Resolvent-Based Optimisation for Approximating the Statistics of a Chaotic Lorenz System

Thomas Burton,[1] Sean Symon,[1] Ati S Sharma,[1,2] and Davide Lasagna[1,*]

[1] *Aerodynamics and Flight Mechanics Research Group*
University of Southampton
[2] *Agalmic Ltd.*

(Dated: February 12, 2025)

We propose a novel framework for approximating the statistical properties of turbulent flows by combining variational methods for the search of unstable periodic orbits with resolvent analysis for dimensionality reduction. Traditional approaches relying on identifying all short, fundamental unstable periodic orbits to compute ergodic averages via cycle expansion are computationally prohibitive for high-dimensional fluid systems. Our framework stems from the observation in Lasagna, Phys. Rev. E (2020), that a single unstable periodic orbit with a period sufficiently long to span a large fraction of the attractor captures the statistical properties of chaotic trajectories. Given the difficulty of identifying unstable periodic orbits for high-dimensional fluid systems, approximate trajectories residing in a low-dimensional subspace are instead constructed using resolvent modes, which inherently capture the temporal periodicity of unstable periodic orbits. The amplitude coefficients of these modes are adjusted iteratively with gradient-based optimisation to minimise the violation of the projected governing equations, producing trajectories that approximate, rather than exactly solve, the system dynamics. A first attempt at utilising this framework on a chaotic system is made here on the Lorenz 1963 equations, where resolvent analysis enables an exact dimensionality reduction from three to two dimensions. Key observables averaged over these trajectories produced by the approach as well as probability distributions and spectra rapidly converge to values obtained from long chaotic simulations, even with a limited number of iterations. This indicates that exact solutions may not be necessary to approximate the system's statistical behaviour, as the trajectories obtained from partial optimisation provide a sufficient "sketch" of the attractor in state space.

I. INTRODUCTION

The high sensitivity to initial conditions coupled with the high number of dynamically significant degrees of freedom pose significant challenges to the detailed prediction of the evolution of turbulent flows. Nonetheless, global behaviour and long-time statistical quantities may be characterised more conveniently by using the notion of a turbulent attractor, a low-dimensional object determining the long-time evolution of turbulent trajectories [1]. The geometry of the attractor may be quite complex and is often fractal in nature [2], allowing solutions to explore it in a complex fashion. For low-dimensional chaotic systems, insight into such geometry may be obtained by examining unstable period orbits (UPOs) densely embedded within the attractor and providing a skeleton that supports the dynamics in state space [3]. Using these UPOs in the form of a weighted sum, a technique known as cycle expansion, the ergodic averages of the dynamical system can be computed [4–6]. For turbulent flows, numerical evidence has been found for the existence of time periodic solutions of the Navier-Stokes equations, starting with the first identification of a nonlinear equilibrium flow by Nagata [7]. A representative set of literature which works on finding these solutions for various flow can be found in Refs. [8–13]. These nonlinear solutions are often referred to as Exact Coherent Structures (ECSs) or Recurrent Flows and are supposed to play the same role of UPOs of low-dimensional systems in shaping the structure of turbulent motion. The exact nature and significance of ECSs has not yet been fully understood [14, 15], but there is evidence that they are repeatedly shadowed by turbulent trajectories [16–18]. Attempts to apply cycle averaging formulae to predict the statistical properties of turbulence from ECSs have also been made [12, 19]. These have, however, faced the challenge that identifying all the structures required for the cycle expansion is not straightforward and is prohibitively expensive using available numerical methods [11, 20–25]. Further, as the dimension of the turbulent attractor increases with the Reynolds number, the likelihood that a turbulent trajectory shadows an ECS over its entire period decreases, thus impacting the quality of initial guesses generated using recurrence analysis techniques [26].

Given the difficulties in identifying a complete hierarchy of structures, we proposed in previous work [27] an alternative heuristic approach whereby computational resources are spent on identify one or few structures having a long time period [28]. Such solutions may span a relatively large fraction of the attractor, visiting the neighbourhoods of a variety of relevant dynamical states. They may thus provide a good approximation of the statistical properties of turbulence, such as averages and probability distributions, potentially as accurate as that obtained from a rigorous but incomplete hierarchy of solutions and lifting the technical burden of having to determine the weights for the cycle expansion [12, 29]. Ow-

* davide.lasagna@soton.ac.uk

ing to the temporal periodicity of these simpler objects, adjoint methods for time-periodic systems [30–32] may be used, despite the instability, to obtain sensitivities of statistical properties with respect to problem parameters of an external forcing [33]. This information obtained from UPOs may then be leveraged to design flow control strategies [34, 35] able to successfully manipulate the turbulent state.

For low-dimensional systems, this programme was carried out by first identifying UPOs with period hundreds or thousands times longer than the shortest UPO. These were found using a global Newton-Raphson search method [22, 33] that is insensitive to the marked sensitivity of such long chaotic trajectories to small perturbations that affects commonly-used shooting techniques. It was shown that period averages calculated on these UPOs appeared to converge to the long-time average of chaotic trajectories. Floquet exponents of such solutions, being the period averages of the local rate of growth of infinitesimal perturbations, also exhibited the same behaviour and appeared to converge to the Lyapunov exponents of the system calculated using standard methods [36]. Nevertheless, extending this programme to high-dimensional fluid systems does not seem feasible at present because of the above-mentioned objective difficulties in locating ECSs, let alone solutions with long period, even using more recent search methods [24, 25]. If, however, the purpose is to obtain approximations of statistical properties and their sensitivity to problem parameters, the question is whether such long solutions really need to be determined exactly or whether it might be computationally advantageous to accept some violation of the governing equations as long as the resulting "quasi-trajectories" provide a sufficiently detailed "sketch" of the attractor whilst, of course, remaining physically relevant. The approach discussed in this paper to construct such quasi-trajectories parallels the ideas recently proposed in McCormack *et al.* [37] and utilises dimensionality reduction as a strategy to trade-off the accuracy of statistical predictions with computational costs required to obtain them. Dimensionality reduction is motivated by the well-established notion that fluid systems display low-dimensional characteristics, and that the attractor lives in a low-dimensional subspace [38]. The dimensionality reduction used in this work demonstrates that the Lorenz attractor also lives in a low-dimensional subspace of the total state-space. Utilising the dimensionality reduction here, although not as pronounced as what can found for fluid systems, provides the framework that can be applied to the Navier-Stokes equations.

In this work, resolvent analysis in the formalism described in Ref. [39] is utilised. Resolvent analysis is an operator-based model reduction technique that has been utilised extensively for the purpose of analysis, control, and modelling of fluid flows [40–46], and has been shown to provide efficient low-dimensional representations of ECSs found for pressure-driven pipe flow and plane Couette flow [47], with the theoretical correspondence between resolvent modes and invariant solutions to the Navier-Stokes equations being shown in Ref. [48]. Resolvent analysis is used here to provide a hierarchy of temporal basis functions onto which the dynamics can be projected, with the assumption that these are able to capture the majority of the dynamics of the turbulent flow with fewer dynamical variables. These basis functions are especially suited to model time-periodic exact solutions, as a hierarchy of modes is obtained at each temporal frequency (and spatial wavenumber for problems with statistically homogeneous spatial directions), so that temporal periodicity is built in explicitly in the modal expansion. Quasi-trajectories are then generated by adapting the variational optimisation methodology developed in Refs. [23, 24, 49] to the present low-dimensional settings. The method is equivalent to optimising an objective that measures the violation of the projected governing equation by a candidate quasi-trajectory. The dimensionality reduction restricts the problem to only the space of solutions that is most dynamically significant, eliminating superfluous search directions of the optimisation and leading to computational savings. Previous attempts at utilising resolvent analysis to construct approximate solutions of the governing equations were made in Ref. [50], although steady solutions of Taylor-Couette flow were sought for, while here we consider time-dependent solutions that model dynamically-relevant processes. In addition, resolvent analysis has been used for a "project-and-search" algorithm [51] that sought new ECSs by projecting known solutions onto a reduced set of resolvent modes. These projected solutions are then used as the initial condition for a new search, which was then successful in finding new equilibria and periodic orbits that had not yet been observed in the literature.

The main contribution of this paper is to lay out the mathematical details of the framework used to generate time-periodic quasi-trajectories with the resolvent-based modelling, showing how variational optimisation methods and resolvent analysis can be combined. For the sake of providing a proof of concept on a prototypical chaotic system, the methodology is applied to the Lorenz 1963 system [52]. Despite being a low-dimensional problem, this system is chosen as it is a computationally accessible test-bed to make a first assessment of the numerical properties of the proposed framework. The application to fluid flows, and a detailed analysis of the role of the modal selection and truncation on the characteristics of the quasi-trajectories obtained is left to future work. The framework is introduced in Section II. In Section III the methodology is applied to the Lorenz 1963 system [52], with results reported in Section III B focusing on how statistics of observables obtained from quasi-trajectories compare to those obtained from long chaotic simulations. To conclude, the findings of the paper are summarised, and the future work is discussed in Section IV.

II. METHODOLOGY

In this section, the main components of the methodology are described. First the variational optimisation is introduced, showing how it can construct exact periodic solutions. This is followed with an explanation of the resolvent analysis, showing how resolvent modes are generated, what they represent, and how they can be truncated to allow the construction of a reduced-order model. This leads to an explanation of how the resolvent modes can be used as a basis for a low-dimensional projection of the dynamics, restricting the variational optimisation to the resolvent subspace. The section is then concluded with the numerical details of the implementation, explaining how the resolvent-based optimisation is constructed and how the main dynamical quantities are computed.

A. Variational Optimiser

To construct quasi-trajectories, an optimisation approach is used. In essence, this methodology aims to minimise an objective that measures the violation of the governing equation by a given state space trajectory. Although the framing is different, this approach is exactly equivalent to the adjoint solver methodology described in Ref. [49].

Consider the general autonomous dynamical system defined with the evolution equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{g}\left(\boldsymbol{x}\right), \quad \boldsymbol{x} \in \mathcal{M} \subseteq \mathbb{R}^d, \tag{1}$$

where $\mathcal{M}$ is the state space (or phase space) for the system. The vector field $\boldsymbol{g}$ is assumed to be smooth, which implies that any solution to Equation 1 is also smooth. The variational optimiser aims to find periodic solutions to Equation 1, such that the following condition is satisfied

$$\int_t^{t+T} \boldsymbol{g}\left(\boldsymbol{x}\left(t'\right)\right) \mathrm{d}t' = 0, \quad \forall t \in \mathbb{R}^{\geq 0},\, T > 0, \tag{2}$$

where T is the period of the solution. In general T is not known a priori and should be included as part of the optimisation. We define the scaled time as $s = 2\pi t/T = \omega t$, which is useful for decoupling the variation due to changes in shape of the trajectory and changes in its period. The variable $\omega = 2\pi/T$ is called the fundamental frequency and corresponds to the smallest frequency that can be permitted in a finite time period. The trajectory $\boldsymbol{x}$ can now be expressed as a function over $s \in [0,\, 2\pi)$. Additionally, the time derivative can now be expressed in terms of the scaled time as $d/dt = \omega d/ds$.

To help in the characterisation of the problem, the space of closed state space loops is defined as

$$\mathscr{P} = \left\{\boldsymbol{x}\left(s\right) \,|\, \boldsymbol{x}\left(0\right) = \boldsymbol{x}\left(2\pi\right)\right\}. \tag{3}$$

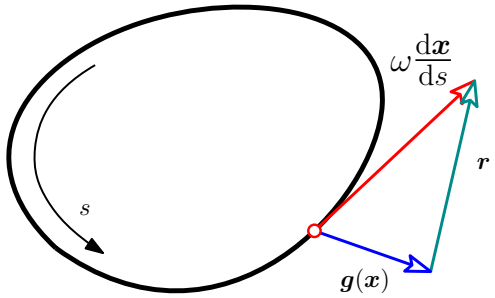

FIG. 1. A state space loop that does not satisfy the governing equations, and thus has a non-zero residual $\boldsymbol{r}$.

