

# Regularized classification and simulation of bifurcation regimes in nonlinear systems <sup>★</sup>

Defne E. Ozan,<sup>\*</sup> Andrea Iannelli,<sup>\*</sup> Mingzhou Yin,<sup>\*</sup>  
Roy S. Smith<sup>\*</sup>

<sup>\*</sup> *Automatic Control Lab, ETH, Zürich 8092, Switzerland.*

---

**Abstract:** The paper proposes a multi-step identification approach to classify a nonlinear system into qualitatively different regimes and then estimate a low-dimensional subspace where predictions of the original state at future times can be obtained by simulation of low-order dynamics. Proper Orthogonal Decomposition is used to build a library of characteristic modes from training data and is combined with regularization techniques for both the classification and estimation problems. Group Lasso is proposed to more effectively perform the former task. Moreover,  $\ell_1$  and  $\ell_2$  regularization problems with singular values weighting of the dynamic modes are suggested to handle the estimation problem in complex scenarios where limited measurement points are available or sensors are noisy. Results obtained on the Rijke tube system, a nonlinear thermoacoustic benchmark problem, demonstrate better classification accuracy and lower prediction error compared with a method from the literature.

*Keywords:* System identification, Regularization, Bifurcations, Thermoacoustics

---

## 1. INTRODUCTION

Models of dynamical systems are traditionally derived from first principles by leveraging domain knowledge. This practice often results in high-order nonlinear partial differential equations (PDEs) that finely describe the interaction among physical variables, e.g. pressure and velocity in an incompressible fluid (Schmid and Henningson, 2001). While this provides valuable information on detailed aspects of the system, the complexity of the model often rules out its use in applications where fast computations are required, e.g. real-time control. Moreover, models always rely on arbitrary assumptions on the system description and are thus susceptible to uncertainty (Iannelli et al., 2020). Unfortunately, assessing robustness of the high-dimensional nonlinear PDE models typically arising from first-principles modeling is often intractable.

Prompted by these shortcomings, increasing interest has been devoted to well-established techniques allowing the extraction of low-order structures from simulations of complex systems (Holmes et al., 2012), with the intent of extending them to the general problem of identifying nonlinear dynamics from experimental data. An interesting work in this research direction was presented in (Brunton et al., 2014), where the identification of nonlinear systems in distinct regimes, each corresponding to qualitatively different response features, is proposed. Proper Orthogonal Decomposition (POD) is used to obtain a library of representative dynamical modes for each bifurcation regime and compressive sensing techniques are used to classify a given partial-state measurement into one of the bifurcation regimes. Following the classification, predictions of future

state trajectories are made using a Galerkin projection-based scheme, whereby the coefficients of the POD modes become the new low-dimensional states. Thus, once the initial value of these fictitious states is estimated from the partial-state measurement, low-order dynamics can be used to predict the time-evolution of the coefficients and, from these, of the original states.

In this work, a further step is taken by combining the POD-based dimensionality reduction approaches with regularization techniques from the statistical learning community (Hastie et al., 2001). First, the classification step is formulated using Group Lasso, which promotes sparsity of groups of modes rather than of the single modes, with tangible benefits in the cases where the different regimes share similar modes. For the second step of the algorithm, consisting of estimating the low-order Galerkin state, regularized regression problems are proposed, particularly suited for underdetermined (number of measurements smaller than number of POD modes) or noisy scenarios. The proposed regularization problems include weighted  $\ell_1$  and  $\ell_2$  norms defined based on the singular values of the POD modes, which provide a priori information (from data) on the relevance of each mode.

Application of the proposed approach is demonstrated on a thermoacoustic problem. This subject is concerned with the interaction between heat transfer and sound waves and it is of well-motivated interest owing to its potential application for energy harvesting (Swift, 2007). We use a model of the Rijke tube described by a set of nonlinear PDEs (Sayadi et al., 2014) to generate the (synthetic) data used by the algorithm. The results confirm the advantages of the proposed regularization solutions, especially for scenarios typically arising in experimental settings.

---

<sup>★</sup> This work is supported by the Swiss National Science Foundation under grant no. 200021\_178890.

Correspondence to: iannelli@control.ee.ethz.ch

## 2. MULTI-STEP IDENTIFICATION APPROACH

This section describes the proposed methodology for the nonlinear identification problem considered in this work.

