

# Adjoint-based optimization with quantized local reduced-order models for spatiotemporally chaotic systems

Defne Ege Ozan<sup>1</sup>[0009–0006–4208–9981], Antonio Colanera<sup>2</sup>[0000–0002–4227–1162],  
and Luca Magri<sup>1,3</sup>[0000–0002–0657–2611]

<sup>1</sup> Department of Aeronautics, Imperial College London, SW7 2AZ, London, UK  
[l.magri@imperial.ac.uk](mailto:l.magri@imperial.ac.uk)

<sup>2</sup> International School for Advanced Studies (SISSA),  
Via Bonomea 265, Trieste 34136, Italy

<sup>3</sup> Politecnico di Torino, DIMEAS, Corso Duca degli Abruzzi, 24 10129 Torino, Italy

**Abstract.** We introduce a computationally efficient and accurate reduced-order modelling approach for the optimization of spatiotemporally chaotic systems. The proposed method combines quantized local reduced-order modelling with adjoint-based optimization. We employ the methodology in a variational data assimilation problem for the chaotic Kuramoto-Sivashinsky equation and show that it successfully reconstructs the full trajectory for up to 0.25 Lyapunov times given full state measurements at the final time. The proposed algorithm provides  $\times 3.5$  speed-up when compared to the full-order model. The proposed method opens up new possibilities for the reduced-order modelling of spatiotemporally chaotic systems.

**Keywords:** Reduced-order modelling · Adjoint methods · Data assimilation · Spatiotemporal chaos.

## 1 Introduction

Fluid systems that are of scientific and engineering interest are typically described by partial differential equations (PDEs), which often have spatiotemporally chaotic solutions [7]. Accurate solutions of these PDEs typically require fine discretizations that yield high-dimensional state representations in order to resolve the underlying multi-scale dynamics. The resulting high computational cost for the simulation of these systems motivates the development of reduced-order modelling (ROM) approaches, which aim to discover low-dimensional models that can be utilized to predict the system behaviour with some approximation error [13]. In dissipative systems, which are common in fluids, the ROM approach becomes viable, as the trajectories are attracted to a low-dimensional manifold.

ROMs are especially appealing in gradient-based optimization, where repeated forward simulations are required to reach convergence. Adjoint methods are central in this context. The adjoint formulation enables the computation of

gradients of an objective functional with respect to many design parameters at a cost that is essentially independent of the number of parameters e.g., [5, 10, 12], unlike tangent-linear or finite-difference approaches whose cost scales with the parameter dimension. In data assimilation, an alternative class of approaches is based on sequential filtering, most notably the ensemble Kalman filter (EnKF); a general EnKF formalism is presented in [11].

In this paper, we leverage the recently developed quantized local reduced-order models (ql-ROMs) [2] for gradient-based optimization, which overcomes the challenge of designing a single global ROM for the entire state space. In this approach, the manifold is quantized via unsupervised clustering, and intrusive local ROMs are built for each cluster. The switching between the local models is enabled via a change of basis using assignment functions. We derive the adjoint of the ql-ROM, where a similar coordinate transformation exists between the local adjoint variables. We demonstrate the application of this method in a variational data assimilation problem of a prototypical spatiotemporally chaotic system, where the objective is to infer the initial condition of a trajectory given observations.

The remainder of the paper is organized as follows. In Section 2, we present the proposed methodology. Specifically, Section 2.1 summarizes the ql-ROM algorithm, Section 2.2 derives the adjoint of the ql-ROM and Section 2.3 embeds it within a gradient-based optimization framework for variational data assimilation problems. The methodology is demonstrated on the spatiotemporally chaotic solutions of the Kuramoto-Sivashinsky (KS) equation in Section 3. Section 4 concludes the paper.