Thus, a trajectory is in $\mathscr{P}$ if and only if it is periodic. If a trajectory $\boldsymbol{x}$ and a given fundamental frequency ω satisfy Equation 2, then $\boldsymbol{x} \in \mathscr{P}$. Thus, there is a subset of $\mathscr{P}$ that represents exact periodic solutions to Equation 1. The general approach is to consider a particular initial loop $\boldsymbol{x}_0 \in \mathscr{P}$ with a fundamental frequency ω_0, and then modify both $\boldsymbol{x}_0$ and ω_0 according to some update law such that $\boldsymbol{x}_n \in \mathscr{P}$ and the limit of $\boldsymbol{x}_n$ and ω_n is a solution to Equation 1

We define an inner-product on the space $\mathscr{P}$ as follows

$$\langle \boldsymbol{x},\, \boldsymbol{y} \rangle := \int_0^{2\pi} \boldsymbol{x} \cdot \boldsymbol{y}\, \mathrm{d}s, \tag{4}$$

which induces the norm $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x},\, \boldsymbol{x} \rangle}$. Then we define a local residual

$$\boldsymbol{r} := \omega \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s} - \boldsymbol{g}\left(\boldsymbol{x}\right), \tag{5}$$

which is a measure of the local violation of the trajectory $\boldsymbol{x}$ with respect to the governing equations. Figure 1 is a schematic for what the local residual represents at each location in state space, given by the vector bridging the distance between the tangent vector to the state space loop $(\mathrm{d}\boldsymbol{x}/\mathrm{d}s)$ and the vector field $\boldsymbol{g}$. We also note that if $\boldsymbol{x} \in \mathscr{P}$ then it is true that $\boldsymbol{r} \in \mathscr{P}$. Thus, the problem can now be understood as finding some way to search for the trajectory $\boldsymbol{x}$ within the space $\mathscr{P}$ such that $\|\boldsymbol{r}\| = 0$, which is only true if $\boldsymbol{r}\left(t\right) = 0$ for all $t \in [0,\, T)$. This motivates the definition of the global residual

$$\mathcal{R}\left[\boldsymbol{x},\, \omega\right] := \frac{1}{2}\|\boldsymbol{r}\|^2, \tag{6}$$

as the measure of global violation of the governing equations by $\boldsymbol{x}$. By minimising $\mathcal{R}$ the alignment between the tangent vector $\mathrm{d}\boldsymbol{x}/\mathrm{d}s$ and the governing vector field $\boldsymbol{g}$ is maximised. Therefore, if $\boldsymbol{x} \in \mathscr{P}$ is a solution to Equation 1 for a given period T then $\mathcal{R} = 0$; otherwise $\mathcal{R} > 0$.

The optimisation problem to find exact periodic solutions can now be summarised as follows

$$\min_{\boldsymbol{x} \in \mathscr{P},\, \omega} \mathcal{R}\left[\boldsymbol{x},\, \omega\right]. \tag{7}$$

To solve this optimisation problem, it is desirable to have access to the gradient information on $\mathcal{R}$ with respect

to variations in $\boldsymbol{x}$ and ω. Variational calculus provides the tools to derive the functional derivative of $\mathcal{R}$ with respect to a given state space trajectory $\boldsymbol{x}$, giving the closed form expression

$$\frac{\delta \mathcal{R}}{\delta \boldsymbol{x}} = -\omega \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} - \boldsymbol{L}^{\top}(\boldsymbol{x})\,\boldsymbol{r}, \tag{8}$$

where $\boldsymbol{L} = \mathrm{d}\boldsymbol{g}/\mathrm{d}\boldsymbol{x}$ is the Jacobian of $\boldsymbol{g}$ evaluated over the trajectory $\boldsymbol{x}$, and $(\cdot)^{\top}$ denotes the matrix transpose. This is a matrix ($\boldsymbol{L} \in \mathbb{R}^{d\times d}$) populated with the partial derivatives of the components of $\boldsymbol{g}$ with respect to the components of $\boldsymbol{x}$. The derivative with respect to ω is given as

$$\frac{\partial \mathcal{R}}{\partial \omega} = \left\langle \frac{\partial \boldsymbol{u}}{\partial s}, \boldsymbol{r} \right\rangle. \tag{9}$$

The details of the derivation of both Equation 8 and Equation 9 are given in Appendix A 1. Thus, by taking an initial state space loop and providing an optimisation algorithm with the objective $\mathcal{R}$ and gradients $\delta\mathcal{R}/\delta\boldsymbol{x}$ and $\partial\mathcal{R}/\partial\omega$, a monotonic decrease in $\mathcal{R}$ is guaranteed.

B. Resolvent Analysis

The variational optimisation approach allows the construction of exact solutions to a set of dynamical equations. However, as discussed in Section I this is a costly and time-consuming procedure for high-dimensional systems encountered when dealing with spatio-temporally varying fields such as those produced by the Navier-Stokes equations. Resolvent analysis is introduced here as a way to construct low-order models of general dynamical systems that would allow for a projection onto a smaller subset reducing the dimensionality of the system while retaining important dynamical information [47, 53, 54]. The formalism described in [39] has been specialised here for finite-dimensional systems.

Given a (not necessarily periodic) state space trajectory $\boldsymbol{x}$ with temporal length T, define the mean $\overline{\boldsymbol{x}}$ as follows

$$\overline{\boldsymbol{x}} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \boldsymbol{x}(t)\,\mathrm{d}t\,. \tag{10}$$

Then the trajectory can be decomposed into the mean and fluctuations

$$\boldsymbol{x}(t) = \overline{\boldsymbol{x}} + \boldsymbol{x}'(t)\,, \tag{11}$$

where $\boldsymbol{x}'$ is the fluctuations of the trajectory. This decomposition can be substituted into Equation 1, noting that $\overline{\boldsymbol{x}}$ is invariant in time, to obtain an evolution equation for the state fluctuations

$$\frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t} = \boldsymbol{c} + \boldsymbol{L}(\overline{\boldsymbol{x}})\,\boldsymbol{x}' + \boldsymbol{f}(\boldsymbol{x}')\,, \tag{12}$$

where $\boldsymbol{c} = \boldsymbol{g}(\overline{\boldsymbol{x}})$, $\boldsymbol{L}(\overline{\boldsymbol{x}})$ is the same Jacobian matrix as given in Equation 8 now evaluated only at $\overline{\boldsymbol{x}}$, and $\boldsymbol{f}$ is all the nonlinear terms in the expansion. Note that the constant $\boldsymbol{c} \neq 0$ since the mean $\overline{\boldsymbol{x}}$ is not an equilibrium of the system.

Since we are concerned only with periodic solutions, it is natural to expand the trajectory as a Fourier series such that the condition $\boldsymbol{x} \in \mathscr{P}$ is automatically enforced, which is done as follows

$$\boldsymbol{x}'(s) = \sum_{n\in\mathbb{Z}^+} (\boldsymbol{x}'_n + \text{c.c.})\, e^{ins}. \tag{13}$$

where the sum is over the positive integers $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$ and c.c. denotes the complex conjugate of each term. Since $\boldsymbol{x}'$ is real valued, the resulting series possesses a Hermitian symmetry i.e. $\boldsymbol{x}'_{-n} = \boldsymbol{x}'^{*}_n$, where $(\cdot)^*$ denotes the complex conjugate, allowing the sum to be expressed only over the positive frequencies. In addition, the zero frequency is not included in Equation 13 since the fluctuations by definition have zero mean component, that is, $\boldsymbol{x}_0 = 0$. The Fourier coefficients are related to the fluctuation with the identity

$$\boldsymbol{x}'_n = \frac{1}{2\pi}\int_0^{2\pi} \boldsymbol{x}'(s)\, e^{-ins}\,\mathrm{d}s. \tag{14}$$

Expanding Equation 12 in terms of the Fourier components provides an algebraic equation governing the Fourier coefficients of the state fluctuations

$$in\omega\boldsymbol{x}'_n = \boldsymbol{L}(\overline{\boldsymbol{x}})\,\boldsymbol{x}'_n + \boldsymbol{f}_n, \quad n \in \mathbb{Z}\setminus\{0\}\,, \tag{15a}$$

$$0 = \boldsymbol{c} + \boldsymbol{f}_0, \tag{15b}$$

where $\boldsymbol{f}_n$ are the Fourier coefficients of the nonlinear terms given by

$$\boldsymbol{f}_n = \frac{1}{2\pi}\int_0^{2\pi} \boldsymbol{f}(\boldsymbol{x}'(s))\, e^{-ins}\,\mathrm{d}s. \tag{16}$$

Equation 15a is the governing equation for the fluctuations' evolution. Equation 15b is a constraint on the fluctuations imposed by the mean state, analogous to the Reynolds averaged Navier-Stokes equations, obtained through the same mean-fluctuation decomposition of a velocity field and averaging.

Rearranging Equation 15a to make the fluctuation coefficients the subject gives the following relationship with the nonlinearity

$$\boldsymbol{x}'_n = \boldsymbol{H}_n\boldsymbol{f}_n, \quad n \in \mathbb{Z}\setminus\{0\}\,. \tag{17}$$

where $\boldsymbol{H}_n \in \mathbb{C}^{d\times d}$ is the resolvent matrix, defined as

$$\boldsymbol{H}_n = (in\omega\boldsymbol{I} - \boldsymbol{L})^{-1}\,, \quad n \in \mathbb{Z}\setminus\{0\} \tag{18}$$

Equation 17 shows that the resolvent matrix acts as a transfer function relating the deviations of the system around the mean state due to the some forcing $\boldsymbol{f}_n$. This forcing term can be any general forcing in different contexts, however here it is known to represent the nonlinear

interactions within the system that act to spread out the spectral content of the solution.

The next step in resolvent analysis is to perform a Singular Value Decomposition (SVD) on the resolvent matrix

$$\boldsymbol{H}_n = \tilde{\boldsymbol{\Psi}}_n \tilde{\boldsymbol{\Sigma}}_n \tilde{\boldsymbol{\Phi}}_n^\dagger, \tag{19}$$

where $\tilde{\boldsymbol{\Psi}}_n \in \mathbb{C}^{d\times d}$ and $\tilde{\boldsymbol{\Phi}}_n \in \mathbb{C}^{d\times d}$ are the response and forcing resolvent matrices, respectively, and $\tilde{\boldsymbol{\Sigma}}_n \in \mathbb{R}^{d\times d}$ is the diagonal matrix of singular values, denoted with σ_i, ordered from largest to smallest. Here the $(\cdot)^\dagger$ represents the conjugate transpose of a matrix. Each column of $\tilde{\boldsymbol{\Psi}}_n$ and $\tilde{\boldsymbol{\Phi}}_n$ are a single response and forcing mode of the resolvent, which form a complete basis for the state of the system and the nonlinear forcing that drives it, respectively. The singular value associated with each mode pair provides the magnitude of response induced by the associated forcing, ordered by the size of said response. If the singular values in $\tilde{\boldsymbol{\Sigma}}_n$ decay quickly, i.e. $\sigma_i \gg \sigma_{i+1}$, then the system is most receptive to certain forcing modes in $\tilde{\boldsymbol{\Phi}}_n$, producing very large responses in the system in certain directions corresponding to the left-most columns of $\tilde{\boldsymbol{\Psi}}_n$. This means the resolvent, and the system as a whole, can be represented accurately with a truncated version of Equation 19 that removes the smaller singular values, as well as the forcing and response modes associated with them. Doing so it is possible to define the truncated SVD

$$\boldsymbol{H}_n \approx \boldsymbol{\Psi}_n \boldsymbol{\Sigma}_n \boldsymbol{\Phi}_n^\dagger, \tag{20}$$

where $\boldsymbol{\Psi}_n \in \mathbb{C}^{d\times d_r}$, $\boldsymbol{\Phi}_n \in \mathbb{C}^{d\times d_r}$, and d_r is the number of modes retained in each of the matrices. The right-most singular vectors of $\boldsymbol{\Psi}_n$ and $\boldsymbol{\Phi}_n$ have therefore been discarded along with their associated singular values. The degree of this truncation depends on the application. In the case of fluid dynamics and turbulence it has been shown that large separations of scale can occur between the singular values for certain wall-bounded flows [39].