It is assumed that the response of the nonlinear system under study can be qualitatively divided in  $J$  distinct regimes denoted by  $\beta_j$  (with  $j = 1, 2, \dots, J$ ). Each regime is associated with different values of some fundamental parameters on which the system dynamics depend, thus they will be informally referred to as *bifurcation regimes*. We consider autonomous vector fields where, in a generic regime  $j$ , the state  $Y \in \mathbb{R}^n$  evolves according to a set of ordinary differential equations (ODE)

$$\dot{Y} = f(Y, \beta_j), \quad Y(0) = Y_0. \quad (1)$$

Given a partial measurement of the state at a certain time, the approach proposed in this work aims at: identifying the bifurcation regime to which the measured state belongs; and predicting the full-state of the system at any future time. The following sections detail the three steps through which this objective is achieved.

### 2.1 Step 1 - Bifurcation library

The first step of the algorithm consists of building, in each bifurcation regime, a library of system's responses via the snapshot matrix  $A_j$  containing  $q$  observations

$$A_j = [Y_j(t_1) \ Y_j(t_2) \ \dots \ Y_j(t_q)], \quad A_j \in \mathbb{R}^{n \times q},$$

where  $Y_j$  denotes the full-state observed in the regime  $\beta_j$ , and typically  $q \ll n$ . The sampling window  $([t_1, t_q])$  and (possibly non-uniform) sampling time are chosen in order to sufficiently capture the main features of the system's response. The singular value decomposition (SVD) of  $A_j$

$$A_j = \Psi_j \Sigma_j \Omega_j^\top, \quad (2)$$

provides the  $q$  POD modes of regime  $\beta_j$ , namely the columns of  $\Psi_j \in \mathbb{R}^{n \times q}$ . The energy captured by each mode is proportional to the square of the corresponding diagonal entry of  $\Sigma_j$  (Holmes et al., 2012). By setting an energy threshold, the least dominant  $(q - r_j)$  modes can be discarded and the matrix of retained POD modes is denoted by  $\Psi_j^{r_j} \in \mathbb{R}^{n \times r_j}$ . The bifurcation library is made of the POD modes obtained in each bifurcation regime

$$\Psi = [\Psi_1 \ \Psi_2 \ \dots \ \Psi_J], \quad \Psi \in \mathbb{R}^{n \times p},$$

where  $p = \sum_{j=1}^J r_j$ . A generic state  $Y$  in any of the  $J$  bifurcation regimes can then be approximated as

$$Y \approx \Psi a, \quad a = [a_1^\top \ \dots \ a_J^\top]^\top, \quad a_j = [a_{j,1} \ \dots \ a_{j,r_j}]^\top, \quad (3)$$

where the coefficient vector  $a$  gives information on the contribution of every POD mode to the state  $Y$ .

### 2.2 Step 2 - Classification of the bifurcation regime

In experimental settings, full-state measurements are not available, and more realistically one has  $m$  measurements taken from a limited number of sensors. The objective here is to identify, from these partial measurements taken at a certain time, the bifurcation regime from which the signal comes. To this aim, a compressive sensing approach is employed. First, the measured state vector  $\hat{Y} \in \mathbb{R}^m$  is written as a function of the full-state  $\hat{Y} = \Phi Y$  by means of the appropriately defined matrix  $\Phi \in \mathbb{R}^{m \times n}$  with 0

and 1 entries. By using the bifurcation library (3), the measurements vector  $\hat{Y}$  can be rewritten

$$\hat{Y} = \Phi \Psi a = \Upsilon a, \quad \Upsilon \in \mathbb{R}^{m \times p}. \quad (4)$$

Notice that in typical applications,  $m < p$ , i.e. the number of measurements is smaller than the number of POD modes across all the bifurcation regimes. In this case, Eq. (4) is underdetermined. Note also that not all modes in  $\Psi$  are required for the reconstruction, but only those from the regime  $\beta_j$  from which the observation comes. Therefore, the coefficients vector  $a$  is sparse. This fact was leveraged in (Brunton et al., 2014) to propose a Lasso regression problem for finding the coefficient vector  $a$  better explaining the measurement

$$\hat{a}^{\text{Lasso}} = \underset{a}{\operatorname{argmin}} \left( \|\hat{Y} - \Upsilon a\|_2^2 + \lambda \|a\|_1 \right), \quad (5)$$

where the  $\ell_1$  regularization, as is well known (Hastie et al., 2001), promotes sparsity of the solution, and the hyperparameter  $\lambda$  can be selected by cross-validation. The bifurcation regime associated with the measurement  $\hat{Y}$  can then be inferred by looking at the largest (in absolute value) coefficient entry (i.e. the most activated mode).

