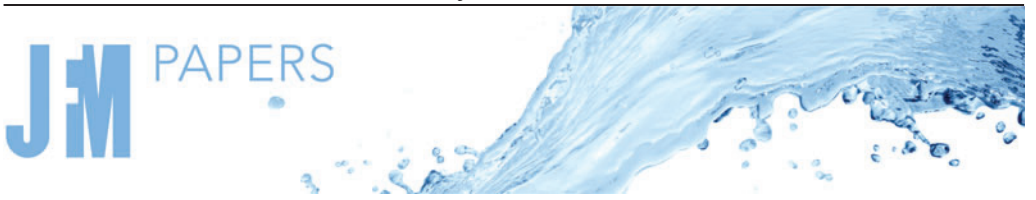
Real-time thermoacoustic data assimilation

J. Fluid Mech. (2022), *vol*. 948, A35, doi:10.1017/jfm.2022.653



Real-time thermoacoustic data assimilation

A. Nóvoa[1] **and L. Magri**[1,2,3,4,†]

[1]Cambridge University Engineering Department, Trumpington St, Cambridge CB2 1PZ, UK

[2]Aeronautics Department, Imperial College London, Exhibition Road, London SW7 2AZ, UK

[3]The Alan Turing Institute, 96 Euston Rd, London NW1 2DB, UK

[4]Institute for Advanced Study, TU Munich, Lichtenbergstraße 2a, 85748 Garching, Germany (visiting)

(Received 11 June 2021; revised 8 June 2022; accepted 20 July 2022)

Low-order thermoacoustic models are qualitatively correct, but typically, they are quantitatively inaccurate. We propose a time-domain bias-aware method to make qualitatively low-order models quantitatively (more) accurate. First, we develop a Bayesian ensemble data assimilation method for a low-order model to self-adapt and self-correct any time that reference data become available. Second, we apply the methodology to infer the thermoacoustic states and heat-release parameters on the fly without storing data (real time). We perform twin experiments using synthetic acoustic pressure measurements to analyse the performance of data assimilation in all nonlinear thermoacoustic regimes, from limit cycles to chaos, and interpret the results physically. Third, we propose practical rules for thermoacoustic data assimilation. An increase, reject, inflate strategy is proposed to deal with the rich nonlinear behaviour; and physical time scales for assimilation are proposed in non-chaotic regimes (with the Nyquist–Shannon criterion) and in chaotic regimes (with the Lyapunov time). Fourth, we perform data assimilation using data from a higher-fidelity model. We introduce an echo state network to estimate in real time the forecast bias, which is the model error of the low-fidelity model. We show that: (i) the correct acoustic pressure, parameters, and model bias can be inferred accurately; (ii) the learning is robust as it can tackle large uncertainties in the observations (up to 50 % of the mean values); (iii) the uncertainty of the prediction and parameters is naturally part of the output; and (iv) both the time-accurate solution and statistics can be inferred successfully. Data assimilation opens up new possibility for real-time prediction of thermoacoustics by combining physical knowledge and experimental data synergistically.

Key words: instability control, acoustics, nonlinear dynamical systems

† Email address for correspondence: lm547@cam.ac.uk

1. Introduction

When the heat released by a heat source, such as a flame, is sufficiently in phase with the acoustic waves of a confined environment, such a gas turbine or a rocket, thermoacoustic oscillations may occur (Rayleigh 1878). Thermoacoustic oscillations manifest themselves as large-amplitude vibrations, which can be detrimental to gas-turbine reliability (e.g. Lieuwen 2012), and can be destructive in high-power-density motors such as rocket engines (e.g. Culick & Kuentzmann 2006). The objective of manufacturers is to design devices that are thermoacoustically stable, which is the goal of optimisation, and suppress a thermoacoustic oscillation if it occurs, which is the goal of control (e.g. Magri 2019). Both optimisation and control rely on a mathematical model, which provides predictions on the key physical variables, such as the acoustic pressure and the heat-release rate. The accurate prediction of thermoacoustic oscillations, however, remains one of the most challenging problems faced by power generation, heating and propulsion manufacturers (e.g. Dowling & Morgans 2005; Noiray *et al.* 2008; Lieuwen 2012; Poinsot 2017; Juniper & Sujith 2018).

The prediction of thermoacoustic dynamics – even in simple systems – is challenging for three reasons. First, thermoacoustics is a multi-physics phenomenon. For a thermoacoustic oscillation to occur, three physical subsystems (flame, acoustics and hydrodynamics) interact constructively with each other (e.g. Lieuwen 2012; Magri 2019). Second, thermoacoustics is a nonlinear phenomenon (e.g. Sujith & Unni 2020). In general, the flame's heat release responds nonlinearly to acoustic perturbations (Dowling 1999), and the hydrodynamics are typically turbulent (e.g. Huhn & Magri 2020*b*). Third, thermoacoustics is sensitive to perturbations to the system. In the linear regime, small changes to the system's parameters, such as the flame time delay, can cause arbitrarily large changes of the eigenvalue growth rates at exceptional points (Mensah *et al.* 2018; Orchini *et al.* 2020). In the nonlinear regime, small changes to the system's parameters can cause a variety of nonlinear bifurcations of the solution. As a design parameter is varied in a small range, thermoacoustic oscillations may become chaotic, by either period doubling or Ruelle–Takens–Newhouse scenarios (Gotoda *et al.* 2011, 2012; Kabiraj & Sujith 2012; Kashinath, Waugh & Juniper 2014; Orchini, Illingworth & Juniper 2015; Huhn & Magri 2020*b*), or by intermittency bifurcations scenarios (Nair, Thampi & Sujith 2014; Nair & Sujith 2015). The rich bifurcation behaviour has an impact on the effectiveness of control strategies, which may work for periodic oscillations with a dominant frequency, but may not work as effectively for multi-frequency oscillations. Additionally, unpredictable changes in the operating conditions and turbulence, which can be modelled as random phenomena (Nair & Sujith 2015; Noiray 2017), increase the uncertainty on the prediction of the quantities of interest.

Thermoacoustics can be modelled with a hierarchy of assumptions and computational costs. Large-eddy simulations make assumptions only on the finer flow scales, which makes the final solution high-fidelity, but computationally expensive (Poinsot 2017). Euler and Helmholtz solvers compute the acoustics that evolve on a prescribed mean flow, which makes the solution medium-fidelity and computationally less expensive than turbulent simulations (e.g. Nicoud *et al.* 2007). Commonly, this is achieved with flame models, which capture the heat-release response to acoustic perturbations with transfer functions (e.g. Noiray *et al.* 2008; Silva *et al.* 2013) and distributed time delays (Polifke 2020). Other medium-fidelity and medium-cost methods are based on flame-front tracking (e.g. Pitsch & de Lageneste 2002) and simple chemistry models (e.g. Magri & Juniper 2014), to name only a few. On the other hand, low-order models based on travelling waves and standing waves (Dowling 1995) provide low-fidelity solutions, but with low computational cost.

These low-order models capture only the dominant physical mechanisms, which are the flame time delay, the flame strength (or index), and the damping. Low-order models, which are the subject of this study, are attractive to practitioners because they provide quick estimates on the quantity of interest. Along with modelling, accurate experimental data are becoming more accessible and available (O'Connor, Acharya & Lieuwen 2015). To monitor the thermoacoustic behaviour in both real engines and academic rig (such as the Rijke tube), the pressure is measured experimentally by microphones (Lieuwen & Yang 2005; Kabiraj *et al.* 2012*a*). Microphones sample the pressure amplitude at typically high rates, which generates large datasets in real time. Except when required for diagnostics and *a posteriori* parameter identification (among many others, Polifke *et al.* 2001; Schuermans 2003; Lieuwen & Yang 2005; Noiray *et al.* 2008; Noiray 2017; Polifke 2020), the data are useful if they can be used in real time, i.e. on the fly, to correct (or update) our knowledge of the thermoacoustic states. The sequential assimilation method that we develop bypasses the need to store data, which enables the real-time assimilation of data as well as on-the-fly parameter estimation.

To summarise, in thermoacoustics, we have three ingredients to improve the design: (i) a human being, who identifies the physical mechanisms that need to be modelled depending on the objectives and resources; (ii) a mathematical model, which provides estimates of the physical states; and (iii) experimental data, which provide a quantitative measure of the system's observables. A model is good if the human being identifies the physical mechanisms needed to formulate a mathematical model that provides the system's states compatibly with the experimental data. The overarching objective of this paper is to propose a method to make qualitatively low-order models quantitatively (more) accurate every time that reference data becomes available. The ingredients for this are: a physical low-order model, which provides the states; data, which provide the observables; and a statistical method, which finds the most likely model by assimilating the data in the model. In weather forecasting, this process is known as data assimilation (Sasaki 1955). Data assimilation techniques have been applied to oceanographic studies (Eckart 1960), aerospace control (Gelb 1974), robotics, geosciences and cognitive sciences (Reich & Cotter 2015), to name only a few. Data assimilation is a principled method, which, in contrast to traditional machine learning, uses a physical model to provide a prediction on the solution (the forecast), which is updated when observations become available to provide a corrected state (the analysis) (Reich & Cotter 2015). The analysis is an estimator of the physical state (the true state), which is more accurate than the forecast.

1.1. *Data assimilation*

Data assimilation methods can be divided into two main approaches (Lewis, Lakshmivarahan & Dhall 2006): (i) variational and (ii) statistical assimilation methods. Variational data assimilation requires the minimisation of a cost functional – e.g. a Mahalanobis (semi)norm – in terms of a control variable to obtain a single optimal analysis state (Bannister 2017). The most common variational methods are 3D-VAR and 4D-VAR, which are used widely in weather centres such as the Met Office in the UK or the European Centre for Medium-Range Weather Forecasts, and in academic research (Bannister 2008). In thermoacoustics, variational data assimilation was introduced by Traverso & Magri (2019), who found the optimal thermoacoustic states given reference data from pressure observations. Because variational methods need batches of data, they are not naturally suited to real-time inference. On the other hand, statistical data assimilation combines concepts of probability and estimation theory. The aim of statistical data assimilation is to compute the probability distribution function of a numerical model to combine it

statistically with data from observations. Because the probability distribution function is high-dimensional, the practitioner is often interested in capturing only the first and second statistical moments of it. In reduced-order modelling, this was achieved in flame tracking methods by Yu, Juniper & Magri (2019), who implemented ensemble Kalman filters and smoothers to learn the flame parameters on the fly. In high-fidelity methods in reacting flows, data assimilation with ensemble Kalman filters have been applied in large-eddy simulation of premixed flames to predict local extinctions in a jet flame (Labahn *et al.* 2019), and under-resolved turbulent simulation to predict autoignition events (Magri & Doan 2020). The ensemble Kalman filter has also been applied successfully to non-reacting flow systems that show high nonlinearities, such as the estimation of turbulent near-wall flows (Colburn, Cessna & Bewley 2011), uncertainties in Reynolds-averaged Navier–Stokes (RANS) equations (Xiao *et al.* 2016), and aerodynamic flows (da Silva & Colonius 2018). Statistical data assimilation based on Bayesian methods, which enable real-time prediction in contrast to variational methods, was introduced in thermoacoustics by Nóvoa & Magri (2020).

1.2. *Model bias*

Commonly, data assimilation methods are derived under the assumption that forecast errors are random with zero mean (Evensen 2009), or, in other words, the error is unbiased. However, in addition to state and parameter uncertainties, low-order models are affected by model uncertainty, which manifests as an error bias. Modelling the model bias is an active research area. To produce an unbiased analysis, both forecast and observation biases need to be estimated (Dee & da Silva 1998). Friedland (1969) developed the separate Kalman filter to estimate the bias, which is a two-stage sequential filtering process that addresses the estimation of a constant bias term, but its application is limited to linear processes. Drécourt, Madsen & Rosbjerg (2006) extended the implementation of the separate Kalman filter and compared it to the coloured noise Kalman filter, which augments the state vector with an auto-retrogressive model that describes the bias. They propose a feedback implementation of these methods, which allows a time-correlated representation of the bias, but the accuracy is limited by the prescribed model of the bias. Houtekamer & Zhang (2016) reviewed techniques that involve multi-physical parametrisation to reduce the model bias in atmospheric data assimilation, which are more computationally expensive than single-model approaches. More recently, there have been studies that estimated model-related errors in the assimilation cycle. Rubio, Chamoin & Louf (2019) proposed a data-based approach to model the error that arises from the truncation of proper generalised decomposition modes, which was integrated into a Bayesian inference method. Da Silva & Colonius (2020) proposed a low-rank representation of the observation discretisation bias, based as well on an auto-regressive model. They performed parameter estimation with an ensemble Kalman filter to calibrate the parameters of the auto-regressive model. These works modelled successfully the truncation and discretisation bias, but they did not model the physical model bias, which is a more general form of error that will be tackled in this paper. The estimation of a nonlinear dynamical state in the presence of a model bias remains an open problem. We propose a framework to obtain an unbiased analysis in thermoacoustic low-fidelity models by inferring the model bias. This consists of the combination of data assimilation with a recurrent neural network, which infers the model error of the low-fidelity model of the thermoacoustic system.

Recurrent neural networks are data-driven techniques that are designed to learn temporal correlations in time series (Rumelhart, Hinton & Williams 1986), with a variety of

applications in time series forecasting. In fluid mechanics, recurrent neural networks have been employed to model unsteady flow around bluff bodies (Hasegawa *et al.* 2020), and as an optimisation tool for gliding control (Novati, Mahadevan & Koumoutsakos 2019). Among recurrent neural networks, echo state networks based on reservoir computing, which are universal approximators (Grigoryeva & Ortega 2018), proved to be successful in learning nonlinear correlations in data (Maass, Natschläger & Markram 2002; Jaeger & Haas 2004) and ergodic properties (Huhn & Magri 2022). Training an echo state network consists of a simple linear regression problem. This is a simple and computationally cheap task with respect to the back propagation required in other architectures (such as long short-term memory networks), which makes echo state networks suited to real-time assimilation. Because no gradient descent is necessary, vanishing/exploding gradient problems do not occur in echo state networks. In chaotic flows, echo state networks have been used, for instance, to learn and optimise the time average of thermoacoustic dynamics (Huhn & Magri 2020*a*, 2022), to predict turbulent dynamics with physical constraints (Doan, Polifke & Magri 2021), and to predict the statistics of extreme events in turbulent flows (Racca & Magri 2021). In this paper, we propose to model the model bias with an echo state network, which is a more versatile and general tool than auto-regressive models (Aggarwal 2018, p. 306).

1.3. *Objectives and structure*

The objectives of this paper are fivefold. First, we develop a sequential data assimilation for a low-order model to self-adapt and self-correct any time that reference data become available. The method, which is based on Bayesian inference, provides the maximum *a posteriori* estimate model prediction, i.e. the most likely prediction. Second, we apply the methodology to infer the thermoacoustic states and heat-release parameters on the fly without storing data. Third, we analyse the performance of the data assimilation algorithm on a twin experiment with synthetic data, and interpret the results physically. Fourth, we propose practical rules for thermoacoustic data assimilation. Fifth, we extend the data assimilation method to account for a biased thermoacoustic model. This method is tested by assimilating observations from a model with non-ideal boundary conditions, a mean flow and the simulation of the flame front with a kinematic model (Dowling 1999). The simulation of the flame dynamics is suitable for a time-domain approach, and it overcomes the limitations of flame response models.

The paper is structured as follows. Section 2 provides a description of the nonlinear thermoacoustic model with the data assimilation technique and its implementation for thermoacoustics. Section 3 presents the method developed for state and parameter estimation. Section 4 presents the method developed for combining data assimilation with an echo state network. Section 5 presents the nonlinear characterisation of the thermoacoustic dynamics. Section 6 shows the results for non-chaotic regimes, whereas § 7 shows and discusses the results for chaotic solutions. Section 8 shows and discusses the results for unbiased state and parameter estimation for a high-fidelity limit cycle. A final discussion and conclusions end the paper in § 9.

2. Thermoacoustic data assimilation

We consider a nonlinear thermoacoustic model, $\mathcal{T}$, as

$$\mathcal{T}(\boldsymbol{\psi}, \boldsymbol{\alpha}, \boldsymbol{y}, \boldsymbol{F}, \boldsymbol{\delta}) = 0, \tag{2.1}$$

$$\mathcal{G}(\boldsymbol{\psi}) = \boldsymbol{y} + \boldsymbol{\epsilon}, \tag{2.2}$$

where $\boldsymbol{\psi}$ is the state of the system, $\boldsymbol{\alpha}$ is the vector of the system's parameters, $\boldsymbol{y}$ are the observables from reference data, $\boldsymbol{F}$ is a nonlinear operator, which, in general, is differential, $\boldsymbol{\delta}$ is a model error, $\mathcal{G}$ is a nonlinear map from the state space to the observable state, and $\boldsymbol{\epsilon}$ is the observation error. The thermoacoustic problem on which we focus is: 'given some data (observables) $\boldsymbol{y}$, and a mathematical model $\mathcal{T}$, what are the most likely physical states $\boldsymbol{\psi}$ and parameters $\boldsymbol{\alpha}$ of the system?' To answer this, we use a Bayesian approach in the well-posed maximum *a posteriori* estimation framework. Although the framework is versatile, in the following subsections, we specify the low-order model $\boldsymbol{F}$, the data $\boldsymbol{y}$, and the data assimilation approach.