C. Dimensionality Reduction

As mentioned, the response matrix $\tilde{\boldsymbol{\Psi}}_n$ provide a complete basis for the state of the system. With the truncation established in Equation 20 it is now possible to use the reduced set of modes defined by the matrix $\boldsymbol{\Psi}_n$ as the basis for a projection that can reduce the dimensionality of the system, restricting the variational optimiser of Section II A to the resolvent subspace. This projection onto the resolvent subspace is defined as

$$\boldsymbol{a}_n = \boldsymbol{\Psi}_n^\dagger \boldsymbol{x}'_n \quad \Leftrightarrow \quad \boldsymbol{x}'_n = \boldsymbol{\Psi}_n \boldsymbol{a}_n. \tag{21}$$

Figure 2 displays a schematic for this projection, showing that $\boldsymbol{a}_n \in \mathbb{C}^{d_r}$ in Equation 21 represents the best approximation to the state of the dynamical system within the subspace defined by the resolvent response modes.

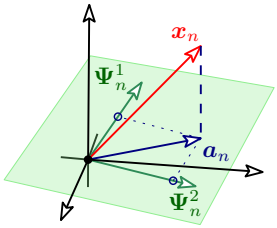

FIG. 2. Schematic for the projection of the state onto the subspace defined by the truncated response modes.

To work in this new reduced space, it is necessary to modify the general optimisation problem given in Equation 7 such that it can be expressed purely in terms of the resolvent subspace. A consequence of working in the reduced space defined by the modes in $\boldsymbol{\Psi}_n$ is that the mean state $\overline{\boldsymbol{x}}$ and the fundamental frequency ω are fixed, since they are a part of the derivation process of the modes themselves. They are required to be prescribed a priori and fixed over the duration of the optimisation of a quasi-trajectory. In principle, resolvent analysis can provide a basis for the mean, allowing for the optimisation to be performed over the whole spectrum. However, this is not implemented in practice since the mean is required regardless to construct the modes. Another consequence of working in the resolvent subspace, fixing the truncated modes $\boldsymbol{\Psi}_n$ and the mean $\overline{\boldsymbol{x}}$, is that it is possible that the residual may not have a zero. That is, it may not be possible to find an exact solution within the resolvent subspace defined by $\boldsymbol{\Psi}_n$ with the prescribed fundamental frequency ω and mean $\overline{\boldsymbol{x}}$. This means the solutions to UPOs is not possible in general given this methodology, however this is not a significant issue as the goal is construct periodic quasi-trajectories that only approximate the geometry of the attractor in state space. To find an exact UPOs, it is necessary to exactly know the mean of that particular UPO before starting the optimisation or allow the optimisation to make changes to the mean, provided some basis $\boldsymbol{\Psi}_0$, and the frequency. It is noted, however, that long UPOs have means that approach the mean obtained from chaotic solutions [27], and therefore the chaotic mean can be used in the optimisation incurring only a small error. In regard to the fundamental frequency, it is known that as the period is increased the density of UPOs increases exponentially, thus for quasi-trajectories with very large periods it is known that solutions exist very close to the prescribed period and thus the convergence of the quasi-trajectory is not affected until small residuals are reached. In other words, if a small enough ω is chosen for a quasi-trajectory, i.e. the length of the trajectory is long, then the quasi-trajectory will be attracted close to a solution that exists with a very similar period.

In summary, the projection of the dynamics onto a the resolvent subspace defined by $\boldsymbol{\Psi}_n$ with a fixed $\overline{\boldsymbol{x}}$ and ω provides a reduced space on which the variational optimisation of Section II A can be performed. In general,

this subspace does not contain zeros of $\mathcal{R}$, representing UPOs of the system. Instead, the low-order model constructed, for large enough periods and using the approximate chaotic mean, can provide a good enough approximation of the dynamics to allow the construction of these quasi-trajectories that can output accurate statistics without being required to have $\mathcal{R} \approx 0$.

Computing the global residual given a set of modal coefficients $\boldsymbol{a}_n$ can be done as follows. First, the definition of the global residual given in Equation 6 can be expressed in spectral space by substituting in the Fourier series of the local residual

$$\mathcal{R} = \frac{1}{2}\boldsymbol{r}_0^\dagger \boldsymbol{r}_0 + \sum_{n\in\mathbb{Z}^+} \boldsymbol{r}_n^\dagger \boldsymbol{r}_n, \tag{22}$$

where

$$\boldsymbol{r}_n = \frac{1}{2\pi}\int_0^{2\pi} \boldsymbol{r}\left(s\right) e^{ins}\,\mathrm{d}s. \tag{23}$$

The coefficients $\boldsymbol{r}_n$ are related to the state-vector as follows

$$\boldsymbol{r}_n = in\omega\boldsymbol{x}_n' - \boldsymbol{L}\left(\overline{\boldsymbol{x}}\right)\boldsymbol{x}_n' + \boldsymbol{f}_n,\, n\in\mathbb{Z}\setminus\{0\}, \tag{24a}$$
$$\boldsymbol{r}_0 = \boldsymbol{c} + \boldsymbol{f}_0, \tag{24b}$$

which are the Fourier transform of the governing equations as given in Equation 15. We remark that the equations for both the fluctuations and the mean are included in Equation 22.

Left multiplying Equation 24 by $\boldsymbol{\Psi}_n^\dagger$ and using Equation 21 a reduced space local residual expression can be derived

$$\boldsymbol{\rho}_n = in\omega\boldsymbol{a}_n - \boldsymbol{L}\left(\overline{\boldsymbol{x}}\right)\boldsymbol{a}_n + \boldsymbol{\Psi}_n^\dagger\boldsymbol{f}_n,\, n\in\mathbb{Z}\setminus\{0\}, \tag{25a}$$
$$\boldsymbol{\rho}_0 = \boldsymbol{\Psi}_0^\dagger\left(\boldsymbol{c}+\boldsymbol{f}_0\right), \tag{25b}$$

where $\boldsymbol{\rho}_n = \boldsymbol{\Psi}_n^\dagger\boldsymbol{r}_n$. Finally the global residual can then be computed with

$$\mathcal{R} = \frac{1}{2}\boldsymbol{\rho}_0^\dagger\boldsymbol{\rho}_0 + \sum_{n\in\mathbb{Z}^+}\boldsymbol{\rho}_n^\dagger\boldsymbol{\rho}_n, \tag{26}$$

utilising the fact that $\boldsymbol{\Psi}_n^\dagger\boldsymbol{\Psi}_n = \boldsymbol{I}$. With this in mind, the resolvent-based optimisation problem can be stated as such

$$\min_{\boldsymbol{a}_n,\,\forall n\in\mathbb{Z}\setminus\{0\}} \quad \mathcal{R}\left(\{\boldsymbol{a}_n\}\right). \tag{27}$$

In effect, the only change between Equation 27 and Equation 7 is that the optimisation variables have been converted from the full-space trajectory $\boldsymbol{x}$ and ω, to just $\boldsymbol{a}_n$. It should also be noted that the coefficient $\boldsymbol{a}_0$ is not included in the optimisation, since $\boldsymbol{a}_n$ is defined in Equation 21 in terms of the fluctuations $\boldsymbol{x}_n'$. This means $\boldsymbol{a}_0 = 0$ by definition and is not included as part of the optimisation problem, which is a consequence of fixing $\overline{\boldsymbol{x}}$. The computational details of computing $\mathcal{R}$ are given in Section II D.

In order to perform gradient-based optimisation within this resolvent subspace, an expression for $\partial\mathcal{R}/\partial\boldsymbol{a}_n$ is required. The gradient of $\mathcal{R}$ with respect to ω is not required in the reduced space as ω is fixed over the duration of the optimisation. Obtaining an expression for $\partial\mathcal{R}/\partial\boldsymbol{a}_n$ is done in two transformation steps. First Fourier transform Equation 8 to get the following expression

$$\frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}}e^{-ins}\mathrm{d}s, \tag{28}$$
$$= -in\omega\boldsymbol{r}_n - \left(\boldsymbol{L}\left(\boldsymbol{x}\right)^\top\boldsymbol{r}\right)_n. \tag{29}$$

Second, we project $\partial\mathcal{R}/\partial\boldsymbol{x}_n$ onto the resolvent subspace in the same way the state is projected to obtain

$$\frac{\partial\mathcal{R}}{\partial\boldsymbol{a}_n} = \boldsymbol{\Psi}_n^\dagger\frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}, \quad n\in\mathbb{Z}\setminus 0. \tag{30}$$

The proof of Equation 29 and Equation 30 is given in Appendix A 2. Note that the mean frequency is not included due to the modes $\boldsymbol{\Psi}_n$ not being properly defined for the mean. This automatically ensures that using Equation 30 for a gradient-based optimisation does not modify the mean coefficient $\boldsymbol{a}_0$. As long as the optimisation is initialised with $\boldsymbol{a}_0 = 0$, the mean $\overline{\boldsymbol{x}}$ is guaranteed to be fixed.

D. Numerical Details

The optimisation process for each iteration is visualised with the flow diagram given in Figure 3. The optimisation is initialised with a mean state $\overline{\boldsymbol{x}}$ and fundamental frequency ω that are fixed throughout the duration of the optimisation. These are used as inputs for resolvent analysis to generate the set of modes $\boldsymbol{\Psi}_n$ for $n\in\mathbb{Z}^+$. Next, the initial trajectory is generated. In the literature it is common to use close recurrences of chaotic trajectories obtained from direct simulations to initialise a given trajectory when trying to find exact periodic solutions [4, 24], which is particularly important for high-dimensional systems as the radius of convergence in such cases is small compared to the space in which the state can inhabit. In this work the results obtained were found to be very robust to the initial guess for the quasi-trajectory, and so the coefficients $\boldsymbol{a}_n$ were initialised randomly with a Gaussian distribution. Once $\overline{\boldsymbol{x}}$, $\boldsymbol{\Psi}_n$, and an initial $\boldsymbol{a}_n$ are known, the optimisation loop can begin.

When computing the residual and its gradient a "pseudo-spectral" method is used to reduce the time complexity of computing the nonlinear (typically quadratic) terms. This takes the form of expanding the coefficients $\boldsymbol{a}_n$ and inverse Fourier transforming the result to obtain the time domain representation of the quasi-trajectory. The nonlinear terms $\boldsymbol{f}\left(\boldsymbol{x}\right)$ can then

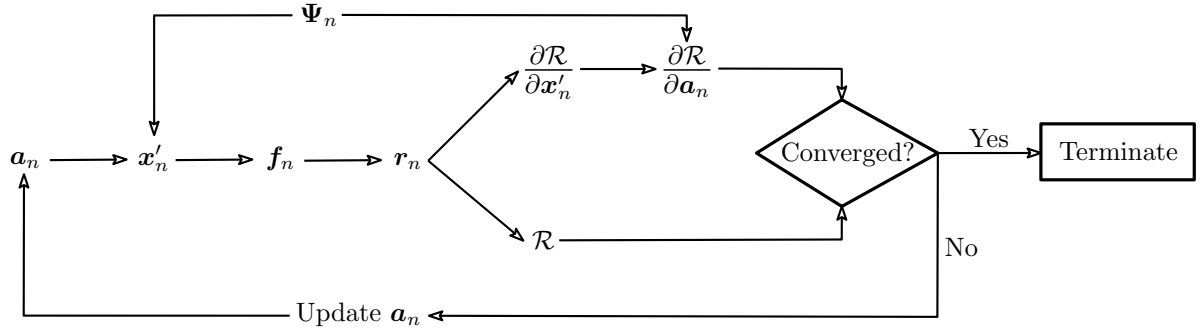

FIG. 3. Flow diagram of a single iteration of the optimisation loop used to compute the residual and its gradient starting from the modal coefficients $\boldsymbol{a}_n$.

be efficiently computed over the length of the quasi-trajectory, after which the result is Fourier transformed to obtain $\boldsymbol{f}_n$. Taking note of Figure 3, $\boldsymbol{f}_n$ is then used directly to compute $\boldsymbol{r}_n$. This means that over the course of the optimisation the trajectory is transformed from reduced to full space and then from spectral to time domain to be able to compute all the terms in the space where it is computationally most efficient. The result is the gradient $\partial\mathcal{R}/\partial\boldsymbol{x}'_n$, which is then projected into the reduced space using Equation 30, providing the required gradient with respect to the coefficients $\boldsymbol{a}_n$.

