Quantifying truncation-related uncertainties in unsteady fluid dynamics reduced order models

Valentin Resseguier, Agustin M Picard, Etienne Mémin, Bertrand Chapron

Quantifying truncation-related uncertainties in unsteady fluid dynamics reduced order models *

Valentin Resseguier†, Agustin M. Picard†, Etienne Mémin‡, and Bertrand Chapron§

Abstract. In this paper, we present a new method to quantify the uncertainty introduced by the drastic dimensionality reduction commonly practiced in the field of computational fluid dynamics, the ultimate goal being to simulate accurate priors for real-time data assimilation. Our key ingredient is a stochastic Navier-Stokes closure mechanism that arises by assuming random unresolved flow components. This decomposition is carried out through Galerkin projection with a Proper Orthogonal Decomposition (POD-Galerkin) basis. The residual velocity fields, model structure and evolution of coefficients of the reduced order's solutions are used to compute the resulting multiplicative and additive noise's correlations. The low computational cost of these consistent correlation estimators makes them applicable to the study of complex fluid flows. This stochastic POD-ROM is applied to 2D and 3D DNS wake flows at Reynolds 100 and 300, respectively, with Uncertainty Quantification (UQ) and forecasting outside the learning interval being the main focus. The proposed stochastic POD-ROM approach is shown to stabilize the unstable temporal coefficients and to maintain their variability under control, while exhibiting an impressively accurate predictive capability.

Key words. Fluid dynamics, reduced order model, uncertainty quantification, stochastic closure, proper orthogonal decomposition

AMS subject classifications. 60H35, 65M60, 65M75, 76M35, 93B11

1. Introduction.

The industrial application of partial differential equations (PDE)-driven processes – fluid dynamics, for instance – can be a daunting task, mainly due to the computational complexity involved with its resolution. This computational burden becomes even excruciatingly difficult to handle when it comes to achieving real-time simulations. To tackle this difficulty, reduced order models (ROM) are commonly employed to speed up deterministic and stochastic design simulations [9, 22, 56], or optimal control problems [19, 36, 45]. The interested community has proposed a plethora of different algorithms to reduce the computational cost of stochastic PDEs for uncertainty quantification (UQ) applications [20, 21, 44, 54, 55, 69, 77]. In particular, when some parameters of deterministic or stochastic PDEs are random, the weighted reduced basis (wRB) method [20, 21] helps choosing the best parameters' value for each new full-order simulation. With the addition of stabilizing terms, [77] applied this method to linear advection-dominated problems, including advection by a random velocity field.

In turbulent fluid dynamics, the system's energy usually spreads out over many degrees of freedom. This prevents low-dimensional approximations of the state vector from being sufficiently accurate, but such a rough approximations can theoretically be sufficient for specific industrial applications, especially when the quantity of interest (QoI) is a spatial average (e.g. lift and drag). However, that multiscale property prevents, in general, the constitution of accurate reduced systems for the dynamics of those low-dimensional state approximations. Indeed, severe modal truncations often usually end up destabilizing the system and overdamping some of the stable coefficients of the reduced order solution [69].

*Submitted to the editors on July 22, 2020.

Funding: This work was supported by the ERC EU project 856408-STUOD, the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project.

†Lab, SCALIAN DS, Rennes, France (valentin.resseguier@scalian.com).

‡Fluminance team, Inria, Rennes, France .

§LOPS, Ifremer, Plouzané, France.

Consequently, to stabilize the ROM, authors commonly introduce an additional deterministic term (typically an eddy viscosity term) [2, 15, 79], along with a possible calibration on available data [12, 23, 80]. This calibration procedure can be extended to the complete set of the ROM's parameters [1, 10, 61, 72]. ROM performances have mainly been evaluated on low Reynolds-number flows (say $Re \leqslant 100$) [e.g. 10, 75, 80], while for flows at much greater Reynolds numbers, evaluations close to the learning time interval and reduced dynamics for two-dimensional flow observables (e.g. particle image velocimetry) have usually been the norm [e.g. 1, 12, 23, 75]. Long time ROM predictions have been otherwise performed with a substantial number of modes [e.g. 14, 74]. Yet, turbulent flow ROMs remain inexact and uncontrolled in the long run, owing to their intrinsic chaotic nature and the growth of accumulated error along time. Outside the learning time interval, predictions becomes less and less accurate. Moreover, these predictions necessitate accurate initial conditions, forcing and geometric information, which are often poorly known.

Ensemble-based data assimilation – such as particle filters or ensemble Kalman filters – can alleviate these issues by forecasting an ensemble of likely future states of the system, while sequentially constraining them with on-coming measurements [27, 28]. Still, this necessitates accurate quantification of their associated simulation's uncertainties. In this context, the aim is not to reduce the dimensionality of UQ, but rather to quantify the uncertainty introduced by the dimensionality reduction.

Note, UQ is a recurrent issue in applied fluid dynamics, and many strategies have been proposed for incorporating randomness in the physical models through some of their parameters [e.g. 44, 47]. However, the error introduced by these noisy parameters is not *a priori* tied to dimensionality reduction and to the contribution of unresolved components. In particular, in fluid dynamics, random initial conditions have first been widely used for both UQ and predictability studies [e.g. 51]. It was later demonstrated this yields underdispersive quantification, i.e. it has a tendency to underestimate the error associated to the dimensionality reduction [8, 31, 35, 53].

Alternatively, authors have considered the introduction of additive noises, most likely beginning with the introduction of EDQNM [46, 57]. Without special care, such a strategy often leads to energy conservation loss, stability issues and radical changes of the underlying attractor [17].

The Modified Quasilinear Gaussian (MQG) method [69, 70, 71] approximates the third-order moment to help redistribute the right amount of energy between the coefficients of the reduced solution. Finally, several techniques related to averaging and homogenization theory exploit a time-scale separation hypothesis [34, 43, 53, 58], one of the most notable being the MTV model [49]. This latter approach can reproduce intermittency and extreme events alike, thanks to its correlated additive and multiplicative noises. However, besides potential energy conservation issues, the noise covariance is often not explicit enough, and has to be simplified and estimated using the available data. Interested readers can refer to [64], and references herein, for more detailed reviews on model error specification in coarse-scale computational fluid dynamics (CFD). For ROM UQ, [73] propose distributions and efficient sampling methods for the projection matrices in Galerkin-projection-based dimensionality reduction methods. Although apparently more pertinent than methods based on randomized parameters, the relation of this technique to errors associated to mode truncation and turbulent chaotic behaviour remains unclear.

In this paper, we propose to quantify the uncertainty introduced by modal dimensionality reduction through the so-called dynamics under location uncertainty (LU) [50, 65]. Specifically, we adapt the aforementioned stochastic closure to the Galerkin-projection-based ROM. Inspired from the theoretical works of [11, 52], the LU closure relies on the

stochastic transport of the flow variables, together with a decorrelation assumption of the unresolved fluctuations with respect to the resolved slow/large scales. More precisely, the residual velocity – i.e. the difference between the usual Navier-Stokes solution $\boldsymbol{v}$ and some large-scale velocity component $\boldsymbol{w}$ – is assumed to be time-uncorrelated at the characteristic time of the large-scale processes. This residual velocity is informally denoted $\boldsymbol{v}' = \boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/\mathrm{d}t$ where $t \mapsto \boldsymbol{\sigma}\boldsymbol{B}_t$ is a $\boldsymbol{Q}$-Wiener process [26, 62], and hence, Gaussian in nature. Note that this apparent simplified Gaussian assumption leads to, as we will see, a non-Gaussian multiplicative noise in the dynamics. Spatial correlations of the residual velocity are then specified through the Hilbert-Schmidt integral operator $\boldsymbol{\sigma}$ with a C^2 kernel $\breve{\boldsymbol{\sigma}}\colon \overline{\Omega}^2 \to \mathbb{R}^{d\times d}$:

$$\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z})\mathrm{d}\boldsymbol{B}_t(\boldsymbol{z})\mathrm{d}\boldsymbol{z} \quad \forall(\boldsymbol{x},t)\in\Omega\times[0,T]. \tag{1.1}$$

This operator – or equivalently the spatial covariance of the residual velocity – can be modeled or learned from data. The review [64] describes some of the many choices that have been explored in this vein. This includes for instance parametric models based on fluid velocity self-similarity or brute-force non-parametric covariance estimation from high-resolution datasets. Here, $\boldsymbol{B}_t$ is an $\mathbb{I}_d$-cylindrical Wiener process and $\mathrm{d}\boldsymbol{B}_t/\mathrm{d}t$ plays the role of spatio-temporal white noise. The above definition enables us to characterize the way physical quantities are transported by the stochastic flow:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t)\mathrm{d}\boldsymbol{B}_t \quad \forall t\in[0,T], \tag{1.2}$$

This resembles the expression for transport in classical fluid dynamics. Material derivatives and other differential operations of fluid dynamics are then derived through the use of stochastic PDEs (SPDE), in particular, by applying the Itō-Wentzell formula [42] and a stochastic version of the Reynolds transport theorem [40, 50, 65]. As such, LU models can be applied to model error quantification [17, 64, 66, 68], to improve large-scale simulations [6, 7, 40, 66, 68], to reduced order modeling and data analysis [67] or to data assimilation purposes [16, 81] in geophysical fluid dynamics and CFD. To note, in the geometric mechanics community [38, 24], the Stochastic Advection by Lie Transport (SALT) method has been derived for large-scale modeling and data assimilation [25]. Both frameworks have been compared, numerically [7, 64] and conceptually, [6, 68], with LU and SALT exhibiting different conservation properties, namely energy preservation and circulation conservation, respectively. Applied to a barotropic Quasi-Geostrophic model, LU leads to improved accuracy when compared to a classical large-scale deterministic framework or to the circulation conservation stochastic setup [6, 7]. The LU setting also fully captures the structural deformation of the large-scale flow component by the spatial inhomogeneity of the small-scales [6]. It is important to note that these properties are independent of the choice of stochastic integral.

In this paper, our focus is to analyse the LU setting to help define efficient, highly reduced order models for real time data assimilation or control applications. For the sake of simplicity, we will deal with Proper Orthogonal Decomposition (POD) [48] as dimensionality reduction technique – where time and space dependency are separated – but the proposed methodology applies to all types of modal decompositions.

In section 2, we introduce the stochastic fluid dynamics closure we will employ throughout the paper, as well as its algebraic structure, followed by the Galerkin projection of this SPDE to derive our stochastic ROM, with a brief recall of the principle behind POD-ROMs in section 3. Several estimations needed to complete our stochastic POD-ROM are detailed in section 4, as well as the efficient and consistent estimators we exploit to rely both on

data and on the closure's physical grounding. In section 5, we discuss the conservative properties of LU dynamics and of its reduced versions, and finally, section 6 is dedicated to the numerical evaluation of the UQ capabilities of our ROM.

2. Navier-Stokes model under location uncertainty.

Galerkin projections of Navier-Stokes equations do not specifically take into account residual velocity contributions, and thus, nor do they offer a precise quantification of their induced errors. To help remedy this issue, we propose to directly project SPDEs – i.e. the LU Navier-Stokes representation – instead of the classical Navier-Stokes equations. The resulting ROM is expected to describe the same physical system, as we postulate that the solution of the SPDE to be statistically similar to a large-scale representation of the original (deterministic) equation's solution.

2.1. The random physical model.

Let us denote by Ω an open bounded subset of $\mathbb{R}^d$, $T \in \mathbb{R}_+^*$ and $L^2(\Omega)$ (resp. $\boldsymbol{L}^2 \triangleq \left(L^2(\Omega)\right)^d$) the space of square-integrable scalar (resp. vector) fields on Ω. In the case of incompressible fluids, the LU Navier-Stokes equations on $\Omega \times [0,T]$ read:

$$\underbrace{\mathbb{D}_t \boldsymbol{w}}_{\substack{\text{Stochastic}\\ \text{transport}}} = \underbrace{-\boldsymbol{\nabla}\left(p\,\mathrm{d}t + \boldsymbol{p}_\sigma \mathrm{d}\boldsymbol{B}_t\right)}_{\substack{\text{Pressure}\\ \text{forcing}}} + \underbrace{\tfrac{1}{Re}\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\substack{\text{Molecular viscous}\\ \text{dissipation}}}, \tag{2.1}$$

$$0 = \underbrace{\boldsymbol{\nabla}\cdot(\boldsymbol{w}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{\text{Mass conservation}}, \tag{2.2}$$

where $p + \boldsymbol{p}_\sigma \frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$ informally represents the pressure also decomposed into a large-scale and a small-scale time-uncorrelated component, Re stands for the Reynolds number, and for every smooth-enough function $\boldsymbol{q} : \Omega \times [0,T] \to \mathbb{R}^d$, we denote:

$$\left(\mathbb{D}_t \boldsymbol{q}\right)_k \triangleq \underbrace{\underbrace{\mathrm{d}_t q_k}_{\triangleq\, q_k(\boldsymbol{x},t+\mathrm{d}t)-q_k(\boldsymbol{x},t)}}_{\text{Time increment}} + \underbrace{\left(\boldsymbol{w}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}q_k}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\tfrac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q_k\right)}_{\substack{\text{Turbulent}\\ \text{diffusion}}}\mathrm{d}t, \tag{2.3}$$

$$\boldsymbol{w}^* \triangleq \boldsymbol{w} - \tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)^T, \tag{2.4}$$

$$\boldsymbol{a} \triangleq \mathbb{E}\left\{\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}/\mathrm{d}t = \boldsymbol{\sigma}\circledast\boldsymbol{\sigma}, \tag{2.5}$$

where operator $\circledast$ is defined later on in this section.

Throughout this paper, we consider various linear integral operators defined on some subspace $\boldsymbol{D} \subset \boldsymbol{L}^2$. Denoted $\boldsymbol{\eta}$, each of those operators can be defined through a matrix kernel $\breve{\boldsymbol{\eta}}$ of $\left(L^2\left(\Omega^2\right)\right)^{d\times d}$ as follows:

$$\left(\boldsymbol{\eta}\boldsymbol{\xi}\right)(\boldsymbol{x}) \triangleq \int_\Omega \underbrace{\breve{\boldsymbol{\eta}}(\boldsymbol{x},\boldsymbol{z})}_{\substack{\text{kernel}\\ \text{of } \boldsymbol{\eta}}}\boldsymbol{\xi}(\boldsymbol{z})d\boldsymbol{z} \quad \forall \boldsymbol{\xi} \in \boldsymbol{D} \subset \boldsymbol{L}^2,\ \forall \boldsymbol{x} \in \Omega. \tag{2.6}$$

Two first examples of such operator are $\boldsymbol{\sigma}$ encoding the small-scale velocity component spatial correlations (defined in (1.1)) and $\boldsymbol{p}_\sigma$ encoding the small-scale pressure component spatial correlations. For two such linear integral operators $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$, we note:

$$(\boldsymbol{\eta}\circledast\boldsymbol{\theta})(\boldsymbol{x}) \triangleq \int_\Omega \breve{\boldsymbol{\eta}}(\boldsymbol{x},\boldsymbol{z})\,\breve{\boldsymbol{\theta}}^{\,T}(\boldsymbol{x},\boldsymbol{z})d\boldsymbol{z} \quad \forall \boldsymbol{x} \in \Omega. \tag{2.7}$$

Theoretical foundations to analyze SPDEs are outside the scope of this paper. Interested readers may refer to [30, 52] for deeper insights. Here, we focus on the stochastic

Physical meaning	Notation	Full-order term	Property	ROM term
Molecular viscous dissipation	$\boldsymbol{L}$	$\frac{1}{Re}\Delta$	symmetric, < 0	$\boldsymbol{l}$
Turbulent diffusion	$\boldsymbol{F}_{\text{dif}}$	$\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\bullet\right)$	symmetric, < 0	$\breve{\boldsymbol{f}}_{\text{dif}}$
Advecting velocity correction	$\boldsymbol{F}_{\text{adv}}$	$\frac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)\boldsymbol{\nabla}$	skew-symmetric (if $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T=0$)	$\breve{\boldsymbol{f}}_{\text{adv}}$
Usual advection	$\boldsymbol{C}(\boldsymbol{w},\bullet)$	$-\left(\boldsymbol{w}\cdot\boldsymbol{\nabla}\right)$	skew-symmetric (if $\boldsymbol{\nabla}\cdot\boldsymbol{w}=0$)	$\boldsymbol{c}$
Advection by the residual velocity	$\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t$	$-\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)$	skew-symmetric noise (if $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}=0$)	$\boldsymbol{\alpha}\mathrm{d}\boldsymbol{B}_t$
Molecular viscous dissipation of the residual velocity	$\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t$	$\frac{1}{Re}\Delta\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$	additive noise	$\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t$
Pressure gradient	$\mathrm{d}\boldsymbol{P}$	$-\boldsymbol{\nabla}\left(p\,\mathrm{d}t+\boldsymbol{p}_\sigma\mathrm{d}\boldsymbol{B}_t\right)$	Potential field	Included in other terms

Table 1: Terms of LU Navier-Stokes equations and their algebraic properties

closure mechanism and its reduced order expression. The stochastic transport operator $\mathbb{D}_t$ involves the usual terms of the deterministic material derivative, on top of three additional new terms: an advecting velocity correction ($\boldsymbol{w}^*$ instead of $\boldsymbol{w}$), a heterogeneous and anisotropic turbulent diffusion, and a multiplicative noise. This last term corresponds to the advection by the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. Finally, we can recover the classical Navier-Stokes equations by setting the residual velocity to zero – i.e. $\boldsymbol{\sigma}=0$. Let us highlight that for conserved scalars, this operator corresponds to the material derivative – i.e. the derivative along the stochastic flow $\mathrm{d}\left(\boldsymbol{w}(\boldsymbol{X}_t,t)\right)$ [65]. This stochastic Navier-Stokes model is generic, and depending on the application, forces and boundary conditions, it may be modified accordingly to adjust to any incompressible flow configurations. For compressible stochastic flows, some new terms appear. Interested readers may refer to [65].