Unfortunately, this strategy can lead to misclassifications when different regimes have similar modes that can equally well explain the data. Measurements can then be wrongly attributed to a regime where a mode similar to the one in the correct regime is activated, even though the other modes in the wrong regime cannot explain them. The issue is that the vector  $a$  is in truth sparse with respect to the bifurcation regimes rather than to the modes. For this reason, in this paper it is proposed to classify the bifurcation regime via a Group Lasso problem (Yuan and Lin, 2006)

$$\hat{a}^{\text{GLasso}} = \underset{a}{\operatorname{argmin}} \left( \|\hat{Y} - \Upsilon a\|_2^2 + \lambda \sum_{j=1}^J \|a_j\|_2 \right). \quad (6)$$

In (6), the cost function is modified so that the regularization term features the sum of  $\ell_2$  norms of the coefficients of each group. It combines  $\ell_1$  (across different groups) and  $\ell_2$  (within the same group) regularizations. The rationale for this is that the predictor  $a$  belongs to some pre-defined group (i.e. the bifurcation regime), as given by its definition in (3). Thus, it is advantageous to select the members of a group together, while leaving the rest of the modes inactive. The hyperparameter selection consists of finding the smallest  $\lambda$  for which only one group is activated.

The classified regime will be denoted by  $\hat{j}$  and the corresponding set of POD modes by  $\hat{\Psi}_{\hat{j}}$  (with coefficients  $\hat{a}_{\hat{j}}$ ).

### 2.3 Step 3 - State estimation and simulation

Once the bifurcation regime to which the system belongs has been detected, the objective is to form, from the data, a low-order approximation of the dynamics (1), which allows the value of the full-state vector  $Y$  at future times to be predicted at less computational cost. In the spirit of Galerkin projections methods (Holmes et al., 2012), this is achieved by considering a linear subspace of the original state-space having size  $r_j$  and spanned by  $\hat{\Psi}_{\hat{j}}$  onto which the dynamic evolution of the coefficients  $\hat{a}_{\hat{j}}$  can be described. The set of POD modes  $\hat{\Psi}_{\hat{j}}$  thus provides

the operator whereby: the full-state  $Y$  is projected onto the subspace where  $a$  evolves; and  $a$  is lifted back to the original state-space. This problem can be decomposed into two sub-steps.

*State estimation* The first step consists of estimating the initial state  $\hat{a}_0 \in \mathbb{R}^{r_j}$  of the low-order dynamics. According to the previous discussion, this can be approximated as  $\hat{Y} \approx \hat{\Psi}_j \hat{a}_0$ , where  $\hat{Y}$  is the measured state vector.

A least-squares (LS) problem was proposed in (Brunton et al., 2014) to find the initial state

$$\hat{a}_0 = \underset{a}{\operatorname{argmin}} \left( \|\hat{Y} - \hat{\Upsilon}_j a\|_2^2 \right), \quad (7)$$

where  $\hat{\Upsilon}_j = \Phi \hat{\Psi}_j \in \mathbb{R}^{m \times r_j}$  is the regressor. This approach however does not prove to be robust in two scenarios commonly encountered in experiments: noisy state measurements (where LS is prone to overfitting); and sparse measurements  $m < r_j$  (resulting in an underdetermined problem). For this reason, we investigate the use of regularization to solve the low-order state estimation problem.

When the presence of noise is of prime concern, a valid alternative to the standard LS is ridge regression

$$\hat{a}_0 = \underset{a}{\operatorname{argmin}} \left( \|\hat{Y} - \hat{\Upsilon}_j a\|_2^2 + \lambda \|a\|_2 \right), \quad (8)$$

which consists of adding an  $\ell_2$  norm penalty term to (7). An alternative is to consider in (8) an  $\ell_1$  penalty to incentivize sparsity of  $\hat{a}_0$ , leading to a Lasso problem. Observe that, differently than in (5), this Lasso problem assumes sparsity within the classified regime.

In the underdetermined case, there exists a subspace of solutions  $\hat{a}_0$  satisfying (7) with zero residual. The classic approach (also suggested in (Brunton et al., 2014)) consists of taking the pseudo-inverse of  $\Psi_j$ , namely  $\hat{a}_0^{\text{PINV}} = \Psi_j^\dagger \hat{Y}$ . Note that  $\hat{a}_0^{\text{PINV}}$  coincides with the minimizer of the following constrained  $\ell_2$  norm projection, or least-norm (LN) problem

$$\begin{aligned} \min_a \quad & \|a\|_2, \\ \text{s.t.} \quad & \hat{Y} - \hat{\Upsilon}_j a = 0. \end{aligned} \quad (9)$$

An alternative, when there is the same sparsity prior commented on earlier, is to penalize the  $\ell_1$  norm of the solution, leading to a basis pursuit (BP) problem.

