



# Multiscale model reduction for incompressible flows

Jared L. Callaham<sup>1,†</sup>, Jean-Christophe Loiseau<sup>2</sup> and Steven L. Brunton<sup>1</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, USA

<sup>2</sup>Arts et Métiers Institute of Technology, CNAM, DynFluid, HESAM Université, F-75013 Paris, France

(Received 29 November 2022; revised 21 April 2023; accepted 8 June 2023)

We introduce a projection-based model reduction method that systematically accounts for nonlinear interactions between the resolved and unresolved scales of the flow in a low-dimensional dynamical systems model. The proposed method uses a separation of time scales between the resolved and subscale variables to derive a reduced-order model with cubic closure terms for the truncated modes, generalizing the classic Stuart–Landau equation. The leading-order cubic terms are determined by averaging out fast variables through a perturbation series approximation of the action of a stochastic Koopman operator. We show analytically that this multiscale closure model can capture both the effects of mean-flow deformation and the energy cascade before demonstrating improved stability and accuracy in models of chaotic lid-driven cavity flow and vortex pairing in a mixing layer. This approach to closure modelling establishes a general theory for the origin and role of cubic nonlinearities in low-dimensional models of incompressible flows.

**Key words:** low-dimensional models

## 1. Introduction

Despite the complex and often chaotic dynamics exhibited by many unsteady fluid flows, in many cases it is dominated by energetic coherent structures evolving on relatively long length and time scales (Holmes, Lumley & Berkooz 1996). This realization made it possible to study fluid flows as high-dimensional dynamical systems evolving on a low-dimensional manifold, providing answers to longstanding questions on topics such as the route to turbulence (Landau 1944; Hopf 1948; Ruelle & Takens 1971; Swinney & Gollub 1981) and the role of nonlinear interactions (Landau 1944; Stuart 1958). The persistent nature of these coherent structures eventually also raised the possibility of

† Email address for correspondence: [jc244@uw.edu](mailto:jc244@uw.edu)

using low-dimensional surrogate models for optimization and control objectives (Noack, Morzynski & Tadmor 2011; Brunton & Noack 2015; Rowley & Dawson 2017).

The fundamental challenges in constructing such reduced-order models may be broken into two categories: approximating the structures themselves, and approximating their evolution. We refer to these as the kinematic and dynamic approximations, respectively. A modal separation of variables assumption is often employed for the former task (Taira *et al.* 2017, 2020), but although this may be suitable for many closed or diffusion-dominated flows, it is not a natural representation of travelling waves or advection-dominated flows (Rowley & Marsden 2000; Reiss *et al.* 2018; Rim, Moe & LeVeque 2018; Grimberg, Farhat & Youkilis 2020; Mendible *et al.* 2020). This issue is inextricably linked to the problem of modelling the coherent structure dynamics; as is well known in many domains, a proper choice of coordinates can greatly simplify the modelling task (Champion *et al.* 2019).

The multiscale nature of fluid flows further complicates both the kinematic and dynamic aspects of low-dimensional modelling. For instance, the effective dimensionality of a chaotic flow well above the threshold of instability may be orders of magnitude greater than that of a laminar flow with one or two instability modes. Meanwhile, the ‘triadic’ structure of the nonlinear interactions in the wavenumber or frequency domain ensures that the dynamics of all scales across the flow is linked. Thus, even if the large, energetic coherent structures can be approximated with a low-dimensional basis, models of their dynamics that do not account for the role played by the unresolved degrees of freedom are often unstable or physically inconsistent (Noack *et al.* 2011; Callaham *et al.* 2021). Similar considerations impact the development of numerical methods (Bazilevs *et al.* 2007), self-consistent mean flow modelling (Meliga 2017) and resolvent analysis (Padovan, Otto & Rowley 2020; Rigas, Sipp & Colonius 2021; Barthel, Zhu & McKeon 2021; Barthel, Gomez & McKeon 2022).

The need for subscale modelling was apparent even in early work combining empirical modal approximations, such as the proper orthogonal decomposition (POD), with physics-based model reduction, such as Galerkin projection. For example, low-dimensional models of vortex shedding in the globally unstable cylinder wake could accurately predict the dynamics over short times, but were subject to structural instability over a longer time horizon (Deane *et al.* 1991; Ma & Karniadakis 2002). Eventually, Noack *et al.* (2003) showed that this was a result of the failure of the standard post-transient POD basis to resolve the Stuart–Landau mechanism of mean-flow deformation associated with nonlinear interactions between the fluctuations (Landau 1944; Stuart 1958), an insight that enabled low-dimensional modelling of natural and actuated flows with an increasingly complex dynamics (Luchtenburg *et al.* 2009; Deng *et al.* 2020; Sieber, Paschereit & Oberleithner 2021; Callaham *et al.* 2022; Deng *et al.* 2021).

Contemporary work on POD–Galerkin modelling of turbulent shear flows also recognized the need for closure models that could approximate the effect of unresolved scales (Aubry *et al.* 1988; Rempfer & Fasel 1994; Ukeiley *et al.* 2001). These efforts targeted two distinct physical mechanisms: mean-flow deformation and subscale dissipation. For flows that are either parallel (Aubry *et al.* 1988) or weakly non-parallel under the assumption of Taylor’s frozen turbulence hypothesis (Ukeiley *et al.* 2001), these authors developed a Boussinesq eddy viscosity relationship between the resolved scales and a slowly varying parallel mean flow, leading to stabilizing cubic terms consistent with the Stuart–Landau description. To capture dissipation due to the unresolved scales, these models also adopted a linear mixing length approximation.

Although these studies remain landmark explorations of low-dimensional coherent structure modelling, there are several opportunities to improve the proposed

closure strategies. First, it is difficult to generalize the Reynolds stress models applied to parallel shear flows by Aubry *et al.* (1988) and Ukeiley *et al.* (2001) to fully inhomogeneous flows. While the ‘shift mode’ approach to mean flow modelling via an augmented POD basis introduced by Noack *et al.* (2003) is agnostic to the geometry of the flow, it requires computation of the unstable steady state of the Navier–Stokes equations, which is not generally experimentally accessible. Second, in the Richardson–Kolmogorov energy cascade description, energy is transferred from large to small scales through nonlinear interactions, where it is finally dissipated. This is not consistent with a linear mixing length model, a fact exploited by later work investigating nonlinear models of subscale dissipation (Wang *et al.* 2012; Cordier *et al.* 2013; Östh *et al.* 2014), particularly the finite time thermodynamics approach, which is centred on modelling unresolved nonlinear energy transfers (Noack *et al.* 2008).

Recent years have seen a surge in interest in data-driven and machine learning-based methods (Brenner, Eldredge & Freund 2019; Duraisamy, Iaccarino & Xiao 2019; Brunton, Noack & Koumoutsakos 2020), including a number of proposed closure and stabilization schemes for reduced-order models, either through regression to additional linear–quadratic terms (Mohebujjaman *et al.* 2017; Mohebujjaman, Rebholz & Iliescu 2018; Xie *et al.* 2018) or by adding a deep learning model to approximate the residual (San & Maulik 2018*a,b*; Menier *et al.* 2022). Alternative work has explored interpretable system identification methods that forego the projection-based model altogether (Brunton, Proctor & Kutz 2016; Peherstorfer & Willcox 2016; Loiseau & Brunton 2018; Qian *et al.* 2020; Callahan *et al.* 2022), but may incorporate physical constraints derived from the Galerkin system (Loiseau, Noack & Brunton 2018*b*; Deng *et al.* 2020; Kaptanoglu *et al.* 2021*a*). If an accurate, non-intrusive model is more important than interpretability or satisfying physical constraints, then the traditional projection-based framework can be eliminated altogether with black-box neural network forecasting methods (Hesthaven & Ubbiali 2018; Wan *et al.* 2018).

These empirical closure models are constructed on the assumption that the influence of the unresolved variables can be approximated based on information from the resolved variables alone. The success of the Reynolds stress–mean-flow models (Aubry *et al.* 1988; Ukeiley *et al.* 2001; Mantič-Lugo, Arratia & Gallaire 2014), invariant or centre manifold reductions (Coulet & Spiegel 1983; Guckenheimer & Holmes 1983; Noack *et al.* 2003; Carini, Auteri & Giannetti 2015) and weakly nonlinear analysis (Stuart 1958; Sipp & Lebedev 2007; Meliga, Chomaz & Sipp 2009; Meliga & Chomaz 2011) suggests that there may be circumstances where this relationship may be derived analytically via traditional analysis. The unifying thread between these methods is the assumption of a scale separation between the resolved and unresolved variables that can be exploited to develop an asymptotically correct closure model.

Beyond the field of fluid dynamics, the method of adiabatic elimination (Haken 1983; Risken 1996) has long been used to discard the fast variables in systems with emergent large-scale coherence when there is a separation in time scales, while heterogeneous multiscale methods have played an important role in simulating physical systems with widely separated scales (Weinan & Engquist 2003; Weinan *et al.* 2007; Weinan 2011). A similar stochastic averaging approach has been successful in climate modelling (Majda, Timofeyev & Vanden-Eijnden 2001), where the primitive equations have the same quadratic nonlinearity as the usual Navier–Stokes equations without rotation, buoyancy, topography, etc. This method, also called homogenization in the multiscale modelling literature, has its roots in the theory of singular perturbations of Markov processes (Kurtz 1973; Papanicolaou 1976) and is rigorously supported for stochastic systems with asymptotic scale separation; see Majda *et al.* (2001), Givon, Kupferman &

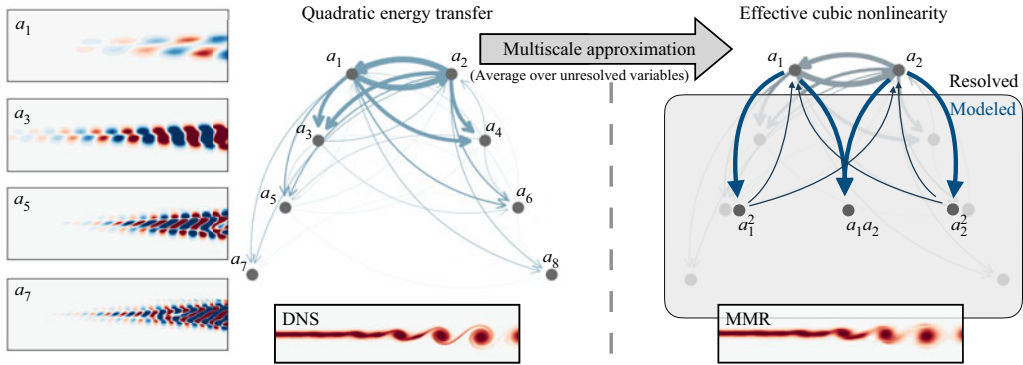


Figure 1. Multiscale closure model applied to a mixing layer. The visualization of the network of average quadratic energy transfer between the leading harmonic modes shows the cascade of energy to higher-order modes. The multiscale model reduction (MMR) method approximates the effects of unresolved higher-order modes via stochastic averaging, which leads to a generalized Stuart–Landau-type equation with cubic nonlinear interactions. The network visualization of quadratic energy transfers (left) is computed from the modal coefficients  $a(t)$  and Galerkin model for the leading harmonics, while that of the multiscale approximation (right) is a notional illustration of the origin of the cubic terms in (4.18*b*). See § 6.3 for details on the construction of this figure and the low-dimensional model of the mixing layer.

Stuart (2004), Weinan (2011) or Pavliotis & Stuart (2012) for in-depth presentations. With some assumptions on ergodicity, a similar approach can also be taken with deterministic systems, even in a regime where the scale separation is not in the asymptotic limit (Majda, Timofeyev & Vanden-Eijnden 2006).

This work explores the application of multiscale stochastic averaging methods developed by Majda *et al.* (2001), Givon *et al.* (2004), Weinan (2011), Pradas *et al.* (2012), Pavliotis & Stuart (2012) and others to the closure problem in reduced-order models of incompressible flow or other systems reducible to a linear–quadratic dynamics (Rowley, Colonius & Murray 2004; Qian *et al.* 2020; Kaptanoglu *et al.* 2021*b*), including introducing an approximation to the form of the fast dynamics that allows for computation of the averaged dynamics in closed form. When applied to a Galerkin model of incompressible flow, this procedure effectively approximates nonlinear energy transfers to unresolved scales by higher-order nonlinearities in the resolved dynamics, as illustrated by figure 1 and illustrated schematically by figure 2. We refer to this secondary dimensionality reduction of the Galerkin system as MMR. Although dynamically complex fluid flows often exhibit structure across a wide range of spatio-temporal scales, in the present context the term multiscale refers specifically to the asymptotic time-scale separation assumed formally by the averaging procedure and does not imply anything about the spectral content of any particular flow.

Since fluid flows generally do not have a true scale separation away from the threshold of instability, we demonstrate via numerical simulations that this method is a robust and systematic approach to stabilizing low-dimensional models. This extends the work of Majda *et al.* (2006) exploring the application of this class of methods to systems beyond the parameter regimes where their validity can be rigorously proven. Multiscale model reduction is a unified framework for understanding the origin and importance of cubic terms in reduced-order models of the linear–quadratic Navier–Stokes equations, capable of capturing both mean-flow deformation and subscale dissipation. Throughout this work we also highlight connections to other modelling methods, including Koopman theory and weakly nonlinear analysis.

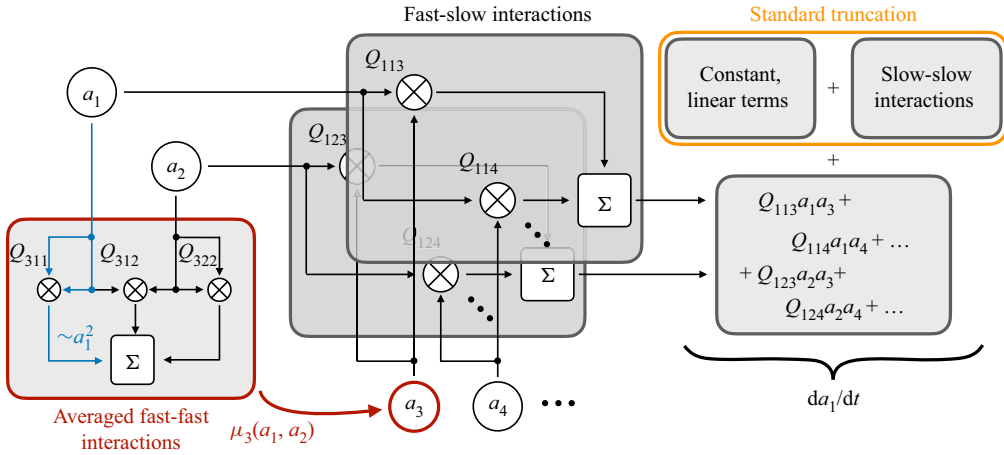


Figure 2. Schematic of nonlinear interactions in the multiscale closure scheme. The dynamics of one variable ( $a_1$ ) in a system with two slow variables involves quadratic interactions between fast and slow variables that would be neglected in a standard truncation. Instead, the proposed method averages over the fast scales, ultimately generating effective cubic nonlinearities in the closed equations. The quadratic interactions  $Q$  represent terms that might arise for instance from Galerkin projection via (2.6c).

## 2. Reduced-order modelling for incompressible flows

For more than three decades, POD and Galerkin projection have been the foundational tools in low-order modelling for nonlinear incompressible fluid flow. The POD analysis identifies dominant coherent structures in the flow and provides an energy-optimal linear modal basis. Galerkin projection approximates both the state of the flow and its time derivative in this finite-dimensional subspace, resulting in a reduced system of ordinary differential equations in terms of the POD mode amplitudes. These topics have been covered extensively elsewhere (see e.g. Holmes *et al.* 1996; Noack *et al.* 2011; Benner, Gugercin & Willcox 2015; Rowley & Dawson 2017; Taira *et al.* 2017); here, we provide a brief overview.

We assume the flow is governed by the unsteady incompressible Navier–Stokes equations

$$\left. \begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u} \otimes \mathbf{u}) &= -\nabla p + \frac{1}{Re} \nabla^2 \mathbf{u} \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned} \right\} \quad (2.1)$$

where  $\mathbf{u}(\mathbf{x}, t)$  is the velocity field,  $p$  is the pressure field and  $Re$  is the Reynolds number,  $Re = UL/\nu$ , based on the kinematic viscosity  $\nu$  and suitable length and velocity scales  $L$  and  $U$ . The nonlinear term is expressed in divergence form with the tensor product  $\otimes$ . In this work, we only consider two dimensional examples for which  $\mathbf{u} = [u \ v]^T$ , although all of the following applies equally well to three-dimensional flows.

In order to reduce the system of partial differential equations (PDEs) (2.1) to a finite-dimensional system of ordinary differential equations, we assume that the space and time dependence can be separated, so that an arbitrary velocity field  $\mathbf{u}(\mathbf{x}, t)$  can be approximated with the  $r$ -dimensional modal representation

$$\mathbf{u}(\mathbf{x}, t) \approx \boldsymbol{\psi}_0(\mathbf{x}) + \sum_{i=1}^r \boldsymbol{\psi}_i(\mathbf{x}) a_i(t), \quad (2.2)$$



where  $\boldsymbol{\psi}_0(\mathbf{x})$  is a fixed base flow around which the expansion is performed and each mode individually satisfies the divergence-free constraint  $\nabla \cdot \boldsymbol{\psi}_i = 0$  for  $i = 0, 1, 2, \dots, r$ . Unless otherwise specified we will take  $\boldsymbol{\psi}_0$  to be the mean flow  $\bar{\mathbf{u}}$  as estimated with a time average.

The POD identifies the set of modes  $\{\boldsymbol{\psi}_i\}_{i=1}^r$  that provide the optimal rank- $r$  approximation for an average velocity field in the energy norm induced by the velocity inner product

$$\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \int_{\Omega} \mathbf{u}_1(\mathbf{x}) \cdot \mathbf{u}_2(\mathbf{x}) \, d\Omega, \tag{2.3}$$

defined on the spatial domain  $\Omega$  for real-valued velocity fields  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . We approximate this integral with a weighted sum  $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle \approx \mathbf{u}_1^T \mathbf{W} \mathbf{u}_2$ , where  $\mathbf{u}_1$  and  $\mathbf{u}_2$  are the discrete approximations to  $\mathbf{u}_1$  and  $\mathbf{u}_2$ , and  $\mathbf{W}$  is a diagonal weight matrix containing the cell volumes. We estimate the POD modes from the two-point temporal correlation matrix using the method of snapshots (Sirovich 1987).