The choice of optimisation algorithm is an important aspect which has been relatively unexplored in the literature. As already mentioned, the choice of Refs. [24, 25, 49] was to use an algorithm equivalent to gradient descent. This can be inefficient as the gradient descent has linear convergence as the minimum is approached [55]. For this reason, the L-BFGS optimisation algorithm is selected for this work [56, 57]. L-BFGS is a gradient-based quasi-Newton algorithm, first used to solve for periodic orbits in Ref. [58]. L-BFGS included approximate Hessian information, significantly improves convergence rates, and generally requires fewer function evaluations compared to conjugate gradient methods, making a generally efficient method for gradient-based optimisation [59]. The L-BFGS iterations are coupled with a line search algorithm detailed in Refs. [55, 60].

The choice of convergence criteria is important in this context. Usually, a small global residual is used as it indicates that an exact solution to Equation 1 has been found. In this work the focus is on quasi-trajectories for which the global residual is not expected to converge to a small value. Hence, it makes more sense here to track some relevant time-averaged observable of the system and terminate the iterations once this observable has converged to some value. This is explored more in Section III B using statistical measures of the Lorenz system.

III. APPLICATION TO THE LORENZ SYSTEM

The Lorenz system from Ref. [52] will be used for the purpose of demonstrating the above methodology on a well known chaotic system. Even though it is an extensively studied problem, we demonstrate that the Lorenz equations permit an elegant, and to the best of the authors' knowledge not previously reported, dimensionality reduction when resolvent analysis is utilised, from three dimensions to two.

The governing equations are given as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma\left(y - x\right), \tag{31a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = x\left(\rho - z\right) - y, \tag{31b}$$

$$\frac{\mathrm{d}z}{\mathrm{d}t} = xy - \beta z. \tag{31c}$$

The standard parameter values of $\sigma = 10$, $\rho = 28$, and $\beta = \frac{8}{3}$ are used, for which it is known the system exhibits chaotic motion confined to a strange attractor. The governing equations are symmetric under the transformation $[x,\, y,\, z] \to [-x,\, -y,\, z]$, which implies that the mean has the form $\overline{\boldsymbol{x}} = \begin{pmatrix} 0 & 0 & \overline{z} \end{pmatrix}^\top$ where $\overline{z}$ denotes the mean in the z-direction.

A. Resolvent Derivation and Low-Oder Model

By applying the mean-fluctuation decomposition of Equation 11 to the Lorenz system, the following evolution equation for the state fluctuations is obtained

$$\frac{\mathrm{d}x'}{\mathrm{d}t} = \sigma\left(y' - x'\right), \tag{32a}$$

$$\frac{\mathrm{d}y'}{\mathrm{d}t} = \left(\rho - \overline{z}\right)x' - y' - x'z', \tag{32b}$$

$$\frac{\mathrm{d}z'}{\mathrm{d}t} = -\beta\left(\overline{z} + z'\right) + x'y'. \tag{32c}$$

Equation 32 can be expressed compactly as

$$\frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t} = \boldsymbol{c} + \boldsymbol{L}\left(\overline{\boldsymbol{x}}\right)\boldsymbol{x}' + \boldsymbol{M}\boldsymbol{f}\left(\boldsymbol{x}'\right), \tag{33}$$

with $\boldsymbol{c} = \begin{pmatrix} 0 & 0 & -\beta\overline{z} \end{pmatrix}^\top$ being the constant mean response of the system, the linearised Lorenz matrix evaluated at $\overline{\boldsymbol{x}}$ given as

$$\boldsymbol{L}\left(\overline{\boldsymbol{x}}\right) = \begin{pmatrix} -\sigma & \sigma & 0 \\ \rho - \overline{z} & -1 & 0 \\ 0 & 0 & -\beta \end{pmatrix},$$

and the nonlinear influence matrix $\boldsymbol{M}$ defined as

$$\boldsymbol{M} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{34}$$

The nonlinear forcing $\boldsymbol{f}\left(\boldsymbol{x}'\right) = \begin{pmatrix} x'z' & x'y' \end{pmatrix}^\top$ is a two-element vector due to the fact that the first equation in Equation 31 is linear.

By decomposing Equation 33 into its Fourier modes and then rearranging, the resolvent operator for the Lorenz system is obtained

$$\boldsymbol{H}_n = \left(in\omega\boldsymbol{I} - \boldsymbol{L}\right)^{-1}\boldsymbol{M} = \begin{pmatrix} \alpha_n & 0 \\ \beta_n & 0 \\ 0 & \gamma_n \end{pmatrix}, \tag{35}$$

where

$$\alpha_n = -\sigma/D_n, \tag{36}$$
$$\beta_n = -\left(in\omega + \sigma\right)/D_n, \tag{37}$$
$$\gamma_n = 1/\left(in\omega + \beta\right), \tag{38}$$
$$D_n = \left(in\omega + 1\right)\left(in\omega + \sigma\right) + \sigma\left(\overline{z} - \rho\right). \tag{39}$$

The matrix $\boldsymbol{M}$ reduces the size of the resolvent from a square matrix to a rank-2 rectangular matrix. The convenient structure allows an explicit expression for the SVD to be obtained as follows

$$\boldsymbol{H}_n = \boldsymbol{\Psi}_n\boldsymbol{\Sigma}_n\boldsymbol{\Phi}_n^\dagger \tag{40}$$
$$= \begin{pmatrix} \zeta_n & 0 \\ \eta_n & 0 \\ 0 & \kappa_n \end{pmatrix}\begin{pmatrix} \sigma_{1,n} & 0 \\ 0 & \sigma_{2,n} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{41}$$

where the coefficients of the left singular matrix $\boldsymbol{\Psi}_n$ are given as $\zeta_n = \alpha_n/\sigma_{1,n}$, $\eta_n = \beta_n/\sigma_{1,n}$, and $\kappa_n = \gamma_n/\sigma_{2,n}$. The rank-2 nature of the resolvent means there are exactly 2 singular values that govern the transfer of nonlinear forcing to the solution of the system. The response modes, defined by the columns of $\boldsymbol{\Psi}_n$, have the property that the first mode contains all the information from the xy-plane, while the second mode contains only the information in the z-direction. As such, retaining only a single pair of the modes restricts the dynamics to only the xy-plane or z-axis.

Due to the right singular vector, $\boldsymbol{\Phi}_n$, being equal to the identity matrix as shown in Equation 41 the following expression for $\boldsymbol{\Sigma}_n$ can be derived

$$\boldsymbol{\Psi}^\dagger\boldsymbol{\Psi}_n = \boldsymbol{\Sigma}_{\boldsymbol{n}}^{-\mathbf{1}}\boldsymbol{H}_n^\dagger\boldsymbol{H}_n\boldsymbol{\Sigma}_n^{-1} = \boldsymbol{I},$$
$$\Rightarrow \boldsymbol{\Sigma}_n^2 = \boldsymbol{H}_n^\dagger\boldsymbol{H}_n,$$

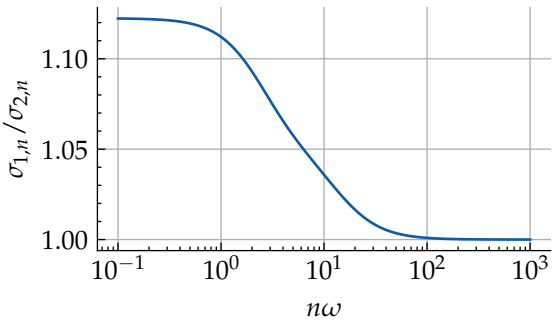

FIG. 4. Ratio of the singular values for the Lorenz system plotted against the frequency.

which gives for the individual singular values

$$\sigma_{1,n} = \sqrt{\frac{\left(n\omega\right)^2 + 2\sigma^2}{|D_n|^2}}, \tag{42}$$

$$\sigma_{2,n} = \sqrt{\frac{1}{\left(n\omega\right)^2 + \beta^2}}. \tag{43}$$

Figure 4 shows the ratio of the singular values as given in Equation 42 and Equation 43 as the frequency $n\omega$ is varied. There is no large separation of scale observed between these singular values and so the system cannot be accurately represented with a modal coefficients $\boldsymbol{a}_n \in \mathbb{C}^{d_r}$. Physically this is obvious if the response modes in Equation 41 are inspected. If one of the response modes is neglected, then the dynamics is constrained to only the xy-plane or the z-axis (depending on which response modes is rejected), which cannot accurately reconstruct the structure of the strange attractor embedded in the full state space.

Therefore, the dimensionality reduction used in this work is from $\mathbb{C}^3$ to $\mathbb{C}^2$. This is an "exact" projection. That is to say, there has been no rejection of any non-zero singular values. This can be considered a special case of the more general (usually higher-dimension) setting, where there are more non-zero singular values as well as a distinctive separation of scales allowing the rejection of a finite number of relatively small singular values.

B. Quasi-Trajectory Statistics

In this section, the quasi-trajectories generated using the low-order resolvent-based model discussed in Section III A for the Lorenz system are assessed in their ability to approximate the statistics of chaotic solutions. To facilitate this, the following observables are defined

$$\mathcal{J}_1\left(\boldsymbol{x}\right) = \sqrt{x^2 + y^2 + z^2}, \qquad \mathcal{J}_2\left(\boldsymbol{x}\right) = xz. \tag{44}$$

These observables will be averaged over the duration of a trajectory with period T, denoted by $\overline{\mathcal{J}_i}^T$.

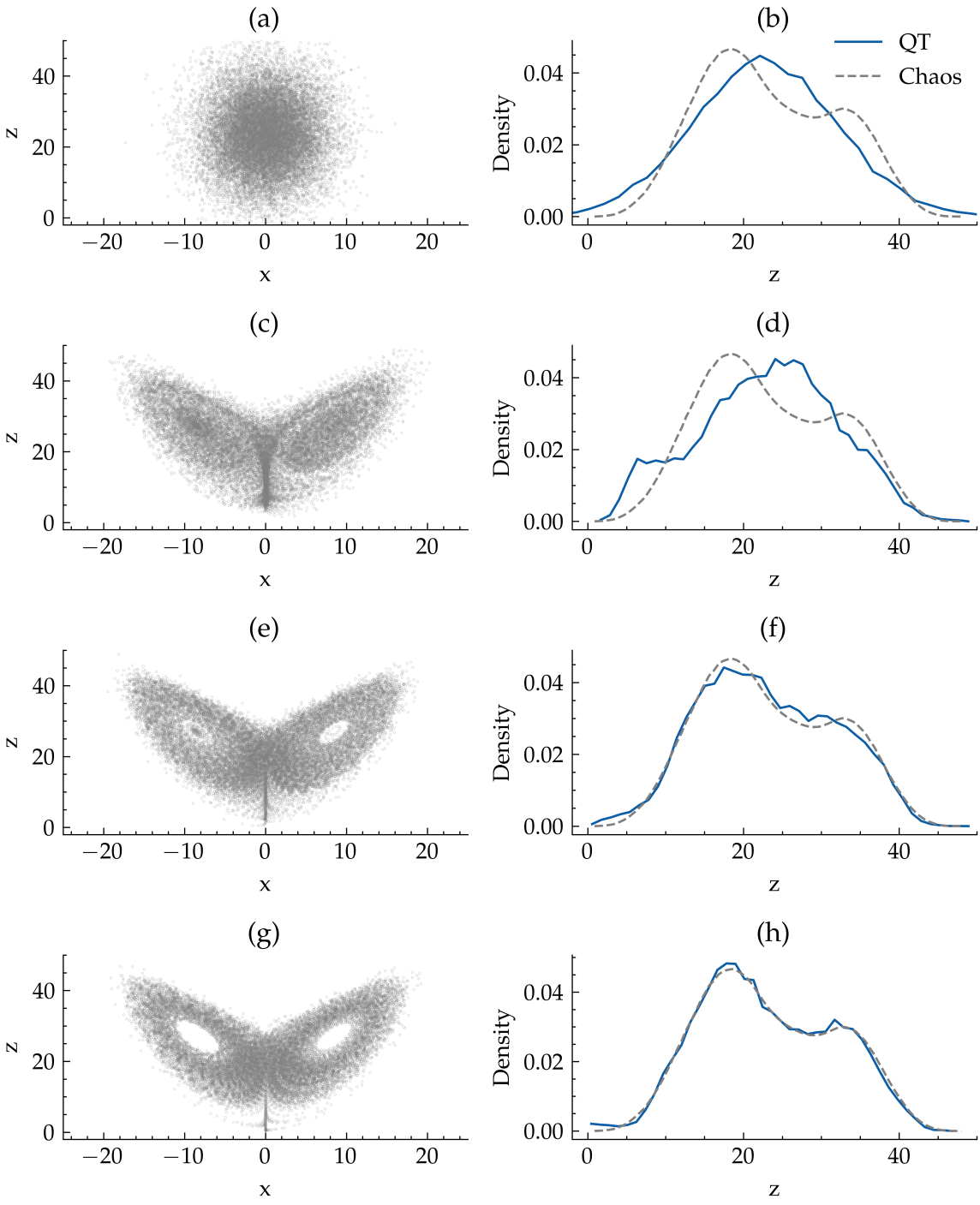

FIG. 5. The optimisation of a quasi-trajectory, with the state space points shown in (a), (c), (e), and (g), and the probability distribution functions over the z-direction shown in (b), (d), (f), and (h), with the blue line corresponding to the Quasi-Trajectory (QT) and the dashed grey line corresponding to the distribution obtained from chaotic data. The initial trajectory (iteration 0) is shown in (a), (b), iteration 100 is shown in (c), (d), iteration 1000 is shown in (e), (f), and iteration 10000 is shown in (g), (h).

