Journal of Computational Physics 474 (2023) 111757

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Sensitivity analysis of chaotic systems using a frequency-domain shadowing approach

Kyriakos D. Kantarakias, George Papadakis *

Department of Aeronautics, Imperial College London, SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 6 April 2022
Received in revised form 30 October 2022
Accepted 4 November 2022
Available online 15 November 2022

Keywords:
Chaotic systems
Sensitivity analysis
Least squares shadowing

ABSTRACT

We present a frequency-domain method for computing the sensitivities of time-averaged quantities of chaotic systems with respect to input parameters. Such sensitivities cannot be computed by conventional adjoint analysis tools, because the presence of positive Lyapunov exponents leads to exponential growth of the adjoint variables. The proposed method is based on the well established least-square shadowing (LSS) approach [1], that formulates the evaluation of sensitivities as an optimisation problem, thereby avoiding the exponential growth of the solution. All existing formulations of LSS (and its variants) are in the time domain. In the present paper, we reformulate the LSS method in the frequency (Fourier) space using harmonic balancing. The resulting system is closed using periodicity. The new method is tested on the Kuramoto-Sivashinsky system and the results match with those obtained using the standard time-domain formulation. The storage and computing requirements of the direct solution grow rapidly with the size of the system. To mitigate these requirements, we propose a resolvent-based iterative approach that needs much less storage. Application to the Kuramoto-Sivashinsky system gave accurate results with low computational cost. Truncating the large frequencies with small energy content from the harmonic balancing operator did not affect the accuracy of the computed sensitivities. Further work is needed to assess the performance and scalability of the proposed method.

1. Introduction

Optimisation of engineering devices is based of the definition of an objective function, usually a time-average quantity $\overline{J}(\boldsymbol{s})$, and the evaluation of the problem parameters, $\boldsymbol{s}$, that minimise or maximise this function, depending on the application. During the optimisation process, the gradient of the objective function with respect to the parameters $d\overline{J}(\boldsymbol{s})/d\boldsymbol{s}$ (also known as sensitivity) is usually required. This is obtained by solving the tangent equation (or the adjoint equations for multiple parameters). In either case, the equations are obtained by linearising the governing non-linear set describing the system around the solution obtained for the reference values of the parameters, $\boldsymbol{s}$. The tangent (or adjoint) equations are then integrated forward (or backward) in time respectively to obtain the desired sensitivities.

The aforementioned approach works very well when the governing set of equations describes a steady or a non-chaotic system. However, when the evolution of the system exhibits chaotic behaviour, this process fails. The reason is that chaotic systems have one or more positive Lyapunov exponents (PLEs), thus two solution trajectories starting from the same initial conditions and evaluated at $\boldsymbol{s}$ and $\boldsymbol{s}+\delta\boldsymbol{s}$ deviate from each other, leading to exponentially growing sensitivities, as ex-

* Corresponding author.
E-mail address: g.papadakis@imperial.ac.uk (G. Papadakis).

https://doi.org/10.1016/j.jcp.2022.111757

plained in [2]. For this reason for example, model predictive control algorithms for transitional or turbulent flows employ the receding horizon approach, whereby the optimisation is performed over a receding window of finite time [3,4], thus sensitivities remain bounded and reliable. The sensitivity of time-average quantity $\overline{J}(\boldsymbol{s})$ to the mesh spacing can be also used for grid adaptation. The specific issues regarding mesh adaptation for chaotic systems and a proposed way for grid refinement/coarsening are discussed in [5].

Several approaches that can compute useful sensitivities in chaotic systems have been proposed. These are based on ensemble schemes [6], the fluctuation dissipation theorem [7], the Fokker-Planck equation [8], cumulant expansions [9], unsteady periodic orbits [10], or the more recent space-split sensitivity (S3) algorithm [11] that transform Ruelle's linear response formula [12] into a well-conditioned ergodic-averaging computation. One of the most promising approaches is Least Squares Shadowing (LSS) [1,13,14], which is based on the shadowing lemma [15,16]. For uniformly hyperbolic systems, this lemma guarantees the existence of a solution trajectory evaluated at $\boldsymbol{s}+\delta\boldsymbol{s}$ that shadows, i.e. remains close to, the reference trajectory evaluated at $\boldsymbol{s}$. This regularises the problem, avoids the exponential growth, and results in meaningful sensitivities.

This lemma is also central in establishing trust into the statistics of numerical solutions of chaotic systems. Due to round off errors, a computed trajectory will deviate from the true trajectory of the system (starting from the same initial condition). The shadowing lemma guarantees the existence of a true trajectory (with different initial condition) that will shadow the numerical one [17–19], thus the statistics of the computed solution can be trusted. Recent work however [20] has shown that the shadowing trajectories may not be physical, thus casting doubt on this central premise. Several chaotic systems (one dimensional perturbed tent maps) were examined; the shadowing trajectories were found to be physical in one system and non-physical in others. This finding raises important fundamental questions, such as under what conditions shadowing solutions are physical, what happens for higher dimensional systems etc. In this paper we take the standard view, that shadowing solutions are physical, and numerical simulations of chaotic systems reproduce the true statistics.

The original LSS method requires the solution of a two-point optimisation problem in time, and the cost and memory requirements grow rapidly with the number of degrees of freedom of the system. An important development in this respect is the non-intrusive Least Squares Shadowing (NILSS), [21–23] that constructs and solves a reduced-size optimization problem. This is achieved by recursive-in-time integration of a set of homogeneous tangent equations, which makes the method applicable to very large systems (the cost varies linearly with the number of degrees of freedom), but the number of equations to be integrated is equal to number of PLEs. Multiple Shooting Shadowing (MSS) is another variant applicable to large systems [24], but the cost of the pre-conditioner [25,26] also grows linearly with the number of PLEs.

The effect of the system parameters on the Lyapunov spectrum is therefore of critical importance. In the area of turbulent flows, this spectrum has been computed for low-Reynolds channel [27] and weakly chaotic Taylor-Couette flows [28]. A more recent study [29] has investigated the variation of the spectrum with Reynolds number for forced homogeneous isotropic turbulence (HIT). The Reynolds numbers examined (based on the Taylor microscale) were $Re_\lambda = 15.5, 21.3, 25.6$. The number of PLEs were found to be about 25, 60 and 100 respectively (see figure 2 of [29]). The decay rate of the Lyapunov spectrum was found to follow a power-law, $\lambda_i \approx \alpha(i-1)^\beta + \lambda_1$, with the exponent β in the region 0.81-0.85 (with the smaller value for the higher Re_τ). The maximum LE, λ_1, is expected to scale with the inverse of the Kolmogorov time scale, τ_η, see theoretical arguments in [30]. This was tested in [31] for HIT and it was found that $\lambda_1\tau_\eta$ is not constant, but instead grows with Re_λ following a power-law, $\lambda_1\tau_\eta \sim Re_\lambda^\gamma$. There is also growing evidence that covariant Lyapunov vectors (CLVs) corresponding to the largest LEs are spatially localised, see [23,32].

An alternative approach to tackle the sensitivity problem relies on exploitation of the underlying physical processes. For example, in the area of fluid flows, it is well known that momentum transfer is dominated by large scale structures, thus smaller scales (that correspond to the largest LEs) can be filtered out and have their effect modelled, hopefully without loss of accuracy in sensitivity. This is exactly what large eddy simulations (LES) are designed for [33]. In standard LES, the equations are filtered in space, but temporal filtering (with filter time scale Δ) is also possible [34]. In the limit of $\Delta \to \infty$, the temporally-averaged LES (TLES) equations tend to the standard Reynolds-Averaged Navier-Stokes (RANS) equations (section 2.3 of [34]). We conjecture therefore that as Δ increases, application of LSS to the TLES equations will recover the sensitivities predicted by the tangent (or adjoint) method applied to RANS. Thus, parameter Δ bridges two limits, LSS applied to unfiltered Navier-Stokes ($\Delta = 0$) and the RANS equations ($\Delta \to \infty$). As Δ increases, the number of PLEs decreases and the problem becomes more tractable. However, accuracy is traded for computational efficiency, because as Δ increases the effect of more scales needs to be modelled.

The above approach thus relies on accurate modelling of the filtered scales. An approach that allows truncation of small scale structures (and thus of the effect of largest PLEs) without the need for modelling is required. This has been achieved recently by projecting out the unstable CLVs from the tangent solutions, [35]. In the present paper, we follow a different approach; we formulate the LSS in the frequency domain (all previous variants have been derived in the time domain) and we show that it allows truncation of the highest frequencies (corresponding to smallest scales) without modelling. Moreover, it provides deep physical insight onto the system.

Frequency domain approaches have been applied from the perspective of linear [36] and non-linear input-output analysis [37], the frequency response of periodically time-varying base flows [38], or model-based design of transverse wall oscillations for turbulent drag reduction in a channel flow [39].

The paper is organised as follows. Section 2 sets scene and presents the standard LSS algorithm in the time domain. The formulation of the algorithm in the frequency domain is derived in section 3 followed by application to the

Kuramoto-Sivashinsky equation in section 4. A resolvent-based iterative algorithm to solve the resulting system is presented in section 5 and the results are further analysed in section 6, where we show that truncating frequencies with small energy content do not affect the accuracy of the results. We conclude in section 7.

2. Sensitivity analysis of chaotic systems using the shadowing approach

Consider a dynamical system governed by a set of ordinary differential equations of the form,

$$\begin{aligned} &\frac{d\boldsymbol{u}}{dt} = f(\boldsymbol{u};\boldsymbol{s}) \\ &\boldsymbol{u}(0;\boldsymbol{s}) = \boldsymbol{u}_0(\boldsymbol{s}), \end{aligned} \tag{1}$$

where $\boldsymbol{u}(t;\boldsymbol{s}) \in \mathbb{R}^{N_u}$ is the vector of state variables and $\boldsymbol{s} \in \mathbb{R}^{N_s}$ is the set of control parameters that define the dynamics of the system. System (1) can arise for example after spatial discretisation of a set of conservation laws that describe mathematically the problem under investigation. We assume that the vector field $\boldsymbol{f}(\boldsymbol{u};\boldsymbol{s}) : \mathbb{R}^{N_u} \times \mathbb{R}^{N_s} \to \mathbb{R}^{N_u}$ varies smoothly with $\boldsymbol{u}$ and $\boldsymbol{s}$.

In many applications we are interested in evaluating the sensitivity of a time-averaged quantity $\overline{J}(\boldsymbol{s}) : \mathbb{R}^{N_s} \to \mathbb{R}$,

$$\overline{J}(\boldsymbol{s}) = \lim_{T\to\infty} \frac{1}{T} \int_0^T J(\boldsymbol{u};\boldsymbol{s}) dt, \tag{2}$$

to the parameters $\boldsymbol{s}$. For example, in the area of aerodynamics, $\overline{J}(\boldsymbol{s})$ can be the drag coefficient and $\boldsymbol{s}$ the set of variables that describe the shape of an airfoil. The gradient of $\overline{J}(\boldsymbol{s})$ with respect to $\boldsymbol{s}$ is defined as

$$\begin{aligned} \frac{d\overline{J}}{d\boldsymbol{s}} &= \lim_{\delta\boldsymbol{s}\to 0} \frac{\overline{J}(\boldsymbol{s}+\delta\boldsymbol{s}) - \overline{J}(\boldsymbol{s})}{\delta\boldsymbol{s}} \\ &= \lim_{\delta\boldsymbol{s}\to 0} \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{J\left(\boldsymbol{u}(t;\boldsymbol{s}+\delta\boldsymbol{s});\boldsymbol{s}+\delta\boldsymbol{s}\right) - J\left(\boldsymbol{u}(t;\boldsymbol{s});\boldsymbol{s}\right)}{\delta\boldsymbol{s}} dt \end{aligned} \tag{3}$$

In chaotic systems, the limit and differentiation operations do not commute, i.e.

$$\begin{aligned} &\frac{d\overline{J}}{d\boldsymbol{s}} \neq \lim_{T\to\infty} \lim_{\delta\boldsymbol{s}\to 0} \frac{1}{T} \int_0^T \frac{J\left(\boldsymbol{u}(t;\boldsymbol{s}+\delta\boldsymbol{s});\boldsymbol{s}+\delta\boldsymbol{s}\right) - J\left(\boldsymbol{u}(t;\boldsymbol{s});\boldsymbol{s}\right)}{\delta\boldsymbol{s}} dt \\ &\left[= \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{dJ\left(\boldsymbol{u}(t;\boldsymbol{s});\boldsymbol{s}\right)}{d\boldsymbol{s}} dt = \lim_{T\to\infty} \frac{1}{T} \int_0^T \left(\frac{\partial J}{\partial \boldsymbol{u}} \boldsymbol{v} + \frac{\partial J}{\partial \boldsymbol{s}} \right) dt \right], \end{aligned} \tag{4}$$

where

$$\boldsymbol{v}(t) = \frac{d\boldsymbol{u}}{d\boldsymbol{s}} = \lim_{\delta\boldsymbol{s}\to 0} \frac{\boldsymbol{u}(t;\boldsymbol{s}+\delta\boldsymbol{s}) - \boldsymbol{u}(t;\boldsymbol{s})}{\delta\boldsymbol{s}} \tag{5}$$

is the sensitivity of the solution $\boldsymbol{u}(t;\boldsymbol{s})$ to a change $\delta\boldsymbol{s}$ of $\boldsymbol{s}$. The reason is that chaotic systems have one or more PLEs. This means that the distance in phase space between $\boldsymbol{u}(t;\boldsymbol{s}+\delta\boldsymbol{s})$ and $\boldsymbol{u}(t;\boldsymbol{s})$, i.e. the Euclidean norm of the vector $\boldsymbol{u}(t;\boldsymbol{s}+\delta\boldsymbol{s}) - \boldsymbol{u}(t;\boldsymbol{s})$ that appears in the nominator of (5), grows exponentially at rate $\sim e^{\lambda_1 t}$, where λ_1 is the maximum of these exponents [2,8,40]. Thus, the quantity $\frac{1}{T}\int_0^T \left(\frac{\partial J}{\partial \boldsymbol{u}}\boldsymbol{v} + \frac{\partial J}{\partial \boldsymbol{s}}\right) dt$ that appears on the right hand side of (4) diverges as $T \to \infty$. On the other hand, assuming that $\overline{J}(\boldsymbol{s})$ varies smoothly with $\boldsymbol{s}$, the sensitivity $\frac{d\overline{J}}{d\boldsymbol{s}}$ is finite.