We observe now that, thus far, the only information that is used from the data is the set of POD modes (upon which the lower-dimensional subspace where the predictor is defined can be computed). However, when the bifurcation library is computed, the energy associated with each mode is also available in the form of the singular values on the diagonal elements of  $\Sigma_j$  (2). They provide a measure of the dominance of the corresponding modes in the measured response. It could then be favourable to leverage this information in the projection step, especially in those circumstances (noisy data or underdetermined instances) where the classification problem is more challenging. Motivated by this, the four regression problems illustrated before are modified by introducing weights on the regularization terms that are inversely proportional to the singular values of the modes. Specifically, we define  $\hat{w}_j = [1/\hat{\Sigma}_{j1,1}, 1/\hat{\Sigma}_{j2,2}, \dots, 1/\hat{\Sigma}_{jr_j,r_j}]$ , where  $\hat{\Sigma}_{ji,i}$  is the  $i$ -th diagonal entry of the matrix of singular values for the

bifurcation regime  $j$ . The *weighted* ridge (8) and least-norm (9) problems can then be rewritten as

$$\hat{a}_0 = \underset{a}{\operatorname{argmin}} \left( \|\hat{Y} - \Upsilon_j a\|_2^2 + \lambda \|\hat{w}_j \circ a\|_2 \right), \quad (10)$$

and

$$\begin{aligned} \min_a \quad & \|\hat{w}_j \circ a\|_2, \\ \text{s.t.} \quad & \hat{Y} - \Upsilon_j a = 0, \end{aligned} \quad (11)$$

respectively, where  $\circ$  denotes the Hadamard product. The weighted penalty terms thus fuse information from data (i.e. the singular values) and regularization. Equivalent modifications lead to the *weighted* Lasso and BP problems.

*Simulation* All the discussion below refers to the bifurcation regime where the response has been classified, and the related subscript  $j$  is dropped for clarity.

Given the initial state  $\hat{a}_0$  of the low-order dynamics estimated at the previous step, the dynamic evolution of  $a$  can be simulated and, from it, the original full-state  $Y$  can be predicted. As also suggested in (Brunton et al., 2014), this can be done by using Galerkin projection

$$\dot{a} = \hat{\Psi}^\top f(\hat{\Psi}a, \beta_j), \quad a(0) = \hat{a}_0, \quad (12)$$

where the right hand side of (12) comes from projecting the original dynamics onto the POD modes and exploiting the fact that they form an orthogonal basis. The full-state is then available at all times by lifting the coefficient vector to the original state space via  $Y = \hat{\Psi}a$ . It is worth noting that, unlike in standard model-based Galerkin methods, the computation of the modes comes here entirely from the data, together with the initial condition. This has the advantage that it is not required to fix a priori the shape for the modes (e.g. sinusoidal functions with pre-defined frequencies). The method however still partially relies on the knowledge of the system since evaluation of the vector field  $f$  is required in (12), but notably it does not require simulation of (1), instead only of lower dimensional ODEs.

### 3. THE RIJKE TUBE SYSTEM

The Rijke tube is the nonlinear thermoacoustic system used to demonstrate the application of the proposed identification method. It consists Fig. (1) of a tube of length  $L$  and a lumped heat source located at  $x_f$ . In the horizontal configuration, there is a uniform air flow with speed  $\tilde{u}$  and the goal is predicting  $u$  and  $p$ , i.e. the deviations of the velocity and pressure fields with respect to  $\tilde{u}$  and ambient pressure, respectively.

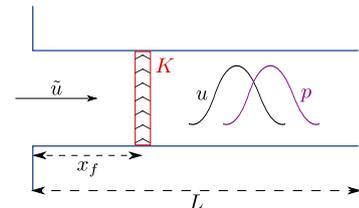


Fig. 1. Schematic of the Rijke tube system.

