, $\frac{a}{\Delta t}$, can be computed by averaging the snapshots $(u-w)(u-w)^T(t_i)$. In other words, the process is ergodic.

$$\frac{a(x)}{\Delta t} = \lim_{N \to +\infty} \frac{1}{N} \sum_{i=1}^{N} (u-w)(x,t_i)((u-w)(x,t_i))^T, \tag{4.8}$$

where N is the number of snapshots after time sub-sampling. The convergence is almost sure by the strong law of large numbers but, here, only the convergence in

probability is used. One can notice that this relation is straightforward when using stochastic calculus, as explained hereafter.
Thanks to the expression above, one can see several advantages of such an assumption. First of all, the construction of a is straightforward and easy to compute. Secondly, σdB_t itself is a a-Wiener process, since $a(x,y)$ is a trace class operator constant in time (see [23] and [22] for more details on Q-Wiener processes). The spectrum of this operator, which is central in the model reduction process, enables us to use a diagonalized version of σ and a as in ([14]):

$$a(x,y) = \sum_{k=n+1}^{\infty} \lambda_k \Delta t \ \phi_k(x)\phi_k(y)^T \approx \sum_{k=n+1}^{N} \lambda_k \Delta t \ \phi_k(x)\phi_k(y)^T, \tag{4.9}$$

$$\sigma(x)dB_t = \sum_{k=n+1}^{\infty} \sqrt{\lambda_k \Delta t} \ \phi_k(x) d\beta_t^{(k)} \approx \sum_{k=n+1}^{N} \sqrt{\lambda_k \Delta t} \ \phi_k(x) d\beta_t^{(k)}, \tag{4.10}$$

where the $(\beta^{(k)})_{k>n}$ are independent standard one-dimensional Brownian motions. As a result, it is very easy in this context to generate realizations of the small-scale uncorrelated component. The knowledge of the leading eigenfunctions of the POD complementary space allows us to access directly to the spectral representation of the diffusion tensor.

4.3.2. The uncertainty variance tensor is time varying

Assuming a constant value for $a(x,x)$ means that the turbulence is not intermittent. In the context of POD, it would mean that all the unresolved modes have a constant variance. It is a good first approximation. But, one may wonder whether it is possible to do better.
If a does depend on time, the estimation is more involved. Since only one realization of the small-scale velocity is available, some time-ergodicity hypothesis would be necessary, at least locally, to use (2.9) as in ([25]). Otherwise, parametric and non parametric estimation of $a(X_t,t)dt = d < X, X >_t$ are studied in the literature ([28], [29], [30], [31], [32] and [33]). Parametric ones use for instance maximum likelihood estimation. Indeed, denoting θ the parameters, the Girsanov theorem ([34]) leads, as explained in [28], to the following log-likelihood:

$$\begin{aligned} l(X_t - X_0|\theta) \quad = \quad & \int_0^t w(t', X_{t'})^T \left(a^{op}(t', X_{t'}|\theta)\right)^{-1} dX_{t'} \\ & -\frac{1}{2}\int_0^t w(t', X_{t'})^T \left(a^{op}(t', X_{t'}|\theta)\right)^{-1} w(t', X_{t'}) dt', \end{aligned}$$

where:

$$a^{op}(t,x|\theta)(f) \triangleq \int_\Omega a(t,x,y|\theta) f(y) dy.$$

It is a very powerful tool because it can be used on Lagrangian data. However, the knowledge of the inverse of the infinite dimensional operator a^{op} is required. Moreover, here, we look rather for a non-parametric technique. These methods assume that σ is either constant in time or in space ([29], [30], [31], [32] and [33]). But, contrary to the main application domain of the literature applications (finance), we have here access to an Eulerian realization of the stochastic flow. For all $x \in \mathbb{R}^d$, it is then possible to build a spatially homogeneous local martingale $\tilde{X}_t^x \overset{\triangle}{=} \int_0^t \sigma(x,t)dB_t$. Its realization, $\int_0^t (u(x,t) - w(x,t))dt$, after time sub-sampling, enables to estimate, for all functions h_k,

$$\begin{aligned}
\int h_k(t)a(x,t)dt &= \int h_k(t)d\left\langle \tilde{X}^x, \left(\tilde{X}^x\right)^T \right\rangle_t, \\
&= \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_i=0}^{T} h^k(t_i)(\tilde{X}^x_{t_{i+1}} - \tilde{X}^x_{t_i})(\tilde{X}^x_{t_{i+1}} - \tilde{X}^x_{t_i})^T, \\
&\approx (\Delta t)^2 \sum_{t_i=0}^{T} h^k(t_i)(u-w)(x,t_i)((u-w)(x,t_i))^T,
\end{aligned}$$

where $\mathbb{P}-\lim_{\Delta t \to 0}$ stands for the limit in probability as the time step, Δt, approaches 0. The functions h^k can be a orthonormal basis of $L^2([0,T])$ such as wavelets ([29]). In [29], it is shown that such estimators have good statistical properties: local asymptotic normality of the integrated square errors, together with the rate of convergence of its bias and variance. Therefore, the influence of the sub-grid tensor on $\frac{d}{dt}b_i$ will be represented by time-dependent, linear coefficients as shown in (4.5).

4.3.3. The uncertainty variance tensor is in the span of the Chronos

The *Chronos* reduced basis $(b_i/\|b_i\|_{L^2([0,T])})_{1\leqslant i\leqslant n}$ provides a much better solution than a wavelet basis. As a matter of fact, this choice has three main advantages. First of all, since we are studying the time evolution of $(b_i)_{1\leqslant i\leqslant n}$, the slow time variations of a, which are consistent with the time variations of $(b_i)_{0\leqslant i\leqslant n}$, are the information most needed. The number of wavelets needed to represent these time variations would be a priori much larger than $n+1$. Secondly, we do not need to reconstruct a. Indeed, noting $z_i(x) = \frac{\overline{b_i(.)a(x,.)}}{\lambda_i}$ and using the fact that $a \to f(a)$ (defined by (4.6)) is linear,

$$\breve{f}(a) = \breve{f}\left(\sum_{k=0}^{n} b_k z_k\right) = \sum_{k=0}^{n} b_k \breve{f}\left(z_k\right). \tag{4.11}$$