If the dynamical system (1) is uniformly hyperbolic, the shadowing lemma [41] guarantees the existence of a solution trajectory evaluated at $\boldsymbol{s}+\delta\boldsymbol{s}$ that remains always close, i.e. shadows indefinitely, the reference trajectory $\boldsymbol{u}(t;\boldsymbol{s})$. We denote this shadowing trajectory as $\boldsymbol{u}(\tau(t);\boldsymbol{s}+\delta\boldsymbol{s})$, where $\tau(t)$ is an appropriate time transformation. The LSS method, proposed in [1], computes the shadowing trajectory by minimising the distance between $\boldsymbol{u}(\tau(t);\boldsymbol{s}+\delta\boldsymbol{s})$ and $\boldsymbol{u}(t;\boldsymbol{s})$ in a least squares sense, i.e.

$$\min_{\boldsymbol{u},\tau} \frac{1}{2} \int_0^T \left(\left\| \boldsymbol{u}\left(\tau(t);\boldsymbol{s}+\delta\boldsymbol{s}\right) - \boldsymbol{u}(t;\boldsymbol{s}) \right\|^2 + \alpha^2 \left\| \frac{d\tau}{dt} - 1 \right\|^2 \right) dt \; s.t. \tag{6a}$$

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{f}\left(\boldsymbol{u}\left(\tau(t);\boldsymbol{s}+\delta\boldsymbol{s}\right)\right), \tag{6b}$$

where α^2 is a constant parameter. Taking the limit $\delta \boldsymbol{s} \to 0$, leads to the following linear minimisation problem,

$$\min_{\boldsymbol{v},\eta} \frac{1}{2} \int_0^T \left(\|\boldsymbol{v}(t)\|^2 + \alpha^2 \|\eta\|^2 \right) dt \; s.t. \tag{7a}$$

$$\frac{d\boldsymbol{v}}{dt} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} \boldsymbol{v} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}} + \eta(t) \boldsymbol{f}, \tag{7b}$$

where

$$\boldsymbol{v}(t) = \frac{d}{d\boldsymbol{s}} \boldsymbol{u}(\tau(t); \boldsymbol{s}) \tag{8a}$$

$$\eta(t) = \frac{d}{d\boldsymbol{s}} \left(\frac{d\tau}{dt} \right). \tag{8b}$$

The second term within the cost functions (6a) and (7a) penalises the deviation of $\tau(t)$ from t. A high value of α^2 results in a small deviation (heavy penalisation), while a small value to light penalisation. The solution of (7) for $\alpha^2 = 0$, leads to the orthogonality condition between the vectors $\boldsymbol{f}(\boldsymbol{u}; s)$ and $\boldsymbol{v}(\boldsymbol{u}; s)$ at each point along the trajectory, i.e. $\langle \boldsymbol{f}(\boldsymbol{u}; \boldsymbol{s}) \boldsymbol{v}(\boldsymbol{u}; \boldsymbol{s}) \rangle = 0$, a constraint from which $\eta(t)$ can be obtained [24]. Thus, problem (7) becomes

$$\min_{\boldsymbol{v},\eta} \frac{1}{2} \int_0^T \|\boldsymbol{v}(t)\|^2 dt \; s.t. \tag{9a}$$

$$\frac{d\boldsymbol{v}}{dt} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} \boldsymbol{v} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}} + \eta(t) \boldsymbol{f} \tag{9b}$$

$$\langle \boldsymbol{f}(\boldsymbol{u}; \boldsymbol{s}) \boldsymbol{v}(\boldsymbol{u}; \boldsymbol{s}) \rangle = 0 \tag{9c}$$

From the solution $\boldsymbol{v}_{lss}(t)$ and $\eta_{lss}(t)$ of (9), the sensitivity $\frac{d\bar{J}}{d\boldsymbol{s}}$ can be easily computed [1].

3. Formulation of the shadowing algorithm in Fourier space

In this section, we formulate the problem in the frequency domain, i.e. in Fourier space, and seek a solution that remains bounded. To this end, we consider a reference trajectory $\boldsymbol{u}(t; \boldsymbol{s})$ of length T and assume that the solution of the minimisation problem (9) is periodic with period T. Thus, it can be written in terms of Fourier series as,

$$\boldsymbol{v}(t) = \sum_{k=-\infty}^{+\infty} \hat{\boldsymbol{v}}_k e^{ik\omega_0 t}, \; \eta(t) = \sum_{k=-\infty}^{+\infty} \hat{\eta}_k e^{ik\omega_0 t}, \tag{10}$$

where $\hat{\boldsymbol{v}}_k, \hat{\eta}_k$ denote the Fourier coefficients, $\omega_0 = \frac{2\pi}{T}$ is the fundamental angular frequency and the index k characterises the harmonics with frequencies $\omega_k = k\omega_0$. We assume similar series expansions for the Jacobian $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t)$ and $\boldsymbol{f}(t)$,

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t) = \sum_{k=-\infty}^{+\infty} \left(\widehat{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} \right)_k e^{ik\omega_0 t}, \; \boldsymbol{f}(t) = \sum_{k=-\infty}^{+\infty} \hat{\boldsymbol{f}}_k e^{ik\omega_0 t}. \tag{11}$$

The matrix-vector product $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t) \boldsymbol{v}(t)$ can be written as

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} \boldsymbol{v} = \sum_{k=-\infty}^{+\infty} \left(\widehat{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} \boldsymbol{v}} \right)_k e^{ik\omega_0 t}, \tag{12}$$

where

$$\left(\widehat{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} \boldsymbol{v}} \right)_k = \sum_{l=-\infty}^{+\infty} \left(\widehat{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} \right)_{k-l} \hat{\boldsymbol{v}}_l, \tag{13}$$

which is the convolution sum between the Fourier coefficients of $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t)$ and $\boldsymbol{v}(t)$. Similarly, the left hand side of the orthogonality condition (9c) can be expanded as

$$\langle \boldsymbol{f}(\boldsymbol{u}; \boldsymbol{s}) \boldsymbol{v}(\boldsymbol{u}; \boldsymbol{s}) \rangle = \boldsymbol{f}^\top \boldsymbol{v} = \sum_{k=-\infty}^{+\infty} \left(\widehat{\boldsymbol{f}^\top \boldsymbol{v}} \right)_k e^{ik\omega_0 t}, \tag{14}$$

where

$$\left(\widehat{\boldsymbol{f}^\top \boldsymbol{v}}\right)_k = \sum_{l=-\infty}^{+\infty} \hat{\boldsymbol{f}}_{k-l}^\top \hat{\boldsymbol{v}}_l, \tag{15}$$

and the notation $()^\top$ denotes the transpose operation. In the above two expressions, we have assumed that the weighting matrix associated with the inner product is the identity matrix, but the analysis below can be easily generalised to an inner product defined as $\langle \boldsymbol{f}(\boldsymbol{u};\boldsymbol{s})\boldsymbol{v}(\boldsymbol{u};\boldsymbol{s})\rangle = \boldsymbol{f}^\top \mathcal{Q}\boldsymbol{v}$. Finally,

$$\eta(t)\boldsymbol{f}(t) = \sum_{k=-\infty}^{+\infty} \left(\widehat{\eta \boldsymbol{f}}\right)_k e^{ik\omega_0 t} \quad \text{with} \quad \left(\widehat{\eta \boldsymbol{f}}\right)_k = \sum_{l=-\infty}^{+\infty} \hat{\boldsymbol{f}}_{k-l}\hat{\eta}_l. \tag{16}$$

In practice, the range of the index k is truncated to lie within the interval $[-q, q]$. For example, if the reference trajectory is sampled every Δt, $q = \frac{T}{2\Delta t} - 1$. The frequency spectrum of $\boldsymbol{u}(t;\boldsymbol{s})$ can also indicate the number of spectral coefficients that must be retained. A finite q amounts to applying a sharp spectral cut-off filter to the above expansions, where all coefficients with $|k|, |l|$ or $|k-l| > q$ are set equal to 0.

Introducing the finite spectral representations to (9b) and (9c), yields the following block set of equations for the k-th pair of coefficients $\hat{\boldsymbol{v}}_k, \hat{\eta}_k$,

$$ik\omega_0 \mathcal{I}_u \begin{bmatrix} \hat{\boldsymbol{v}}_k \\ \hat{\eta}_k \end{bmatrix} - \sum_{l=-q}^{q} T_{k-l} \begin{bmatrix} \hat{\boldsymbol{v}}_l \\ \hat{\eta}_l \end{bmatrix} = \begin{bmatrix} \frac{d\hat{\boldsymbol{f}}_k}{d\boldsymbol{s}} \\ 0 \end{bmatrix} \tag{17}$$

where

$$\mathcal{I}_u = \begin{bmatrix} \mathcal{I}_{N_u} & 0 \\ 0 & 0 \end{bmatrix}, \tag{18}$$

$$T_m = \begin{bmatrix} \left(\widehat{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\right)_m & \hat{\boldsymbol{f}}_m \\ \hat{\boldsymbol{f}}_m^\top & 0 \end{bmatrix}, \tag{19}$$

and $\mathcal{I}_{N_u}$ is the identity matrix of dimension N_u. Each block consists of $N_u + 1$ equations and there are in total $2q+1$ blocks, resulting in $(N_u+1)\times(2q+1)$ equations and unknowns.

Due to the sharp spectral cut-off filter mentioned earlier, the starting and final values of index l in the summation are slightly modified when $k \neq 0$. For example, for $k = -q$ equation (17) becomes

$$i(-q)\omega_0 \mathcal{I}_u \begin{bmatrix} \hat{\boldsymbol{v}}_{-q} \\ \hat{\eta}_{-q} \end{bmatrix} - \sum_{l=-q}^{0} T_{-q-l} \begin{bmatrix} \hat{\boldsymbol{v}}_l \\ \hat{\eta}_l \end{bmatrix} = \begin{bmatrix} \frac{d\hat{\boldsymbol{f}}_{-q}}{d\boldsymbol{s}} \\ 0 \end{bmatrix}, \tag{20}$$

while for $k = +q$,

$$iq\omega_0 \mathcal{I}_u \begin{bmatrix} \hat{\boldsymbol{v}}_q \\ \hat{\eta}_q \end{bmatrix} - \sum_{l=0}^{q} T_{q-l} \begin{bmatrix} \hat{\boldsymbol{v}}_l \\ \hat{\eta}_l \end{bmatrix} = \begin{bmatrix} \frac{d\hat{\boldsymbol{f}}_q}{d\boldsymbol{s}} \\ 0 \end{bmatrix}. \tag{21}$$

Stacking the blocks together one below the other results in a linear system of equations that takes the form

$$\left[\mathcal{D} - \mathcal{T}(T_m)\right]\widehat{\mathcal{V}} = \widehat{\mathcal{R}}, \tag{22}$$

where $\widehat{\mathcal{V}} = \left[\widehat{\mathcal{V}}_{-q}, \ldots, \widehat{\mathcal{V}}_0, \ldots, \widehat{\mathcal{V}}_q\right]^\top$ is the block vector of unknowns, with $\widehat{\mathcal{V}}_k = [\hat{\boldsymbol{v}}_k, \hat{\eta}_k]^\top$. The right hand side is $\widehat{\mathcal{R}} = [\widehat{\mathcal{R}}_{-q}, \ldots, \widehat{\mathcal{R}}_0, \ldots, \widehat{\mathcal{R}}_q]^\top$ with $\widehat{\mathcal{R}}_k = \left[\frac{d\hat{\boldsymbol{f}}_k}{d\boldsymbol{s}}, 0\right]^\top$, while $\mathcal{D}$ is the block diagonal matrix $\mathcal{D} = diag[\mathcal{D}_{-q}, \ldots, \mathcal{D}_0, \ldots, \mathcal{D}_q]$, with $\mathcal{D}_k = ik\omega_0\mathcal{I}_u$. Matrix $\mathcal{T}(T_m)$ has a block Toeplitz form,

$$\mathcal{T}(T_m) = \begin{bmatrix} T_0 & T_{-1} & \ldots & T_{-q} & & & \\ T_1 & T_0 & T_{-1} & \ldots & T_{-q} & & \\ & \ddots & \ddots & \ddots & \ldots & \ddots & \\ T_q & \ldots & T_1 & T_0 & T_{-1} & \ldots & T_{-q} \\ & \ldots & \ddots & \ddots & \ddots & \ddots & \ldots \\ & & T_q & \ldots & T_1 & T_0 & T_{-1} \\ & & & T_q & \ldots & T_1 & T_0 \end{bmatrix}, \tag{23}$$

with the same blocks in each diagonal.