We follow (Sayadi et al., 2014) to implement a first-principles model of the system used to generate the data for testing the algorithm. A set of PDEs coming from momentum and energy balances and linearized about a

steady-state solution is used to describe the velocity and pressure field inside the tube. The states thus represent deviations from the steady-state values and are non-dimensionalized. The heat released from the wire acts as source term in the energy balance, and depends on the gas velocity via the King's Law (Epperlein et al., 2015). This describes a saturation-type nonlinearity and is the key element triggering the nonlinear responses described later. There is also a delay  $\tau$  between the gas velocity and the heat release (modelled with a first order approximation). The spatial discretization of the PDEs is done using a staggered scheme. Denoting the number of grid points by  $N$ , this results in a state vector  $Y \in \mathbb{R}^{2N-1}$  composed of  $N$  velocity and  $N - 1$  pressure states, specifically  $Y = [U_{1/2}, \dots, U_{i-1/2}, \dots, U_{N-1/2}, P_1, \dots, P_i, \dots, P_{N-1}]^T$ . In all the analyses shown later  $N = 64$ , and thus the states are  $n = 127$ . The resulting model is in the ODE form (1).

The effect of two fundamental parameters, namely the heat strength  $K$  and location  $x_f$ , is investigated. The other parameters are set to the same values considered in (Sayadi et al., 2014), and the time delay  $\tau = 0.05$ . All variables are adimensionalized with respect to velocity, pressure, length and time reference scales. The reference time scale is  $\frac{L_0}{c_0}$ , i.e. the ratio between the tube length and the velocity of sound in air at standard conditions. The nonlinear behaviour of the system is qualitatively investigated via time-domain simulations by perturbing the initial condition of the velocity at the heat source. The effect of increasing the heat strength for a fixed heat source location and delay is analysed in Fig. 2(a)-(c). For this analysis, the heat source location is chosen as  $x_f = 0.25$ . In Fig. 2(a), where  $K = 0.3$ , the system is stable and the response is visibly damped; in Fig. 2(b), where  $K = 0.75$ , the system is still stable, but the response is visibly less damped. At approximately  $K = 0.8$  the system undergoes a Hopf bifurcation. The case of  $K = 1.5$  shown in Fig. 2(c) shows an initially diverging response that eventually settles onto a Limit Cycle Oscillation. The effect of changing the heat source location can be observed in Fig. 2(d) where the time series for the case  $K = 2.5$  and  $x_f = 0.64$  is provided. The heat source location affects the dominant mode of the system and hence the frequency of the response. Precisely, the frequency of the LCO in Fig. 2(c) is  $\pi$ , while the one in Fig. 2(d) is  $2\pi$ . Note that the heat source typically lies in the first quarter of the tube (Epperlein et al., 2015), and the latter configuration is illustrated here for the sake of exploring different nonlinear responses of the system. The considered bifurcation regimes are summarized in Table 1.

Table 1. Bifurcation regimes of the Rijke tube.

	$K$	$x_f$	Description
$\beta_1$	0.3	0.25	stable (highly damped)
$\beta_2$	0.75	0.25	stable (close to bifurcation point)
$\beta_3$	1.5	0.25	LCOs (first mode dominant)
$\beta_4$	2.5	0.64	LCOs (second mode dominant)

## 4. RESULTS

In this section, the identification method developed in Section 2 is applied to the Rijke tube system for the regimes defined in Table 1. In the first step (Section 2.1), sampling is done via 20 uniformly spaced points ( $q = 20$ ).

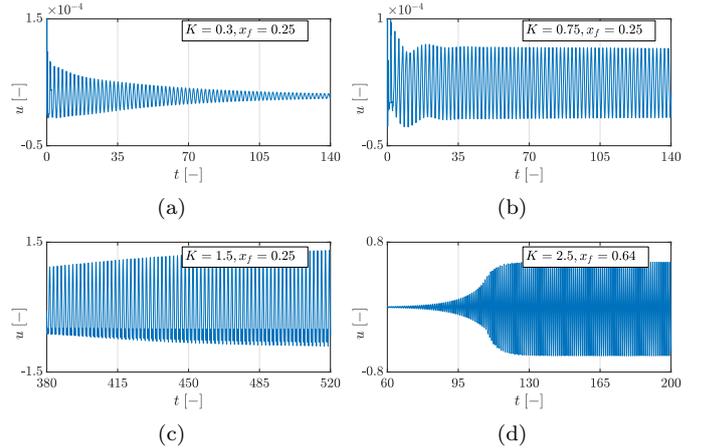


Fig. 2. Velocity at  $x = 0.125$  for different parameters.