The POD expansion (2.2) is often only applied to the velocity fields. This can pose a problem for model reduction methods approximating the pressure gradient term in the Navier–Stokes equations. While this term vanishes for closed flows and is often negligible for flows with a localized global instability, it is important in open shear flows, such as the mixing layer examined in §6 (Noack, Papas & Monkewitz 2005). There are various methods for approximating the pressure term (see e.g. Caiazzo *et al.* 2014), but in this work we use a velocity–pressure expansion with the same inner product (2.3). Defining  $\mathbf{q}(\mathbf{x}, t) = [\mathbf{u} \ p]^T$ , we replace (2.2) and (2.3) with

$$\mathbf{q}(\mathbf{x}, t) \approx \begin{bmatrix} \boldsymbol{\psi}_0^u(\mathbf{x}) \\ \boldsymbol{\psi}_0^p(\mathbf{x}) \end{bmatrix} + \sum_{i=1}^r \begin{bmatrix} \boldsymbol{\psi}_i^u(\mathbf{x}) \\ \boldsymbol{\psi}_i^p(\mathbf{x}) \end{bmatrix} a_i(t) \tag{2.4a}$$

$$= \boldsymbol{\psi}_0(\mathbf{x}) + \sum_{i=1}^r \boldsymbol{\psi}_i(\mathbf{x}) a_i(t), \tag{2.4b}$$

and energy inner product  $\langle \mathbf{q}_1, \mathbf{q}_2 \rangle_u \equiv \langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ . Numerically, we simply carry the pressure fields through the method of snapshots by setting the diagonal entries of  $\mathbf{W}$  corresponding to pressure to zero. The pressure components of the POD modes are therefore estimated from the same linear combination of snapshots as the velocity components.

The derivation of the Galerkin system proceeds by substituting the velocity–pressure POD approximation (2.4) into the Navier–Stokes equations (2.1). In order for the dynamics to be an optimal continuous-time approximation, the residual error should be orthogonal to the POD subspace. Using the inner product (2.3) and orthonormality properties of the POD, this optimality condition leads to the linear–quadratic system of ordinary differential equations (ODEs) (Holmes *et al.* 1996; Noack *et al.* 2011)

$$\dot{a}_i = F_i + L_{ij} a_j + Q_{ijk} a_j a_k, \quad i, j, k = 1, 2, \dots, r \tag{2.5}$$

with constant, linear and quadratic terms given by

$$F_i = \langle \boldsymbol{\psi}_i, -\nabla \cdot (\boldsymbol{\psi}_0^u \otimes \boldsymbol{\psi}_0^u) - \nabla \boldsymbol{\psi}_0^p + Re^{-1} \nabla^2 \boldsymbol{\psi}_0^u \rangle_u \tag{2.6a}$$

$$L_{ij} = \langle \boldsymbol{\psi}_i, -\nabla \cdot (\boldsymbol{\psi}_0^u \otimes \boldsymbol{\psi}_j^u + \boldsymbol{\psi}_j^u \otimes \boldsymbol{\psi}_0^u) - \nabla \boldsymbol{\psi}_j^p + Re^{-1} \nabla^2 \boldsymbol{\psi}_j^u \rangle_u \tag{2.6b}$$

$$Q_{ijk} = \langle \boldsymbol{\psi}_i, -\nabla \cdot (\boldsymbol{\psi}_j^u \otimes \boldsymbol{\psi}_k^u) \rangle_u. \tag{2.6c}$$

A key feature of the quadratic nonlinearity of the Navier–Stokes equations is that it is energy preserving, in the sense that it has no net contribution to the evolution equation for

kinetic energy in the spectral domain (Kraichnan & Chen 1989), with some restrictions on the boundary conditions. This is the foundation of the energy cascade picture of turbulence, in which the role of the nonlinearity is to transfer energy from the large scales to the small, dissipative scales. For the Galerkin system, the energy preservation condition implies that (Schlegel & Noack 2015)

$$Q_{ijk} + Q_{ikj} + Q_{jik} + Q_{jki} + Q_{kij} + Q_{kji} = 0. \tag{2.7}$$

This is often approximately true numerically for POD–Galerkin systems, but enforcing it explicitly can improve the stability of the model (Cordier *et al.* 2013). Based on the structure of the linear–quadratic system (2.5), the quadratic tensor can also be symmetrized in the last two indices

$$Q_{ijk} = Q_{ikj}. \tag{2.8}$$

In this work we explicitly enforce conditions (2.7) and (2.8) after construction of the quadratic tensor (2.6c).

### 3. Generators of deterministic and stochastic processes

In the study of mechanics there are often several equivalent representations of the same system (e.g. Lagrangian, Hamiltonian, etc.), each of which may be useful depending on the application. The operator-theoretic perspective will be especially useful in the development of the multiscale closure model as a framework for abstract formal manipulations of the dynamical systems model. In particular, the closure model will be derived via a perturbation series approximation to the action of a stochastic Koopman operator. Critically, the infinite-dimensional operator itself need not be computed or directly approximated; instead, the closed model appears as a solvability condition analogous to the derivation of the amplitude equation in weakly nonlinear analysis (Sipp & Lebedev 2007).

This section gives a brief overview of topics necessary for the development of the closure models in § 4; for a more in-depth presentation of these topics see, e.g. Risken (1996), Weinan (2011), Pavliotis & Stuart (2012), Klus *et al.* (2018, 2020) and Brunton *et al.* (2022). Koopman theory is perhaps the most widely discussed operator-theoretic method in recent work on the analysis of fluid flows and other large-scale nonlinear dynamics (Mezić 2005; Rowley *et al.* 2009; Mezić 2013; Klus *et al.* 2018; Brunton *et al.* 2022), making it a convenient place to begin the discussion.

Suppose the state  $\mathbf{x}(t)$  is governed by an autonomous ODE

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathcal{X}. \tag{3.1}$$

We will assume, unless otherwise specified, that states are real valued, e.g.  $\mathcal{X} \subseteq \mathbb{R}^N$ , and that all other functions are  $L^2$ -integrable with inner product

$$\langle f(\mathbf{x}), g(\mathbf{x}) \rangle = \int_{\mathcal{X}} f(\mathbf{x})g(\mathbf{x}) \, d\mathbf{x}, \tag{3.2}$$

where the usual complex conjugation is omitted since  $g$  is real valued.

The Koopman operator  $\mathcal{K}^t$  acts on the  $L^2$ -integrable space of scalar observables  $g(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}$ , advancing them forward time  $t$ . More precisely, let  $\mathbf{x}$  be the solution to the initial value problem of (3.1) with  $\mathbf{x}(0) = \mathbf{x}_0$ . Then the Koopman operator is defined as

$$\mathcal{K}^t g(\mathbf{x}_0) = g(\mathbf{x}(t)). \tag{3.3}$$

Although the Koopman operator is generally difficult to either represent explicitly or approximate in a useful finite-dimensional subspace, analysis of its spectral properties

has drawn great interest in recent work. For our purposes, the infinitesimal generator  $\mathcal{L}$ , defined by

$$\mathcal{L}g = \lim_{t \rightarrow 0} \frac{\mathcal{K}^t g - g}{t}, \tag{3.4}$$

and sometimes called the Lie operator, is more theoretically useful. If we consider  $g(\mathbf{x}(t))$  to be an explicit function of time, then  $\mathcal{L}$  can be derived by applying the chain rule to  $g$ . With a slight abuse of notation, if  $g(\mathbf{x}, t)$  is taken to be a function of both time and initial state  $\mathbf{x}$  (i.e. in the Lagrangian frame of reference in state space), then this can be extended to a PDE over all of state space

$$\frac{\partial g}{\partial t} = \mathcal{L}g = \mathbf{f}(\mathbf{x}) \cdot \nabla_{\mathbf{x}} g. \tag{3.5}$$

Thus the generator  $\mathcal{L}$  is a linear advection operator governing the evolution of scalar observables of the system described by (3.1). Similarly, the adjoint of (3.5), defined by  $\langle \mathbf{f}, \mathcal{L}g \rangle = \langle \mathcal{L}^\dagger \mathbf{f}, g \rangle$  with the inner product (3.2)

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^\dagger \rho = -\nabla_{\mathbf{x}} \cdot (\rho \mathbf{f}(\mathbf{x})), \tag{3.6}$$

is a continuity equation governing the evolution of densities  $\rho(\mathbf{x}, t)$  in phase space. As a point of reference, (3.6) reduces to the Liouville equation from classical mechanics under the incompressibility condition  $\nabla_{\mathbf{x}} \cdot \mathbf{f} = 0$ , and  $\mathcal{L}^\dagger$  is also known as the generator of the Perron–Frobenius operator (Froyland 2005; Froyland & Padberg 2009; Froyland, Santitissadeekorn & Monahan 2010; Klus, Koltai & Schütte 2016).

This description of the dynamics can readily be extended to systems governed by stochastic differential equations (SDEs) of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\Sigma} \mathbf{w}(t), \tag{3.7}$$

where the deterministic component  $\mathbf{f}(\mathbf{x})$  is known as the drift function and the diffusion matrix  $\boldsymbol{\Sigma}$  modifies a vector-valued Gaussian white noise process  $\mathbf{w}(t)$ . In this case the evolution of the probability distribution  $\rho(\mathbf{x}, t)$  is governed by the stochastic analogue of (3.6), known as the forward Kolmogorov, or Fokker–Planck, equation

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^\dagger \rho = -\nabla_{\mathbf{x}} \cdot (\rho \mathbf{f}) + \nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T : (\rho \mathbf{D}), \tag{3.8}$$

where  $\mathbf{D} = \boldsymbol{\Sigma} \boldsymbol{\Sigma}^T / 2$  is the diffusion tensor and the colon denotes tensor contraction. The associated backwards Kolmogorov equation is the adjoint of (3.8)

$$\frac{\partial g}{\partial t} = \mathcal{L}g = \mathbf{f} \cdot \nabla_{\mathbf{x}} g + \mathbf{D} : \nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T g. \tag{3.9}$$

To interpret (3.9), consider the expectation of a scalar observable  $g(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}$  defined by the inner product (3.2) of  $g$  with the probability distribution  $\rho_0(\mathbf{x})$

$$\mathbb{E}[g] = \langle \rho_0, g \rangle = \int_{\mathcal{X}} \rho_0(\mathbf{x}) g(\mathbf{x}) \, d\mathbf{x}. \tag{3.10}$$

The probability distribution is advanced in time by  $\rho(\mathbf{x}, t) = e^{\mathcal{L}^\dagger t} \rho_0(\mathbf{x})$ . Using the definition of the adjoint,  $\langle \rho, \mathcal{L}g \rangle = \langle \mathcal{L}^\dagger \rho, g \rangle$ , the expectation of  $g$  at time  $t$  can be written



equivalently as

$$\mathbb{E}[g](t) = \langle \rho, g \rangle = \int_{\mathcal{X}} \rho_0(x) e^{\mathcal{L}t} g(x) dx. \quad (3.11)$$

Therefore, just as the Koopman operator advances an observable in time, the backwards Kolmogorov equation describes the evolution of the expectation of an observable in the case of statistically stationary system where  $\rho = \rho_0 \equiv \rho^\infty$ .

As a simple example relevant for the closure modelling presented below, consider the Ornstein–Uhlenbeck process defined by the SDE

$$\dot{x} = v(\mu - x) + \sigma w(t), \quad (3.12)$$

for positive constants  $v$ ,  $\mu$  and  $\sigma$ . The stationary distribution  $\rho^\infty(x)$  is given by the steady-state Fokker–Planck equation

$$0 = \mathcal{L}^\dagger \rho^\infty = -\frac{\partial}{\partial x} \rho^\infty(x) v(\mu - x) + \frac{\sigma^2}{2} \frac{d^2}{dx^2} \rho^\infty(x), \quad (3.13)$$

along with the usual normalization condition on  $\rho^\infty$ . This can be solved analytically by a Gaussian distribution with mean  $\mu$  and variance  $\sigma^2/2v$ . Defining an observable that is the state itself  $g(x) = x$ , the evolution is given by the backwards Kolmogorov equation (3.9), which in this case reduces to the initial value problem

$$\frac{\partial}{\partial g} t = v(\mu - g), \quad g(0) = x. \quad (3.14a,b)$$

The expectation at time  $t$  is the solution  $g(t)$ , which in this case is exponential decay towards the mean  $\mu$

$$\mathbb{E}[x](t) = g(t) = \mu - (x - \mu) e^{-vt}. \quad (3.15)$$

Given that the model reduction procedure described in § 2 aims to reduce the physics model from an infinite-dimensional PDE to a finite-dimensional system of ODEs, it may be counterintuitive that it would be helpful to return to an infinite-dimensional function space and represent the dynamics with a partial differential equation. The primary advantage in doing so is that these generators are linear operators, which in some cases can be amenable to approaches that are unavailable for the nonlinear dynamics of the ODE (3.1) or SDE (3.7).

For this reason, the prospect of using Koopman operators or Kolmogorov-type equations to analyse nonlinear and/or stochastic systems is appealing. However, aside from simple one-dimensional or linear examples like the Ornstein–Uhlenbeck process, it is typically very difficult to solve these PDEs analytically. Multiscale modelling approaches use operator representations to construct simplified approximations of the underlying systems, addressing their intractability with perturbation methods. As will be seen in § 4, an expansion can be used to derive a solvability condition based on the Fredholm alternative. Using the manipulations reviewed in this section, this condition can be interpreted as an average over subscale variables, which can be computed explicitly under suitable assumptions. For a Galerkin model of incompressible flow, the averaged slow dynamics takes the form of a cubic generalized Stuart–Landau equation that accounts for the leading-order effect of the unresolved fast variables.

#### 4. Multiscale closure modelling

One of the primary difficulties of reduced-order modelling for fluid flows is that the nonlinear interactions in the Navier–Stokes equations transfer energy between all scales of the flow. Thus, restricting the dynamics to a low-dimensional subspace can lead to significant approximation errors. Generally speaking, the goal of closure modelling is to augment this truncated model with additional terms that account for the effect of the unresolved scales.

The multiscale approach to closure modelling accomplishes this by first partitioning the dynamics into fast and slow variables, and then approximating the solution to the associated backwards Kolmogorov equation with a perturbation series expansion. The result can be interpreted as a Koopman generator of the form of (3.5) corresponding to the coarse-grained dynamics for the slow variables. As we will show, when applied to the linear–quadratic Galerkin system (2.5), this leads to a generalized Stuart–Landau equation including cubic terms.

An interesting feature of cubic Stuart–Landau-type models of fluid flow is that they are often more accurate than linear–quadratic models, even though the nonlinearity in the underlying governing equations is quadratic (Noack *et al.* 2003; Loiseau & Brunton 2018). One reason for this is the artificial space–time separation of variables introduced by the modal representation (2.2), which can introduce spurious degrees of freedom that represent phase-locked harmonics of travelling waves, for instance (Callaham, Brunton & Loiseau 2022). As a result, POD–Galerkin models are vulnerable to decoherence and can fail to improve with increasing rank, even when the kinematic approximation of the flow field becomes nearly perfect. As illustrated by the examples in § 6, the introduction of cubic terms can mitigate this, either by modifying the dynamics to resemble phase-locked nonlinear oscillators or by eliminating spurious degrees of freedom altogether.

Although multiscale methods can be rigorously justified when there is a strict separation of time scales (Majda *et al.* 2001; Weinan & Engquist 2003; Weinan 2011; Pavliotis & Stuart 2012), there is no such spectral gap in most fluid flows of practical interest. The proposed method for model reduction should therefore properly be viewed as an approximate closure model motivated by the asymptotic limit. Nevertheless, it provides a systematic method for stabilizing low-dimensional models and also highlights a generic mechanism by which higher-order nonlinearities can arise from the quadratic term in the governing equations.

##### 4.1. Averaging over unresolved variables

Beginning with the Galerkin dynamics (2.5) for the state consisting of POD coefficients  $\mathbf{a}(t) \in \mathbb{R}^r$ , where we assume that  $r$  is large enough for an accurate kinematic reconstruction of a typical flow field, we partition the system into slow variables  $\mathbf{x}(t) \in \mathcal{X} = \mathbb{R}^{r_0}$  and fast variables  $\mathbf{y}(t) \in \mathcal{Y} = \mathbb{R}^{r-r_0}$ , so that  $\mathbf{a} = [\mathbf{x}^T \ \mathbf{y}^T]^T$ . Since the POD modes are sorted according to typical energy content and not dominant time scale, in principle, this partition could lead to ‘fast’ variables included in  $\mathbf{x}$ , or *vice versa*. However, due to the mechanism of harmonic generation via quadratic nonlinear interaction, lower-order modes are often related to dominant instability modes, while higher-order modes tend to represent higher wavenumber and frequency content.

For concise notation we will use the Einstein convention that repeated indices imply summation and we will omit explicit summation unless not doing so would lead to ambiguity. We will index the slow variables with Roman subscripts  $i, j, k, \ell$  ranging from 1 to  $r_0$  and the fast variables with Greek subscripts  $\alpha, \beta, \gamma$  ranging from 1 to  $r - r_0$ .

Without loss of generality we will also assume that the quadratic term has been symmetrized in the last two indices, so that  $Q_{ijk} = Q_{ikj}$ . Then the partitioned Galerkin system is

$$\dot{x}_i = f_i^x(\mathbf{x}, \mathbf{y}) = F_i^x + L_{ij}^{xx}x_j + L_{i\beta}^{xy}y_\beta + Q_{ijk}^{xxx}x_jx_k + 2Q_{ij\beta}^{xxy}x_jy_\beta + Q_{i\beta\gamma}^{xyy}y_\beta y_\gamma \tag{4.1a}$$

$$\dot{y}_\alpha = f_\alpha^y(\mathbf{x}, \mathbf{y}) = F_\alpha^y + L_{\alpha j}^{yx}x_j + L_{\alpha\beta}^{yy}y_\beta + Q_{\alpha jk}^{yxx}x_jx_k + 2Q_{\alpha j\beta}^{yyx}x_jy_\beta + Q_{\alpha\beta\gamma}^{yyy}y_\beta y_\gamma. \tag{4.1b}$$

Standard truncation of this system is equivalent to retaining only the terms  $\mathbf{F}^x$ ,  $\mathbf{L}^{xx}$  and  $\mathbf{Q}^{xxx}$ . Here, we will attempt to approximate the terms involving the fast variables in  $f^x(\mathbf{x}, \mathbf{y})$  in an average sense in order to derive a closed system  $\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x})$ , largely following the approach of Pavliotis & Stuart (2012).

