

Data-driven stability analysis of a chaotic time-delayed system ^{*}

Georgios Margazoglou¹ and Luca Magri^{1,2,3}

¹ Department of Aeronautics, Imperial College London, London SW7 2AZ, UK

² The Alan Turing Institute, London NW1 2DB, UK

³ Isaac Newton Institute for Mathematical Sciences, Cambridge, CB3 0EH, UK (visiting)
g.margazoglou@imperial.ac.uk l.magri@imperial.ac.uk

Abstract. Systems with time-delayed chaotic dynamics are common in nature, from control theory to aeronautical propulsion. The overarching objective of this paper is to compute the stability properties of a chaotic dynamical system, which is time-delayed. The stability analysis is based only on data. We employ the echo state network (ESN), a type of recurrent neural network, and train it on timeseries of a prototypical time-delayed nonlinear thermoacoustic system. By running the trained ESN autonomously, we show that it can reproduce (i) the long-term statistics of the thermoacoustic system's variables, (ii) the physical portion of the Lyapunov spectrum, and (iii) the statistics of the finite-time Lyapunov exponents. This work opens up the possibility to infer stability properties of time-delayed systems from experimental observations.

Keywords: Echo State Networks · Chaos · Time-delayed systems.

1 Introduction

Chaotic systems with time-delayed dynamics appear in a range of scientific fields [8]. Because of their dependence on both the present and past states, these systems have rich and intricate dynamics. Their chaotic behaviour can be assessed with stability analysis, which is a mathematical tool that quantifies the system's response to infinitesimal perturbations. Stability analysis relies on the linearization of the time-delayed dynamical equations, which spawns the Jacobian of the system. From the Jacobian, we compute the Lyapunov Exponents (LEs), which are the key quantities to quantifying chaos [2].

A data-driven method, which considers the sequential nature of the dataset (e.g. time-series) to infer chaotic dynamics, is the recurrent neural network (RNN). Such networks have been successfully applied to learn chaotic dynamics (with no time delay) for different applications [12, 13, 17]. The majority of RNNs require backpropagation through time for training, which can lead to vanishing or exploding gradients, as well as long training times [17]. Instead, the echo state network (ESN), which we employ here, is trained via ridge regression, which eliminates backpropagation and provides a faster training [7, 9, 14]. The objective of this paper is to train an ESN with data from a prototypical chaotic thermoacoustic system, which is a nonlinear time-delayed wave equation. We further assess the capabilities of the ESN to accurately learn the ergodic and

^{*} This research has received financial support from the ERC Starting Grant No. PhyCo 949388.

stability properties of the thermoacoustic system, by calculating fundamental quantities, such as the Lyapunov exponents. Other RNN architectures for time-series forecasting are the Long-Short term memory and the Gated Recurrent units [3], which require backpropagation. In this paper, we take advantage of the simple training of ESNs, which do not require backpropagation. We briefly review stability analysis for time-delayed systems in Sec. 2 and present the considered thermoacoustic system in Sec. 2.1. In Sec. 3, we discuss the ESN architecture and properties. We present the results in Sec. 4 and conclude in Sec. 5.

2 Stability Analysis for time-delayed systems

We consider a physical state $\mathbf{x}(t) \in \mathbb{R}^D$, which is the solution of a nonlinear time-delayed dynamical system