System (22) can be also written in expanded matrix form as

$$\begin{bmatrix} \mathcal{D}_{-q}-T_0 & -T_{-1} & \dots & -T_{-q} & & & \\ -T_1 & \mathcal{D}_{-q+1}-T_0 & -T_{-1} & \dots & -T_{-q} & & \\ & \ddots & \ddots & \ddots & \dots & \ddots & \\ -T_q & \dots & -T_1 & \mathcal{D}_0-T_0 & -T_{-1} & \dots & -T_{-q} \\ & \dots & \ddots & \ddots & \ddots & \ddots & \dots \\ & & -T_q & \dots & -T_1 & \mathcal{D}_{q-1}-T_0 & -T_{-1} \\ & & & -T_q & \dots & -T_1 & \mathcal{D}_q-T_0 \end{bmatrix} \begin{bmatrix} \widehat{\mathcal{V}}_{-q} \\ \widehat{\mathcal{V}}_{-q+1} \\ \vdots \\ \widehat{\mathcal{V}}_0 \\ \vdots \\ \widehat{\mathcal{V}}_{q-1} \\ \widehat{\mathcal{V}}_q \end{bmatrix} = \begin{bmatrix} \widehat{\mathcal{R}}_{-q} \\ \widehat{\mathcal{R}}_{-q+1} \\ \vdots \\ \widehat{\mathcal{R}}_0 \\ \vdots \\ \widehat{\mathcal{R}}_{q-1} \\ \widehat{\mathcal{R}}_q \end{bmatrix}. \tag{24}$$

The above matrix, known as Hill matrix [42,43], contains square blocks with dimensions $(N_u+1)\times(N_u+1)$. The solution of system (22) can be written symbolically as

$$\widehat{\mathcal{V}} = \mathcal{H}\widehat{\mathcal{R}}, \tag{25}$$

where

$$\mathcal{H} = \left[\mathcal{D} - \mathcal{T}(T_m)\right]^{-1}, \tag{26}$$

is the matrix that maps the input $\widehat{\mathcal{R}}$ to the output $\widehat{\mathcal{V}}$. As $q\to\infty$, $\mathcal{H}$ becomes an operator, termed here shadowing harmonic operator. Note that equations (9b) and (9c) form a linear, time-varying periodic system,

$$\begin{bmatrix} \mathcal{I}_{N_u} & 0 \\ 0 & 0 \end{bmatrix} \frac{d}{dt}\begin{bmatrix} \boldsymbol{v} \\ \eta \end{bmatrix} = \begin{bmatrix} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} & \boldsymbol{f} \\ \boldsymbol{f}^\top & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{v} \\ \eta \end{bmatrix} + \begin{bmatrix} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}} \\ 0 \end{bmatrix}. \tag{27}$$

Thus, the shadowing harmonic operator is identical to the standard harmonic operator, defined in [43], applied to the above system. The properties of this operator will be examined in the next section for the Kuramoto-Sivashinsky equation.

The sensitivity $\frac{d\overline{J}}{ds}$ is computed from [1],

$$\frac{d\overline{J}}{ds} = \frac{1}{T}\int_0^T \frac{\partial J}{\partial \boldsymbol{u}}\boldsymbol{v} + \frac{\partial J}{\partial \boldsymbol{s}} + \eta(t)(J(\boldsymbol{u};\boldsymbol{s};t) - \overline{J})dt, \tag{28}$$

which in the frequency domain can be written as,

$$\frac{d\overline{J}}{ds} = \left(\widehat{\frac{\partial J}{\partial \boldsymbol{s}}}\right)_0 + \sum_{k=-q}^{q}\left(\widehat{\frac{\partial J}{\partial \boldsymbol{u}}}\right)_{-k}\hat{\boldsymbol{v}}_k + \sum_{\substack{k=-q \\ k\neq 0}}^{q}\hat{\eta}_{-k}\hat{J}_k \tag{29}$$

In the above formulation, the reference trajectory $\boldsymbol{u}(t;\boldsymbol{s})$ was expanded using the same number of Fourier modes as the solution, $2q+1$, and each diagonal of the block Toeplitz matrix $\mathcal{T}(T_m)$ corresponds to one harmonic of $\boldsymbol{u}(t;\boldsymbol{s})$. This is however not necessary. If the quantity of interest $\overline{J}(\boldsymbol{s})$ is known to depend on a particular frequency region, then only the relevant diagonals of $\mathcal{T}(T_m)$ need to be retained (we apply this idea in section 6). For example integral quantities, such as lift and drag, are known to depend predominantly on the effect of large scale structures that usually have a footprint at low frequencies. Retaining only the relevant frequencies and truncating the rest can lead to significant savings in the storage requirements and solution time of system (24). In the limiting case, where only the time-average of $\boldsymbol{u}(t;\boldsymbol{s})$ is retained, the equations decouple and the k-th component $\widehat{\mathcal{V}}_k = [\hat{\boldsymbol{v}}_k, \hat{\eta}_k]^\top$ can be obtained from

$$(\mathcal{D}_k - T_0)\,\widehat{\mathcal{V}}_k = \widehat{\mathcal{R}}_k \tag{30}$$

In this case, the harmonic balancing method becomes identical to the standard resolvent analysis [36].

Some comments are warranted here to clarify an underlying assumption of the above formulation. Suppose that the reference trajectory $\boldsymbol{u}(t;\boldsymbol{s})$, and thus $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t)$ and $\boldsymbol{f}(t)$, are exactly periodic with period T and the minimisation problem (9) is solved in the time domain using the multiple shooting shadowing method, [24,25]. If the trajectory is sampled at points $t_0\dots t_K$ during T, then the solution $\boldsymbol{v}(t_i), \eta(t_i)$ is sought at $K+1$ points $i=0\dots K$, thus there are $2(K+1)$ unknowns. There are K intervals, thus (9b) provides K equations, while (9c) provides additional $K+1$ equations. The remaining equation arises from the solution of the minimisation problem (9a). The solution of this problem will not necessarily yield a periodic solution, i.e. $\boldsymbol{v}(t_0)$ will not necessarily be equal to $\boldsymbol{v}(t_K)$.

In the above formulation, we have not explicitly considered the minimisation of the cost function (9a); instead we closed the system assuming periodicity, i.e. $\boldsymbol{v}(t_0)=\boldsymbol{v}(t_K)$ and $\eta(t_0)=\eta(t_K)$. The same assumption is made in the periodic

shadowing method of Lasagna [44], and leads to a sensitivity error that initially decays at a rate $1/T$, followed by the asymptotic rate $1/\sqrt{T}$ (the latter dictated by the central limit theorem). There is however an important difference compared to the present method: In [44], the time transformation $\tau(t)$ is linear with respect to t, leading to a constant $\eta(t)$. In the present method, we do not prescribe any form of $\tau(t)$, so $\eta(t)$ is unknown and is obtained by imposing the orthogonality constraint (9c) at every point along the trajectory. Another difference is that our method is formulated in the frequency domain instead of the time domain. As will be seen later, this allows one to obtain deep physical insight on the dominant factors that determine the sensitivity, which is not possible in the time domain.

An approach closely related to the present one was proposed recently by Padovan et al. [38]. The authors perform a frequency-domain analysis of periodic perturbations about a periodically time-varying base flow. If the base flow satisfies the governing equations, the perturbations are governed by $\frac{d\boldsymbol{q}'}{dt} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}\boldsymbol{q}' + \boldsymbol{g}'$, which is similar to (9b). The authors restrict the harmonic operator to a subspace that is orthogonal to the direction of the phase shift given by $\hat{\boldsymbol{f}}_k$. This is achieved by projecting out of $\hat{\boldsymbol{g}}'_k$ (the Fourier coefficients of the forcing $\boldsymbol{g}'(t)$) the component that would lead to a non-zero projection of the solution $\hat{\boldsymbol{q}}'_k$ to $\hat{\boldsymbol{f}}_k$. In the present paper, we seek the sensitivity with respect to $\boldsymbol{s}$, thus the forcing vector $\boldsymbol{g}'(t)$ takes the particular form $\boldsymbol{g}'(t) = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}(t)$. This vector is subsequently modified by adding $\eta(t)\boldsymbol{f}(t)$, where the time dilation $\eta(t)$ is computed so that the orthogonality constraint is satisfied at all time instants, as already mentioned.

In the above frequency-domain formulation, we seek a harmonic solution that remains bounded, see expansions (10), and does not suffer from exponentially growing terms. There is a price to pay however; the frequencies are all coupled together leading to large storage requirements for the Hill matrix, see (24). The properties of the shadowing harmonic operator will be examined in the next section for the Kuramoto-Sivashinsky equation. This test case is small enough that the Hill matrix can be stored and the linear system (24) is solved directly with LU decomposition. In section 5 we propose an iterative method that mitigates the storage and solution time requirements.

4. Application to the Kuramoto-Sivashinsky equation

We apply the method proposed in the previous section to the Kuramoto Sivashinsky (KS) equation, which displays complex spatio-temporal chaos and is frequently used in the literature as a test case for chaotic systems [45]. The equation takes the form

$$\begin{aligned} &\frac{\partial u}{\partial t} = -\left[(u+c)\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial^4 u}{\partial x^4}\right] \text{ with } x \in [0, L], \\ &u(0,t) = u(L,t) = 0, \\ &\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(L,t) = 0, \end{aligned} \tag{31}$$

where c is an artificially introduced parameter [46]. The term $\frac{\partial^2 u}{\partial x^2}$ is responsible for energy production, while $\frac{\partial^4 u}{\partial x^4}$ adds dissipation to the system. We set $L = 128$ to generate chaotic solutions [45] and discretise (31) with a second order finite difference scheme with $\delta x = 1$. For $c = 0$, the dynamical system has 16 positive Lyapunov exponents, the largest of which is $\lambda_1 = 0.093$ [46].

Two objective functions are considered, the space-time average of the state $u(x,t;c)$

$$\overline{J_1}(c) = \frac{1}{TL}\int_0^T\int_0^L u(x,t;c)dx\,dt = \frac{1}{L}\int_0^L \overline{u}(x;c)dx \tag{32}$$

and of the total kinetic energy

$$\overline{J_2}(c) = \frac{1}{TL}\int_0^T\int_0^L u^2(x,t;c)dx\,dt = \frac{1}{L}\int_0^L \overline{u^2}(x;c)dx = \frac{1}{L}\int_0^L \left(\overline{u}^2 + \overline{u'^2}\right)dx \tag{33}$$

where an overbar $\overline{()}$ denotes time-average and a prime $()'$ the fluctuation around the average, i.e. $u = \overline{u} + u'$. We seek their sensitivities with respect to c, i.e. $\frac{d\overline{J_1}(c)}{dc}$ and $\frac{d\overline{J_2}(c)}{dc}$, thus $s = c$.

The time-averages $\overline{u}(x)$, $\overline{u^2}(x)$ and $\overline{\frac{df}{dc}}(x)$ for $c = 0$ are plotted against x in Fig. 1. It can be seen that $\overline{u}(x)$, shown in panel (a), takes both positive and negative values, so it is not intuitively evident which part of the domain contributes most to the integral $\overline{J_1}(c)$ and its sensitivity. The kinetic energy $\overline{u^2}(x)$, shown in panel (b), has an almost symmetric shape with peaks close to the boundaries, but remains flat in the middle of the domain. In the latter region, $\overline{u}(x)$ attains low values, so the total kinetic energy consists mainly of the fluctuation energy, $\overline{u'^2}$. Finally, the time-average forcing term $\overline{\frac{df}{dc}}(x)$, shown in panel (c), fluctuates around 0 close to the boundaries and is almost negligible elsewhere. Again, it is not trivial to foresee how this term will affect the sensitivities of $\overline{J_1}(c)$ and $\overline{J_2}(c)$ with respect to c.