## 2 Methodology

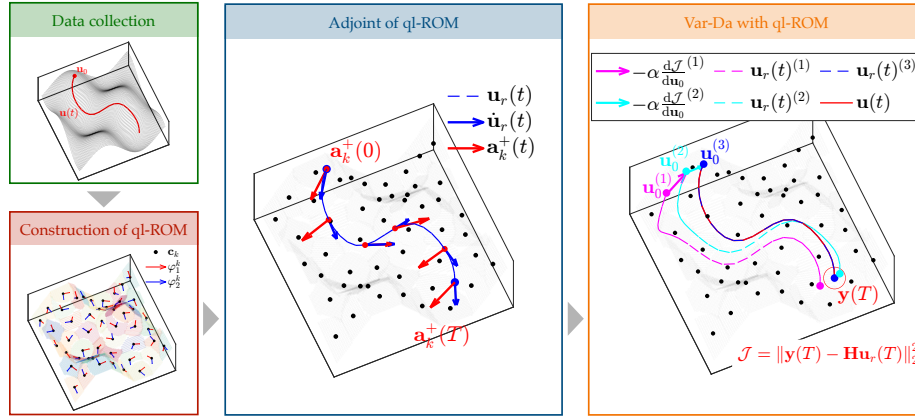
In this section, we outline the proposed workflow to enable adjoint-based optimization with quantized local reduced-order models (ql-ROMs). The approach consists of three building blocks: the construction of local intrusive ROMs on a quantized state space, the derivation of the corresponding adjoint with appropriate jump/coordinate transformations at cluster switches, and the integration of this adjoint within a variational data assimilation loop to update the initial condition. The overall workflow is shown in Figure 1.

### 2.1 Quantized reduced-order models (ql-ROMs)

We consider the spatial discretization of a PDE which results in the following system of ordinary differential equations

$$\dot{\mathbf{u}} = \mathbf{f}(\mathbf{u}), \quad \mathbf{u} \in \mathbb{R}^N, \quad \mathbf{u}(t = 0) = \mathbf{u}_0 \quad (1)$$

where  $\mathbf{u}$  is the state vector,  $\dot{\mathbf{u}} = d\mathbf{u}/dt$  is its time derivative,  $\mathbf{f}$  is the nonlinear operator that describes the evolution of the state vector in time, and  $N$  the dimension of the full-order state. Given a dataset of  $M$  snapshots  $\{\mathbf{u}(t_m)\}_{m=0}^M$ ,



**Fig. 1.** Overview of adjoint-based optimization with quantized local reduced-order models (ql-ROMs): (1) data collection; (2) phase-space quantization and local basis/model construction; (3) ql-ROM adjoint integrated backward in time with jump/coordinate-transformation at cluster switches; (4) variational data assimilation using the ql-ROM direct-adjoint loop to update the initial condition.

where  $t_m = m\Delta t$  with  $\Delta t$  being the time step, the method first partitions this data into  $K$  clusters with centroids  $\mathbf{c}_k \in \mathbb{R}^N$ ,  $k = 1, 2, \dots, K$ . Here, we employ K-means for this purpose, which is an unsupervised machine learning method for clustering data based on a distance metric, e.g., the Euclidean distance from the centroid. The cluster affiliation function then assigns a snapshot to the closest cluster using this metric, where the centroid acts as the mean of the cluster. In the next step, we project  $\mathbf{u}$  on a basis defined by a centroid  $\mathbf{c}_k$  and orthonormal basis vectors  $\mathbf{V}_k \in \mathbb{R}^{N \times r_k}$  such that

$$\mathbf{u} = \mathbf{c}_k + \mathbf{V}_k \mathbf{a}_k, \quad (2)$$

where  $\mathbf{a}_k \in \mathbb{R}^{r_k}$  are the coordinates in this new basis. The objective of reduced-order modelling is to build an accurate model with  $r_k \ll N$ . (When we set the number of clusters  $K = 1$ , this formulation becomes equivalent to the global modelling approach.) We determine

$$\mathbf{a}_k = \mathbf{V}_k^\top (\mathbf{u} - \mathbf{c}_k). \quad (3)$$

Here, we employ proper orthogonal decomposition (POD) [7] to obtain the local bases  $\mathbf{V}_k$  in each cluster. Using Galerkin projection of the dynamics onto the chosen basis vector, we obtain a model for the time evolution of  $\mathbf{a}_k$  for each cluster

$$\dot{\mathbf{a}}_k = \mathbf{g}_k(\mathbf{a}_k), \quad \mathbf{a}_k \in \mathbb{R}^{r_k}, \quad \mathbf{a}_k(0) = \mathbf{V}_k^\top (\mathbf{u}_0 - \mathbf{c}_k). \quad (4)$$

The advantage of the ql-ROM approach is that the dynamics are modelled locally, which means that as the trajectory moves along the state space, the cluster

affiliations change. The coordinate transformation for switch from cluster  $i$  to cluster  $j$  is given by

$$\mathbf{a}_j(t) = \mathbf{V}_j^\top \mathbf{V}_i \mathbf{a}_i(t) + \mathbf{V}_j^\top (\mathbf{c}_i - \mathbf{c}_j). \quad (5)$$

For further details about the algorithm, we refer the reader to [2, 3].