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}(t), \mathbf{x}(t - \tau)), \quad \mathbf{x}(t) = \mathbf{x}_0, \quad \forall t \leq 0, \quad (1)$$

where τ is a constant time-delay. We analyse the system's stability by perturbing the state with infinitesimal perturbations $\mathbf{u} \sim \mathcal{O}(\varepsilon)$, $\varepsilon \rightarrow 0$, as $\mathbf{x} + \mathbf{u}$, with $\mathbf{x} \sim \mathcal{O}(1)$. Hence, we obtain the tangent linear equation

$$\frac{d\mathbf{U}}{dt} = \mathbf{J}(\mathbf{x}(t), \mathbf{x}(t - \tau)) \mathbf{U}, \quad (2)$$

which involves the time-marching of $K \leq D$ tangent vectors, $\mathbf{u}_i \in \mathbb{R}^D$, as columns of the matrix $\mathbf{U} \in \mathbb{R}^{D \times K}$, $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K]$. This is a linear basis of the tangent space. The linear operator $\mathbf{J} \in \mathbb{R}^{D \times D}$ is the Jacobian of the system, which is time-dependent in chaotic attractors. As shown in [12], we can extract the Jacobian of a reservoir computer by linearizing Eqs. (7), which are the evolution equations of the ESN.

We periodically orthonormalize the tangent space basis during time evolution by using a QR-decomposition of \mathbf{U} , as $\mathbf{U}(t) = \mathbf{Q}(t)\mathbf{R}(t, \Delta t)$ and by updating the columns of \mathbf{U} with the columns of \mathbf{Q} , i.e. $\mathbf{U} \leftarrow \mathbf{Q}$ [2]. The matrix $\mathbf{R}(t, \Delta t) \in \mathbb{R}^{K \times K}$ is upper-triangular and its diagonal elements $[\mathbf{R}]_{i,i}$ are the local growth rates over a time span Δt of \mathbf{U} . The LEs are the time averages of the logarithms of the diagonal of $[\mathbf{R}]_{i,i}$, i.e.,

$$\lambda_i = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^T \ln[\mathbf{R}(t, \Delta t)]_{i,i} dt. \quad (3)$$

The FTLEs are defined as $\Lambda_i = \frac{1}{\Delta t} \ln[\mathbf{R}]_{i,i}$, which quantify the expansion and contraction rates of the tangent space on finite-time intervals, $\Delta t = t_2 - t_1$.

2.1 Time-delayed thermoacoustic system

As a practical application of time-delayed systems, we consider a thermoacoustic system, which is composed of three interacting subsystems, the acoustics, the flame and the hydrodynamics (see, e.g., [10]). The interaction of these sub systems can result in a

positive feedback loop, which manifests itself as a thermoacoustic instability. If uncontrolled, this instability can lead to structural failure. Typical paradigms from engineering include gas-turbine and rocket-motor engines, and their configuration requires design optimization to prevent instabilities [10]. We consider a prototypical time-delayed thermoacoustic system with a longitudinal acoustic cavity and a heat source modelled with a time-delayed model, following the same setup as in [5, 6, 11]. The system is governed by the conservation of momentum, mass, and energy. Upon re-arrangement [10], thermoacoustic dynamics are governed by the nondimensional partial differential equations

$$\frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} = 0, \quad \frac{\partial p}{\partial t} + \frac{\partial u}{\partial x} + \zeta p - \dot{q}\delta(x - x_f) = 0, \quad (4)$$

where u is the non-dimensional velocity, p the pressure, \dot{q} the heat-release rate, $x \in [0, 1]$ the axial coordinate and t the time; ζ is the damping coefficient, which takes into account all the acoustic dissipation. The heat source is assumed to be small compared to the acoustic wavelength, and it is modelled as a point in the grid, via the Dirac delta distribution $\delta(x - x_f)$, located at $x_f = 0.2$. The heat-release rate is provided by a modified King's law, $\dot{q}(t) = \beta (\sqrt{|1 + u_f(t - \tau)|} - 1)$, which is a nonlinear time-delayed model. For the numerical studies of this paper, we set $\beta = 7.0$ for the heat parameter, and $\tau = 0.2$ for the time delay. Those values ensure chaotic evolution, and encapsulate all information about the heat source, base velocity and ambient conditions.