We assume that the time scales of the fast and slow dynamics are separated by a parameter  $\epsilon \ll 1$ , so that we may define  $f^x(\mathbf{x}, \mathbf{y}) \equiv \epsilon \tilde{f}^x(\mathbf{x}, \mathbf{y})$ . Furthermore, we note that the role of the fast self-interaction term  $\mathbf{Q}^{yyy}(\mathbf{y}, \mathbf{y})$  is to transfer energy between the unresolved scales. Since this mechanism is of secondary importance to the transfers between slow and fast scales, we apply a version of the ‘working assumption of stochastic modelling’ (Majda *et al.* 2001)

$$\sum_{\beta, \gamma=1}^{r-r_0} Q_{\alpha\beta\gamma}^{yyy}y_\beta y_\gamma \approx \sigma_\alpha w_\alpha(t), \tag{4.2}$$

where  $w(t)$  is a Wiener process and  $\sigma$  is an as-yet-undefined constant forcing amplitude. This approximation of the fast self-interactions as uncorrelated additive white noise is the simplest possible stochastic closure. More sophisticated models such as state-dependent noise or non-diagonal covariance could also be considered. For the results shown in § 6 we ultimately take the noise amplitude to be zero, so that (4.2) is sufficiently expressive for the present purposes.

Finally, we coarse grain the dynamics on the time scale  $\tau = \epsilon t$ , leading to the slow/fast system

$$\frac{\partial x_i}{\partial \tau} = \tilde{f}_i^x(\mathbf{x}, \mathbf{y}) \tag{4.3a}$$

$$\frac{\partial y_\alpha}{\partial \tau} = \frac{1}{\epsilon} \tilde{f}_\alpha^y(\mathbf{x}, \mathbf{y}) + \frac{1}{\sqrt{\epsilon}} \sigma_\alpha w_\alpha(\tau), \tag{4.3b}$$

where  $\tilde{f}^y(\mathbf{x}, \mathbf{y})$  are the fast dynamics (4.1b) excluding the self-interaction term  $\mathbf{Q}^{yyy}(\mathbf{y}, \mathbf{y})$  modelled by (4.2) and we have also used the scaling property of Wiener processes that  $w(\epsilon t) = \sqrt{\epsilon} w(t)$ .

Defining an arbitrary scalar-valued observable  $g^\epsilon(\mathbf{x}, \mathbf{y})$ , the backwards Kolmogorov equation (3.9) associated with (4.3) is

$$\frac{\partial g^\epsilon}{\partial \tau} = \frac{1}{\epsilon} \underbrace{\left[ \tilde{f}_\alpha^y(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial y_\alpha} + \frac{\sigma_\alpha^2}{2} \frac{\partial^2}{\partial y_\alpha^2} \right]}_{\mathcal{L}_0} g^\epsilon + \underbrace{\tilde{f}_i^x(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial x_i}}_{\mathcal{L}_1} g^\epsilon. \tag{4.4}$$

Note that  $\mathcal{L}_0$  can be viewed as the generator of a stochastic process in  $\mathbf{y}$  with  $\mathbf{x}$  as a fixed parameter. We make the following assumptions related to the ergodicity of  $\mathbf{y}$ :

- (i) The operator  $\mathcal{L}_0$  has a one-dimensional nullspace spanned by constants in  $\mathbf{y}$

$$\mathcal{L}_0 g(\mathbf{x}) = 0. \tag{4.5}$$

- (ii) The Fokker–Planck operator  $\mathcal{L}_0^\dagger$  has a one-dimensional nullspace corresponding to the stationary distribution  $\rho_x^\infty$ , where again  $x$  is treated as a fixed parameter

$$\mathcal{L}_0^\dagger \rho_x^\infty(\mathbf{y}) = 0, \tag{4.6}$$

along with the usual normalization condition  $\int_{\mathbf{y}} \rho_x^\infty \, d\mathbf{y} = 1$ .

Since  $\mathcal{L}_0$  is the generator of a stochastic Koopman operator, (4.5) states that only observables that do not depend on the fast variables  $\mathbf{y}$  are constant with respect to the fast dynamics. Equation (4.6) requires that there be a unique stationary probability density function in  $\mathbf{y}$  for each value of  $x$ . As shown in §4.2, in this work  $\mathcal{L}_0$  corresponds to a multi-dimensional Ornstein–Uhlenbeck process for which both assumptions hold and  $\rho_x^\infty$  can be expressed analytically. These assumptions obviate the need to express the fast scale  $\mathbf{y}$  as an instantaneous function of  $x$ , as in an invariant manifold model (Guckenheimer & Holmes 1983; Pavliotis & Stuart 2012). Instead, the fast variable is modelled with a simplified distribution that can be averaged over as follows. We assume that the solution to the backwards Kolmogorov equation (4.4) can be approximated by means of an asymptotic expansion

$$g^\epsilon(\mathbf{x}, \mathbf{y}, \tau) = g_0 + \epsilon g_1 + O(\epsilon^2). \tag{4.7}$$

This expansion holds for small values of the scale separation parameter  $\epsilon$ , implying wide time-scale separations between the fast and slow dynamics. Substituting into (4.4) and equating powers of  $\epsilon$  gives the consistency conditions

$$0 = \mathcal{L}_0 g_0, \tag{4.8a}$$

$$\frac{\partial g_0}{\partial \tau} = \mathcal{L}_0 g_1 + \mathcal{L}_1 g_0. \tag{4.8b}$$

By virtue of the assumption of a one-dimensional nullspace, (4.8a) is satisfied if  $g_0$  is not a function of  $\mathbf{y}$ , or  $g_0 = g_0(\mathbf{x}, \tau)$ . As expected, the leading-order solution does not depend on the fast variable.

An effective evolution equation for  $g_0$  can be derived by considering (4.8b) as a linear equation for  $g_1(\mathbf{x}, \mathbf{y}, \tau)$ . The Fredholm alternative specifies that, for any equation of the form  $\mathcal{L}_0 g_1 = b$  to have a unique solution, all functions  $\rho$  in the nullspace of the adjoint operator  $\mathcal{L}_0^\dagger$  must be orthogonal to  $b$ . Since we have assumed that the nullspace of  $\mathcal{L}_0^\dagger$  is one-dimensional and spanned by the stationary distribution  $\rho_x^\infty(\mathbf{y})$ , this implies that  $\langle \rho_x^\infty, b \rangle = 0$ . In other words, the solvability condition for

$$\mathcal{L}_0 g_1 = \frac{\partial g_0}{\partial \tau} - \mathcal{L}_1 g_0, \tag{4.9}$$

is that the right-hand side of (4.9) has zero mean with respect to the stochastic process generated by  $\mathcal{L}_0$

$$\int_{\mathbf{y}} \rho_x^\infty(\mathbf{y}) \left[ \frac{\partial}{\partial \tau} - \mathcal{L}_1 \right] g_0(\mathbf{x}, \tau) \, d\mathbf{y} = 0. \tag{4.10}$$

Using the normalization condition for  $\rho_x^\infty$ , this simplifies to a closed evolution equation for  $g_0(\mathbf{x}, \tau)$

$$\frac{\partial g_0}{\partial \tau} = \left[ \int_{\mathbf{y}} \rho_x^\infty(\mathbf{y}) \tilde{f}_i^x(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \right] \frac{\partial g_0}{\partial x_i}. \tag{4.11}$$

Since the observable  $g_\epsilon$  is arbitrary, (4.11) must hold for any  $g_0$ . This PDE has the form of the generator of a deterministic Koopman operator corresponding to the coarse-grained

dynamics in  $\mathbf{x}$  alone. Undoing the  $\epsilon$  scaling in  $\tau$  and  $\tilde{\mathbf{f}}^x$ , (4.11) can be written as

$$\frac{\partial g_0}{\partial t} = \hat{\mathbf{f}}(\mathbf{x}) \cdot \nabla g_0, \tag{4.12}$$

corresponding to the averaged dynamics

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}), \tag{4.13a}$$

$$\hat{f}_i(\mathbf{x}) = \int_{\mathcal{Y}} \rho_{\mathbf{x}}^\infty(\mathbf{y}) f_i^x(\mathbf{x}, \mathbf{y}) \, d\mathbf{y}. \tag{4.13b}$$

The distribution  $\rho_{\mathbf{x}}^\infty(\mathbf{y})$  specifies the probability distribution of the fast variables  $\mathbf{y}$  and is stationary on the fast time scale, but implicitly time varying since it is parameterized by the state  $\mathbf{x}$  of the slow variables. In the simplest case  $\rho_{\mathbf{x}}^\infty$  might be proportional to a delta function in  $\mathbf{x}$ , indicating that the fast variables are a direct function of the slow variables. Physically this might correspond to the case where  $\mathbf{y}$  represents a slow amplitude-dependent deformation of the base flow or phase-locked higher harmonics, for instance.

More generally the functional form of this distribution might be complicated, but it is not necessary to specify it in closed form provided the integral in (4.13b) can be evaluated. For instance, when the dynamics is linear–quadratic then (4.13b) simplifies to first and second moments of the distribution. Then the key to the MMR closure for the POD–Galerkin is deriving an approximation of the fast dynamics that is consistent with the original system but allows for computation of these moments in closed form.

As an aside, although the disappearance of the fictitious small parameter  $\epsilon$  is necessary for consistency with the original Galerkin system, it does call into question the validity of the perturbation series approximation (4.7). In this work we do not attempt to make this more rigorous, but instead demonstrate by example that it is a useful heuristic capable of resolving important features of the flow physics. In contrast to typical multiscale modelling applications, the underlying linear–quadratic Galerkin systems often poorly approximate the true dynamics, so it would not be useful to perfectly match the original model even if it was possible to do so.

#### 4.2. Application to the Galerkin system

In order to make practical use of the averaging procedure for the partitioned Galerkin system (4.1), we must be able to perform the integral over the distribution  $\rho_{\mathbf{x}}^\infty(\mathbf{y})$  of the fast variables. This task is somewhat simplified for the case of the linear–quadratic Galerkin dynamics since only the means  $\mathbb{E}[y_\alpha]$  and covariances  $\mathbb{E}[y_\alpha y_\beta]$  are necessary.

The distribution  $\rho_{\mathbf{x}}^\infty(\mathbf{y})$  is the solution for fixed  $\mathbf{x}$  to the steady-state Fokker–Planck equation

$$\mathcal{L}_0^\dagger \rho = -\frac{\partial(\rho f_\alpha^y)}{\partial y_\alpha} + \frac{\sigma_\alpha^2}{2} \frac{\partial^2 \rho}{\partial y_\alpha^2} = 0, \tag{4.14}$$

corresponding to the linear stochastic process

$$\dot{y}_\alpha = \left[ F_\alpha^y + L_{\alpha j}^{yx} x_j + Q_{\alpha j k}^{yxx} x_j x_k \right] + \left[ L_{\alpha \beta}^{yy} + 2Q_{\alpha j \beta}^{yxy} x_j \right] y_\beta + \sigma_\alpha w_\alpha(t). \tag{4.15}$$

The stationary distribution is a multivariate Gaussian with mean and covariance that can be determined from the solution of a Lyapunov equation (Risken 1996), but this would

need to be done at each value of  $\mathbf{x}$ . Instead, we propose the diagonal drift approximation

$$\dot{y}_\alpha \approx \nu_\alpha(\mathbf{x}) (\mu_\alpha(\mathbf{x}) - y_\alpha) + \sigma_\alpha w_\alpha(t), \tag{4.16}$$

for which the stationary distribution is a product of univariate Gaussians solving the one-dimensional Fokker–Planck equation (3.13), i.e.  $y_\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2/2\nu_\alpha)$ . The integrals in (4.13b) can then be evaluated easily using  $\mathbb{E}[y_\alpha] = \mu_\alpha$  and  $\mathbb{E}[y_\alpha y_\beta] = \delta_{\alpha\beta} \sigma_\alpha^2/2\nu_\alpha$ , where  $\delta_{\alpha\beta}$  is the Kronecker delta symbol.

By comparison with (4.15), the conditional mean is naturally defined as

$$\mu_\alpha(\mathbf{x}) = \sum_{j,k=1}^r \nu_\alpha^{-1} \left[ L_{\alpha j}^{yx} x_j + Q_{\alpha jk}^{yxx} x_j x_k \right]. \tag{4.17}$$

Here, we have omitted the contribution of the constant forcing term  $\mathbf{F}^y$  so that the approximate fast process  $\mathbf{y}$  preserves the zero-mean property of the POD coefficients for  $\mathbf{x} = \mathbf{0}$ . In this work we will make the simplifying assumption that the effective damping coefficients  $\nu_\alpha$  are constant, although more generally they could be functions of  $\mathbf{x}$ . Appropriate values for  $\nu_\alpha$  can be determined from energy balance, as described below.

With this approximation, the fast variables  $y_\alpha$  are each an independent Ornstein–Uhlenbeck process with mean  $\mu_\alpha$  and damping  $\nu_\alpha$ . Since the first and second moments of the stationary distribution for this process are given by  $\mu_\alpha$  and  $\sigma_\alpha^2/2\nu_\alpha$ , the average in (4.13) is then a straightforward calculation resulting in the generalized Stuart–Landau model

$$\hat{f}_i(\mathbf{x}) = F_i^1 + L_{ij}^{11} x_j + L_{i\alpha}^{12} \mu_\alpha(\mathbf{x}) + Q_{ijk}^{111} x_j x_k + 2Q_{ij\alpha}^{112} x_j \mu_\alpha(\mathbf{x}) + \nu_\alpha^{-1} Q_{i\alpha\alpha}^{122} \sigma_\alpha \sigma_\alpha / 2 \tag{4.18a}$$

$$\equiv \hat{F}_i + \hat{L}_{ij} x_j + \hat{Q}_{ijk} x_j x_k + \hat{C}_{ijkl} x_j x_k x_l, \tag{4.18b}$$

with the following closed quantities denoted by a hat:

$$\hat{F}_i = F_i^x + \nu_\alpha^{-1} Q_{i\alpha\alpha}^{xyy} \sigma_\alpha \sigma_\alpha / 2 \tag{4.19a}$$

$$\hat{L}_{ij} = L_{ij}^{xx} + \nu_\alpha^{-1} L_{i\alpha}^{xy} L_{\alpha j}^{yx} \tag{4.19b}$$

$$\hat{Q}_{ijk} = Q_{ijk}^{xxx} + \nu_\alpha^{-1} \left( L_{i\alpha}^{xy} Q_{\alpha jk}^{yxx} + 2Q_{ij\alpha}^{xy} L_{\alpha k}^{yx} \right) \tag{4.19c}$$

$$\hat{C}_{ijkl} = 2\nu_\alpha^{-1} Q_{ij\alpha}^{xy} Q_{\alpha kl}^{yxx}. \tag{4.19d}$$

We will refer to the secondary reduction of a standard rank- $r$  POD–Galerkin model to the cubic model (4.18b) as the MMR approach to closure modelling.

The terms in the closure model are computed by summing the Galerkin tensors over the fast variables. As a mnemonic for these averaged quantities, the superscripts for the partitioned system could be thought of as tensor contractions with the modification of the damping  $\nu_\alpha^{-1}$ . For example, the cubic term  $Q_{ij\alpha}^{xy} \nu_\alpha^{-1} Q_{\alpha kl}^{yxx}$  is suggestive of the slow variables ‘filtering through’ the fast dynamics via the quadratic interactions  $\mathbf{Q}^{yxx}$  and  $\mathbf{Q}^{xy}$ .

Although the constant, linear and quadratic terms are all modified as a result of the stochastic averaging procedure, the appearance of this cubic term is perhaps the most noteworthy. It represents the leading-order contribution of the fast–slow nonlinear interaction in the slow dynamics due to the slow–slow interaction in the fast dynamics. We will explore this in more detail in the following section, but it is a generalization of the weakly nonlinear Stuart–Landau mechanism (Landau 1944; Stuart 1958) that is capable of resolving the stabilizing influences of both mean-flow deformation and the energy cascade.



### Multiscale model reduction

The final ingredient in the multiscale closure model (4.18b) is the determination of the damping and diffusion coefficients  $\nu_\alpha$  and  $\sigma_\alpha$ . A reasonable approximation for the diffusion could be the mean of the neglected term  $\sigma_\alpha = \sum_\beta Q_{\alpha\beta\beta}^{yyy} \overline{y_\beta^2}$ , although in our numerical examples we find little difference from neglecting the diffusion altogether. For the damping term we apply an energy-balance condition on the closed model, derived from the assumption that the system is statistically stationary so that the average variation of kinetic energy vanishes (Noack *et al.* 2011)

$$\frac{\partial}{\partial t} \overline{x_i^2} = \overline{\hat{f}_i(\mathbf{x}) x_i} = 0. \quad (4.20)$$

When the closed tensors (4.19) are substituted for  $\hat{f}(\mathbf{x})$ , this simplifies to a linear system of equations for  $\mathbf{v}^{-1}$ . The energy-balance approximation does not require statistics of the fast variables, but due to the cubic term it does require well-converged fourth moments of the slow variables. In particular, the vector  $\mathbf{v}_{inv}$ , which is the element-wise inverse of  $\mathbf{v}$ , is the solution to the linear system

$$\mathbf{A} \mathbf{v}_{inv} + \mathbf{b} = 0, \quad (4.21a)$$

$$\begin{aligned} A_{i\alpha} = & \sum_{j=1}^{r_0} L_{i\alpha}^{xy} L_{\alpha j}^{yx} \overline{a_i a_j} + \sum_{j,k=1}^{r_0} L_{i\alpha}^{xy} Q_{\alpha j k}^{yxx} \overline{a_i a_j a_k} \\ & + 2 \sum_{j,k=1}^{r_0} Q_{ij\alpha}^{xxy} L_{\alpha k}^{yx} \overline{a_i a_j a_k} + 2 \sum_{j,k,\ell=1}^{r_0} Q_{ij\alpha}^{xxy} Q_{\alpha k \ell}^{yxx} \overline{a_i a_j a_k a_\ell} \end{aligned} \quad (4.21b)$$

$$b_i = L_{ii}^{xx} \overline{a_i^2} + \sum_{j,k=1}^{r_0} Q_{ijk}^{xxx} \overline{a_i a_j a_k}. \quad (4.21c)$$

In general, the damping should be positive since the linear stochastic approximation (4.15) and its diagonal simplification are only plausible if the mean  $\boldsymbol{\mu}(\mathbf{x})$  is linearly stable for all  $\mathbf{x}$ . This is physically consistent with the energy cascade picture, in which we expect that energy will primarily flow ‘downhill’ from the slow variables, representing large coherent structures and global instabilities, to the fast variables, representing the smaller, dissipative scales of the flow. On this basis, negative values of the damping coefficients can also be set to zero to avoid introducing unphysical instabilities in the cubic term, for example.