Thus, (4.5) becomes:

$$\forall i \in [\![1,n]\!], \frac{db_i}{dt} = i_i + l_{.i}^T b + b^T\left(c_{..i} + f_{..i}\right)b, \tag{4.12}$$

$$\text{with } f_{pqi} \overset{\triangle}{=} \breve{f}_{qi}\left(z_p\right). \tag{4.13}$$

If one chooses the basis $(b_i)_{0\leqslant i\leqslant n}$, only $(z_i)_{0\leqslant i\leqslant n}$ is needed to compute f and, hence, to simulate the influence of a in the evolution of b_i. The calculation is thus more direct. One may notice that equations (4.12) define again a quadratic evolution system of *Chronos*. Thus, the criterion for time step choice, developed in 4.2, remains the same. The third advantage of this basis is that f does not depend on time, unlike the term $\breve{f}(a)$ in equation (4.5). It is thus faster to compute than $\breve{f}(a(t))$ at all time t, and requires a lower memory capacity to store it. But, above all and unlike any other basis, although the variance tensor is time dependent, the evolution system of $(b_i)_{1\leqslant i\leqslant n}$ remains autonomous. This is an unavoidable requirement for a forecasting task ([7]).
To simplify the equation and to be more precise in what follows, we will remove the constant balance in (4.12). In permanent regime, since the system is stable, one can assume that $\overline{\frac{db_i}{dt}} = 0$. Thus, noting $b_{1:n} = (b_i)_{1\leqslant i\leqslant n}$, we get:

$$\begin{aligned}(4.14)\quad \forall i \in [\![1,n]\!],\ \frac{db_i}{dt} = &-\sum_{k=1}^{n} \lambda_k \left(c+f\right)_{k,k,i} + b_{1:n}^T \left(c+f\right)_{1:n,1:n,i} b_{1:n} \\ &+ \left(l_{1:n,i} + \left(c+f\right)_{0,1:n,i}^T + \left(c+f\right)_{1:n,0,i}\right)^T b_{1:n}.\end{aligned}$$

This model leads, as you will see later on, to an improved accuracy and stability of the system. We however lose the direct sampling capability of the previous simple ergodic assumption.

5. Numerical results

The different variations of the proposed approach have been assessed and compared numerically on numerical data of a wake behind a cylinder at Reynolds 300 and 3900 ([5]).

5.1. Characteristics of the data

The fluid is incompressible: $\nabla \cdot u = 0$. At $x = 0$, there is a constant velocity $U = 1$ directed along $x > 0$. At $(x, y) = (5, 0)$, there is a motionless cylinder with an axis along the z axis. In permanent regime, it creates a Von Kármán vortex street behind the cylinder. A clockwise vortex is created at the bottom right of the cylinder, it breaks away from it and moves downstream. Then, a counter-clockwise vortex is created at the top right of the cylinder, breaks away from the first one and moves downstream, and so on. This periodic physical process makes the two first *Chronos* almost sinusoidal.
Figures 1 and 2 show the z component of the vorticity $\nabla \times u$ on horizontal section of the fluid. In Figure 1, the cylinder is cropped. The vorticity is a measurement of

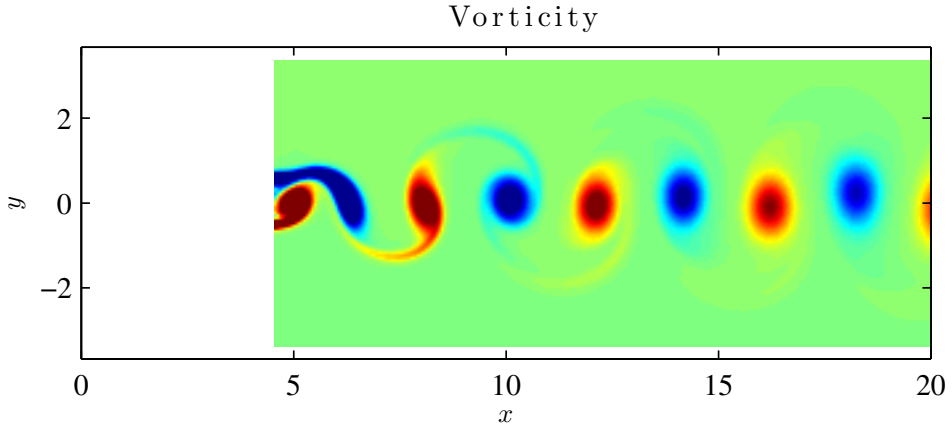

Figure 1: Vorticity along z of a wake behind a cylinder at Reynolds 300.

the rotation of the fluid on this plane. A positive vorticity (in red) means a counter-clockwise rotation. A negative vorticity (in blue) means a clockwise rotation. At this point, one can see, in both Figures, a counter-clockwise is breaking away at the right of the cylinder and a clockwise one is enlarging at the bottom right.

In Figure 2, Kelvin-Helmholtz instabilities at the top right and bottom right of the cylinder can be observed just before the vortex creation zone. At the top and bottom of the cylinder, the velocity is close to the inflow velocity $U = 1$ along x whereas, at the right of the cylinder, close to it, the velocity is close to zero. Thus, there are two mixing layers at the boundaries, at the top right and bottom right of the cylinder. These Kelvin-Helmholtz instabilities as well as the Von Kármán vortex street creates a turbulent wake downstream of the cylinder.
At Reynolds 300, there are only few small-scales features. Most of the energy and most of the dynamic is in large-scale structures. We use 80 vortex shedding. At Reynolds 3900, the turbulence is relatively important. Therefore, the spectrum support of the velocity is quite large, meaning that the velocity exists at several space and time scales. Indeed, one can see both small and large structures on Figure 2. Thus, in the context of POD, the *Chronos* live at different time scales. Since the spectrum is more energetic for lower wave-number, the first *Chronos*, i.e. the most energetic ones, have larger time-scale. Due to the quasi-periodic behavior of the flow along time, the *Chronos* are closed to the Fourier modes but not exactly equal. This analysis of the *Chronos* time scale is hence just a rough tendency. It explains nevertheless why our stochastic model, based on a separation between smooth and highly oscillating parts of the velocity is relevant.
Compared to the data of [5], we slightly filtered and sub-sampled them spatially in order to reduce by two the number of gridded points by axes. The Gaussian filtering is used here only to reduce a potential spatial aliasing. To speed-up and facilitate the