In addition to the classical physical assumptions pertaining to the establishment of the Navier-Stokes equations for incompressible fluids, the main assumption of the LU setting is to consider the unresolved fluctuation velocity component uncorrelated in time. To note, the incompressible character of the random fluctuations can be relaxed at the price of additional terms in the transport operator [65]

2.2. Algebraic structure of the model. The algebraic properties of the different terms can be quickly described. We can formally rewrite the velocity evolution law (2.1) for $t\in[0,T]$ as follows:

$$\mathrm{d}_t\boldsymbol{w}=(\mathrm{d}\mathbb{M})(\boldsymbol{w})\overset{\triangle}{=}\left((\boldsymbol{L}+\boldsymbol{F})(\boldsymbol{w})+\boldsymbol{C}(\boldsymbol{w},\boldsymbol{w})\right)\mathrm{d}t+(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w})+(\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t)+\mathrm{d}\boldsymbol{P}, \tag{2.8}$$

where table 1 details each term. Operator $\boldsymbol{L}$ represents the molecular viscosity term, $\boldsymbol{F}=\boldsymbol{F}_{\text{dif}}+\boldsymbol{F}_{\text{adv}}$, the turbulent diffusion plus the advecting velocity correction, and $\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t$, the advection by the random residual velocity. All of them are linear differential operators, while $\boldsymbol{C}$, the term representing the usual non-linear advection effect, is a bilinear differential operator. The additive noise $\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t$ corresponds to the molecular viscous dissipation of the time-uncorrelated velocity component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, while the last term on the right-hand side,

197 $\mathrm{d}\boldsymbol{P}$, is the pressure forcing.
198 Additionally, under suitable boundary conditions, the algebraic structures of the dif-
199 ferent operators can be further specified. For instance, $\boldsymbol{L}$ and $\boldsymbol{F}_{\text{dif}}$ are symmetric negative
200 operators. For the other terms, additional incompressibility conditions are needed. The
201 mass conservation equation (2.2) implies $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}=0$[1], which is actually the case in practice,
202 either because of the parametric model used for $\boldsymbol{\sigma}$ or because of the incompressibility of
203 the velocity data used to estimate $\boldsymbol{\sigma}$. Accordingly, $\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t$ is a skew-symmetric operator.
204 As a matter of fact, if $(\boldsymbol{\zeta},\boldsymbol{\xi})\overset{\triangle}{=}\int_\Omega \boldsymbol{\zeta}\cdot\boldsymbol{\xi}$ denotes the scalar product of $\boldsymbol{L}^2$, for every $\boldsymbol{\zeta}$ and
205 $\boldsymbol{\xi}$ in the Sobolev space $H_0^1(\Omega)=\{\boldsymbol{f}\in\boldsymbol{L}^2:\int_\Omega\left\|\nabla\boldsymbol{f}^T\right\|^2<\infty;\quad \boldsymbol{f}_{|\partial\Omega}=0\}$ (with partial
206 derivative taken in the weak sense and $\|\bullet\|$ stands for the Euclidean norm on $\mathbb{R}^{d\times d}$), an
207 integration by parts gives:

208 $$(2.9)\qquad (\boldsymbol{\zeta},(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\xi}))=-\sum_{k=1}^{d}\int_\Omega\zeta_k\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\xi_k=-\sum_{k=1}^{d}\int_\Omega\zeta_k\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\xi_k\right),$$

209 $$(2.10)\qquad =\sum_{k=1}^{d}\int_\Omega\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\zeta_k\right)\xi_k=-((\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\zeta}),\boldsymbol{\xi}).$$

210 The mass conservation equation also implies $\boldsymbol{\nabla}\cdot\boldsymbol{w}^*=0$ which may or may not be the
211 case in practice. If $\boldsymbol{\nabla}\cdot\boldsymbol{w}=0$, operator $\boldsymbol{C}$ is skew-symmetric with respect to the second
212 argument – *i.e.* $\boldsymbol{g}\mapsto\boldsymbol{C}(\boldsymbol{f},\boldsymbol{g})$ is skew-symmetric – whereas $\boldsymbol{F}_{\text{adv}}$ is skew-symmetric if the
213 drift correction $\boldsymbol{w}^*-\boldsymbol{w}=-\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T$ is divergence-free.
214 Moreover, the turbulent diffusion $\boldsymbol{F}_{\text{dif}}$ is related to the random skew-symmetric operator
215 $\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t$. Indeed, for every process $\boldsymbol{\xi}$ in $H^2(\Omega)=\{\boldsymbol{f}\in\boldsymbol{L}^2:\frac{\partial\boldsymbol{f}}{\partial x_i},\frac{\partial\partial\boldsymbol{f}}{\partial x_i\partial x_j}\in\boldsymbol{L}^2,1\leqslant i,j\leqslant d\}$,

216 $$(2.11)\qquad (\boldsymbol{F}_{\text{dif}}(\boldsymbol{\xi}))_k\overset{\triangle}{=}\boldsymbol{\nabla}\cdot\left(\tfrac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\xi_k\right)=\boldsymbol{\nabla}\cdot\left(\tfrac{1}{2}(\boldsymbol{\sigma}\circledast\boldsymbol{\sigma})\boldsymbol{\nabla}\xi_k\right)=\tfrac{1}{2}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)^T\left(\circledast\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)^T\xi_k\right),$$

217 $$(2.12)\qquad =\tfrac{1}{2}\boldsymbol{G}\left(\circledast\boldsymbol{G}(\xi_k)\right)=\left(\tfrac{1}{2}\boldsymbol{G}\left(\circledast\boldsymbol{G}(\boldsymbol{\xi})\right)\right)_k=\left(-\tfrac{1}{2}\boldsymbol{G}^*\left(\circledast\boldsymbol{G}(\boldsymbol{\xi})\right)\right)_k,$$

218 where $\boldsymbol{G}^*$ denotes the adjoint of $\boldsymbol{G}$. It also shows that $\boldsymbol{F}_{\text{dif}}(\boldsymbol{w})\mathrm{d}t=\frac{1}{2}\mathrm{d}\left\langle\boldsymbol{G}(\boldsymbol{w}),\boldsymbol{B}_t\right\rangle$, where
219 $\langle\boldsymbol{g},\boldsymbol{h}\rangle$ denotes the quadratic covariation of any functions $\boldsymbol{g}$ and $\boldsymbol{h}$. The diffusion term
220 explicitly appears when working with the Itō stochastic integral and is only implicitly
221 taken into account with Stratonovich integral [6, 68].
222 As discussed in more detail in section 5, these algebraic properties make the LU Navier-
223 Stokes model – and to a certain extent, its reduced order versions – conservative (up to
224 molecular viscosity and boundary conditions effects).

225 **3. Galerkin projection.** To sample good priors for future Bayesian estimation algo-
226 rithms, we aim at deriving a computationally efficient fluid dynamics ROM able to quantify
227 its own errors with respect to the true fluid dynamics (i.e. the Navier-Stokes equations).
228 As previously mentioned, standard Galerkin techniques – even with the best determinis-
229 tic closures – are hardly capable of such a goal as they were not originally designed for
230 it. Hence, we propose to perform Galerkin projections on the LU Navier-Stokes model
231 instead, and to study its appropriateness for this sort of tasks.

232 **3.1. A ROM with correlated additive and multiplicative noise.** Let $\boldsymbol{v}$ be the real
233 velocity field (i.e. the Navier-Stokes equations' solution) and $\boldsymbol{\phi}_0$, a background velocity
234 field, typically the velocity temporal mean $\overline{\boldsymbol{v}}\overset{\triangle}{=}\frac{1}{T}\int_0^T\boldsymbol{v}$. To reduce the state space dimen-
235 sion, we project the fluid velocity anomaly, $\boldsymbol{v}-\boldsymbol{\phi}_0$, in a subspace spanned by a number of

[1] i.e. $\boldsymbol{\nabla}_{\boldsymbol{x}}\cdot\breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z})=0,\forall(\boldsymbol{x},\boldsymbol{z})\in\Omega^2$

orthonormal spatial modes $(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}$.

$$\boldsymbol{v}(\boldsymbol{x},t,\omega)=\underbrace{\sum_{i=0}^{n} b_i(t,\omega)\boldsymbol{\phi}_i(\boldsymbol{x})}_{\triangleq \boldsymbol{w}^R}+\underbrace{\text{Residual}}_{\triangleq \boldsymbol{v}'} \qquad \forall(\boldsymbol{x},t)\in\Omega\times[0,T], \tag{3.1}$$

where $b_0=1$ and $\boldsymbol{\phi}_0=\overline{\boldsymbol{v}}$ by convention. The associated temporal coefficients, b_i, are possibly random and depend on a realization ω of a sample space $\breve{\Omega}$ whereas the reduced basis functions $\boldsymbol{\phi}_i$, are assumed to be deterministic and stationary.

As typically done while working with ROMs, we aim at specifying the evolution of the projected velocity field $\boldsymbol{w}^R$. A standard technique for that is the Galerkin projection of the physical PDE – here, the Navier-Stokes equations – onto the reduced basis' functions $\boldsymbol{\phi}_i$, where the resolved component $\boldsymbol{w}^R$ is approximated by the solution of these projected equations. Mode truncation can create many problems. For moderately turbulent to turbulent flows and small dimension n, a closure model to handle the truncated modes is unavoidable. With the LU setting – described in section 2 – an elegant stochastic alternative for this problem can be derived. Accordingly, we will (i) assume that the residual velocity $\boldsymbol{v}'$ is time-uncorrelated and is denoted $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t/dt$ and (ii) approximate the resolved component $\boldsymbol{w}^R$ to a realization of the solution of the Galerkin projection of the stochastic Navier-Stokes representation (2.1)-(2.2). The former hypothesis is a debatable choice with respect to some ROM's applications whose pertinence and limitations will be discussed further in section 4.5.

To obtain this stochastic ROM, the SPDE (2.8) is first projected onto the divergence-free function space through the non-local Leray operator $\boldsymbol{\mathcal{P}}=\mathbb{I}_d-\boldsymbol{\nabla}\boldsymbol{\nabla}^T\Delta^{-1}$. This projection, which requires the resolution of a Poisson equation is used to simplify the system by removing the pressure term. Since divergence-free solution is considered (see section 3.2), the resolved velocity component w^R is naturally incompressible and we get

$$\boldsymbol{\mathcal{P}}\mathrm{d}_t\boldsymbol{w}^R=\mathrm{d}_t\boldsymbol{w}^R, \qquad \boldsymbol{\mathcal{P}}\boldsymbol{L}(\boldsymbol{w}^R)=\boldsymbol{L}(\boldsymbol{w}^R), \qquad \boldsymbol{\mathcal{P}}\mathrm{d}\boldsymbol{P}=0. \tag{3.2}$$

Moreover, $\boldsymbol{\mathcal{P}}\left(\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t\right)=\left(\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t\right)$ because of the necessary incompressibility of the Brownian term in the continuity equation (2.2)[2]. Then, the resulting SPDE is projected onto each of the reduced basis' functions:

$$\mathrm{d}b_i=(\boldsymbol{\phi}_i,\mathrm{d}_t\boldsymbol{w}^R)=(\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b})\overset{\triangle}{=}\left(\boldsymbol{\phi}_i,\boldsymbol{\mathcal{P}}(\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right), \qquad 1\leqslant i\leqslant n, \tag{3.3}$$

where $\boldsymbol{b}=(b_i)_{0\leqslant i\leqslant n}$ and

$$\begin{aligned}(\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b})=&\sum_{p=0}^{n}\underbrace{\left(\boldsymbol{\phi}_i,\boldsymbol{L}(\boldsymbol{\phi}_p)\right)}_{\triangleq l_{pi}}b_p\mathrm{d}t+\sum_{p=0}^{n}\underbrace{\left(\boldsymbol{\phi}_i,\boldsymbol{\mathcal{P}}\boldsymbol{F}(\boldsymbol{\phi}_p)\right)}_{\triangleq \breve{f}_{pi}}b_p\mathrm{d}t+\sum_{p,q=0}^{n}\underbrace{\left(\boldsymbol{\phi}_i,\boldsymbol{\mathcal{P}}\boldsymbol{C}(\boldsymbol{\phi}_p,\boldsymbol{\phi}_q)\right)}_{\triangleq c_{pqi}}b_pb_q\mathrm{d}t\\&+\underbrace{\left(\boldsymbol{\phi}_i,(\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t)\right)}_{\triangleq(\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t)}+\sum_{p=0}^{n}\underbrace{\left(\boldsymbol{\phi}_i,\boldsymbol{\mathcal{P}}(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\phi}_p)\right)}_{\triangleq(\boldsymbol{\alpha}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t)}b_p, \qquad 1\leqslant i\leqslant n.\end{aligned} \tag{3.4}$$

The terms $(\boldsymbol{\alpha}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t)_{1\leqslant p,i\leqslant n}$ and $((\boldsymbol{\theta}_{i\bullet}+\boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_t)_{1\leqslant i\leqslant n}$ correspond to a Gaussian skew-symmetric matrix and a Gaussian vector respectively, both with correlated coefficients. If the functions $\boldsymbol{\phi}_i$ were spatial Fourier modes associated with small wave-numbers, the ROM (3.3) would represent a (stochastic) Large-Eddy-Simulation-like model expressed in Fourier space and $\boldsymbol{b}$ would be the set of Fourier coefficients of the solution. Hereafter, we will solely focus on the POD.

[2] which is coherent with the fact that $\boldsymbol{\nabla}\cdot\boldsymbol{v}'=0$ from (3.1) because $\boldsymbol{\nabla}\cdot\boldsymbol{v}=0$.

3.2. Proper Orthogonal Decomposition. In the POD framework, the reduced basis' functions are computed through a set of velocity snapshots $(\boldsymbol{v}_{\text{obs}}(\bullet,t_i))_{0\leqslant i\leqslant N-1}$. More precisely, they are a solution to the constrained optimization problem:

$$\underset{(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}}{\text{Maximize}} \sum_{i=1}^{n} \overline{(\boldsymbol{\phi}_i, \boldsymbol{v}_{\text{obs}}-\overline{\boldsymbol{v}})^2} \quad \text{subject to } \left(\boldsymbol{\phi}_i,\boldsymbol{\phi}_j\right)=\delta_{ij},\quad 1\leqslant i,j\leqslant n. \tag{3.5}$$

Reduced basis functions $\boldsymbol{\phi}_i$ are thus the n orthonormal functions which can best explain the snapshots' temporal variability. Similarly to a principal component analysis (PCA), the solutions of this optimization problem are the eigenfunctions of the velocity anomalies' $(\boldsymbol{v}_{\text{obs}}-\overline{\boldsymbol{v}})$ spatial covariance. Numerically, this matrix is extremely large, and we generally opt to solve the dual problem instead: we seek the eigenvalue decomposition of the velocity anomalies' temporal covariance:

$$C^v_{ij} = \left((\boldsymbol{v}_{\text{obs}}-\overline{\boldsymbol{v}})(\bullet,t_i),(\boldsymbol{v}_{\text{obs}}-\overline{\boldsymbol{v}})(\bullet,t_j)\right),\quad 0\leqslant i,j\leqslant N-1. \tag{3.6}$$

This method is often referred to as the snapshot method. In the POD framework, the mode $\boldsymbol{\phi}_0=\overline{\boldsymbol{v}}$ is set to the time averaged velocity, and the temporal coefficients energies $\overline{b_i^2}$ are denoted λ_i. Furthermore, if the snapshots describe a divergence-free velocity field $\boldsymbol{v}_{\text{obs}}$, the spatial bases $\boldsymbol{\phi}_i$ are divergence-free as well.