As in [5], we transform the time-delayed problem into an initial value problem. This is mathematically achieved by modelling the advection of a perturbation v with velocity τ^{-1} as

$$\frac{\partial v}{\partial t} + \frac{1}{\tau} \frac{\partial v}{\partial X} = 0, \quad 0 \leq X \leq 1, \quad v(X = 0, t) = u_f(t). \quad (5)$$

We discretise Eqs. (4) by a Galerkin method. First, we separate the acoustic variables in time and space as $u(x, t) = \sum_{j=1}^{N_g} \eta_j(t) \cos(j\pi x)$, and $p(x, t) = -\sum_{j=1}^{N_g} \mu_j(t) \sin(j\pi x)$, in which the spatial functions are the acoustic eigenfunctions of the configuration under investigation. Then, we project Eqs. (4) onto the Galerkin spatial basis $\{\cos(\pi x), \cos(2\pi x), \dots, \cos(N_g \pi x)\}$ to obtain

$$\dot{\eta}_j - j\pi\mu_j = 0, \quad \dot{\mu}_j + j\pi\eta_j + \zeta_j\mu_j + 2\dot{q}\sin(j\pi x_f) = 0. \quad (6)$$

The system has $2N_g$ degrees of freedom. The time-delayed velocity becomes $u_f(t - \tau) = \sum_{k=1}^{N_g} \eta_k(t - \tau) \cos(k\pi x_f)$, and the damping, ζ_j , is modelled by $\zeta_j = c_1 j^2 + c_2 j^{1/2}$, where $c_1 = 0.1$ and $c_2 = 0.06$. The equation for linear advection, Eq. (5), is discretised using $N_c + 1$ points with a Chebyshev spectral method. This discretisation adds N_c degrees of freedom, thus a total of $D = 2N_g + N_c = 30$ in our case, as $N_g = N_c = 10$. We integrate Eqs. (6) with a fourth order Runge-Kutta scheme and timestep $dt = 0.01$.

3 Echo State Network

By applying the method of [12] to time-delayed problems, we linearize the Echo State Network (ESN) [7] to calculate the stability properties of chaotic systems. ESNs are proven effective for accurate learning of chaotic dynamics (see e.g. [1, 4, 12–14, 17]).

The ESN is a reservoir computer [7]. It has a sparsely-connected single-layer hidden state, which is termed “reservoir”. The reservoir weights, \mathbf{W} , as well as the input-to-reservoir weights, \mathbf{W}_{in} , are randomly assigned and remain fixed through training and testing. The reservoir-to-output weights, \mathbf{W}_{out} , are trained via ridge regression. The evolution equations of the reservoir and output are, respectively

$$\mathbf{r}(t_{i+1}) = \tanh([\hat{\mathbf{y}}_{\text{in}}(t_i); b_{\text{in}}]^T \mathbf{W}_{\text{in}} + \mathbf{r}(t_i)^T \mathbf{W}), \quad \mathbf{y}_{\text{p}}(t_{i+1}) = [\mathbf{r}(t_{i+1}); 1]^T \mathbf{W}_{\text{out}}, \quad (7)$$

where at any discrete time t_i the input vector, $\mathbf{y}_{\text{in}}(t_i) \in \mathbb{R}^{N_y}$, is mapped into the reservoir state $\mathbf{r} \in \mathbb{R}^{N_r}$, by the input matrix, \mathbf{W}_{in} , where $N_r \gg N_y$ [7, 14]. Here, $(\hat{\cdot})$ indicates normalization by the component-wise maximum-minus-minimum range of the target in training set, T indicates matrix transposition, and the semicolon indicates array concatenation. The dimensions of the weight matrices are $\mathbf{W}_{\text{in}} \in \mathbb{R}^{(N_y+1) \times N_r}$, $\mathbf{W} \in \mathbb{R}^{N_r \times N_r}$ and $\mathbf{W}_{\text{out}} \in \mathbb{R}^{(N_r+1) \times N_y}$. The hyperparameter input bias, $b_{\text{in}} = 1$, is selected to have the same order of magnitude as the normalized inputs, $\hat{\mathbf{y}}_{\text{in}}$. The dimensions of the input and output vectors are equal to the dimension of the dynamical system; here described by Eqs. (6), i.e. $N_y \equiv D$. Furthermore, \mathbf{W}_{out} is trained via the minimization of the mean square error $\text{MSE} = \frac{1}{N_{\text{tr}} N_y} \sum_{i=0}^{N_{\text{tr}}} \|\mathbf{y}_{\text{p}}(t_i) - \mathbf{y}_{\text{in}}(t_i)\|^2$ between the outputs and the data over the training set, where $\|\cdot\|$ is the L_2 norm, $N_{\text{tr}} + 1$ is the total number of data in the training set, and \mathbf{y}_{in} the input data on which the ESN is trained.