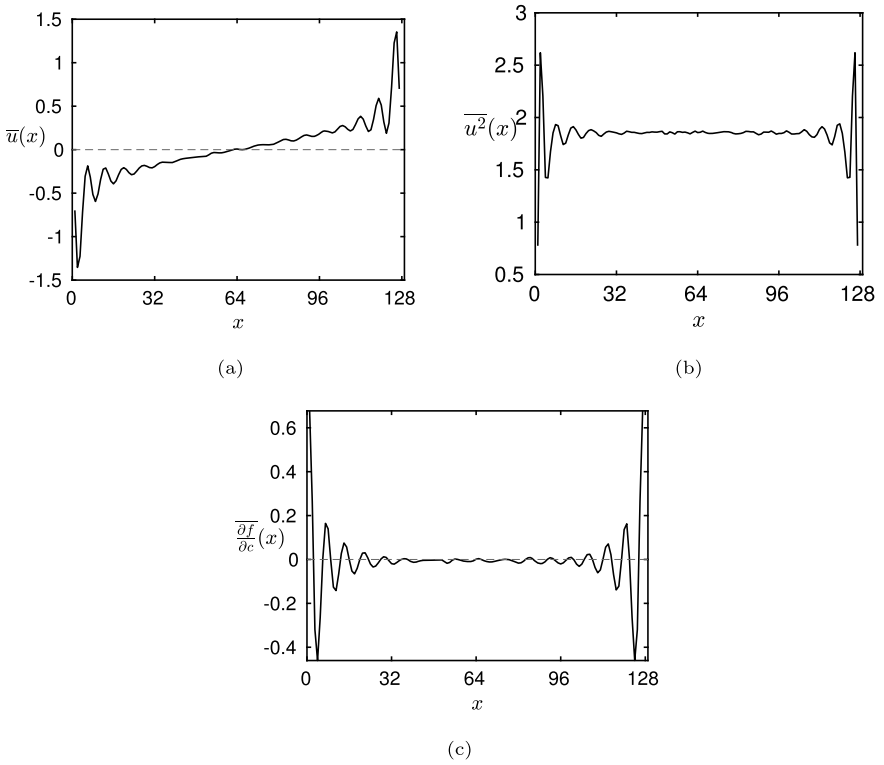


Fig. 1. Spatial distribution of the time-averages (a) $\overline{u}(x)$, (b) $\overline{u^2}(x)$ and (c) $\overline{\frac{df}{dc}}(x)$ for $c = 0$ and $T = 100$. The results are averaged over 1000 random initial conditions.

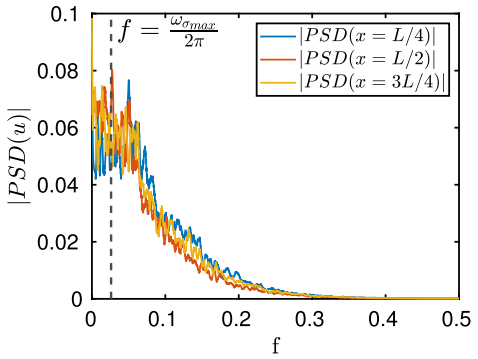


Fig. 2. Spectra of $u(x,t)$ at $x = \frac{L}{4}$, $\frac{L}{2}$ and $\frac{3L}{4}$ for $c = 0$ obtained from a trajectory with length $T = 10000$. The spectra were smoothed with a $5-th$ order Savitzky-Golay convolution filter with 5 averaging windows. The vertical dashed line indicates the angular frequency $\omega \approx 0.164$ that gives the maximum gain $\sigma(\omega)$ in Fig. 13.

The Power Spectral Densities (PSD) of $u(x,t)$ against frequency at three locations for $c = 0$ are shown in Fig. 2. The plot indicates that the solution has energy footprint in the frequency range $[0., 0.3]$, with the highest energy content in the region $[0, 0.07]$. We exploit this feature in section 6 of the paper.

From equation (29), we have

$$\frac{d\overline{J_1}}{dc} = \frac{1}{L}\int_0^L \left(\hat{\boldsymbol{v}}_0 + \sum_{\substack{k=-q \\ k\neq 0}}^{q} \hat{\eta}_{-k}\hat{\boldsymbol{u}}_k \right) dx, \tag{34}$$

and

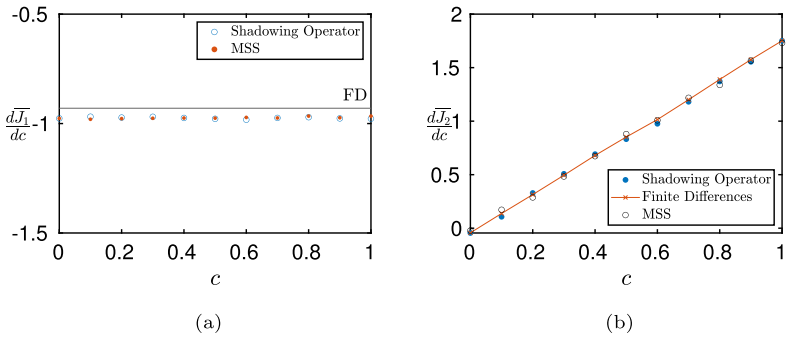


Fig. 3. Sensitivities $\frac{d\overline{J_1}}{dc}$ (a) and $\frac{d\overline{J_2}}{dc}$ (b) for various values of parameter c, obtained with $T = 100$ and $f \in [-0.3, 0.3]$. The finite difference (FD) data were taken from [46]. The MSS values were obtained using $K = 10$ time segments of length $\Delta T = 10$. Values were averaged over 100 random initial conditions.

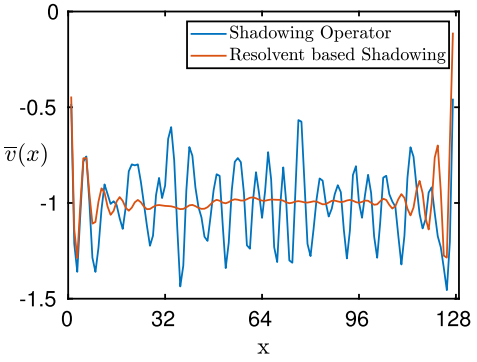


Fig. 4. Comparison of $\overline{\boldsymbol{v}}(x) = \hat{\boldsymbol{v}}_0(x)$ between the harmonic shadowing operator and the resolvent-based approximation (described in section 5), for $c = 0$ and $T = 100$, averaged over 100 random initial conditions.

$$\frac{d\overline{J_2}}{dc} = \frac{1}{L}\int_0^L \left(2\sum_{k=-q}^{q} \boldsymbol{u}_{-k}\hat{\boldsymbol{v}}_k + \sum_{\substack{k=-q \\ k\neq 0}}^{q} \hat{\eta}_{-k}\widehat{\boldsymbol{u}^2}_k \right) dx. \tag{35}$$

The above analytical expressions can be interpreted physically. The distribution of the integrands over x reveals the parts of the domain that mostly affect the quantity of interest; this information can be useful for control applications for example.

As can be seen from Fig. 3, the sensitivities obtained using the shadowing harmonic operator match very well with those obtained using the preconditioned MSS [25]. A comparison with finite difference (FD) data is also presented. For $\frac{d\overline{J_1}}{dc}$, there is a small bias, which has also been observed in the time-domain formulation of the method [25,26,46]; for $\frac{d\overline{J_2}}{dc}$ the matching with FD is very good.

Fig. 4 shows the variation of $\overline{\boldsymbol{v}}(x) = \hat{\boldsymbol{v}}_0(x)$ against x. The values fluctuate around -1 almost uniformly in the whole domain, thus changing c to $c + \delta c$ reduces $\overline{u}(x)$ everywhere by the same amount on average.

In Fig. 5, we plot the normalised difference in the sensitivity $\frac{d\overline{J_1}}{dc}$ from MSS and the shadowing harmonic method, $\epsilon = \frac{\left(\frac{d\overline{J_1}}{dc}\right)_{SH} - \left(\frac{d\overline{J_1}}{dc}\right)_{MSS}}{\left(\frac{d\overline{J_1}}{dc}\right)_{MSS}}$, against T. The difference initially decays at a rate $\frac{1}{T}$ (similarly to [44]) and then at $\frac{1}{\sqrt{T}}$ (as dictated by the central limit theorem). The frequencies considered are $f \in [-0.3, 0.3]$, for $c = 0$. Each point was obtained by averaging over a large number of initial conditions dictated by the value of T and enough for the sensitivity to converge to two significant digits.

Evaluating the sensitivities using the preconditioned MSS with $K = 10$ segments requires approximately 55 s in a 3.2 GHz Intel i7-8700 CPU. On the other hand, the shadowing harmonic method requires an order of magnitude less time (approximately 3 s). This comparison however is case dependent. The storage requirements are larger, while the cost of the LU decomposition of the block-diagonal system (24) scales linearly with the number of diagonals and the cost of inversion of a single block. Below we investigate in more detail the properties of the shadowing harmonic operator, while in the next section 5 we explore an approach that can mitigate the rapid growth of computational cost and storage requirements for larger systems.

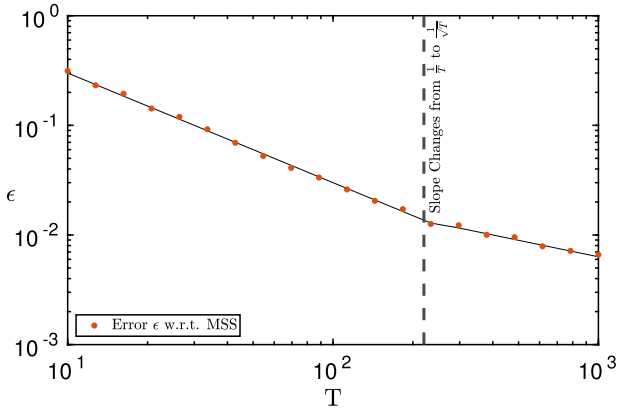


Fig. 5. Normalised difference in sensitivity $\frac{d\overline{J_1}}{dc}$ as computed by the shadowing harmonic method and MSS. The vertical dashed line marks the value of T at which the decay rate changes from $\frac{1}{T}$ to $\frac{1}{\sqrt{T}}$.

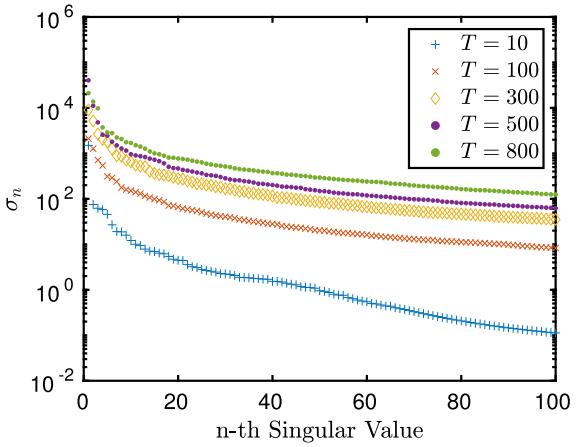


Fig. 6. Leading singular values of $\mathcal{H}$ for different values of T and $c = 0$. The value of q varies with T such that the resolved frequencies are always in the interval $[-0.3, 0.3]$.

4.1. Singular value decomposition of the shadowing harmonic operator

The singular values σ_i of the shadowing resolvent matrix $\mathcal{H}$, defined in (26), are obtained from the solution of the eigenvalue problem,

$$\mathcal{H}^*\mathcal{H}\mathbf{\Phi} = \sigma^2\mathbf{\Phi}. \tag{36}$$

The solution maximises the system gain, defined as the ratio of the (squared) 2-norm of the output (response $\widehat{\mathcal{V}}$) to that of the input (forcing $\widehat{\mathcal{R}}$), i.e.

$$\sigma^2 = \max_{\widehat{\mathcal{R}}} \frac{\|\widehat{\mathcal{V}}\|_2^2}{\|\widehat{\mathcal{R}}\|_2^2}, \tag{37}$$

where $\|\widehat{\mathcal{V}}\|_2^2 = \widehat{\mathcal{V}}^*\widehat{\mathcal{V}}$ and $\|\widehat{\mathcal{R}}\|_2^2 = \widehat{\mathcal{R}}^*\widehat{\mathcal{R}}$. Since $\delta x = 1$, the norms represent the discrete values of the integrals over the domain, for example $\|\widehat{\mathcal{V}}\|_2^2 = \widehat{\mathcal{V}}^*\widehat{\mathcal{V}} = \sum_{k=-q}^{q} \hat{\boldsymbol{v}}_k^*\hat{\boldsymbol{v}}_k + \hat{\eta}_k^*\hat{\eta}_k = \sum_{k=-q}^{q} \hat{\boldsymbol{v}}_{-k}\hat{\boldsymbol{v}}_k + \hat{\eta}_{-k}\hat{\eta}_k \approx \int_0^L \overline{\boldsymbol{v}(t)^2}dx + \overline{\eta(t)^2}$. The first term represents physically (twice) the time-average kinetic energy of the response integrated over the domain. We could have defined a weighted 2-norm that eliminates the presence of $\overline{\eta(t)^2}$, but we did not pursue this. Other values of δx could have been easily accounted for (again by defining an appropriately weighted 2-norm). The eigenvectors $\mathbf{\Phi}_i$ are the right singular vectors of $\mathcal{H}$, and represent the optimal forcings, while the corresponding responses are the left singular vectors $\mathbf{\Psi}_i = \sigma_i^{-1}\mathcal{H}\mathbf{\Phi}_i$.