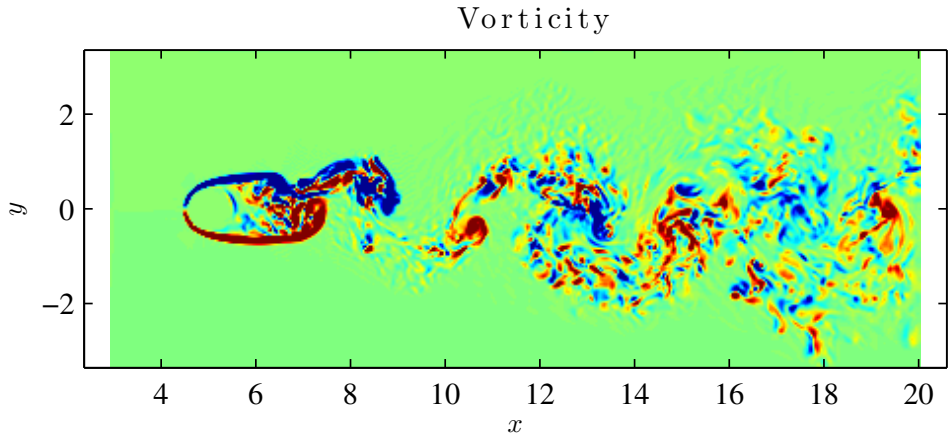

Figure 2: Vorticity along z of a wake behind a cylinder in the horizontal section $z = 0$, at Reynolds 3900.

computations, we also removed part of the space where the vorticity is negligible: at $|y| > 3.5$ and $x < 3$, as seen in Figure 2. We use $N = 251$ time steps to observe 3 vortex shedding.

5.2. Reconstruction of Chronos

To reconstruct the *Chronos*, the reduced order dynamical system (4.14) is used. The modes mean energy, $(\lambda_i)_{1\leqslant i\leqslant n}$, and the *Topos*, $(\phi_i)_{0\leqslant i\leqslant n}$, are computed from the whole sequence of snapshots ($N = 3999$ for Reynolds 300 and $N = 251$ for Reynolds 3900). As for the initial condition, we used the referenced values of the Chronos computed from the scalar product of initial velocity with the *Topos*. Then, regarding the *Chronos* spectra, an optimal time sub-sampling is chosen, as explained in subsection 4.2. Afterwards, using the *Topos*, the residual velocity and possibly the *Chronos*, the variance tensor, a, or its decomposition is estimated. The coefficients of the reduced order dynamical system of *Chronos* (see equation 4.14) are computed, using discrete derivation schemes and integration. Finally, the *Chronos* are recomputed, integrating (4.14) with a 4-th order Runge-Kutta method, with $(b_i^{ref})(t = 0)_{1\leqslant i\leqslant n}$ as initial condition.

Figures 3 and 4 show examples of the reconstruction of the *Chronos* for $n = 2$, at Reynolds 300, and $n = 10$, at Reynolds 3900, with the classical POD method (blue lines) and our method with a variance tensor defined as a linear combination of *Chronos* (red lines). At Reynolds 300, the first two modes explain most of the energy. That is why we consider only $n = 2$. The reference $(b_i^{ref})_{1\leqslant i\leqslant n}$ (black dots) are superimposed for comparison purpose. It can be observed that our model follows the references quite well whereas the deterministic model blows up. The divergence

occurs very quickly at Reynolds 3900. It may be pointed out that here both reduced models are parameter free. No constant had to be tuned to adapt any viscosity model.

Figures 5 and 6 show the error of the solution along time. The error is defined as follows:

$$\begin{aligned} err(t) &= T\frac{\left\|u^{ref}-u\right\|_{L^2(\Omega)}}{\|u^{ref}\|_{L^2(\Omega\times[0,T])}}, \\ &= T\frac{\left\|\sum_{i=1}^{n}\left(b_i^{ref}-b_i\right)\phi_i+\sum_{i=n+1}^{N}b_i^{ref}\phi_i\right\|_{L^2(\Omega)}}{\left\|\sum_{i=0}^{n}b_i^{ref}\phi_i\right\|_{L^2(\Omega\times[0,T])}}, \\ &= \left(\frac{\sum_{i=1}^{n}\left(b_i^{ref}-b_i\right)^2+\sum_{i=n+1}^{N}\left(b_i^{ref}\right)^2}{\|\bar{w}\|^2_{L^2(\Omega)}+\sum_{i=1}^{N}\lambda_i}\right)^{1/2}. \end{aligned}$$

Approximating the square of the real unresolved modes, $\left(\left(b_i^{ref}\right)^2\right)_{n+1\leqslant i\leqslant N}$, by their time average, $(\lambda_i)_{n+1\leqslant i\leqslant N}$, the error simplifies to:

$$err(t) \approx \left(\frac{\sum_{i=1}^{n}\left(b_i^{ref}-b_i\right)^2+\sum_{i=n+1}^{N}\lambda_i}{\|\bar{w}\|^2_{L^2(\Omega)}+\sum_{i=1}^{N}\lambda_i}\right)^{1/2}, \tag{5.1}$$

which is greater than the minimal error associated to the modal truncation:

$$err(t) \geqslant \left(\frac{\sum_{i=n+1}^{N}\lambda_i}{\|\bar{w}\|^2_{L^2(\Omega)}+\sum_{i=1}^{N}\lambda_i}\right)^{1/2}. \tag{5.2}$$

Equation (5.1) defines the criterion error plotted in Figures 5 and 6, whereas (5.2) constitutes a lower bound of this error.
Here, we use $u^{ref} \triangleq u - U$ as the reference solution, and U is the constant inflow velocity. The reference velocity is null, far from the cylinder and the integration of its energy, on the domain, does not depend on the size of the domain.
Figures 5 and 6 illustrate the error obtained for the standard POD Galerkin model without sub-grid dissipative term, and our model for a variance tensor which is either fixed constant along time or expressed as a linear combination of the *Chronos*. For the Reynolds 300, only the model with constant variance has been used. Indeed, this fluid dynamics system has only few degrees of freedom. For this Reynolds number, the model with variance tensor varying in time is overparameterized. The doted line represents the minimal error associated to the reduced subspace truncation error. The black solid line is the error considering only the time mean velocity – if we set