Training the ESN is performed by solving with respect to \mathbf{W}_{out} via ridge regression of the equation $(\mathbf{R}\mathbf{R}^T + \beta\mathbb{I})\mathbf{W}_{\text{out}} = \mathbf{R}\mathbf{Y}_{\text{d}}^T$. In the previous expression $\mathbf{R} \in \mathbb{R}^{(N_r+1) \times N_{\text{tr}}}$ and $\mathbf{Y}_{\text{d}} \in \mathbb{R}^{N_y \times N_{\text{tr}}}$ are the horizontal concatenation of the reservoir states with bias, $[\mathbf{r}(t_i); 1]$, $t_i \in [0, T_{\text{train}}]$, and of the output data, respectively; \mathbb{I} is the identity matrix and β is the Tikhonov regularization parameter [16]. Therefore the ESN does not require backpropagation. The ESN can run in two configurations, either open-loop or closed-loop. In open-loop, which is necessary for the training stage, the input data is given at each step, allowing for the calculation of the reservoir timeseries $\mathbf{r}(t_i)$, $t_i \in [0, T_{\text{train}}]$. In closed-loop the output \mathbf{y}_{p} at time step t_i , is recurrently used as an input at time step t_{i+1} , allowing for the autonomous temporal evolution of the network. The closed-loop configuration is used for validation (i.e. hyperparameter tuning) and testing, but not for training.

Regarding validation, we use the chaotic recycle validation (RVC), as introduced in [14]. It has proven to be a robust strategy, providing enhanced performance of the ESN, compared to standard strategies, as recently successfully applied in [12, 15]. Briefly, in RVC the network is trained only once on the entire training dataset (in open-loop), and validation is performed on multiple intervals already used for training (but now in closed-loop). The validation interval simply shifts as a small multiple of the first Lyapunov exponent, $N_{\text{val}} = 3\lambda_1$ here. The key hyperparameters that we tune are the input scaling σ_{in} of the input matrix \mathbf{W}_{in} , the spectral radius ρ of the matrix \mathbf{W} , and the Tikhonov parameter β . Furthermore, σ_{in} and ρ are tuned via Bayesian Optimization in the hyperparameter space $[\sigma_{\text{in}}, \rho] = [-1, 1] \times [0.1, 1.4]$, while for β we perform a grid search $\{10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}\}$ within the optimal $[\sigma_{\text{in}}, \rho]$. The reservoir size is $N_r = 300$. The connectivity of matrix \mathbf{W} is set to $d = 80$. We further add to the training and validation data a Gaussian noise with zero mean and standard deviation, $\sigma_n = 0.0006\sigma_y$, where σ_y is the standard deviation of the data component-wise (noise

regularizes the problem, see [12, 14, 17] for more details). The ESN is trained on a training set (open-loop) of size $500\tau_\lambda$, and is tested on a test set (closed-loop) of size $4000\tau_\lambda$, where $\tau_\lambda = 1/\lambda_1$ is the Lyapunov time, which is the inverse of the maximal Lyapunov exponent $\lambda_1 \approx 0.13$ in our case.