The data presented in this section is obtained from two sources. The first, denoted as "chaos" in the figures legends, is from chaotic simulations via black-box solvers supplied by the *solve_ivp* function from *SciPy* with an explicit Runge-Kutta 45 method with adaptive time-stepping, described in [61], with the output trajectory being uniformly sampled in time. The second source is from quasi-trajectories. These are initialised randomly,

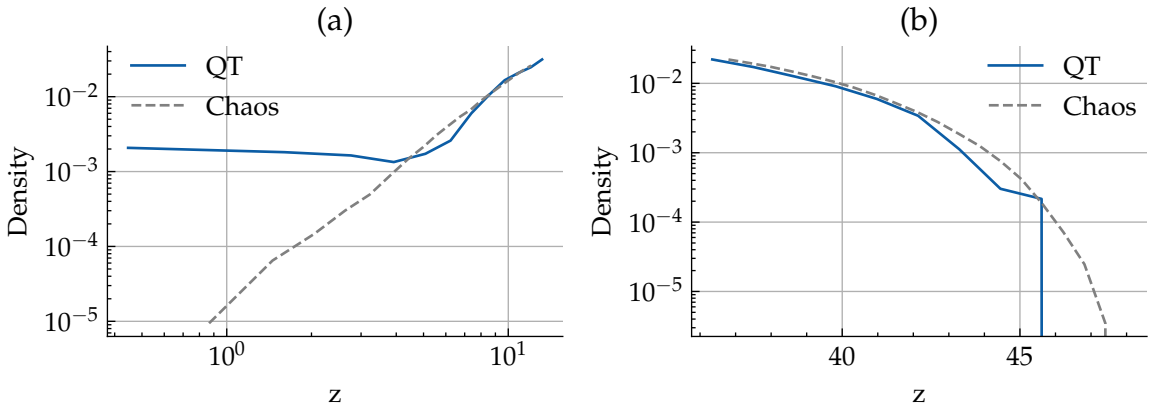

FIG. 6. Tails of the PDF shown in panel (h) of Figure 5, panel (a) near $z = 0$, and panel (b) near the extreme RHS of the distribution.

generating the points in the time domain around the given mean $\overline{\boldsymbol{x}}$ with a standard deviation of 10. The coefficients $\boldsymbol{a}_n$ are then determined from the time domain representation of the initial quasi-trajectory so that the optimisation can begin. The mean state used for the optimisation is set to $\overline{\boldsymbol{x}} = \begin{pmatrix} 0 & 0 & 23.64 \end{pmatrix}^\top$, obtained from independent chaotic simulations of the system using Equation 10.

Figure 5 displays the optimisation of a quasi-trajectory with a period of $T = 1000$ starting from a random distribution around the mean and how it evolves over 100, 1000, and 10000 iterations. On the left side the xz-projection of the quasi-trajectory sampled for 10000 points is shown, reconstructed in the time domain from the coefficients $\boldsymbol{a}_n$. On the right side the corresponding probability distribution function (PDF) over the z-direction, with the solid line representing the PDF obtained from the quasi-trajectory and the dashed line being obtained from a chaotic solution. We first note that there is a qualitative resemblance between the quasi-trajectory and what is expected from a chaotic simulation of the Lorenz system. The noted resemblance is achieved after only roughly 1000 iterations. This is a result of the optimisation seeking out the strange attractor very early, guiding the quasi-trajectory into a shape that lies on the attractor as well as possible for the given iteration. This is reinforced with the PDFs at each iteration, with the PDF shown in panels (f) and (h) of Figure 5 agreeing well with the PDF obtained from a chaotic solution. Both the quasi-trajectory and chaotic solutions display a bimodal distribution. The PDF of the quasi-trajectory at each iteration is computed using 40 bins over the range of z values obtained by the quasi-trajectory. The coarseness of the bins used is due to the fact that as the number of bins is increased the PDF would display peaks that do not subside as the number of bins is increased. The presence of these peaks in the PDFs of periodic orbits was observed and discussed in Ref. [62] and is a result of the turning points in the orbits.

There is a notable feature of the quasi-trajectory that is not present in chaotic trajectories. For the Lorenz system, there exists an unstable fixed point at the origin ($\boldsymbol{x} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$), which repels any trajectory that approaches it very quickly along its unstable manifold. However, as a consequence of the way in which the variational methodology is constructed, the residual is small around all fixed points regardless of their stability. This means that a quasi-trajectory is not heavily penalised for drifting away from the strange attractor towards the unstable fixed point at the origin. This can be observed in panels (g,h) of Figure 5, where there is a small increase in the density of the quasi-trajectory near the origin compared to the chaotic PDF that approaches zero as z goes to zero. In panel (a) of Figure 6 a slice of the PDF of panel (h) from Figure 5 is taken, plotted on a log-log scale, to show this trend more clearly. The chaotic PDF continues down as z decreases in the fashion of a power law, while the quasi-trajectory PDF plateaus at a particular distance from $z = 0$ after which it does not decrease any further. The effect on the statistics, however, is minimal. The quantitative effect of an unstable fixed point attracting quasi-trajectories under the variational optimisation is dependent on the stability characteristic of the particular fixed point. If the fixed point is close to neutrally stable, and therefore close to a bifurcation of some kind, it is difficult for the optimiser to converge towards it. The magnitude of attraction exerted on the optimiser is determined by the magnitude of the stability/instability of the point. There is also the possibility of so-called "ghost" points, discussed in Ref. [63], which are fixed points that have bifurcated and no longer exist in the state-space, corresponding to a local minimum for the optimiser. Despite these fixed points no longer existing, there are remnants of their presence on the state-space which can have an effect on the optimisation, especially as the optimisers convergence is strongly affected by near neutral structures. In fluid problems, where there may be many more unstable fixed points scattered around the state-space, the exact effect this would have on the op-

timisation is not known and is a planned future topic of work.

Panel (b) of Figure 6 shows the right tail of the PDF from panel (h) of Figure 5, showing the probability of the quasi-trajectory to undergo a particularly large loop around a lobe of the strange attractor. It shows that the quasi-trajectory is able to capture some of the more unlikely/extreme events of the chaotic motion. There is an upper limit to the PDF that is smaller than that predicted from the chaotic PDF, however a larger quasi-trajectory would lead to extreme events at larger values of z to be captured. With this in mind, extreme events can be important features in certain turbulent flows that display intermittency, and it is a topic for future work how well quasi-trajectories can capture them.

In Figure 7, the power spectra of the final quasi-trajectory achieved in Figure 5 at 100, 1000, and 10000 iterations is compared with that obtained from the chaotic data. All spectra are obtained using Welch's method, with Hann windowing to reduce spectral leakage. Welch's method was used to compute the average power spectra of the quasi-trajectory. This is done since the raw spectrum obtained from each optimisation is not deterministic due to the random initialisation, which results in power spectra with seemingly random values. Using Welch's method averages this randomness out and reveals the deterministic statistics achieved by a given quasi-trajectory. This is fundamentally the same reason Welch's method is used when computing the power spectrum of chaotic signals [64]. A similar trend is seen

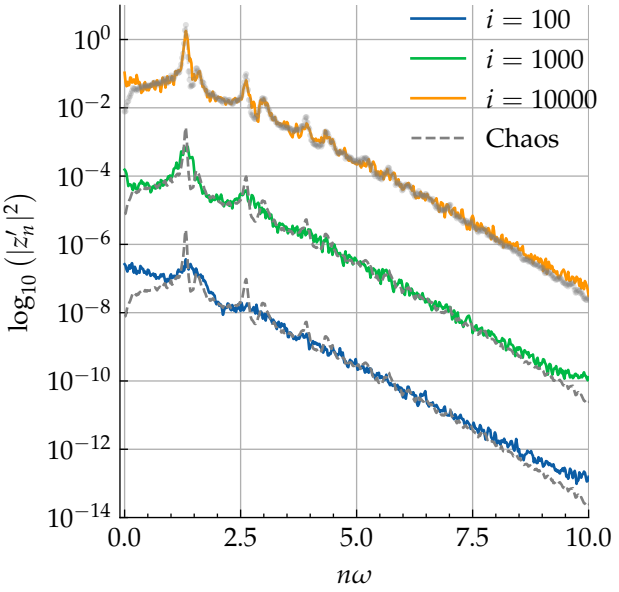

FIG. 7. Comparisons of the power spectra obtained from the final quasi-trajectory in Figure 5 at 100, 1000, and 10000 iterations, and chaotic data. The spectra at each iteration (along with the corresponding chaotic spectra) is plotted offset from each other to improve readability.

in Figure 5, in that the power spectra after 100 does not particularly resemble the chaotic spectra, missing the important peak at $n\omega \approx 1.3$ and its harmonics. After 1000 iterations the quasi-trajectory spectra displays a spread out version of this peak with a couple of its harmonics, and finally 10000 iterations shows the best agreement with the spectral peaks more clearly defined and multiple of its harmonics resolved in agreement with the chaotic data. The high frequency component of the spectra also gradually reduces over the duration of the optimisation, gradually approaching the spectral decay observed from the chaotic spectra. Notably, there is an increase in the spectral energy in the low frequencies that persists in the quasi-trajectory. This is possibly an artefact of the previously mentioned fixed point at the origin dragging part of the quasi-trajectory towards it. Near this point the quasi-trajectory moves rather slowly, approaching near a marginally unstable manifold, which adds an extra low frequency component to the spectra.

Panel (a) of Figure 8 the trace of the global residual, normalised by $\|\boldsymbol{g}(\boldsymbol{x}_f)\|^2$ for the final quasi-trajectory obtained over the duration of the optimisation of the same quasi-trajectory as in Figure 5. It can be seen that the global residual has not yet converged to either a zero $\mathcal{R} = 0$ or a non-zero minimum after 10000 iterations. Nevertheless, the close qualitative resemblance observed in Figure 5 is achieved after only a moderate number of iterations. On the right of Figure 8 are the period-averaged observables defined in Equation 44 computed on the quasi-trajectory at each iteration, plotted with horizontal lines corresponding to the values of the mean observables computed from a long chaotic trajectory obtained with a numerical integration of the equations of motion. The values of the period averaged observables over the quasi-trajectory approach the chaotic values, displaying a convergence of the statistics well before the residual itself has converged, with most of the improvement being done between 10 and 100 iterations. Thus, it is reasonable to say that this quasi-trajectory has converged to the point of providing useful approximations to the statistics of the chaotic dynamics at around 1000 iterations.

The local residual $\boldsymbol{r}$ in equation 5 can be viewed as a small perturbation imposed on the governing equations, and therefore a quasi-trajectory can be viewed as an exact solution to this slightly perturbed system. The ratio of the global residual to the norm of the system's right-hand side, $\mathcal{R}/\|\boldsymbol{g}(\boldsymbol{x})\|^2$, can then be viewed as a measure of the closeness of this forced system to original system. As such, these results show that nearby systems to the Lorenz system, or equivalently a lightly forced Lorenz system, have very similar statistics to each other.