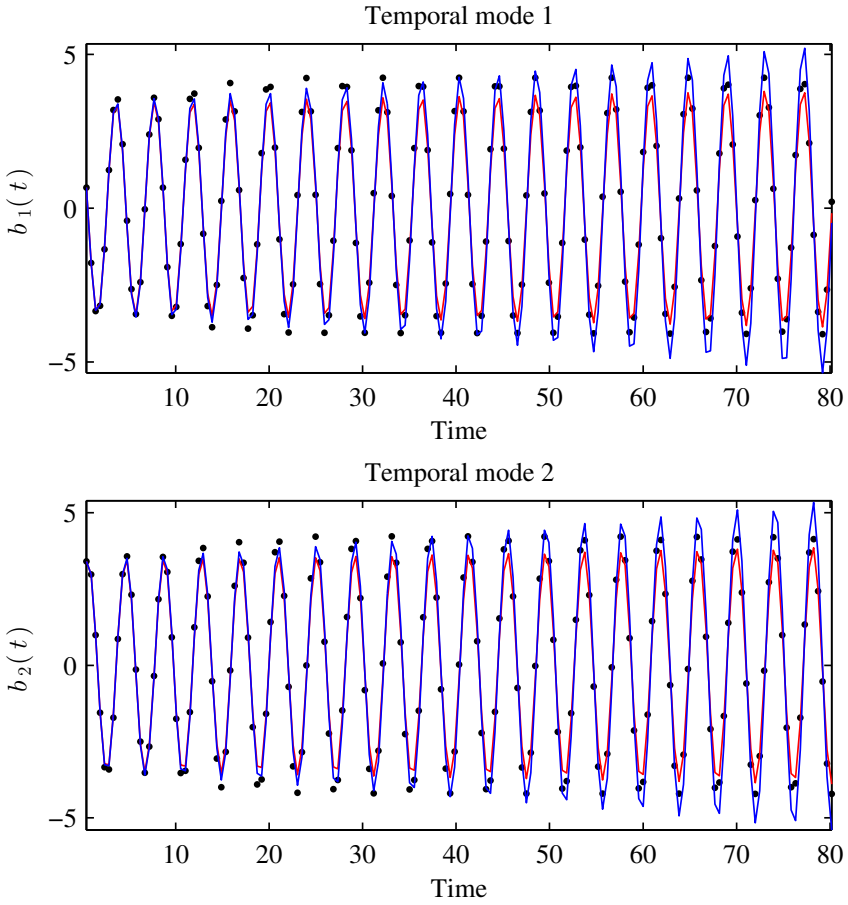

Figure 3: Reconstruction of the first two modes ($n = 2$), of a wake flow at Reynolds 300, with a variance tensor constant in time. The black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation, without any corrective coefficient. The initial condition, at $t = 0$, is common.

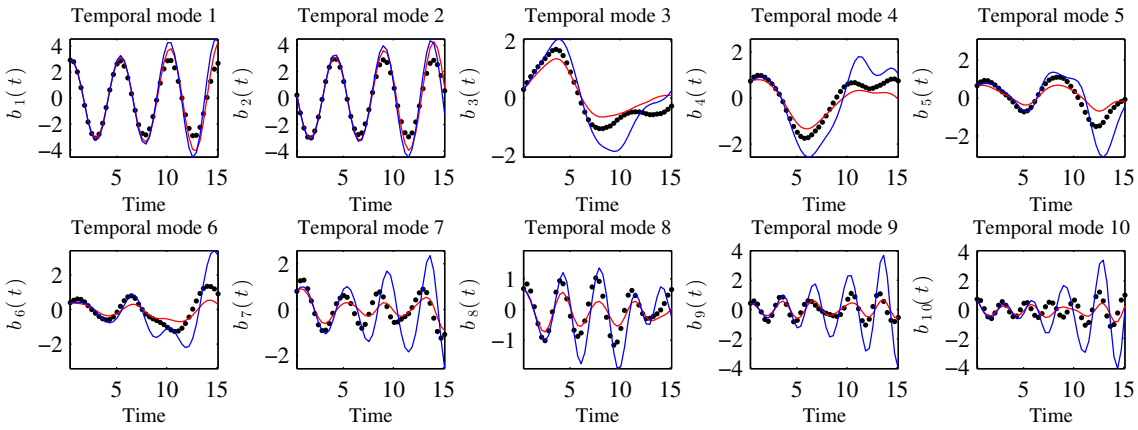

Figure 4: Reconstruction of the first ten modes ($n = 10$), of a wake flow at Reynolds 3900, with a variance tensor expressed as a linear function of the *Chronos*. The black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation, without any corrective coefficient. The initial condition, at $t = 0$, is common.

all the *Chronos* to 0.
In this case:

$$\begin{aligned} err_{|b=0}(t) &= T\frac{\left\|u^{ref} - \bar{u}\right\|_{L^2(\Omega)}}{\|u^{ref}\|_{L^2(\Omega\times[0,T])}}, \\ &= \left(\frac{\sum_{i=1}^{N}\left(b_i^{ref}\right)^2}{\|\bar{w}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}, \end{aligned}$$

can be finally approximated as

$$err_{|b=0}(t) \approx \left(\frac{\sum_{i=1}^{N}\lambda_i}{\|\bar{w}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}.$$

This term does not constitute an upper bound of the error. However, if this limit is crossed it means that the model is completely useless. In Figures 5 and 6, the fast exponential divergence of the standard POD reduced order (in blue) is clearly visible. Conversely, our methods, based on a physically relevant stochastic representation of the small scale component, have much weaker errors, without tuning any additional parameters on the data. There is only a slight difference between a constant and a linear representation of the variance tensor. A drawback of the second method is that $a(x,t)$ is not ensured to be a positive definite matrix. When the number of

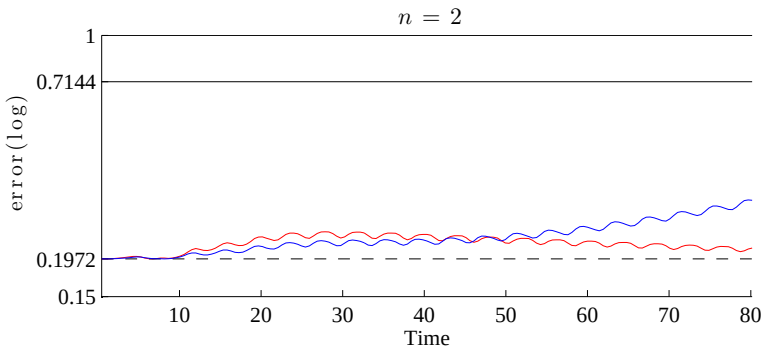

Figure 5: Normalized error for $n = 2$ modes, of a wake flow at Reynolds 300. The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue line corresponds to the standard POD Galerkin. The red one stands for our model with a constant variance tensor along time. The magenta one represents our model with linear representation of the variance tensor. The doted line indicates the error associated to the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error considering only the time mean velocity.

modes increases, the basis used for the projection of a, $(b_i)_{0 \leqslant i \leqslant n}$, is larger. Thus, the projection approximates better the identity, and the estimation of $a(x,t)$ becomes closer to a positive matrix and close to a. This may explain the difference between the two methods.
Whatever their differences, both methods provide very encouraging results. These representations clearly enable the construction of autonomous sub-grid models. This constitutes an essential point for the devising of autonomous reduced order dynamical systems.