4 Results

We analyse the statistics produced by the autonomous temporal evolution of the ESN and the target time-delayed system. The selected observables are the statistics of the system's chaotic variables, the Lyapunov exponents (LEs), and the statistics of the finite-time Lyapunov exponents (FTLEs).

First, we test the capabilities of the ESN to learn the long-term statistical properties of the thermoacoustic system by measuring the probability density function (PDF) of the learned variables, \mathbf{y} . In Fig. 1, we show the PDF of the first three components of the Galerkin modes (see Eq. (6)) η_i, μ_i for $i = 1, 2, 3$, in which the black line corresponds to the target and the dashed red to the ESN. The ESN predictions are in agreement with the target, including the variables that are not shown here.

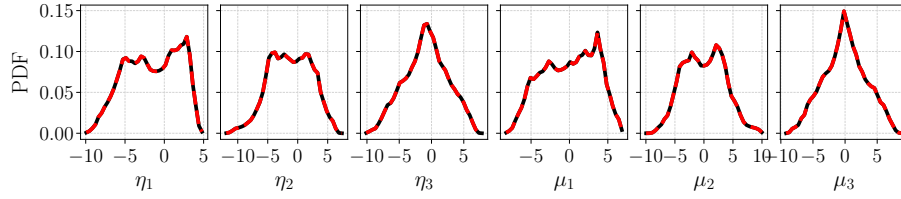


Fig. 1: Probability density functions (via histograms) of the three first Galerkin modes, $\eta_i, \mu_i, i = 1, 2, 3$. Black is for target and dashed red for ESN. The statistics for ESN are collected in autonomous evolution on unseen data, after training and validation.

Testing the accuracy of the calculated LEs in autonomous evolution is a harder consistency check for the ESN. Indeed, the ESN has been trained only on timeseries of the variables η_i, μ_i , and v . Therefore, a good agreement of the LEs means that the ESN is capable to accurately reproduce intrinsic chaotic properties of the system's attractor. The LEs of the ESN are calculated following [12], and in Fig. 2 we compare the first $K = 14$ LEs. We also add an inner plot showing, in additional detail, the first 6 LEs. Each LE is the average of the measured LEs from seven selected independent ESNs used for the analysis. We train the ESNs independently on different chaotic target sets. The shaded region corresponds to the standard deviation per λ_i from those seven ESNs. There is close agreement for the first 8 exponents. In particular, we measure the leading, and only positive, exponent $\lambda_1^{\text{targ}} \approx 0.130$ and $\lambda_1^{\text{ESN}} \approx 0.127$ for ESN, which gives a 2.3% absolute error. The ESN also provides an accurate estimate of the neutral exponent ($\lambda_2 = 0$) with $\lambda_2^{\text{ESN}} \approx 0.008$. The rest of the exponents, $\lambda_i, i \geq 3$, are negative and the ESN achieves a small 5.4% mean absolute percentage error for all. Note

that a gradual disagreement of the negative exponents, which are sensitive due to the numerical method, between ESN and target has also been reported in [13, 17] for the one-dimensional Kuramoto-Sivashinsky equation.

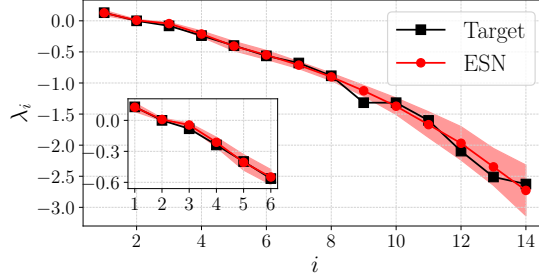


Fig. 2: The first 14 (outer plot) and the first 6 (inner plot) Lyapunov exponents. The shaded red region indicates the error based on the ensemble of 5 ESNs. The statistics for ESN are collected in closed-loop mode.