Fig. 6 shows the 100 leading singular values for different T. Note that the largest singular value of $\mathcal{H}$, $\sigma_{max}^{\mathcal{H}}$, grows as T increases, thus the smallest singular value of $\mathcal{H}^{-1} = \mathcal{D} - \mathcal{T}(T_m)$, $\sigma_{min}^{\mathcal{H}^{-1}} = \frac{1}{\sigma_{max}^{\mathcal{H}}}$, is reduced. The same behaviour is observed when the problem is formulated in the time domain [24] and it is related to the lack of uniform hyperbolicity of the

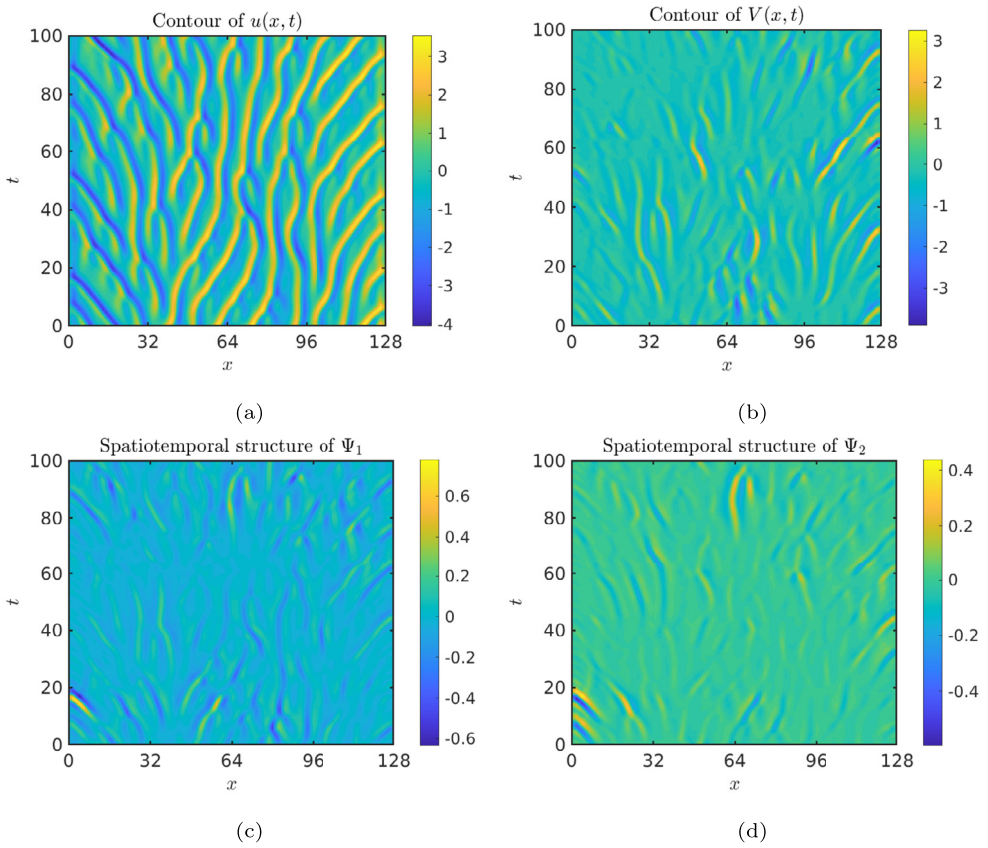


Fig. 7. Contours of (a) $\boldsymbol{u}(x,t)$, (b) the sensitivity $\boldsymbol{v}(x,t)$, (c) the first $\boldsymbol{\Psi}_1$ and (d) the second $\boldsymbol{\Psi}_2$ optimal responses in the $x-t$ plane. Results are shown for $T=100$, $c=0$ and $f\in[-0.3,0.3]$.

system. For such systems, the covariant Lyapunov vectors (CLVs) align at different locations along the trajectory and they are no longer linearly independent (this is equivalent to the eigenvectors of a matrix becoming parallel). This leads to local tangencies that result in the reduction of $\sigma_{min}^{\mathcal{H}^{-1}}$ at T increases. On the other hand, for uniform hyperbolic systems, the angle between CLVs stays away from 0 and $\sigma_{min}^{\mathcal{H}^{-1}}$ remains bounded. For a discussion of the angles between CLVs for the turbulent flow around a cylinder see [23]. Note also that as T grows, apart from the largest singular values, the rest start to converge. This indicates that they represent the true behaviour of the system, i.e. they are physically meaningful.

Contour plots of $u(x,t)$ in the $x-t$ plane for one realisation (with $c=0$ and $T=100$) are shown in Fig. 7a. Note the oscillatory back and forth motion in the middle region of the domain $[L/4,3L/4]$ that is captured in the spectra of Fig. 2. The sensitivity $v(x,t)$ (obtained with $q=30$) for the same realisation is shown in 7b. It follows a pattern similar to $u(x,t)$, but it is spotty (note the highly localised large positive and negative values), probably due to the aforementioned tangencies. Figs. 7c and 7d show contours of the spatio-temporal distribution of the optimal responses corresponding to the first (i.e. largest) and second singular values, i.e. $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$. The maps show again a wavy behaviour and are also locally spotty. There is evidence that the CLVs corresponding to largest LEs are spatially localised [35,23], and it is interesting to see that $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ exhibit similar behaviour.

Contours of the actual forcing $\frac{df}{ds}(x,t)$ in the $x-t$ plane are shown in Fig. 8a for the same realisation as in the previous figure. The spatio-temporal distributions of the optimal forcings that correspond to the first three singular values, i.e. $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$ and $\boldsymbol{\Phi}_3$, are shown in panels (b)-(d) respectively. The distributions are more difficult to interpret physically, but note the significant differences with respect to the actual forcing. It is interesting to note for example that they do not exhibit the wavy pattern of $\frac{df}{ds}(x,t)$, instead they are relatively smooth but with some local peaks and valleys.

Using the r largest singular values of $\mathcal{H}$, an approximate solution of system (24) can be written as

$$\widehat{\mathcal{V}}_{(r)}=\sum_{i=1}^{r}\sigma_i\langle\widehat{\mathcal{R}}\boldsymbol{\Phi}_i\rangle\boldsymbol{\Psi}_i, \tag{38}$$

where $\langle\widehat{\mathcal{R}}\boldsymbol{\Phi}_i\rangle=\boldsymbol{\Phi}_i^*\widehat{\mathcal{R}}$ is the projection of the right hand side $\widehat{\mathcal{R}}$ onto the optimal forcing $\boldsymbol{\Phi}_i$. For $r=(2q+1)(N_u+1)$ the approximation is exact. Using $\widehat{\mathcal{V}}_{(r)}$, approximations $\left(\frac{d\overline{J_1}}{dc}\right)_{(r)}$ and $\left(\frac{d\overline{J_2}}{dc}\right)_{(r)}$ can be computed from (34) and (35) respectively;

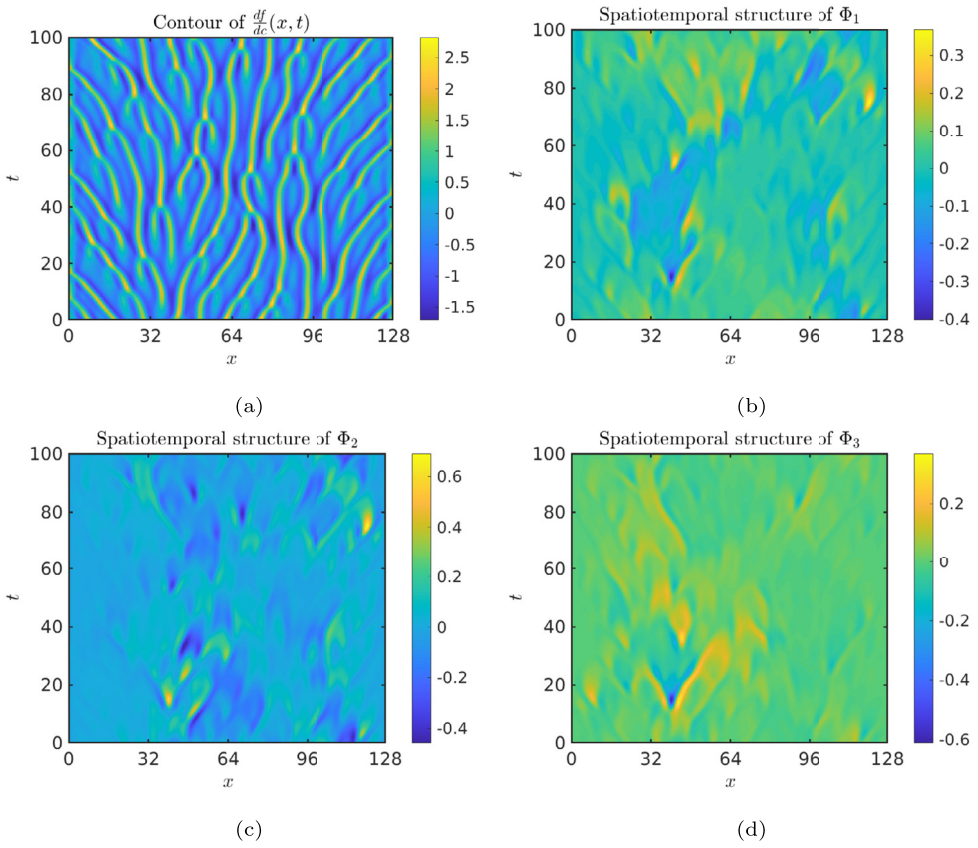


Fig. 8. Contour plots of (a) the forcing $\frac{df}{dc}(x,t)$, the spatio-temporal structure of optimal forcings (b) $\mathbf{\Phi}_1$, (c) $\mathbf{\Phi}_2$ and (d) $\mathbf{\Phi}_3$ for $T=100$, $c=0$ and $f\in[-0.3,0.3]$.

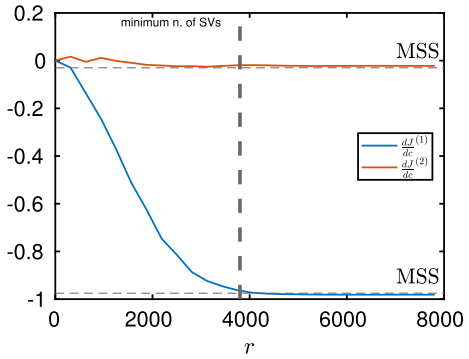


Fig. 9. Approximations $\left(\frac{d\overline{J_1}}{dc}\right)_{(r)}$ and $\left(\frac{d\overline{J_2}}{dc}\right)_{(r)}$ as a function of the number of retained singular values, r, for $c=0$, $q=30$ and $T=100$.

these are plotted as functions of r in Fig. 9. It can be seen that a relatively large value of r is required to obtain an accurate result. Although the first few singular values σ_i are large, the component of $\widehat{\mathcal{R}}$ along the $\mathbf{\Phi}_i$ direction is weak, thus a substantial number of terms are required to obtain the correct sensitivity. Bearing in mind that $\widehat{\mathcal{R}}_k=\left[\frac{d\hat{\mathbf{f}}_k}{d\mathbf{s}},0\right]^\top$, this is related to the different spatio-temporal patterns between $\frac{df}{ds}(x,t)$ and the optimal forcings (at least for the first 3 modes) as shown in the previous Fig. 8.

The linear system (24) has large storage requirements and is time consuming to solve. Below we propose an approach that can mitigate these requirements.

5. A resolvent-based iterative method for the shadowing direction

Instead of solving directly system (24), an iterative method can be devised, where only a few diagonals are retained in the left hand side and treated implicitly, and the rest are moved to the right hand side and updated at every iteration. For example, by retaining only the blocks of the main diagonal, the Hill matrix becomes diagonal and the blocks decouple. In this case, the iterative method takes the form

$$\begin{bmatrix} ik\omega_0\mathcal{I}_{N_u} - \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} & -\overline{\boldsymbol{f}} \\ -\overline{\boldsymbol{f}}^{\top} & 0 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{v}}_k \\ \hat{\eta}_k \end{bmatrix}^{(m)} = \begin{bmatrix} \frac{d\hat{\boldsymbol{f}}_k}{d\boldsymbol{s}} \\ 0 \end{bmatrix} + \begin{bmatrix} \hat{\boldsymbol{g}}_k \\ \hat{\boldsymbol{h}}_k \end{bmatrix}^{(m-1)}, \tag{39}$$

where m is the iteration number, and $\hat{\boldsymbol{g}}_k$, $\hat{\boldsymbol{h}}_k$ denote the explicitly treated terms. This is block-Jacobi type of iterative algorithm.