4. Estimations of subgrid terms. To close our stochastic ROM (3.3)-(3.4), we need to estimate the variance tensor $\boldsymbol{a}$ (involved in the ROM matrix $\breve{\boldsymbol{f}}=\breve{\boldsymbol{f}}_{\text{dif}}+\breve{\boldsymbol{f}}_{\text{adv}}$ defined in table 1 and in equation (3.4) under the braces) as well as the ROM noise variances and correlations. Firstly, if we recall that $b_0=1$, we note that:

$$\left(\boldsymbol{\alpha}_{\bullet i\bullet}\mathrm{d}\boldsymbol{B}_t\right)^T\boldsymbol{b}+\left(\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t\right)=\sum_{k=1}^{n}\left(\boldsymbol{\alpha}_{ki\bullet}\mathrm{d}\boldsymbol{B}_t\right)b_k+\left((\boldsymbol{\theta}_{i\bullet}+\boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_t\right),\quad 1\leqslant i\leqslant n. \tag{4.1}$$

The multiplicative and additive noises of the ROM correspond to the first and second term of the right-hand side, respectively. To simplify notations, we write:

$$\tilde{\boldsymbol{\alpha}}_{pi\bullet}=\boldsymbol{\alpha}_{pi\bullet}+\delta_{p0}\boldsymbol{\theta}_{i\bullet},\quad 1\leqslant i\leqslant n,\quad 0\leqslant p\leqslant n. \tag{4.2}$$

To fully specify the ROM, the following correlations must be estimated:

$$\left\{\begin{array}{rcl}\boldsymbol{a}(\boldsymbol{x}) &=& \mathbb{E}\left\{\left(\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t\right)^T\right\}/\mathrm{d}t \quad \forall \boldsymbol{x}\in\Omega,\\ \Sigma^{\boldsymbol{\alpha}}_{pi,qj} &=& \mathbb{E}\left\{\left(\tilde{\boldsymbol{\alpha}}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t\right)\left(\tilde{\boldsymbol{\alpha}}_{qj\bullet}\mathrm{d}\boldsymbol{B}_t\right)\right\}/\mathrm{d}t,\quad 1\leqslant i,j\leqslant n,\quad 0\leqslant p,q\leqslant n.\end{array}\right. \tag{4.3}$$

4.1. The curse of dimensionality. Computing correlations of Gaussian noises $(\tilde{\boldsymbol{\alpha}}\mathrm{d}\boldsymbol{B}_t)$ involves the two-point quadratic cross-variation tensor of the small-scale velocity:

$$\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})=\mathbb{E}\left\{\left(\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\sigma}(\boldsymbol{y})\mathrm{d}\boldsymbol{B}_t\right)^T\right\}/\mathrm{d}t \quad \forall \boldsymbol{x},\boldsymbol{y}\in\Omega. \tag{4.4}$$

The coefficients (i,j) of the covariance matrix of the Gaussian vector $(\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t)$ can be expressed as follows:

$$\mathbb{E}\left\{\left(\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\theta}_{j\bullet}\mathrm{d}\boldsymbol{B}_t\right)\right\}=\mathbb{E}\left\{\left(\boldsymbol{\phi}_i,\nu\boldsymbol{\nabla}^2\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\left(\nu\boldsymbol{\nabla}^2\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t,\boldsymbol{\phi}_j\right)\right\}, \tag{4.5}$$

$$=\iint_{\Omega^2}\boldsymbol{\phi}_i(\boldsymbol{x})^T\nu^2\boldsymbol{\nabla}_x^2\boldsymbol{\nabla}_y^2\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}t\;\boldsymbol{\phi}_j(\boldsymbol{y})\,d\boldsymbol{x}d\boldsymbol{y}. \tag{4.6}$$

Since, by definition, $\forall \boldsymbol{x}\in\Omega,\ \boldsymbol{a}(\boldsymbol{x})=\boldsymbol{Q}(\boldsymbol{x},\boldsymbol{x})$, the whole stochastic ROM is closed by specifying the tensor $\boldsymbol{Q}$. In practice, this tensor is often heterogeneous – i.e. $\boldsymbol{\sigma}\boldsymbol{B}_t$ is

non-stationary in space. Thus, its spatially-discrete version is expected to be so large that its estimation or even its storage becomes prohibitive. To overcome this difficulty, a first strategy is to assume a model structure for the covariance as per [41, 66, 64, 68]. If the small-scale velocity is observed, other techniques can be considered. Indeed, one can build a POD representation of the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, learned from available snapshots or realizations as in [7, 24, 64, 68]. Depending on the desired accuracy for the noise representation, a great number of modes would have to be estimated and the number of coefficients involved in the ROM can quickly grow out of control. Here, we rely on a method specifically devised for this kind of ROM frameworks: the noise structure is again learned from observed residual velocity snapshots, but without passing through the covariance $\boldsymbol{Q}$. It enables us to directly estimate the correlations of the random ROM's coefficients, as it will be explained later on.

4.2. Variance tensor estimation. Considering the variance tensor $\boldsymbol{a}$ stationary leads to a simple estimator through time averaging as in [67]:

$$\widehat{\boldsymbol{a}} = \Delta t \; \overline{\boldsymbol{v}'_{\text{obs}} \left(\boldsymbol{v}'_{\text{obs}}\right)^T}, \tag{4.7}$$

where

$$\boldsymbol{v}'_{\text{obs}} = \boldsymbol{v}_{\text{obs}} - \boldsymbol{w}^R_{\text{obs}} = \boldsymbol{v}_{\text{obs}} - \sum_{i=0}^{n} b_i^{\text{obs}} \boldsymbol{\phi}_i, \tag{4.8}$$

is the observed POD's residual velocity. Since $\boldsymbol{a}\mathrm{d}t = \mathrm{d}\left\langle \boldsymbol{\sigma}\boldsymbol{B}, (\boldsymbol{\sigma}\boldsymbol{B})^T \right\rangle$, the estimator (4.7) is consistent thanks to the quadratic covariation definition.

4.3. Noise correlation estimation. Besides the variance tensor, there are $n^2(n+1)^2$ correlations to estimate according to (4.3). For any function $\boldsymbol{\xi}$ in $H^2(\Omega)$, let us introduce the linear functional:

$$K_{jq}[\boldsymbol{\xi}] \stackrel{\triangle}{=} \left(\boldsymbol{\phi}_j, -\boldsymbol{\mathcal{P}}\left[(\boldsymbol{\xi}\cdot\boldsymbol{\nabla})\,\boldsymbol{\phi}_q\right] + \delta_{q0}\,\nu\Delta\boldsymbol{\xi}\right), \quad 1 \leqslant j \leqslant n, \quad 0 \leqslant q \leqslant n. \tag{4.9}$$

Using this notation, the noise's covariance can be estimated as follows:

$$\widehat{\Sigma^{\boldsymbol{\alpha}}_{pi,qj}} = \tfrac{\Delta t}{\lambda_p^{\text{obs}}}\; K_{jq}\left[\; \overline{b_p^{\text{obs}} \left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'' \boldsymbol{v}'_{\text{obs}}} \;\right], \quad 1 \leqslant i,j \leqslant n, \quad 0 \leqslant p,q \leqslant n, \tag{4.10}$$

where $b_0^{\text{obs}} = \lambda_0^{\text{obs}} = 1$ and for $1 \leqslant i \leqslant n$,

$$b_i^{\text{obs}} = \left(\boldsymbol{\phi}_i, \boldsymbol{v}_{\text{obs}}\right), \tag{4.11}$$

$$\lambda_i^{\text{obs}} = \overline{(b_i^{\text{obs}})^2}, \tag{4.12}$$

$$\left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'' = \left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)' - \overline{\left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'}, \tag{4.13}$$

$$\left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)' = \left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right) - \left(\left(\boldsymbol{b}^{\text{obs}}\right)^T \left(\boldsymbol{l} + \breve{\boldsymbol{f}}\right)_{\bullet i} + \left(\boldsymbol{b}^{\text{obs}}\right)^T \boldsymbol{c}_{\bullet\bullet i}\, \boldsymbol{b}^{\text{obs}}\right), \tag{4.14}$$

$$\left(\tfrac{\Delta b_i^{\text{obs}}}{\Delta t}\right)(t_k) = \frac{b_i^{\text{obs}}(t_k + \Delta t) - b_i^{\text{obs}}(t_k)}{\Delta t}, \quad 0 \leqslant k \leqslant N-1. \tag{4.15}$$

To ensure the noise's covariance matrix to have the desired symmetric non-negative structure, we only keep the symmetric part of the estimated tensor (4.10) and set its possible negative eigenvalues to zero. The part inside the functional K_{jq} in the estimator (4.10) is inspired from [32], where products of martingale time increments are projected onto

orthogonal functions of $L^2([0,T])$. For ROMs with a small number of dimensions n, the estimator's computational cost is remarkably low as the observed coefficients of the reduced solution $\boldsymbol{b}^{\text{obs}}$ were already computed by the method of snapshots (see section 3.2). Hence, the computational cost for $b_p^{\text{obs}}\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)''$ is negligible and the part inside K_{jq} only involves a projection of the residual velocity along the time dimension. Finally, the $n(n+1)$ differential operators K_{jq} are computed only on $n(n+1)$ functions that do not depend on the time variable.

We prove the consistency of this estimator in Appendix A by capitalising on the quadratic covariation's definition and the orthogonality of the observed temporal coefficients b_i^{obs}. In addition, this can also be extended to non-orthogonal coefficients by solving the linear system engendered by the matrix $\left(\overline{b_p b_k}\right)_{pk}$.

4.4. Noise dimension reduction. Using the Cholesky decomposition $\boldsymbol{\sigma}^{\boldsymbol{\alpha}}$ of the noise covariance tensor:

$$\Sigma^{\boldsymbol{\alpha}}_{pi,qj} = \sum_{lk} \sigma^{\boldsymbol{\alpha}}_{pi,lk}\sigma^{\boldsymbol{\alpha}}_{qj,lk}, \quad 1 \leqslant i,j \leqslant n, \quad 0 \leqslant p,q \leqslant n, \tag{4.16}$$

and $O(n^2)$ independent white noises, one can sample realizations of the ROM Gaussian noise terms $\tilde{\boldsymbol{\alpha}}\mathrm{d}\boldsymbol{B}_t$. However, $\boldsymbol{\sigma}^{\boldsymbol{\alpha}}$ has $O(n^4)$ coefficients, while the ROM works with only n modes and involves – leaving the noise terms aside – $O(n^3)$ coefficients. Therefore, we propose to reduce the noise dimension through a tensorial PCA of $\boldsymbol{\Sigma}^{\boldsymbol{\alpha}}$, eventually only keeping the n first eigenvectors. This leads to the following sampling strategy:

$$\tilde{\boldsymbol{\alpha}}\mathrm{d}\boldsymbol{B}_t \approx \sum_{k=1}^{n} \tilde{\boldsymbol{\alpha}}^R_k \mathrm{d}\beta^{(k)}_t, \tag{4.17}$$

where $\left(\tilde{\boldsymbol{\alpha}}^R_k\right)_k \in \mathbb{R}^{(n+1)\times n}$ are the matrix forms of the first n eigenvectors (weighted by the corresponding eigenvalues' square roots) and $(\beta^{(k)})_k$ are n independent one-dimensional Brownian motions. Since $\tilde{\boldsymbol{\alpha}}\mathrm{d}\boldsymbol{B}_t$ is a multiplicative noise and the temporal coefficients b_i have various amplitudes $\sqrt{\lambda_i}$, the covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{\alpha}}$ is adequately re-normalized by the amplitudes $\sqrt{\lambda_i}$ before applying the PCA.

It is important to note that the methodology described here is different from a more usual methodology based on PCA decomposition of the residual velocity $\boldsymbol{v}'$ keeping n modes and assuming that the corresponding temporal coefficients $(b_i)_{n+1\leqslant i\leqslant 2n}$ are time-decorrelated. The complexity of the final ROM is the same in both methodologies but our ROM maximizes the noise's variance instead of the residual velocity's variance. Thus, our proposed method is better suited in terms of ROM UQ.

4.5. Time down-sampling rate. Under the LU Navier-Stokes model hypothesis, the unresolved term of the velocity field corresponds to a noise uncorrelated in time. This assumption is consistent with the fact that the higher-order coefficients of the reduced order solution often tend to have shorter correlation time in fluid dynamics systems. However, in practice, this assumption is not found to be true, and it is a recurrent issue for the data-driven modeling of systems combining fast and slowly evolving components [4, 5, 59, 60]. Consequently, a time down-sampling scheme is proposed to force the noise terms to be as decorrelated as possible.

By assuming that the spatially averaged covariance function has a Gaussian form with a standard deviation equals to the correlation time τ, a simple expression allows us to

compute it. For a given unresolved velocity correlation matrix we write:

$$C_{ij}^{v'} = \left(\boldsymbol{v}'_{\text{obs}}(\bullet, t_i), \boldsymbol{v}'_{\text{obs}}(\bullet, t_j)\right) = C_{ij}^{v} - \sum_{k=1}^{n} b_k^{\text{obs}}(t_i) b_k^{\text{obs}}(t_j), \quad 0 \leqslant i, j \leqslant N-1, \tag{4.18}$$

and its associated stationary covariance function

$$\text{Cov}_s(t_p) = \tfrac{1}{N-p} \sum_{q=0}^{N-1-p} C_{q,q+p}^{v'}, \quad 0 \leqslant p \leqslant N-1. \tag{4.19}$$

We propose the following correlation time estimation:

$$\widehat{\tau} = \sqrt{2 \frac{\overline{\text{Cov}_s^2}}{\overline{\left(\frac{\Delta \text{Cov}_s}{\Delta t}\right)^2}}}, \tag{4.20}$$

using a forward Euler temporal discretization of the stationary covariance:

$$\frac{\Delta \text{Cov}_s}{\Delta t}(t_p) \triangleq \frac{Cov_s(t_p + \Delta t) - Cov_s(t_p)}{\Delta t}, \quad 0 \leqslant p \leqslant N-1. \tag{4.21}$$

These estimations follow over-simplified assumptions known to have a restricted validity. Experimentally, they still systematically provide the best simulation results when compared to other more complex estimators (whose derivation is outside the scope of this paper). Moreover, with a white unresolved velocity with a dirac stationary covariance function, we obtain $\widehat{\tau} = \mathrm{d}t$. This result is exactly what we would expect from a white unresolved velocity, to prevent an overly aggressive down-sampling.

Before computing the estimations presented in sections 4.2, 4.3 and 4.4, we use this estimated correlation time $\widehat{\tau}$ to down-sample both the entire dataset and the observed coefficients of the reduced order solution, leaving us with a time step $\Delta t \approx \widehat{\tau}$. On top of the considerable gain in accuracy, this time down-sampling reduces the amount of data to process, and hence, makes the off-line ROM building process faster.

Having estimated and specified all of the stochastic ROM's parameters in equations (3.3) and (3.4), it is now possible to forecast ensembles of realizations of the ROM through Monte-Carlo simulations. Before presenting numerical results, we discuss some important properties of LU models and of their reduced order versions.

5. Kinetic energy budget.

First, the conservative properties of the LU Navier-Stokes representation are recalled, followed by a proof that by combining Galerkin projections and the advecting velocity correction compressibility, an intrinsic energy dissipation appears.

5.1. Full-order model budget.

As derived in [65, 68], since the pressure does not influence the energy budget, by neglecting the molecular viscosity, the divergence of $\boldsymbol{w}$ and boundary conditions effects, and by applying Itō's lemma, the expression for the kinetic energy budget writes:

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}\|\boldsymbol{w}\|_{\boldsymbol{L}^2}^2\right) = \underbrace{\left(-\tfrac{1}{2}\boldsymbol{G}^*\left(\circledast \boldsymbol{G}(\boldsymbol{w})\right), \boldsymbol{w}\right)}_{\text{Loss by diffusion}} + \underbrace{\tfrac{1}{2}\left\|\boldsymbol{G}(\boldsymbol{w})\right\|_{\text{HS}}^2}_{\substack{\text{Energy flux} \\ \text{from the noise}}} = 0 \quad \forall t \in [0, T], \tag{5.1}$$

where $\|\boldsymbol{\xi}\|_{\boldsymbol{L}^2}^2 = (\boldsymbol{\xi}, \boldsymbol{\xi})$ is the squared norm of $\boldsymbol{L}^2$ and

$$\|\boldsymbol{\eta}\|_{\text{HS}}^2 = \iint_{\Omega^2} \sum_{i,j=1}^{d} \breve{\eta}_{i,j}^2(\boldsymbol{x}, \boldsymbol{z}) d\boldsymbol{x} d\boldsymbol{z} \tag{5.2}$$

is the squared Hilbert-Schmidt norm of the integral operator $\boldsymbol{\eta}$. This enables us to state that the energy is conserved for each realization of the stochastic process, and as a direct consequence of this, the energy intake of the noise and the dissipation by the turbulent diffusion must exactly compensate each other. The latter dissipates the kinetic energy of the mean $\|\mathbb{E}\{\boldsymbol{w}\}\|^2_{\boldsymbol{L}^2}$, while the former only releases random energy $\|\boldsymbol{w}-\mathbb{E}\{\boldsymbol{w}\}\|^2_{\boldsymbol{L}^2}$ into the system. Thus, the time-uncorrelated component of the velocity field drains energy from the mean field to the random component of $\boldsymbol{w}$.