For the stable regimes ( $\beta_1, \beta_2$ ), the sampling window is designated as the time interval  $t = 25 - 45$  and for the limit cycle regimes ( $\beta_3, \beta_4$ ), as one period at steady state. Table 2 reports the number of POD modes selected for each regime using different energy thresholds for the truncation of the singular values. Unless otherwise specified, an energy threshold of 99.9% is used in the analyses.

Table 2. Number of modes ( $r_j$ ) of each parameter set ( $\beta_j$ ) for different energy thresholds

Energy threshold	$r_1$	$r_2$	$r_3$	$r_4$
80%	2	3	3	2
92%	3	3	4	2
99.9%	3	3	6	9
99.99%	4	5	11	11

### 4.1 Classification

The performance of the second step of the proposed algorithm (Section 2.2) is investigated here. As a first illustrative example, Fig. 3 shows the output of the classification methods for a measurement vector consisting of  $m = 30$  measurements (15 equally spaced pressure and velocity states) sampled at time instant  $t_m = 970$  in the  $\beta_3$  regime. The implementations of Lasso and Group Lasso are taken respectively from (Friedman et al., 2010) and (Boyd et al., 2011).

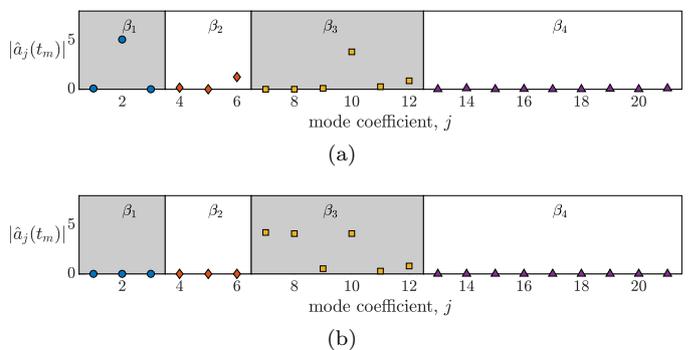


Fig. 3. Classification using (a) Lasso, (b) Group Lasso.

Results indicate that Lasso not only activates modes from different regimes, but also estimates the highest

activated mode in the wrong regime and thus leads to misclassifications, while this is avoided with Group Lasso by activating modes in only one regime.

To quantitatively compare the performance of the two classification approaches, measurements at 100 random time instants are tested for each regime. The resulting confusion matrices (Fig. 4) show the distribution of correct and incorrect classifications in the four bifurcation regimes. It can be noticed that there is a general tendency for misclassification among the regimes  $\beta_1$ ,  $\beta_2$  and  $\beta_3$ , due to the fact that they share similar modes. This is ameliorated in the case of Group Lasso, where the ambiguity is limited to the two stable regimes. Overall, Group Lasso achieves 86% of classification accuracy, while Lasso has a classification accuracy of 76%.

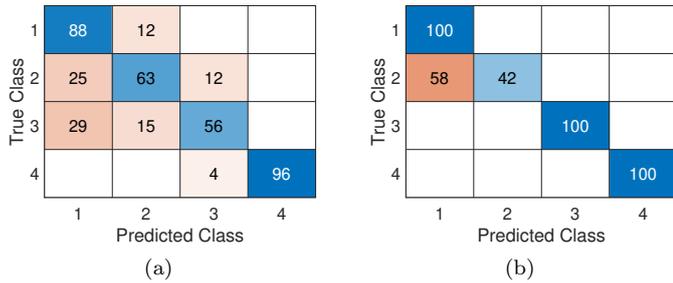


Fig. 4. Confusion matrices for (a) Lasso (b) Group Lasso.

Finally, the effects of the energy threshold used to build the bifurcation library and the number of measurements on the classification accuracy are presented in Table 3. If the number of measurements is low, a lower threshold is better for high classification accuracy, but if the number of measurements is high, a higher threshold can help overcoming the misclassification issues discussed above.

Table 3. Classification accuracy.

Lasso	Energy Threshold			
	80%	92%	99.9%	99.99%
Number of measurements	80%	92%	99.9%	99.99%
6	62%	58%	51%	40%
14	75%	71%	62%	57%
30	73%	62%	76%	61%
Group Lasso	Energy Threshold			
	80%	92%	99.9%	99.99%
Number of measurements	80%	92%	99.9%	99.99%
6	81%	63%	52%	40%
14	75%	64%	82%	77%
30	76%	70%	86%	67%

#### 4.2 State estimation

Having classified the bifurcation regime to which the system's response belongs, the initial condition  $\hat{a}_0$  can be computed (Section 2.3). The accuracy of this procedure is investigated by looking at the reconstruction of the full-state vector using the relationship  $\hat{Y} \approx \hat{\Psi}_j \hat{a}_0$ , which can be compared with the true one. The analyses will focus on the  $\beta_3$  regime with six modes ( $r_3 = 6$ ). These are also representative of the results for the other regimes.