It is instructive to derive the form of vectors $\boldsymbol{g}(t)$ and $\boldsymbol{h}(t)$ in the time domain. To this end, using Reynolds decomposition,

$$\begin{aligned} &\boldsymbol{v} = \overline{\boldsymbol{v}} + \boldsymbol{v}', \quad \eta = \overline{\eta} + \eta', \quad \boldsymbol{f} = \overline{\boldsymbol{f}} + \boldsymbol{f}', \\ &\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} = \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}, \quad \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}} = \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}} + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}} \end{aligned} \tag{40}$$

and substituting in (9b) and (9c) we get

$$\frac{d\boldsymbol{v}'}{dt} = \left(\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\right)(\overline{\boldsymbol{v}} + \boldsymbol{v}') + \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}} + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}} + (\overline{\eta} + \eta')\left(\overline{\boldsymbol{f}} + \boldsymbol{f}'\right) \tag{41a}$$

$$\langle\left(\overline{\boldsymbol{f}} + \boldsymbol{f}'\right)(\overline{\boldsymbol{v}} + \boldsymbol{v}')\rangle = \left(\overline{\boldsymbol{f}}^{\top} + \boldsymbol{f}'^{\top}\right)(\overline{\boldsymbol{v}} + \boldsymbol{v}') = 0 \tag{41b}$$

Taking the time-average we obtain

$$\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\overline{\boldsymbol{v}} + \overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'} + \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}} + \overline{\eta}\,\overline{\boldsymbol{f}} + \overline{\eta'\boldsymbol{f}'} = 0, \tag{42a}$$

$$\overline{\boldsymbol{f}}^{\top}\overline{\boldsymbol{v}} + \overline{\boldsymbol{f}'^{\top}\boldsymbol{v}'} = 0 \tag{42b}$$

and subtracting the two sets, we get

$$\frac{d\boldsymbol{v}'}{dt} = \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\boldsymbol{v}' + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\overline{\boldsymbol{v}} + \left(\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}' - \overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'}\right) + \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}} + \overline{\eta}\,\boldsymbol{f}' + \eta'\overline{\boldsymbol{f}} + \left(\eta'\boldsymbol{f}' - \overline{\eta'\boldsymbol{f}'}\right) \tag{43a}$$

$$\overline{\boldsymbol{f}}^{\top}\boldsymbol{v}' + \boldsymbol{f}'^{\top}\overline{\boldsymbol{v}} + \left(\boldsymbol{f}'^{\top}\boldsymbol{v}' - \overline{\boldsymbol{f}'^{\top}\boldsymbol{v}'}\right) = 0 \tag{43b}$$

After some rearrangement,

$$\frac{d\boldsymbol{v}'}{dt} - \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\boldsymbol{v}' - \eta'\overline{\boldsymbol{f}} = \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}} + \underbrace{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\overline{\boldsymbol{v}} + \left(\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}' - \overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'}\right) + \overline{\eta}\,\boldsymbol{f}' + \left(\eta'\boldsymbol{f}' - \overline{\eta'\boldsymbol{f}'}\right)}_{=\boldsymbol{g}'(t)} \tag{44a}$$

$$\overline{\boldsymbol{f}}^{\top}\boldsymbol{v}' = \underbrace{-\boldsymbol{f}'^{\top}\overline{\boldsymbol{v}} - \left(\boldsymbol{f}'^{\top}\boldsymbol{v}' - \overline{\boldsymbol{f}'^{\top}\boldsymbol{v}'}\right)}_{=\boldsymbol{h}'(t)} \tag{44b}$$

Thus we get

$$\frac{d\boldsymbol{v}'}{dt} - \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\boldsymbol{v}' - \eta'\overline{\boldsymbol{f}} = \frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}} + \boldsymbol{g}'(t) \tag{45a}$$

$$\overline{\boldsymbol{f}}^{\top}\boldsymbol{v}' = \boldsymbol{h}'(t) \tag{45b}$$

and taking the Fourier transform leads to (39) for $k \neq 0$. Similarly, the time-average system (42) can be written as

$$-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\overline{\boldsymbol{v}}-\overline{\eta}\,\overline{\boldsymbol{f}}=\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}}+\underbrace{\overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'+\eta'\boldsymbol{f}'}}_{=\overline{\boldsymbol{g}(t)}} \tag{46a}$$

$$\overline{\boldsymbol{f}}^{\top}\overline{\boldsymbol{v}}=\underbrace{-\overline{\boldsymbol{f}'^{\top}\boldsymbol{v}'}}_{=\overline{\boldsymbol{h}(t)}}, \tag{46b}$$

from which we obtain the form

$$-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\overline{\boldsymbol{v}}-\overline{\eta}\,\overline{\boldsymbol{f}}=\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}}+\overline{\boldsymbol{g}(t)}, \tag{47a}$$

$$\overline{\boldsymbol{f}}^{\top}\overline{\boldsymbol{v}}=\overline{\boldsymbol{h}(t)}, \tag{47b}$$

that corresponds to $k=0$ for system (39).

As can be seen from (44a), $\boldsymbol{g}'(t)$ consists of two groups of three terms; the first group involves the fluctuating Jacobian and sensitivity $\boldsymbol{v}'$, and the second the fluctuating η' and $\boldsymbol{f}'$. It is possible to get a simplified system by assuming that η is constant (for the physical interpretation see [44]). In this case we get

$$\frac{d\boldsymbol{v}'}{dt}-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\boldsymbol{v}'=\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{s}}+\eta\boldsymbol{f}'+\underbrace{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\overline{\boldsymbol{v}}+\left(\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'-\overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'}\right)}_{=\tilde{\boldsymbol{g}}'(t)} \tag{48a}$$

$$-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\overline{\boldsymbol{v}}-\eta\,\overline{\boldsymbol{f}}=\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}}+\underbrace{\overline{\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}}\boldsymbol{v}'}}_{=\overline{\tilde{\boldsymbol{g}}(t)}} \tag{48b}$$

This system requires an additional constraint to obtain the constant η. To this end, we require that $\boldsymbol{f}(t)$ and $\boldsymbol{v}(t)$ are perpendicular in a time-average sense, i.e.

$$\frac{1}{T}\int_0^T \boldsymbol{f}^{\top}(t)\boldsymbol{v}(t)dt=\sum_{k=-q}^{q}\hat{\boldsymbol{f}}_{-k}^{\top}\hat{\boldsymbol{v}}_k=\sum_{k=-q}^{q}\hat{\boldsymbol{f}}_k^{*}\hat{\boldsymbol{v}}_k=0, \tag{49}$$

because $\boldsymbol{f}(t)$ is a real variable. Note that this condition couples together all Fourier components. Taking the Fourier transform of (48) we can form the following iterative method

$$\left(ik\omega_0\mathcal{I}_{N_u}-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\right)\hat{\boldsymbol{v}}_k^{(m)}=\frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}}+\eta^{(m)}\hat{\boldsymbol{f}}_k+\tilde{\boldsymbol{g}}_k^{(m-1)}\quad[k=-q\dots q] \tag{50a}$$

$$\sum_{k=-q}^{q}\hat{\boldsymbol{f}}_k^{*}\hat{\boldsymbol{v}}_k^{(m)}=0 \tag{50b}$$

In the first iteration $m=1$, we set $\tilde{\boldsymbol{g}}_k^{(0)}=0$ and we get

$$\left(ik\omega_0\mathcal{I}_{N_u}-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\right)\hat{\boldsymbol{v}}_k^{(1)}=\frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}}+\eta^{(1)}\hat{\boldsymbol{f}}_k\quad[k=-q\dots q] \tag{51a}$$

$$\sum_{k=-q}^{q}\hat{\boldsymbol{f}}_k^{*}\hat{\boldsymbol{v}}_k^{(1)}=0 \tag{51b}$$

These two equations can be combined together to obtain $\eta^{(1)}$. Denoting the standard resolvent operator as

$$\mathcal{R}(k\omega_0)=\left(ik\omega_0\mathcal{I}_{N_u}-\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}\right)^{-1}, \tag{52}$$

solving for $\hat{\boldsymbol{v}}_k^{(1)}$ and substituting in (51b) we get

$$\sum_{k=-q}^{q}\hat{\boldsymbol{f}}_k^{*}\left[\mathcal{R}(k\omega_0)\left(\frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}}+\eta^{(1)}\hat{\boldsymbol{f}}_k\right)\right]=0 \tag{53}$$

from which we obtain

$$\eta^{(1)} = -\frac{\sum_{k=-q}^{q} \hat{\boldsymbol{f}}_k^* \mathcal{R}(k\omega_0) \frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}}}{\sum_{k=-q}^{q} \hat{\boldsymbol{f}}_k^* \mathcal{R}(k\omega_0) \hat{\boldsymbol{f}}_k} \tag{54}$$

or

$$\eta^{(1)} = -\frac{\sum_{k=-q}^{q} \hat{\boldsymbol{f}}_k^* \boldsymbol{\lambda}_k}{\sum_{k=-q}^{q} \hat{\boldsymbol{f}}_k^* \boldsymbol{\mu}_k} \tag{55}$$

where

$$\mathcal{R}(k\omega_0) \frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}} = \boldsymbol{\lambda}_k \Rightarrow \left(ik\omega_0 \mathcal{I}_{N_u} - \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} \right) \boldsymbol{\lambda}_k = \frac{\partial \hat{\boldsymbol{f}}_k}{\partial \boldsymbol{s}} \tag{56a}$$

$$\mathcal{R}(k\omega_0) \hat{\boldsymbol{f}}_k = \boldsymbol{\mu}_k \Rightarrow \left(ik\omega_0 \mathcal{I}_{N_u} - \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}} \right) \boldsymbol{\mu}_k = \hat{\boldsymbol{f}}_k \tag{56b}$$

Thus the solution of two linear systems that involve the standard resolvent operator, $\mathcal{R}(k\omega_0)$, is required. Due to the linearity of (50a),

$$\hat{\boldsymbol{v}}_k^{(1)} = \boldsymbol{\lambda}_k + \eta^{(1)} \boldsymbol{\mu}_k. \tag{57}$$

Applying inverse Fourier transform to $\hat{\boldsymbol{v}}_k^{(1)}$ yields $v^{(1)}(t)$, from which $\tilde{\boldsymbol{g}}(t)$ can be obtained from (48), and Fourier transformed to find $\tilde{\boldsymbol{g}}_k^{(1)}$. The right hand side of (50a) can then be assembled and the second iteration performed. Note that system (56b) does not change with m, thus $\boldsymbol{\mu}_k$ is computed once. In this process, the key variables $\boldsymbol{\lambda}_k$, η, and $\boldsymbol{\mu}_k$ required for the evaluation of the sensitivity $\hat{\boldsymbol{v}}_k$ (also known as shadowing direction) are obtained with the aid of the Resolvent operator $\mathcal{R}(k\omega_0)$. This is therefore a resolvent-based iterative method for computing the shadowing direction, called Resolvent-based Shadowing (RbS).

The simplest and most efficient approach is to perform LU decomposition of $\mathcal{R}(k\omega_0)$ once at the start of the sensitivity analysis and solve systems (56) with forward and backward substitution for each k. Note that the decomposition (57) of $\hat{\boldsymbol{v}}_k$ into a linear combination of $\boldsymbol{\lambda}_k$ and $\boldsymbol{\mu}_k$ is valid for all iterations, and of course the final converged solution.

In (48a), the term $\frac{\partial \boldsymbol{f}'}{\partial \boldsymbol{u}} \overline{\boldsymbol{v}}$ was considered as part of $\tilde{\boldsymbol{g}}'(t)$ and treated explicitly. This is not necessary, but it simplifies the algebra. A better approach would be to obtain $\overline{\boldsymbol{v}}$ from (48b) and substitute in (48a); this would lead to a form very similar to (51a), albeit more complex. The process to extract η remains the same. Such an approach would couple better the time-average and the fluctuating components of the solution with an expected beneficial effect on the convergence rate.

The storage requirements of the proposed iterative method are much reduced compared with the harmonic balancing approach presented in section 3. Only the time-average Jacobian $\overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}}$ needs to be stored in sparse matrix form in order to assemble the standard resolvent operator $\mathcal{R}(k\omega_0)$. The computational bottleneck is the application of LU decomposition to $\mathcal{R}(k\omega_0)$, which needs however to be performed only once. If the method is applied to fluid flow problems that possess at least one direction of homogeneity (for example boundary layers, jets, pipe flows, flows around bluff bodies etc), this decomposition can be performed very efficiently using existing linear algebra packages, such as MUMPS [47], that exploit the sparse structure of the Jacobian. For three-dimensional inhomogeneous flows this is more challenging, but doable (at least for moderate-size systems).

The proposed method was applied to compute the sensitivities of the objective functions $\overline{J_1}(c)$ and $\overline{J_2}(c)$, defined in (32) and (33), for the Kuramoto-Sivashinsky equation. Since η is constant, equations (34) and (35) are simplified to

$$\frac{d\overline{J_1}}{dc} = \frac{1}{L} \int_0^L \hat{\boldsymbol{v}}_0 \, dx, \tag{58}$$

and

$$\frac{d\overline{J_2}}{dc} = \frac{2}{L} \int_0^L \sum_{k=-q}^{q} \hat{\boldsymbol{u}}_{-k} \hat{\boldsymbol{v}}_k \, dx = \sum_{k=-q}^{q} \frac{2}{L} \int_0^L \hat{\boldsymbol{u}}_k^* \hat{\boldsymbol{v}}_k \, dx, \tag{59}$$

respectively. A comparison of the sensitivities produced by the harmonic balance approach of section 3, finite differences, and the methodology presented in this section (applied with a single iteration) is shown in Fig. 10. The sensitivities match well with the those obtained with shadowing harmonic for both objective functions. The explanation for this (perhaps surprisingly) good result for $\frac{d\overline{J_1}}{dc}$ is provided in Fig. 4, where we compare the distribution of $\hat{\boldsymbol{v}}_0(x) = \overline{\boldsymbol{v}}(x)$ between the two

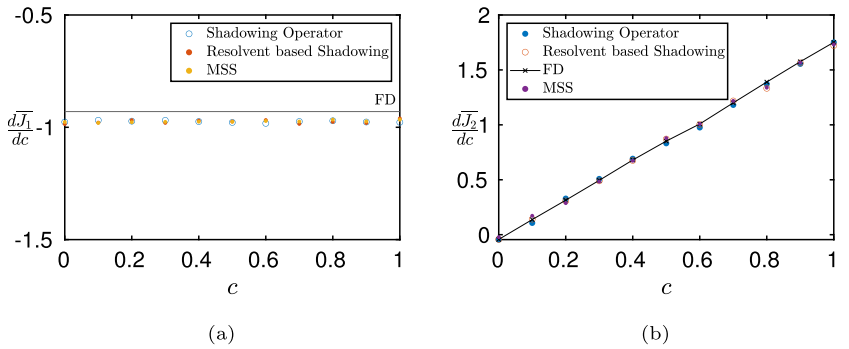


Fig. 10. Comparison of sensitivities obtained using full harmonic balancing, finite differences, MSS and Resolvent based Shadowing (RbS) for (a) $\frac{d\overline{J_1}}{dc}$, (b) $\frac{d\overline{J_2}}{dc}$. Results are averaged over 100 random initial conditions and the time-average Jacobian was obtained using $T = 500$.