2.1. *Qualitative nonlinear thermoacoustic model*

The system consists of an open-ended tube containing a heat source, such as a flame or an electrically heated gauze. Because the tube is sufficiently long with respect to the diameter, the cut-on frequency is such that only longitudinal acoustic waves propagate. This is known as the Rijke tube, which is a common laboratory-scale device that has been employed in a variety of fundamental studies (Heckl 1990; Balasubramanian & Sujith 2008; Juniper 2011; Magri *et al.* 2013). This device is represented in figure 1. The Rijke model used in this work is described by Balasubramanian & Sujith (2008) and Juniper (2011). The flow is assumed to be a perfect gas, the mean flow is sufficiently slow such that its effects are neglected in the acoustic propagation, and viscous and body forces are neglected. The acoustics are governed by the dimensionless linearised momentum and energy conservation equations

$$\frac{\partial u'}{\partial t} + \frac{\partial p'}{\partial x} = 0, \tag{2.3a}$$

$$\frac{\partial p'}{\partial t} + \frac{\partial u'}{\partial x} = \dot{Q}\delta(x - x_f) - \zeta p', \tag{2.3b}$$

where u' is the acoustic velocity, p' is the acoustic pressure, $\dot{Q}$ is the heat-release rate, x_f is the flame location, δ is the Dirac delta distribution, which models the heat source as a point source (compact assumption), and ζ is the damping factor, which encapsulates the acoustic energy radiated from both ends of the duct, and the thermoviscous losses in boundary layers. The non-dimensional heat-release rate perturbation, $\dot{Q}$, is modelled with a qualitative nonlinear time-delayed model (Heckl 1990):

$$\dot{Q} = \beta\left[\sqrt{\left|\tfrac{1}{3} + u'_f(t-\tau)\right|} - \sqrt{\tfrac{1}{3}}\right], \tag{2.4}$$

where β is the strength of the source, u'_f is the acoustic velocity at the flame location, and τ is the time delay. The heat-release rate is a key thermoacoustic parameter for the system's stability. The dimensionless variables in (2.3)–(2.4) and the dimensional variables (with $\tilde{}$) are related as $x = \tilde{x}/\tilde{L}_0$, where $\tilde{L}_0$ is the length of the tube, $t = \tilde{t}\tilde{c}_0/\tilde{L}_0$ (where $\tilde{c}_0$ is the mean speed of sound), $u' = \tilde{u}'/\tilde{c}_0$, $\rho' = \tilde{\rho}'/\tilde{\rho}_0$ (where $\tilde{\rho}_0$ is the mean density), $p' = \tilde{p}'/(\tilde{\rho}_0\,\tilde{c}_0^2)$, $\dot{Q} = \tilde{\dot{Q}}'(\gamma-1)/(\tilde{\rho}_0\,\tilde{c}_0^3)$ (where γ is the heat capacity ratio), and $\delta(x - x_f) = \tilde{\delta}(\tilde{x} - \tilde{x}_f)\tilde{L}_0$. The open-ended boundary conditions are ideal, which means that the acoustic pressure is zero, i.e. $p' = 0$ at $x = \{0, 1\}$. By separation of variables, the acoustic velocity and pressure

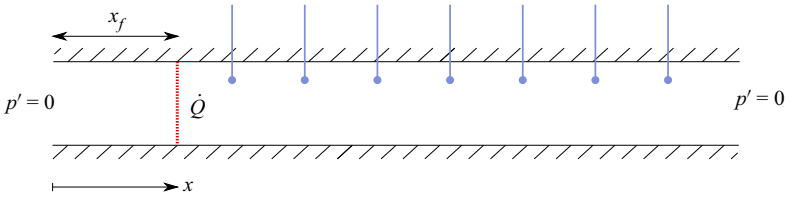


Figure 1. Schematic of an open-ended duct with a heat source (also known as the Rijke tube). The heat released by the compact heat source is indicated by the vertical dotted line. The blue vertical lines indicate microphones located equidistantly.

are decomposed into Galerkin modes as (Zinn & Lores 1971)

$$u'(x,t) = \sum_{j=1}^{N_m} \eta_j(t)\cos(j\pi x), \quad p'(x,t) = -\sum_{j=1}^{N_m} \frac{\dot{\eta}_j(t)}{j\pi}\sin(j\pi x), \tag{2.5a,b}$$

where $\cos(j\pi x)$ and $\sin(j\pi x)$ are the eigenfunctions of the acoustic velocity and pressure, respectively, when $\zeta = 0$ and $\dot{Q} = 0$, and N_m is the number of acoustic modes kept in the decomposition. Substituting (2.5*a*,*b*) into (2.3), multiplying (2.3*b*) by $\sin(k\pi x)$, and integrating over $x = [0, 1]$, yields the governing ordinary differential equations, which represent physically a set of nonlinearly coupled oscillators:

$$\frac{\mathrm{d}\eta_j}{\mathrm{d}t} - j\pi\left(\frac{\dot{\eta}_j}{j\pi}\right) = 0, \tag{2.6a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{\eta}_j}{j\pi}\right) + j\pi\eta_j + \zeta_j\frac{\dot{\eta}_j}{j\pi} + 2\dot{Q}\sin(j\pi x_f) = 0, \tag{2.6b}$$

where the damping term is defined by modal components $\zeta_j = C_1 j^2 + C_2\sqrt{j}$, which is motivated physically in Landau & Lifshitz (1987). The damping coefficients C_1 and C_2 are assumed to be constant. For reasons that will be explained in § 2.3, we introduce an advection equation to eliminate mathematically the time-delayed velocity term (Huhn & Magri 2020*b*):

$$\frac{\partial v}{\partial t} + \frac{1}{\tau}\frac{\partial v}{\partial X} = 0, \quad 0 \leqslant X \leqslant 1, \tag{2.7}$$

where v is a dummy variable that travels with non-dimensional velocity τ^{-1} in a dummy spatial domain X such that

$$u'_f(t-\tau) = v(X=1,t), \quad u'_f(t) = v(X=0,t). \tag{2.8a,b}$$

Equations (2.8*a*,*b*) are discretised with a Chebyshev method (Trefethen 2000) with $N_c + 1$ points in the interval $0 \leqslant X \leqslant 1$.

In a state-space notation, the thermoacoustic problem is governed by

$$\left.\begin{aligned} \frac{\mathrm{d}\boldsymbol{\psi}}{\mathrm{d}t} &= \boldsymbol{F}(\boldsymbol{\alpha};\boldsymbol{\psi}), \quad \boldsymbol{\psi}(t=0) = \boldsymbol{\psi}_0, \\ \boldsymbol{y} &= \boldsymbol{M}(\boldsymbol{x})\,\boldsymbol{\psi}, \end{aligned}\right\} \tag{2.9}$$

where the state vector $\boldsymbol{\psi} \equiv (\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}, \boldsymbol{v})^{\mathrm{T}} \in \mathbb{R}^{2N_m+N_c}$ is the column concatenation of the acoustic amplitudes $\boldsymbol{\eta} \equiv (\eta_1, \eta_2, ..., \eta_{N_m})^{\mathrm{T}} \in \mathbb{R}^{N_m}$ and $\dot{\boldsymbol{\eta}} \equiv (\dot{\eta}_1/\pi, \dot{\eta}_2/(2\pi), \ldots, \dot{\eta}_{N_m}/$

$(N_m\pi))^{\mathrm{T}} \in \mathbb{R}^{N_m}$, and the dummy velocity variables $\boldsymbol{v} \equiv (v_1, v_2, \ldots, v_{N_c})^{\mathrm{T}} \in \mathbb{R}^{N_c}$, which arise from the discretisation of (2.7). Also, the thermoacoustic parameters are contained in the vector $\boldsymbol{\alpha} = (\beta, \tau, \zeta)^{\mathrm{T}} \in \mathbb{R}^{N_P}$; $\boldsymbol{F}$ represents the nonlinear operator that consists of (2.6*a*), (2.6*b*) and (2.7), $\boldsymbol{F} : \mathbb{R}^{2N_m+N_c+N_P} \to \mathbb{R}^{2N_m+N_c}$; and $\boldsymbol{M}(\boldsymbol{x})$ is the measurement operator, which maps the state to the observable space at $\boldsymbol{x}$. The expression of the measurement operator depends on the nature of the observables being assimilated, as explained in § 3. To work with a reduced-order model that captures qualitatively the essential dynamics, we use $N_m = 10$ acoustic modes. For the advection equation, $N_c = 10$ ensures numerical convergence (Huhn & Magri 2020*b*). The number of degrees of freedom of the reduced-order model is $N = 2N_m + N_c = 30$. The initial value problem (2.9) is solved with an automatic step size control method that combines fourth- and fifth-order Runge–Kutta methods (Shampine & Reichelt 1997).

2.2. *Qualitative and quantitative accuracy*

We say that a low-order model is qualitatively correct when it captures the key physical parameters/mechanisms (e.g. the time delay). Although a low-order model may be physically motivated, it is subject to three sources of errors: (i) uncertainty in the state; (ii) uncertainty in the parameters; and (iii) bias in the model, i.e. the low-order equation does not contain all the terms necessary to model the phenomenon (the model bias is equivalently referred to as the model error). Data assimilation methods combine the forecast of a low-order model with observations from either real experiments or high-fidelity simulations, which reduces the bias in the state (§ 3.1) and/or in the parameters of the model (§ 3.2). However, traditional data assimilation methods do not tackle the model bias because they assume that the forecast model is unbiased. In § 4, we propose an echo state network as a method to estimate the model bias, thereby closing the low-order model equations in the data assimilation. In summary, the aim of data assimilation is to make a qualitative accurate model more quantitatively correct.

2.3. *Data assimilation*

Data assimilation optimally combines the prediction from an imperfect model with data from observations to improve the knowledge of the system's state. The updated solution (analysis) optimally combines the information from the observations $\boldsymbol{y}$ and the model solution (forecast) with their uncertainties. In order to (i) update the system's knowledge any time that data become available, and (ii) not store the data during the entire operation, we assimilate sequentially assuming that the process is a Markovian process. The concept of Bayesian update is key to this process, as explained in § 2.3.1.

2.3.1. *Bayesian update*

In a Bayesian framework, we quantify our confidence in a model by a probability measure. Hence we update our confidence in the model predictions every time we have reference data from observations. The rigorous framework to achieve this is probability theory, as explained in Cox's theorem (Jaynes 2003).

To set a probabilistic framework at time $t = t_k$, the state $\boldsymbol{\psi}_k$ and reference observations $\boldsymbol{y}_k$ are assumed to be realisations of their corresponding random variables acting on the sample spaces $\Omega_{\boldsymbol{\psi}} = \mathbb{R}^{2N_m+N_c}$ and $\Omega_{\boldsymbol{y}} = \mathbb{R}^{N_y}$. Because we transformed the time-delayed problem into an initial value problem, the solution of (2.9) at the present depends on the solution at the previous time step only. In other words, we transformed a non-Markovian

system into a Markovian system, which simplifies the design of the Bayesian update. We quantify our confidence in a quantity through a probability $\mathcal{P}$:

$$\boldsymbol{\psi}_k \sim \mathcal{P}(\boldsymbol{\psi}_k|\boldsymbol{\psi}_{k-1}, \boldsymbol{\alpha}, \boldsymbol{F}), \quad \boldsymbol{y}_k \sim \mathcal{P}(\boldsymbol{y}_k|\boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{F}), \tag{2.10a,b}$$

where | denotes that the quantity on the left is conditioned on the knowledge of the quantities on the right. The leftmost probability answers the question: 'given a model $\boldsymbol{F}$, a set of parameters $\boldsymbol{\alpha}$, and the state $\boldsymbol{\psi}_{k-1}$, what is the probability that the state takes the value $\boldsymbol{\psi}_k$?' The rightmost probability answers the question: 'if we forecast the state $\boldsymbol{\psi}_k$ from the model, what is the probability that we observe $\boldsymbol{y}_k$?' We assume that the observations are statistically independent and uncorrelated with respect to the forecast. To update our knowledge of the system, the prior knowledge from the reduced-order model and the reference observations are combined through Bayes' rule

$$\mathcal{P}(\boldsymbol{\psi}_k|\boldsymbol{y}_k, \boldsymbol{\alpha}, \boldsymbol{F}) = \frac{\mathcal{P}(\boldsymbol{y}_k|\boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{F})\,\mathcal{P}(\boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{F})}{\mathcal{P}(\boldsymbol{y}_k, \boldsymbol{\alpha}, \boldsymbol{F})}. \tag{2.11}$$

First, $\mathcal{P}(\boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{F})$ is the prior, which measures the knowledge of our system prior to observing $\boldsymbol{y}_k$. The prior evolves through the Chapman–Kolmogorov equation (Jazwinski 2007), which involves multi-dimensional integrals. To solve the Chapman–Kolmogorov equation numerically, we use an ensemble method by integrating the model equations (§ 2.3.2), which provide a forecast on the state. Second, $\mathcal{P}(\boldsymbol{y}_k|\boldsymbol{\psi}_k, \boldsymbol{\alpha}, \boldsymbol{F})$ is the likelihood (2.10*b*), which measures the confidence that we have in our model prediction. The likelihood is prescribed (see § 2.3.2). Third, $\mathcal{P}(\boldsymbol{y}_k, \boldsymbol{\alpha}, \boldsymbol{F})$ is the evidence, which is the probability that the observable takes on the value $\boldsymbol{y}_k$. This can be prescribed from the knowledge of the experimental uncertainties. Finally, $\mathcal{P}(\boldsymbol{\psi}_k|\boldsymbol{y}_k, \boldsymbol{\alpha}, \boldsymbol{F})$ is the posterior, which measures the knowledge that we have on the state, $\boldsymbol{\psi}_k$, after we have observed $\boldsymbol{y}_k$. Here, we will select the most probable value of $\boldsymbol{\psi}_k$ in the posterior (i.e. the mode) as the best estimator of the state (maximum *a posteriori* approach, which is a well-posed approach in inverse problems). The best estimator is called analysis in weather forecasting (Tarantola 2005). Equation (2.11) provides the Bayesian update, which is key to this work and sequential data assimilation.

2.3.2. *Stochastic ensemble filtering for sequential assimilation*

For brevity, we will omit the subscript k, unless it becomes necessary for clarity. We focus on a qualitative reduced-order model in which (i) the bias on the solution is negligible, (ii) the uncertainty on the state is represented by a covariance, (iii) the probability density function of the state is assumed to be symmetrical around the mean, and (iv) the dynamics at regime do not present frequent extreme events, i.e. the tails of the probability density function are not heavy. In § 4, we relax assumption (i) by introducing a methodology to estimate the bias of the solution, i.e. the model error.

The probability distribution to employ is the distribution that maximises the information entropy (Jaynes 1957), which, in this scenario, is the Gaussian distribution. Therefore, the system's forecast and the observations are assumed to follow Gaussian distributions, i.e. $\boldsymbol{\psi}^f \sim \mathcal{N}(\boldsymbol{\psi}, \boldsymbol{C}^f_{\psi\psi})$ and $\boldsymbol{y} \sim \mathcal{N}(M\boldsymbol{\psi}, \boldsymbol{C}_{\epsilon\epsilon})$, respectively, where $\mathcal{N}$ denotes the normal distribution, with the first argument being the mean, and the second argument being the covariance matrix. The forecast and observation covariance matrices are $\boldsymbol{C}^f_{\psi\psi}$ and $\boldsymbol{C}_{\epsilon\epsilon}$, respectively. If the dynamics were linear, then the Bayesian update (2.11) would be solved

exactly by the Kalman filter equations (Kalman 1960)

$$\boldsymbol{\psi}^a = \boldsymbol{\psi}^f + \left(\boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\right)^{\mathrm{T}}\left[\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\,\boldsymbol{M}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{y} - \boldsymbol{M}\,\boldsymbol{\psi}^f\right), \tag{2.12a}$$

$$\boldsymbol{C}^a_{\psi\psi} = \boldsymbol{C}^f_{\psi\psi} - \left(\boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\right)^{\mathrm{T}}\left[\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\,\boldsymbol{M}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\right), \tag{2.12b}$$

where the superscripts a and f denote analysis and forecast, respectively. Equation (2.12a) corrects the model prediction by weighting the statistical distance between the observations (data) and the forecast, according to the prediction and observation covariances (Evensen 2003). The observation error covariance has to be prescribed based on the knowledge of the experimental methodology used.

In an ensemble method, the distribution is represented by the sample statistics

$$\bar{\boldsymbol{\psi}} \approx \frac{1}{m}\sum_{i=1}^{m}\boldsymbol{\psi}^i, \quad \boldsymbol{C}_{\psi\psi} \approx \frac{1}{m-1}\boldsymbol{\Psi}\boldsymbol{\Psi}^{\mathrm{T}}, \tag{2.13a,b}$$

where the ith column of the matrix $\boldsymbol{\Psi}$ is the deviation from the mean of the ith realisation, $\boldsymbol{\psi}^i - \bar{\boldsymbol{\psi}}$, and m is the number of ensemble members. Because (2.13a,b) is a Monte Carlo Markov chain integration, the sampling error scales as $O(N^{-1/2})$. The key idea of ensemble filters is to group forecast states from a numerical model (the ensemble) to obtain, on filtering, the analysis state. Ensemble methods describe the state's uncertainty by the spread in the ensemble at a given time to avoid the explicit formulation of the covariance matrices (Livings, Dance & Nichols 2008). The algorithmic procedure is as follows. First, the initial condition is integrated forward in time to provide the forecast state $\boldsymbol{\psi}^f$. Second, experimental observations $\boldsymbol{y}$ are assimilated statistically into the forecast to obtain the analysis state $\boldsymbol{\psi}^a$, which, in turn, becomes the initial condition for the next time step. The forecast accumulates errors over the integration period, which is reduced in the assimilation stage through observations with their experimental uncertainties. If the model is qualitatively correct and unbiased, after a sufficient number of assimilations, the ensemble concentrates around the true value. This sequential filtering process on one ensemble member is shown in figure 2. The process is repeated in parallel for the other ensemble members.

2.3.3. *Ensemble square-root Kalman filter*

In the ensemble Kalman filter (2.12), each ensemble member is updated with the assimilation of independently perturbed observation data. However, this method provides a sub-optimal solution that, in some cases, does not preserve the ensemble mean and is affected by sampling errors of the observations (Evensen 2003). Moreover, the ensemble Kalman filter may require a fairly large ensemble to compensate the sampling errors of the observations (Sakov & Oke 2008). The ensemble square-root Kalman filter (EnSRKF), which is an ensemble-transform Kalman filter, overcomes these issues (Livings *et al.* 2008). The key idea of the EnSRKF is to update the ensemble mean and deviations instead of each ensemble member. The EnSRKF for m ensemble members and a state vector of

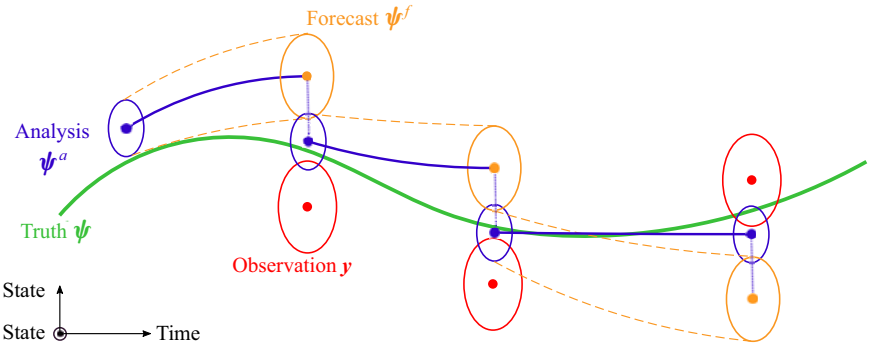


Figure 2. Conceptual schematic of a sequential filtering process: truth (green); observations and their uncertainties (red); forecast state and uncertainties (orange); and analysis state and uncertainties (blue). The circles represent pictorially the spread of the probability density functions: the larger the circles, the larger the uncertainty.

size N reads

$$\boldsymbol{A}^a = \bar{\boldsymbol{A}}^a + \boldsymbol{\Psi}^a, \tag{2.14a}$$

$$\bar{\boldsymbol{A}}^a = \bar{\boldsymbol{A}}^f + \boldsymbol{\Psi}^f \left(\boldsymbol{M}\boldsymbol{\Psi}^f\right)^{\mathrm{T}} \left[(m-1)\,\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\boldsymbol{\Psi}^f \left(\boldsymbol{M}\boldsymbol{\Psi}^f\right)^{\mathrm{T}}\right]^{-1} \left(\boldsymbol{Y} - \boldsymbol{M}\bar{\boldsymbol{A}}^f\right), \tag{2.14b}$$

$$\boldsymbol{\Psi}^a = \boldsymbol{\Psi}^f \boldsymbol{V} \left(\mathbb{I} - \boldsymbol{\Sigma}\right)^{1/2} \boldsymbol{V}^{\mathrm{T}}, \tag{2.14c}$$

$$\boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}} = \left(\boldsymbol{M}\boldsymbol{\Psi}^f\right)^{\mathrm{T}} \left[(m-1)\,\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\boldsymbol{\Psi}^f \left(\boldsymbol{M}\boldsymbol{\Psi}^f\right)^{\mathrm{T}}\right]^{-1} \left(\boldsymbol{M}\boldsymbol{\Psi}^f\right), \tag{2.14d}$$

where $\boldsymbol{A} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_m) \in \mathbb{R}^{N\times m}$ is the matrix that contains the ensemble members as columns, $\bar{\boldsymbol{A}} = (\bar{\boldsymbol{\psi}}, \ldots, \bar{\boldsymbol{\psi}}) \in \mathbb{R}^{N\times m}$ contains the mean analysis states in each column, $\boldsymbol{Y} = (\boldsymbol{y}, \ldots, \boldsymbol{y}) \in \mathbb{R}^{q\times m}$ is the matrix containing the q observations repeated m times, the identity matrix is represented by $\mathbb{I}$, and $\boldsymbol{V}$ and $\boldsymbol{\Sigma}$ are the orthogonal matrices of eigenvectors and a diagonal matrix of eigenvalues, respectively, from singular value decomposition. The largest matrices required in the EnSRKF algorithm have dimension $N \times m$ and $m \times m$, therefore the storage requirements are significantly smaller than those of non-ensemble based filters. In addition, this filter is non-intrusive and suitable for parallel computation. A derivation of the EnSRKF can be found in Appendix A.