We conclude the discussion of the application to Galerkin-type systems with a note on the computational scaling of the closure model. While the simulation of the original linear–quadratic model is dominated by the quadratic term, which requires  $O(r^3)$  operations to evaluate, the cubic term in the closure model (4.18b) requires  $O(r_0^4)$  operations to evaluate. In some cases this may mean that the cubic model is actually more expensive to simulate than the original Galerkin system, although whether or not this is the case will depend on the specific values of  $r$ , the dimension of the linear–quadratic system and  $r_0$ , the dimension of the closed model. For the examples given in § 6 either  $r_0^4 < r^3$ , as in §§ 6.1 and 6.2, or the two are the same order of magnitude, as in § 6.3. As we will show, the primary advantage of this approach is not necessarily that it is faster to simulate than the usual POD–Galerkin system, but that it is more stable and physically faithful with many fewer modes.

## 5. Flow configurations

In this work we consider models of three fluid flows: the canonical flow past a cylinder at Reynolds number 100, a lid-driven cavity flow and an incompressible mixing layer with two domain extents. We construct projection-based models of these flows based on direct numerical simulation (DNS) using the open-source Nek5000 spectral element solver (Fischer, Lottes & Kerkemeir 2008). The semi-implicit time stepping integrates diffusive terms with third-order backwards differentiation and convective terms with a third-order extrapolation.

Once the DNS is complete, we estimate the POD modes and coefficients using the `modred` library, which provides a set of parallelized, high-performance algorithms in Python for linear modal decompositions and model reduction (Belson, Tu & Rowley 2014). We then compute the components of the Galerkin system (2.5) by extracting the weight matrix and necessary gradients from the POD modes with the built-in Nek5000 post-processor. Both the linear–quadratic Galerkin model (2.5) and the cubic closure (4.18*b*) are simulated using the LSODA time-stepping algorithm available from `scipy` with analytic Jacobian matrices for each system. The initial conditions are the orthogonal projections of the DNS flow fields onto the POD basis.

### 5.1. Flow past a circular cylinder

The vortex shedding in the wake behind a cylinder at Reynolds number 100, based on the free-stream velocity and cylinder diameter, is a canonical flow configuration for reduced-order modelling (Noack *et al.* 2003) and it is shown in figure 3. We simulate this flow on a domain of 2600 sixth-order spectral elements on  $x, y \in (-5, 15) \times (-5, 5)$  refined close to the cylinder wall; further details and analysis can be found in Loiseau, Brunton & Noack (2018*a*) and Loiseau *et al.* (2018*b*). We first perform a global stability analysis of the the unstable steady state, determined by solving the stationary Navier–Stokes equations with selective frequency damping (Åkervik *et al.* 2006), using a Krylov–Schur time-stepping algorithm (Loiseau *et al.* 2019). The transient simulation is then initialized with the unstable steady state perturbed by the least-stable global eigenmode, normalized so that its energy is  $10^{-4}$ .

The flow regime at Reynolds number 100 is beyond the Hopf bifurcation leading to vortex shedding but below the onset of three-dimensional instability (Noack & Eckelmann 1994; Barkley & Henderson 1996; Ma & Karniadakis 2002). The flow is also far enough from the threshold of bifurcation that weakly nonlinear analyses fail to predict the amplitude and frequency of the vortex shedding limit cycle, although a variety of techniques have been used to derive accurate models for the transient and post-transient wake (Noack *et al.* 2003, 2011; Mantič-Lugo *et al.* 2014; Loiseau & Brunton 2018; Loiseau *et al.* 2018*b*).

Following Noack *et al.* (2003) and Loiseau *et al.* (2018*a*), we compute POD modes from 100 snapshots of the mean-subtracted velocity field, equally spaced over one period of vortex shedding; the pressure gradient term can be shown to be negligible for Galerkin models of this flow. In order to construct a basis that can span the deformation between the unstable steady state and mean flow, we augment the POD basis with the ‘shift mode’, visualized in figure 3. This additional mode is computed from the difference between the base and mean flows and then orthonormalized with respect to the rest of the POD modes with the Gram–Schmidt process. We denote the coefficient of this mode by  $a_{\Delta}(t)$  to distinguish it from the usual POD modes.

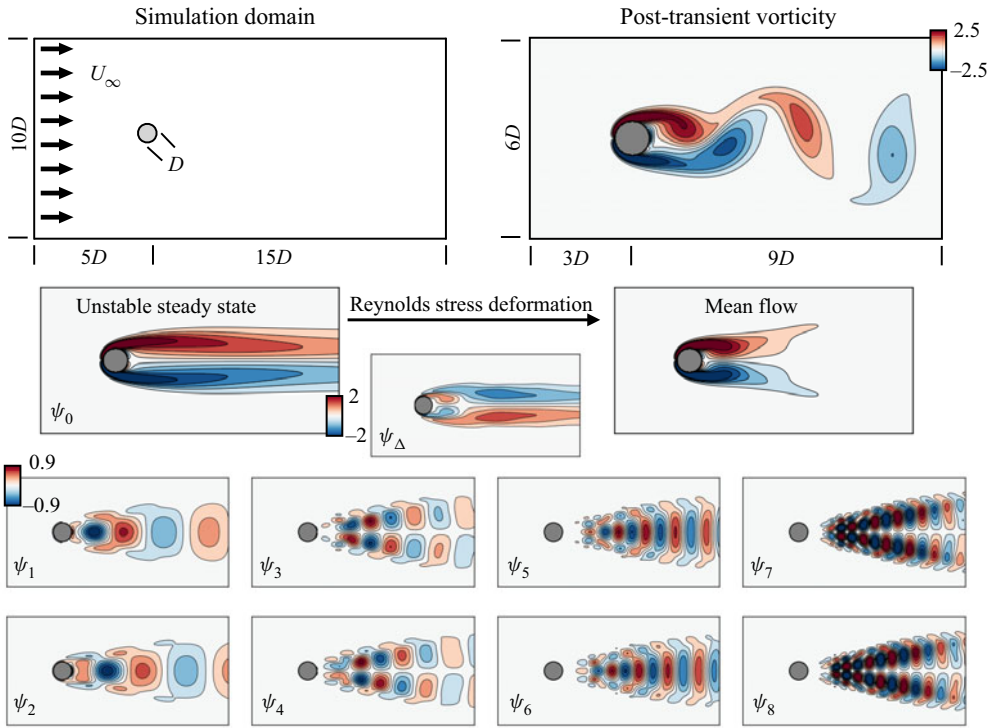


Figure 3. Vortex shedding in the wake of a circular cylinder at  $Re = 100$ . En route to the periodic post-transient flow, Reynolds stresses deform the background flow from the unstable steady state to the marginally stable mean flow. This can be approximated in a model by augmenting the POD basis with the ‘shift mode’  $\psi_\Delta$ . The usual POD modes show the typical structure of spatial and temporal harmonics describing periodically advecting flow features.

The full expansion basis then consists of eight POD modes capturing the limit-cycle vortex shedding, plus the shift mode. We take the unstable steady state to be the base flow of the modal expansion so that the origin is a fixed point of the POD–Galerkin system. This basis accurately represents the kinematics of the post-transient flow, with a normalized energy residual of  $O(10^{-4})$  (Loiseau *et al.* 2018a).

### 5.2. Lid-driven cavity flow

Lid-driven cavity flow is another idealized geometry that serves as a benchmark problem for numerical methods and model reduction (Arbabi & Mezić 2017; Arbabi & Sapsis 2022; Lee, Dowell & Balajewicz 2019; Rubini, Lasagna & Ronch 2020). The flow, shown in figure 4, takes place on a square domain with  $(x, y) \in (-L, L) \times (-L, L)$  with  $L = 1$ . A velocity profile is imposed on the upper boundary, given by

$$u(x, L) = U(1 - x^2)^2. \tag{5.1}$$

This driving velocity profile roughly approximates the shear-driven flow over an open cavity at a much lower computational cost, although the dynamics of the shear- and lid-driven cavities is different. The regularized fourth-order velocity profile on the lid avoids numerical problems where the lid meets the no-slip boundaries on the other walls.

The standard Reynolds number for this flow is defined as  $Re = UL/\nu$  based on the cavity half-height and forcing velocity. The flow follows the stereotypical Ruelle–Takens route

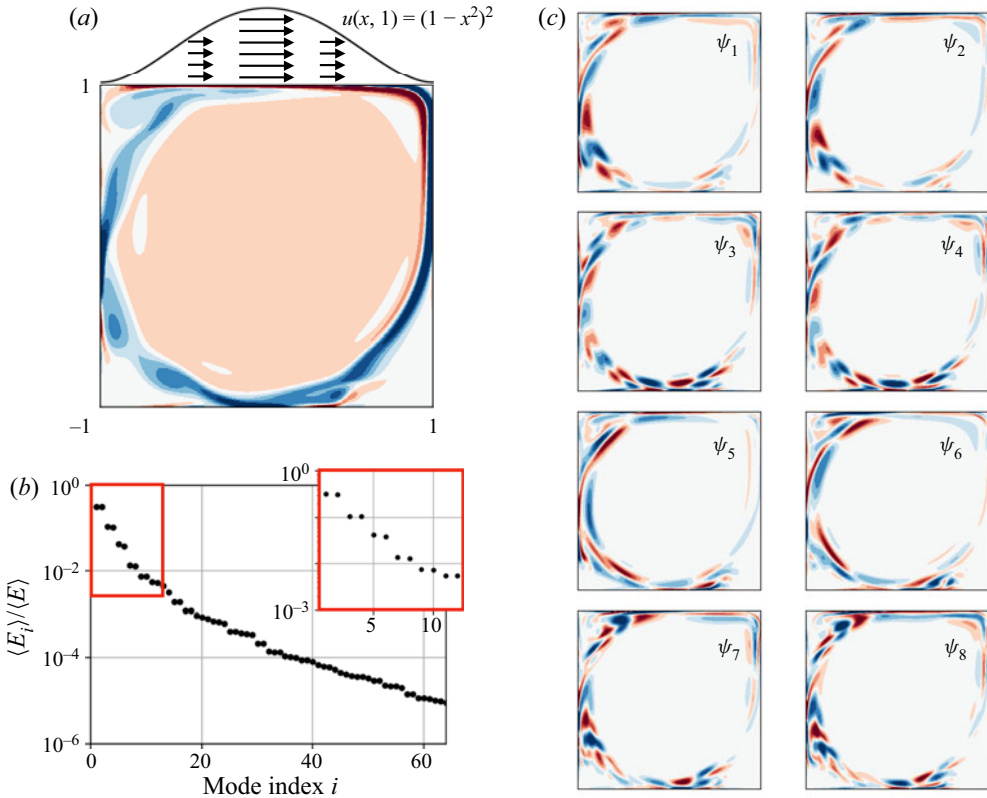


Figure 4. Regularized lid-driven cavity flow at  $Re = 20000$ . The singular value spectrum is slow to converge (a,b), indicating that relatively many modes are necessary for an accurate kinematic approximation. The leading POD modes themselves appear roughly in pairs with similar spatial structure and frequency content (c). As for the cylinder wake shown in figure 3, this usually indicates oscillatory dynamics, although the evolution of the temporal coefficients is much more irregular in this chaotic flow than for the periodic wake.

to chaos parameterized by Reynolds number, undergoing a Hopf bifurcation at  $Re_c^{(2)} \approx 10\,250$ , a secondary Hopf bifurcation to quasiperiodic flow at  $Re_c^{(2)} \approx 15\,500$ , and a final bifurcation to chaos near  $Re_c^{(3)} \approx 18\,000$  (Lee *et al.* 2019). We simulate the flow in the chaotic regime at  $Re = 20\,000$ .

The numerical configuration is similar to that described by Arbabi & Sapsis (2022). The domain is discretized with 50 seventh-order spectral elements in each direction, refined towards the walls. The flow is integrated for 3000 time units, saving snapshots at a sampling rate  $\Delta t = 0.1$ . Again, the pressure gradient term vanishes for this closed flow (Noack *et al.* 2005), so the velocity POD modes are computed from the first 5000 snapshots, with the remaining 25000 snapshots retained for the statistical comparison in § 6.2. Figure 4 shows a snapshot of the DNS along with the leading POD modes and the singular value spectrum.

### 5.3. Incompressible mixing layer

As a more difficult test of the proposed MMR closure method, we also examine models of an incompressible mixing layer, pictured in figure 11. The mixing layer is a canonical example of a free shear flow exhibiting convective instability (Monkewitz & Huerre 1982;

Huerre & Monkewitz 1985, 1990). Inlet forcing is used to excite Kelvin–Helmholtz instability waves, which roll up into approximately discrete vortices. These vortices are subject to a secondary instability (Brancher & Chomaz 1997) and undergo successive helical pairing at the subharmonic of the primary Kelvin–Helmholtz frequency, a process that drives the linear growth of the mixing layer (Winant & Browand 1974). At higher Reynolds numbers and in three dimensions, the turbulent mixing layer is dominated by the linear growth of coherent structures similar to those seen in two-dimensional simulation (Brown & Roshko 1974; Stanley & Sarkar 1997), which play a major role in mixing, transport and entrainment of the turbulent flow (Hussain 1981). From a global perspective, the flow behaves as an amplifier, where strong non-normality in the linear operator results in transient algebraic energy growth of small perturbations (Chomaz 2005).

By analogy to numerical methods, modelling an advection-dominated flow with a POD–Galerkin approach is generally ill advised in much the same way that pseudospectral methods are not ideal for solving hyperbolic (advection) equations. Although the vortices shed in the flow past a cylinder are also travelling waves in a sense, the Kelvin–Helmholtz instability in the mixing layer is convective in nature in contrast with the absolute instability in the cylinder wake. For these reasons, Galerkin projection onto a global modal basis such as POD is not a natural way to approximate shear flows like the mixing layer. Nevertheless, both free and wall-bounded shear layers are present in a wide variety of flows, so in order to employ a low-dimensional dynamical systems model it is important that it be able to capture these features. The challenge of modelling advection-driven phenomena in a global reduced-order modelling framework motivates the use of similar configurations in several studies of stabilization methods for Galerkin-type systems (Ukeiley *et al.* 2001; Noack *et al.* 2005; Balajewicz, Dowell & Noack 2013; Cordier *et al.* 2013).

Since the flow becomes increasingly complex downstream, the difficulty of modelling the flow can be tuned by varying the streamwise extent of the domain. In contrast, the observed dynamics of the cylinder wake and lid-driven cavity is not strongly dependent on domain extent. We construct models on both short ( $L_x = 150$ ) and long ( $L_x = 250$ ) domains, as shown in figure 5. Over the extent of the short domain, the flow is mainly characterized by the initial instability and vortex rollup, with the earliest signs of vortex pairing at  $x \gtrsim 130$ . On the full domain, the vortex pairing accounts for a much larger fraction of the fluctuation kinetic energy, as shown by the POD analysis below.

The inlet profile is a standard hyperbolic tangent

$$U(y) = \bar{U} + (\Delta U/2) \tanh 2y, \quad (5.2)$$

where  $\bar{U} = (U_1 + U_2)/2 \equiv 1$  and  $\Delta U = U_1 - U_2$  in terms of the free-stream velocities  $U_1$  and  $U_2$  above and below the layer, respectively, and the length scale is non-dimensionalized by the initial vorticity thickness  $\delta_\omega = \Delta U/U'(0)$ . We define a Reynolds number  $Re = \delta_\omega(0)\Delta U/\nu$  based on initial vorticity thickness, velocity difference  $\Delta U$  across the layer and kinematic viscosity  $\nu$  and set  $Re = 500$  in the simulation. The domain consists of 5200 eleventh-order spectral elements on  $x, y \in (-10, 300) \times (-50, 50)$ , but the full streamwise extent is masked to  $x, y \in (0, 250) \times (-20, 20)$  for modelling in order to discount any boundary effects. The domain is shown in figure 11.

This configuration roughly follows that described by Balajewicz *et al.* (2013) and Cordier *et al.* (2013), with two key differences. First, we simulate incompressible flow, while these authors simulate isothermal compressible flow at low Mach numbers ( $Ma \sim 0.1$ ). Consequently, we replace the ‘sponge’ boundary conditions for far-field acoustic



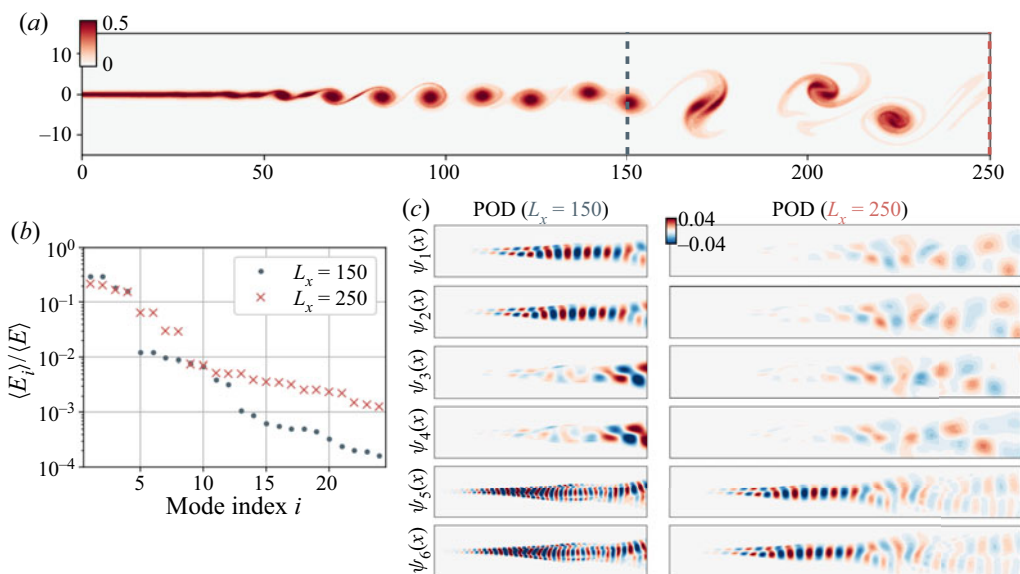


Figure 5. Proper orthogonal decomposition applied to the mixing layer. The primary shear layer instability forms Kelvin–Helmholtz waves that roll up into vortices, which in turn undergo successive vortex pairing (a). Modes computed on both the short and long domains (c) reveal modes related to the dominant flow features: the shear layer instability and the downstream vortex pairing. Although the vortex pairing is secondary to the upstream instability, on the longer domain it accounts for a larger proportion of the fluctuation kinetic energy (b). Higher-order modes not pictured here include harmonics, nonlinear cross-talk and modes related to irregularity in the location of the vortex pairing events.

waves with a region of increased viscous damping corresponding to  $Re = 50$  for  $x > 250$  to avoid numerical instability at the outflow. Second, rather than disturbing the inlet with random solenoidal perturbations, we use eigenfunction forcing described in Appendix A, similar to that employed by Colonius, Lele & Moin (1997).