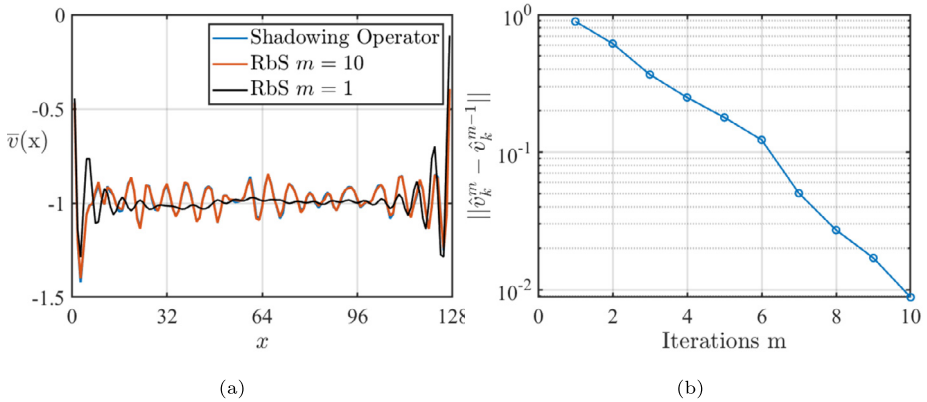


Fig. 11. (a) Comparison of $\overline{v}(x) = \hat{v}_0(x)$, as computed by the shadowing harmonic approach for $\eta = const$ and by Resolvent-based Shadowing (RbS) for $m = 1$ and $m = 10$ iterations. (b) Convergence of the $\mathcal{L}_2$ norm of the difference between the Fourier coefficients against the number of iterations, m.

approaches. As can be seen, after a single iteration, the average value over x is approximated well, but the details of the distribution in the middle of the domain are not captured. However, these differences cancel out when integrating $\hat{\boldsymbol{v}}_0(x)$, see (58), leading to an accurate sensitivity value. If $\overline{J_1}(c)$ had been formulated as the average of $\overline{u}(x;c)$ over a smaller part of the domain, then the sensitivities with a single iteration would not have been accurate. Of course, in this case additional iterations can be performed.

To prove this is indeed the case, we carried out more iterations and compared with the result of the shadowing harmonic operator with $\eta(t)$ constant; this required a modification of the system (24) because only $\hat{\eta}_0$ exists. The solution $\overline{v}(x) = \hat{v}_0(x)$ obtained with $m = 1$ and $m = 10$ iterations is shown in Fig. 11a below. As can be seen, the iterative solution with 10 iterations and the harmonic shadowing operator result are almost indistinguishable. Note that no averaging over realisations is performed here, only one set of initial conditions was used. In Fig. 11b, the convergence of the norm of the difference of the Fourier coefficients $\hat{\boldsymbol{v}}_k$ between two successive iterations is shown; it decreases by 2 orders of magnitude within $m = 10$ iterations. As a final confirmation, we computed the sensitivity $\frac{d\overline{J_2}}{dc}$ in the subdomain $x \in [30, 40]$ and plot the results with $m = 1$ and $m = 10$ iterations in Fig. 12 below for different values of c. As can be seen, the sensitivity of $\overline{J_2}$ computed in part of the domain is accurate, provided that additional iterations are performed.

In terms of computational cost, a single iteration took approximately 0.15 s of CPU time. This is about one order of magnitude faster with respect to the shadowing harmonic approach. Again, we stress that these results are case dependent. More research is need to investigate the performance of the algorithm in other flow cases. In the next section, we explore the properties of the standard resolvent operator $\mathcal{R}(k\omega_0)$ and explore the effect of frequency truncation.

6. Sensitivities using of the standard resolvent operator and effect of frequency truncation

The five largest singular values of the resolvent operator $\mathcal{R}(\omega)$ are plotted in Fig. 13 as a function of ω. The maximum singular value has a peak at $\omega|_{\sigma=\sigma_{max}} \approx 0.164$ and in this frequency region it is more than 3 orders of magnitude larger than the second. This result is compatible with the spectra of Fig. 2; large gains appear in the frequency range with strong spectral content. For angular frequencies $\omega > 2\pi\, 0.3 \approx 1.9$ there is no coupling with the time-average Jacobian, leading to $\sigma(\omega) \approx 1/\omega$ (indicated by a black dashed line); again this is consistent with the frequency spectra.

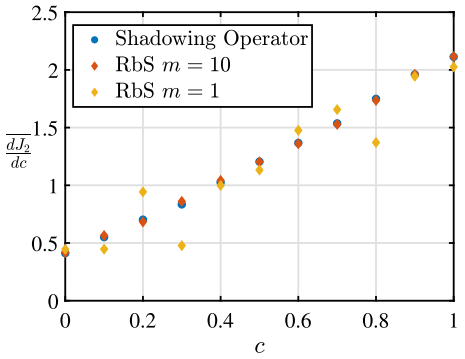


Fig. 12. Sensitivity $\frac{d\overline{J_2}}{dc}$ against c, with $\overline{J_2}$ evaluated in the interval $x \in [30, 40]$. A comparison is made between results obtained with the Shadowing Harmonic Operator and RbS with $m = 1$ and $m = 10$ iterations.

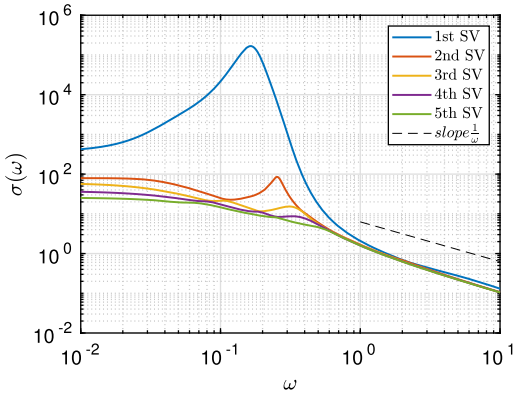


Fig. 13. Variation of the five largest singular values against ω, for $c = 0$, $T = 100$.

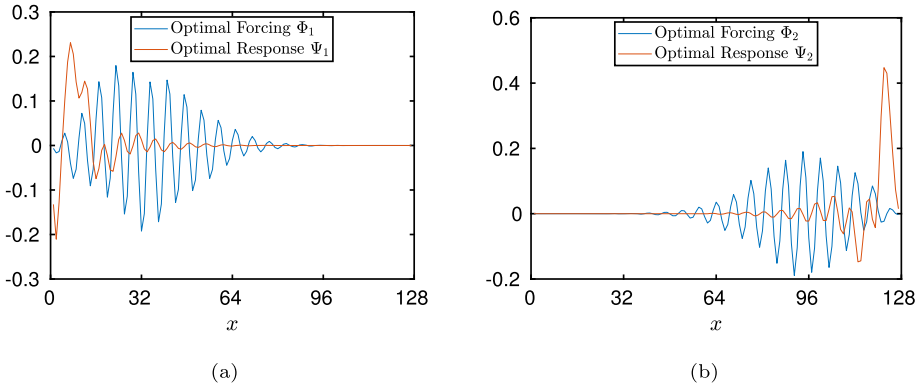
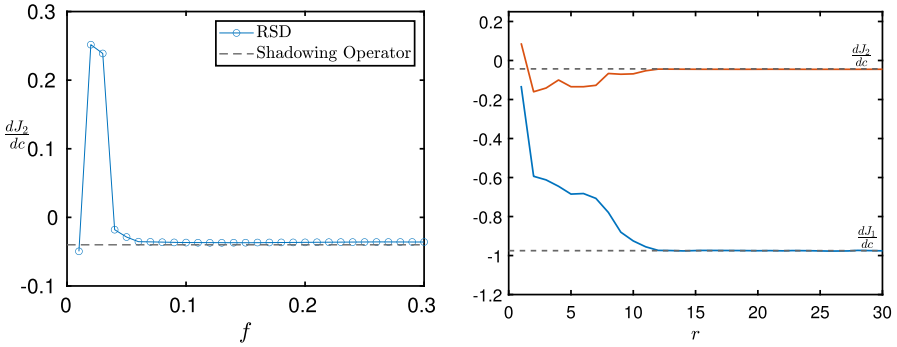


Fig. 14. Optimal forcing and response corresponding to the largest (left) and second largest (right) singular values for $\omega = 0.164$, $T = 20000$ and $c = 0$.

In Fig. 14 the optimal forcing and response that correspond to the two largest singular values for $\omega = 0.164$ are plotted. These plots are easier to interpret compared to the spatio-temporal maps shown in Figs. 7 and 8. For example, note that both distributions are spatially localised, on the left half of the domain for the largest singular value, and on the right half for the second largest. In both cases, the optimal forcing is located upstream of the optimal response, as expected due to the convective nature of the KS (recall that $\overline{u}(x)$, shown in Fig. 1, changes sign in the middle of the domain).

Substituting expression (57) into (58) we get

$$\frac{d\overline{J_1}}{dc} = \frac{1}{L}\int_0^L \left(\boldsymbol{\lambda}_0 + \eta\boldsymbol{\mu}_0\right)dx. \tag{60}$$



(a) (b)

Fig. 15. (a) Convergence of $\frac{d\overline{J_2}}{dc}$ computed using frequencies in the interval $[-f, f]$ as f increases, for $c = 0$ and $T = 100$. (b) Effect of the number of retained singular values, r, on the convergence of sensitivities (32) and eq. (33) for $T = 100$, $q = 30$ and $c = 0$.

The solutions λ_0 and μ_0 of the linear systems (56a) and (56b) can be written in terms of the optimal forcings and responses evaluated for $\omega = 0$ as

$$\lambda_{0(r)} = \sum_{i=1}^{r} \sigma_i^{(\omega=0)} \left\langle \overline{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{s}}} \Phi_i^{(\omega=0)} \right\rangle \Psi_i^{(\omega=0)}, \quad \mu_{0(r)} = \sum_{i=1}^{r} \sigma_i^{(\omega=0)} \left\langle \overline{\boldsymbol{f}} \Phi_i^{(\omega=0)} \right\rangle \Psi_i^{(\omega=0)}, \tag{61}$$

where r is the number of retained singular values. This expression is analogous to equation (38) presented earlier for the harmonic balance method. Substituting in (60) we get

$$\left(\frac{d\overline{J_1}}{dc}\right)_{(r)} = \sum_{i=1}^{r} \sigma_i^{(\omega=0)} \left(\left\langle \overline{\frac{\partial \boldsymbol{f}}{\partial c}} \Phi_i^{(\omega=0)} \right\rangle + \eta \left\langle \overline{\boldsymbol{f}} \Phi_i^{(\omega=0)} \right\rangle \right) \frac{1}{L} \int_0^L \Psi_i dx. \tag{62}$$

The sensitivity can therefore be written as the weighted sum of the spatial averages of the optimal responses, Ψ_i.

Similarly for $\frac{d\overline{J_2}}{dc}$, we get from (59),

$$\left(\frac{d\overline{J_2}}{dc}\right)_{(r)} = \sum_{k=-q}^{q} \frac{2}{L} \int_0^L \boldsymbol{u}_k^* (\lambda_k + \eta \mu_k) \, dx, \tag{63}$$

where

$$\lambda_{k(r)} = \sum_{i=1}^{r} \sigma_i^{(\omega=k\omega_0)} \left\langle \frac{\partial \hat{\boldsymbol{f}}_k}{\partial c} \Phi_i^{(\omega=k\omega_0)} \right\rangle \Psi_i^{(\omega=k\omega_0)}, \tag{64a}$$

$$\mu_{0(r)} = \sum_{i=1}^{r} \sigma_i^{(\omega=k\omega_0)} \left\langle \hat{\boldsymbol{f}}_k \Phi_i^{(\omega=k\omega_0)} \right\rangle \Psi_i^{(\omega=k\omega_0)}. \tag{64b}$$

Expression (63) is useful because it allows us to find the contribution of each frequency on the sensitivity; we investigate this in Fig. 15a. More specifically, we compute $\frac{d\overline{J_2}}{dc}$ using frequencies in the interval $[-f, f]$, where $f = \omega/(2\pi)$, and we plot the result against f. Note the convergence of $\frac{d\overline{J_2}}{dc}$ to the value predicted by the harmonic resolvent as f increases, i.e. the range of k in (63) expands. It can be seen that only the frequency range $[0.01, 0.07]$ contributes, which is consistent with the spectra shown in Fig. 2. Outside this range, the spectral content is small, and does not contribute to the sensitivity. Note that this is the result of a single realisation; no averaging over initial conditions has been performed to obtain this plot. This example justifies the premise made in the introduction that accurate results can be obtained if only frequencies with the high energy footprint are retained. In this section, we applied the idea to the iterative method, but equally well this can be applied to the full system (24), where only the relevant diagonals need to be retained.