6. Conclusion

In this paper, a fluid dynamics model built from fundamental physical principles applied to a stochastic representation of the flow has been used. In this representation, the fluid velocity is random and partially decorrelated in time. This time decorrelation can be interpreted as coming from a subsampling in time of a fast oscillating part of the velocity. In this framework, mass and momentum conservation principles can be constituted from stochastic calculus to derive a complete fluid flow dynamics model. This framework brings a strong theoretical support to classical empirical models, while generalizing them through the incorporation of an anisotropic, inhomogeneous and time-dependent diffusion. Compared to the original stochastic model, introduced in [14], some initial assumptions have been removed. The diffusion tensor σ does not need to be symmetric anymore, w can be any semimartingale for the stochastic version of the Reynolds transport theorem and any finite variation process for the stochastic Navier-Stokes model. It has also been proved that the

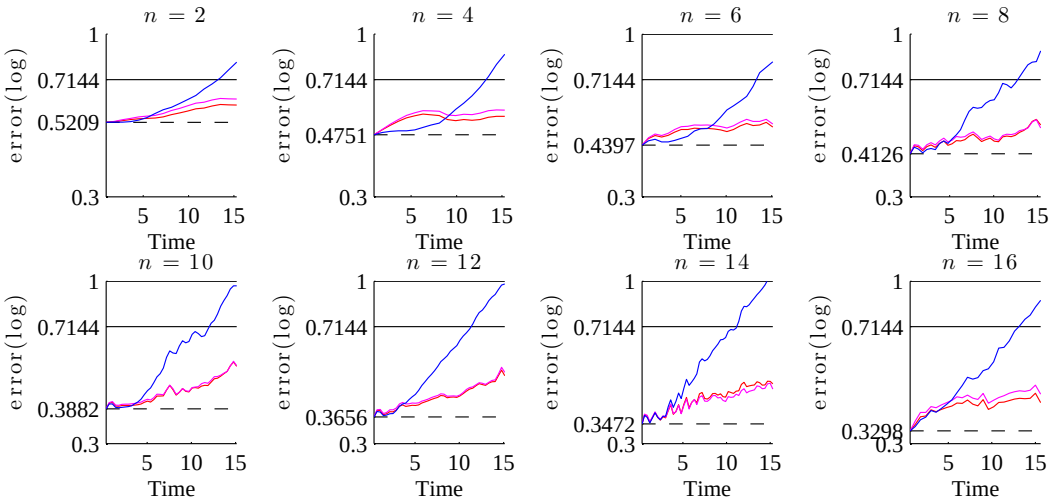

Figure 6: Normalized error for $n = 2$, 4, 6, 8, 10, 12, 14 and 16 modes, of a wake flow at Reynolds 3900. The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue line corresponds to the standard POD Galerkin. The red one stands for our model with a constant variance tensor along time. The magenta one represents our model with linear representation of the variance tensor. The doted line indicates the error associated to the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error considering only the time mean velocity.

sub-grid tensor is diffusive when the density is constant. Thanks to our stochastic representation of fluid dynamics, a reduced model, describing the resolved modes evolution, has been derived. This model takes explicitly into account the unresolved modes influence. Since our stochastic model enables to deal with aliasing effects, we have chosen a time step as large as possible to simplify the reduced model simulation. A criterion based on Shannon-Nyquist theorem has been proposed to set the time step. Two different methods have been proposed to estimate the variance tensor. The first one relies on the assumption of a constant variance tensor along time, whereas the second one decomposes this tensor as a linear combination of the *Chronos* basis. From both methods, closed autonomous reduced systems have been derived. Finally, in section 5, both methods have been tested on numerical data from DNS simulation at Reynolds 300 and LES simulation at Reynolds 3900 of wake flow. The two kinds of reduced models have been compared to POD Galerkin reduced system. The standard reduced system exhibits very fast diverging trajectories. On the contrary, our models have shown to provide much better results without any parameter tuning.

Those results are very encouraging. Indeed, we have written basic physical conservation laws in a stochastic framework where Itō formalism is interpreted as scale separation. This new methodology has yielded to a powerful stochastic fluid dynamics model. It is true that the variance tensor a remains to be estimated or modeled. Nevertheless, we have proposed two estimation methods of this tensor, based on stochastic calculus, in the context of reduced order model. These simple estimation methods were sufficient to illustrate the potential of our new fluid dynamics model. Needless to say, a lot of improvements are possible and may be considered. Here, the variance has been assumed to be constant in time or in the span of *Chronos*. However, it can also be assumed that it is a quadratic or a cubic function of the temporal modes. To obtain a more sophisticated time dependence for the variance tensor, a dynamical model can also be used. Many dynamical models of the sub-grid velocity variance exist in the literature, based mainly on heuristic observations or statistical estimations. For instance, the RANS equations are closed by empirical turbulent kinetic energy evolution equations ([35]). More recent works reveal new sub-grid dynamical models. Stochastic superparametrization ([36]) is one example of such models. Using the so-called point approximation, the large-scale influence on the evolution equation of the small-scale velocity becomes constant and uniform. Then, a Gaussian closure decouples the small-scale Fourier modes and enables solving the small-scale variance dynamic. To go further, the sub-grid velocity can be non-Gaussian. Modified Quasilinear Gaussian (MQG) closure ([37, 38]) can be used instead of a Gaussian one. The Quasilinear Gaussian (QG) method neglects third order moments. The MQG algorithm enables simulating accurately the two first moments by modeling the third order moments ([37]). This model is based on energy transfer principles and estimations on long time. The MQG-DO algorithm manages to also simulate accurately the other moments in a reduced subspace ([38]). QG, MQG and MQG-DO closures will also lead to a dynamical model of the small-scale