The DNS is run until a final time of  $t = 2820$ , corresponding to fifty periods of the lowest-frequency component of the eigenfunction forcing, saving at every tenth time step with  $\Delta t = 0.00705$ . The POD modes are then computed from the first 10% of these snapshots (4000 fields), with the remainder reserved for statistical comparison with the models. The domain truncation can be accomplished by simply setting elements of the weight matrix  $\mathbf{W}$  in (2) to zero for mesh locations with  $x > L_x$  so that these locations do not contribute to the statistics in the correlation matrix. For the mixing layer we compute modes with both velocity and pressure fields, as described in § 2, since the pressure gradient term is not necessarily negligible in Galerkin models of free shear flows (Noack *et al.* 2005).

Figure 5 shows the leading POD eigenvalues and modes for both domains. In both cases the bulk of the fluctuation kinetic energy is contained in the first four modes, with the remainder of the spectrum decaying relatively slowly. As expected, the dominant mode pair for the shorter domain corresponds to the Kelvin–Helmholtz waves, while the second pair represents the onset of vortex pairing at roughly half the spatial wavenumber. Higher-order modes are either harmonics (such as the third mode pair) or less regular downstream structures related to slight variations in the vortex pairing.

The directed graph of energy transfers in figure 1 is constructed from the modes computed on the short domain. Specifically, the leading mode pair and its first three harmonics are shown for a simple visualization of the energy cascade; the nodes



labelled  $a_1$ – $a_8$  thus actually correspond to mode pairs  $(\varphi_1, \varphi_2)$ ,  $(\varphi_5, \varphi_6)$ ,  $(\varphi_{17}, \varphi_{18})$  and  $(\varphi_{24}, \varphi_{25})$ . The thickness of the edge from node  $a_i$  to  $a_j$  is proportional to  $\sum_k Q_{ijk} \overline{a_i a_j a_k}$  for the quadratic Galerkin tensor  $\mathbf{Q}$ , which is roughly related to average nonlinear energy transfer from one mode to the other, although it should not necessarily be interpreted in any rigorous way beyond the purposes of conceptual visualization of the energy cascade.

On the longer domain the vortex pairing makes up a larger proportion of the fluctuation kinetic energy. Consequently, the leading four modes are spatially localized downstream, representing vortex pairing, while the upstream linear instability is not evident until the third mode pair, even though the vortex pairing is secondary to the shear layer instability from a physical perspective.

The slowly decaying POD spectrum and wave-like spatial structure of the modes are a consequence of the advection-driven nature of this flow. It is well known that the space–time separation of variables assumed by POD is not a natural representation of travelling waves. This is one reason that free shear flows have long posed difficulties for POD–Galerkin modelling, making this a challenging test case for the MMR approach.

## 6. Results

In this section we evaluate the proposed MMR closure method on several numerical test problems. In § 6.1 we analyse vortex shedding behind a circular cylinder, showing that the method reproduces the classic invariant manifold model due to Noack *et al.* (2003), which must account for both mean-flow deformation and the energy cascade in order to accurately model the transient dynamics. As a more challenging test in § 6.2, we show that the multiscale closure stabilizes a reduced-order model of the chaotic lid-driven cavity flow, while closely matching the spectral energy content of the DNS. Finally, in § 6.3 we develop a model for convective instability and subharmonic vortex pairing in the spatially evolving mixing layer.

By means of these examples, we show that the proposed approach for multiscale closure is a systematic method for eliminating degrees of freedom from reduced-order models that can resolve critical subscale phenomena and produce stable models of dynamically complex unsteady flows. More broadly, the multiscale framework represents a general description of the means by which effective cubic nonlinearity arises in low-dimensional models.

As a general comment on the following figures, in order to evaluate the quality of the dynamical approximation (rather than that of detailed time-series forecasts) we choose visualizations that capture the overall statistical behaviour of the system more broadly than time series. In particular, we focus on fluctuation energy levels, estimated power spectral densities and phase portraits, which may be interpreted as approximate slices of joint probability distributions. However, for the sake of completeness we provide a comparison of forecasting errors in [Appendix C](#).

### 6.1. Vortex shedding in the cylinder wake

In this section we revisit the canonical two-dimensional flow past a circular cylinder to show that the MMR closure can naturally resolve the stabilizing effects, both of mean-flow deformation and of energy transfer to higher-order modes, without assuming proximity to a bifurcation.

In keeping with previous work demonstrating that the cylinder wake can be modelled accurately with as few as two degrees of freedom, we reduce the 9-mode POD–Galerkin system to a 2-mode generalized Stuart–Landau model (4.18*b*) with the closure model (4.19)

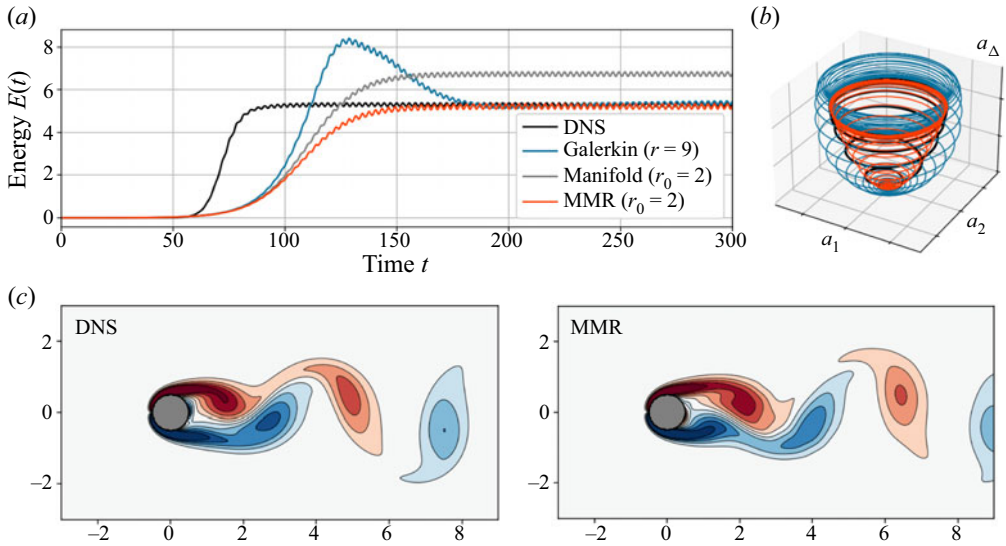


Figure 6. Reduced-order models of the cylinder wake. The standard 9-mode Galerkin model accurately estimates the stable limit cycle (a), but the transient dynamics deviates from the slow manifold, leading to an energy overshoot (b). Both the two-dimensional invariant manifold reduction (Noack *et al.* 2003) and MMR closure models prevent this by eliminating the rapidly equilibrating variable associated with mean-flow deformation, but MMR more accurately estimates the limit-cycle amplitude. The full flow field can be reconstructed by approximating the harmonics with polynomial regression (c).

and damping derived from the energy-balance relation (4.20). The first 9 modes capture >99.9% of fluctuation kinetic energy for this flow. In this case, the elimination of the higher-order modes does not require sacrificing kinematic resolution, since the unresolved coefficients can be approximated as sparse polynomial functions of the resolved variables (Loiseau *et al.* 2018a; Callaham *et al.* 2022).

Figure 6(a) compares the energy of the first mode pair,  $E(t) = a_1^2(t) + a_2^2(t)$ , predicted by the POD–Galerkin model and the 2-mode closure model, with DNS of the transient flow initialized from the unstable steady state and perturbed by the least-stable eigenmode. Also included for comparison is the invariant manifold model introduced by Noack *et al.* (2003). Although the Galerkin model is stable and accurately predicts the amplitude of the vortex shedding, it exhibits a transient energy overshoot before settling onto the limit cycle. Figure 6(b) shows the reason for this; whereas the DNS is restricted to an approximately parabolic invariant manifold, the finite relaxation time in the shift mode dynamics results in the energy of the leading modes continuing its exponential growth for longer before the energy can be absorbed into the mean-flow deformation.

In contrast, the cubic term in the MMR and invariant manifold models approximates the relaxation of the shift mode dynamics as instantaneous, pinning its amplitude to that of the primary mode pair. While the structure and interpretation of the two models is the same, there are quantitative differences in their accuracy. In particular, the manifold model slightly overestimates the energy of the limit cycle, with mode amplitudes larger than the DNS by  $\sim 13\%$ , while the multiscale closure accurately predicts the limit-cycle amplitude. This discrepancy can most likely be attributed to the energy-balance procedure used to estimate the subscale damping  $\nu$  in the MMR model, which is not used in the invariant manifold model.

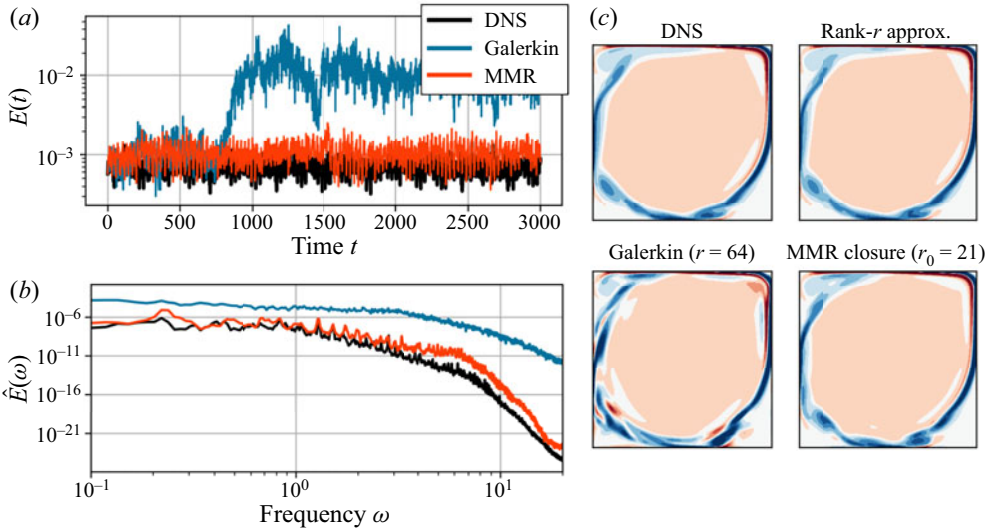


Figure 7. Reduced-order models of the lid-driven cavity flow. The Galerkin system eventually equilibrates at much higher energy levels, while the MMR closure model produces stable predictions that closely match the power spectrum of the chaotic flow (a,b). The reconstructed fields from the MMR model are also more physically consistent with the DNS than the Galerkin model (c). For comparison, the top right panel shows the reconstruction of the DNS flow field using an equal number of modes as the POD–Galerkin system ( $r = 64$ ).

Finally, the full flow field can also be reconstructed from the two-degree-of-freedom multiscale model by means of the sparse polynomial approximation to the nonlinear correlations in the higher-order modes (figure 6c). The flow field predicted by the model gives a good approximation of the DNS solution, although over long times the phase of the vortex shedding tends to drift from that of the high-dimensional solution. By systematically averaging over the fast variables, the multiscale closure method is thus able to reduce the dimensionality of the model while also improving its fidelity.

### 6.2. Chaotic lid-driven cavity flow

The Stuart–Landau equation derived for the cylinder wake in § 6.1 is an example of a relatively simple nonlinear oscillator with either a stable fixed point or periodic limit cycle as an equilibrium. Here, we show that the MMR closure approach can also be applied to systems with a more complex dynamics, as illustrated on the chaotic lid-driven cavity flow introduced in § 5.2.

First, we construct a standard POD–Galerkin model with  $r = 64$  modes, sufficient to capture  $\sim 99.8\%$  of the fluctuation kinetic energy. The Galerkin system is chaotic and energetically stable. Despite its precise kinematic resolution, the Galerkin model is inaccurate, overestimating the kinetic energy by roughly one order of magnitude, as shown by the blue traces in figure 7.

Reducing the model to  $r_0 = 21$  degrees of freedom via a multiscale closure stabilizes it with an equilibrium energy that approximately matches the power spectrum of the DNS (figure 7, orange). The MMR model still retains enough modes for a  $\sim 98\%$  accurate POD reconstruction. The predicted flow fields are coherent and qualitatively consistent with the DNS fields, even though they do not match in every detail due to the chaotic nature of the flow.

Varying the dimension  $r_0$  of the approximate slow variables results in models that are energetically stable, but not necessarily chaotic. Instead, many such models eventually settle onto a post-transient dynamics that is periodic or quasi-periodic. The persistence of the dominant frequencies suggests that, even for models with relatively many degrees of freedom ( $r_0 > 2$ ), the generalized Stuart–Landau equations produced by the MMR closure scheme can be viewed as coupled nonlinear oscillators, especially when the spatial components of the POD modes suggest that they may be grouped in pairs with similar temporal frequency content.

### 6.3. Convective instability and vortex pairing in a mixing layer

Advection-dominated flows have historically posed a challenge for reduced-order models, primarily because the global mode ansatz of POD is a poor representation of travelling waves. As a final numerical example, we apply the MMR closure to model the mixing layer introduced in § 5.3, a canonical example of convective instability in free shear flows.

In one sense, the dynamics of the mixing layer is relatively simple. The flow close to the inlet is dominated by the shear layer instability and subsequent vortex roll-up. These vortices are advected downstream at approximately the midline flow velocity, until the secondary vortex pairing begins at the subharmonic of the primary instability frequency. However, since this behaviour unfolds as the disturbances are carried downstream, a principal challenge for low-dimensional models based on global POD modes is to preserve phase coherence between the harmonics and subharmonics, so that the reconstructed flow fields consist of approximately discrete vortices undergoing the vortex pairing.

As discussed in § 5.3, both the POD mode basis and the difficulty of modelling this flow depend on the streamwise extent of the domain. Since the subharmonic vortex pairing is energetically dominant downstream, the leading POD modes computed from a longer domain will be related to vortex pairing, even though the primary shear layer instability drives the dynamics from a physically causal perspective. To illustrate this effect, we consider two domain lengths: one that is dominated by the primary instability and vortex roll-up, and the full modelling domain with vortex pairing.

In this case the POD–Galerkin models do not necessarily improve with an increasing number of modes. For both domains, the best dynamical approximation is found for models constructed from  $r = 24$  modes, corresponding to 99.8% of fluctuation kinetic energy on the short domain and 98.4% on the long domain. On the short domain ( $L_x = 150$ ), this model is energetically stable but eventually approaches an unphysical fixed point (figure 8). On the longer domain ( $L_x = 250$ ), the model initially follows the DNS trajectory before eventually becoming unstable (figure 9). In either case, increasing the dimension of the POD–Galerkin model resulted in worse models. We then reduce the 24-dimensional Galerkin systems to  $r_0 = 16$ -dimensional generalized Stuart–Landau models with the MMR closure.

Figure 8 shows the comparison of the closed model with the DNS and Galerkin systems for the short domain. The MMR model closely matches the phase portraits of the DNS, indicating that it preserves the phase relationships between the various modes, including the early subharmonic vortex pairing ( $a_3, a_4$ ) and higher harmonics, e.g. ( $a_5, a_6$ ). This remains true even when integrated to very long times (the final integration time is  $t = 2820$  in this case, as with the DNS described in § 5.3). As a result, the approximated fields are highly coherent, with the same pattern of vortex roll-up, advection, and pairing as seen in the DNS (figure 8*b–d*).

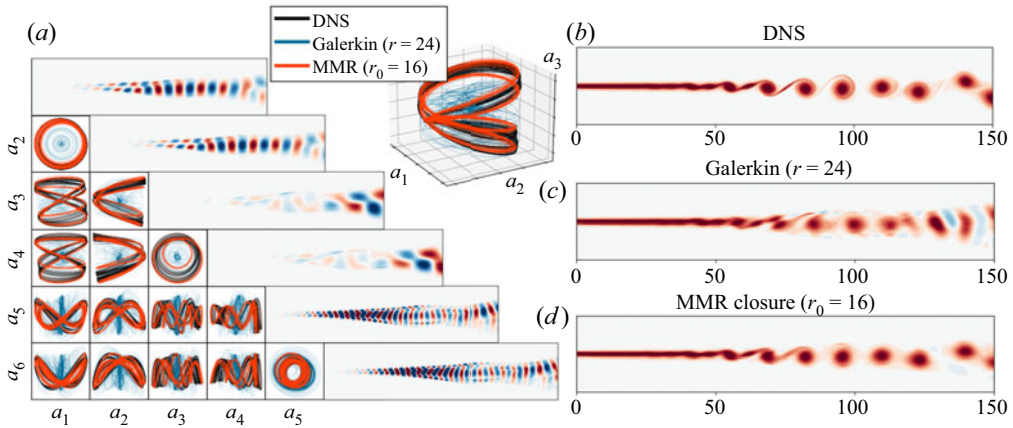


Figure 8. Models of the mixing layer on a short domain. The phase portraits of the mode pairs (a) show phase locking between the fundamental mode pair ( $a_1, a_2$ ), the subharmonic vortex pairing ( $a_3, a_4$ ), and the first harmonic ( $a_5, a_6$ ). These phase relationships are preserved by the MMR closure, resulting in physically consistent flow field estimates (b–d).