We can also express this energy transfer with the expectation of equation (5.1):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{Var}(\boldsymbol{w}) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\|\boldsymbol{w}-\mathbb{E}\{\boldsymbol{w}\}\|^2_{\boldsymbol{L}^2} = -\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{E}\{\boldsymbol{w}\}\|^2_{\boldsymbol{L}^2} \qquad \forall t\in[0,T]. \tag{5.3}$$

Besides the physical relevance of energy conservation, variance inflation and its relation to the mean field are also of primary interest for data assimilation or ensemble forecasting issues.

5.2. Reduced-order model budget. Following (5.1), the full-order LU Navier-Stokes model (2.1)-(2.2) conserves the kinetic energy, up to molecular viscosity and boundary condition effects. With reduced order models, the advecting velocity correction is expected however to create either energy compression or dilation, and the mode truncation, to introduce a small energy leak. The ROM (3.3) does not exactly solve the global LU Navier-Stokes model (2.1)-(2.2) but the Galerkin projection of the divergence-free component of equation (2.1) :

$$\mathrm{d}_t\boldsymbol{w}^R = \sum_{i=0}^{n}\mathrm{d}b_i\ \boldsymbol{\phi}_i = \sum_{i=1}^{n}(\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b})\ \boldsymbol{\phi}_i = \Pi_\phi\left[\boldsymbol{\mathcal{P}}(\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right] \qquad \forall t\in[0,T], \tag{5.4}$$

where Π_ϕ is the projection onto the reduced subspace. Specifically, for any function $\boldsymbol{\zeta}\in\boldsymbol{L}^2$, the projection Π_ϕ is defined as $\Pi_\phi[\boldsymbol{\zeta}] \triangleq \sum_{p=1}^{n}\left(\boldsymbol{\phi}_p,\boldsymbol{\zeta}\right)\boldsymbol{\phi}_p$. As previously stated, we can evaluate the variation of kinetic energy with the Itō formula. The following result is proved in Appendix B for every $t\in[0,T]$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}\|\boldsymbol{w}^R\|^2_{\boldsymbol{L}^2}\right) = \underbrace{-\tfrac{1}{2}\left\|\boldsymbol{G}(\boldsymbol{w}^R)\right\|^2_{\mathrm{HS}}}_{\text{Loss by diffusion}} + \underbrace{\tfrac{1}{2}\left\|\Pi_\phi\left[(\boldsymbol{G})(\boldsymbol{w}^R)\right]\right\|^2_{\mathrm{HS}}}_{\text{Energy flux from the noise}} - \underbrace{\tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T,\|\boldsymbol{w}^R\|^2\right)}_{\substack{\text{Advecting velocity correction}\\ \text{compressibility}}}, \tag{5.5}$$

$$= -\tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T,\|\boldsymbol{w}^R\|^2\right) \underbrace{-\tfrac{1}{2}\left\|\Pi_\phi^\perp\left[(\boldsymbol{G})(\boldsymbol{w}^R)\right]\right\|^2_{\mathrm{HS}}}_{<0}, \tag{5.6}$$

where $\|\bullet\|$ stands for the Euclidean norm on $\mathbb{R}^d$ and $\Pi_\phi^\perp = \mathbb{I}_d - \Pi_\phi$ is the projector onto the orthogonal complement of the space spanned by the reduced basis' functions.

The first term of (5.6) corresponds to the fluid compression/dilation created by the velocity correction $\boldsymbol{w}^* - \boldsymbol{w} = -\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T$. By construction, the ROM meets the incompressible condition – i.e. $\boldsymbol{\nabla}\cdot\boldsymbol{w}=0$ – but unlike the LU Navier-Stokes model, it does not satisfy the finite-variation part of the mass conservation (2.2): $\boldsymbol{\nabla}\cdot\boldsymbol{w}^*=0$. The velocity correction divergence is associated to turbulence heterogeneity [6, 16, 67], corresponding to the spatial maxima (positive energy fluxes) and minima (negative energy fluxes) of the turbulence's kinetic energy $\|\boldsymbol{v}'\|^2$.

The second term of (5.6) informs about the energy flux due to mode truncation. The ROM's subgrid diffusion extracts energy from some temporal coefficients of the reduced order solution, while the multiplicative noise distributes it to others. These stabilizing and

destabilizing effects recreate a large part of the energy fluxes between coefficients of the reduced solution, otherwise lost in deterministic ROMs. Thus, the uncorrelated velocity component drains energy from the coefficients of the reduced solution to give it back to the temporal coefficients of the truncated modes – i.e. modes orthogonal to the reduced space. This energy flux is exactly characterized by the second term of (5.6). But, our stochastic ROM cannot transfer energy from the truncated modes to the reduced solution's coefficients, as the former cannot be specified due to the truncation.

Such a dissipation could be perfectly prevented by considering $-\frac{1}{2}\left(\Pi_\phi\left[\boldsymbol{G}\right]\right)^*\left(\circledast\Pi_\phi\left[\boldsymbol{G}\right]\right)$ represented in the ROM by $\left(-\frac{1}{2}\sum_{i=1}^{n}\Sigma^{\boldsymbol{\alpha}}_{pi,qi}\right)_{p,q}$ (see equations (B.11) and (B.12) in Appendix B), instead of the reduced version of the full turbulent diffusion operator $\left(-\frac{1}{2}\left(\boldsymbol{\phi}_p,\ \boldsymbol{G}^*\left(\circledast\boldsymbol{G}(\boldsymbol{\phi}_q)\right)\right)\right)_{p,q}$. This correction would have been implicit using the Galerkin projection with Stratonovich calculus. As demonstrated in [6, 64, 68], using Stratonovich integral, the turbulent dissipation term does not explicitly appear, neither when deriving the ROM from scratch nor when switching notations after the Galerkin projection. Let us stress that the modified advection is still present in the Stratonovich form of the LU model. Changing the ROM notations from Itō to Stratonovich [42], the following correction term appears:

$$\frac{1}{2}\mathrm{d}\left\langle \boldsymbol{b}^T\boldsymbol{\alpha}_{\bullet p\bullet},\boldsymbol{B}\right\rangle = \sum_{q=0}^{n}\left(\frac{1}{2}\sum_{i=1}^{n}\Sigma^{\boldsymbol{\alpha}}_{ip,qi}\right)b_q\,\mathrm{d}t = \sum_{q=0}^{n}\left(-\frac{1}{2}\sum_{i=1}^{n}\Sigma^{\boldsymbol{\alpha}}_{pi,qi}\right)b_q\,\mathrm{d}t. \tag{5.7}$$

In Stratonovich form, the ROM also still implicitly includes the diffusion term ensuing from the truncature. Notwithstanding, this additional dissipation in (5.6) is not a problem as it enables the setting of an energy dissipation mechanism representative of a direct energy cascade in the truncated modes' coefficients. Hence, for very turbulent flows described by few modes, the energy leak appears to be necessary as our main concern is to restore energy fluxes between the coefficients of the reduced order system. Indeed, those energy fluxes can be very difficult to model correctly in reduced versions of non-linear systems. As an example, [69] observe that a basis encoding 50% of the energy can lead to a ROM missing more than 98% of energy transfers. These missing energy fluxes create instabilities in some coefficients of the reduced solution (missing negative energy fluxes) while overdamping others (missing positive energy fluxes). Restoring the energy fluxes between coefficients of the reduced system is a significant challenge LU models tackle through the combination of multiplicative noise and turbulent diffusion.

To note, the noise dimensionality reduction of section 4.4 introduces an additional energy leak due to the fact that we are only keeping n eigenvectors over $n(n+1)$ in the noise covariance matrix's PCA. This introduces another negative energy flux, as the noise variance is reduced while the dissipation is maintained.

6. Ensemble forecasting.

As a means to measure the performance of the proposed ROM in UQ scenarios, ensemble simulations were carried out using the data available from DNS simulations of wake flows at Reynolds 100 and 300, the former being quasi two dimensional while the latter, fully three dimensional. Wake flows are well-studied, non-linear, oscillatory flows which are physically produced by a uniform-velocity inflow facing a solid obstacle – here, a cylinder. Vortices are thus created behind the obstacle and periodically detach from it.

The dynamics, and in particular their intrinsic dimensionality, strongly differ between Reynolds 100 and 300. Indeed, as illustrated by figure 1, at Reynolds 100, most of the energy is concentrated in just a few coefficients of the reduced order solution. In contrast, at Reynolds 300, the two first modes are meaningful for a rough approximation, but energy

spreads over many degrees of freedom. The top panels of figures 3 and 7 confirm a strong difference in complexity between the two flows.

The ROM's results $\boldsymbol{w}^R$ will be compared directly to DNS simulations $\boldsymbol{v}_{\text{ref}}$ on a time interval outside intervals on which the ROM's coefficients and basis functions have been estimated.

6.1. Baseline ROMs. To better appreciate our stochastic ROM's capabilities, results are compared to two different state-of-the-art algorithms.

6.1.1. Deterministic baseline ROM. The first state-of-art ROM is a prototype of deterministic ROM widely used in fluid dynamics. After a POD-Galerkin on the classical Navier-Stokes equations, the molecular viscosity coefficient $1/Re$ is replaced by an eddy viscosity coefficient $1/Re^{\text{ev}} \geqslant 1/Re$ [15]:

$$\frac{\mathrm{d}}{\mathrm{d}t} b_i = \frac{Re}{Re^{\text{ev}}}\, \boldsymbol{b}^T \boldsymbol{l}_{\bullet i} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i}\, \boldsymbol{b}, \quad 1 \leqslant i \leqslant n. \tag{6.1}$$

The eddy viscosity coefficient is typically fitted by least squares using the dataset b^{obs}. Without that correction, some coefficients of the reduced solution might become unstable because of the missing negative energy fluxes due to mode truncation mentioned in 5.2. The eddy viscosity term generally manages to stabilize those coefficients, but this does not always make this data-driven method accurate. In order to emulate that unstable behavior, we also simulate a simple DNS POD-Galerkin with no eddy viscosity model, i.e. the ROM (6.1) with $Re^{\text{ev}} = Re$.

6.1.2. Stochastic baseline ROM. As a deterministic model, the ROM above can hardly be used for dynamics-error UQ nor ensemble-based data assimilation. As a matter of fact, a simple randomization of the initial conditions are known to usually lead to error underestimation (see section 1). So, to keep things simple, we define a second baseline ROM by adding a white noise term to the first baseline ROM. This constitutes a simple *adhoc* randomization technique for a given dynamical system through a Gaussian additive forcing. Despite its potential lack of physical relevance, such a strategy is very often adopted in data-assimilation applications [17].

$$\mathrm{d}b_i = \left(\frac{Re}{Re^{\text{ev}}}\, \boldsymbol{b}^T \boldsymbol{l}_{\bullet i} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i}\, \boldsymbol{b}\right) \mathrm{d}t + \boldsymbol{\sigma}^{\text{ev}}_{i\bullet} \mathrm{d}\boldsymbol{W}_t, \quad 1 \leqslant i \leqslant n, \tag{6.2}$$

where $\boldsymbol{\sigma}^{\text{ev}} \in \mathbb{R}^{n\times n}$ is the Cholesky decomposition of the ROM's noise covariance $\boldsymbol{\Sigma}^{\text{ev}} = \boldsymbol{\sigma}^{\text{ev}} \left(\boldsymbol{\sigma}^{\text{ev}}\right)^T$ and $t \mapsto \boldsymbol{W}_t \in \mathbb{R}^n$ is a vector of n independent Brownian motions. Since no prior physical information is available for these baseline models, the noise's covariance is also learned from data as follows:

$$\widehat{\Sigma^{\text{ev}}_{ij}} = \Delta t\, \overline{\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)''_{\text{ev}} \left(\frac{\Delta b_j^{\text{obs}}}{\Delta t}\right)''_{\text{ev}}}, \quad 1 \leqslant i,j \leqslant n, \tag{6.3}$$

where for $1 \leqslant i \leqslant n$,

$$\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)''_{\text{ev}} = \left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'_{\text{ev}} - \overline{\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'_{\text{ev}}}, \tag{6.4}$$

$$\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)'_{\text{ev}} = \left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right) - \left(\frac{Re}{Re^{\text{ev}}}\, \left(\boldsymbol{b}^{\text{obs}}\right)^T \boldsymbol{l}_{\bullet i} + \left(\boldsymbol{b}^{\text{obs}}\right)^T \boldsymbol{c}_{\bullet\bullet i}\, \boldsymbol{b}^{\text{obs}}\right). \tag{6.5}$$

As for the LU ROM, the implicit assumption of white noise residual $\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)''_{\text{ev}}$ may not be valid. Thus, similarly to section 4.5, an optimal down-sampling time step is estimated with $\sum_{k=1}^{n} \left(\frac{\Delta b_k^{\text{obs}}}{\Delta t}\right)''_{\text{ev}}(t_i) \left(\frac{\Delta b_k^{\text{obs}}}{\Delta t}\right)''_{\text{ev}}(t_j)$ instead of $C^{v'}_{ij}$, before the noise covariance estimation.

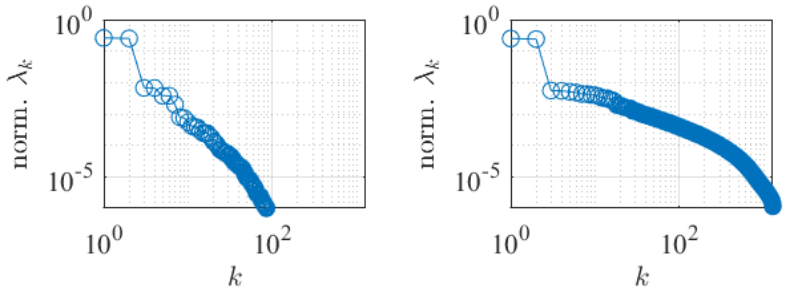

Figure 1: Energies of the reduced solution coefficients normalized by the total solution energy temporal mean $\overline{\|\boldsymbol{v}\|^2_{\boldsymbol{L}^2}}$ for the 2D flow at Reynolds 100 (left panel) and the 3D flow at Reynolds 300 (right panel).

6.2. 2D wake flow at Reynolds 100. Starting with the simpler of the two test cases, 800 seconds of DNS simulation data (about 160 pseudo-periods) were generated for a 2D wake flow behind a circular cylinder at $Re = 100$. These data were then split into training and test sets, the former (140 pseudo-periods) being utilized for the construction of the ROMs, and the latter (20 pseudo-periods), as a reference to which the ROMs will be compared. Performance will thus be measured through four different metrics: comparison of temporal coefficients in the time domain, comparison of reconstructed velocity fields in the spatial domain, global prediction accuracy as measured by the evolution in time of RMSE, and finally ensemble minimum error.

6.2.1. Temporal coefficients forecast. The 2-dimensional LU ROM of the Reynolds-100 flow is simulated 100 times. In figure 2, the coefficients of the reduced solution $b_i^{(k)}$ (where k stands for the k-th simulation) are compared to the coefficients of the reference solution $b_i^{\text{ref}} = (\boldsymbol{\phi}_i, \boldsymbol{v}_{\text{ref}})$ (plotted in black). The ensemble mean (in green) follows almost perfectly the reference's phase, and exhibits a slightly damped magnitude. The damping effect can be seen as a consequence of the exchange of energy between the mean and the variance (see equation (5.3)). This interplay is also evidenced by the growth of the confidence interval. The realization we have (randomly) singled-out shows a slowly divergent behavior.