Figure 3 shows the PDF of the first 14 FTLEs of the target (black line) and ESN (dashed red line). We collect the statistics from 7 independently trained ESNs, thus creating 7 histograms of FTLEs. We then average those histograms bin-wise, and the standard deviation of each averaged bin is given by the shaded regions. Note that the mean of each PDF should correspond to each Lyapunov exponent λ_i (i.e. of Fig. 2), which is indeed the case. In Fig. 4, we quantify the difference of the estimated FTLE statistics by computing the Wasserstein distance between the two distributions for each Λ_i . The difference is small. We observe a gradual increase in Λ_i , $i \geq 10$, which is in agreement with the trend of the Lyapunov exponents in Fig. 2.

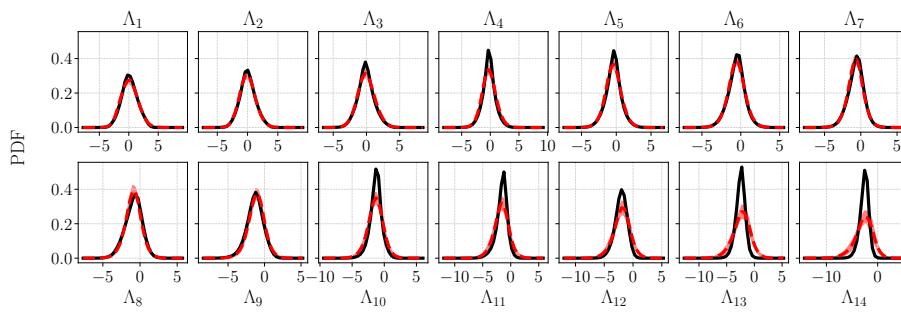


Fig. 3: Probability density functions of the first 14 finite-time Lyapunov exponents. Black is for target and dashed red for ESN.

We also report the values for the Kaplan-Yorke dimension for both ESN and target. This dimension is an upper bound of the attractor's fractal dimension [2]. It is given by $D_{KY} = k + \frac{\sum_{i=1}^k \lambda_i}{|\lambda_{k+1}|}$, where k is such that the sum of the first k LEs is positive and the sum of the first $k + 1$ LEs is negative. We obtain $D_{KY}^{\text{targ}} \approx 3.37$ for target and $D_{KY}^{\text{ESN}} \approx 3.43$ for ESN, which results in a low 1.8% absolute error. This observation further confirms the ability of the ESN to accurately learn and reproduce the properties of the chaotic attractor when running in autonomous closed loop mode.

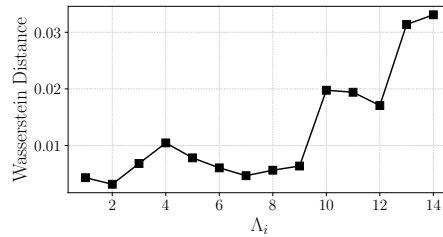


Fig. 4: Wasserstein distance between the PDFs of target versus ESN of the first 14 finite-time Lyapunov exponents (see Fig. 3).

5 Conclusion

We propose a method to compute the stability properties of chaotic time-delayed systems using only data. We use the echo state network (ESN) as a surrogate model for learning the chaotic dynamics from time series observations of the system and calculate its long-term statistical and stability properties. Viewing the ESN as a discrete dynamical system, we linearize the map (7) to derive the tangent evolution of the attractor through the Jacobian. When running the ESN in a long autonomous mode (closed-loop), we show that (i) the long-term statistics of the variables are correctly learned, (ii) the physical portion of the Lyapunov spectrum is correctly predicted, and (iii) the finite-time Lyapunov exponents and their statistics are correctly inferred. This work opens up the possibility to infer the stability of nonlinear and time-delayed dynamical systems from data. Future directions include the application of this approach to experimental observations with the objective of aiding sensitivity analysis [5, 6].