2.4. *Discussion*

An ensemble method enables us to: (i) work with high-dimensional systems because we do not need to propagate the covariance matrix, which has $O(N^2)$ components; (ii) work with nonlinear systems, such as the thermoacoustic system under investigation; (iii) work with time-dependent problems; (iv) not store the data because we sequentially assimilate (real-time, i.e. on-the-fly, assimilation); and (v) avoid implementing tangent or adjoint solvers, which are required, for example, in variational data assimilation methods (Traverso & Magri 2019). On the one hand, if the system were linear, then a Gaussian prior would remain Gaussian under time integration. This makes the ensemble filter the exact Bayesian update in the limit of an infinite number of samples. On the other hand, if the system were nonlinear (e.g. in the present study), then a Gaussian prior would not necessarily remain

Gaussian under time integration. This makes the ensemble filter an approximate Bayesian update. The update of the first and second statistical moments, however, remains exact. In other words, we cannot capture the skewness, kurtosis, and other higher moments. (Particle filter methods overcome this limitation, but they may be expensive computationally; Pham 2001.)

3. State and parameter estimation

This work considers both state estimation, in which the state is the uncertain quantity (§ 3.1), and combined state and parameter estimation, in which both the state and model parameters are uncertain (§ 3.2).

3.1. *State estimation*

State estimation is the process of using a series of noisy measurements into an estimation of the state of the dynamical system, $\boldsymbol{\psi}$. This paper considers two different scenarios in assimilating acoustic data in thermoacoustics: (i) assimilation of the acoustic modes; and (ii) assimilation of pressure measurements from N_{mic} microphones, which are located equidistantly from the flame location up to the end of the Rijke tube (figure 1). The assimilation of acoustic modes assumes that observation data are available for the pressure and velocity acoustic modes, $\{\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}\}$. Hence the state equations are

$$\left.\begin{gathered}\frac{\mathrm{d}\boldsymbol{\psi}}{\mathrm{d}t} = \boldsymbol{F}(\boldsymbol{\alpha}; \boldsymbol{\psi}), \quad \boldsymbol{\psi}(t=0) = \boldsymbol{\psi}_0 = (\boldsymbol{\eta}_0, \dot{\boldsymbol{\eta}}_0, \boldsymbol{v}_0)^{\mathrm{T}}, \\ \boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})\,\boldsymbol{\psi} = (\boldsymbol{\eta}, \dot{\boldsymbol{\eta}})^{\mathrm{T}}.\end{gathered}\right\} \tag{3.1}$$

Alternatively, in scenario (ii), from (2.5*b*), the reference pressure measurements are

$$\boldsymbol{p}'_{mic} = \begin{pmatrix} p'_1(t) \\ p'_2(t) \\ \vdots \\ p'_{N_{mic}}(t) \end{pmatrix} = - \begin{pmatrix} \sin(\pi x_1) & \sin(2\pi x_1) & \dots & \sin(N_m \pi x_1) \\ \sin(\pi x_2) & \sin(2\pi x_2) & \dots & \sin(N_m \pi x_2) \\ \vdots & \vdots & & \vdots \\ \sin(\pi x_{N_{mic}}) & \sin(2\pi x_{N_{mic}}) & \dots & \sin(N_m \pi x_{N_{mic}}) \end{pmatrix} \begin{pmatrix} \dfrac{\dot{\eta}_1(t)}{\pi} \\ \dfrac{\dot{\eta}_2(t)}{2\pi} \\ \vdots \\ \dfrac{\dot{\eta}_{N_m}(t)}{N_m \pi} \end{pmatrix}. \tag{3.2}$$

The statistical errors of the microphones are assumed to be independent and Gaussian. In the twin experiment, the pressure observations are created from the true state, with a standard deviation σ_{mic} that mimics the measurement error. Pressure data cannot be assimilated directly with the EnSRKF because the state vector contains the acoustic modes, i.e. it does not contain the acoustic pressure. To circumvent this, we augment the state vector with the acoustic pressure at the microphones' locations according to (3.2). Therefore, the new state vector includes the Galerkin acoustic modes, the dummy velocity variables and the pressure at the different microphone locations, i.e. $\boldsymbol{\psi}' \equiv (\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}, \boldsymbol{v}, \boldsymbol{p}'_{\boldsymbol{mic}})^{\mathrm{T}}$, with dimension $N' = 2N_m + N_c + N_{mic}$. The augmented state equations are

$$\left.\begin{gathered}\frac{\mathrm{d}\boldsymbol{\psi}'}{\mathrm{d}t} = \boldsymbol{F}(\boldsymbol{\alpha}; \boldsymbol{\psi}), \quad \boldsymbol{\psi}'(t=0) = \boldsymbol{\psi}'_0 = (\boldsymbol{\eta}_0, \dot{\boldsymbol{\eta}}_0, \boldsymbol{v}_0, \boldsymbol{p}'_{\boldsymbol{mic}0})^{\mathrm{T}}, \\ \boldsymbol{y} = \boldsymbol{M}(\boldsymbol{x})\,\boldsymbol{\psi}' = \boldsymbol{p}'_{\boldsymbol{mic}}(x).\end{gathered}\right\} \tag{3.3}$$

With this, the modes will be updated indirectly during the assimilation step using the microphone data and their experimental error.

3.2. *Combined state and parameter estimation*

Combined state and parameter estimation is the process of using a series of noisy measurements into an estimation of the state of the dynamical system $\boldsymbol{\psi}$ and the parameters $\boldsymbol{\alpha}$. The parameters are regarded as variables of the dynamical system so that they are updated in every analysis step. This is achieved by combining the governing equations of the thermoacoustic model with the equations that describe the evolution of parameters, which are constant in time, but can change when observations are assimilated. The equations for the augmented state of combined state and parameter estimation are

$$\left.\begin{gathered}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\boldsymbol{\psi}\\ \boldsymbol{\alpha}\end{bmatrix}=\begin{bmatrix}\boldsymbol{F}(\boldsymbol{\alpha};\boldsymbol{\psi})\\ 0\end{bmatrix}, \quad \begin{gathered}\boldsymbol{\psi}(t=0)=\boldsymbol{\psi}_0,\\ \boldsymbol{\alpha}(t=0)=\boldsymbol{\alpha}_0,\end{gathered}\\ \boldsymbol{y}=\boldsymbol{M}(\boldsymbol{x})\,\boldsymbol{\psi}.\end{gathered}\right\} \tag{3.4}$$

With a slight abuse of notation, the state vector $\boldsymbol{\psi}$ in (3.4) is equal to $\boldsymbol{\psi}\equiv(\boldsymbol{\eta},\dot{\boldsymbol{\eta}},\boldsymbol{v})^{\mathrm{T}}$ in (3.1) for the assimilation of acoustic modes, and equal to $\boldsymbol{\psi}'\equiv(\boldsymbol{\eta},\dot{\boldsymbol{\eta}},\boldsymbol{v},\boldsymbol{p}'_{mic})^{\mathrm{T}}$ in (3.3) for the assimilation of pressure measurements.

The data assimilation algorithm is applied to the augmented system for both the forecast state and the parameters to be updated at every analysis step. The parameters need to be initialised for each ensemble member from a uniform distribution with width 25 % of the mean parameter value. In other words, we assume that the parameters are uncertain by $\pm 25\,\%$.

3.3. *Performance metrics*

The performance of the state estimation and combined state and parameter estimation are evaluated with three metrics: (i) the trace of the forecast covariance, $\boldsymbol{C}^{f}_{\psi\psi}$, which measures globally the spread of the ensemble; (ii) the relative difference between the true pressure oscillations at the flame location and the filtered solution, which measures the instantaneous error; and (iii) for the combined state and parameter assimilation, the convergence of the filtered parameters normalised to their true values, as well as the root-mean-square error with respect to the true solution.

4. Data assimilation with bias estimation

In this section, we analyse the case of state, parameter and model bias estimation. Sources of model bias in the model of § 2.1 include (i) idealised boundary conditions, (ii) a simple heat-release law with no simulation of the flame, and (iii) zero mean flow effects. In this paper, a higher-fidelity model produces data that account for these three sources of model bias. We infer the model bias to correct the biased forecast state prior to the analysis step. By performing state and parameter estimation on the unbiased forecast, we increase the quantitative accuracy of the model prediction. First, we define the time-dependent model bias $\boldsymbol{U}(t)$ as the difference between the true pressure state (from the higher-fidelity model) at the microphone locations $\boldsymbol{p}'^{\,t}_{mic}$, and the expected biased pressure $\langle\boldsymbol{p}'_{mic}\rangle$, i.e. the mean of the ensemble of pressures

$$\boldsymbol{U}(t)=\boldsymbol{p}'^{\,t}_{mic}(t)-\left\langle\boldsymbol{p}'_{mic}(t)\right\rangle. \tag{4.1}$$

We propose an echo state network to predict the evolution of the model bias.

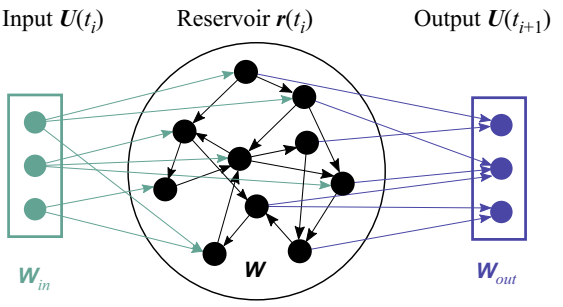


Figure 3. Schematic representation of an echo state network.

4.1. *Echo state networks*

An echo state network (ESN) is a type of recurrent neural network based on reservoir computing (Lukoševičius 2012). The ESNs learn temporal correlations in data by nonlinearly expanding the data into a high-dimensional reservoir, which acts as the memory of the system, and is a framework more versatile than auto-regressive models (Aggarwal 2018).

Figure 3 shows a pictorial representation of an ESN. The reservoir is defined by a high-dimensional vector $\boldsymbol{r}(t_i) \in \mathbb{R}^{N_r}$ and a state matrix $\boldsymbol{W} \in \mathbb{R}^{N_r \times N_r}$, where N_r is the number of neurons in the reservoir. We use $N_r = 100$ neurons, which are sufficient to represent the dynamics of the thermoacoustic system (Huhn & Magri 2022). The inputs and outputs from the reservoir are vectors of dimension $\mathbb{R}^{N_{mic}}$ because we define the bias as the pressure error at each microphone (4.1). Their input and output matrices are $\boldsymbol{W}_{in} \in \mathbb{R}^{N_r \times (N_{mic}+1)}$ and $\boldsymbol{W}_{out} \in \mathbb{R}^{N_{mic} \times (N_r+1)}$, respectively. At every time t_i, the input bias $\boldsymbol{U}(t_i)$ and the state of the reservoir at the previous time step $\boldsymbol{r}(t_{i-1})$ are combined to predict the reservoir state at the current time as well as the bias at the next time step $\boldsymbol{U}(t_{i+1})$ such that

$$\boldsymbol{r}(t_i) = \tanh\Big(\boldsymbol{W}_{in}[\tilde{U}(t_i); 0.1] + \boldsymbol{W}\boldsymbol{r}(t_{i-1})\Big) \quad \text{and} \quad U(t_{i+1}) = \boldsymbol{W}_{out}[\boldsymbol{r}(t_i); 1], \tag{4.2a,b}$$

where $\tilde{U}$ is the input bias normalised by the range, component-wise, and the constants 0.1 and 1 are used to break the symmetry of the ESN (Huhn & Magri 2020*a*). The operator $[\,;\,]$ indicates vertical concatenation. The matrices $\boldsymbol{W}_{in}$ and $\boldsymbol{W}$ are predefined as fixed, sparse and randomly generated. Specifically, $\boldsymbol{W}_{in}$ has only one non-zero element per row, which is sampled from a uniform distribution in $[-\sigma_{in}, \sigma_{in}]$, where σ_{in} is the input scaling. Matrix $\boldsymbol{W}$ is an Erdős–Rényi matrix with average connectivity $d = 5$, in which each neuron (each row of $\boldsymbol{W}$) has on average d connections (non-zero elements), which are obtained by sampling from a uniform distribution in $[-1, 1]$; the entire matrix is then rescaled by a multiplication factor to set the spectral radius ρ. The weights of $\boldsymbol{W}_{out}$ are determined through training, which consists of solving a linear system for a training set of length N_{tr}:

$$(\boldsymbol{R}\boldsymbol{R}^{\mathrm{T}} + \gamma_t \mathbb{I})\boldsymbol{W}_{out}^{\mathrm{T}} = \boldsymbol{R}\boldsymbol{U}_{train}^{\mathrm{T}}, \tag{4.3}$$

where $\boldsymbol{R} \in \mathbb{R}^{(N_r+1)\times N_{tr}}$ is the horizontal concatenation of the augmented reservoir state for each time in the training set $[\boldsymbol{r}(t_i); 1]$ with $i = 1, \ldots, N_{tr}$, $\boldsymbol{U}_{train} \in \mathbb{R}^{N_{mic}\times N_{tr}}$ is the time concatenation of the output data, $\mathbb{I}$ is the identity matrix, and γ_t is the Tikhonov regularisation parameter (Lukoševičius 2012). We employ recycle validation (Racca &

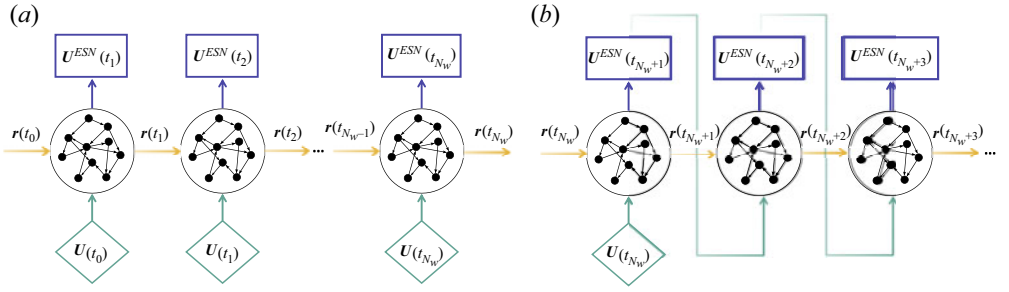


Figure 4. Schematic representation of the ESN forecast methods: (*a*) open-loop forecast, and (*b*) closed-loop estimation of the bias.

Magri 2021) to select the input scaling $\sigma_{in} = 0.0126$, the spectral radius $\rho = 0.9667$, and the Tikhonov parameter $\gamma_t = 1 \times 10^{-16}$.

4.2. *Thermoacoustic echo state network*

On the one hand, in open loop (figure 4*a*), the true bias data (4.1) are fed into every forecast step to compute (4.2*a*,*b*). The output bias from the ESN is disregarded in open loop, which is used to initialise the network (the washout), such that $\boldsymbol{U}_w \in \mathbb{R}^{N_{mic} \times N_w}$. State and parameters are not updated during the washout. On the other hand, in closed loop (figure 4*b*), the true bias data (4.1) are fed in the first step, then the output bias from a forecast step (4.2*b*) is used as the initial condition of the next step. The ESN forecast frequency is set to be five times smaller than the thermoacoustic model time step, to reduce the additional computation cost associated with the bias estimation.

In detail, the pseudo-algorithm 1 summarises the procedure that we propose for bias-aware data assimilation with an ESN: (1) the ensemble of acoustic modes is initialised and forecast for the washout time; (2) we run the ESN in open loop to initialise the reservoir; and (3) we perform data assimilation. When measurements become available, we compute the ensemble of pressures by adding the estimated bias from the ESN, U, to the expectation of the forecast pressure at the microphones, such that the unbiased ensemble is centred around the unbiased pressure state $\hat{\boldsymbol{p}}'_{mic\,j} = \boldsymbol{U} + \boldsymbol{p}'_{mic\,j}$ for $j = 1, \ldots, m$, where $(\hat{})$ indicates a statistically unbiased quantity. Subsequently, we perform the analysis step. The EnSRKF obtains the optimal combination between the unbiased pressures and the observations, which updates indirectly the biased ensemble of acoustic modes and parameters. The resulting analysis ensemble is used to re-initialise the ESN for the next forecast in closed loop, such that the bias is the difference between the observations and the expectation of the analysis pressures, i.e. $\boldsymbol{U} = \boldsymbol{p}'^{\,t}_{mic} - \langle \boldsymbol{p}'^{\,a}_{mic} \rangle$. Finally, the Rijke model and the ESN are time-marched until the next measurement becomes available for assimilation.

4.3. *Test case*

The higher-fidelity model that we use is based on the travelling-wave approach of Dowling (1999), which is described in detail by Aguilar Pérez (2019). The acoustic pressure and velocity are written as functions of two acoustic waves that propagate downstream and upstream of the tube (see (3.5) and (3.7) in Dowling 1999). As shown in figure 5, the waves are defined as f and g, with convective velocities $\tilde{c}_0 \pm \tilde{u}_{0,u}$ in the region $\tilde{x} \leqslant \tilde{x}_f$, and h and j, with convective velocities $\tilde{c}_0 \pm \tilde{u}_{0,d}$ in $\tilde{x} \geqslant \tilde{x}_f$. This model uses dimensional quantities, so we transform our dimensionless thermoacoustic system (2.3) into its dimensional form.

Algorithm 1 Data assimilation with ESN bias estimation

$A^f(t_0) \leftarrow$ InitialiseEnsemble
while $t \leq t_{washout}$ **do** ▷ Get ESN running
  $p^f(t) \leftarrow$ ComputePressure($A^f(t)$) (3.2)
  $U(t) = p^{true} - \langle p^f \rangle$
  $A^f(t + \Delta t) \leftarrow$ ForecastRijkeModel($A^f(t)$) (2.9)
  $t = t + \Delta t$
$[U(t), r(t)] \leftarrow$ OpenLoopESN($U(t_0 : t_{washout})$) (4.2*a*,*b*)
loop: for every t when observations p^{true} are available ▷ Start data assimilation
  $p^f(t) \leftarrow$ ComputePressure($A^f(t)$)
  $\hat{p}^f_j(t) = p^f_j(t) + U(t)$, for $j = 1 : m$
  $A^\star \leftarrow$AugmentState($A^f, \hat{p}^f$)
  $A^a \leftarrow$ EnSRKF($A^\star, p^{true}$) (2.14*a*)
  $p^a(t) \leftarrow$ ComputePressure($A^a(t)$)
  $U(t) = p^{true} - \langle p^a \rangle$
  while no observations are available at time t **do**
    $[U(t + \Delta t), r(t + \Delta t)] \leftarrow$ ClosedLoopESN($U(t), r(t)$) (4.2*a*,*b*)
    $A^f(t + \Delta t) \leftarrow$ ForecastRijkeModel($A^f(t)$)
    $t = t + \Delta t$

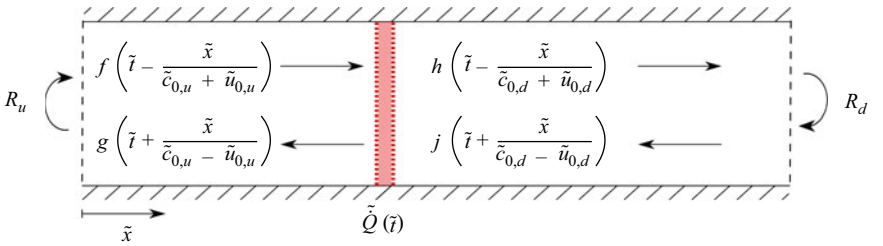


Figure 5. Schematic of the acoustic waves travelling upstream and downstream of a flame in an open-ended tube.