6.2.2. Velocity forecasts. Having analyzed the model's forecast capabilities in terms of the ROM's temporal coefficient, we now focus on the analysis of the velocity fields' forecast and compare them to the reference – i.e. the projection of full DNS solution onto the POD basis : $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}] \overset{\triangle}{=} \Pi_\phi[\boldsymbol{v}_{\text{ref}} - \overline{\boldsymbol{v}}] + \overline{\boldsymbol{v}} = \sum_{i=0}^{n} b_i^{\text{ref}} \boldsymbol{\phi}_i$ (with $b_0^{\text{ref}} = 1$). Here, our goal will be to assess qualitatively the potential limitations of the ROM. Through equation (3.1), we will compute the velocity fields for the mean, the lowest error realization, and the prediction generated by the eddy viscosity ROM (6.1) of same dimension. At each time step, we define the lowest-error realization as follows:

$$\boldsymbol{b}^{\min} = \underset{(\boldsymbol{b}^{(k)})_k}{\operatorname{argmin}} \left\| \boldsymbol{w}^R - \widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}] \right\|^2_{\boldsymbol{L}^2} = \underset{(\boldsymbol{b}^{(k)})_k}{\operatorname{argmin}} \|\boldsymbol{b}^{(k)} - \boldsymbol{b}^{\text{ref}}\|^2. \tag{6.6}$$

We plot in figure 3 the vorticity fields of the ROMs of dimension $n = 8$. A system of dimension 8 enables us to reproduce fairly well the topology of the velocity fields, and from figure 2 we conclude that our model's best realization is capable of staying in phase with the reference, even 50 seconds into the validation set, while the ensemble realization's mean and the deterministic baseline start losing accuracy. Indeed, the eddies of the latter predictions

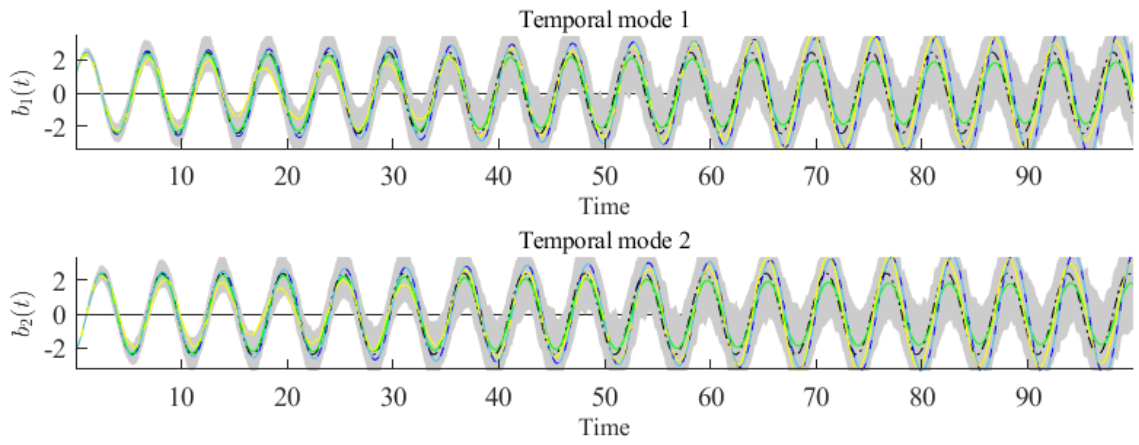

Figure 2: LU POD-ROM forecast for $n = 2$ coefficients of the reduced solution for a 2D wake flow at Reynolds 100: ensemble mean (green line), one random realization (yellow line), confidence interval (gray shade). Blue lines correspond to deterministic ROMs: baseline eddy-viscosity POD-ROM (6.1) (light-blue dashed line) and DNS POD-Galerkin (dark-blue solid line). The dash-dot black plots represent the observed references.

are late compared to the reference. This misalignment of eddies is a major type of model error in CFD, and it is very problematic for data assimilation because misalignment can hardly be corrected by linear Eulerian algorithms [64]. The LU model's cornerstone is precisely the random locations of transported quantities, and in particular, the locations of eddies. This random location idea is encoded in the random flow equation (1.2) and gave rise to the name "Location Uncertainty" (LU). Here, the LU POD-ROM generates an ensemble of simulations. Some of them have eddies in advance compared to the mean and some have late eddies. This variety of eddies' locations allows some ensemble realizations – in particular the best one – to match the reference's eddies' locations.

6.2.3. Global prediction accuracy. To compare more precisely the performance of the proposed ROM, we systematically plot the RMSE, as well as the ensemble's bias, variance and minimum error. These quantities will also be compared to those obtained by the stochastic baseline ROM (6.2). This enables us to inspect the UQ capabilities of the proposed ROM.

Although the random energy transfers are a necessary feature when the number of modes is not large enough, it increases the temporal coefficients' variance while keeping their biases constant, which yields an increase of the Root Mean Square Error (RMSE) of the whole ensemble:

$$\text{RMSE} \triangleq \hat{\mathbb{E}} \left\| \boldsymbol{w}^R - \widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}] \right\|^2_{\boldsymbol{L}^2} = \left\| \hat{\mathbb{E}}\{\boldsymbol{w}^R\} - \widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}] \right\|^2_{\boldsymbol{L}^2} + \int_\Omega \widehat{\text{Var}}(\boldsymbol{w}^R). \tag{6.7}$$

In all the following normalized error plots, the blue line corresponds to the solution computed using a deterministic POD-Galerkin ROMs (with and without eddy viscosity) and the red one, to the RMSE. The green line represents the ensemble bias, whereas the magenta line is the error incurred by the ensemble's solution closest to the reference (6.6), and the gray shade corresponds to 1.96 $\times$ the standard-deviation. These quantities are computed for the whole set of generated realizations. The initial condition is common to all and the values are normalized by the square root of the solution's energy averaged over the training set : $\left(\|\overline{\boldsymbol{v}}\|^2_{\boldsymbol{L}^2} + \sum_{i=1}^n \lambda_i \right)^{1/2}$.

Reference simulation $\boldsymbol{v}_{\text{ref}}$ (2D DNS at Reynolds 100: state space dimension of about 10^4)

Projection of the reference onto the POD basis $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}]$ (State space of dimension 8)

LU POD-ROM model best prediction (ROM state space of dimension 8)

LU POD-ROM model ensemble mean (ROM state space of dimension 8)

Prediction of the deterministic baseline POD-ROM (6.1) (fitted eddy viscosity) (ROM state space of dimension 8)

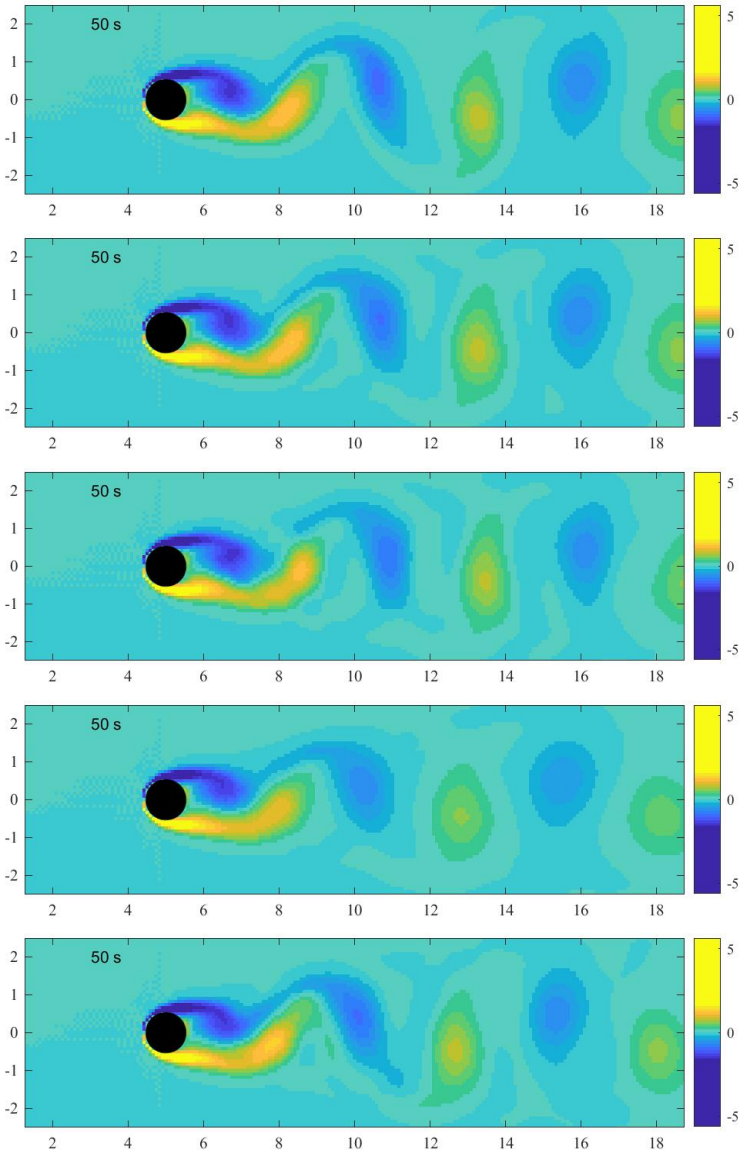

Figure 3: Vorticity fields – 10 vortex shedding cycles after the learning period – from (from top to bottom) the reference simulation $\boldsymbol{v}_{\text{ref}}$ (2D DNS at Reynolds 100: state space dimension of about 10^4), its projection onto the POD basis $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}]$, the LU POD-ROM model's best prediction, the LU POD-ROM model's ensemble mean, and the prediction of the deterministic baseline POD-ROM (6.1) (fitted eddy viscosity) (ROM state spaces are of dimension 8). The vorticity is the velocity curl and is the usual and a convenient way of visualizing 2D flows and their vortices.

In figure 4, the aforementioned error curves generated by the LU POD-ROM are compared to the stochastic POD with eddy viscosity model (6.2) for $n \in \{2, 4, 8\}$. The LU POD-ROM's best realization incurs a rather low error, and for the first two cases, a low bias too. The same cannot be said for the stochastic eddy viscosity model, which sees its variance grow out of control at the very beginning of the simulation, making the model's predictions diverge almost instantly. This divergence can be explained by the very large additive noise variance, fitted on the large ROM residuals $\left(\frac{\Delta b_i^{\text{obs}}}{\Delta t}\right)''_{\text{ev}}$ through equation

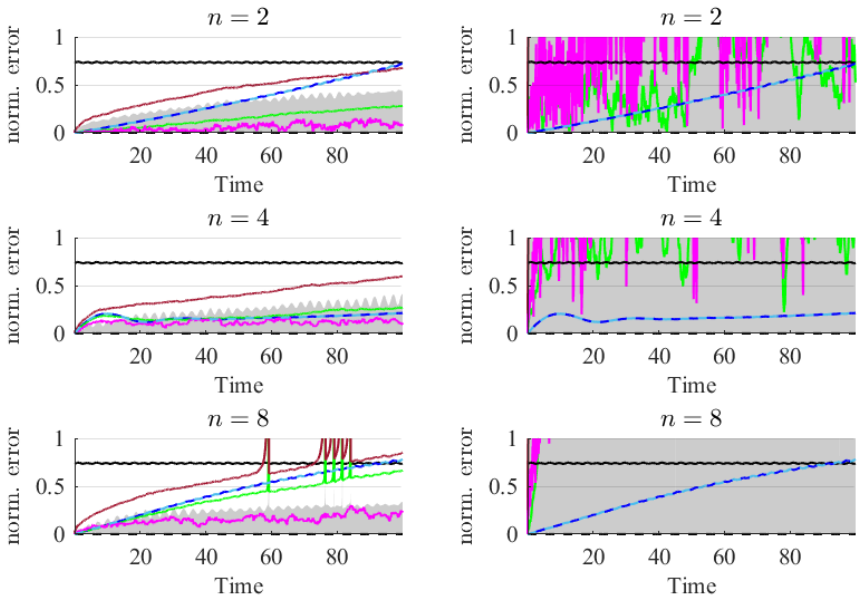

Figure 4: Normalized error for $n = 2$, 4 and 8 coefficients of the reduced solution of a 2D wake flow at Reynolds 100 – with the projection of the DNS onto the POD basis $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}]$ as reference : RMSE (brown line), bias (green line), ensemble minimum error (magenta line) and 1.96 $\times$ the standard-deviation (shaded gray) for the LU POD-ROM (left panel) and stochastic baseline POD-ROM (right panel) . Blue lines correspond to deterministic ROMs : baseline eddy-viscosity POD-ROM (6.1) (light-blue dashed line) and DNS POD-Galerkin (dark-blue solid line). The black solid line at the top is the error considering only the time mean velocity, i.e. $b_i = 0, \forall i > 0$.

(6.3). This large additive noise hence quickly fills the reduced subspace overwhelming the physical terms of the reduced dynamics. As for the $n = 8$ simulation, we observe that the LU POD-ROM differs from the other two simulations by having unstable realizations, thence the sudden peaks in the bias and mean curves. When a realisation diverges, we re-sample it uniformly from the other members of the ensemble.

Focusing on our stochastic POD-ROM, for any $n \in \{2, 4, 8\}$, even though the ensemble bias grows as time passes by, the best realization error does not. This is a crucial property for a UQ model to have, as this means that an appropriate filtering technique could eventually retrieve it at each observation, leading to a stable, low error data assimilation system.

6.3. 3D wake flow at Reynolds 300. For a more challenging test, a 3D wake flow behind a circular cylinder at $Re = 300$ was simulated, constructing the ROM just like in the case of the wake flow at Reynolds 100. The 440 seconds of DNS simulation data (about 88 vortex shedding pseudo-periods) were split into training (80 vortex shedding pseudo-periods) and test sets (8 pseudo-periods). Similarly, the ROM was subjected to the same benchmarks as before, with the equivalent eddy-viscosity UQ-enabled model for comparison. As it is a more complex flow, we expect the different temporal coefficients to be harmonically richer and to capture as much of the small-scale interactions as possible, but yet not to be able to fully reproduce the intricacies of the DNS-simulated flow.

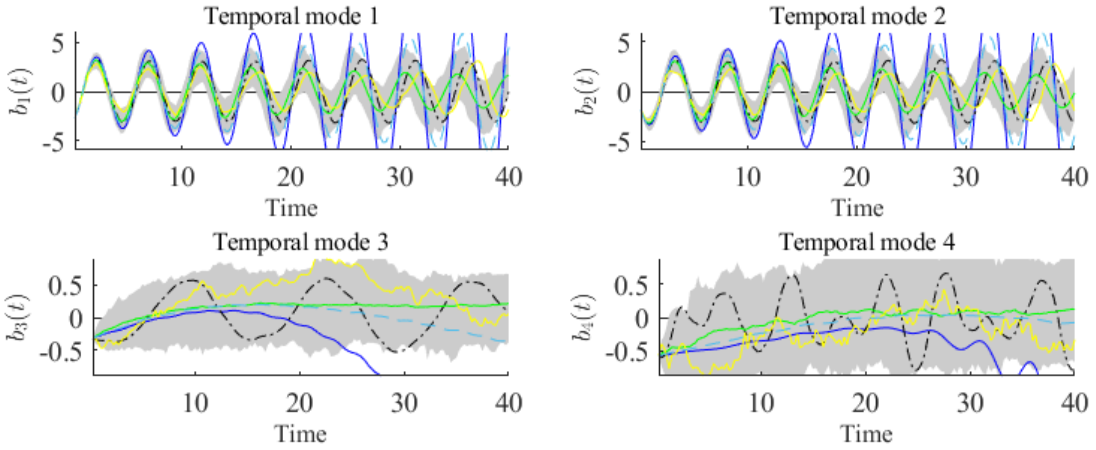

Figure 5: LU POD-ROM forecast for $n = 4$ coefficients of the reduced solution of a 3D wake flow at Reynolds 300: ensemble mean (green line), one random realization (yellow line), confidence interval (gray shade). Blue lines correspond to deterministic ROMs: baseline eddy-viscosity POD-ROM (6.1) (light-blue dashed line) and DNS POD-Galerkin (dark-blue solid line). The dash-dot black plots represent the observed references.

6.3.1. Temporal coefficients forecast. This new scenario being fully 3D and more complex, the advantages of the proposed model are clear. This can be seen from the temporal coefficient forecasts, where the proposed LU stochastic ROM manages to stabilize the system while the deterministic PODs starts diverging after 5 seconds (i.e. before one complete vortex shedding), as evidenced by the first temporal coefficients in figure 5. Intermittency probably fools the learning procedure of the eddy viscosity method and makes it less robust and not adapted to the test set. In contrast, the LU learning procedure – section 4 – shows a greater degree of robustness and leads to more accurate results. It is also interesting to note how the random energy contribution increases with the order of the temporal coefficient in the form of ensemble variance. This is attributed to energy transfers between temporal coefficients, facilitated by the interplay of multiplicative noise and turbulent diffusion. This effect is in fact amplified in the 8-dimension ROM in figure 6, where the temporal coefficients' means get more damped as the amount of random energy in the system increases. This energy dissipation mechanism helps the system to remain bounded, unlike its deterministic counterparts.

Interestingly enough, in the $n = 8$ dimension ROM, even when the ensemble variance increases with the temporal coefficient order, coefficients 5, 6, 7 and 8 of the random realization (yellow line) also attempt to follow the right reference amplitude and variability, as it can be observed in figure 6.

6.3.2. Velocity forecasts. As in the previous case, we now analyze qualitatively the performance of our proposed method in terms of velocity field predictions.