Figure 9 shows the results of a number of batch optimisations at increasing periods T performed for 100, 1000, and 10000 iterations, with each batch consisting of 50 quasi-trajectories. For reference, the shortest UPO of the Lorenz equations has a period of about 1.55 time units. Shown on the top are the ensemble averages of the

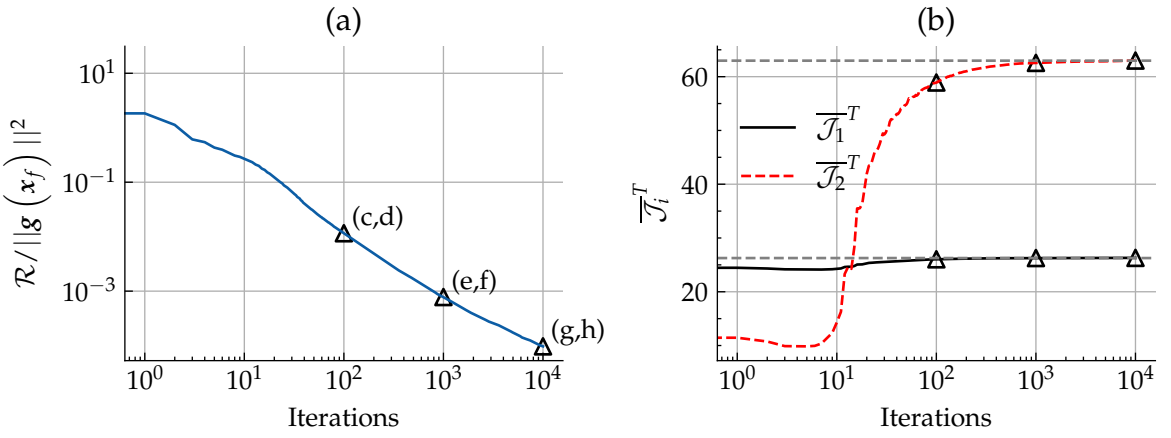

FIG. 8. Trace plots of the global residual normalised by the size of the system response $\boldsymbol{g}$ quasi-trajectory at the end of the optimisation $\boldsymbol{x}_f$ (a) and the mean observables over the quasi-trajectory at each iteration of the optimisation (b), shown with the values obtained from a long chaotic trajectory as horizontal grey lines. Annotated on the plots are the iterations corresponding to the reconstructions shown in Figure 5.

period averaged observables within the batch, denoted by $\left\langle \overline{\mathcal{J}_i}^T \right\rangle$, and in the bottom is the associated standard deviation of the period averaged observables within each batch, denoted with $\sigma\left(\overline{\mathcal{J}_i}^T\right)$. The corresponding values for the period averaged observables obtained from a long chaotic trajectory are shown with the dashed grey lines. The trend for 1000 and 10000 is for the period averaged observables to approach the long chaotic values. The values of the period averaged observables for 100 iterations exhibit poorer convergence towards the chaotic values, although the relative error is still rather small, being on the order of 1% and 6% for $\left\langle \overline{\mathcal{J}_1}^T \right\rangle$ and $\left\langle \overline{\mathcal{J}_2}^T \right\rangle$, respectively. The period averaged observables for the 10000 iterations case are close to the long chaotic value even for the shortest periods shown. Panels (c,d) illustrate the change in the standard deviation of the period averaged observables with the quasi-trajectory period exhibiting a steady decline as the period increases. The rate of this decrease is roughly proportional to the inverse square root of the period shown in Figure 9 with the grey dashed line, which is a consequence of the central limit theorem. The larger period therefore produces quasi-trajectories that become more similar from a statistical point of view. The standard error of the estimation of $\sigma\left(\overline{J_i}^T\right)$ ranges from roughly 10% to 0.1% between the time-averaged observables [65]. The result that the longer quasi-trajectories (for $T \gtrsim 20$) better reflect the statistics of the chaotic trajectories stem from their ability to explore the larger fractions of the strange attractor governing the chaotic dynamics. Therefore, a trade-off exists between the accuracy of the statistical predictions obtained and the speed at which the result can be achieved by varying the period of a quasi-trajectory. It should be noted that $\left\langle \overline{\mathcal{J}_1}^T \right\rangle$ approaches the chaotic value more closely for the optimisations that terminate at 1000 iterations, whereas $\left\langle \overline{\mathcal{J}_2}^T \right\rangle$ is closer to the chaotic value for the optimisations that terminate at 10000 iterations. This result suggests that certain observables may be most accurately captured during a quasi-trajectory optimisation.

Figure 10 is similar to Figure 9, but instead shows the standard deviation (panels (a) and (b)), skewness (panels (c) and (d)), and kurtosis (panels (e) and (f)) of the observables taken over the period T of a given quasi-trajectory, denoted with $\mathrm{std}_T(\cdot)$, $\mathrm{skew}_T(\cdot)$, and $\mathrm{kurt}_T(\cdot)$ respectively. These statistical moments are then averaged over the ensemble of 50 optimisations. The ensemble standard deviation follows the same descent trend as in Figure 9, so is omitted for the sake of compactness. The expected trend is that for the same period quasi-trajectory, the higher-order statistical moments will be less accurate. Figure 10 agrees with this, albeit with the estimation of the higher-order modes only degrading for the first observable, J_1. In fact, the quasi-trajectory statistics of J_2 appears to accurately reconstruct the statistical moments for periods $T \gtrsim 20$, implying the total distribution is faithfully reconstructed including the more extreme parts which would be emphasised by the kurtosis. The deviation of the moments of J_1 are difficult to directly attribute to a particular source, although it appears that more iterations leads to slightly better agreement with the value obtained from chaotic simulation. Additionally, the moments of J_2 appear to converge to a final value at a modest period, whereas the skewness and kurtosis of J_1 have seemingly not converged to the same degree. The optimisation then prioritises capturing certain aspects of the dynamics first, in this case balancing the nonlinear interactions over the total magnitude which takes larger periods and more iterations to accurately capture. In other words, the cross-correlations,

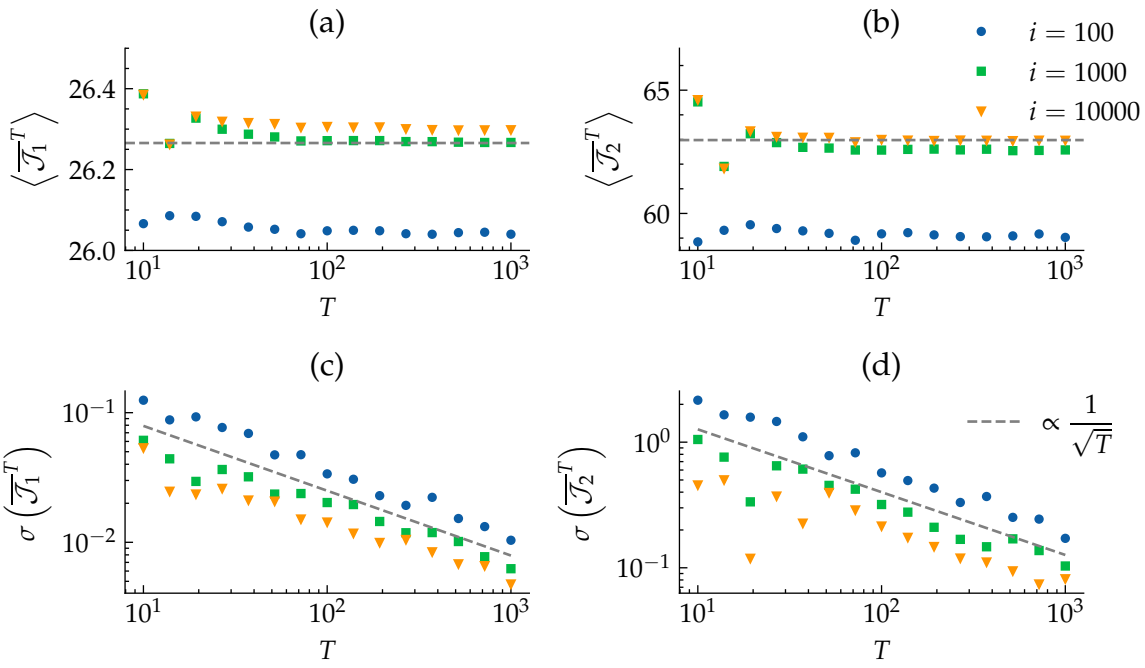

FIG. 9. The ensemble average and standard deviations of the period averaged observables over a batch optimisation of 50 quasi-trajectories, performed over a range of periods T. The optimisations were terminated at 100, 1000, and 10000 iterations. The left side (panels (a,c)) show the observable $\overline{\mathcal{J}_1}^T$, and the right side (panels (b,d)) show the observable $\overline{\mathcal{J}_2}^T$. The top (panels (a,b)) show the ensemble averages, and the bottom (panels (c,d)) show the standard deviations.

represented by the observable J_2, are very accurately captured by quasi-trajectories with a modest period and without having converged the global residual.

C. Time Cost of Computing Quasi-Trajectories

In this small test case for the Lorenz system, the computation of the quasi-trajectories is much heavier than using a simple ODE solver. There are two primary reasons for this. The first is that the variational optimisation inherently scales worse than a time-stepping approach to solving the governing equations since the whole temporal evolution of the trajectory is considered simultaneously. The second, and more pertinent reason for this work, is the low-dimensionality of the system not admitting a large dimensionality reduction. For the algorithm implemented here, depicted in Figure 3, each iteration of the optimisation is dominated by the computation of the FFTs in full-space and the projection/expansion steps between full- and reduced-space. The time complexities of each of these operations are given by $\mathcal{O}(N \log(N))$ and $\mathcal{O}(d_r N)$ respectively, where N are the degrees of freedom of the system (original dimension of the system multiplied by the temporal modes used), and d_r are the number of retained modes for the projection. For the variational optimiser to arrive at a sufficiently accurate estimate of the statistics more rapidly than a time-stepping method, it is required that the optimiser requires fewer degrees-of-freedom to accurately approximate the statistics, which can be best achieved through a large dimensionality reduction, i.e. $d_r \ll d$.

To illustrate for the Lorenz system, obtaining 1000 optimisation iterations of a quasi-trajectory with a period of $T = 20$ using 200 temporal modes ($N = 600$) with the dimensionality reduction $d_r = 2$ takes $\mathcal{O}(1)$ seconds. Obtaining a chaotic ODE solution of the same period using *solve_ivp* takes $\mathcal{O}(0.1)$ seconds. In short, for low-dimensional systems, where the degrees of freedom are already small and the dimensionality reduction is not large, computing quasi-trajectories is more expensive than an ODE computation.