variance tensor a. Another improvement of our algorithm could consist in using several time steps. One time step by resolved mode will involve one value of the variance tensor by resolved mode. This should make the most of the time-decorraleted unresolved velocity explicit influence. Therefore, the variance tensor would be different for each resolved modes. Finally, it could be suitable to remove the finite variations assumption for the large-scale drift, w. This yields to a new model, that will be exposed in future works. In such a model, an evolution equation determines the partially time-correlated sub-grid velocity component: the martingale part of w. This component is random, centered and not differentiable w.r.t. time. Therefore, it lives at a smaller scale than the finite-variation velocity component, studied previously in this paper. However, since the martingale component is continuous w.r.t. time, its evolution is smoother than the time-uncorrelated component. This new evolution equation, on the martingale part of the drift, may bring several advantages. First, this is a linear stochastic partial differential equation with additive and multiplicative noises, which is both easy to handle and physically pertinent ([13]). Then, the noises covariances are naturally linked to the covariance of the time-uncorrelated velocity component $a(x, y)$ and to the sub-grid tensor. It implies a lot of interesting properties such as energy conservation, up to molecular viscous effect. Moreover, this stochastic equation is only inferred from fundamental physical laws. Therefore, neither tuning nor ad hoc model assumption is needed. For all these reasons, this new stochastic fluid dynamics model should be very helpful to built relevant uncertainty quantification (UQ) and sub-grid stochastic dynamic models, with reduced complexity. Such UQ methods can be used for stochastic reduced order models, filtering or probabilistic closures. And the associated sub-grid stochastic dynamics can be used, for instance, for stochastic superparametrization. Some work on this new model is currently ongoing.

A. Basic notions of stochastic calculus

We recall here some basic definitions and properties of stochastic calculus. Here, for simplicity, we only deal with functions of a compact set of time: $t \in [0, T]$ with $T \in \mathbb{R}_+^*$. However, everything can be generalized easily to functions of $\mathbb{R}_+ \times \Omega$ with $\Omega \subset \mathbb{R}^d$ (see [24], [22] and [23]).
We use a sample space $\breve{\Omega}$, a probability measure $\mathbb{P}$, a Wiener process, $(B_t)_{t\geqslant 0}$, its filtration $(\mathcal{F})_{t\geqslant 0}$ (the set of σ-algebra generated by each B_t), the whole σ-algebra, $\mathcal{F} \overset{\triangle}{=} \mathcal{F}_\infty \overset{\triangle}{=} \bigcup_{t\geqslant 0} \mathcal{F}_t$, and the resulting filtered probability space $(\breve{\Omega}, \mathcal{F}, (\mathcal{F})_{t\geqslant 0}, \mathbb{P})$

Defintion A.1. *Finite variation function*
$t \to f(t)$ is a finite variation function if and only if for all $a < b$ and all partition $a = t_0 < ... < t_n = b$ of $[a, b]$, $\sum_{i=1}^n |f(t_i) - f(t_{i-1})| < \infty$.

Defintion A.2. *Finite variation process*
$(t, \omega) \to f(t, \omega)$ has finite variations if and only if:

- *f is adapted (i.e. $f(t,.)$ is $\mathcal{F}_t$ measurable),*
- *For each trajectory ω, $f(.,\omega)$ is a finite variation function .*

Characterization:
f is a finite variation process if and only if $\exists g, f(t,.) = f(0,.) + \int_0^t g(t',.)dt'$.

Defintion A.3. *Martingale*
$(t,\omega) \to f(t,\omega)$ is a martingale if and only if:

- *f is adapted ,*
- $f(t,.) \in L^1_{\Omega} \stackrel{\triangle}{=} \{Y : \mathbb{E}|Y| < \infty\}$,
- $\forall s < t,\ \mathbb{E}(f(t,.)|\mathcal{F}_s) = f(s,.)$.

In particular, if $f = 0$ at $t = 0$, then f is a centered process.
Characterization:
f is a martingale if and only if $\exists g, f(t,.) = f(0,.) + \int_0^t g(t',.)dB_{t'}$.

Defintion A.4. *Continuous semimartingale*
f is a continuous semimartingale if and only if it the sum of a finite variation process and a martingale

Stochastic calculus deals only with semimartingales. In our fluid dynamics representation, we also deal with time-decorrelated processes, formally, the differentiation along time of a martingale.

Defintion A.5. *Quadratic variation and quadratic cross-variation*
If f and g are semimartingale and $f(t=0) = g(t=0) = 0$, then, their quadratic cross-variation, noted $< f,g >$, is the unique finite variation process such $fg- < f,g >$ is a martingale and $< f,g >_{t=0}= 0$.

Characterization:

- If $f(t,.) = \int_0^t f_1(t')dt' + \int_0^t f_2(t')dB_{t'}$ and $g(t,.) = \int_0^t g_1(t')dt' + \int_0^t g_2(t')dB_{t'}$, then $< f,g >_t= \int_0^t f_2(t')g_2(t')dt'$.
  It should be noticed that, if f_2 and g_2 are random, $< f,g >$ is also random.
- $< f,g >_t= \mathbb{P}-\lim_{\Delta t\to 0} \sum_{t_i=0}^{t_n=t}(f(t_i) - f(t_{i-1}))(g(t_i) - g(t_{i-1}))$.

Thus, $< f,g >$ may be interpreted as a kind of "covariance along time".