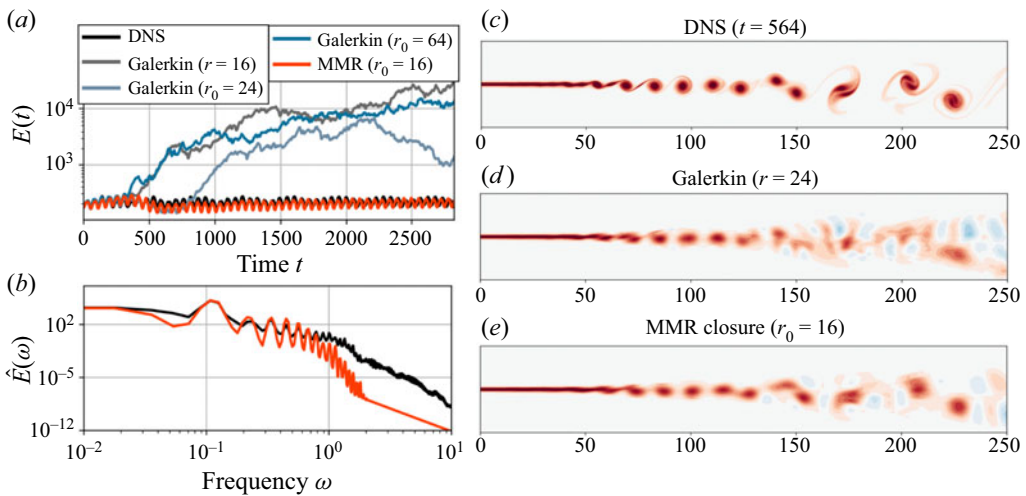


Figure 9. Mixing layer models on the long domain. While standard POD–Galerkin models are energetically unstable until at least  $r = 64$ , the multiscale closure stabilizes the model with only 16 variables (a,b). The model also remains coherent on long integration times (see also figure 10), producing flow field predictions that are consistent with the large-scale structure of the DNS (c–e).

On the longer domain, the multiscale closure also stabilizes the Galerkin model. Figure 9 compares the true energy content of the POD coefficients with that predicted by the model, both in the time and frequency domain. Although the evolution of the energy matches at the dominant frequencies, the MMR model does not capture the weakly chaotic dynamics exhibited by the DNS as a result of slight irregularity in the vortex pairing. Instead, after a brief transient the solution becomes approximately periodic with phase locking between the modes (figure 10). As with the short domain, this synchronization ensures that the predicted flow fields remain coherent, even when integrated for long times (figure 9c–e).

As might be anticipated based on the more complex dynamics on the longer domain, the fields predicted by the MMR model do not match the DNS quite as closely as on the



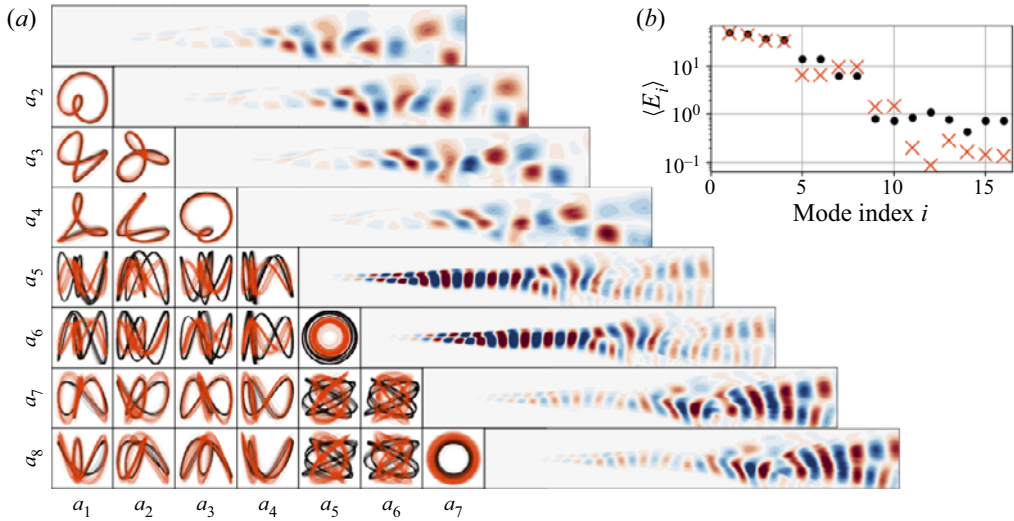


Figure 10. Phase portraits of the mixing layer model on the long domain. The MMR closure (orange) closely matches the DNS trajectories (black) for the most energetic modes, even though the structure is more complicated than the Lissajous-type harmonics on the short domain (figure 8). This figure does not show a Galerkin model for comparison because they are energetically unstable (figure 9).

shorter domain, particularly closer to the inlet. This is likely because fine-scale resolution of the upstream region of the flow requires higher harmonics of the primary instability that do not appear in the POD modes of the longer domain until much higher-order mode number. Higher-resolution predictions might therefore be constructed by applying nonlinear correlations analysis and polynomial regression to selected higher-order modes, as in the cylinder reconstruction in § 6.1; see also Loiseau *et al.* (2018a) and Callaham *et al.* (2022). Still, the dominant vortices are still evident in the prediction of the low-dimensional model, even in the downstream pairing region.

On both domains, the synchronization and long-time phase coherence between the modes further reinforces the picture outlined in § 6.2 of the closure model as consisting of coupled nonlinear oscillators. Although the description of this flow in terms of spatially fixed global modes is not an ideal representation of the advection-driven dynamics, the proposed MMR closure method is able to successfully stabilize the POD–Galerkin models and accurately capture the phase relationships necessary for accurate predictions of the full flow fields.

### 7. Conclusion

In this work, we have developed a MMR approach to improving linear–quadratic dynamical systems derived from POD–Galerkin projection of the Navier–Stokes equations, with the addition of systematically computed cubic closure terms. These cubic closure terms are derived through an adaptation and application of a multiscale stochastic averaging method that originated in singular perturbation theories of Markov processes (Kurtz 1973; Papanicolaou 1976; Majda *et al.* 2001; Weinan 2011; Pavliotis & Stuart 2012). Whereas the standard truncation of the Galerkin system disregards the influence of unresolved variables, the proposed MMR method accounts for their effect through averaging via a stochastic Koopman operator. In particular, this approach is able to model nonlinear interactions between resolved and unresolved variables, capturing key



mechanisms such as mean-flow deformation and the energy cascade. The closed model includes cubic terms, taking the form of generalized Stuart–Landau equations that often act as coupled nonlinear oscillators.

Since the derivation of the MMR closure method relies on an asymptotic time-scale separation between resolved and subscale variables that is generally absent in fluid flows, we have not attempted to prove the general validity of this approach when the flow is not close to a bifurcation. Instead, we demonstrated by numerical example that the method dramatically improves the stability and accuracy of low-dimensional models of unsteady fluid flows compared with the standard POD–Galerkin models. After a comparison with the well-known benchmark problem of vortex shedding in a cylinder wake, we have shown that the MMR method can reproduce the chaotic dynamics in lid-driven cavity flow. Finally, we developed models of an incompressible mixing layer and showed that the coupled-oscillator MMR closures not only stabilize the models, but preserve phase relationships between the modes, which is critical for physically consistent predictions of advection-driven flow.

These results raise several possibilities for interesting future work. For example, the mixing layer is an ‘amplifier’ flow characterized by convective instability and transient energy growth. It is therefore highly sensitive to the nature of the upstream flow; models developed for a specific inlet forcing may not generalize. Further consideration of this issue could be the foundation of a multiscale approach to input/output models in the resolvent framework (McKeon & Sharma 2010; Sharma & McKeon 2013; Zare, Jovanović & Georgiou 2017; Padovan *et al.* 2020; Pickering *et al.* 2021). In this work we neglect the time-varying boundary condition in the region upstream of the modelling domain, but an input/output model could more properly treat this explicitly as forcing, as is done for instance in control-oriented reduced-order models (Barbagallo, Sipp & Schmid 2009; Sipp *et al.* 2010). Recent work on trajectory-based optimal oblique projections has also shown promise for capturing dynamically important, but typically low-energy, modes in an input–output framework for model reduction of open shear flows (Otto, Padovan & Rowley 2022).

Amplifier flows are therefore one potential failure mode for the MMR method as presented here. This may be understood as a reflection of the unsuitability of standard POD–Galerkin for input–output modelling. By extension, without further development MMR would also most likely struggle to model configurations with time-varying base flows or actuation. As a more fundamental limitation, this approach does depend on the presumption of underlying low-dimensional structure in the dynamics of the flow. As a limiting case, MMR would likely fail to produce useful models of homogeneous isotropic turbulence.

As discussed for the chaotic lid-driven cavity flow in § 6.2, we have observed that the MMR models often tend towards phased-locked periodic or quasiperiodic solutions rather than chaos, consistent with the idea of a system of weakly coupled oscillators. However, we expect that in the development of models of chaotic or turbulent flows it may be helpful to re-introduce some degree of stochasticity to better match the flow statistics. Fortunately, this is straightforward in the multiscale modelling framework (Weinan 2011; Pavliotis & Stuart 2012). For instance, the homogenization method (Majda *et al.* 2001) begins from slightly different scaling assumptions and results in a closed model in which the fast-scale stochastic forcing appears as a new diffusion term. A careful investigation of different scale separation assumptions and stochastic approximations could be valuable in extending this framework to turbulent flows. The diagonal drift approximation could also be revisited; numerical estimation of the ‘stationary’ distribution of the fast Ornstein–Uhlenbeck

process by solving the Lyapunov equation at each time step may improve the accuracy of averaging enough to justify the additional computational cost.

Moreover, since successful POD–Galerkin models of chaotic or turbulent flows often depend on explicit models of mean-flow deformation (Aubry *et al.* 1988; Ukeiley *et al.* 2001) or subscale dissipation (Rempfer & Fasel 1994; Noack *et al.* 2008; Wang *et al.* 2012; Cordier *et al.* 2013; Östh *et al.* 2014), both of which can be approximated naturally in the multiscale modelling framework, the MMR approach may be useful as a general approach to subscale and mean-flow modelling that is independent of the flow geometry.

The potential application to flows with continuous spectral components still relies conceptually on the existence of low-dimensional structure in the flow, although the success of filtering in flows with broadband turbulence to perform large-eddy simulation (LES) raises the possibility of a subscale LES-type closure model based on multiscale analysis. This might take the form of a multiple-scale expansion in both space and time or of a filter designed for the isotropic scales of turbulence, for instance. In either case, the Stuart–Landau form of the MMR closure model suggests that a spatio-temporal analysis would be akin to a Ginzburg–Landau theory of turbulence modelling.

In this work we have only considered a linear–quadratic dynamics derived from POD–Galerkin projection. This has the advantage of being based on the governing equations, but recent work has shown that data-driven model discovery can be a powerful alternative (Brunton *et al.* 2016; Loiseau & Brunton 2018; Loiseau *et al.* 2018*a,b*; Deng *et al.* 2020; Qian *et al.* 2020; Rubini *et al.* 2020; Callaham *et al.* 2022). Data-driven modelling is especially useful when the governing equations are unknown or projection-based modelling is not as straightforward, as for incompressible fluid flow (Loiseau 2020; Qian *et al.* 2020; Guan, Brunton & Novosselov 2021). In a sense, data-driven modelling might circumvent the need for closure modelling by fitting directly to the time series, including leveraging the intuition that cubic terms will appear in the effective dynamics (Loiseau & Brunton 2018). However, the linear–quadratic system has certain symmetries and conservation properties that can be enforced in constrained regression to a quadratic model (Schlegel & Noack 2015; Loiseau & Brunton 2018; Kaptanoglu *et al.* 2021*a,b*), while extending these properties to cubic nonlinearity is not straightforward. An interesting future direction could be exploring a two-step process, first identifying a quadratic model that satisfies the constraints, and then applying the MMR closure to reduce it to a lower-dimensional cubic model. Since the empirical models are often sparse, this could potentially result in models with improved stability and scaling compared with the projection-based approach. Alternatively, a more flexible deep-learning model could be used to replace the proposed diagonal drift approximation, for instance with a modified variational autoencoder (Kingma & Welling 2013) that parameterizes the distribution of subscale variables conditioned on the slow variables.

It will also be interesting to extend this analysis to systems where the leading-order cubic terms tend to destabilize the system, such as the subcritical bifurcation in plane Poiseuille flow (Stewartson & Stuart 1971). This is an important scenario in the transition to turbulence, particularly for shear flows in which non-normal energy amplification can activate the nonlinearity even when the flow is globally stable. Higher-order effective nonlinearity could potentially be incorporated into the framework via state-dependent diffusion or an alternative fast drift model, for instance.

Beyond the context of POD–Galerkin reduced-order modelling, the MMR framework represents a general theory of effective cubic nonlinearity arising in amplitude equations for global modes in quadratically nonlinear fluid flow. The system of generalized Stuart–Landau equations produced by the closure model presents a picture of the flow as

a set of coupled nonlinear oscillators describing the evolution of mode pairs. Close to the threshold of instability the MMR closure is consistent with a weakly nonlinear analysis, but it does not rely on proximity to a bifurcation. The multiscale modelling methodology is therefore an alternative to standard asymptotic expansions that may have implications for a variety of theoretical and numerical approaches to modelling the large, slow scales of unsteady fluid flows.

**Funding.** S.L.B. acknowledges funding support from the Army Research Office (ARO W911NF-19-1-0045), the Air Force Office of Scientific Research (AFOSR FA9550-21-1-0178) and National Science Foundation AI Institute in Dynamic Systems (grant number 2112085). J.L.C. acknowledges funding support from the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program.

**Declaration of interests.** The authors report no conflict of interest.

**Author ORCIDs.**

 Jared L. Callaham <https://orcid.org/0000-0002-1972-1233>;

 Jean-Christophe Loiseau <https://orcid.org/0000-0002-7244-8416>;

 Steven L. Brunton <https://orcid.org/0000-0002-6565-5118>.

**Appendix A. Stability analysis of the mixing layer inlet profile**

This appendix describes the eigenfunction forcing applied at the inlet of the mixing layer configuration described in § 5.3. The flow acts an amplifier (Huerre & Monkewitz 1985, 1990; Chomaz 2005), so the downstream flow is highly sensitive to the inlet conditions; eigenfunction forcing accentuates the natural dynamics of the flow.

Since the dominant Kelvin–Helmholtz instability is inviscid (Michalke 1964), we derive the forcing from a temporal stability analysis of the inviscid equations linearized about the parallel hyperbolic tangent inlet profile  $\mathbf{u}(\mathbf{x}) = (U(y), 0)$ . Defining a perturbation streamfunction  $\psi(\mathbf{x}, t)$  with real wavenumber  $\alpha$  and complex velocity  $c = c_r + ic_i$

$$\psi(\mathbf{x}, t) = \text{Re} \left\{ \hat{\psi}(y) \exp(i\alpha(x - ct)) \right\}, \tag{A1}$$

the evolution of the perturbation is given by the Rayleigh equation

$$(U - c) \left( \frac{d^2}{dy^2} - \alpha^2 \right) \hat{\psi} - U'' \hat{\psi} = 0. \tag{A2}$$

For each wavenumber  $\alpha$ , this can be written as a generalized eigenvalue problem

$$\mathcal{A}(\alpha) \hat{\psi}_j(y; \alpha) = c^{(j)} \mathcal{B}(\alpha) \hat{\psi}_j(y; \alpha), \tag{A3}$$

for the  $j$ th velocity  $c$  and perturbation streamfunction  $\hat{\psi}$ , where we use the convention of sorting by decreasing  $\text{Im}\{c^{(j)}\}$ .

The base profile is odd, so the real phase velocity is the midline value  $c_r = \bar{U}$  (Michalke 1964). With  $\bar{U} = 1$ , perturbations oscillate with frequency  $\omega = \alpha$  and growth rate  $\sigma(\alpha) = \alpha c_i(\alpha)$ . For further details on linear stability analysis of shear flows, see for example Huerre & Monkewitz (1985) and Schmid & Henningson (2012). We choose a spectral discretization of (A2) using a Hermite collocation scheme from a Python implementation of the `dmsuite` library (Weidemann & Reddy 2000). We find the least stable wavenumber  $\alpha^*$  by maximizing the growth rate  $\sigma$  (i.e. minimizing  $-\alpha c_i^{(1)}$ ) using an implementation of the BFGS algorithm available in `scipy`. Since this departs from the standard shooting method, we validate it against Michalke (1964), who uses

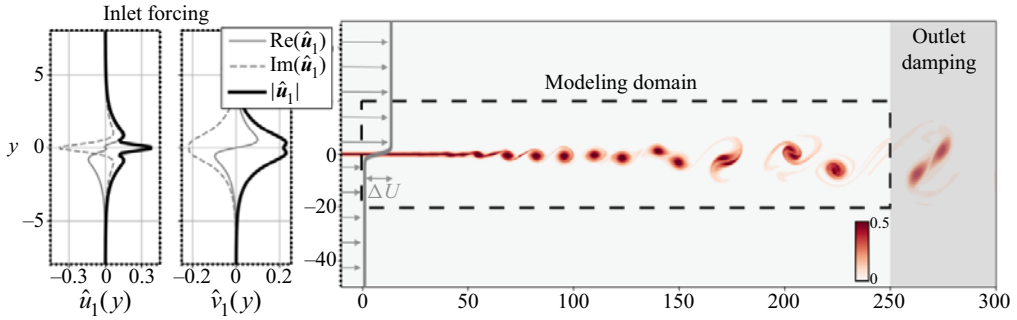


Figure 11. An incompressible mixing layer at  $Re = 500$ . The hyperbolic tangent base flow is convectively unstable, amplifying small perturbations as they are carried downstream. We force the flow at the inlet with eigenfunctions of inviscid flow equations linearized around the base flow (left). The central part of the domain is used for further modelling to avoid boundary effects, while the downstream extent  $x \in (250, 300)$  is strongly damped to prevent numerical instability at the outlet.

the profile  $U(y) = 0.5(1 + \tanh y)$  based on non-dimensionalizing with the shear layer thickness  $\delta_\omega/2$  rather than vorticity thickness. With that profile we find a least-stable wavenumber  $\alpha^* = 0.4449$  and growth rate  $\sigma = 0.0949$  compared with the reported results of  $\alpha^* = 0.4446$  and  $\sigma = 0.0949$ . For the present base profile (5.2) we find  $\alpha^* = 0.8912$  and  $\sigma = 0.2129$ .