The equations that relate the travelling waves $g(\tilde{t})$ and $h(\tilde{t})$ to the heat-release rate $\tilde{Q}(\tilde{t})$ are, with a slight abuse of notation,

$$\boldsymbol{\mathcal{X}} \begin{pmatrix} g(\tilde{t}) \\ h(\tilde{t}) \end{pmatrix} = \boldsymbol{\mathcal{Y}} \begin{pmatrix} g(\tilde{t} - \tilde{\tau}_u) \\ h(\tilde{t} - \tilde{\tau}_d) \end{pmatrix} + \begin{pmatrix} 0 \\ \left(\tilde{Q}(t) - \bar{\tilde{Q}}_0\right) / \left(A\tilde{c}_{0,u}\right) \end{pmatrix}, \tag{4.4}$$

where A is the cross-sectional area of the duct, and $\tilde{\tau}_u$ and $\tilde{\tau}_d$ are the times taken for the acoustic waves to travel from the flame to the upstream and downstream boundaries, respectively. For completeness, the matrices $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are provided in the supplementary material available at https://doi.org/10.1017/jfm.2022.653.

In contrast to the low-fidelity model of § 2.1, the travelling-wave approach accounts for (i) mean flow effects, and (ii) non-ideal open boundary conditions, such that $f(\tilde{t}) = R_u\, g(\tilde{t} - \tilde{\tau}_u)$ and $j(\tilde{t}) = R_d\, h(\tilde{t} - \tilde{\tau}_d)$, where R_u and R_d are the reflection coefficients. We set the mean flow velocity upstream of the flame to $\tilde{u}_{0,u} = 10$ m s^{-1}, and the mean heat release rate to $\bar{\tilde{Q}}_0 = 2000$ W. The velocity downstream of the flame is computed

by applying the jump conditions in the energy and momentum equations (see (3.2) and (3.3) in Dowling 1999). We set dissipative boundary condition to $R_u = R_d = -0.999$. In addition, the higher-fidelity model employs a flame kinematic model, which simulates the flame dynamics and provides the heat release as a function of the area enclosed by the flame front (detailed code in Aguilar Pérez 2019). On the other hand, the low-order model employs a simple time-delayed model in which the heat released by the flame is modelled by two parameters (β and τ) and a nonlinear law that links the acoustic fluctuation with the heat release (adapted King's law (2.4)).

We test the bias-aware data assimilation with the ESN for a limit cycle. We perform unbiased state estimation, and combined unbiased state and parameter estimation using synthetic pressure data obtained from the travelling-wave and flame kinematic model as the truth (higher-fidelity data). For brevity, we perform parameter estimation to the heat-source strength $\tilde{\beta}$ only, which we expect physically to be $\tilde{\beta} \sim O(10^6)$ (W s$^{1/2}$ m$^{-5/2}$)((2.7) in Juniper 2011). The results are discussed in § 8.

5. Nonlinear characterisation

In order to assess the performance of data assimilation, we first characterise the nonlinear dynamics of the low-order model of § 2.1 by analysing the solutions at regime (after the initial transient) with bifurcation analysis and nonlinear time series analysis (Kantz & Schreiber 2003; Kabiraj, Sujith & Wahi 2012*b*; Guan, Gupta & Li 2020). The system's parameters are $x_f = 0.2$, $C_1 = 0.1$, $C_2 = 0.06$ and $N_m = 10$.

In bifurcation analysis, we examine the topological changes in the pressure oscillations at the flame location p'_f as the control parameters vary. First, we study the two-dimensional bifurcation diagram, which is shown in figure 6. The classification in the two-dimensional diagram is obtained following the procedure of Huhn & Magri (2020*b*). This method consists of obtaining the Lyapunov exponents λ_i through covariant-vector analysis. With this, the dynamical motions are identified as: (i) fixed point if $\lambda_1 < 0$; (ii) limit cycle if $\lambda_1 = 0$ and $\lambda_2 < 0$; (iii) quasi-periodic if $\lambda_1 = 0$ and $\lambda_2 = 0$; and (iv) chaotic if $\lambda_1 > 0$. For small β and τ, the system converges to a fixed point because the thermoacoustic energy is smaller than damping. As the heat-source strength increases, the Rayleigh criterion is fulfilled and self-excited oscillations arise as limit cycles. When β reaches values over 2.5, different types of solution appear, such as quasi-periodic or chaotic attractors. The refined region in figure 6 shows that the type of solution is sensitive to small changes in the control parameters, which has implications for data assimilation, as argued in the remainder of the paper.

These topological changes are investigated further with a one-dimensional bifurcation diagram for a fixed time delay ($\tau = 0.2$), shown in figure 7. Because the nonlinear solutions at regime may vary with the initial condition, two sets of results are shown for a small initial condition ($\eta_j = \dot{\eta}_j/j\pi = 0.005$) and a large initial condition ($\eta_j = \dot{\eta}_j/j\pi = 5$) to capture subcritical behaviours. The bifurcation diagram is obtained by marching forward in time the governing equations of the nonlinear dynamical system until the system reaches a statistically stationary state. For each value of the control parameter, the bifurcation diagram shows the peaks and troughs of the acoustic pressure at the flame location. (The nonlinear time series analysis results are shown in the supplementary material.) From left to right, first, the solution is the fixed point (region A), which is the case of no oscillations. Second, the appearance of periodic oscillations from a fixed point is observed with a large initial condition at $\beta = 0.26$, with a small region of hysteresis from $\beta = 0.26$ to $\beta = 0.34$. This first self-sustained state is a period-1 limit cycle

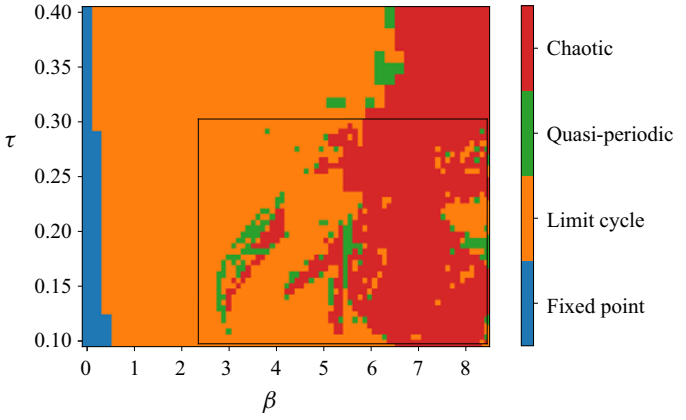


Figure 6. Two-dimensional bifurcation diagram. Classification of the attractor of the thermoacoustic system. The area enclosed by the black rectangle corresponds to a refined grid. The coarse and fine sweeps are performed with resolutions $(\Delta\beta, \Delta\tau) = (0.2, 0.01)$ and $(\Delta\beta, \Delta\tau) = (0.1, 0.005)$, respectively.

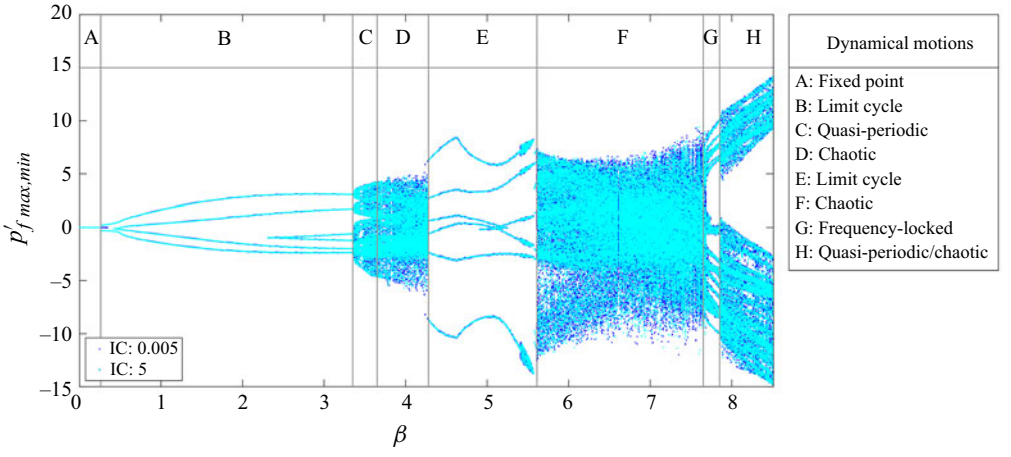


Figure 7. One-dimensional bifurcation diagram. Maxima and minima of the pressure oscillations at the flame location versus the heat-source strength. The solutions obtained for small/large initial conditions (IC) are shown in a dark/light blue. This diagram identifies different nonlinear behaviours, which have implications for data assimilation.

(region B), which originates from a subcritical Hopf bifurcation. Within region B, the system undergoes a period-doubling bifurcation at $\beta = 0.6$ from period-1 to period-2 oscillations. Third, the period-2 limit cycle bifurcates into a 3-torus quasi-periodic motion at $\beta = 3.35$ (region C). A quasi-periodic oscillation is an aperiodic solution that results from the interaction between two or more incommensurate frequencies (also known as a Neimark–Sacker bifurcation) (Kabiraj *et al.* 2012*b*). Fourth, the solution becomes chaotic at $\beta = 3.65$ (region D). In summary, the evolution from region A to region D shows that the system reaches a chaotic state via a quasi-periodic route to chaos, i.e. via a Ruelle–Takens scenario (Kabiraj *et al.* 2012*a*). Fifth, after this first route to chaos, changes in the control parameter drive the system back to a periodic limit cycle through a tangent bifurcation (Kantz & Schreiber 2003) at approximately $\beta = 4.25$ (region E), with a second region of hysteresis from $\beta = 4.24$ to $\beta = 4.28$. This high-amplitude limit cycles region

Parameter	Value	Parameter	Value
x_f	0.2	β	[0.2, 0.4, 3.6, 7.7, 7.0]
N_m	10	τ	0.2
N_c	10	m	10
Δt	0.001	σ_{frac}	0.25
$\eta_j(t=0)$	0.005	$\Delta t_{analysis}$	[2 (Non-chaotic), 0.5 (Chaotic)]
$\dot{\eta}_j/j\pi(t=0)$	0.005	N_{mic}	6
$v_i(t=0)$	0	σ_{mic}	0.01

Table 1. Parameters and initial conditions for generating the true solution and the assimilated observations.

becomes again chaotic at $\beta = 5.61$ (region F). Sixth, when β reaches 7.65, the system evolves towards a frequency-locked state (region G). Frequency-locked solutions arise from the competition between two or more frequencies, but in contrast to quasi-periodic signals, these frequencies are commensurate. Seventh, at $\beta = 7.85$, the frequency-locked solution bifurcates into a quasi-periodic solution (region H). Region H solutions show a two-dimensional toroidal structure, in contrast to the three-dimensional torus from region C. In region H, some of the simulations showed that there are areas of chaotic dynamics, which can be appreciated by the difference in the solutions from the small and large initial conditions in figure 7. (A higher region refinement could be performed to fully understand the bifurcations within this region, however, that is beyond the scope of this work.) The qualitative bifurcation behaviour of this reduced-order model is observed in experiments (Kabiraj *et al.* 2012*b*; Kabiraj & Sujith 2012), which means that the reduced-order model captures qualitatively the nonlinear thermoacoustic dynamics.

The bifurcation analysis shows a rich variety of solutions in a relatively small range of parameters, i.e. small changes of a parameter, or a state, can generate solutions that are topologically different. This nonlinear sensitivity has implications in the design of an ensemble data assimilation framework, as discussed in § 6.

6. Twin experiments in non-chaotic regimes

We perform a series of experiments with synthetic data, which are generated by the low-order model (§ 2.1). To mimic an experiment, we add stochastic uncertainty to the synthetic data by prescribing an observation covariance matrix. This approach is also known as the twin experiment (e.g. Traverso & Magri 2019). The EnSRKF algorithm is tested in the different regions of figure 7, for the different nonlinear regimes: fixed point, limit cycle, frequency-locked, quasi-periodic and chaotic. The filter is first tested in the non-chaotic regimes for the assimilation of (i) acoustic modes (§ 6.1), and (ii) acoustic pressure from microphones (§ 6.2). The assimilation of chaotic solutions, which presents further challenges, is investigated in § 7. Different simulations are performed to determine suitable values for the number of ensemble members (m), the time between analysis ($\Delta t_{analysis}$), the standard deviation (σ_{frac}), i.e. the observations' uncertainties during the acoustic modes assimilation, and the standard deviation of the microphone measurements (σ_{mic}). Table 1 shows the parameters and initial conditions of the reference (i.e. 'true') solution. This range of parameters is justified from the literature in thermoacoustic data assimilation (Traverso & Magri 2019). Computational time is discussed in the supplementary material.

6.1. *Assimilation of the acoustic modes*

This subsection includes results for state estimation (§ 6.1.1) and combined state and parameter estimation (§ 6.1.2).

6.1.1. *State estimation*

This subsection presents simulations performed assuming that there are observations available for all acoustic modes, i.e. the number of observations is $q = 2N_m = 20$. (Including observations for the dummy velocity variables would improve further the filter convergence; however, they are not considered because the velocity advection field in the heat-source region is not measured in a real engine.) Figure 8 shows the acoustic pressure before assimilation (unfiltered solution) and after assimilation (filtered solution), and the data at the assimilation steps (analysis steps), for the transient of a fixed point, a period-1 limit cycle, a frequency-locked motion, and a quasi-periodic motion. In the filtered solution, data assimilation is performed during the first 50 time units, and it is marched in time without further assimilation for 10 more time units. The EnSRKF learns successfully (i.e. infers) the true solution for all the nonlinear regimes. As expected, the convergence is faster for the fixed point and limit cycle cases (figures 8*a*,*b*) because they are simpler dynamical motions. (The unfiltered solution also converges to the same value for these simple cases. This is due to the stable nature of their attractors, and because their regions are unaffected by the chaotic butterfly effect.) For multi-frequency dynamical regimes, figures 8(*c*,*d*) show that the Bayesian update can learn the frequency-locked and quasi-periodic states of regions C and G in figure 7. However, these show more discrepancies between the filtered and true solutions. Physically, this is due to the multiple bifurcations that occur in a small range of parameters, which is typical of thermoacoustic systems. In reference to figure 7, region C is next to the chaotic region D; and region G is a short-range region surrounded by the chaotic region F and the mixed quasi-periodic-chaotic region H. Therefore, the discrepancy in these cases is caused by some ensemble members falling in different basins of attraction. To overcome this issue, we propose a strategy in § 6.2.2.

The data assimilation process depends on the observation's uncertainty σ_{frac} and ensemble size m. Figure 9 shows the performance metrics (§ 3) for the quasi-periodic solution of figure 8(*d*). As expected, the filtered solution is more accurate for a smaller standard deviation because the observations are closer to the truth. Importantly, the algorithm is capable of learning the reference solution for an ensemble having an error as large as 50 % of the mean of the acoustic modes, which means that the data assimilation algorithm is robust. For the pressure performance metric, the algorithm brings the relative error below 10 % after 15 time units (in the worst case scenario, figure 9*a*). For the covariance matrix trace performance metric, the EnSRKF continuously reduces the initial ensemble variance up to a final plateau, which cannot be zero because of the non-zero observation and forecast background noise (figure 9*c*). The evolution of the trace is an indicator of the spread of the forecast ensemble, which informs on the uncertainty of the solution. The ensemble size does not have a strong influence in the ensemble uncertainty during the assimilation because the trace of the covariance matrix remains of the same magnitude independently of the value of m (figure 9*d*). Nevertheless, the relative error is significantly higher for a small ensemble with $m = 4$ (figure 9*c*). This means that four ensemble members are not sufficient to give a sufficient ensemble distribution, therefore the solution converges to an incorrect state, but with a small spread around it. Comparing the errors for 10 and 50 ensemble members, we see no significant differences between the

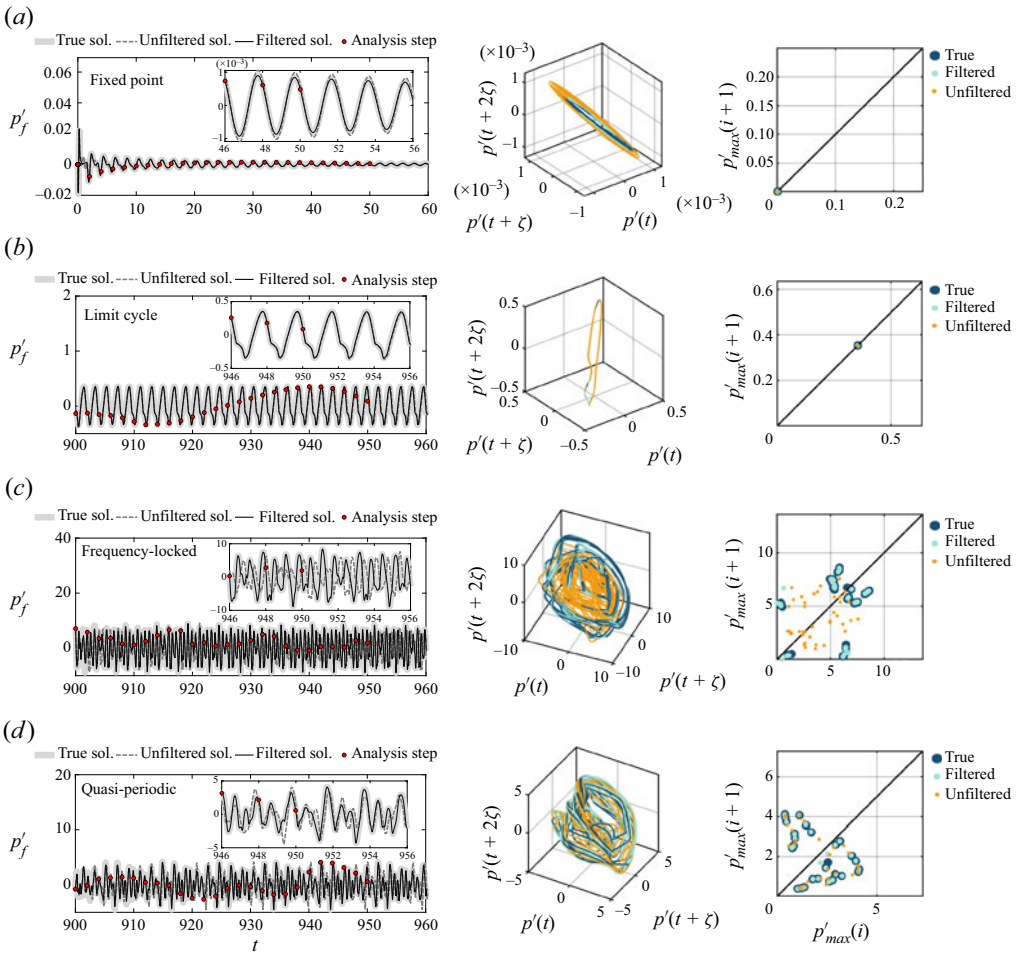


Figure 8. Real-time learning of the state. Assimilation of acoustic modes for state estimation of non-chaotic regimes: (*a*) transient towards a fixed point ($\beta = 0.2$); (*b*) limit cycle ($\beta = 0.4$); (*c*) frequency-locked ($\beta = 7.7$); and (*d*) quasi-periodic ($\beta = 3.6$). Left-hand plots: true pressure oscillations at the flame location (light grey), unfiltered solution (dashed dark grey) and filtered solution (black). The analysis time steps are indicated with red circles. Centre and right-hand plots: phase portrait and first return map of the true (dark blue), filtered (light blue) and unfiltered (orange) solutions. Here, $m = 10$, $\sigma_{frac} = 0.25$ and $\Delta t_{analysis} = 2$.

solutions, which shows that having an ensemble size larger than the number of degrees of freedom is not required. This is one of the benefits of using the square-root filter (in the standard ensemble Kalman filter, larger ensembles are needed to avoid sampling errors Livings *et al.* 2008). However, the computational time required for 50 ensemble members was approximately 4 times longer than that for 10. Therefore, an ensemble size $m = 10$ provides a good approximation of the true state for the assimilation of acoustic modes, while keeping the computation time minimal.