In Fig. 15b we plot the convergence of the two sensitivities against the number of retained singular values r (same r for every k) in the summations (62) and (64). Approximately 10-12 singular values are required for convergence. Hence, although the maximum singular value is significantly larger compared with the rest as evidenced in 13, keeping just one contribution will not provide accurate results. In order to explain this behaviour for $\frac{d\overline{J_1}}{dc}$, the optimal forcing corresponding to σ_{max} at $\omega = 0$ is plotted together with the true forcing $\overline{\frac{\partial \boldsymbol{f}}{\partial c}}$ in Fig. 16. The latter has strong footprint in the boundaries of

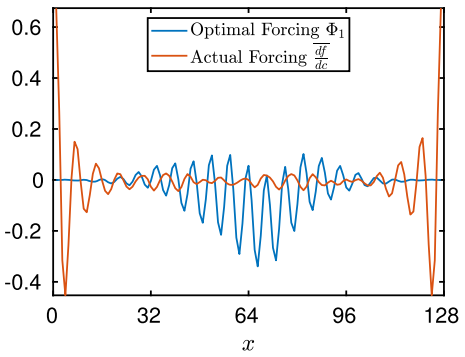


Fig. 16. Optimal forcing Φ_1 for $\omega = 0$ and time-average forcing $\overline{\frac{\partial f}{\partial c}}$ against x, for $T = 20000$, $c = 0$.

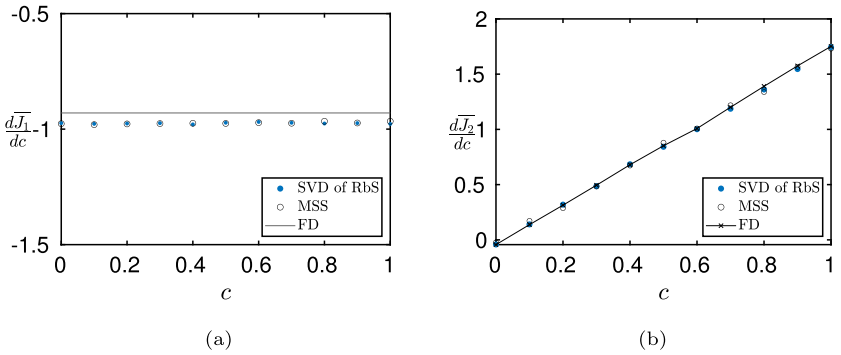


Fig. 17. Comparison of the sensitivities computed with SVD of RbS, FD and MSS for (a) eq. (32) and (b) eq. (33), for $T = 100$, $q = 7$ and $r = 12$. Results are averaged over 100 random initial conditions.

the domain, while the former in the centre, thus the projection $\left\langle \overline{\frac{\partial \boldsymbol{f}}{\partial c}} \Phi_1^{(\omega=0)} \right\rangle$ appearing in expression (61) for $\lambda_{0(r)}$ is weak and mitigates the effect of the large σ_{max}.

In Fig. 17 we compare the sensitivities computed with MSS, RbS and finite differences for different values of c, using $T = 100$, $q = 7$ and $r = 12$. This combination of T and q captures the frequency band where fluctuations have large spectral footprint, while at each frequency a number of singular values are retained based on the evidence from Fig. 15b. Using these settings, the matching with the reference MSS values is very good.

7. Conclusions

Least-squares shadowing (LSS) is a very promising method for the computation of sensitivities of time-average quantities of chaotic dynamical systems to system parameters. In this paper, we reformulate LSS in the frequency domain using the harmonic balancing approach. This formulation leads naturally to the definition of the shadowing harmonic operator. The proposed method allows truncation of the frequencies that do not carry significant spectral content without any modelling assumptions. We apply this operator and study its properties for the Kuramoto-Sivashinsky equation.

The sensitivities computed with the time- and frequency-domain formulations match. We propose also an iterative approach to solve the resulting system of equations, where only the diagonal blocks of the shadowing harmonic operator need to be stored and inverted. These blocks correspond to the standard resolvent operator, thus this is a resolvent-based iterative method to compute the shadowing direction. We show that this approach provides very accurate results even with a single iteration. Furthermore we show that this is dependent on the particular objective function considered and that keeping only the dominant frequencies did not affect the accuracy of the results.

Clearly more work is needed to assess the accuracy, cost and scalability of the proposed method. Many questions arise for example, how does the convergence rate depend on the system parameters, for example Reynolds number? How the quantity of interest affects the frequency range that must be retained and the number of iterations required? For instance, dissipation in turbulent flows is determined by the smallest scale structures; does this mean that evaluating the sensitivity of dissipation will require a wider frequency range compared to the sensitivity of the forces acting on a body surface (that are mostly determined by large scale structures)? Can the sensitivity direction $\boldsymbol{v}(t)$ be approximated accurately by a reduced order model? Work towards answering these questions will form part of future research.

CRediT authorship contribution statement

Kyriakos D. Kantarakias: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **George Papadakis:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The first author wishes to acknowledge the financial support of the President's Scholarship Award from Imperial College London. The second author is grateful for the financial support provided by EPSRC, grant No. EP/P020194/1.

References

[1] Q. Wang, R. Hu, P. Blonigan, Least squares shadowing sensitivity analysis of chaotic limit cycle oscillations, J. Comput. Phys. 267 (2014) 210–224.
[2] D.J. Lea, M.R. Allen, T.W. Haine, Sensitivity analysis of the climate of a chaotic system, Tellus, Ser. A Dyn. Meteorol. Oceanogr. 52 (5) (2000) 523–532.
[3] T. Bewley, P. Moin, R. Temam, DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms, J. Fluid Mech. 447 (2001) 179–225, https://doi.org/10.1017/S0022112001005821.
[4] D. Xiao, G. Papadakis, Nonlinear optimal control of transition due to a pair of vortical perturbations using a receding horizon approach, J. Fluid Mech. 861 (2019) 524–555.
[5] J. Larsson, Grid-adaptation for chaotic multi-scale simulations as a verification-driven inverse problem, in: AIAA Aerospace Sciences Meeting, 2018, pp. 1–17.
[6] G.L. Eyink, T.W.N. Haine, D.J. Lea, Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights, Nonlinearity 17 (5) (2004) 1867–1889.
[7] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. 29 (1) (1966) 255.
[8] J. Thuburn, Climate sensitivities via a Fokker–Planck adjoint approach, Q. J. R. Meteorol. Soc. 131 (605) (2005) 73–92, https://doi.org/10.1256/qj.04.46.
[9] J. Craske, Adjoint sensitivity analysis of chaotic systems using cumulant truncation, Chaos Solitons Fractals 119 (2019) 243–254.
[10] D. Lasagna, Sensitivity analysis of chaotic systems using unstable periodic orbits, SIAM J. Appl. Dyn. Syst. 17 (1) (2018) 547–580.
[11] N. Chandramoorthy, Q. Wang, Efficient computation of linear response of chaotic attractors with one-dimensional unstable manifolds, SIAM J. Appl. Dyn. Syst. 21 (2) (2022) 735–781.
[12] D. Ruelle, Differentiation of SRB states, Commun. Math. Phys. 187 (1) (1997) 227–241, https://doi.org/10.1007/s002200050134.
[13] Q. Wang, Convergence of the least squares shadowing method for computing derivative of ergodic averages, SIAM J. Numer. Anal. 52 (1) (2014) 156–170, https://doi.org/10.1137/130917065.
[14] P.J. Blonigan, Q. Wang, E.J. Nielsen, B. Diskin, Least-squares shadowing sensitivity analysis of chaotic flow around a two-dimensional airfoil, AIAA J. 56 (2) (2018) 658–672.
[15] S.Y. Pilyugin, Shadowing in Dynamical Systems, Lecture Notes in Mathematics, Springer-Verlag, 1999.
[16] R. Bowen, ω-Limit sets for Axiom A diffeomorphisms, J. Differ. Equ. 18 (2) (1975) 333–339.
[17] S.M. Hammel, J.A. Yorke, C. Grebogi, Do numerical orbits of chaotic dynamical processes represent true orbits?, J. Complex. 3 (2) (1987) 136–145.
[18] T. Sauer, J.A. Yorke, Rigorous verification of trajectories for the computer simulation of dynamical systems, Nonlinearity 4 (3) (1991) 961–979.
[19] T. Sauer, C. Grebogi, J.A. Yorke, How long do numerical chaotic solutions remain valid?, Phys. Rev. Lett. 79 (1997) 59–62.
[20] N. Chandramoorthy, Q. Wang, On the probability of finding nonphysical solutions through shadowing, J. Comput. Phys. 440 (2021) 110389.
[21] A. Ni, Q. Wang, Sensitivity analysis on chaotic dynamical systems by Non-Intrusive Least Squares Shadowing (NILSS), J. Comput. Phys. 347 (2017) 56–77.
[22] A. Ni, C. Talnikar, Adjoint sensitivity analysis on chaotic dynamical systems by Non-Intrusive Least Squares Adjoint Shadowing (NILSAS), J. Comput. Phys. 395 (2019) 690–709.
[23] A. Ni, Hyperbolicity, shadowing directions and sensitivity analysis of a turbulent three-dimensional flow, J. Fluid Mech. 863 (2019) 644–669.
[24] P.J. Blonigan, Q. Wang, Multiple shooting shadowing for sensitivity analysis of chaotic dynamical systems, J. Comput. Phys. 354 (2018) 447–475.
[25] K. Shawki, G. Papadakis, A preconditioned multiple shooting shadowing algorithm for the sensitivity analysis of chaotic systems, J. Comput. Phys. 398 (2019) 108861, https://doi.org/10.1016/j.jcp.2019.108861.
[26] K. Kantarakias, K. Shawki, G. Papadakis, Uncertainty quantification of sensitivities of time-average quantities in chaotic systems, Phys. Rev. E 101 (2020) 022223.
[27] L. Keefe, P. Moin, J. Kim, The dimension of attractors underlying periodic turbulent Poiseuille flow, J. Fluid Mech. 242 (1992) 1–29.
[28] J.A. Vastano, R.D. Moser, Short-time Lyapunov exponent analysis and the transition to chaos in Taylor–Couette flow, J. Fluid Mech. 233 (1991) 83–118.
[29] M. Hassanaly, V. Raman, Lyapunov spectrum of forced homogeneous isotropic turbulent flows, Phys. Rev. Fluids 4 (2019) 114608.
[30] A. Crisanti, M.H. Jensen, A. Vulpiani, G. Paladin, Intermittency and predictability in turbulence, Phys. Rev. Lett. 70 (1993) 166–169.
[31] P. Mohan, N. Fitzsimmons, R.D. Moser, Scaling of Lyapunov exponents in homogeneous isotropic turbulence, Phys. Rev. Fluids 2 (2017) 114606.
[32] M. Xu, M.R. Paul, Covariant Lyapunov vectors of chaotic Rayleigh-Bénard convection, Phys. Rev. E 93 (2016) 062208.
[33] S.B. Pope, Turbulent Flows, Cambridge University Press, 2000.
[34] C. Pruett, Temporal large-eddy simulation: theory and implementation, Theor. Comput. Fluid Dyn. 22 (2008) 275–304.
[35] A.A. Sliwiak, Q. Wang, Approximating linear response of physical chaos, arXiv:2205.03734, 2022.
[36] B.J. McKeon, A.S. Sharma, A critical-layer framework for turbulent pipe flow, J. Fluid Mech. 658 (2010) 336–382.
[37] G. Rigas, D. Sipp, T. Colonius, Nonlinear input/output analysis: application to boundary layer transition, J. Fluid Mech. 911 (2021) A15.
[38] A. Padovan, S.E. Otto, C.W. Rowley, Analysis of amplification mechanisms and cross-frequency interactions in nonlinear flows via the harmonic resolvent, J. Fluid Mech. 900 (2020) A14.

[39] R. Moarref, M.R. Jovanovic, Model-based design of transverse wall oscillations for turbulent drag reduction, J. Fluid Mech. 707 (2012) 205–240.
[40] Q. Wang, Forward and adjoint sensitivity computation of chaotic dynamical systems, J. Comput. Phys. 235 (2013) 1–13, https://doi.org/10.1016/j.jcp.2012.09.007.
[41] S.Y. Pilyugin, Shadowing in Dynamical Systems, Lecture Notes in Mathematics, vol. 1706, Springer Berlin Heidelberg, 1999.
[42] A. Lazarus, O. Thomas, A harmonic-based method for computing the stability of periodic solutions of dynamical systems, C. R., Méc. 338 (9) (2010) 510–517, https://doi.org/10.1016/j.crme.2010.07.020.
[43] N.M. Wereley, Analysis and control of linear periodically time varying systems, Ph.D. thesis, Massachusetts Institute of Technology, Dept. of Aeronautics and Astronautics, 1991.
[44] D. Lasagna, A. Sharma, J. Meyers, Periodic shadowing sensitivity analysis of chaotic systems, J. Comput. Phys. 391 (2019) 119–141.
[45] J.M. Hyman, B. Nicolaenko, The Kuramoto-Sivashinsky equation: a bridge between PDE's and dynamical systems, Phys. D, Nonlinear Phenom. 18 (1) (1986) 113–126.
[46] P.J. Blonigan, Q. Wang, Least squares shadowing sensitivity analysis of a modified Kuramoto–Sivashinsky equation, Chaos Solitons Fractals 64 (2014) 16–25.
[47] P.R. Amestoy, I.S. Duff, J.-Y. L'Excellent, J. Koster, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM J. Matrix Anal. Appl. 23 (1) (2001) 15–41.