## 2.2 Adjoint of ql-ROM

Given a model of the system's dynamics, adjoint-based optimization relies on a direct-adjoint loop: the system is integrated forward in time, the dynamics are linearized along the resulting trajectory, and the corresponding adjoint equations are integrated backward to evaluate gradients e.g., [6, 10].

Consider a scalar objective function

$$\mathcal{J} = \tilde{\mathcal{J}}(\mathbf{u}(T)), \quad (6)$$

and the ql-ROM dynamics (4). Our goal is to compute the sensitivity of the objective function to the initial condition, i.e., to compute the gradient  $d\mathcal{J}/d\mathbf{u}_0$ . Since  $\mathbf{a}_0 = \mathbf{V}_{k(0)}^\top (\mathbf{u}_0 - \mathbf{c}_{k(0)})$ , it is sufficient to compute  $d\mathcal{J}/d\mathbf{a}_0$  and then map it to  $d\mathcal{J}/d\mathbf{u}_0$  by a constant transformation. We define the Lagrangian

$$\mathcal{L} = \mathcal{J} - \langle \mathbf{a}^+, \dot{\mathbf{a}} - \mathbf{g}(\mathbf{a}) \rangle = \tilde{\mathcal{J}}(\mathbf{a}(T)) - \int_0^T \mathbf{a}^{+\top} (\dot{\mathbf{a}} - \mathbf{g}(\mathbf{a})) dt, \quad (7)$$

where  $\mathbf{a}^+$  are the adjoint variables. Stationarity of  $\mathcal{L}$  with respect to  $\mathbf{a}_0$  yields, within each active cluster  $k$ ,

$$\dot{\mathbf{a}}_k^+ = - \left( \frac{d\mathbf{g}_k(\mathbf{a}_k)}{d\mathbf{a}_k} \right)^\top \mathbf{a}_k^+, \quad (8)$$

with terminal condition  $\mathbf{a}^+(T) = d\tilde{\mathcal{J}}/d\mathbf{a}(T)$ . The term  $d\mathbf{g}_k(\mathbf{a}_k)/d\mathbf{a}_k \in \mathbb{R}^{r_k \times r_k}$  in Eq. (8) is the Jacobian of the local reduced-order model.

When the direct trajectory switches from cluster  $i$  to cluster  $j$  at time  $T_s$ , the adjoint satisfies the corresponding jump/coordinate transformation

$$\mathbf{a}_i^+(T_s^-) = \mathbf{V}_i^\top \mathbf{V}_j \mathbf{a}_j^+(T_s^+), \quad (9)$$

consistent with the direct coordinate change (5). This transformation follows from splitting the integral in Eq. (7) at the switching time, and collecting the boundary terms resulting from integration by parts of each subinterval. The reduced gradient follows as

$$\frac{d\mathcal{J}}{d\mathbf{a}_0} = \mathbf{a}^+(0), \quad (10)$$

and  $d\mathcal{J}/d\mathbf{u}_0$  is obtained via the fixed mapping  $d\mathbf{a}_0/d\mathbf{u}_0$ .

In this derivation we assume that the switching time is insensitive to the initial condition, i.e.  $dT_s/d\mathbf{a}_0 \approx 0$ . We verify this assumption in Figure 2(a). Accounting for event-time sensitivity would require a differentiable switching mechanism and additional terms [1, 8].