6.1.2. *Combined state and parameter estimation*

In this subsection, we analyse the combined state and parameter estimation to calibrate both the state and parameters. The two uncertain parameters ($\boldsymbol{\alpha} = (\beta, \tau)^{\mathrm{T}}$ in (3.4)) are added to the state vector and updated simultaneously with the acoustic modes and dummy

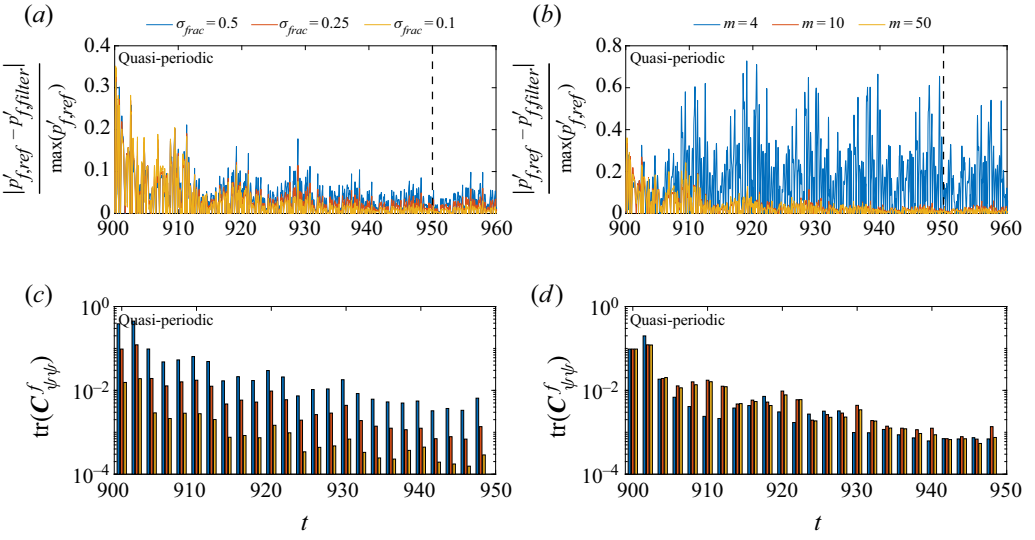


Figure 9. Assimilation of acoustic modes for state estimation of a quasi-periodic regime – performance metrics. (*a*,*c*) Effect of the standard deviation with $m = 10$. (*b*,*d*) Effect of the ensemble size with modes measurement uncertainty $\sigma_{frac} = 0.25$. The error evolution is shown with the relative difference between the filtered solutions and truth (*a*,*b*) and the trace of the ensemble covariance (*c*,*d*). The dashed vertical line indicates when data assimilation ends. Here, $\beta = 3.6$ and $\Delta t_{analysis} = 2$.

velocity variables, as detailed in § 3.2. Figure 10 shows the evolution of the parameters, normalised to their true value, for the four non-chaotic solutions. The convergence shows that the EnSRKF update is capable of learning the true β and τ values for the four dynamical motions.

For a comparison of combined state and parameter estimation with state estimation, we compute the root-mean-square (r.m.s.) error. The r.m.s. error at each time step is defined as the square root of the trace of the covariance matrix of the filtered ensemble, relative to the true solution:

$$\text{r.m.s. error} = \sqrt{\mathrm{tr}\left(\frac{1}{m-1}\sum_{j=1}^{m}\left(\boldsymbol{\psi}_j - \boldsymbol{\psi}^{true}\right)\left(\boldsymbol{\psi}_j - \boldsymbol{\psi}^{true}\right)^{\mathrm{T}}\right)}. \tag{6.1}$$

The r.m.s. error is evaluated for the state estimation and the combined state and parameter estimation cases, using different initial uncertainties for β and τ. This is achieved in state estimation by defining $\beta = c\,\beta^{true}$ and $\tau = c\,\tau^{true}$, where c is the defined initial uncertainty. For the combined state and parameter estimation, the initial β and τ of each member in the ensemble are taken from a uniform distribution centred around $c\,\beta^{true}$ and $c\,\tau^{true}$, with sample standard deviation 25 %. Figure 11(*a*) shows the r.m.s. error for the initial parameters set to their true value. The state estimation outperforms the combined state and parameter estimation only in this case, as the state estimation model works with constant true parameters while the combined state and parameter estimation updates the parameters in each analysis step with the EnSRKF update. The true parameters are perturbed by 5 %, 25 % and 50 % in figures 11(*b*,*c*,*d*), respectively. The combined state and parameter estimation simulations are capable of learning the true state up to a 25 % error in the parameters initialisation, as the r.m.s. error is reduced by two orders of magnitude

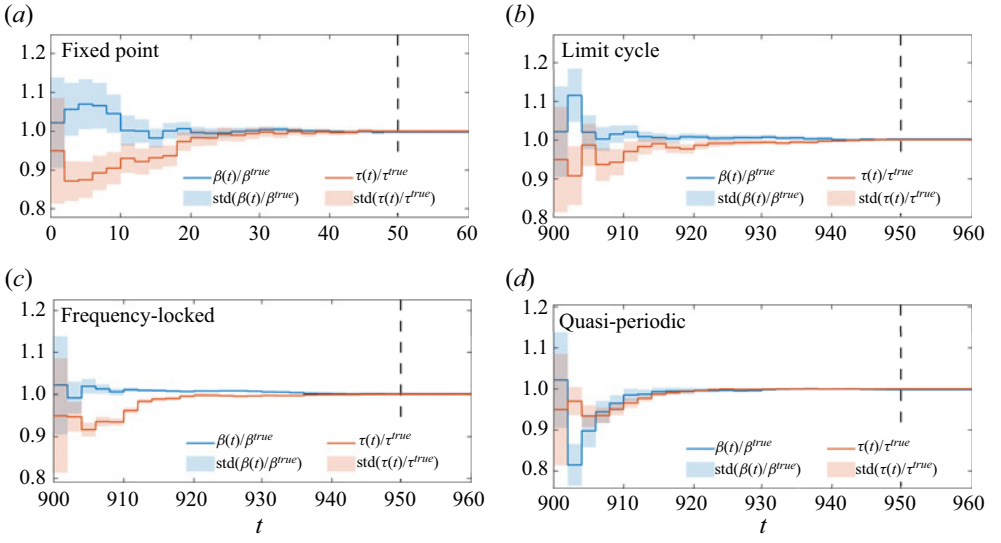


Figure 10. Real-time learning of the parameters and the state. Assimilation of acoustic modes for combined state and parameter estimation of non-chaotic regimes: (*a*) transient towards a fixed point ($\beta^{true} = 0.2$); (*b*) limit cycle ($\beta^{true} = 0.4$); (*c*) frequency-locked solution ($\beta^{true} = 7.7$); and (*d*) quasi-periodic solution ($\beta^{true} = 3.6$). The dashed vertical line indicates when data assimilation ends. Here, $\tau^{true} = 0.2$, $m = 10$, $\sigma_{frac} = 0.25$ and $\Delta t_{analysis} = 2$.

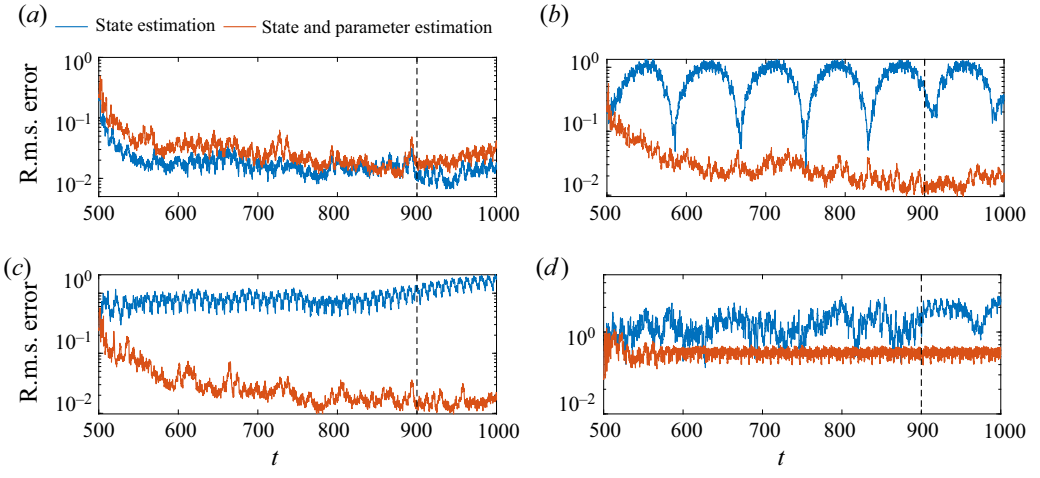


Figure 11. Assimilation of acoustic modes of a quasi-periodic regime. Performance of state estimation (blue) versus combined state and parameter estimation (orange) in a quasi-periodic regime. Initial conditions $\beta = c\,\beta^{true}$ and $\tau = c\,\tau^{true}$, with (*a*) $c = 1$, (*b*) $c = 1.05$, (*c*) $c = 1.25$, (*d*) $c = 1.5$; and $\beta^{true} = 3.6$, $\tau^{true} = 0.2$. The dashed vertical line indicates when data assimilation ends.

from the initial state, such as in the case of figure 11(*a*). Combined state and parameter estimation provides an improved approximation of the solution for the highly uncertain case of 50 % error (figure 11*d*).

6.2. *Assimilation of the acoustic pressure from microphones*

As detailed in § 3.1, we consider the scenario of assimilation of pressure measurements from N_{mic} microphones, located equidistantly from the flame location. This subsection includes results for state estimation (§ 6.2.1) and combined state and parameter estimation (§ 6.2.2).

6.2.1. *State estimation*

We consider a tube that is equipped with $N_{mic} = 6$ microphones, which measure multiple frequency contributions in the signal. This value is chosen from the literature in thermoacoustic experiments (Garita, Yu & Juniper 2021). Figure 12 shows the acoustic pressure at the flame location of the true solution, the unfiltered solution and the filtered solution, as well as their phase space reconstructions and first return maps (the calculation procedure can be found in the supplementary material). In nonlinear regimes, the algorithm learns successfully the pressure state. The accuracy of the solution is lower than in the assimilation of the acoustic modes of § 6.1.1 because here, less information on the state is assimilated. (The filter is not designed for statistically non-stationary problems, which is why the transient fixed point solution is not fully learnt by the filter.)

The effect of the experimental uncertainty is analysed by varying the microphones' standard deviation. Physically, the errors are larger than those in figure 9 because here, we are assimilating 6 components of the augmented state vector out of 36 components, whereas in § 6.1.1, the filter assimilates 20 out of the 30 components of the state vector. Figures 13(*a*,*c*) show that after about 20 analysis steps, the filter follows the model more closely than the observations for larger observation uncertainties. (In other words, the filtered solution 'trusts' the prediction from the model more than the observations when the experimental uncertainty is high.) We set $\sigma_{mic} = 0.01$ in the following simulations, which models experimental microphone uncertainties (de Domenico, Rolland & Hochgreb 2017). The relative error is higher than 20 % for this case (figure 13*a*). Increasing the frequency of analysis allows for a faster convergence with a smaller relative error (figures 13*b*,*d*). With a time between analysis $\Delta t_{analysis} = 1.5$ or 1, the relative error of the filtered solution becomes less than 10 % in only 10 time units, approximately. Thus for the assimilation of microphone pressure data, a higher frequency of analysis is more suitable. We choose the time between analysis as 1.5 time units. The evolution of the trace of the forecast covariance matrix indicates that the spread of the ensemble shrinks rapidly (figures 13*c*,*d*). Besides, the spread is two orders of magnitude smaller than in the assimilation of the modes (figures 9*c*,*d*), and remains small even with large relative errors. Physically, this is because the acoustic modes are updated directly in the modes assimilation, but in this case, the acoustic modes are unobserved variables that are updated indirectly through the microphone pressure observations.

6.2.2. *Combined state and parameter estimation*

The parameters β and τ are updated by the EnSRKF at each analysis step, which occurs every 1.5 time units. Figures 14(*a*,*b*) show that for an ensemble of ten members, the solution converges to the parameters $\beta \approx 6.6$ and $\tau \approx 0.4$, which correspond to a chaotic solution (see figure 6). Nevertheless, the true solution is a quasi-periodic oscillator with $\beta = 3.6$ and $\tau = 0.2$. This means that the filtered solution not only converges to different parameters, but also belongs to a different nonlinear regime than that of the true solution. Physically, this occurs because thermoacoustic dynamics experiences several

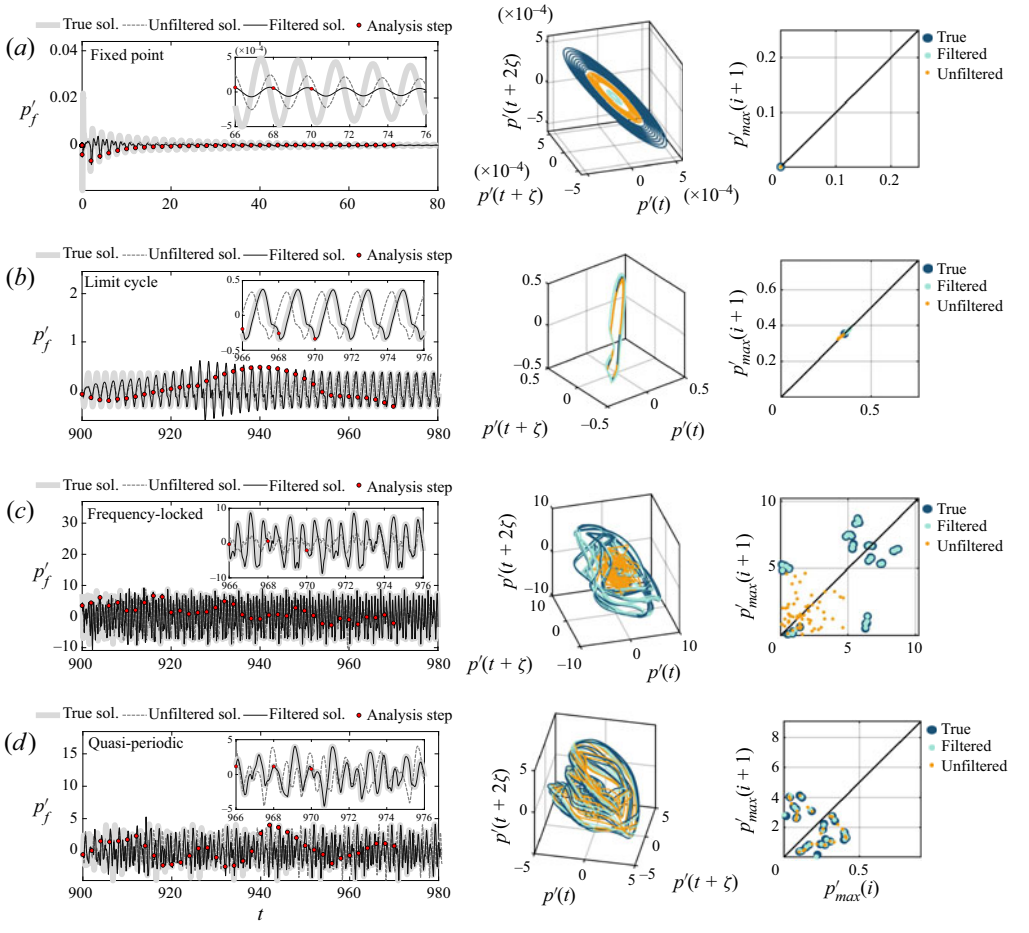


Figure 12. Real-time learning of the state. Assimilation of acoustic pressure from microphones for state estimation of non-chaotic regimes: (*a*) transient towards a fixed point ($\beta = 0.2$); (*b*) limit cycle ($\beta = 0.4$); (*c*) frequency-locked ($\beta = 7.7$); and (*d*) quasi-periodic ($\beta = 3.6$). Left-hand plots: true pressure oscillations at the flame location (light grey), unfiltered solution (dashed dark grey) and filtered solution (black). The analysis time steps are indicated with red circles. Centre and right-hand plots: phase portrait and first return map of the true (dark blue), filtered (light blue) and unfiltered (orange) solutions. Here, $m = 10$, $\sigma_{mic} = 0.01$ and $\Delta t_{analysis} = 2$.

bifurcations in short ranges of β and τ (figure 7). This makes the sampling of nonlinear thermoacoustics challenging. A way to circumvent this is to increase the ensemble size. A parametric study of the effect of the number of realisations is shown in figure 14. Ten ensemble members are not sufficient to learn the reference solution; however, the larger the ensemble, the faster the EnSRKF converges to the true solution.