The potential power of the method becomes more apparent when applied to high-dimensional systems, specifically fluid turbulence, where the possible dimensionality reduction can be very large. This property, combined with the possibly smaller required degrees of the freedom in the full-space, and potentially relaxed time-step constraints, could reduce the time taken to find a statistically meaningful solution. The method described here performs the majority of the computations in the full state-space, only projecting back to the reduced space for the update to the optimisation state. It is possible to perform a complete optimisation loop within the reduced-space by converting the pseudo-spectral approach to computing the nonlinear terms to convolutions in the reduced-space, resulting in projected nonlinear terms which are computed in terms of the coefficients $\boldsymbol{a}_n$. Although con-

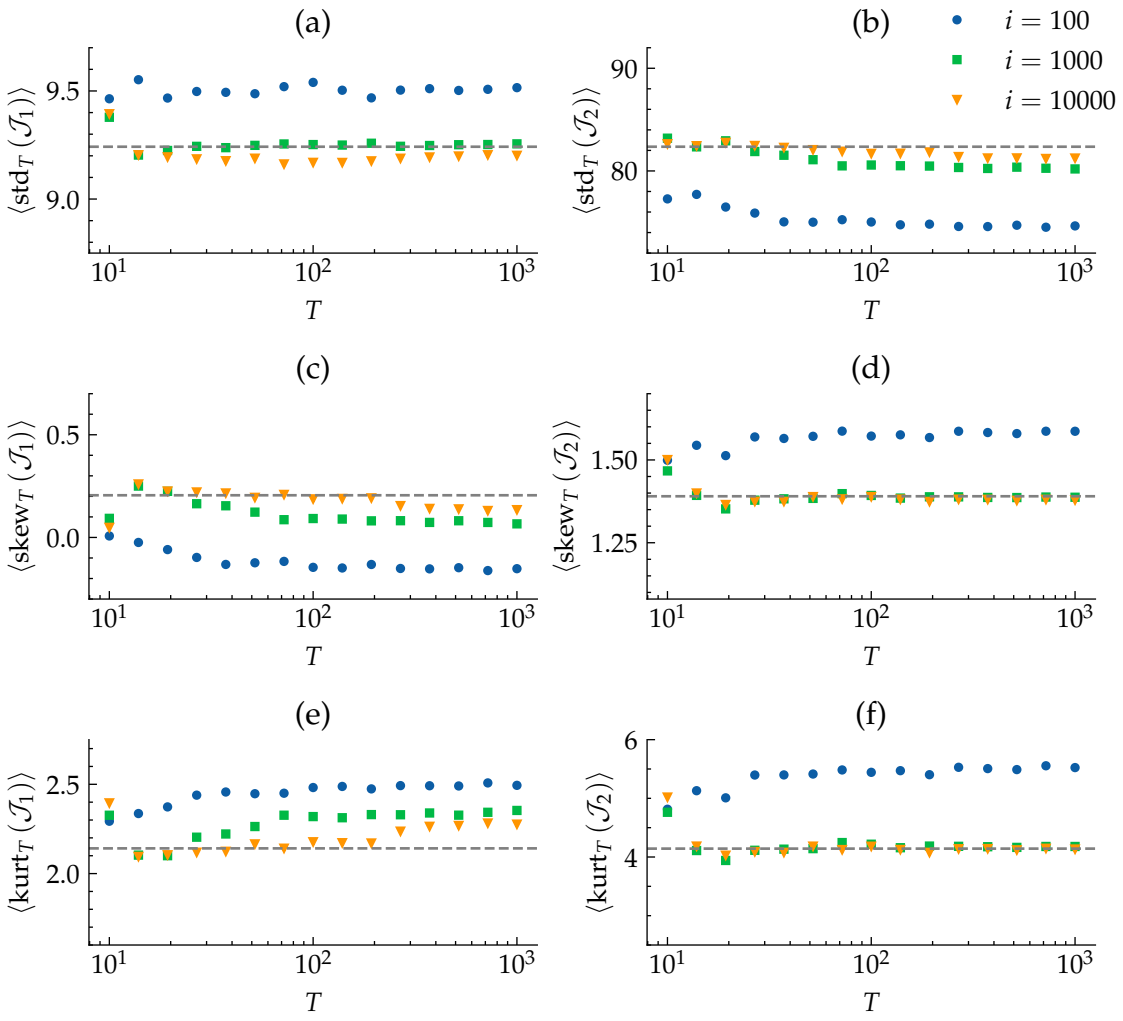

FIG. 10. The ensemble average of the standard deviation (panels (a) and (b)), skewness (panels (c) and (d)), and kurtosis (panels (e) and (f)) of the observables J_1 and J_2 over the same set of batch optimisation as in Figure 9.

volutions are known to scale worse with the size of the computation, the dimensionality reduction could actually make the convolutions in reduced-space faster than FFTs in full-space. To finish this discussion, the global-in-time description of the flow that is implicit in the current framework would also permit temporal parallelism to be exploited, which may lead to a further speed-up that is not possible in direct time-stepping methods used by most flow solvers.

IV. SUMMARY AND CONCLUSIONS

In this work we proposed a modelling framework combining variational methods for the search of UPOs with resolvent analysis for dimensionality reduction. The fundamental idea is to construct candidate trajectories from a low-dimensional expansion of space-time basis functions with initially unknown amplitude coefficients. These coefficients are then found by optimising an objective function describing the overall violation of the governing equations along the trajectory. Using a reduced set of resolvent modes as a basis, where only the most receptive response modes (associated with the largest singular values of the resolvent matrix), the temporal periodicity of UPOs is built into the expansion. For the general case, the projection onto such a set of modes forces the resulting trajectories to live in a lower-dimensional space that rejects some of the dynamical information present in full state space. In addition, in using resolvent analysis the mean $\overline{\boldsymbol{x}}$ and period T of the sought solution is fixed during the course of the optimisation. Thus, in general the governing equations can only be approximately solved. The Lorenz system, however, is a special case that allows for an exact projection of the 3-dimensional dynamics onto a 2-dimensional subspace that fully captures the dynamics. It was shown that no further truncation of the modes was possible while ensuring any accu-

racy of the resulting solutions, due to the two remaining singular values of the resolvent not having a large enough scale separation. Due to the truncation, solving this optimisation problem, may be computationally cheaper than identifying exact UPOs of the original system, especially for high-dimensional fluid systems. Despite this, these solutions, dubbed quasi-trajectories, may provide an adequate "sketch" of the attractor in state space and may thus have the ability to approximate to a sufficient degree of accuracy the statistical properties of the original system.

This framework is demonstrated on the Lorenz 1963 system, chosen as a computationally accessible test-bed for developing and testing the numerical techniques. For such system, resolvent analysis provides a natural dimensionality reduction that can be exploited to reduce the degrees of freedom of the system from three to two. Observables averaged over the period of a quasi-trajectory are monitored as the optimisation progresses. One key finding of this study is that such observables approach the values obtained from long chaotic trajectories rather quickly, after only 100 - 1000 iterations, with a great robustness to the initial condition used. This suggests that the governing equations projected on the low-dimensional subspace need not be solved exactly to obtain close approximations of the statistical properties of the original system. Instead, the gradient-based iteration procedure "adjusts" a candidate state space loop relatively quickly, pushing it towards the region of state space occupied by the attractor. The accuracy of the period averaged observables obtained from a quasi-trajectory increases with its period, with the variance in the value obtained from a given quasi-trajectory decreasing at the same time. This shows that longer quasi-trajectories better approximate the chaotic statistics, as a result of a larger portion of the attractor being explored.

The application of this framework to fluid systems remains to be explored. The dimensionality reduction afforded by resolvent analysis is much more pronounced in many flows of practical interest, compare to what can be obtained here for the Lorenz system. Here the dimensionality achieved is exact, losing no information in the projection step. This is not true for the fluid dynamics case since the mode truncation is generally motivated by a sufficiently large separation of scales between resolvent singular values. The expectation is that this truncation will not have a significant effect on the efficacy of the model, which is still required to be shown. The modal truncation could lead to large computational savings, relatively speaking, although its role on the quality of the approximation needs to be assessed, because there is no guarantee that the leading resolvent modes can capture the majority of the kinetic energy. This could result in approximations of the velocity field that miss important features. A second important task is to ascertain the impact of the attractor dimension on the convergence rate of observables computed from quasi-trajectories to the long-time statistics computed from chaotic trajectories. As shown, few iterations were sufficient here to obtain relatively good approximations of averages and probability distributions, but in a multi-scale problem such as turbulence the convergence may be slower. An answer to such question would then provide insight into how computational costs required to obtain estimates of the statistical properties depend on the problem dimension, as computational costs scale linearly with the number of iterations performed. These aspects are currently being considered and will be reported in future work.

ACKNOWLEDGMENTS

This paper is dedicated to the memory of Bruno Eckhardt. The work presented here was stimulated by discussions between B.E. and A.S. at the *Recurrence, Self-Organization, and the Dynamics Of Turbulence* meeting, held at the Kavli Institute of Theoretical Physics, University of California, Santa Barbara in January 2017.

Appendix A: Residual Gradient

The gradient of the global residual $\mathcal{R}$ with respect to the state space loop $\boldsymbol{x}$ and its fundamental frequency ω is given in Equation 8 and Equation 9. Here these expressions are derived using variational calculus. In addition, the proof of the projection of the gradients onto the reduced subspace is provided.

1. Full-Space Derivation

We will begin with Equation 8. Let's, consider some perturbation to the state space loop $\boldsymbol{x} \to \boldsymbol{x} + \epsilon\boldsymbol{\delta x}$, such that $\boldsymbol{\delta x}|_{t=0} = \boldsymbol{\delta x}|_{t=T}$ and $\epsilon > 0$ is some small real number. This perturbation propagates through all the variables that depend on the shape of the trajectory. The resulting perturbation in the local residual can be expressed as

$$\begin{aligned}\boldsymbol{r}\left(\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right) &= \boldsymbol{r} + \boldsymbol{\delta r} \\ &= \omega\frac{\mathrm{d}}{\mathrm{d}s}\left(\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right) - \boldsymbol{g}\left(\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right) \\ &= \omega\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s} - \boldsymbol{g}\left(\boldsymbol{x}\right) + \\ &\qquad \epsilon\left(\omega\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\delta x} - \boldsymbol{L}\left(\boldsymbol{x}\right)\boldsymbol{\delta x}\right),\end{aligned} \tag{A1}$$

which when rearranged gives

$$\boldsymbol{\delta r} = \epsilon\left(\omega\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\delta x} - \boldsymbol{L}\left(\boldsymbol{x}\right)\boldsymbol{\delta x}\right), \tag{A2}$$

where $\boldsymbol{L}$ is the Jacobian matrix of the vector-valued function $\boldsymbol{g}$ with respect to the state space variables $\boldsymbol{x}$. The

perturbation of the global residual is given as

$$\begin{aligned}\mathcal{R}\left[\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right] &= \frac{1}{2}\|\boldsymbol{r}+\boldsymbol{\delta r}\|^2 \\ &= \frac{1}{2}\|\boldsymbol{r}\|^2 + \langle \boldsymbol{r},\, \boldsymbol{\delta r}\rangle + \frac{1}{2}\|\boldsymbol{\delta r}\|^2 \\ &= \frac{1}{2}\|\boldsymbol{r}\|^2 + \\ &\quad \epsilon\left\langle \boldsymbol{r},\, \omega\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\delta x} - \boldsymbol{L}\left(\boldsymbol{x}\right)\boldsymbol{\delta x}\right\rangle + \\ &\quad \frac{\epsilon^2}{2}\left\|\omega\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\delta x} - \boldsymbol{L}\left(\boldsymbol{x}\right)\boldsymbol{\delta x}\right\|^2.\end{aligned} \tag{A3}$$

Now, variational calculus provides the following identities

$$\left[\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{R}\left[\boldsymbol{x}+\epsilon\delta\boldsymbol{x}\right]\right]_{\epsilon=0} = \left\langle \frac{\delta\mathcal{R}}{\boldsymbol{\delta x}},\, \boldsymbol{\delta x}\right\rangle, \tag{A4}$$

Applying Equation A4 to the last equality of Equation A3 the following is obtained

$$\left\langle \boldsymbol{r},\, \omega\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{\delta x} - \boldsymbol{L}\left(\boldsymbol{x}\right)\boldsymbol{\delta x}\right\rangle = \left\langle \frac{\delta\mathcal{R}}{\boldsymbol{\delta x}},\, \boldsymbol{\delta x}\right\rangle. \tag{A5}$$

Thus, to obtain a closed-form expression for $\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}}$ it is necessary to rearrange the left-hand side of Equation A5 such that it resembles the form of the right-hand side. Leveraging the bi-linearity of the inner-product, and performing integration by parts of the time derivative (noting that the boundary term disappears due to the periodicity of $\boldsymbol{\delta x}$) we get

$$\left\langle -\omega\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} - \boldsymbol{L}^\top\left(\boldsymbol{x}\right)\boldsymbol{r},\, \boldsymbol{\delta x}\right\rangle = \left\langle \frac{\delta\mathcal{R}}{\boldsymbol{\delta x}},\, \boldsymbol{\delta x}\right\rangle, \tag{A6}$$

which, because $\boldsymbol{\delta x}$ is free to be any function we want, implies that

$$\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}} = -\omega\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} - \boldsymbol{L}^\top\left(\boldsymbol{x}\right)\boldsymbol{r}. \tag{A7}$$