As can be observed from figure 7, the best prediction remains stable and manages to stay close to the theoretically optimal solution in 8-dimensions on a period of four shedding cycles after the learning period. Conversely, the baseline model starts to diverge, as evidenced by the red zones inside the vortices. This result is quite impressive as, even if the mean velocity field is not as close to the reference as one would like, it still reveals the potential of a method combining our model with a data assimilation technique to retrieve the best realization and provide corrections on the fly.

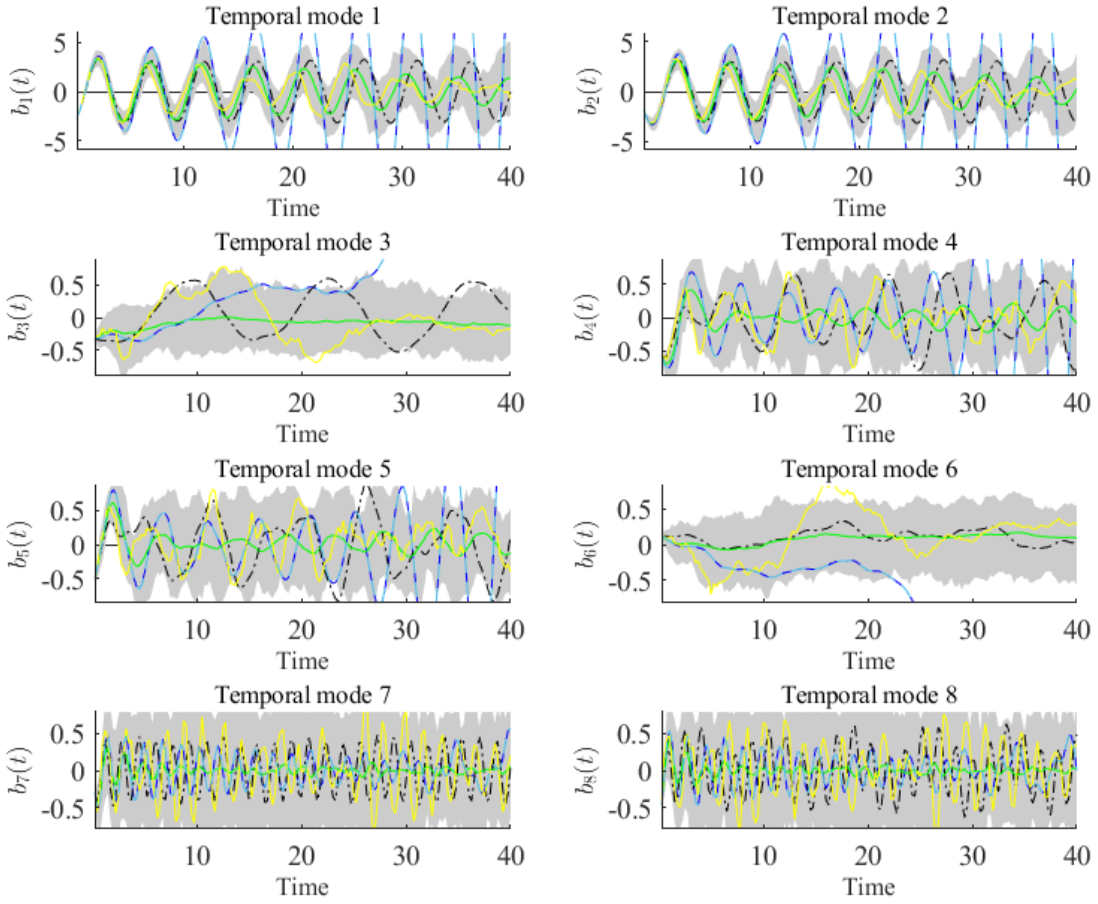

Figure 6: LU POD-ROM forecast for $n = 8$ coefficients of the reduced solution of a 3D wake flow at Reynolds 300: ensemble mean (green line), one random realization (yellow line), confidence interval (gray shade). Blue lines correspond to deterministic ROMs: baseline eddy-viscosity POD-ROM (6.1) (light-blue dashed line) and DNS POD-Galerkin (dark-blue solid line). The dash-dot black plots represent the observed references.

6.3.3. Global prediction accuracy. Just like in the two-dimensional flow, the error curves are plotted and compared to the results obtained with the stochastic eddy viscosity model (6.2). The curves in figure 8 evidence the predictive power of our proposed model: it is capable of great stability and accuracy even after 40 seconds while the deterministic and eddy viscosity models prove to be quite unstable by rapidly diverging. These properties are showcased by the bias and mean curves that converge to the zero-temporal-coefficient error curve instead of growing in an unbounded manner, while still having at least one realization with relatively low error that could eventually be identified through data assimilation techniques.

7. Conclusion. This paper proposed a stochastic ROM derived from a stochastic fluid dynamics modelling standpoint, called dynamics under location uncertainty (LU), which formulates unresolved small-scale parameterization through SPDEs. The stochastic ROM is obtained through a classical POD-Galerkin projection of these SPDEs, and the basis functions are defined from a high-resolution deterministic simulation of the target flow. The resulting model bears similarities with the Navier-Stokes equations, but also encompasses an advection velocity correction, a turbulent diffusion and a skew-symmetric multi-

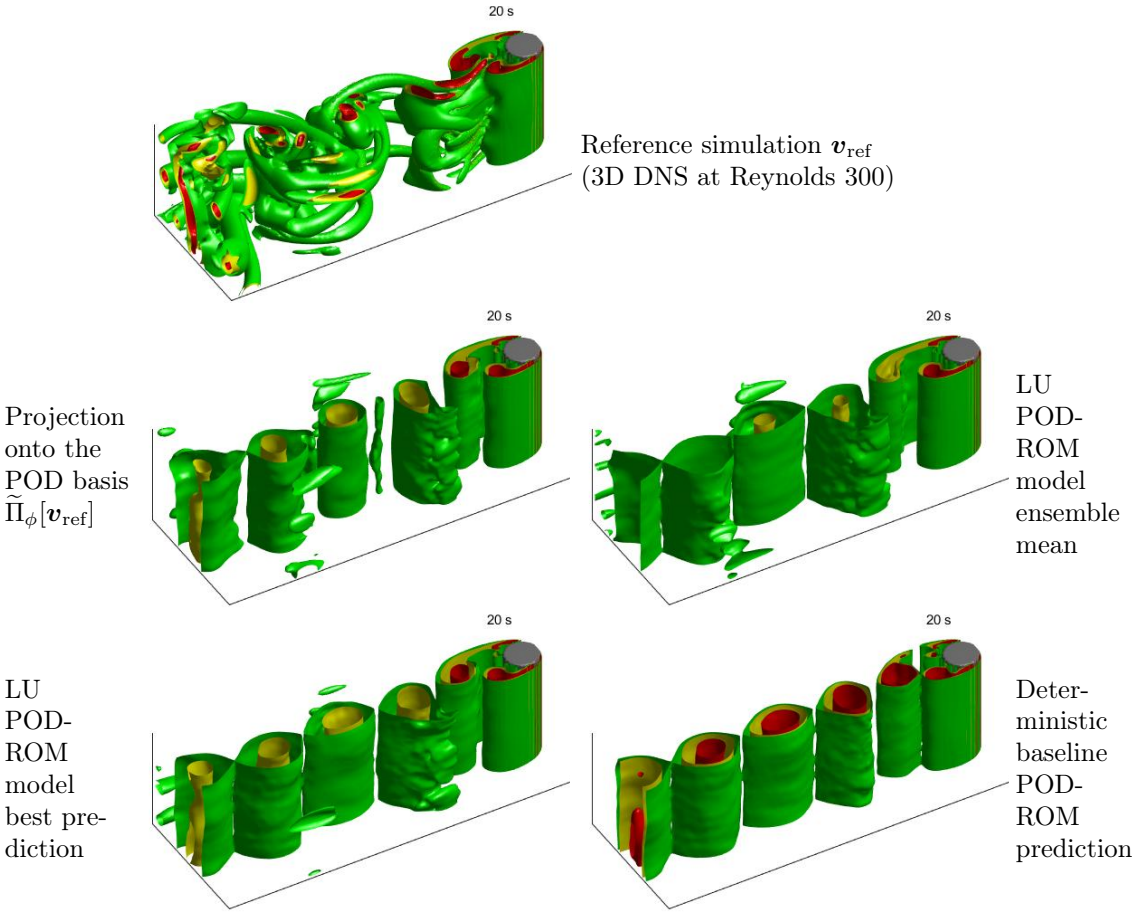

Figure 7: Q-criterion iso-surfaces – 4 vortex shedding cycles after the learning period – from (from left to right and from top to bottom) the reference simulation $\boldsymbol{v}_{\text{ref}}$ (3D DNS at Reynolds 300: state space dimension of about 10^7), its projection onto the POD basis $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}]$, the LU POD-ROM model's ensemble mean, the LU POD-ROM model's best prediction, and the prediction of the deterministic baseline POD-ROM (6.1) (fitted eddy viscosity) (POD-ROM state spaces are of dimension 8). The Q-criterion [39] is a quadratic function of the velocity gradient and is the usual and a convenient way for visualizing 3D flows and their vortices.

plicative noise terms. From the statistics of the POD's residuals – i.e. the data component orthogonal to the POD's modes – these new terms can be fully characterized.

The implementation necessitates additional off-line computations on top of the classical POD-Galerkin procedure: estimations of a variance tensor and of the reduced multiplicative noise's covariance matrix. The former, proportional to the point-wise $d \times d$ covariance matrix of POD's residual, is readily computed. The latter can constitute a formidable computational challenge. Moreover, with a fully data-driven method, the risk of over-fitting is very high. To circumvent these issues, an intermediate solution is proposed with an easy-to-compute estimator, along with consistency proofs under the LU setting assumptions. Since the latter may not always be met – specifically, the time decorrelation assumption of the POD residual – measures are taken to enforce them on the model, as part of its constitutive steps. Namely, to force the decorrelation – and in doing so, improve the accuracy of the estimators and the ROM as a whole – a down-sampling of the dataset

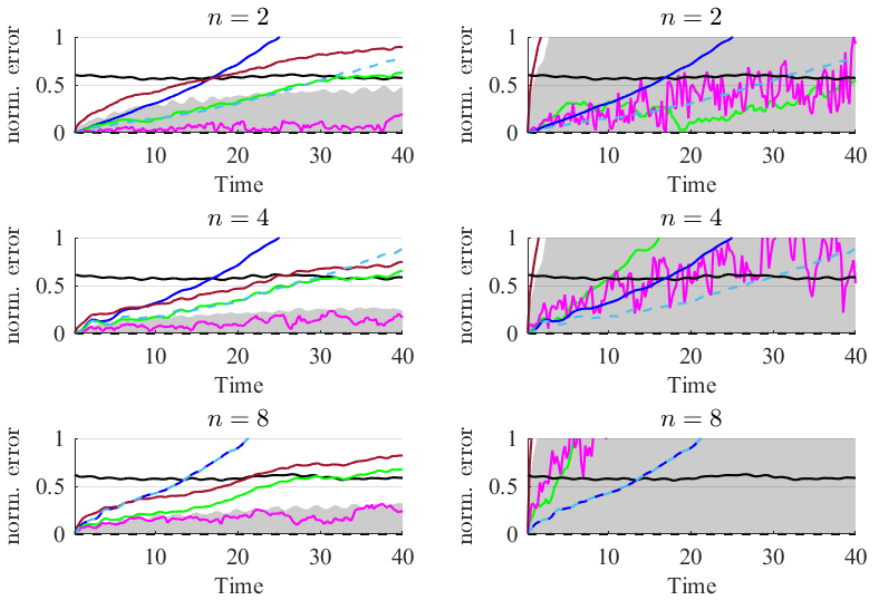

Figure 8: Normalized error for $n = 2$, 4 and 8 coefficients of the reduced solution of a 3D wake flow at Reynolds 300 – with the projection of the DNS onto the POD basis $\widetilde{\Pi}_\phi[\boldsymbol{v}_{\text{ref}}]$ as reference : RMSE (red line), bias (green line), ensemble minimum error (magenta line) and 1.96 $\times$ the standard-deviation (shaded gray) for the LU POD-ROM (left panel) and stochastic baseline POD-ROM (right panel) . Blue lines correspond to deterministic ROMs: baseline eddy-viscosity POD-ROM (6.1) (light-blue dashed line) and DNS POD-Galerkin (dark-blue solid line). The black solid line at the top is the error considering only the time mean velocity, i.e. $b_i = 0, \forall i > 0$.

is applied at a rate equal to the residual's correlation time. Finally, in order to restrict the number of coefficients needed to characterize the ROM to the usual $O(n^3)$, a technique to reduce the dimensionality of the noise covariance matrix is also proposed.

The conservative properties of the LU closure are discussed, and we demonstrate that, when performing Galerkin projections of the Itō form of LU SPDEs, we end up forfeiting these properties. In the reduced order version, two energy fluxes appear: the first one is attributed to the possible advecting velocity correction divergence, whilst the second one is negative and is directly associated with mode truncation. We argue that this energy loss is understandable, and even desirable, with an interplay between noise and turbulent diffusion to maintain non-linear energy fluxes between coefficients of the reduced solution despite mode truncation.

Numerical comparisons are performed between our stochastic ROM and state-of-the-art deterministic and stochastic ROMs. As test cases, we chose a two-dimensional wake flow at Reynolds 100 with few degrees of freedom, and a more complex three-dimensional wake flow at Reynolds 300 with many more degrees of freedom. Deterministic Navier-Stokes simulations stood for references and all of the ROMs were initialized with their values, and the forecasts, compared to them. For the stochastic ROMs, ensembles of 100 realizations were simulated. The state-of-the-art stochastic ROM's solutions quickly diverged in time, whilst the temporal coefficients of our ROM proved to be neither unstable nor over-damped. The LU POD-ROM solution's biases were even found to be smaller than those of each of the

other stochastic ROMs' that we have considered. Moreover, at any given moment, the LU POD-ROM ensembles managed to remain very close to the reference, suggesting that they could be efficient priors for Bayesian inverse problems.

Following this work, we have applied the LU POD-ROM with unknown initial conditions in conjunction with particle filter algorithms to estimate velocity flows in real-time from few local measurements [63].

Finally, when working at very large Reynolds numbers, DNS simulations are no longer an option, but LES (Large Eddy Simulation), DDES (Delayed Detached Eddy Simulation) and RANS (Reynolds Averaged Navier Stokes) can still provide useful data to build ROMs, at the cost of some approximation error. To tackle this, the closure mechanism's estimation procedure may be adjusted to address the small-scale velocity statistics neglected by LES-like approaches. Another approach that we are currently exploring is to consider the Galerkin projection on a LU version of LES and DDES. Non-polynomial terms of LES and DDES complicate the dimensional reduction, but discrete empirical interpolation methods (DEIM) [18] may be able to handle it. As illustrated in this paper, our stochastic closure naturally prevents truncation-induced instabilities, and we expect a similar behavior at higher Reynolds. If the proposed in-house stabilization is not sufficient, existing stabilisation and/or data-driven methods [33, 37, 77] could be added. Furthermore, CFD outputs may also be at a resolution different from the resolution of the measurements. Hence, depending on the relative precision of CFD and measurements, there could be an additional or a missing very-small-scale residual turbulence in the assimilated measures. Accordingly, we will need a simple statistical model for this very-small-scale turbulence and include its effects in the observation model through either an additional "measurement noise" or additional smoothing.

Besides, in realistic applications, Reynolds number, initial and boundary conditions, as well as a plethora of other parameters, are often only approximately known. So, combinations with other state-of-the-art dimensionality reduction methods are most likely necessary. Typically, we could use wRB [20, 21, 77] for uncertain mean inflow (Reynolds number and angle of attack) and a reduced pressure equation [75, 76] for time dependent boundary conditions. Our new implementation of LU terms in the ITHACA-FV library [75, 76] will surely aid on these practical aspects.

Nevertheless, note that a wRB method to simulate random velocity effect [77] does not directly apply to the purpose of the current paper. Indeed, we try to quantify the uncertainty induced by the dimensional reduction of the deterministic Navier-Stokes equations and not to reduce the dimension of an uncertainty quantification problem. The random residual velocity can hardly be considered a wRB parameter because we do not choose its value. There is only one "realization" and this is the one given by the deterministic Navier-Stokes equations. We could specify some statistics for the random residual velocity and simulate the stochastic Navier-Stokes equations (2.1)-(2.2) but this would make our method much more complex, much more CPU demanding, and could possibly lead to over-fitting. Moreover, that random velocity is a residual, and thus, its statistics depend on the reduced dimension n. That would also greatly complicate the wRB algorithm.

Finally, relying purely on data can offer the advantage of directly tackling very high Reynolds number configurations. It would require us to filter noisy measurements, and most likely, to rely on simplified (often 2D) data-driven models [1, 13, 29, 78]. From this latter point of view, modeling under location location offers a great flexibility which might be of great interest. As a matter of fact, the LU formalism enables us to derive random dynamics close to the data but still based on stochastic transport principles [3].