Qualitative aspects are first analyzed by means of Fig. 5, where the estimation at  $t_m = 970$  is investigated by considering three different scenarios. Fig. 5(a) considers a non-

nal case where 30 noise-free measurements are available. The LS estimate (7) matches almost perfectly the true state vector, similarly for ridge (8) and weighted ridge (10) formulations. Fig. 5(b) shows the reconstruction results when  $\hat{Y}$  is contaminated with i.i.d. zero-mean Gaussian noise with variance  $\sigma^2 = 0.36$ . In this case, the weighted ridge method outperforms the LS and ridge methods by producing the closest reconstruction to the true state, and preserving the discontinuity of the velocity field at the heat source location. Fig. 5(c) examines the noise-free underdetermined case where only  $m = 4$  measurements are available. The estimate obtained with the LN approach (9) is markedly inaccurate, while solving the same problem by using the singular values as weights (11) considerably improves the result.

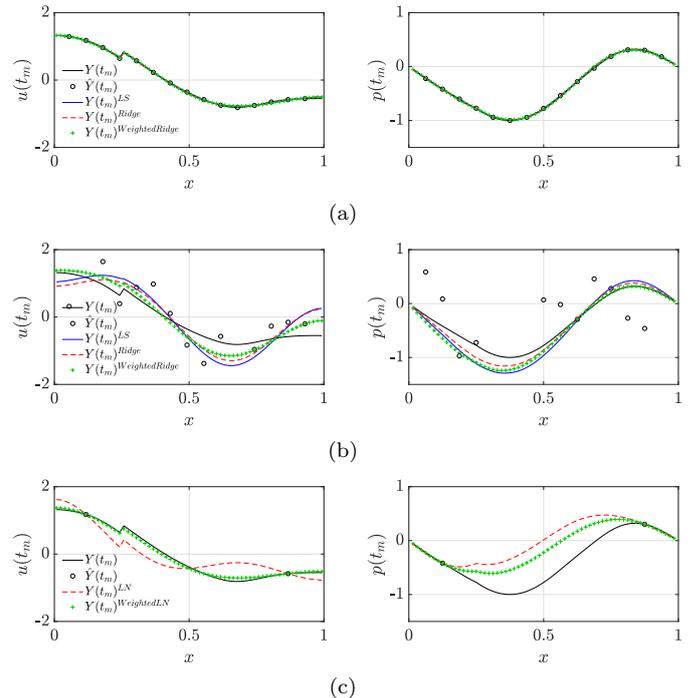


Fig. 5. State reconstruction (a)  $m=30$  noise-free; (b)  $m=30$  noisy; (c)  $m=4$  noise-free.

A quantitative analysis of the accuracy of the algorithms for the state reconstruction is performed using the last two scenarios, because in the noise-free overdetermined case the methods produce, as already commented, the same results and the error monotonically decreases with increasing number of modes. The effect of changing the number of modes  $r_j$  in the library is studied by randomly sampling 100 time instants at which the reconstruction is performed, and plotting the relative error given by the Euclidean distance between estimated and true state normalized by the Euclidean norm of the true state. In the following analyses, variations due to the different time instants are captured using the box plot representation. In the noisy overdetermined case (Fig. 6(a)), the LS method shows overfitting, as observed by the increase in error as  $r_j$  increases. This is to some extent ameliorated by the regularized methods, especially when weighted extensions are employed. The advantage of using weighted regularizations is further showcased in Fig. 6(b), where the

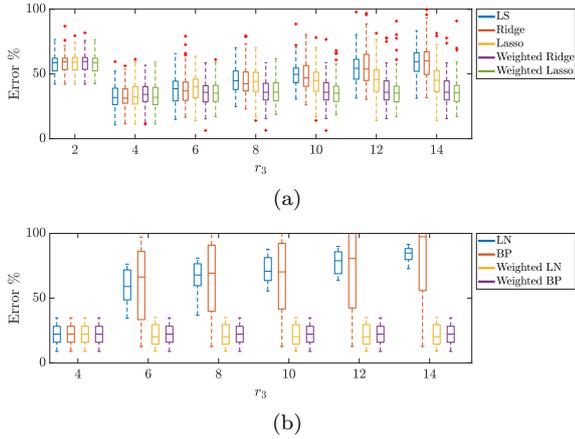


Fig. 6. State reconstruction error as a function of  $r_j$ : (a)  $m=30$  noisy; (b)  $m=4$  noise-free.

underdetermined case is analyzed. The standard LN and BP approaches suffer high errors when a larger number of POD modes is employed, while the solutions using weighted norms prove successful. The reason is that less dominant modes will have a higher relative impact on the penalized norm, and thus a priori (data-driven) information on the relevance of the single modes further regularize the underdetermined problem.