Occasionally, the EnSRKF provides unphysical parameters as the solution of the optimisation problem, such as negative heat-source strength as the solution of the optimisation problem. To avoid this, we reject the analysis steps that give unphysical solutions and continue the forecast with no assimilation. This means that we are left-truncating the Gaussian. Thus the parameters remain constant until the EnSRKF gives a physical solution to the optimisation problem. (*Ad hoc* ways to bound parameters can be designed (Li *et al.* 2019). This is beyond the scope of this work.) The thresholds

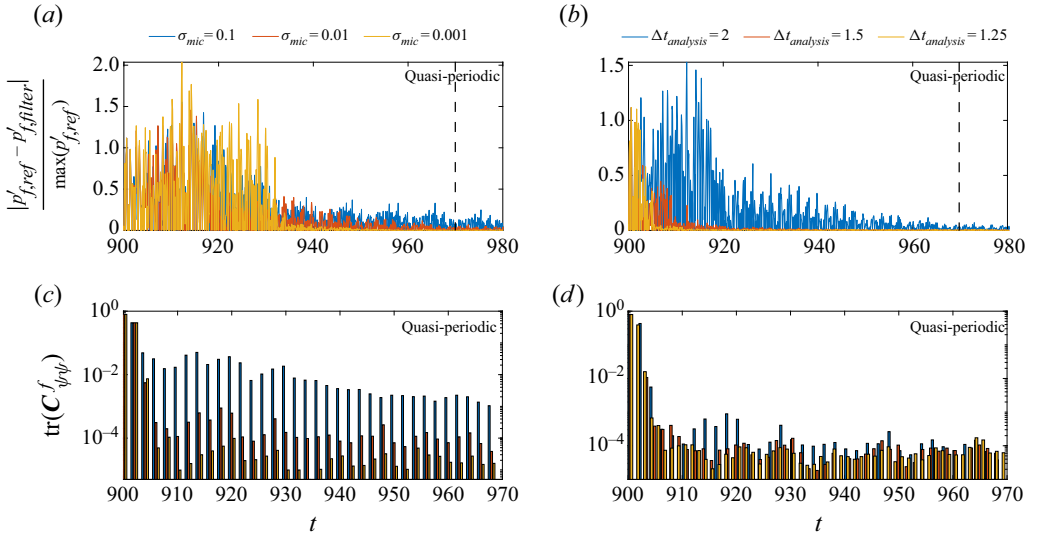


Figure 13. Assimilation of pressure from microphones for state estimation of a quasi-periodic regime. Performance metrics. (*a*,*c*) Effect of the microphones' standard deviation with $\Delta t_{analysis} = 2$. (*b*,*d*) Effect of the assimilation frequency with $\sigma_{mic} = 0.01$. The error evolution is shown with the relative difference between the true and filtered solutions (*a*,*b*) and the trace of the ensemble covariance (*c*,*d*). The dashed vertical line indicates when data assimilation ends. Here, $\beta = 3.6$ and $m = 10$.

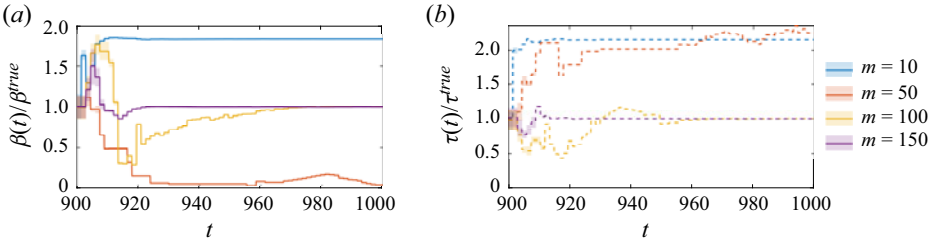


Figure 14. Real-time learning of the parameters. Assimilation of acoustic pressure from microphones for combined state and parameter estimation of a quasi-periodic solution: (*a*) normalised β; (*b*) normalised τ. Here, $N_{mic} = 6$, $\beta^{true} = 3.6$, $\tau^{true} = 0.2$, $\Delta t_{analysis} = 1.5$ and $\sigma_{mic} = 0.01$. The shaded areas show the standard deviation, which becomes smaller as more data are assimilated.

for rejection are defined as $\beta \in [0.1, 10]$ and $\tau \in [0.005, 0.8]$. Because the rejection is effectively reducing the amount of information that can be assimilated, the ensemble convergence slows down. This increase and reject approach is not always sufficient to reach convergence. Figures 15(*a*,*b*) show the same simulation as in figure 14 with more microphones, $N_{mic} = 15$, and $\Delta t_{analysis} = 1$. In this case, the filtered solution is not converging even for 150 ensemble members, which is caused by covariance collapse. To accelerate the convergence and overcome the spurious correlations of finite-sized ensembles (Evensen 2009), we introduce a covariance inflation to the ensemble forecast when the solution of the analysis step provides unfeasible parameters. The inflation method can be used to counteract the variance reduction due to the spurious correlations, and force the model to explore more states. Here, we include the model uncertainty as stochastic

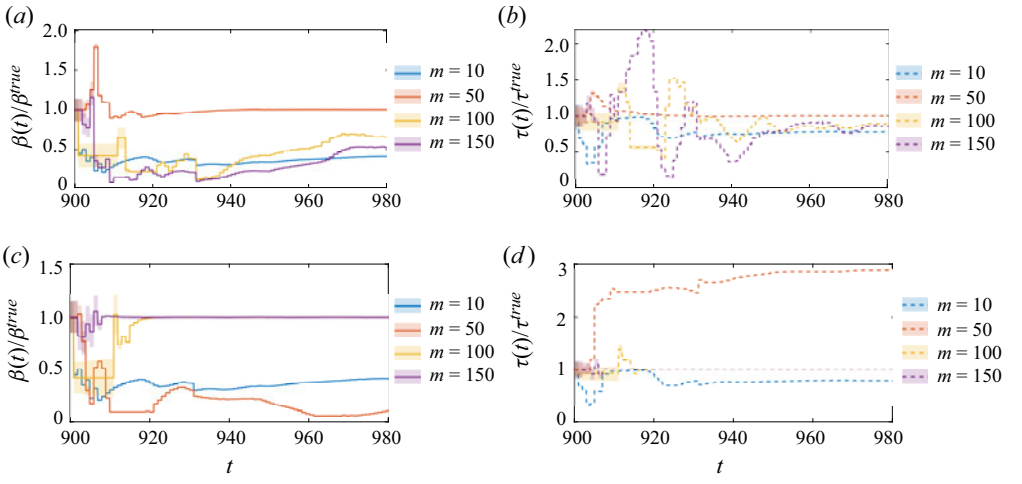


Figure 15. Real-time learning of the parameters. Assimilation of acoustic pressure from microphones for combined state and parameter estimation of a quasi-periodic solution: (*a*,*c*) normalised β; (*b*,*d*) normalised τ. Effect of ensemble size is shown without inflation (*a*,*b*), and with inflation using $\rho = 1.02$ (*c*,*d*). Here, $N_{mic} = 15$, $\beta^{true} = 3.6$, $\tau^{true} = 0.2$, $\Delta t_{analysis} = 1$ and $\sigma_{mic} = 0.01$. The shaded areas show the standard deviation, which becomes smaller as more data are assimilated.

noise by adding an inflation factor ρ to the ensemble forecast:

$$\boldsymbol{A}^f = \bar{\boldsymbol{A}}^f + \rho \boldsymbol{\Psi}^f. \tag{6.2}$$

In this case, $\rho = 1.02$ improves the analysis for the quasi-periodic solution. If necessary, adaptive strategies can be designed following Evensen (2009). Figures 15(*c*,*d*) show the parameters' convergence for the same ensemble sizes as figures 15(*a*,*b*), but with covariance inflation. This is sufficient to remove the plateau caused by the divergence of the EnSRKF to unphysical parameters in large ensembles, thereby speeding up the convergence.

To summarise, we propose an increase, reject, inflate strategy to learn the nonlinear dynamics and parameters of thermoacoustics.

7. Twin experiments in chaotic regimes

This section addresses the assimilation in chaotic regimes. We perform a series of twin experiments with synthetic data using the base parameters of table 1 and the obtained suitable parameters in § 6. Both state estimation and combined state and parameter estimation are tested in the chaotic region F. In the combined state and parameter estimation, the initial conditions for β and τ are sampled from uniform distributions with an upper bound 25 % larger than their true value, and a lower bound 25 % smaller than the true parameters. Different simulations are performed to analyse the predictability of the solutions and to determine a suitable time between analysis ($\Delta t_{analysis}$), which is not trivial in chaotic oscillations.

Figure 16 shows the comparison between the combined state and parameter assimilation solution, an unfiltered solution, and the true state in the chaotic region F of the bifurcation diagram with the same time between analysis as the previous non-chaotic studies. The assimilation does not perform as well as in non-chaotic regimes. This is due physically to the short predictability of chaotic systems.

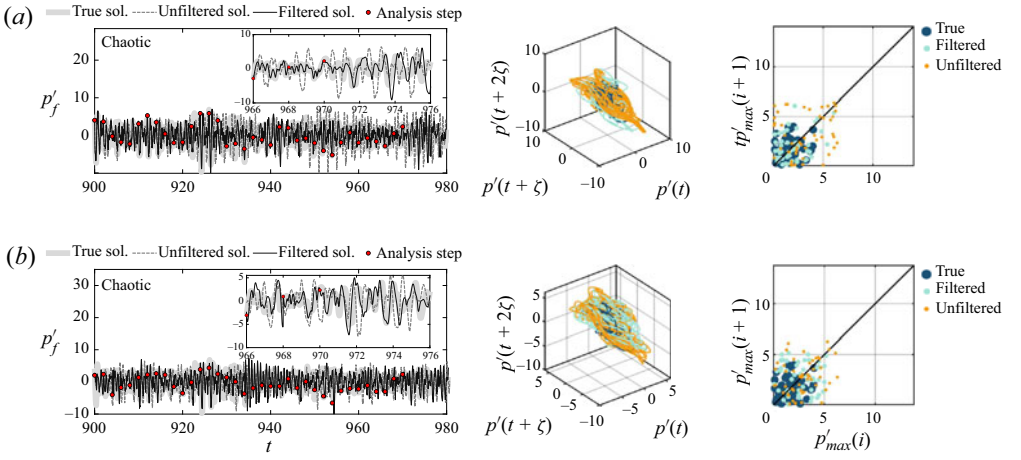


Figure 16. Real-time learning of the state. Assimilation of (*a*) acoustic modes and (*b*) pressure from microphones for state estimation of a chaotic regime ($\beta = 7.0$). Left-hand plots: true pressure oscillations at the flame location (light grey), unfiltered solution (dashed dark grey) and filtered solution (black). The analysis time steps are indicated with red circles. Centre and right-hand plots: phase portrait and first return map of the true (dark blue), filtered (light blue) and unfiltered (orange) solutions. Here, $m = 10$, $\sigma_{mic} = 0.01$, $\sigma_{frac} = 0.25$ and $\Delta t_{analysis} = 2$.

There are several ways to estimate the predictability of a chaotic system (Boffetta *et al.* 2002). Here, the predictability is computed as the inverse of the maximal Lyapunov exponent, which provides a time scale after which two nearby trajectories diverge (linearly) due to the butterfly effect. The methodology followed is described in Magri & Doan (2020). The maximal Lyapunov exponent is determined by analysing the growth of the distance between two nearby trajectories. In a logarithmic scale, the Lyapunov exponent is the slope of the linear region, which is computed by linear regression. Figure 17(*a*) shows two trajectories that are the same until $t_1 = 980$, when they are set apart by $\epsilon = 10^{-6}$. After 10 time units, the two instantaneous solutions are completely different, which is a manifestation of chaos. The logarithmic evolution of the distance between the two trajectories is shown in figure 17(*b*), where the slope of the linear region gives the dominant Lyapunov exponent. This method is carried out for several initial conditions in the attractor. The resulting maximal Lyapunov exponent is $\lambda_1 = 0.74 \pm 0.30$, which corresponds to a predictability time scale $t_\lambda = \lambda_1^{-1} = 1.62 \pm 0.78$. Physically, the predictability t_λ is key to the implementation of the EnSRKF for time-accurate predictions because if the time between analysis is too large, then the forecast ensemble will already be far from the truth. Figure 16 shows how the filtered chaotic solution with an assimilation time on the high end of the time scale t_λ is completely different to the true solution.

Figure 18 shows the effect of the time between analysis $\Delta t_{analysis}$ in the chaotic assimilation. The EnSRKF time-accurately learns the true solution for $\Delta t_{analysis} < t_\lambda$ only as the relative error and the trace of the covariance are reduced significantly and converge. Therefore, we consider a time between analysis $\Delta t_{analysis} = 0.5$ for chaotic regions. (The butterfly effect is not present in non-chaotic behaviours, therefore the time between analysis considered in the fixed point, limit cycle, frequency-locked and quasi-periodic cases can be increased to reduce the computation time, as long as the Nyquist–Shannon criterion is fulfilled (Traverso & Magri 2019).)

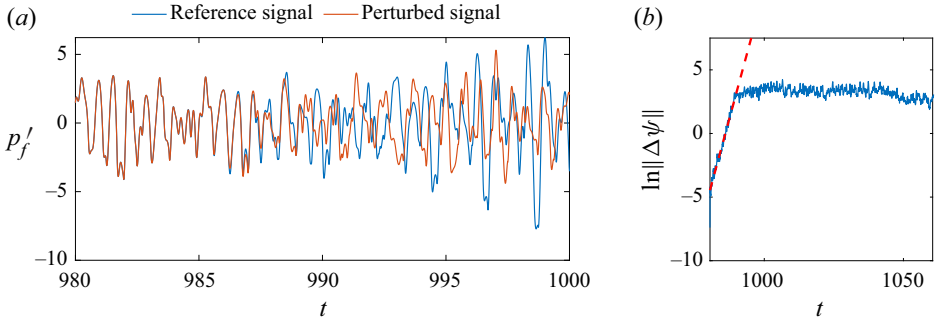


Figure 17. Calculation of the Lyapunov exponent to select the analysis time in data assimilation. (*a*) Time evolution of the pressure oscillations at the flame location of two nearby chaotic solutions. (*b*) Logarithmic growth of the trajectory separation.

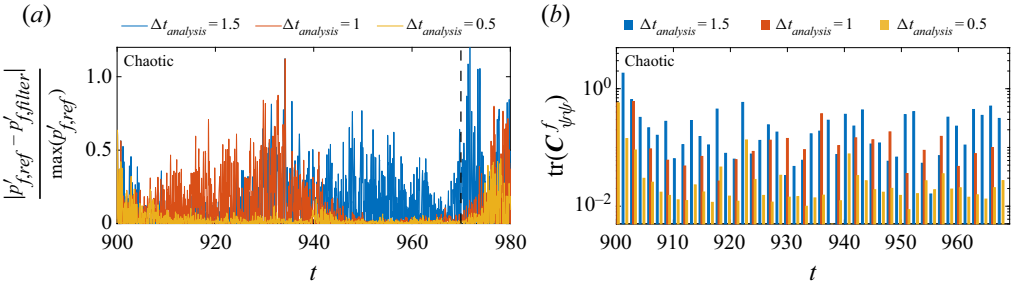


Figure 18. Assimilation of acoustic modes for state estimation of a chaotic regime. Performance metrics. Effect of the assimilation frequency. (*a*) Relative difference between the filtered solutions and truth. (*b*) The trace of the ensemble covariance. The dashed vertical line indicates when data assimilation ends. Here, $\beta = 7.0$, $m = 10$ and $\sigma_{frac} = 0.25$.

Figure 19 shows the results of state estimation in a chaotic regime. The assimilation of the acoustic modes is shown in figure 19(*a*), while the assimilation of pressure observations is shown in figure 19(*b*). The results are generated with an ensemble size $m = 100$. The results indicate that the filter learns the pressure state in chaotic regimes for the two assimilation approaches. Because of the butterfly effect, the filtered pressure and the true signal start differing after removing the filtering due to the chaotic nature of the solutions. The agreement is also evident in the phase space reconstruction and first return map. Figure 20 shows the results of state estimation in the form of the power spectral density (PSD). The top PSDs are computed during the assimilation window ($t \in [900, 1200]$), and the bottom PSDs are computed after removing the filter and propagating the filtered solution without data assimilation ($t \in [1200, 1500]$). The PSDs during the assimilation indicate that the filter learns as well almost exactly the frequency spectrum of the solution, while the unfiltered solution exhibits significant discrepancies. After removing the filter, the PSDs of the true and filtered solutions remain qualitatively similar, but differ slightly due to the chaotic divergence of the solution.

Finally, the data assimilation algorithm is able to estimate β and τ in the combined state and parameter estimation in chaotic regimes for the assimilation of both acoustic modes and pressure from microphones (figures 21*a*,*b*, respectively). The results indicate that there is a successful convergence of the parameters even though their initial uncertainty is large. These simulations are performed with a large ensemble of 300 members and by

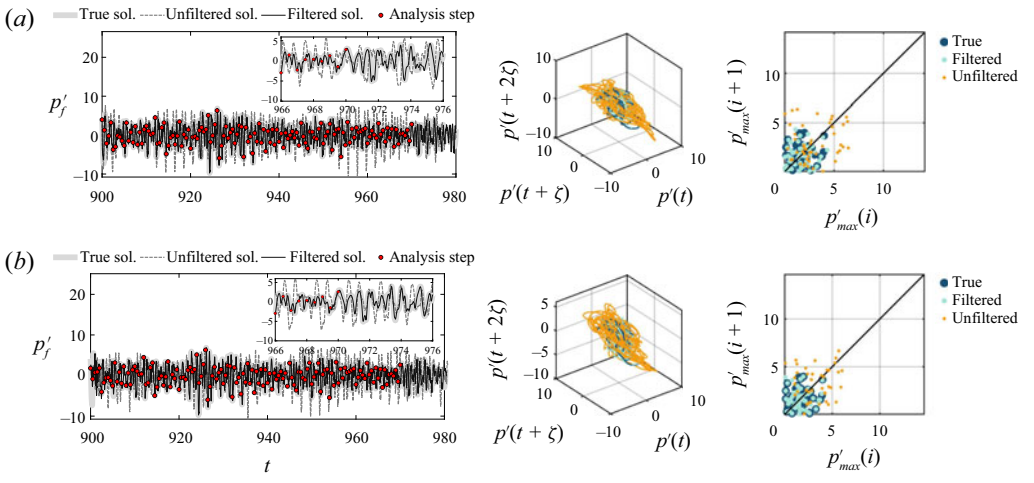


Figure 19. Real-time learning of the state. Assimilation of (*a*) acoustic modes and (*b*) pressure from microphones for state estimation of a chaotic regime ($\beta = 7.0$). Left-hand plots: true pressure oscillations at the flame location (light grey), unfiltered solution (dashed dark grey) and filtered solution (black). The analysis time steps are indicated with red circles. Centre and right-hand plots: phase portrait and first return map of the true (dark blue), filtered (light blue) and unfiltered (orange) solutions. Here, $m = 100$, $\sigma_{mic} = 0.01$, $\sigma_{frac} = 0.25$ and $\Delta t_{analysis} = 0.5$.