Once the wavenumber  $\alpha^*$  and eigenfunction  $\hat{\psi}_1$  corresponding to the maximum growth rate is identified, the inlet velocity perturbation can be computed from

$$\hat{u}_1(y; \alpha^*) = \hat{\psi}'_1(y; \alpha^*), \quad \hat{v}_1(y; \alpha^*) = i\alpha^* \hat{\psi}_1(y; \alpha^*). \quad (\text{A4a,b})$$

The optimal perturbations are shown in figure 11. Following Colonius *et al.* (1997), we also apply forcing at the first three subharmonics of the least-stable mode computed from solving (A2) at  $\alpha^*/2$ ,  $\alpha^*/4$ , and  $\alpha^*/8$ , rescaling each so that the maximum value of  $\hat{u}_1$  is  $10^{-3} \Delta U$ . We perturb the inlet profile by  $\text{Re}\{\hat{u}_1(y, \alpha_n) \exp(i\omega_n t)\}$  for  $n = 1, 2, 4, 8$  and  $\alpha_n = \omega_n = \alpha^*/n$  and use a simulation time step of  $\Delta t = 0.00705 = 10^{-3} \times 2\pi/\alpha^*$ , so that the sampling rate of the simulation is commensurate with the forcing frequency.

### Appendix B. Variable truncation of the mixing layer model

One of the limitations of projection-based reduced-order modelling for systems without an explicit scale separation is a lack of principled criteria for selecting the truncation rank, or dimension of the projection subspace. A common approach is to use *ad hoc* criteria such as the number of modes required to resolve some fraction of the fluctuation kinetic energy (e.g. 90% or 99%) on average. To some extent this simply replaces the arbitrary choice of rank truncation with an equally arbitrary choice of approximation accuracy. This is especially the case when the dynamic approximation accuracy of the reduced-order model does not monotonically improve with increasing kinematic approximation accuracy (as can be seen in figure 9, for instance).

Unfortunately, this problem is compounded in the secondary reduction step of the MMR method, since one must select both the rank  $r$  of the original POD–Galerkin system and the rank  $r_0$  of the final MMR system. Our approach in this work was to follow the heuristic of simulating the POD–Galerkin system at various values of  $r$  and choosing that which performs best or remains stable for longest (e.g.  $r = 24$  for the mixing layer model in figure 9). The final MMR models then typically behave similarly for a range of final ranks  $r_0$ .

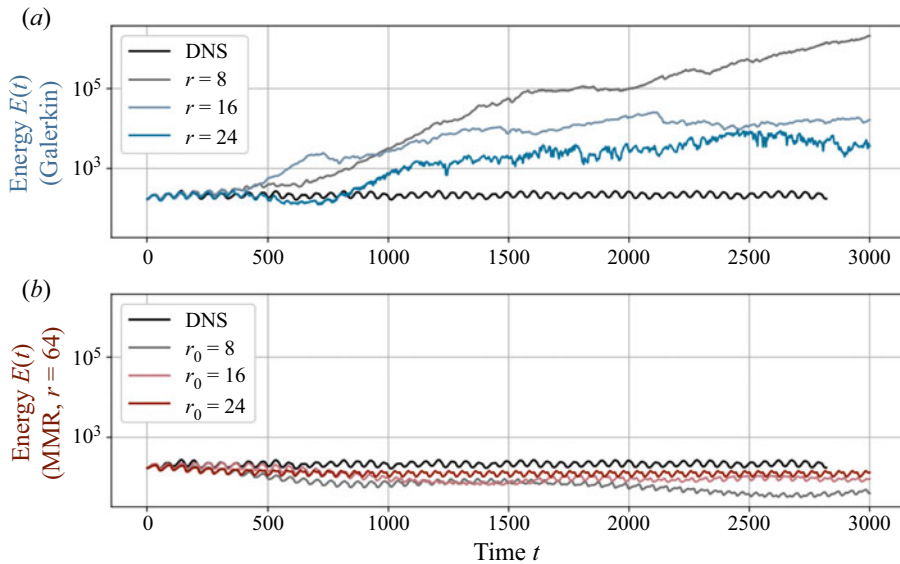


Figure 12. Alternative truncations of the mixing layer model in § 6.3. Rather than the ‘optimal’ value of  $r = 24$ , the MMR models shown here are constructed from Galerkin systems of rank  $r = 64$ . While the POD–Galerkin models are inaccurate and energetically unstable for any rank truncation, while the MMR models are less sensitive to the specific values of rank truncation.

However, in order to show that the method is relatively insensitive to the specific choices of rank truncation, here we repeat the model reduction for the mixing layer on the longest domain shown in figure 9. These models begin with  $r = 64$  rather than the ‘optimal’ POD–Galerkin rank of  $r = 24$ . Although less accurate than the model reported in § 6.3, the resulting models are stable for all values  $r_0$  of the final rank and significantly outperform any POD–Galerkin model as shown by figure 12.

As a more principled example, figure 13 shows a model constructed by choosing a kinematic reconstruction accuracy threshold of 99 % for the POD–Galerkin rank  $r$ . The MMR rank  $r_0$  is then selected with the classic ‘elbow’ criterion, which selects a truncation level based on the inflection point of the singular value distribution. Again, the model is less accurate than the results shown in § 6.3, but significantly more stable and accurate than any Galerkin system. Thus, while a principled approach to rank selection is an open problem for both POD–Galerkin modelling and the proposed multiscale reduction, the latter appears generally less sensitive to this choice.

### Appendix C. Forecasting accuracy

As discussed in § 6, MMR is not primarily intended to produce accurate forecasts over long time horizons. The underlying POD–Galerkin model is developed based on optimizing the estimate of the instantaneous time derivative, which does not necessarily result in optimal finite-time predictions. Since MMR is developed based on asymptotic consistency with the Galerkin system it is also not optimized for forecasting accuracy. In this, autoregressive methods (Billings 2013) that are specifically tailored to forecasting are more appropriate. On the other hand, dynamical accuracy is often more important than long-term forecasting in applications such as the design of feedback controllers (Noack *et al.* 2011; Brunton & Noack 2015).



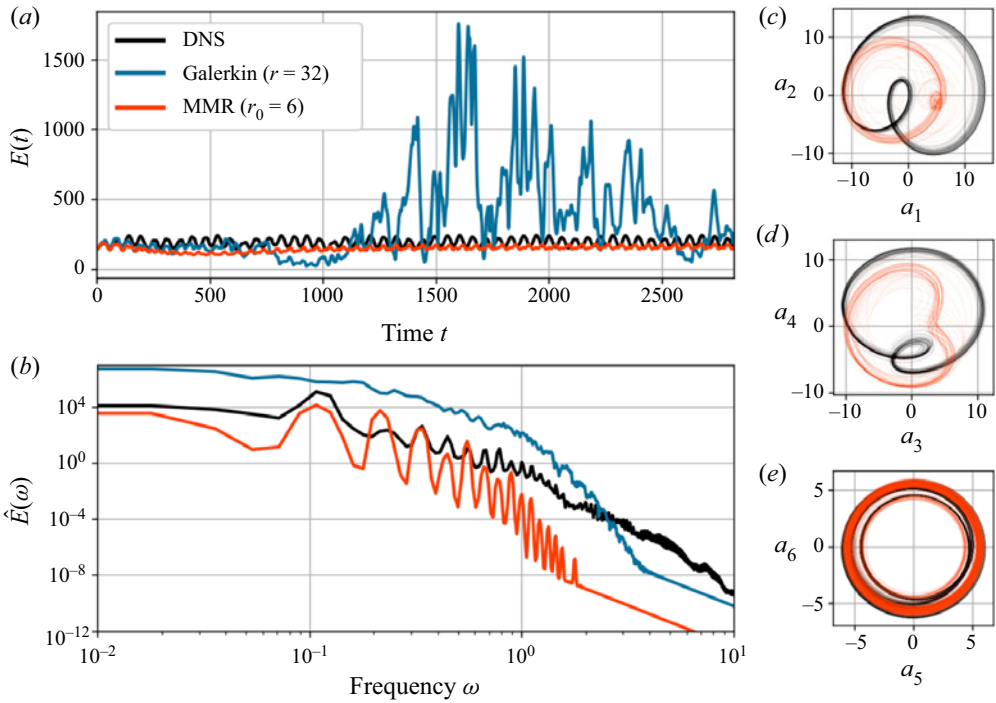


Figure 13. Truncation of the mixing layer in § 6.3 based on the singular value spectrum. In this case the initial truncation level  $r = 32$  is selected based on a residual of 1 %, while the MMR rank  $r_0 = 6$  is based on the ‘knee’ criterion (see text). As with the models shown in figure 12, the resulting model is stable and roughly matches the true dynamics, but is not as accurate as the model reported in § 6.3.

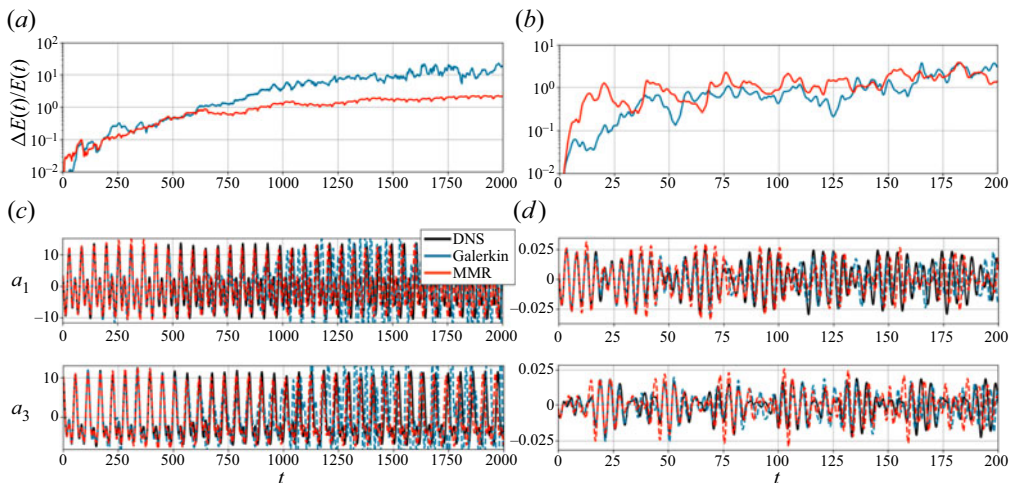


Figure 14. Time-series forecasts for the mixing layer (*a,c*) and lid-driven cavity (*b,d*). Due to the complex dynamics of these flows, neither modelling approach produces accurate forecasts over long time horizons. The normalized energy error  $\Delta E(t)/E(t)$  is comparable between the standard POD–Galerkin and MMR models (*a,b*). Predictions for leading individual coefficients show significant errors after several periods of the dominant oscillatory frequency (*c,d*).

However, for the sake of completeness figure 14 gives a comparison of the forecasting accuracy of both the standard POD–Galerkin and MMR models. The error metric is the normalized energy error

$$\frac{\Delta E(t)}{E(t)} = \frac{\|\mathbf{a}_{DNS}(t) - \mathbf{a}_{ROM}(t)\|^2}{\|\mathbf{a}_{DNS}(t)\|^2}. \quad (\text{C1})$$

In this metric neither POD–Galerkin nor MMR produces accurate long-term forecasts.

In flows with a relatively simple dynamics (e.g. the cylinder wake or short-domain mixing layer), small errors in the phase of oscillations tend to accumulate over time, even if the dynamical behaviour is modelled accurately. For instance, figure 14(c) shows that errors in MMR-based predictions of individual POD coefficients can largely be attributed to phase drift, while the Galerkin forecast diverges due to dynamical instability.

However, if the flow is chaotic (e.g. the lid-driven cavity) then detailed forecasting on times much longer than that determined by the Lyapunov exponent is impossible even in principle. Hence, while MMR more accurately captures the statistical behaviour of the lid-driven cavity flow (figure 7), it does not improve upon the finite-time predictions of the POD–Galerkin model (figure 14d).

## REFERENCES

- ÅKERVIK, E., BRANDT, L., HENNINGSON, D.S., HÖPFNER, J., MARXEN, O. & SCHLATTER, P. 2006 Steady solutions of the Navier–Stokes equations by selective frequency damping. *Phys. Fluids* **18**, 068102.
- ARBABI, H. & MEZIĆ, I. 2017 Study of dynamics in post-transient flows using Koopman mode decomposition. *Phys. Rev. Fluids* **2**, 124402.
- ARBABI, H. & SAPSIS, T. 2022 Data-driven modeling of strongly nonlinear chaotic systems with non-Gaussian statistics. *SIAM/ASA J. Uncertainty Quant.* **10** (2), 555–583.
- AUBRY, N., HOLMES, P., LUMLEY, J.L. & STONE, E. 1988 The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.* **192**, 115.
- BALAJEWICZ, M., DOWELL, E. & NOACK, B.R. 2013 Low-dimensional modelling of high-Reynolds-number shear flows incorporating constraints from the Navier–Stokes equation. *J. Fluid Mech.* **729**, 285–308.
- BARBAGALLO, A., SIPP, D. & SCHMID, P.J. 2009 Closed-loop control of an open cavity flow using reduced-order models. *J. Fluid Mech.* **641**, 1.
- BARKLEY, D. & HENDERSON, R.D. 1996 Three-dimensional Floquet stability analysis of the wake of a circular cylinder. *J. Fluid Mech.* **322**, 215–241.
- BARTHEL, B., GOMEZ, S. & MCKEON, B.J. 2022 Variational formulation of resolvent analysis. *Phys. Rev. Fluids* **7**, 013905.
- BARTHEL, B., ZHU, X. & MCKEON, B.J. 2021 Closing the loop: nonlinear Taylor vortex flow through the lens of resolvent analysis. *J. Fluid Mech.* **924**, A9.
- BAZILEVS, Y., CALO, V.M., COTTRELL, J.A., HUGHES, T.J.R., REALI, A. & SCOVAZZI, G. 2007 Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. *Comput. Meth. Appl. Mech. Engng* **197** (1), 173–201.
- BELSON, B.A., TU, J.H. & ROWLEY, C.W. 2014 Algorithm 945: modred - a parallelized model reduction library. *ACM Trans. Math. Softw.* **40** (4), 1–23.
- BENNER, P., GUGERCIN, S. & WILLCOX, K. 2015 A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev.* **57** (4), 483–531.
- BILLINGS, S.A. 2013 *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Wiley.
- BRANCHER, P. & CHOMAZ, J.-M. 1997 Absolute and convective instabilities in spatially periodic shear flows. *Phys. Rev. Lett.* **78** (4), 658–661.
- BRENNER, M.P., ELDREDGE, J.D. & FREUND, J.B. 2019 Perspective on machine learning for advancing fluid mechanics. *Phys. Rev. Fluids* **4**, 100501.
- BROWN, G.L. & ROSHKO, A. 1974 On density effects and large structures in turbulent mixing layers. *J. Fluid Mech.* **64** (4), 775–816.
- BRUNTON, S.L., BUDIŠIĆ, M., KAISER, E. & KUTZ, J.N. 2022 Modern Koopman theory for dynamical systems. *SIAM Rev.* **64** (2), 229–340.

- BRUNTON, S.L. & NOACK, B.R. 2015 Closed-loop turbulence control: progress and challenges. *Appl. Mech. Rev.* **67**, 050801.
- BRUNTON, S.L., NOACK, B.R. & KOUMOUTSAKOS, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52** (1), 477–508.
- BRUNTON, S.L., PROCTOR, J.L. & KUTZ, J.N. 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Natl Acad. Sci. USA* **113** (15), 3932–3937.
- CAIAZZO, A., ILIESCU, T., JOHN, V. & SCHYSCHLOWA, S. 2014 A numerical investigation of velocity-pressure reduced order models for incompressible flows. *J. Comput. Phys.* **259**, 598–616.
- CALLAHAM, J.L., BRUNTON, S.L. & LOISEAU, J.-C. 2022 On the role of nonlinear correlations in reduced-order modeling. *J. Fluid Mech.* **938**, A1.
- CALLAHAM, J.L., LOISEAU, J.-C., RIGAS, G. & BRUNTON, S.L. 2021 Nonlinear stochastic modeling with Langevin regression. *Proc. R. Soc. A* **477** (2250), 20210092.
- CALLAHAM, J.L., RIGAS, G., LOISEAU, J.-C. & BRUNTON, S.L. 2022 An empirical mean-field model of symmetry-breaking in a turbulent wake. *Sci. Adv.* **8**, eabm4786.
- CARINI, M., AUTERI, F. & GIANNETTI, F. 2015 Centre-manifold reduction of bifurcating flows. *J. Fluid Mech.* **767**, 109–145.
- CHAMPION, K., LUSCH, B., KUTZ, J.N. & BRUNTON, S.L. 2019 Data-driven discovery of coordinates and governing equations. *Proc. Natl Acad. Sci. USA* **116** (45), 22445–22451.
- CHOMAZ, J.-M. 2005 Global instabilities in spatially developing flows: non-normality and nonlinearity. *Annu. Rev. Fluid Mech.* **37**, 357–392.
- COLONIUS, T., LELE, S.K. & MOIN, P. 1997 Sound generation in a mixing layer. *J. Fluid Mech.* **330**, 375–409.
- CORDIER, L., NOACK, B.R., TISSOT, G., LEHNASCH, G., DELVILLE, J., BALAJEWICZ, M., DAVILLER, G. & NIVEN, R.K. 2013 Identification strategies for model-based control. *Exp. Fluids* **54**, 1580.
- COULLET, P.H. & SPIEGEL, E.A. 1983 Amplitude equations for systems with competing instabilities. *SIAM J. Appl. Maths* **43** (4), 776–820.
- DEANE, A.E., KEVREKEDIS, I.G., KARNIADAKIS, G.E. & ORSZAG, S.A. 1991 Low-dimensional models for complex-geometry flows - application to grooved channels and circular-cylinders. *Phys. Fluids A* **3**, 2337–2354.
- DENG, N., NOACK, B.R., MORZYNSKI, M. & PASTUR, L.R. 2020 Low-order model for successive bifurcations of the fluidic pinball. *J. Fluid Mech.* **884**, A37.
- DENG, N., NOACK, B.R., MORZYŃSKI, M. & PASTUR, L.R. 2021 Galerkin force model for transient and post-transient dynamics of the fluidic pinball. *J. Fluid Mech.* **918**, A4.
- DURAISAMY, K., IACCARINO, G. & XIAO, H. 2019 Turbulence modeling in the age of data. *Annu. Rev. Fluid Mech.* **51**, 357–377.
- FISCHER, P.F., LOTTES, J.W. & KERKEMEIR, S.G. 2008 Nek5000 web pages. <http://nek5000.mcs.anl.gov>.
- FROYLAND, G. 2005 Statistically optimal almost-invariant sets. *Physica D* **200** (3), 205–219.
- FROYLAND, G. & PADBERG, K. 2009 Almost-invariant sets and invariant manifolds – connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D* **238**, 1507–1523.
- FROYLAND, G., SANTITISSADEEKORN, N. & MONAHAN, A. 2010 Transport in time-dependent dynamical systems: finite-time coherent sets. *Chaos* **20** (4), 043116.
- GIVON, D., KUPFERMAN, R. & STUART, A.M. 2004 Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity* **17**, R55–R127.
- GRIMBERG, S., FARHAT, C. & YOUKILIS, N. 2020 On the stability of projection-based model order reduction for convection-dominated laminar and turbulent flows. *J. Comput. Phys.* **419**, 109681.
- GUAN, Y., BRUNTON, S.L. & NOVOSSELOV, I. 2021 Sparse nonlinear models of chaotic electroconvection. *R. Soc. Open Sci.* **8** (8), 202367.
- GUCKENHEIMER, J. & HOLMES, P. 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer.
- HAKEN, H. 1983 *Synergetics: an Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry and Biology*. Springer.
- HETHAVEN, J.S. & UBBIALI, S. 2018 Non-intrusive reduced order modeling of nonlinear problems using neural networks. *J. Comput. Phys.* **363**, 55–78.
- HOLMES, P., LUMLEY, J.L. & BERKOOZ, G. 1996 *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge Monographs on Mechanics.
- HOPF, E. 1948 A mathematical example displaying features of turbulence. *Commun. Pure Appl. Maths* **1** (4), 302–322.
- HUERRE, P. & MONKEWITZ, P.A. 1985 Absolute and convective instabilities in free shear layers. *J. Fluid Mech.* **159**, 151–168.