#### 4.3 Simulation

Once estimated the initial state  $\hat{a}_0$  of the low-order projected dynamics, full-state trajectories can be predicted (12). For the same two scenarios analyzed earlier (but only considering ridge and LN approach for clarity), starting from 100 random time instants, the dynamic evolution in an interval of 5 time units is computed and the error with respect to the true state sequence is evaluated (Fig. 7). To compute the relative error, the snapshot matrices of the true and predicted state sequences are constructed and the ratio between the Frobenius norm of their differences and the Frobenius norm of the matrix with the true state sequence is computed. The results labelled *Galerkin* are obtained using the model-based Galerkin projection proposed in (Juniper, 2011) and making use of pre-defined sinusoidal modes. The observed superior performance of the proposed solutions testifies to the advantages of using a POD-based Galerkin projection where the modes are computed from data, rather than using a pre-defined set of modes which might be less representative since they do not take into account any specific feature of the system.

## 5. CONCLUSION

An approach to identify the nonlinear regimes of a dynamical system from partially observed states and predict the future state evolution using low-order dynamics is proposed. It combines Proper Orthogonal Decomposition-based dimensionality reduction techniques and regularization approaches from statistical learning. The novel combination of these two methodologies is believed to be particularly useful when studying complex nonlinear dynamics featuring similar modes across different response regimes and/or when only very limited or noisy measurements are available. Detailed investigations carried out

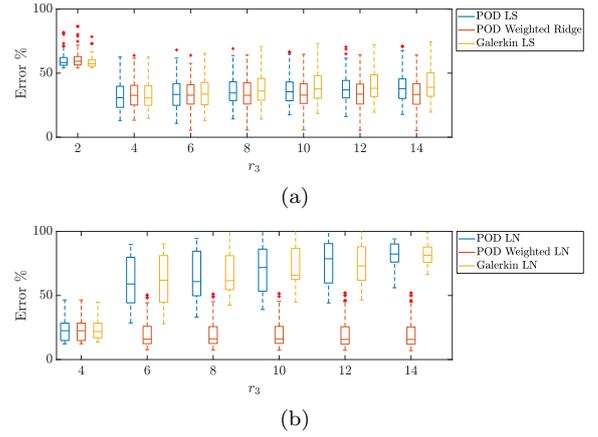


Fig. 7. State simulation error as a function of  $r_j$ : (a)  $m=30$  noisy; (b)  $m=4$  noise-free.

on a thermoacoustic benchmark problem confirm these claims and showcase improved performance with respect to model-based or standard estimation techniques.

## REFERENCES

- Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1).
- Brunton, S.L., Tu, J.H., Bright, I., and Kutz, J.N. (2014). Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems. *SIAM Journal on Applied Dynamical Systems*, 13(4).
- Epperlein, J.P., Bamieh, B., and Astrom, K.J. (2015). Thermoacoustics and the rijke tube: Experiments, identification, and modeling. *IEEE Control Systems*, 35(2).
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1).
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning*. Springer New York.
- Holmes, P., Lumley, J.L., Berkooz, G., and Rowley, C.W. (2012). *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press.
- Iannelli, A., Lowenberg, M., and Marcos, A. (2020). Computation of bifurcation margins based on robust control concepts. *SIAM Journal of Applied Dynamical Systems*, 19(3).
- Juniper, M.P. (2011). Triggering in the horizontal rijke tube: non-normality, transient growth and bypass transition. *Journal of Fluid Mechanics*, 667, 272–308.
- Sayadi, T., Le Chenadec, V., Schmid, P.J., Richecoeur, F., and Massot, M. (2014). Thermoacoustic instability - a dynamical system and time domain analysis. *Journal of Fluid Mechanics*, 753.
- Schmid, P. and Henningson, D. (2001). *Stability and Transition in Shear Flows*. Springer-Verlag New York.
- Swift, G. (2007). *Thermoacoustics*. Springer.
- Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1).