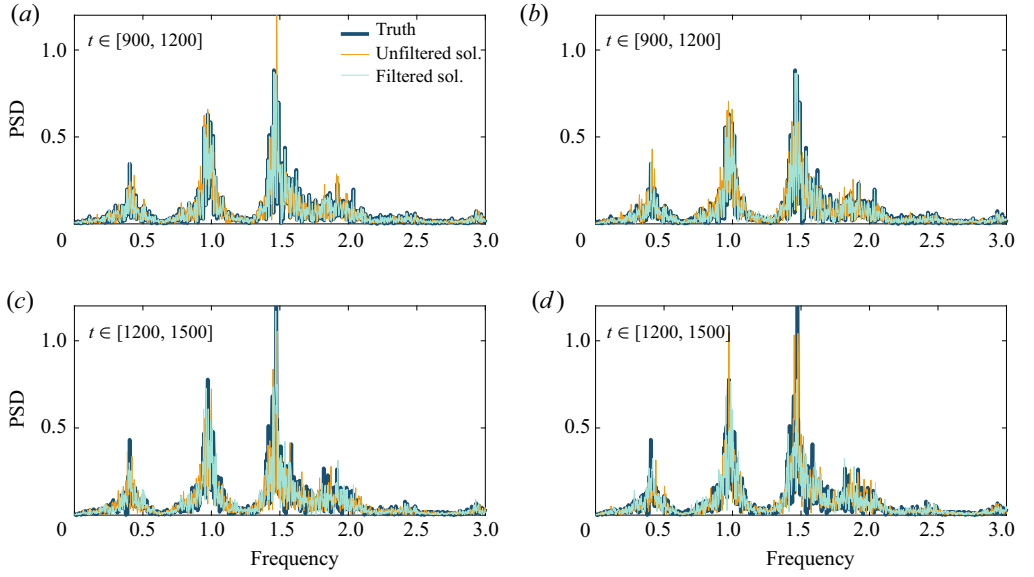


Figure 20. Power spectral density (PSD) during (*a*,*b*) and after (*c*,*d*) assimilation of the true pressure oscillations (dark blue), unfiltered solution (orange) and filtered solution (light blue), during state estimation in a chaotic regime ($\beta = 7.0$). (*a*,*c*) Assimilation of acoustic modes. (*b*,*d*) Assimilation of pressure from microphones. Here, $m = 100$, $\sigma_{mic} = 0.01$, $\sigma_{frac} = 0.25$ and $\Delta t_{analysis} = 0.5$.

inflating the ensemble when the assimilation is neglected due to unphysical parameters. The inflation parameter required for convergence in the assimilation of pressure data (figure 21*b*) is large ($\rho = 1.2$). Figure 22(*b*) shows that the convergence is significantly faster and requires a smaller inflation ($\rho = 1.02$) if the number of microphones is increased

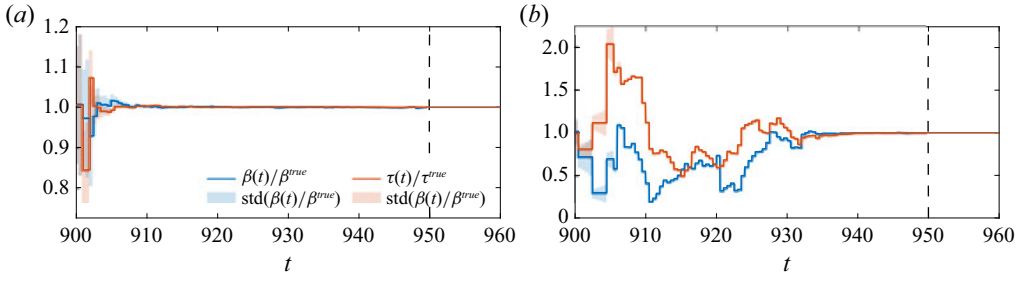


Figure 21. Real-time learning of the parameters. Assimilation of (*a*) acoustic modes and (*b*) pressure from microphones for combined state and parameter estimation of a chaotic regime. Time evolution of the parameters and their standard deviation. Chaotic solution ($\beta = 7.0$). The dashed vertical line indicates when data assimilation ends. Here, $m = 300$, $\rho = 1.2$, $\sigma_{mic} = 0.01$, $\Delta t_{analysis} = 0.5$ and $N_{mic} = 6$.

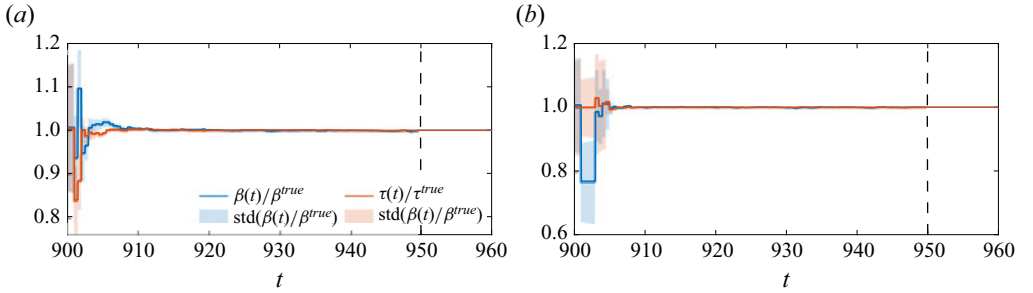


Figure 22. Real-time learning of the parameters. Assimilation of (*a*) acoustic modes and (*b*) pressure from microphones for combined state and parameter estimation of a chaotic regime. Time evolution of the parameters and their standard deviation. Chaotic solution ($\beta = 7.0$). The dashed vertical line indicates when data assimilation ends. Here, $m = 300$, $\rho = 1.02$, $\sigma_{mic} = 0.01$, $\Delta t_{analysis} = 0.5$ and $N_{mic} = 15$.

to 15, as they provide a greater amount of information on the system, i.e. the problem is less ill-conditioned.

The data assimilation successfully learns the true state and parameters for chaotic regimes in the twin experiments by increasing the assimilation frequency, the ensemble size and the inflation parameter.

8. Bias-aware data assimilation with echo state networks

The ESN is trained with the bias, which is the difference between the statistically stationary solution of the travelling-wave model with the flame kinematic model (§ 4.3) and a single statistically stationary realisation of the low-order model with $\tilde{\beta}_{train} = 10^6$ (W s$^{1/2}$ m$^{-5/2}$). The training set is short (1.2 s), and sampled at 2000 Hz. This sampling frequency is consistent with experimental works (e.g. Garita *et al.* 2021). The washout consists of 0.025 s of observations sampled at the same frequency. Following the results from § 6, the simulations in this section use ensemble size $m = 100$, inflation factor $\rho = 1.02$, microphone standard deviation $\sigma_{mic} = 0.01$, and time between analysis $\Delta\tilde{t}_{analysis} = 3 \times 10^{-3}$ s.

Figure 23 shows the unbiased state estimation results. The biased filtered solution represents the expectation of the forecast pressure, while the unbiased filtered solution is the resulting pressure state after correcting the model bias with the ESN. The amplitude

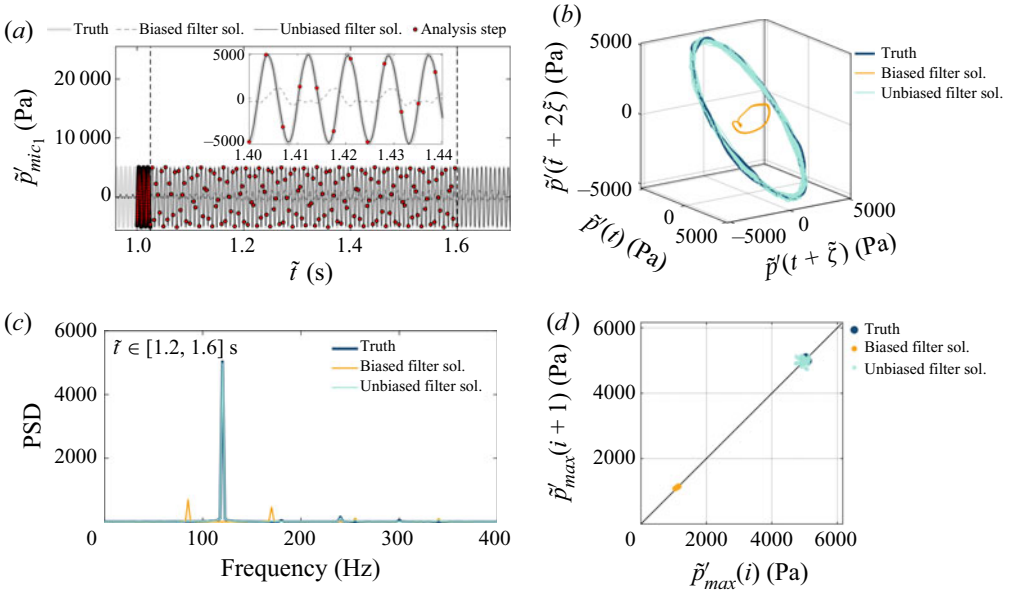


Figure 23. Real-time learning of the model bias and unbiased state. (*a*) True pressure oscillations at $\tilde{x}/\tilde{L}_0 = 0.18$ (light grey), biased filtered solution (dashed dark grey) and unbiased filtered solution (black). The analysis time steps are indicated with red circles; the vertical dashed lines show the assimilation window. (*b*) Phase portrait, (*c*) power spectral density (PSD), and (*d*) first return map of the true pressure oscillations (dark blue), biased filtered solution (orange) and unbiased filtered solution (light blue).

of the pressure oscillations outputted by the higher-fidelity model is about 5 % of the atmospheric pressure, which is a typical pressure level for oscillations in a ducted flame (Bloxsidge, Dowling & Langhorne 1988). Because the Mach number and acoustic pressure are small, the linear assumption on the acoustics is justified. After 1 s of assimilation-free forecasting, when both the truth and the low-order solutions are statistically stationary, the ESN is initialised with 0.025 s of washout. At $\tilde{t} = 1.025$ s, the state estimation begins. The results indicate that the ESN estimates the model bias favourably. This allows the EnSRKF to recover the true pressure time series, as well as to learn its frequency spectrum and the attractor (figure 23*b*). Figure 24 shows the performance of the ESN at the start of unbiased state estimation, and after the data assimilation ends. After the open-loop initialisation, the agreement between the estimated bias and the actual bias is favourable. In the autonomous evolution (closed loop), a re-initialisation every 3×10^{-3} s is sufficient to maintain the accuracy on the inferred bias. The ESN is trained with the bias resulting from a simulation using $\tilde{\beta} = \tilde{\beta}_{train}$, so it is expected to provide good estimates of the bias when initialising the ensemble to the same value of $\tilde{\beta}$, provided that there is a long enough washout.

The results for combined unbiased state and parameter estimation are shown in figure 25. The values of heat-source strength for the ensemble members are initialised far from the training $\tilde{\beta}_{train}$, as a uniform random distribution with 10 % standard deviation and mean value $\tilde{\beta} = 3\,\tilde{\beta}_{train}$. The data assimilation with the ESN algorithm converges to a physical value of heat-source strength ($\tilde{\beta} = 1.70 \times 10^6$ (W s$^{1/2}$ m$^{-5/2}$) in figure 25(*a*), which recovers the dominant frequencies of the higher-fidelity simulation (figure 25*b*). By comparing figures 26 and 24, it can be seen that the time series of the model

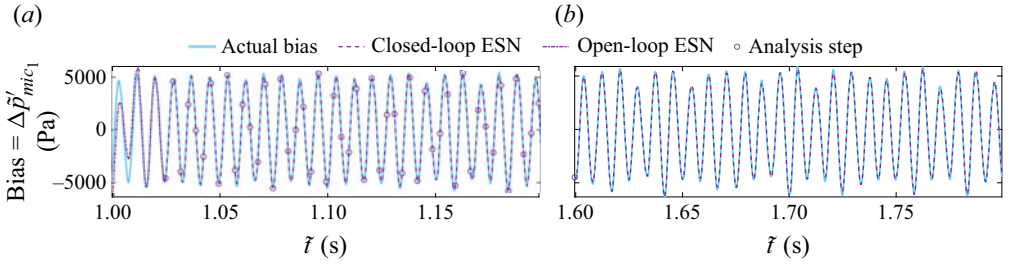


Figure 24. Real-time learning of the model bias and unbiased state. Comparison of the actual bias and the estimated model bias from the ESN during state estimation. (*a*) Open-loop initialisation and sequential closed loops re-initialised at analysis steps. (*b*) Echo state network's closed-loop forecast after data assimilation ends.

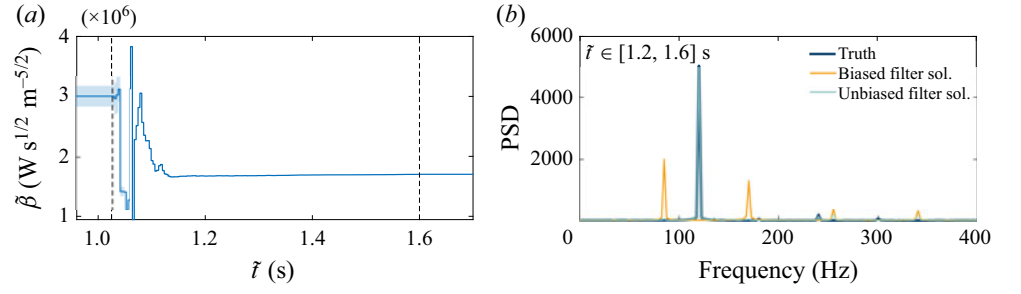


Figure 25. Real-time learning of the model bias, unbiased state and parameters. (*a*) Time evolution of the parameter $\tilde{\beta}$ with the standard deviation. (*b*) Power spectral density (PSD) of the pressure oscillations at $\tilde{x}/\tilde{L}_0 = 0.18$ of the true pressure oscillations (dark blue), biased filtered solution (orange) and unbiased filtered solution (light blue). The vertical dashed lines show the assimilation window.

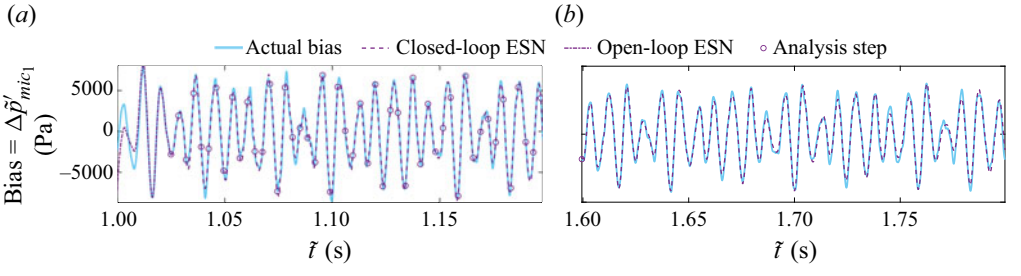


Figure 26. Real-time learning of the model bias, unbiased state and parameters. Comparison of the actual bias and the estimated model bias from the ESN during state and parameter estimation. (*a*) Open-loop initialisation and sequential closed loops re-initialised at analysis steps. (*b*) Echo state network's closed-loop forecast after data assimilation ends.

bias for a low-order model with $\tilde{\beta} = 1.70 \times 10^6$ (W s$^{1/2}$ m$^{-5/2}$) is significantly different from that of $\tilde{\beta}_{train}$. This means that, although the ESN is trained on data with a fixed $\tilde{\beta}_{train} = 10^6$ (W s$^{1/2}$ m$^{-5/2}$), it is able to infer adaptively the bias of unseen data (with a different $\tilde{\beta}$).

9. Conclusions

Low-order thermoacoustic models are qualitatively correct, but they can be quantitatively incorrect. In this work, we introduce data assimilation to make qualitative models quantitatively (more) accurate. This is achieved by combining the knowledge from observations, such as experimental data, and a physical model prediction. Data and model predictions are combined with a Bayesian data assimilation method. The algorithm learns the state, such as the acoustic pressure, and the parameters of the model, every time that reference data become available (real-time).

First, we develop a sequential data assimilation algorithm based on the ensemble square-root Kalman filter in the time domain. This nonlinear filter selects the most likely state and set of physical parameters that are compatible with model predictions and their uncertainties, and with observations and their uncertainties. The filter is physical, i.e. it is not a purely machine learning technique, as it provides estimates that are compatible with the conservation laws, which makes it robust and principled. The data, once assimilated, do not need to be stored. For the data assimilation, which is based on a Markov assumption, we transform the time-delayed dynamics (non-Markovian) into an initial value problem (Markovian). Second, we perform twin experiments in each region of the bifurcation diagram with reference data on (i) the acoustic Galerkin modes, and (ii) the acoustic pressure taken from multiple microphones. On the one hand, in non-chaotic oscillations, the frequency with which data should be assimilated needs to fulfil the Nyquist–Shannon criterion with respect to the dominant acoustic mode. On the other hand, in chaotic oscillations, we highlight that the assimilation frequency should scale with the Lyapunov exponent. During the combined state and parameter estimation with pressure observations, it is observed that occasionally, the filter provides unphysical solutions, such as negative time delays, that lead to convergence to incorrect solutions. This is due to the bifurcations and hystereses that occur in a small range of parameters. Hence we propose an increase, reject, inflate strategy to overcome this. In detail, we increase the ensemble size to better capture the correct dynamics, we reject the analysis steps that provide unphysical parameters, e.g. negative time delays, and we inflate the ensemble covariance by adding noise as a regularisation term. With the twin experiments in data assimilation, we show that: (i) the correct acoustic pressure and parameters can be learnt accurately (i.e. inferred); (ii) the ensemble size is small (in contrast to standard Kalman filters), from 10 to 100 depending on the multi-frequency content of the solution; (iii) the learning is robust because it can tackle large uncertainties in the observations (up to 50 % of the mean values); (iv) the uncertainty of the prediction and parameters is naturally part of the output; and (v) both the time-accurate solution and statistics (through the power spectral density function) can be learnt successfully. Third, we propose a data assimilation framework to learn the model error (bias). The model bias is inferred by an ESN, which is a data-driven tool that is more general than an auto-regressive model. We perform data assimilation using reference data from a higher-fidelity acoustic model, which contains a mean flow, non-ideal boundary conditions, and a kinematic model for the flame. The ESN is trained *a priori*, and then it is run in parallel with the sequential data assimilation algorithm. We show that with a short training set, the reservoir learns the dynamics of the thermoacoustic model error. The proposed methodology learns successfully in real time both the time-accurate solution and the statistics of it.

The technology developed in this paper is being applied to improve the quantitative accuracy of reduced-order models with experimental data from pressure sensors, to learn different model parameters, and to provide estimates of the model error. Data assimilation

with an ESN opens up new opportunities for real-time prediction of thermoacoustics by combining physical knowledge and data synergistically, as well as for estimating the model bias beyond the field of thermoacoustics.

Supplementary material. Supplementary material is available at https://doi.org/10.1017/jfm.2022.653.

Acknowledgements. The authors are grateful to F. Huhn, who helped the authors to produce figure 6, A. Racca, who helped to implement the echo state network algorithm, and J. Aguilar, who helped to set up the kinematic flame code.

Funding. A.N. is supported financially by Rolls-Royce, the EPSRC-DTP and the Cambridge Commonwealth, European and International Trust. L.M. gratefully acknowledges support from the RAEng Research Fellowships Scheme and the ERC Starting Grant PhyCo 949388, and the visiting fellowship at the Technical University of Munich – Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763.

Declaration of interests. The authors report no conflict of interest.

Code availability. The code developed for in this study is openly available at MagriLab/Real-time-TA-DA.

Author ORCIDs.
A. Nóvoa https://orcid.org/0000-0003-0597-8326;
L. Magri https://orcid.org/0000-0002-0657-2611.