Theorem A.1. *Itô-Wentzell Formula*
If $(t,x) \to f(t,x)$ and $(t,y) \to g(t,y)$ are semimartingale (as function of time), and $x \to f(.,x)$ is twice differentiable, then

$$d_t\left(f(t,g(t,y))\right) = d_t f + \partial_x f d_t g + \frac{1}{2}\partial^2_{xx} f d_t < g,g > + d_t < \partial_x f, g > .$$

B. Equivalence of formulation for the stochastic transport and Navier-Stokes theorem

B.1. Stochastic Reynolds transport model

[14] describes the stochastic transport theorem as follows:

$$d\int_{V(t)} q(x,t)dx = \int_{V(t)} \left(d_t q + \left[\nabla\cdot(qw) + \frac{1}{2}\|\nabla\cdot\sigma\|^2 q - \frac{1}{2}\sum_{i,j}\partial_{x_i}\partial_{x_j}(a_{ij}q)_{|\nabla\cdot\sigma=0} \right] dt + \nabla\cdot(q\sigma dB_t) \right) dx,$$

where:

$$\begin{aligned}
\sum_{i,j}\partial_{x_i}\partial_{x_j}(a_{ij}q)_{|\nabla\cdot\sigma=0} &= \sum_{i,j}(\partial_{x_i}\partial_{x_j}(\sigma_{i.}\sigma^t_{j.}))_{|\nabla\cdot\sigma=0}q + 2(\partial_{x_i}(\sigma_{i.}\sigma^t_{j.}))_{|\nabla\cdot\sigma=0})\partial_{x_j}q + a_{ij}\partial_{x_i}\partial_{x_j}q, \\
&= \sum_{i,j}\partial_{x_j}\sigma_{i.}\partial_{x_i}\sigma^t_{j.}q + 2\sigma_{i.}\partial_{x_i}\sigma^t_{j.}\partial_{x_j}q + a_{ij}\partial_{x_i}\partial_{x_j}q, \\
&= \sum_{i,j}\partial_{x_i}\partial_{x_j}(a_{ij}q) - 2\partial_{x_i}\partial_{x_j}\sigma_{i.}\sigma^t_{j.}q - \partial_{x_i}\sigma_{i.}\partial_{x_j}\sigma^t_{j.}q - 2\partial_{x_i}\sigma_{i.}\sigma^t_{j.}\partial_{x_j}q.
\end{aligned}$$

So,

$$\begin{aligned}
\frac{1}{2}\|\nabla\cdot\sigma\|^2 q - \frac{1}{2}\sum_{i,j}\partial_{x_i}\partial_{x_j}(a_{ij}q)_{|\nabla\cdot\sigma=0} &= \sum_{i,j} -\frac{1}{2}\partial_{x_i}\partial_{x_j}(a_{ij}q) + \partial_{x_i}\partial_{x_j}\sigma_{i.}\sigma^t_{j.}q \\
&\quad + \partial_{x_i}\sigma_{i.}\partial_{x_j}\sigma^t_{j.}q + \partial_{x_i}\sigma_{i.}\sigma^t_{j.}\partial_{x_j}q, \\
&= -\frac{1}{2}\nabla\cdot(\nabla\cdot(aq)^T) + \nabla\cdot(\sigma(\nabla\cdot\sigma)^T q), \\
&= \nabla\cdot\left(-\frac{1}{2}(\nabla\cdot(aq)^T) + \sigma(\nabla\cdot\sigma)^T q\right).
\end{aligned}$$

B.2. Stochastic Navier-Stokes model

[14] describes the r-th coordinate of the diffusion tensor of the stochastic Navier-Stokes model as:

$$\begin{aligned}
\tau_r(w) &= \sum_{i,j}\frac{1}{2}\rho a_{ij}\partial_{x_i}\partial_{x_j}(w_r) + \partial_{x_i}(\rho a_{ij})_{|\nabla\cdot\sigma=0}\partial_{x_j}w_r, \\
&= \sum_{i,j}\frac{1}{2}\rho a_{ij}\partial_{x_i}\partial_{x_j}(w_r) + \partial_{x_i}(\rho a_{ij})\partial_{x_j}w_r - \rho\partial_{x_i}(\sigma_{i.})\sigma_{j.}\partial_{x_j}w_r, \\
&= \frac{1}{2}(\nabla\cdot(\nabla\cdot(\rho a w_r)^T)) - \nabla\cdot(\nabla\cdot(\rho a)^T)))w_r - \rho\nabla\cdot\sigma\sigma^T\nabla w_r.
\end{aligned}$$

C. Dissipative effect of sub-grid tensor τ

If the density, ρ, is assumed to be constant, then $\nabla \cdot \sigma = 0$ by the martingale part of the mass conservation. Moreover, a or w and its derivatives are assumed to be null in the border of Ω, then with two integrations by parts,

$$\begin{aligned}
\sum_k \int_\Omega w_k \nabla \cdot (\nabla \cdot (a w_k))^T \, dx &= \int_\Omega w \cdot \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} \left(a_{i,j} w\right) dx \\
&= -\int_\Omega \sum_{i,j} \frac{\partial}{\partial x_i} w^t \frac{\partial \left(a_{i,j} w\right)}{\partial x_j} dx, \\
&= -\int_\Omega \sum_{i,j} \left(\frac{\partial w^T}{\partial x_i} \frac{\partial a_{i,j}}{\partial x_j} w + \frac{\partial w^T}{\partial x_i} a_{i,j} \frac{\partial w}{\partial x_j} \right) dx, \\
&= -\int_\Omega \sum_{i,j} \left(\frac{1}{2} \frac{\partial \|w\|_2^2}{\partial x_i} \frac{\partial a_{i,j}}{\partial x_j} + \frac{\partial w^T}{\partial x_i} a_{i,j} \frac{\partial w}{\partial x_j} \right) dx, \\
&= \int_\Omega \left(\frac{1}{2} \nabla \cdot (\nabla \cdot a)^T \|w\|_2^2 - \|\nabla w^T\|_a^2 \right) dx,
\end{aligned}$$

where $\|\nabla w^t\|_a^2 \stackrel{\triangle}{=} \sum_k \|\nabla w_k\|_a^2 \stackrel{\triangle}{=} \sum_k \nabla w_k^T a \nabla w_k = tr\left((\nabla w^T)^T a \nabla w^T\right)$.

$$\begin{aligned}
2 \int_\Omega w \cdot \tau \, dx &= \rho \int_\Omega \left(\frac{1}{2} \nabla \cdot (\nabla \cdot a)^T \|w\|_2^2 - \|\nabla w^T\|_a^2 - \nabla \cdot (\nabla \cdot a)^T \|w\|_2^2 \right) dx, \\
&= -\rho \int_\Omega \left(\frac{1}{2} \nabla \cdot (\nabla \cdot a)^T \|w\|_2^2 + \|\nabla w^T\|_a^2 \right) dx.
\end{aligned}$$

Using now the finite variation part of the mass conservation, which is $\nabla \cdot w = \frac{1}{2} \nabla \cdot (2 \nabla \cdot a)^T$, we get:

$$\begin{aligned}
2 \int_\Omega w \cdot \tau \, dx &= -\rho \int_\Omega \left(\nabla \cdot w \|w\|_2^2 + \|\nabla w^T\|_a^2 \right) dx, \\
&= \rho \int_\Omega \left((w \cdot \nabla) \|w\|_2^2 - \|\nabla w^T\|_a^2 \right) dx, \\
&= \rho \int_\Omega \left(2 \, w \cdot ((w \cdot \nabla) w) - \|\nabla w^T\|_a^2 \right) dx.
\end{aligned}$$

Considering together the advection term and the sub-grid term of the Navier Stokes equation we have for the energy:

$$\int_\Omega w \cdot \left(-\rho (w \cdot \nabla) w + \tau\right) dx = -\frac{\rho}{2} \int_\Omega \|\nabla w^T\|_a^2 dx < 0, \tag{C.1}$$

which is the sought result. It should be noted that in the incompressible deterministic equation, $\nabla \cdot w = 0$ and thus the advection term $(w \cdot \nabla) w$ does not influence the global energy. Here however, it is not the case anymore and this term has to be taken into account, as above.