### 2.3 Variational data assimilation with ql-ROMs

We consider an application in variational data assimilation e.g., [15], where we plug the adjoint of the ql-ROM developed in the previous section in a gradient-based optimization scheme. We consider the problem of inferring the unknown initial condition  $\mathbf{u}_0$  of a trajectory given some observations at the final time  $T$ . We want to minimize the objective function

$$\mathcal{J} = \|\mathbf{y}(T) - \mathbf{H}\mathbf{u}(T)\|_2^2, \quad (11)$$

where  $\mathbf{y} \in \mathbb{R}^p$  are the observations and  $\mathbf{H} \in \mathbb{R}^{p \times N}$  is the observation matrix that maps the full state to the observations space. (In this formulation, uncertainties can be captured by using a covariance matrix and bias in the objective function e.g., [11].) Starting with an initial guess for  $\mathbf{u}_0$ , we propagate the ql-ROM's equations (4) forward in time. We linearize the ql-ROM around this trajectory, i.e., compute the local Jacobian at each time step, and integrate the adjoint equations (8) backward in time, obtaining  $d\mathcal{J}/d\mathbf{u}_0$ . We update the initial condition using steepest gradient descent

$$\mathbf{u}_0^{(i+1)} = \mathbf{u}_0^{(i)} - \alpha \frac{d\mathcal{J}^{(i)}}{d\mathbf{u}_0}, \quad (12)$$

where  $i$  is the iteration step and  $\alpha$  is the step size. The procedure is repeated until convergence. We perform the optimization in the full space by updating  $\mathbf{u}_0$  (rather than the reduced coordinates  $\mathbf{a}(0)$ ). This allows the cluster-affiliation map to be re-evaluated at each iteration, so that the sequence of active local models can adapt during the optimization and the algorithm can converge to the correct trajectory.

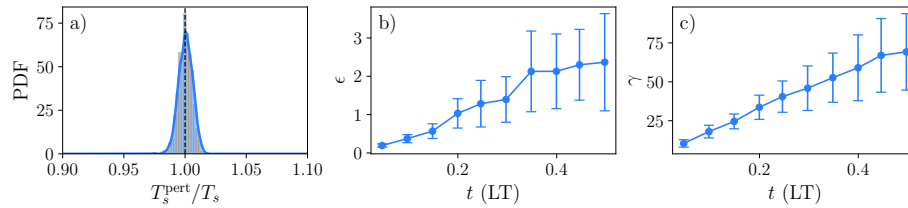
## 3 Application to spatio-temporal chaos

We demonstrate the developed methodology on a prototypical spatiotemporally chaotic system. The Kuramoto-Sivashinsky (KS) equation is a 1D PDE that arises in a variety of physical phenomena including flame front instabilities [14] and chemical reaction-diffusion equations [9]. The equation is described by

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + \nu \frac{\partial^4 u}{\partial x^4} = 0 \quad (13)$$

where  $u(x, t) : [0, L) \times [0, \infty) \rightarrow \mathbb{R}$  is the velocity on a periodic domain of length  $L$ , i.e.,  $u(x, t) = u(x+L, t)$ . Letting  $\nu = 1$ , the dynamical behaviour of the system is determined by the domain length  $L$ . Following a spectral discretization with 128 Fourier modes, the time integration is performed by a fourth-order exponential time differencing Runge-Kutta method (ETDRK4) [4] with a time step of  $\Delta t = 0.05$ . For the adjoint equations, we employ a discrete time formulation of the equations derived in Section 2.2.

For data generation, we choose  $L = 20\pi$ , which results in a chaotic regime with a Lyapunov (LT) of approximately 11.8 time units (235 time steps). The