The noise parameters can be updated online from incoming data and simulations. A full characterization of the noise dynamics through statistical learning procedures would then probably enable us to devise advanced data-driven models that are able to handle even more complex situations.

Appendix A. Estimation formulas. In this appendix, we consider a full probability space $(A, \mathcal{F}, P)$, and a filtration of σ-algebra $\{\mathcal{F}_t\}_{t\geq 0}$. For the Hilbert space $H^2(\Omega)$, we understand $\mathcal{H}^2 = \mathcal{H}^2(0,T;\Omega,\mathcal{F})$ as the space of strongly measurable, $\{\mathcal{F}_t\}_{t\geq 0}$ adapted processes $\boldsymbol{u} : \mathbb{R}^d \times [0,T] \to \mathbb{R}^d$ such that $(\boldsymbol{x} \mapsto \boldsymbol{u}(\boldsymbol{x},t))$ belong to $H^2(\Omega)$ for every $t \in [0,T]$ and

$$\|\boldsymbol{u}\|^2_{\mathcal{H}^2} = \mathbb{E}\int_0^T \|\boldsymbol{u}(\bullet,t)\|^2_{H^2(\Omega)}\,\mathrm{d}t < \infty. \tag{A.1}$$

Similarly, we can define $\mathcal{L}^2(0,T;\Omega,\mathcal{F})$ from $\boldsymbol{L}^2$ and its norm. We recall below a classical proposition related to quadratic variation process:

Proposition: Stochastic integration and quadratic variations. If $\boldsymbol{M}$ is a continuous martingale and $\boldsymbol{X} \in \mathcal{L}^2(0,T;\Omega,\mathcal{F})$, then there exists a unique bounded continuous martingale $\int_0^t \boldsymbol{X}\mathrm{d}\boldsymbol{M}$ such that for every continuous martingale $\boldsymbol{N}$ (with zero initial condition)

$$\left\langle \int_0^t \boldsymbol{X}\mathrm{d}\boldsymbol{M}, \boldsymbol{N}\right\rangle = \int_0^t \boldsymbol{X}\mathrm{d}\langle \boldsymbol{M},\boldsymbol{N}\rangle. \tag{A.2}$$

As stated in the introduction, we assume that the operator $\boldsymbol{\sigma}$ is Hilbert-Schmidt and that its kernel $\breve{\boldsymbol{\sigma}} : \overline{\Omega}^2 \to \mathbb{R}^{d\times d}$ is C^2 over the compact bounded set $\overline{\Omega}^2$. Therefore,

$$\|\boldsymbol{\sigma}\boldsymbol{B}\|^2_{\mathcal{H}^2} \leqslant \max_{i,j}\sup_{\overline{\Omega}^2}\left(\|\breve{\boldsymbol{\sigma}}\|^2 + \|\partial_{x_i}\breve{\boldsymbol{\sigma}}\|^2 + \left\|\partial_{x_i}\partial_{x_j}\breve{\boldsymbol{\sigma}}\right\|^2\right)\left|\Omega^2\right| T < \infty, \tag{A.3}$$

where $\|\bullet\|$ stands for the Euclidean norm on $\mathbb{R}^{d\times d}$ and $|\Omega^2|$ the Lebesgue measure of Ω^2. As such, $((\boldsymbol{x},t) \mapsto \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{B}_t)$ belong to $\mathcal{H}^2$. In particular, $(\boldsymbol{x} \mapsto \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{B}_t)$ belong to $H^2(\Omega)$ for every $t \in [0,T]$ and $\partial_{x_i}\boldsymbol{\sigma}, \partial_{x_i}\partial_{x_j}\boldsymbol{\sigma}$ are also Hilbert-Schmidt for every $1 \leqslant i,j \leqslant d$. Additionally, we assume that the POD modes $\boldsymbol{\phi}_i$ belong to $H^2(\Omega)$ and have bounded gradients.

After defining boundary conditions, $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{K}$ are well defined on $H^1(\Omega)$ and $H^2(\Omega)$ respectively. Then, for every $\boldsymbol{\zeta} \in H^1(\Omega)$, $\|\boldsymbol{\mathcal{P}}\boldsymbol{\zeta}\|_{\boldsymbol{L}^2} \leqslant \|\boldsymbol{\zeta}\|_{\boldsymbol{L}^2}$ since $\boldsymbol{\mathcal{P}}$ is a projection and then, for every $\boldsymbol{\xi} \in H^2(\Omega)$, by Cauchy-Schwarz and triangular inequalities:

$$|K_{jq}(\boldsymbol{\xi})| \leqslant \|\boldsymbol{\phi}_j\|_{\boldsymbol{L}^2}(\|\boldsymbol{\nabla}\boldsymbol{\phi}_q^T\|_\infty + \nu)\|\boldsymbol{\xi}\|_{H^2(\Omega)}, \tag{A.4}$$

where $\|\boldsymbol{\nabla}\boldsymbol{\phi}_i^T\|_\infty = \sup_{\Omega}\|\boldsymbol{\nabla}\boldsymbol{\phi}_i^T\| < \infty$. The bound above gives the continuity of $\boldsymbol{K}$ on $H^2(\Omega)$. We also assume that the observed coefficients of the reduced solution $b_i^{\mathrm{obs}} = (\boldsymbol{\phi}_i, \boldsymbol{v}_{\mathrm{obs}})$ are continuous semi-martingales and solutions of the ROM (3.3)-(3.4). From there, the orthogonality of the coefficients of the reduced solution yields, for $1 \leqslant i,j \leqslant n$ and $0 \leqslant p,q \leqslant n$:

$$\int_0^T b_p \mathrm{d}\Big\langle b_i, \int_0^t (\tilde{\boldsymbol{\alpha}}_{qj\bullet}\mathrm{d}\boldsymbol{B}_s)\Big\rangle = \int_0^T \sum_{k=0}^n b_p \mathrm{d}\Big\langle \int_0^t (\tilde{\boldsymbol{\alpha}}_{ki\bullet}\mathrm{d}\boldsymbol{B}_s)\, b_k, \int_0^t \tilde{\boldsymbol{\alpha}}_{qj\bullet}\mathrm{d}\boldsymbol{B}_s\Big\rangle, \tag{A.5}$$

$$= \sum_{k=0}^n \left(\int_0^T b_p b_k\right)\Sigma^{\boldsymbol{\alpha}}_{ki,qj}, \tag{A.6}$$

$$= T\lambda_p\, \Sigma^{\boldsymbol{\alpha}}_{pi,qj}. \tag{A.7}$$

Now, let us note that for $1 \leqslant j \leqslant n$ and $0 \leqslant q \leqslant n$:

$$\tilde{\boldsymbol{\alpha}}_{qj\bullet}\mathrm{d}\boldsymbol{B}_s = \left(\boldsymbol{\phi}_j, -\boldsymbol{\mathcal{P}}\left[(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s \cdot \boldsymbol{\nabla})\,\boldsymbol{\phi}_q\right] + \delta_{q0}\,\nu\Delta\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s\right) = \mathrm{d}\left(K_{jq}\left[\boldsymbol{\sigma}\boldsymbol{B}\right]\right). \tag{A.8}$$

Then, with the definition of the quadratic covariation and the increment notation $\Delta\xi(t_k) = \xi(t_{k+1}) - \xi(t_k)$, we obtain the estimator's expression and its consistency for every $1 \leqslant i, j \leqslant n$ and $0 \leqslant p, q \leqslant n$:

$$\Sigma^{\boldsymbol{\alpha}}_{pi,qj} = \frac{1}{\lambda_p T}\int_0^T b_p \mathrm{d}\langle b_i, \int_0^t (\tilde{\boldsymbol{\alpha}}_{qj\bullet}\mathrm{d}\boldsymbol{B}_s)\rangle, \tag{A.9}$$

$$= \frac{1}{\lambda_p T}\int_0^T b_p \mathrm{d}\langle b_i, K_{jq}\left[\boldsymbol{\sigma}\boldsymbol{B}\right]\rangle, \tag{A.10}$$

$$= \frac{1}{\lambda_p T}\,\mathbb{P}\text{-}\lim_{\Delta t\to 0}\sum_{t_k=0}^{T} b_p(t_k)(\Delta b_i)(t_k) K_{jq}\left[\boldsymbol{\sigma}\Delta\boldsymbol{B}_{t_k}\right], \tag{A.11}$$

$$= K_{jp}\left[\frac{1}{\lambda_p T}\,\mathbb{P}\text{-}\lim_{\Delta t\to 0}\sum_{t_k=0}^{T} b_p(t_k)(\Delta b_i)(t_k)\,\boldsymbol{\sigma}\Delta\boldsymbol{B}_{t_k}\right], \tag{A.12}$$

where the continuity of the operator K_{jp} on $H^2(\Omega)$ enabled us to switch the limit in probability and the operator K_{jp}. The martingale flow increments $\boldsymbol{\sigma}\Delta\boldsymbol{B}_{t_k}$ are approximated by $\boldsymbol{v}'(\bullet, t_k)\Delta t$. In practice, for each $1 \leqslant i \leqslant n$ we can replace Δb_i by

$$\Delta b_i'' = \Delta b_i' - \overline{\Delta b_i'}, \tag{A.13}$$

where

$$\Delta b_i' = \Delta b_i - \left(\boldsymbol{b}^T(\boldsymbol{l} + \breve{\boldsymbol{f}})_{\bullet i} + \boldsymbol{b}^T\boldsymbol{c}_{\bullet\bullet i}\,\boldsymbol{b}\right)\Delta t. \tag{A.14}$$

Mathematically, this is still correct for very large values of T. Indeed, $\overline{\Delta b_i'} = \frac{b_i'(T+\Delta T) - b_i'(0)}{T}\Delta t \underset{T\to\infty}{\longrightarrow} 0$ and $b_i' - b_i$ has finite variations. Thus, $b_i'' - b_i$ approaches a finite variation process (for large T). Numerically, this formulation improves the accuracy of the model as it allows us to remove the smooth-in-time part of Δb_i, thus minimizing estimation error as well.

Appendix B. Energy dissipation.

On top of assumptions of Appendix A, we neglect the viscosity ($\boldsymbol{L} = 0$) and assume zero Dirichlet boundary conditions for the reduced basis' functions and for $\boldsymbol{x} \mapsto \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{B}_t$ for every $t \in [0, T]$. Note that these fields are divergence-free (since they are learned from a set of incompressible velocity fields). As such, we have: $\forall i \leqslant n$, $\|\boldsymbol{G}(\boldsymbol{\phi}_i)\|^2_{\mathrm{HS}} \leqslant \|\boldsymbol{\nabla}\boldsymbol{\phi}_i^T\|^2_\infty \|\boldsymbol{\sigma}\|^2_{\mathrm{HS}} < \infty$, i.e. $\boldsymbol{G}(\boldsymbol{\phi}_i)$ is Hilbert-Schmidt. This makes $\boldsymbol{G}(\boldsymbol{w}^R)$ and $\Pi_\phi\left[\boldsymbol{G}(\boldsymbol{w}^R)\right]$ also Hilbert-Schmidt.

Using the projected Navier-Stokes model (5.4), for every $t \in [0, T]$ we can formally remove the orthogonal projection by moving it into the divergence-free functions space $\boldsymbol{\mathcal{P}}$ through integration by parts:

$$\mathrm{d}_t\boldsymbol{w}^R = \sum_{i=1}^n \left(\boldsymbol{\phi}_i, \boldsymbol{\mathcal{P}}(\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right)\,\boldsymbol{\phi}_i = \sum_{i=1}^n \left(\boldsymbol{\mathcal{P}}\boldsymbol{\phi}_i, (\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right)\,\boldsymbol{\phi}_i, \tag{B.1}$$

$$= \sum_{i=1}^n \left(\boldsymbol{\phi}_i, (\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right)\,\boldsymbol{\phi}_i = \Pi_\phi\left[\left(\boldsymbol{C}(\boldsymbol{w}^R, \boldsymbol{w}^R) + \boldsymbol{F}(\boldsymbol{w}^R)\right)\mathrm{d}t + (\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w}^R)\right]. \tag{B.2}$$

Then, upon applying the Itō formula to the local kinetic energy we obtain

$$\mathrm{d}\left(\tfrac{1}{2}\|\boldsymbol{w}^R\|^2_{\boldsymbol{L}^2}\right) = \int_\Omega \left(\mathrm{d}_t(\boldsymbol{w}^R)^T\boldsymbol{w}^R + \tfrac{1}{2}\mathrm{d}_t\langle(\boldsymbol{w}^R)^T,\boldsymbol{w}^R\rangle\right) \qquad \forall t\in[0,T]. \tag{B.3}$$

To remove the (orthogonal) projection operator Π_ϕ from the first term, we exploit its symmetry, and afterwards, the fact that $\boldsymbol{w}^R$ is already in the reduced subspace:

$$\begin{aligned}
\int_\Omega \mathrm{d}_t(\boldsymbol{w}^R)^T\boldsymbol{w}^R &= \left(\Pi_\phi\left[(\mathrm{d}\mathbb{M})(\boldsymbol{w}^R)\right],\boldsymbol{w}^R\right), && \text{(B.4)}\\
&= \left((\mathrm{d}\mathbb{M})(\boldsymbol{w}^R),\Pi_\phi\left[\boldsymbol{w}^R\right]\right), && \text{(B.5)}\\
&= \left((\mathrm{d}\mathbb{M})(\boldsymbol{w}^R),\boldsymbol{w}^R\right), && \text{(B.6)}\\
&= \left(((\boldsymbol{F}_{\mathrm{dif}}+\boldsymbol{F}_{\mathrm{adv}})(\boldsymbol{w}^R)+\boldsymbol{C}(\boldsymbol{w}^R,\boldsymbol{w}^R))\mathrm{d}t+(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w}^R),\boldsymbol{w}^R\right), && \text{(B.7)}\\
&= \underbrace{\left(-\tfrac{1}{2}\boldsymbol{G}^*\left(\circledast\boldsymbol{G}(\boldsymbol{w}^R)\right),\boldsymbol{w}^R\right)\mathrm{d}t}_{\text{from } (\boldsymbol{F}_{\mathrm{dif}}(\boldsymbol{w}^R),w^R) \text{ using (2.12)}} + \underbrace{\left(\tfrac{1}{2}((\boldsymbol{\nabla}\cdot\boldsymbol{a})\boldsymbol{\nabla})\boldsymbol{w}^R,\boldsymbol{w}^R\right)}_{\text{from } (\boldsymbol{F}_{\mathrm{adv}}(\boldsymbol{w}^R),w^R)}\\
&\quad + \underbrace{\left(\boldsymbol{C}(\boldsymbol{w}^R,\boldsymbol{w}^R)\mathrm{d}t+(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w}^R),\boldsymbol{w}^R\right)}_{=0 \text{ by skew-symmetry of } \boldsymbol{\xi}\mapsto\boldsymbol{C}(\boldsymbol{w}^R,\boldsymbol{\xi}) \text{ and } \boldsymbol{G}\mathrm{d}\boldsymbol{B}_t}, && \text{(B.8)}\\
&= \tfrac{1}{2}\left\|\boldsymbol{G}(\boldsymbol{w}^R)\right\|^2_{\mathrm{HS}} - \tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T,\|\boldsymbol{w}^R\|^2\right), && \text{(B.9)}
\end{aligned}$$

where the second term comes from integration by parts. Besides, the Itō term of the energy budget is straightforward to compute from equation (B.2):

$$\begin{aligned}
\int_\Omega \tfrac{1}{2}\mathrm{d}_t\langle(\boldsymbol{w}^R)^T,\boldsymbol{w}^R\rangle &= \tfrac{1}{2}\left\|\Pi_\phi\left[(\boldsymbol{G})(\boldsymbol{w}^R)\right]\right\|^2_{\mathrm{HS}}\mathrm{d}t, && \text{(B.10)}\\
&= \sum_{p,q=0}^{n}\left(\tfrac{1}{2}\int_\Omega \Pi_\phi\left[\boldsymbol{G}(\boldsymbol{\phi}_p)\right]\circledast\Pi_\phi\left[\boldsymbol{G}(\boldsymbol{\phi}_q)\right]\right)b_pb_q\mathrm{d}t. && \text{(B.11)}
\end{aligned}$$

From the definition of the projection operator Π_ϕ and using the extended notation $(\boldsymbol{\eta}\circledast\boldsymbol{\theta}) \triangleq \int_\Omega \breve{\boldsymbol{\eta}}(\boldsymbol{z})\,\breve{\boldsymbol{\theta}}^{\,T}(\boldsymbol{z})d\boldsymbol{z}$, we can express the above quadratic operator with the noise statistics as follows:

$$\int_\Omega \Pi_\phi\left[\boldsymbol{G}(\boldsymbol{\phi}_p)\right]\circledast\Pi_\phi\left[\boldsymbol{G}(\boldsymbol{\phi}_q)\right] = \sum_{i=1}^{n}\left(\boldsymbol{\phi}_i,\boldsymbol{G}(\boldsymbol{\phi}_p)\right)\circledast\left(\boldsymbol{\phi}_i,\boldsymbol{G}(\boldsymbol{\phi}_q)\right) = \sum_{i=1}^{n}\Sigma^{\boldsymbol{\alpha}}_{pi,qi}. \tag{B.12}$$

Finally, by orthogonality, the kinetic energy budget (B.3) simplifies to:

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\left(\tfrac{1}{2}\|\boldsymbol{w}^R\|^2_{\boldsymbol{L}^2}\right) &= \tfrac{1}{2}\left\|\boldsymbol{G}(\boldsymbol{w}^R)\right\|^2_{\mathrm{HS}} - \tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T,\|\boldsymbol{w}^R\|^2\right) - \tfrac{1}{2}\left\|\Pi_\phi\left[\boldsymbol{G}(\boldsymbol{w}^R)\right]\right\|^2_{\mathrm{HS}}, && \text{(B.13)}\\
&= -\tfrac{1}{2}\left\|\Pi^\perp_\phi\left[\boldsymbol{G}(\boldsymbol{w}^R)\right]\right\|^2_{\mathrm{HS}} - \tfrac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T,\|\boldsymbol{w}^R\|^2\right), && \text{(B.14)}
\end{aligned}$$

where $\Pi^\perp_\phi = \mathbb{I}_d - \Pi_\phi$ is the projector into the orthogonal complement of the reduced subspace.