References

1. Doan, N.A.K., Polifke, W., Magri, L.: Short- and long-term predictions of chaotic flows and extreme events: a physics-constrained reservoir computing approach. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **477**(2253), 20210135 (2021). <https://doi.org/10.1098/rspa.2021.0135>
2. Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617–656 (Jul 1985). <https://doi.org/10.1103/RevModPhys.57.617>

3. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016), <http://www.deeplearningbook.org>
4. Huhn, F., Magri, L.: Learning ergodic averages in chaotic systems. In: Krzhizhanovskaya, V.V., Závodszy, G., Lees, M.H., Dongarra, J.J., Sloot, P.M.A., Brissos, S., Teixeira, J. (eds.) Computational Science – ICCS 2020. pp. 124–132. Springer International Publishing, Cham (2020)
5. Huhn, F., Magri, L.: Stability, sensitivity and optimisation of chaotic acoustic oscillations. *Journal of Fluid Mechanics* **882**, A24 (2020). <https://doi.org/10.1017/jfm.2019.828>
6. Huhn, F., Magri, L.: Gradient-free optimization of chaotic acoustics with reservoir computing. *Phys. Rev. Fluids* **7**, 014402 (Jan 2022), <https://doi.org/10.1103/PhysRevFluids.7.014402>
7. Jaeger, H., Haas, H.: Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science* **304**(5667), 78–80 (2004). <https://doi.org/10.1126/science.1091277>
8. Lakshmanan, M., Senthilkumar, D.V.: Dynamics of nonlinear time-delay systems. Springer Science & Business Media (2011). <https://doi.org/10.1007/978-3-642-14938-2>
9. Lukoševičius, M.: A Practical Guide to Applying Echo State Networks, pp. 659–686. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35289-8_36
10. Magri, L.: Adjoint Methods as Design Tools in Thermoacoustics. *Applied Mechanics Reviews* **71**(2) (03 2019). <https://doi.org/10.1115/1.4042821>, 020801
11. Magri, L., Juniper, M.P.: Sensitivity analysis of a time-delayed thermo-acoustic system via an adjoint-based approach. *Journal of Fluid Mechanics* **719**, 183–202 (2013). <https://doi.org/10.1017/jfm.2012.639>
12. Margazoglou, G., Magri, L.: Stability analysis of chaotic systems from data. *Nonlinear Dynamics* **111**, 8799–8819 (2023). <https://doi.org/10.1007/s11071-023-08285-1>
13. Pathak, J., Lu, Z., Hunt, B.R., Girvan, M., Ott, E.: Using machine learning to replicate chaotic attractors and calculate lyapunov exponents from data. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**(12), 121102 (2017). <https://doi.org/10.1063/1.5010300>
14. Racca, A., Magri, L.: Robust optimization and validation of echo state networks for learning chaotic dynamics. *Neural Networks* **142**, 252–268 (2021). <https://doi.org/https://doi.org/10.1016/j.neunet.2021.05.004>
15. Racca, A., Magri, L.: Data-driven prediction and control of extreme events in a chaotic flow. *Phys. Rev. Fluids* **7**, 104402 (Oct 2022). <https://doi.org/10.1103/PhysRevFluids.7.104402>
16. Tikhonov, A.N., Goncharsky, A., Stepanov, V., Yagola, A.G.: Numerical methods for the solution of ill-posed problems, vol. 328. Springer Science & Business Media (1995). <https://doi.org/doi.org/10.1007/978-94-015-8480-7>
17. Vlachas, P., Pathak, J., Hunt, B., Sapsis, T., Girvan, M., Ott, E., Koumoutsakos, P.: Back-propagation algorithms and Reservoir Computing in Recurrent Neural Networks for the forecasting of complex spatiotemporal dynamics. *Neural Networks* **126**, 191–217 (2020). <https://doi.org/https://doi.org/10.1016/j.neunet.2020.02.016>