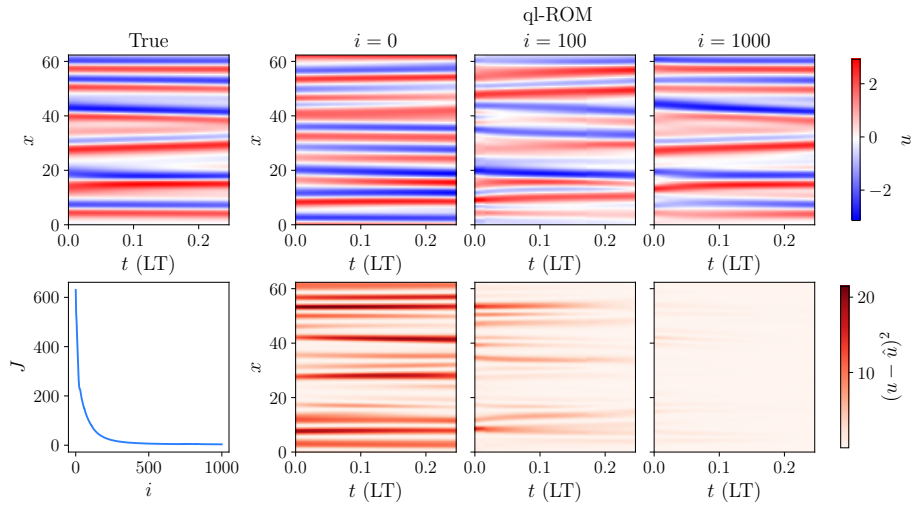


**Fig. 2.** Sensitivity of the switching time and accuracy of the adjoint sensitivity computed using the quantized local reduced model in comparison to the full-order model. (a) Estimated PDF of the normalized  $T_s$  in the ql-ROM, by perturbing the initial condition with random noise with standard deviation  $10^{-3}$ . (b) the relative  $\ell_2$ -norm of the error between these two gradient vectors, denoted by  $\epsilon$ , and (c) the cosine similarity between these two gradient vectors, denoted by  $\gamma$ . The error bars indicate the mean and  $\pm 1$  standard deviation of 50 randomly initialized runs within the test set.

system is initialized by setting  $u(x, 0) = \cos(x)$ . We discard the first  $5 \times 10^3$  time steps as transient, and we utilize the next  $10^5$  time steps in training and the next  $10^4$  time steps in testing. Following the analysis of [2], we set the number of clusters  $K = 10$  and build local ROMs using  $r_k = 30$  POD modes for each cluster.

First, we verify the accuracy of the sensitivity computation using the ql-ROM. For this purpose, we use an objective function  $\mathcal{J} = \|\mathbf{u}(T)\|_2^2$ , which is the  $\ell_2$ -norm of the final state. Figure 2 reports (a) the probability density function (PDF) of the normalized  $T_s$  obtained from randomly perturbed initial conditions, the accuracy of the ql-ROM adjoint sensitivity  $d\mathcal{J}/d\mathbf{u}_0$  with respect to the full-order model one via two error metrics (b) the relative  $\ell_2$ -norm of the error between these two gradient vectors, denoted by  $\epsilon$ , and (c) their cosine similarity, denoted by  $\gamma$ . Since gradient-based optimization primarily depends on the search direction, the cosine similarity is particularly informative in practice. As expected in chaotic systems, for short time horizons  $\sim 0.25$  LT, the gradient remains accurate, while it gradually loses accuracy as the time horizon increases.

Second, we consider the variational data assimilation problem described in 2.3 for trajectories of length  $T = 0.25$  LT. We assume full state measurements, i.e., the observation operator is the identity matrix  $\mathbf{H} = I_{N \times N}$ . The ql-ROM is initialized from a random snapshot on the chaotic attractor, and the initial condition is updated via gradient descent using the adjoint-based gradient. The results are shown in Figure 3, where we plot the predicted trajectory at the beginning of the optimization, after 100 iterations, and at the end of the optimization after 1000 iterations. The method successfully recovers the true trajectory with decreasing error near the measurement time. Furthermore, the ql-ROM methodology is significantly more computationally efficient compared to the full-order model. Adjoint-based optimization with the full-order model takes approximately 26 seconds for 100 iterations, whereas with the model reduction enabled by the ql-ROM, this computation time is reduced to approximately 7.4



**Fig. 3.** Variational data assimilation on the Kuramoto-Sivashinsky equation using the adjoint of the quantized local reduced-order model. The method reconstructs the true trajectory from measurements at the final time  $T = 0.25$  LT following the convergence of the optimization objective,  $\mathcal{J}$ .

seconds for 100 iterations, overall resulting in  $\times 3.5$  speed-up (results obtained on Intel Xeon(R) Gold 5218R CPU at  $2.10 \text{ GHz} \times 80$ ). We expect the cost of the adjoint method to scale with the cost of running two forward simulations. Reducing the state dimension from 128 to 30 would ideally yield a speed-up of about  $\times 4.3$ , which, in our numerical experiments, is reduced to  $\times 3.5$  due to the overhead associated with the cluster switching operations.