Acknowledgments. We warmly thank Pranav Chandramouli for the generation of the three-dimensional wake flow data, Pierre-Louis Lee for his help in the generation of the two-dimensional wake flow data, and Reda Bouaida for his help in the computation of the ROM matrix $\breve{\boldsymbol{f}}$. The authors also acknowledge the support of the ERC EU project 856408-STUOD, the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project. Finally, we also thank Darryl D. Holm, Dan Crisan, Wei Pan and Igor Shevchenko for the insightful discussions.

References.

[1] G. Artana, A. Cammilleri, J. Carlier, and E. Mémin, *Strong and weak constraint variational assimilations for reduced order fluid flow modeling*, Journal of Computational Physics, 231 (2012), pp. 3264–3288.

[2] N. Aubry, P. Holmes, J. Lumley, and E. Stone, *The dynamics of coherent structures in the wall region of a turbulent boundary layer*, J. Fluid Mech., 192 (1988), pp. 115–173.

[3] C. Avenel, E. Mémin, and P. Pérez, *Stochastic level set dynamics to track closed curves through image data*, Journal of mathematical imaging and vision, 49 (2014), pp. 296–316.

[4] R. Azencott, A. Beri, A. Jain, and I. Timofeyev, *Sub-sampling and parametric estimation for multiscale dynamics*, Communications in Mathematical Sciences, 11 (2013), pp. 939–970.

[5] R. Azencott, A. Beri, and I. Timofeyev, *Adaptive sub-sampling for parametric estimation of Gaussian diffusions*, Journal of Statistical Physics, 139 (2010), pp. 1066–1089.

[6] W. Bauer, P. Chandramouli, B. Chapron, L. Li, and E. Mémin, *Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach*, Journal of Physical Oceanography, 50 (2020), pp. 983–1003.

[7] W. Bauer, P. Chandramouli, L. Li, and E. Mémin, *Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models*, Ocean Modelling, (2020).

[8] J. Berner, S.-Y. Ha, J. Hacker, A. Fournier, and C. Snyder, *Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations*, Monthly Weather Review, 139 (2011), pp. 1972–1995.

[9] S. Boyaval, C. Le Bris, T. Lelievre, Y. Maday, N. C. Nguyen, and A. T. Patera, *Reduced basis techniques for stochastic problems*, Archives of Computational methods in Engineering, 17 (2010), pp. 435–454.

[10] S. L. Brunton, J. L. Proctor, and J. N. Kutz, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proceedings of the national academy of sciences, 113 (2016), pp. 3932–3937.

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

[12] M. Buffoni, S. Camarri, A. Iollo, and M. V. Salvetti, *Low-dimensional modelling of a confined three-dimensional wake flow*, Journal of Fluid Mechanics, 569 (2006), pp. 141–150.

[13] A. Cammilleri, F. Gueniat, J. Carlier, L. Pastur, E. Mémin, F. Lusseyran, and G. Artana., *POD-spectral decomposition for fluid flow analysis and model reduction*, Theor. and Comp. Fluid Dyn., (2013).

[14] 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 (2011), pp. 155–181.

[15] W. Cazemier, R. Verstappen, and A. Veldman, *Proper orthogonal decomposition and low-dimensional models for driven cavity flows*, Physics of fluids, 10 (1998), pp. 1685–1699.

[16] P. Chandramouli, D. Heitz, S. Laizet, and E. Mémin, *Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty*, Computers & Fluids, 168 (2018), pp. 170–189.

[17] B. Chapron, P. Dérian, E. Mémin, and V. Resseguier, *Large-scale flows under*

location uncertainty: a consistent stochastic framework, Quarterly Journal of the Royal Meteorological Society, 144 (2018), pp. 251–260.

[18] S. Chaturantabut and D. C. Sorensen, *Nonlinear model reduction via discrete empirical interpolation*, SIAM Journal on Scientific Computing, 32 (2010), pp. 2737–2764.

[19] G. Chen, J. Sun, and Y.-M. Li, *Adaptive reduced-order-model-based control-law design for active flutter suppression*, Journal of Aircraft, 49 (2012), pp. 973–980.

[20] P. Chen, A. Quarteroni, and G. Rozza, *A weighted reduced basis method for elliptic partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 3163–3185.

[21] P. Chen, A. Quarteroni, and G. Rozza, *Reduced basis methods for uncertainty quantification*, SIAM/ASA Journal on Uncertainty Quantification, 5 (2017), pp. 813–869.

[22] F. Chinesta, P. Ladeveze, and E. Cueto, *A short review on model order reduction based on proper generalized decomposition*, Archives of Computational Methods in Engineering, 18 (2011), pp. 395–404.

[23] L. Cordier, B. R. Noack, G. Tissot, G. Lehnasch, J. Delville, M. Balajewicz, G. Daviller, and R. K. Niven, *Identification strategies for model-based control*, Experiments in fluids, 54 (2013), p. 1580.

[24] C. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko, *Numerically modeling stochastic lie transport in fluid dynamics*, Multiscale Modeling & Simulation, 17 (2019), pp. 192–232.

[25] C. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko, *Sequential monte carlo for stochastic advection by Lie transport (SALT): A case study for the damped and forced incompressible 2d stochastic euler equation*, Journal on Uncertainty Quantification, (under review).

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

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

[28] G. Evensen, *Data assimilation: The ensemble Kalman filter*, Springer-Verlag, Newyork, 2006.

[29] L. Fick, Y. Maday, A. T. Patera, and T. Taddei, *A stabilized pod model for turbulent flows over a range of reynolds numbers: Optimal parameter sampling and constrained projection*, Journal of Computational Physics, 371 (2018), pp. 214–243.

[30] F. Flandoli, *The interaction between noise and transport mechanisms in PDEs*, Milan Journal of Mathematics, 79 (2011), pp. 543–560.

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

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

[33] S. Georgaka, G. Stabile, K. Star, G. Rozza, and M. J. Bluck, *A hybrid reduced order method for modelling turbulent heat transfer problems*, Computers & Fluids, (2020), p. 104615.

[34] G. Gottwald, D. Crommelin, and C. Franzke, *Stochastic climate theory*, in Nonlinear and Stochastic Climate Dynamics, Cambridge University Press, 2015.

[35] G. Gottwald and J. Harlim, *The role of additive and multiplicative noise in fil-*

tering complex dynamical systems, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 469 (2013), p. 20130096.

[36] M. Gunzburger and J. Ming, *Optimal control of stochastic flow over a backward-facing step using reduced-order modeling*, SIAM Journal on Scientific Computing, 33 (2011), pp. 2641–2663.

[37] S. Hijazi, G. Stabile, A. Mola, and G. Rozza, *Data-driven pod-galerkin reduced order model for turbulent flows*, Journal of Computational Physics, (2020), p. 109513.

[38] D. Holm, *Variational principles for stochastic fluid dynamics*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 471 (2015).

[39] J. Hunt, A. Wray, and P. Moin, *Eddies, stream, and convergence zones in turbulent flows*, Center for turbulence research report CTR-S88, (1988), pp. 193–208.

[40] S. Kadri Harouna and E. Mémin, *Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling*, Computers & Fluids, 156 (2017), pp. 456–469.

[41] R. Kraichnan, *Small-scale structure of a scalar field convected by turbulence*, Physics of Fluids (1958-1988), 11 (1968), pp. 945–953.

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

[43] T. Kurtz, *A limit theorem for perturbed operator semigroups with applications to random evolutions*, Journal of Functional Analysis, 12 (1973), pp. 55–67.

[44] O. Le Maitre, M. Reagan, H. Najm, R. Ghanem, and O. Knio, *A stochastic projection method for fluid flow. II. random process*, Journal of Computational Physics, 181 (2002), pp. 9–44.

[45] C. Leclercq, F. Demourant, C. Poussot-Vassal, and D. Sipp, *Linear iterative method for closed-loop control of quasiperiodic flows*, Journal of Fluid Mechanics, 868 (2019), pp. 26–65.

[46] C. Leith, *Atmospheric predictability and two-dimensional turbulence*, Journal of the Atmospheric Sciences, 28 (1971), pp. 145–161.

[47] V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters, *Mathematical and physical ideas for climate science*, Reviews of Geophysics, 52 (2014), pp. 809–859.

[48] J. L. Lumley, *Coherent structures in turbulence*, in Transition and turbulence, Elsevier, 1981, pp. 215–242.

[49] A. Majda, I. Timofeyev, and E. Vanden Eijnden, *A mathematical framework for stochastic climate models*, Communications on Pure and Applied Mathematics, 54 (2001), pp. 891–974.

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

[51] O. Métais and M. Lesieur, *Statistical predictability of decaying turbulence*, Journal of the atmospheric sciences, 43 (1986), pp. 857–870.

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

[53] L. Mitchell and G. Gottwald, *Data assimilation in slow-fast systems using homogenized climate models*, Journal of the atmospheric sciences, 69 (2012), pp. 1359–1377.

[54] A. Nouy, *Generalized spectral decomposition method for solving stochastic finite element equations: invariant subspace problem and dedicated algorithms*, Computer Methods in Applied Mechanics and Engineering, 197 (2008), pp. 4718–4736.

998 [55] A. Nouy, *Recent developments in spectral stochastic methods for the numerical solu-*
999 *tion of stochastic partial differential equations*, Archives of Computational Methods in
1000 Engineering, 16 (2009), pp. 251–285.
1001 [56] A. Nouy, *A priori model reduction through proper generalized decomposition for solv-*
1002 *ing time-dependent partial differential equations*, Computer Methods in Applied Me-
1003 chanics and Engineering, 199 (2010), pp. 1603–1626.
1004 [57] S. Orszag, *Analytical theories of turbulence*, Journal of Fluid Mechanics, 41 (1970),
1005 pp. 363–386.
1006 [58] G. Papanicolaou and W. Kohler, *Asymptotic theory of mixing stochastic ordinary*
1007 *differential equations*, Communications on Pure and Applied Mathematics, 27 (1974),
1008 pp. 641–668.
1009 [59] A. Papavasiliou, G. Pavliotis, and A. Stuart, *Maximum likelihood drift estima-*
1010 *tion for multiscale diffusions*, Stochastic Processes and their Applications, 119 (2009),
1011 pp. 3173–3210.
1012 [60] G. Pavliotis and A. Stuart, *Parameter estimation for multiscale diffusions*, Jour-
1013 nal of Statistical Physics, 127 (2007), pp. 741–781.
1014 [61] L. Perret, E. Collin, and J. Delville, *Polynomial identification of POD based*
1015 *low-order dynamical system*, Journal of Turbulence, (2006), p. N17.
1016 [62] C. Prévôt and M. Röckner, *A concise course on stochastic partial differential*
1017 *equations*, vol. 1905, Springer, 2007.
1018 [63] V. Resseguier, M. Ladvig, and D. Heitz, *Real-time algorithm to estimate and*
1019 *predict unsteady flows from few measurements using reduced-order models*, Submitted
1020 to Journal of Computational Physics, (Submitted).
1021 [64] V. Resseguier, L. Li, G. Jouan, P. Dérian, E. Mémin, and C. Bertrand,
1022 *New trends in ensemble forecast strategy: uncertainty quantification for coarse-grid*
1023 *computational fluid dynamics*, Archives of Computational Methods in Engineering,
1024 (2020), pp. 1–82.
1025 [65] V. Resseguier, E. Mémin, and B. Chapron, *Geophysical flows under location*
1026 *uncertainty, part I random transport and general models*, Geophysical & Astrophysical
1027 Fluid Dynamics, 111 (2017), pp. 149–176.
1028 [66] V. Resseguier, E. Mémin, and B. Chapron, *Geophysical flows under location*
1029 *uncertainty, part II quasi-geostrophy and efficient ensemble spreading*, Geophysical &
1030 Astrophysical Fluid Dynamics, 111 (2017), pp. 177–208.
1031 [67] V. Resseguier, E. Mémin, D. Heitz, and B. Chapron, *Stochastic modelling*
1032 *and diffusion modes for proper orthogonal decomposition models and small-scale flow*
1033 *analysis*, Journal of Fluid Mechanics, 826 (2017), pp. 888–917.
1034 [68] V. Resseguier, W. Pan, and B. Fox-Kemper, *Data-driven versus self-similar*
1035 *parameterizations for stochastic advection by lie transport and location uncertainty*,
1036 Nonlinear Processes in Geophysics, 27 (2020), pp. 209–234.
1037 [69] T. Sapsis and A. Majda, *Blending modified Gaussian closure and non-Gaussian re-*
1038 *duced subspace methods for turbulent dynamical systems*, Journal of Nonlinear Science,
1039 23 (2013), pp. 1039–1071.
1040 [70] T. Sapsis and A. Majda, *Statistically accurate low-order models for uncertainty*
1041 *quantification in turbulent dynamical systems*, Proceedings of the National Academy
1042 of Sciences, 110 (2013), pp. 13705–13710.
1043 [71] T. Sapsis and A. Majda, *A statistically accurate modified quasilinear Gaussian clo-*
1044 *sure for uncertainty quantification in turbulent dynamical systems*, Physica D: Non-
1045 linear Phenomena, 252 (2013), pp. 34–45.
1046 [72] P. Schmid, *Dynamic mode decomposition of numerical and experimental data.*, J.

Fluid Mech., 656 (2010), pp. 5–28.
[73] C. Soize and C. Farhat, *A nonparametric probabilistic approach for quantifying uncertainties in low-dimensional and high-dimensional nonlinear models*, International Journal for Numerical Methods in Engineering, 109 (2017), pp. 837–888.
[74] G. Stabile, F. Ballarin, G. Zuccarino, and G. Rozza, *A reduced order variational multiscale approach for turbulent flows*, Advances in Computational Mathematics, (2019), pp. 1–20.
[75] G. Stabile, S. Hijazi, A. Mola, S. Lorenzi, and G. Rozza, *POD-Galerkin reduced order methods for CFD using finite volume discretisation: vortex shedding around a circular cylinder*, Communications in Applied and Industrial Mathematics, 8 (2017), pp. 210–236.
[76] G. Stabile and G. Rozza, *Finite volume POD-Galerkin stabilised reduced order methods for the parametrised incompressible navier–stokes equations*, Computers & Fluids, 173 (2018), pp. 273–284.
[77] D. Torlo, F. Ballarin, and G. Rozza, *Stabilized weighted reduced basis methods for parametrized advection dominated problems with random inputs*, SIAM/ASA Journal on Uncertainty Quantification, 6 (2018), pp. 1475–1502.
[78] A. Towne, O. Schmidt, and T. Colonius, *Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis*, Journal of Fluid Mechanics, 847 (2018), p. 821âĂŞ867.
[79] Z. Wang, I. Akhtar, J. Borggaard, and T. Iliescu, *Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison*, Computer Methods in Applied Mechanics and Engineering, 237 (2012), pp. 10–26.
[80] X. Xie, M. Mohebujjaman, L. G. Rebholz, and T. Iliescu, *Data-driven filtered reduced order modeling of fluid flows*, SIAM Journal on Scientific Computing, 40 (2018), pp. B834–B857.
[81] Y. Yang and E. Mémin, *High-resolution data assimilation through stochastic subgrid tensor and parameter estimation from 4DEnVar*, Tellus A: Dynamic Meteorology and Oceanography, 69 (2017), p. 1308772.