Next, an expression for $\partial\mathcal{R}/\partial\omega$ can be obtained using standard calculus since ω is just a real number rather than a function like $\boldsymbol{x}$. Taking the definition of the global residual Equation 6, substituting in Equation 5, and then rearranging to make ω the subject, gives the following

$$\begin{aligned}\mathcal{R} = \frac{1}{2}\|\boldsymbol{r}\|^2 &= \frac{1}{2}\left\|\omega\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s} - \boldsymbol{g}\left(\boldsymbol{x}\right)\right\|^2 \\ &= \frac{\omega^2}{2}\left\|\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s}\right\|^2 - \omega\left\langle \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s},\, \boldsymbol{g}\left(\boldsymbol{x}\right)\right\rangle + \frac{1}{2}\|\boldsymbol{g}\left(\boldsymbol{x}\right)\|^2.\end{aligned} \tag{A8}$$

Taking the derivative of this expression with respect to ω results in

$$\frac{\partial\mathcal{R}}{\partial\omega} = \omega\left\|\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s}\right\|^2 - \left\langle \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s},\, \boldsymbol{g}\left(\boldsymbol{x}\right)\right\rangle, \tag{A9}$$

which can be rearranged to a simpler form

$$\frac{\partial\mathcal{R}}{\partial\omega} = \left\langle \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s},\, \boldsymbol{r}\right\rangle. \tag{A10}$$

2. Projection onto the Resolvent Subspace

The projection onto the reduced subspace is performed in two steps: first the expansion in terms of a Fourier series, and then a least-squares projection onto a subspace defined by the set of response modes obtained from resolvent analysis. First, to prove the expression given in Equation 29, we express the change in the global residual due to a change in the Fourier coefficients of $\boldsymbol{x}$ as

$$\delta\mathcal{R} = \sum_{n\in\mathbb{Z}} \boldsymbol{\delta x}_n^\dagger \frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}. \tag{A11}$$

Thus, the gradient of $\mathcal{R}$ is expressed as

$$\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}} = \sum_{n\in\mathbb{Z}} \frac{\delta\boldsymbol{x}_n}{\delta\boldsymbol{x}}^\dagger \frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}. \tag{A12}$$

To obtain an expression for the functional derivative of the Fourier coefficients $\boldsymbol{x}_n$ with respect to the trajectory $\boldsymbol{x}$, we use the definition Equation 14 and introduce a perturbation $\epsilon\boldsymbol{\delta x}$

$$\begin{aligned}\boldsymbol{x}_n\left(\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right) &= \frac{1}{2\pi}\int_0^{2\pi}\left(\boldsymbol{x}+\epsilon\boldsymbol{\delta x}\right)e^{-ins}\,\mathrm{d}s \\ &= \boldsymbol{x}_n + \frac{\epsilon}{2\pi}\int_0^{2\pi}\boldsymbol{\delta x}e^{-ins}\,\mathrm{d}s.\end{aligned} \tag{A13}$$

Applying Equation A4 to the above gives the following

$$\frac{\delta\boldsymbol{x}_n}{\delta\boldsymbol{x}} = \boldsymbol{I}e^{-ins}. \tag{A14}$$

Substituting this into Equation A12 gives

$$\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}} = \sum_{n\in\mathbb{Z}} \frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}e^{ins}, \tag{A15}$$

which when the identity Equation 14 is applied to provides

$$\frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}}e^{-ins}\,\mathrm{d}s = \left(\frac{\delta\mathcal{R}}{\delta\boldsymbol{x}}\right)_n. \tag{A16}$$

The proof for Equation 30 is quite similar, if a little simpler. Using the chain rule, we have

$$\frac{\partial\mathcal{R}}{\partial\boldsymbol{a}_n} = \frac{\partial\boldsymbol{x}_n}{\partial\boldsymbol{a}_n}^\dagger \frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}. \tag{A17}$$

Using the definition Equation 21, the following derivative can be obtained

$$\frac{\partial\boldsymbol{x}_n}{\partial\boldsymbol{a}_n} = \boldsymbol{\Psi}_n, \quad n\in\mathbb{Z}\setminus\{0\}. \tag{A18}$$

which when substituted into Equation A17 gives the final result

$$\frac{\partial\mathcal{R}}{\partial\boldsymbol{a}_n} = \boldsymbol{\Psi}_n^\dagger \frac{\partial\mathcal{R}}{\partial\boldsymbol{x}_n}, \quad n\in\mathbb{Z}\setminus\{0\}. \tag{A19}$$

[1] D. Ruelle and F. Takens, Communications in Mathematical Physics **20**, 167 (1971).
[2] Y. Aizawa, Progress of Theoretical Physics **68**, 64 (1982).
[3] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[4] D. Auerbach, P. Cvitanovic, J. P. Eckmann, G. Gunaratne, and I. Procaccia, Phys. Rev. Lett. **58**, 2387 (1987).
[5] P. Cvitanović, Physical Review Letters **61**, 2729 (1988).
[6] P. Cvitanović, Physica D: Nonlinear Phenomena **83**, 109 (1995).
[7] M. Nagata, Journal of Fluid Mechanics **217**, 519 (1990).
[8] F. Waleffe, Journal of Fluid Mechanics **435**, 93 (2001).
[9] F. Waleffe, in *Turbulence and Interactions*, edited by M. Deville, T.-H. Lê, and P. Sagaut (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009) pp. 139–158.
[10] G. Kawahara and S. Kida, Journal of Fluid Mechanics **449**, 291 (2001).
[11] D. Viswanath, Journal of Fluid Mechanics **580**, 339 (2007).
[12] G. J. Chandler and R. R. Kerswell, Journal of Fluid Mechanics **722**, 554 (2013).
[13] N. B. Budanur, K. Y. Short, M. Farazmand, A. P. Willis, and P. Cvitanović, Journal of Fluid Mechanics **833**, 274 (2017).
[14] G. Kawahara, M. Uhlmann, and L. V. Veen, Annual Review of Fluid Mechanics **44**, 203 (2012).
[15] M. D. Graham and D. Floryan, Annual Review of Fluid Mechanics **53**, 227 (2021).
[16] B. Suri, L. Kageorge, R. O. Grigoriev, and M. F. Schatz, Physical Review Letters **125**, 64501 (2020).
[17] M. C. Krygier, J. L. Pughe-sanford, and R. O. Grigoriev, Journal of Fluid Mechanics **923**, A7 (2021), arXiv:arXiv:2105.02126v1.
[18] C. J. Crowley, J. L. Pughe-Sanford, W. Toler, M. C. Krygier, R. O. Grigoriev, and M. F. Schatz, Proceedings of the National Academy of Sciences **119**, e2120665119 (2022).
[19] J. Page, P. Norgaard, M. P. Brenner, and R. R. Kerswell, Proceedings of the National Academy of Sciences **121**, e2320007121 (2024).
[20] R. L. Davidchack and Y. C. Lai, Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics **60**, 6172 (1999), arXiv:9906007 [chao-dyn].
[21] F. K. Diakonos, P. Schmelcher, and O. Biham, Phys. Rev. Lett. **81**, 4349 (1998).
[22] Y. Lan and P. Cvitanović, Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics **69**, 10 (2004), arXiv:0308008 [nlin].
[23] L. Fazendeiro, B. M. Boghosian, P. V. Coveney, and J. Lätt, Journal of Computational Science **1**, 13 (2010).
[24] S. Azimi, O. Ashtari, and T. M. Schneider, Phys. Rev. E **105**, 14217 (2022).
[25] O. Ashtari and T. M. Schneider, Journal of Fluid Mechanics **977**, A7 (2023).
[26] J. Page and R. R. Kerswell, Journal of Fluid Mechanics **886**, A28 (2020).
[27] D. Lasagna, Phys. Rev. E **102**, 052220 (2020), arXiv:arXiv:1910.06706v3.
[28] Y. Saiki and M. Yamada, Phys. Rev. E **79**, 15201 (2009).
[29] D. Lucas and R. R. Kerswell, Physics of Fluids **27**, 45106 (2015).
[30] Y. Hwang and H. Choi, Physics of Fluids **20**, 071703 (2008).
[31] F. Giannetti, S. Camarri, and P. Luchini, Journal of Fluid Mechanics **651**, 319 (2010).
[32] J. Sierra, P. Jolivet, F. Giannetti, and V. Citro, Journal of Computational Physics **440**, 110403 (2021).
[33] D. Lasagna, SIAM Journal on Applied Dynamical Systems **17**, 547 (2018).
[34] P. Meliga, E. Boujo, and F. Gallaire, Journal of Fluid Mechanics **800**, 327 (2016).
[35] F. Giannetti, S. Camarri, and V. Citro, Journal of Fluid Mechanics **864**, 45 (2019).
[36] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn, Meccanica **15**, 9 (1980).
[37] M. McCormack, A. V. G. Cavalieri, and Y. Hwang, Journal of Fluid Mechanics **983**, A33 (2024).
[38] C. Foias, G. R. Sell, and R. Temam, Journal of Differential Equations **73**, 309 (1988).
[39] B. J. McKeon and A. S. Sharma, Journal of Fluid Mechanics **658**, 336 (2010), arXiv:arXiv:1001.3100v1.
[40] F. Gómez, H. M. Blackburn, M. Rudman, A. S. Sharma, and B. J. McKeon, Journal of Fluid Mechanics **798**, R2 (2016), arXiv:1606.04735.
[41] S. Beneddine, R. Yegavian, D. Sipp, and B. Leclaire, Journal of Fluid Mechanics **824**, 174 (2017).
[42] B. Jin and S. J. Illingworth, Theoretical and Computational Fluid Dynamics **36**, 491 (2021), arXiv:arXiv:2105.04927v1.
[43] D. F. Gayme and B. A. Minnick, Phys. Rev. Fluids **4**, 110505 (2019).
[44] X. Garnaud, L. Lesshafft, P. J. Schmid, and P. Huerre, Journal of Fluid Mechanics **716**, 189 (2013).
[45] D. Gayme, *A robust control approach to understanding nonlinear mechanisms in shear flow turbulence*, Ph.D. thesis, California Institute of Technology (2010).
[46] S. Symon, K. Rosenberg, S. T. M. Dawson, and B. J. McKeon, Phys. Rev. Fluids **3**, 053902 (2018), arXiv:1712.05473.
[47] A. S. Sharma, R. Moarref, B. J. Mckeon, J. S. Park, M. D. Graham, and A. P. Willis, Phys. Rev. E **93**, 021102 (2016).
[48] A. S. Sharma, I. Mezić, and B. J. McKeon, Physical Review Fluids **1**, 32402 (2016).
[49] M. Farazmand, Journal of Fluid Mechanics **795**, 278 (2016).
[50] B. Barthel, X. Zhu, and B. J. McKeon, Journal of Fluid Mechanics **924**, A9 (2021), arXiv:2102.09141.
[51] M. A. Ahmed and A. S. Sharma, Physical Review E **101**, 12213 (2020).
[52] E. Lorenz, Journal of the Atmospheric Sciences **20**, 130 (1963).
[53] A. S. Sharma and B. J. McKeon, Journal of Fluid Mechanics **728**, 196 (2013), arXiv:1301.7580.
[54] B. J. McKeon, A. S. Sharma, and I. Jacobi, Physics of Fluids **25**, 31301 (2013).
[55] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. (Springer-Verlag New York, 2006).
[56] D. C. Liu and J. Nocedal, Mathematical Programming **45**, 503 (1989).

[57] J. L. Morales and J. Nocedal, ACM Trans. Math. Softw. **38**, 10.1145/2049662.2049669 (2011).
[58] J. Otero, *Development and application of an adjoint-based optimal flow control framework for compressible direct numerical simulations*, Ph.D. thesis, University of Southampton (2017).
[59] H. Badreddine, S. Vandewalle, and J. Meyers, Journal of Computational Physics **256**, 1 (2014).
[60] R. Fletcher, in *Practical Methods of Optimization* (John Wiley & Sons, Ltd, 2000) Chap. 2, pp. 12–43.
[61] J. R. Dormand and P. J. Prince, Journal of Computational and Applied Mathematics **6**, 19 (1980).
[62] S. M. Zoldi, Phys. Rev. Lett. **81**, 3375 (1998).
[63] Z. Zheng, P. Beck, T. Yang, O. Ashtari, J. P. Parker, and T. M. Schneider, arXiv preprint arXiv:2411.10320 (2024), arXiv:2411.10320.
[64] P. L. Andrews and R. E. Waltz, The Physics of Fluids **31**, 3168 (1988).
[65] E. L. Lehmann and G. Casella, *Theory of point estimation* (Springer Science & Business Media, 2006).