- HUERRE, P. & MONKEWITZ, P.A. 1990 Local and global instabilities in spatially developing shear flows. *Annu. Rev. Fluid Mech.* **22**, 473–537.
- HUSSAIN, A.K.M.F. 1981 Role of coherent structures in turbulent shear flows. *Proc. Indian Acad. Sci C* **4** (2), 129–175.
- KAPTANOGLU, A.A., CALLAHAM, J.L., HANSEN, C.J., ARAVKIN, A. & BRUNTON, S.L. 2021a Promoting global stability in data-driven models of quadratic nonlinear dynamics. *Phys. Rev. Fluids* **6** (9), 094401.
- KAPTANOGLU, A.A., MORGAN, K.D., HANSEN, C.J. & BRUNTON, S.L. 2021b Physics-constrained, low-dimensional models for magnetohydrodynamics: first-principles and data-driven approaches. *Phys. Rev. E* **104** (1), 015206.
- KINGMA, D.P. & WELLING, M. 2013 Auto-encoding variational Bayes. [arXiv:1312.6114](https://arxiv.org/abs/1312.6114).
- KLUS, S., KOLTAI, P. & SCHÜTTE, C. 2016 On the numerical approximation of the Perron-Frobenius and Koopman operator. *J. Comput. Dyn.* **3** (1), 51–77.
- KLUS, S., NÜSKE, F., KOLTAI, P., WU, H., KEVREKIDIS, I., SCHÜTTE, C. & NOÉ, F. 2018 Data-driven model reduction and transfer operator approximation. *J. Nonlinear Sci.* **28** (3), 985–1010.
- KLUS, S., NÜSKE, F., PEITZ, S., NIEMANN, J.-H., CLEMENTI, C. & SCHÜTTE, C. 2020 Data-driven approximation of the Koopman generator: model reduction, system identification, and control. *Physica D* **406**, 132416.
- KRAICHNAN, R.H. & CHEN, S. 1989 Is there a statistical mechanics of turbulence? *Physica D* **37**, 160–172.
- KURTZ, T.G. 1973 A limit theorem for perturbed operator semigroups with applications to random evolutions. *J. Funct. Anal.* **12**, 55–67.
- LANDAU, L.D. 1944 On the problem of turbulence. *Dokl. Akad. Nauk SSSR* **44**, 339–342.
- LEE, M.W., DOWELL, E.H. & BALAJEWICZ, M.J. 2019 A study of the regularized lid-driven cavity's progression to chaos. *Commun. Nonlinear Sci. Numer. Simul.* **71**, 50–72.
- LOISEAU, J.-C. 2020 Data-driven modeling of the chaotic thermal convection in an annular thermosyphon. *Theor. Comput. Fluid Dyn.* **34** (4), 339–365.
- LOISEAU, J.-C. & BRUNTON, S.L. 2018 Constrained sparse Galerkin regression. *J. Fluid Mech.* **838**, 42–67.
- LOISEAU, J.-C., BRUNTON, S.L. & NOACK, B.R. 2018a From the POD-Galerkin method to sparse manifold models. In *Handbook on Model Order Reduction*. De Gruyter GmbH.
- LOISEAU, J.-C., BUCCI, M.A., CHERUBINI, S. & ROBINET, J.-C. 2019 Time-stepping and Krylov methods for large-scale instability problems. In *Computational Modelling of Bifurcations and Instabilities in Fluid Dynamics*, pp. 33–73. Springer.
- LOISEAU, J.-C., NOACK, B.R. & BRUNTON, S.L. 2018b Sparse reduced-order modelling: sensor-based dynamics to full-state estimation. *J. Fluid Mech.* **844**, 454–490.
- LUCHTENBURG, D.M., GÜNTHER, B., NOACK, B.R., KING, R. & TADMOR, G. 2009 A generalized mean-field model of the natural and high-frequency actuated flow around a high-lift configuration. *J. Fluid Mech.* **623**, 283.
- MA, X. & KARNIADAKIS, G.E. 2002 A low-dimensional model for simulating three-dimensional cylinder flow. *J. Fluid Mech.* **458**, 181–190.
- MAJDA, A.J., TIMOFEYEV, I. & VANDEN-EIJNDEN, E. 2001 A mathematical framework for stochastic climate models. *Commun. Pure Appl. Maths* **54** (8), 891–974.
- MAJDA, A.J., TIMOFEYEV, I. & VANDEN-EIJNDEN, E. 2006 Stochastic models for selected slow variables in large deterministic systems. *Nonlinearity* **19**, 769–794.
- MANTIC-LUGO, V., ARRATIA, C. & GALLAIRE, F. 2014 Self-consistent mean flow description of the nonlinear saturation of the vortex shedding in the cylinder wake. *Phys. Rev. Lett.* **113**, 084501.
- MCKEON, B.J. & SHARMA, A.S. 2010 A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.* **658**, 336–382.
- MELIGA, P. 2017 Harmonics generation and the mechanics of saturation in flow over an open cavity: a second-order self-consistent description. *J. Fluid Mech.* **826**, 503–521.
- MELIGA, P. & CHOMAZ, J.-M. 2011 An asymptotic expansion for the vortex-induced vibrations of a circular cylinder. *J. Fluid Mech.* **671**, 137–167.
- MELIGA, P., CHOMAZ, J.-M. & SIPP, D. 2009 Global mode interaction and pattern selection in the wake of a disk: a weakly nonlinear expansion. *J. Fluid Mech.* **633**, 159–189.
- MENDIBLE, A., BRUNTON, S.L., ARAVKIN, A.Y., LOWRIE, W. & KUTZ, J.N. 2020 Dimensionality reduction and reduced-order modeling for traveling wave physics. *Theor. Comput. Fluid Dyn.* **34** (4), 385–400.
- MENIER, E., BUCCI, M.A., YAGOUBI, M., MATHELIN, L. & SCHCOENAUER, M. 2022 CD-ROM: Complementary deep - reduced order model. Euromech colloquium on Machine learning methods for turbulent separated flows, June 2021, Paris, France.



- MEZIĆ, I. 2005 Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn.* **41** (1–3), 309–325.
- MEZIĆ, I. 2013 Analysis of fluid flows via spectral properties of the Koopman operator. *Annu. Rev. Fluid Mech.* **45** (1), 357–378.
- MICHALKE, A. 1964 On the inviscid instability of the hyperbolic-tangent velocity profile. *J. Fluid Mech.* **19** (4), 543–556.
- MOHEBUJJAMAN, M., REBHOLZ, L.G. & ILIESCU, T. 2018 Physically constrained data-driven correction for reduced-order modeling of fluid flows. *Intl J. Numer. Meth. Fluids* **89** (3), 103–122.
- MOHEBUJJAMAN, M., REBHOLZ, L.G., XIE, X. & ILIESCU, T. 2017 Energy balance and mass conservation in reduced order models of fluid flows. *J. Comput. Phys.* **346** (1), 262–277.
- MONKEWITZ, P.A. & HUERRE, P. 1982 Influence of the velocity ratio on the spatial instability of mixing layers. *Phys. Fluids* **25** (7), 1137.
- NOACK, B.R., AFANASIEV, K., MORZYNSKI, M., TADMOR, G. & THIELE, F. 2003 A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.* **497**, 335–363.
- NOACK, B.R. & ECKELMANN, H. 1994 A low-dimensional Galerkin method for the three-dimensional flow around a circular cylinder. *Phys. Fluids* **6** (1), 124–143.
- NOACK, B.R., MORZYNSKI, M. & TADMOR, G. 2011 *Reduced-Order Modelling for Flow Control*, vol. 528. Springer Science & Business Media.
- NOACK, B.R., PAPAS, P. & MONKEWITZ, P.A. 2005 The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.* **523**, 339–365.
- NOACK, B.R., SCHLEGEL, M., AHLBORN, B., MUTSCHKE, G., MORZYNSKI, M., COMTE, P. & TADMOR, G. 2008 A finite-time thermodynamics of unsteady fluid flows. *J. Non-Equilib. Thermodyn.* **33**, 103–148.
- ÖSTH, J., NOACK, B.R., KRAJNOVIĆ, S., BARROS, D. & BORÉE, J. 2014 On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body. *J. Fluid Mech.* **747**, 518–544.
- OTTO, S.E., PADOVAN, A. & ROWLEY, C.W. 2022 Optimizing oblique projections for nonlinear systems using trajectories. *SIAM J. Sci. Comput.* **44** (3), A1681–A1702.
- PADOVAN, A., OTTO, S.E. & ROWLEY, C.W. 2020 Analysis of amplification mechanisms and cross-frequency interactions in nonlinear flows via the harmonic resolvent. *J. Fluid Mech.* **900**, A14.
- PAPANICOLAOU, G. 1976 Some probabilistic problems and methods in singular perturbations. *Rocky Mt. J. Maths* **6** (4), 653–673.
- PAVLIOTIS, G.A. & STUART, A.M. 2012 *Multiscale Methods: Averaging and Homogenization*. Springer.
- PEHERSTORFER, B. & WILLCOX, K. 2016 Data-driven operator inference for nonintrusive projection-based model reduction. *Comput. Meth. Appl. Mech. Engng* **306**, 196–215.
- PICKERING, E., RIGAS, G., SCHMIDT, O.T., SIPP, D. & COLONIUS, T. 2021 Optimal eddy viscosity for resolvent-based models of coherent structures in turbulent jets. *J. Fluid Mech.* **917**, A29.
- PRADAS, M., PAVLIOTIS, G.A., KALLIADASIS, S., PAPAGEORGIOU, D.T. & TSELUIKO, D. 2012 Additive noise effects in active nonlinear spatially extended systems. *Eur. J. Appl. Maths* **23**, 563–591.
- QIAN, E., KRAMER, B., PEHERSTORFER, B. & WILLCOX, K. 2020 Lift and learn: physics-informed machine learning for large-scale nonlinear dynamical systems. *Physica D* **406**, 132401.
- REISS, J., SCHULZE, P., SESTERHENN, J. & MEHRMANN, V. 2018 The shifted proper orthogonal decomposition: a mode decomposition for multiple transport phenomena. *SIAM J. Sci. Comput.* **40** (3), A1322–A1344.
- REMPFER, D. & FASEL, H.F. 1994 Dynamics of three-dimensional coherent structures in a flat-plate boundary layer. *J. Fluid Mech.* **275**, 257–282.
- RIGAS, G., SIPP, D. & COLONIUS, T. 2021 Nonlinear input/output analysis: application to boundary layer transition. *J. Fluid Mech.* **911**, A15.
- RIM, D., MOE, S. & LEVEQUE, R.J. 2018 Transport reversal for model reduction of hyperbolic partial differential equations. *SIAM/ASA J. Uncert. Quant.* **6** (1), 118–150.
- RISKEN, H. 1996 *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer.
- ROWLEY, C., MEZIĆ, I., BAGHERI, S., SCHLATTER, P. & HENNINGSON, D.S. 2009 Spectral analysis of nonlinear flows. *J. Fluid Mech.* **641**, 115–127.
- ROWLEY, C.W., COLONIUS, T. & MURRAY, R.M. 2004 Model reduction for compressible flows using POD and Galerkin projection. *Physica D* **189**, 115–129.
- ROWLEY, C.W. & DAWSON, S.T.M. 2017 Model reduction for flow analysis and control. *Annu. Rev. Fluid Mech.* **49**, 387–417.
- ROWLEY, C.W. & MARSDEN, J.E. 2000 Reconstruction equations and the Karhunen–Loève expansion for systems with symmetry. *Physica D* **142** (1–2), 1–19.



- RUBINI, R., LASAGNA, D. & RONCH, A.D. 2020 The  $\ell_1$ -based sparsification of energy interaction in unsteady lid-driven cavity flow. *J. Fluid Mech.* **905**, A15.
- RUELLE, D. & TAKENS, F. 1971 On the nature of turbulence. *Commun. Math. Phys.* **20** (3), 167–192.
- SAN, O. & MAULIK, R. 2018a Extreme learning machine for reduced order modeling of turbulent geophysical flows. *Phys. Rev. E* **97**, 042322.
- SAN, O. & MAULIK, R. 2018b Neural network closures for nonlinear model order reduction. *Adv. Comput. Maths* **44** (6), 1717–1750.
- SCHLEGEL, M. & NOACK, B.R. 2015 On long-term boundedness of Galerkin models. *J. Fluid Mech.* **765**, 325–352.
- SCHMID, P.J. & HENNINGSON, D.S. 2012 *Stability and Transition in Shear Flows*. Springer.
- SHARMA, A.S. & MCKEON, B.J. 2013 On coherent structure in wall turbulence. *J. Fluid Mech.* **728**, 196–238.
- SIEBER, M., PASCHEREIT, C.O. & OBERLEITHNER, K. 2021 Stochastic modelling of a noise driven global instability in a turbulent swirling jet. *J. Fluid Mech.* **916**, A7.
- SIPP, D. & LEBEDEV, A. 2007 Global stability of base and mean flows: a general approach and its applications to cylinder and open cavity flows. *J. Fluid Mech.* **593**, 333–358.
- SIPP, D., MARQUET, O., MELIGA, P. & BARBAGALLO, A. 2010 Dynamics and control of global instabilities in open-flows: a linearized approach. *Appl. Mech. Rev.* **63** (3), 030801.
- SIROVICH, L. 1987 Turbulence and the dynamics of coherent structures, parts I–III. *Q. J. Appl. Maths* **45** (3), 561–90.
- STANLEY, S. & SARKAR, S. 1997 Simulations of spatially developing two-dimensional shear layers and jets. *Theor. Comput. Fluid Dyn.* **9** (2), 121–147.
- STEWARTSON, K. & STUART, J.T. 1971 A non-linear instability theory for a wave system in plane Poiseuille flow. *J. Fluid Mech.* **48** (3), 529–545.
- STUART, J.T. 1958 On the non-linear mechanics of hydrodynamic stability. *J. Fluid Mech.* **4** (1), 1–21.
- SWINNEY, H.L. & GOLLUB, J.P. 1981 *Hydrodynamic Instabilities and the Transition to Turbulence*. Springer.
- TAIRA, K., BRUNTON, S.L., DAWSON, S.T.M., ROWLEY, C.W., COLONIUS, T., MCKEON, B.J., SCHMIDT, O.T., GORDEYEV, S., THEOFILIS, V. & UKEILEY, L.S. 2017 Modal analysis of fluid flows: an overview. *AIAA J.* **55** (12), 4013–4041.
- TAIRA, K., HEMATI, M.S., BRUNTON, S.L., SUN, Y., DURAISAMY, K., BAGHERI, S., DAWSON, S. & YEH, C.-A. 2020 Modal analysis of fluid flows: applications and outlook. *AIAA J.* **58** (3), 998–1022.
- UKEILEY, L., CORDIER, L., MANCEAU, R., DELVILLE, J., GLAUSER, M. & BONNET, J.P. 2001 Examination of large-scale structures in a turbulent plane mixing layer. Part 2. Dynamical systems model. *J. Fluid Mech.* **441**, 67–108.
- WAN, Z.Y., VLACHAS, P., KOUMOUTSAKOS, P. & SAPSIS, T. 2018 Data-assisted reduced-order modeling of extreme events in complex dynamical systems. *PLoS ONE* **13** (5), e0197704.
- WANG, Z., AKHTAR, I., BORGGARD, J. & ILIESCU, T. 2012 Proper orthogonal decomposition closure models for turbulent flows: a numerical comparison. *Comput. Meth. Appl. Mech. Engng* **237**, 10–26.
- WEIDEMANN, J.A.C. & REDDY, S.C. 2000 A MATLAB differentiation matrix suite. *ACM Trans. Math. Softw.* **26**, 465–519.
- WEINAN, E. 2011 *Principles of Multiscale Modeling*. Cambridge University Press.
- WEINAN, E. & ENGQUIST, B. 2003 The heterogeneous multiscale methods. *Commun. Math. Sci.* **1**, 87–132.
- WEINAN, E., ENGQUIST, B., LI, X., REN, W. & VANDEN-EIJNDEN, E. 2007 The heterogeneous multiscale method: a review. *Commun. Comput. Phys.* **2**, 367–450.
- WINANT, C.D. & BROWAND, F.K. 1974 Vortex pairing: the mechanism of turbulent mixing-layer growth at moderate Reynolds number. *J. Fluid Mech.* **63** (2), 237–255.
- XIE, X., MOHEBUJAMAN, M., REBHOLZ, L.G. & ILIESCU, T. 2018 Data-driven filtered reduced order modeling of fluid flows. *SIAM J. Sci. Comput.* **40** (3), B834–B857.
- ZARE, A., JOVANOVIĆ, M.R. & GEORGIU, T.T. 2017 Colour of turbulence. *J. Fluid Mech.* **812**, 636–680.