Acknowledgements

The authors would like to kindly thank Johan Carlier and Dominique Heitz for providing data and their expertise in physical interpretations.

References

[1] J. Slingo and T. Palmer. Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1956):4751–4767, 2011.

[2] T. Palmer and P. Williams. Introduction. stochastic physics and climate modelling. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366(1875):2419–2425, 2008.

[3] G. Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2006.

[4] R. Kraichnan. Eddy viscosity and diffusivity: exact formulas and approximations. *Complex Systems*, 1(4-6):805–820, 1987.

[5] P. Parnaudeau, J. Carlier, D. Heitz, and E. Lamballais. Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900. *Physics of Fluids*, 20(8):085101, 2008.

[6] M. Lesieur and O. Metais. New trends in large-eddy simulations of turbulence. *Annual Review of Fluid Mechanics*, 28(1):45–82, 1996.

[7] B. Protas, B. Noack, and J. Östh. Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows. *arXiv preprint arXiv:1406.1912*, 2014.

[8] A. Kupiainen. Statistical theories of turbulence. *workshop Random media 2000*, June 2000.

[9] C. Franzke, T. O'Kane, J. Berner, P. Williams, and V. Lucarini. Stochastic climate theory and modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 6(1):63–78, 2015.

[10] A. Doucet, N. De Freitas, and N. Gordon. *Sequential Monte Carlo methods in practice*. Springer, 2001.

[11] A. Doucet and A. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12:656–704, 2009.

[12] J. Candy. *Bayesian signal processing: Classical, modern and particle filtering methods*, volume 54. John Wiley & Sons, 2011.

[13] A. Majda, I. Timofeyev, and E. Eijnden. Models for stochastic climate prediction. *Proceedings of the National Academy of Sciences*, 96(26):14687–14691, 1999.

[14] E. Mémin. Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, 108(2):119–146, 2014.

[15] Z. Brzeźniak, M. Capiński, and F. Flandoli. Stochastic partial differential equations and turbulence. *Mathematical Models and Methods in Applied Sciences*, 1(01):41–59, 1991.

[16] R. Mikulevicius and B. Rozovskii. Stochastic Navier–Stokes equations for turbulent flows. *SIAM Journal on Mathematical Analysis*, 35(5):1250–1310, 2004.

[17] P. Holmes, J. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press, 1998.

[18] N. Aubry, P. Holmes, J. L Lumley, and E. Stone. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *Journal of Fluid Mechanics*, 192:115–173, 1988.

[19] D. Rempfer and H. Fasel. Evolution of three-dimensional coherent structures in a flat-plate boundary layer. *Journal of Fluid Mechanics*, 260:351–375, 1994.

[20] J. Östh, B. Noack, S. Krajnović, D. Barros, and J. Borée. On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body. *Journal of Fluid Mechanics*, 747:518–544, 2014.

[21] K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations. *International Journal for Numerical Methods in Engineering*, 86(2):155–181, 2011.

[22] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992.

[23] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905. Springer, 2007.

[24] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

[25] S. Harouna and E. Mémin. A wavelet based numerical simulation of Navier-Stokes equations under uncertainty. Preprint 2014.

[26] N. Nakamura. A new look at eddy diffusivity as a mixing diagnostic. *Journal of the atmospheric sciences*, 58(24):3685–3701, 2001.

[27] I. Mezić, S. Loire, V. Fonoberov, and P. Hogan. A new mixing diagnostic and gulf oil spill movement. *Science*, 330(6003):486–489, 2010.

[28] P. Rao. *Statistical inference for diffusion type processes*. Arnold, 1999.

[29] V. Genon-Catalot, C. Laredo, and D. Picard. Non-parametric estimation of the diffusion coefficient by wavelets methods. *Scandinavian Journal of Statistics*, pages 317–335, 1992.

[30] D. Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *Journal of applied probability*, pages 790–804, 1993.

[31] V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'institut Henri Poincar (B) Probabilits et Statistiques*, 29(1):119–151, 1993.

[32] M. Hofmann et al. Lp estimation of the diffusion coefficient. *Bernoulli*, 5(3):447–481, 1999.

[33] F. Comte, V. Genon-Catalot, and Y. Rozenholc. Penalized nonparametric mean square estimation of the coefficients of diffusion processes. *Bernoulli*, pages 514–

543, 2007.

[34] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[35] F. Menter. Improved two-equation k-omega turbulence models for aerodynamic flows. *NASA STI/Recon Technical Report N*, 93:22809, 1992.

[36] I. Grooms and A. Majda. Stochastic superparameterization in quasigeostrophic turbulence. *Journal of Computational Physics*, 271:78–98, 2014.

[37] T. Sapsis and A. Majda. Statistically accurate low-order models for uncertainty quantification in turbulent dynamical systems. *Proceedings of the National Academy of Sciences*, 110(34):13705–13710, 2013.

[38] T. Sapsis and A. Majda. Blending modified Gaussian closure and non-Gaussian reduced subspace methods for turbulent dynamical systems. *Journal of Nonlinear Science*, 23(6):1039–1071, 2013.

Inria Rennes - Bretagne Atlantique
Campus de Beaulieu, 263 Avenue Général Leclerc,
35042 Rennes, FRANCE
e-mail: valentin.resseguier@inria.fr
etienne.memin@inria.fr
url: http://www.irisa.fr/fluminance

Ifremer,
Pointe du Diable,
29280 Plouzan, FRANCE
e-mail: bertrand.chapron@ifremer.fr