## 4 Conclusion

We introduced a computationally efficient reduced-order framework for adjoint-based optimization of spatiotemporally chaotic systems. The proposed method combines quantized local reduced-order modelling with adjoint-based optimization at a fraction of the full-order cost. We demonstrated the methodology on a variational data assimilation problem for the chaotic Kuramoto-Sivashinsky equation, showing that it can successfully reconstruct the full trajectory over horizons up to  $0.25$  Lyapunov times from full state measurements at the final time. The proposed algorithm achieves an overall speed-up of  $\times 3.5$  compared to the full-order model. In this paper, we used K-means for clustering and obtained the dynamics of the local reduced-order variables using Galerkin projection of the original PDE on local proper orthogonal decomposition modes. These choices are convenient but not essential: alternative clustering strategies and reduced modelling paradigms can be incorporated within the same quantized local and adjoint-based workflow. This method opens up new possibilities for efficient optimization, control, and data assimilation of spatiotemporally chaotic systems.

**Acknowledgments.** The authors acknowledge funding from the ERC Starting Grant No. PhyCo 949388.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. di Bernardo, M., Budd, C.J., Champneys, A.R., Kowalczyk, P.: Piecewise-smooth Dynamical Systems: Theory and Applications. Applied Mathematical Sciences, Springer London (2008). <https://doi.org/10.1007/978-1-84628-708-4>
2. Colanera, A., Magri, L.: Quantized local reduced-order modeling in time (ql-ROM). *Computer Methods in Applied Mechanics and Engineering* **447**, 118393 (Dec 2025). <https://doi.org/10.1016/j.cma.2025.118393>
3. Colanera, A., Magri, L.: Towards extreme event prediction of turbulent flows with quantized local reduced-order models (2025). <https://doi.org/10.48550/ARXIV.2511.04586>
4. Cox, S.M., Matthews, P.C.: Exponential Time Differencing for Stiff Systems. *Journal of Computational Physics* **176**(2), 430–455 (Mar 2002). <https://doi.org/10.1006/jcph.2002.6995>
5. Giles, M.B., Pierce, N.A.: An Introduction to the Adjoint Approach to Design. *Flow, Turbulence and Combustion* **65**(3/4), 393–415 (2000). <https://doi.org/10.1023/A:1011430410075>
6. Gunzburger, M.D.: Perspectives in Flow Control and Optimization. Society for Industrial and Applied Mathematics (Jan 2002). <https://doi.org/10.1137/1.9780898718720>
7. Holmes, P., Lumley, J.L., Berkooz, G.: Turbulence, Coherent Structures, Dynamical Systems and Symmetry. Cambridge University Press, 1 edn. (Oct 1996). <https://doi.org/10.1017/CBO9780511622700>
8. Kong, N.J., Payne, J.J., Zhu, J., Johnson, A.M.: Saltation matrices: The essential tool for linearizing hybrid dynamical systems. *arXiv* (2024). <https://doi.org/10.1109/JPROC.2024.3440211>
9. Kuramoto, Y.: Diffusion-Induced Chaos in Reaction Systems. *Progress of Theoretical Physics Supplement* **64**, 346–367 (1978). <https://doi.org/10.1143/PTPS.64.346>
10. Magri, L.: Adjoint Methods as Design Tools in Thermoacoustics. *Applied Mechanics Reviews* **71**(2), 020801 (Mar 2019). <https://doi.org/10.1115/1.4042821>
11. Nóvoa, A., Racca, A., Magri, L.: Inferring unknown unknowns: Regularized bias-aware ensemble kalman filter. *Computer Methods in Applied Mechanics and Engineering* **418**, 116502 (Jan 2024). <https://doi.org/10.1016/j.cma.2023.116502>
12. Ozan, D.E., Magri, L.: Data-driven computation of adjoint sensitivities without adjoint solvers: An application to thermoacoustics. *Physical Review Fluids* **9**(10), 103902 (Oct 2024). <https://doi.org/10.1103/physrevfluids.9.103902>
13. Quarteroni, A., Rozza, G., Manzoni, A.: Certified reduced basis approximation for parametrized partial differential equations and applications. *Journal of Mathematics in Industry* **1**(1), 3 (Dec 2011). <https://doi.org/10.1186/2190-5983-1-3>
14. Sivashinsky, G.I.: On Flame Propagation Under Conditions of Stoichiometry. *SIAM Journal on Applied Mathematics* **39**(1), 67–82 (Aug 1980). <https://doi.org/10.1137/0139007>
15. Zaki, T.A.: Turbulence from an Observer Perspective. *Annual Review of Fluid Mechanics* **57**(1), 311–334 (Jan 2025). <https://doi.org/10.1146/annurev-fluid-030424-114735>