Appendix A. Derivation of the EnSRKF

Before starting with the derivation of the filter, some definitions are introduced. For m ensemble members and a state vector $\boldsymbol{\psi}_i \in \mathbb{R}^{N\times 1}$, the matrix that encapsulates the ensemble members and the ensemble mean are defined as

$$\boldsymbol{A} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_m) \in \mathbb{R}^{N\times m} \quad \text{and} \quad \bar{\boldsymbol{\psi}} \approx \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{\psi}_i. \tag{A1a,b}$$

With these, the following definition for the ensemble perturbation matrix applies:

$$\boldsymbol{\Psi} = (\boldsymbol{\psi}_1 - \bar{\boldsymbol{\psi}}, \boldsymbol{\psi}_2 - \bar{\boldsymbol{\psi}}, \ldots, \boldsymbol{\psi}_m - \bar{\boldsymbol{\psi}}). \tag{A2}$$

The ensemble covariance matrix can be determined from (A3), introducing a factor $(m-1)$ to avoid a sample bias – the covariance matrix is defined as an approximation because it is derived from a statistical sample:

$$\boldsymbol{C}_{\psi\psi} \approx \frac{1}{m-1}\boldsymbol{\Psi}\,\boldsymbol{\Psi}^{\mathrm{T}}. \tag{A3}$$

Accounting for these definitions, the Kalman filter update (2.12a) for the ensemble is, in matrix form,

$$\boldsymbol{A}^a = \boldsymbol{A}^f + \left(\boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\right)^{\mathrm{T}}\left[\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\,\boldsymbol{C}^f_{\psi\psi}\,\boldsymbol{M}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{Y} - \boldsymbol{M}\,\boldsymbol{A}^f\right), \tag{A4}$$

where $\boldsymbol{Y} \in \mathbb{R}^{q\times m}$ is the matrix containing the q observations of each member in the ensemble, $\boldsymbol{M} \in \mathbb{R}^{q\times N}$ is the measurement operator matrix, and $\boldsymbol{C}_{\epsilon\epsilon} \in \mathbb{R}^{q\times q}$ is the observations' error covariance matrix.

Using the definition for the ensemble covariance in (A3), the ensemble mean of (A4) is

$$\bar{\boldsymbol{A}}^a = \bar{\boldsymbol{A}}^f + \boldsymbol{\Psi}^f(\boldsymbol{M}\boldsymbol{\Psi}^f)^{\mathrm{T}}[(m-1)\,\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\boldsymbol{\Psi}^f(\boldsymbol{M}\boldsymbol{\Psi}^f)^{\mathrm{T}}]^{-1}(\boldsymbol{Y} - \boldsymbol{M}\bar{\boldsymbol{A}}^f), \tag{A5}$$

where $\bar{\boldsymbol{A}}$ is an $N \times m$ matrix of identical mean analysis states in each column. Now introducing the covariance expression into the analysis error update (see (2.12*b*)) yields the analysis covariance matrix

$$\boldsymbol{C}^a_{\psi\psi} = \frac{\boldsymbol{\Psi}^f\boldsymbol{\Psi}^{f\mathrm{T}}}{m-1} - \left(\boldsymbol{M}\frac{\boldsymbol{\Psi}^f\boldsymbol{\Psi}^{f\mathrm{T}}}{m-1}\right)^{\mathrm{T}}\left[\boldsymbol{C}_{\epsilon\epsilon} + \boldsymbol{M}\frac{\boldsymbol{\Psi}^f\boldsymbol{\Psi}^{f\mathrm{T}}}{m-1}\boldsymbol{M}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{M}\frac{\boldsymbol{\Psi}^f\boldsymbol{\Psi}^{f\mathrm{T}}}{m-1}\right). \tag{A6}$$

Equations (A5) and (A6) can be simplified by introducing the matrices

$$\boldsymbol{S} = \boldsymbol{M}\boldsymbol{\Psi}^f \quad \text{and} \quad \boldsymbol{W}_s = \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} + (m-1)\,\boldsymbol{C}_{\epsilon\epsilon}, \tag{A7a,b}$$

leading to

$$\bar{\boldsymbol{A}}^a = \bar{\boldsymbol{A}}^f + \boldsymbol{\Psi}^f\boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}_s^{-1}(\boldsymbol{Y} - \boldsymbol{M}\bar{\boldsymbol{A}}^f), \tag{A8}$$

$$\boldsymbol{C}^a_{\psi\psi} = \frac{1}{m-1}\boldsymbol{\Psi}^f(\mathbb{I} - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}_s^{-1}\boldsymbol{S})\boldsymbol{\Psi}^{f\mathrm{T}}, \quad \therefore \quad \boldsymbol{\Psi}^a\boldsymbol{\Psi}^{a\mathrm{T}} = \boldsymbol{\Psi}^f(\mathbb{I} - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{W}_s^{-1}\boldsymbol{S})\boldsymbol{\Psi}^{f\mathrm{T}}. \tag{A9a,b}$$

The key idea of the EnSRKF is to find a matrix $\boldsymbol{\Psi}^a$ with the covariance of (A9), which is added to the mean ensemble in (A8) to compute the full ensemble. First, the matrix $\boldsymbol{W}_s$ defined in (A7*a*,*b*) can be eigen-decomposed such that $\boldsymbol{W}_s = \boldsymbol{Z}\boldsymbol{\Lambda}\boldsymbol{Z}^{\mathrm{T}}$ because it is a symmetric square matrix, where $\boldsymbol{\Lambda}$ and $\boldsymbol{Z}$ are the matrices of eigenvalues (diagonal) and eigenvectors (orthogonal), respectively. Substituting the eigen-decomposition into the definition of the analysis perturbation matrix, (A9) is rewritten as

$$\boldsymbol{\Psi}^a\boldsymbol{\Psi}^{a\mathrm{T}} = \boldsymbol{\Psi}^f(\mathbb{I} - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{S})\boldsymbol{\Psi}^{f\mathrm{T}} = \boldsymbol{\Psi}^f(\mathbb{I} - \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})\boldsymbol{\Psi}^{f\mathrm{T}}, \tag{A10}$$

where $\boldsymbol{X} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{S}$. Similarly to the decomposition of $\boldsymbol{W}_s$, the symmetric matrix given by the product $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ can be expressed as $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{V}$ is an orthogonal matrix of eigenvectors, and $\boldsymbol{\Sigma}$ is a diagonal matrix of eigenvalues. Next, introducing this decomposition into (A10) yields

$$\begin{aligned}\boldsymbol{\Psi}^a\boldsymbol{\Psi}^{a\mathrm{T}} &= \boldsymbol{\Psi}^f(\mathbb{I} - \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}})\boldsymbol{\Psi}^{f\mathrm{T}}\\ &= \boldsymbol{\Psi}^f\boldsymbol{V}\,(\mathbb{I} - \boldsymbol{\Sigma})\,\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Psi}^{f\mathrm{T}}\\ &= [\boldsymbol{\Psi}^f\,\boldsymbol{V}(\mathbb{I} - \boldsymbol{\Sigma})^{1/2}\boldsymbol{V}^{\mathrm{T}}][\boldsymbol{\Psi}^f\,\boldsymbol{V}(\mathbb{I} - \boldsymbol{\Sigma})^{1/2}\boldsymbol{V}^{\mathrm{T}}]^{\mathrm{T}}.\end{aligned} \tag{A11}$$

Hence a solution for the analysis ensemble perturbations, which preserves the zero mean in the updated perturbations and keeps the EnSRKF unbiased, is (Sakov & Oke 2008)

$$\boldsymbol{\Psi}^a = \boldsymbol{\Psi}^f\,\boldsymbol{V}\,(\mathbb{I} - \boldsymbol{\Sigma})^{1/2}\,\boldsymbol{V}^{\mathrm{T}}. \tag{A12}$$

Finally, the analysis state of the ensembles is determined by adding the analysis ensemble perturbations to the mean analysis ensembles. This analysis state is then

propagated in time using the nonlinear forecast model, i.e.

$$\boldsymbol{A}^a = \bar{\boldsymbol{A}}^a + \boldsymbol{\Psi}^a, \tag{A13a}$$

$$\boldsymbol{A}^f(t + \Delta t) = \mathcal{F}(\boldsymbol{A}^a(t)), \tag{A13b}$$

where $\mathcal{F}$ is a compact representation of the nonlinear thermoacoustic equations. Note that in the absence of observations, there would be no data assimilation and the initial conditions for the next forecast are the forecast states rather than the analysis states, hence

$$\boldsymbol{A}^f(t + \Delta t) = \mathcal{F}(\boldsymbol{A}^f(t)). \tag{A14}$$

References

Aggarwal, C.C. 2018 *Neural Networks and Deep Learning*, vol. 10, pp. 978–983. Springer.

Aguilar Pérez, J.G. 2019 Sensitivity analysis and optimization in low order thermoacoustic models. PhD thesis, University of Cambridge.

Balasubramanian, K. & Sujith, R.I. 2008 Thermoacoustic instability in a Rijke tube: non-normality and nonlinearity. *Phys. Fluids* **20** (4), 044103.

Bannister, R.N. 2008 A review of forecast error covariance statistics in atmospheric variational data assimilation. I: characteristics and measurements of forecast error covariances. *Q. J. R. Meteorol. Soc.* **134** (637), 1951–1970.

Bannister, R.N. 2017 A review of operational methods of variational and ensemble-variational data assimilation. *Q. J. R. Meteorol. Soc.* **143** (703), 607–633.

Bloxsidge, G.J., Dowling, A.P. & Langhorne, P.J. 1988 Reheat buzz: an acoustically coupled combustion instability. Part 2. Theory. *J. Fluid Mech.* **193**, 445–473.

Boffetta, G., Cencini, M., Falcioni, M. & Vulpiani, A. 2002 Predictability: a way to characterize complexity. *Phys. Rep.* **356** (6), 367–474.

Colburn, C.H., Cessna, J.B. & Bewley, T.R. 2011 State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *J. Fluid Mech.* **682**, 289–303.

Culick, F.E.C. & Kuentzmann, P. 2006 Unsteady motions in combustion chambers for propulsion systems. *Tech. Rep.* RTO-AG-AVT-039. NATO Research and Technology Organization Neuilly-Sur-Seine.

Dee, D.P. & da Silva, A.M. 1998 Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **124** (545), 269–295.

Doan, N.A.K., Polifke, W. & Magri, L. 2021 Short- and long-term predictions of chaotic flows and extreme events: a physics-constrained reservoir computing approach. *Proc. R. Soc. Lond.* A **477** (2253), 20210135.

de Domenico, F., Rolland, E.O. & Hochgreb, S. 2017 Detection of direct and indirect noise generated by synthetic hot spots in a duct. *J. Sound Vib.* **394**, 220–236.

Dowling, A.P. 1995 The calculation of thermoacoustic oscillations. *J. Sound Vib.* **180** (4), 557–581.

Dowling, A.P. 1999 A kinematic model of a ducted flame. *J. Fluid Mech.* **394**, 51–72.

Dowling, A.P. & Morgans, A.S. 2005 Feedback control of combustion oscillations. *Annu. Rev. Fluid Mech.* **37** (1), 151–182.

Drécourt, J.-P., Madsen, H. & Rosbjerg, D. 2006 Bias aware Kalman filters: comparison and improvements. *Adv. Water Resour.* **29** (5), 707–718.

Eckart, C. 1960 *Hydrodynamics of Ocean and Atmosphere*. Pergamon.

Evensen, G. 2003 The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.* **53** (4), 343–367.

Evensen, G. 2009 *Data Assimilation: The Ensemble Kalman Filter*. Springer.

Friedland, B. 1969 Treatment of bias in recursive filtering. *IEEE Trans. Autom. Control* **14** (4), 359–367.

Garita, F., Yu, H. & Juniper, M.P. 2021 Assimilation of experimental data to create a quantitatively accurate reduced-order thermoacoustic model. *Trans. ASME J. Engng Gas Turbines Power* **143** (2), 021008.

Gelb, A. 1974 *Applied Optimal Estimation*. MIT.

Gotoda, H., Ikawa, T., Maki, K. & Miyano, T. 2012 Short-term prediction of dynamical behavior of flame front instability induced by radiative heat loss. *Chaos* **22** (3), 033106.

Gotoda, H., Nikimoto, H., Miyano, T. & Tachibana, S. 2011 Dynamic properties of combustion instability in a lean premixed gas-turbine combustor. *Chaos* **21** (1), 013124.

Grigoryeva, L. & Ortega, J.-P. 2018 Echo state networks are universal. *Neural Netw.* **108**, 495–508.

Guan, Yu., Gupta, V. & Li, L.K.B. 2020 Intermittency route to self-excited chaotic thermoacoustic oscillations. *J. Fluid Mech.* **894**, R3.

HASEGAWA, K., FUKAMI, K., MURATA, T. & FUKAGATA, K. 2020 Machine-learning-based reduced-order modeling for unsteady flows around bluff bodies of various shapes. *Theor. Comput. Fluid Dyn.* **34** (4), 367–383.

HECKL, M.A. 1990 Non-linear acoustic effects in the Rijke tube. *Acustica* **72**, 63–71.

HOUTEKAMER, P.L. & ZHANG, F. 2016 Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Weath. Rev.* **144** (12), 4489–4532.

HUHN, F. & MAGRI, L. 2020*a* Learning ergodic averages in chaotic systems. In *International Conference on Computational Science* (ed. V.V. Krzhizhanovskaya, G. Závodszky, M.H. Lees, J.J. Dongarra, P.M.A. Sloot, S. Brissos & J. Teixeira), vol. 12142. pp. 124–132. Springer.

HUHN, F. & MAGRI, L. 2020*b* Stability, sensitivity and optimisation of chaotic acoustic oscillations. *J. Fluid Mech.* **882**, A24.

HUHN, F. & MAGRI, L. 2022 Gradient-free optimization of chaotic acoustics with reservoir computing. *Phys. Rev. Fluids* **7** (1), 014402.

JAEGER, H. & HAAS, H. 2004 Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* **304** (5667), 78–80.

JAYNES, E.T. 1957 Information theory and statistical mechanics. *Phys. Rev.* **106** (4), 620.

JAYNES, E.T. 2003 *Probability Theory: The Logic of Science*. Cambridge University Press.

JAZWINSKI, A.H. 2007 *Stochastic Processes and Filtering Theory*. Courier Corporation.

JUNIPER, M.P. 2011 Triggering in the horizontal Rijke tube: non-normality, transient growth and bypass transition. *J. Fluid Mech.* **667**, 272–308.

JUNIPER, M.P. & SUJITH, R.I. 2018 Sensitivity and nonlinearity of thermoacoustic oscillations. *Annu. Rev. Fluid Mech.* **50** (1), 661–689.

KABIRAJ, L., SAURABH, A., WAHI, P. & SUJITH, R.I. 2012*a* Route to chaos for combustion instability in ducted laminar premixed flames. *Chaos* **22** (2), 023129.

KABIRAJ, L. & SUJITH, R.I. 2012 Nonlinear self-excited thermoacoustic oscillations: intermittency and flame blowout. *J. Fluid Mech.* **713**, 376–397.

KABIRAJ, L., SUJITH, R.I. & WAHI, P. 2012*b* Bifurcations of self-excited ducted laminar premixed flames. *Trans. ASME J. Engng Gas Turbines Power* **134** (3), 031502.

KALMAN, R.E. 1960 A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Engng* **82** (1), 35–45.

KANTZ, H. & SCHREIBER, T. 2003 *Nonlinear Time Series Analysis*, 2nd edn. Cambridge University Press.

KASHINATH, K., WAUGH, I.C. & JUNIPER, M.P. 2014 Nonlinear self-excited thermoacoustic oscillations of a ducted premixed flame: bifurcations and routes to chaos. *J. Fluid Mech.* **761**, 399–430.

LABAHN, J.W., WU, H., CORITON, B., FRANK, J.H. & IHME, M. 2019 Data assimilation using high-speed measurements and LES to examine local extinction events in turbulent flames. *Proc. Combust. Inst.* **37** (2), 2259–2266.

LANDAU, L.D. & LIFSHITZ, E.M. 1987 *Fluid Mechanics*, 2nd edn. Pergamon.

LEWIS, J.M., LAKSHMIVARAHAN, S. & DHALL, S. 2006 *Dynamic Data Assimilation: A Least Squares Approach*. Encyclopedia of Mathematics and its Applications, vol. 13. Cambridge University Press.

LI, R., JAN, N.M., HUANG, B. & PRASAD, V. 2019 Constrained ensemble Kalman filter based on Kullback–Leibler divergence. *J. Process Control* **81**, 150–161.

LIEUWEN, T.C. 2012 *Unsteady Combustor Physics*. Cambridge University Press.

LIEUWEN, T.C. & YANG, V. 2005 *Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modelling*. American Institute of Aeronautics and Astronautics.

LIVINGS, D.M., DANCE, S.L. & NICHOLS, N.K. 2008 Unbiased ensemble square root filters. *Physica* D **237** (8), 1021–1028.

LUKOŠEVIČIUS, M. 2012 A practical guide to applying echo state networks. In *Neural Networks: Tricks of the Trade* (ed. G. Montavon, G.B. Orr & K.-R. Müller), pp. 659–686. Springer.

MAASS, W., NATSCHLÄGER, T. & MARKRAM, H. 2002 Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* **14** (11), 2531–2560.

MAGRI, L. 2019 Adjoint methods as design tools in thermoacoustics. *Appl. Mech. Rev.* **71** (2), 020801.

MAGRI, L., BALASUBRAMANIAN, K., SUJITH, R.I. & JUNIPER, M.P. 2013 Non-normality in combustion–acoustic interaction in diffusion flames: a critical revision. *J. Fluid. Mech.* **719**, 183–202.

MAGRI, L. & DOAN, N.A.K. 2020 Physics-informed data-driven prediction of turbulent reacting flows with Lyapunov analysis and sequential data assimilation. In *Data Analysis for Direct Numerical Simulations of Turbulent Combustion* (ed. P. Heinz & A. Antonio), pp. 177–196. Springer.

MAGRI, L. & JUNIPER, M. 2014 Global modes, receptivity, and sensitivity analysis of diffusion flames coupled with duct acoustics. *J. Fluid Mech.* **752**, 237–265.

MENSAH, G.A., MAGRI, L., SILVA, C.F., BUSCHMANN, P.E. & MOECK, J.P. 2018 Exceptional points in the thermoacoustic spectrum. *J. Sound Vib.* **433**, 124–128.
NAIR, V. & SUJITH, R.I. 2015 A reduced-order model for the onset of combustion instability: physical mechanisms for intermittency and precursors. *Proc. Combust. Inst.* **35** (3), 3193–3200.
NAIR, V., THAMPI, G. & SUJITH, R.I. 2014 Intermittency route to thermoacoustic instability in turbulent combustors. *J. Fluid Mech.* **756**, 470–487.
NICOUD, F., BENOIT, L., SENSIAU, C. & POINSOT, T. 2007 Acoustic modes in combustors with complex impedances and multidimensional active flames. *AIAA J.* **45** (2), 426–441.
NOIRAY, N. 2017 Linear growth rate estimation from dynamics and statistics of acoustic signal envelope in turbulent combustors. *Trans. ASME J. Engng Gas Turbines Power* **139** (4), 041503.
NOIRAY, N., DUROX, D., SCHÜLLER, T. & CANDEL, S. 2008 A unified framework for nonlinear combustion instability analysis based on the flame describing function. *J. Fluid Mech.* **615**, 139–167.
NOVATI, G., MAHADEVAN, L. & KOUMOUTSAKOS, P. 2019 Controlled gliding and perching through deep-reinforcement-learning. *Phys. Rev. Fluids* **4** (9), 093902.
NÓVOA, A. & MAGRI, L. 2020 A Bayesian approach for predicting and filtering linear and nonlinear thermoacoustic oscillations. *Bull. Am. Phys. Soc.* **65** (13), E04.007.
O'CONNOR, J., ACHARYA, V. & LIEUWEN, T.C. 2015 Transverse combustion instabilities: acoustic, fluid mechanic, and flame processes. *Prog. Energy Combust. Sci.* **49**, 1–39.
ORCHINI, A., ILLINGWORTH, S.J. & JUNIPER, M.P. 2015 Frequency domain and time domain analysis of thermoacoustic oscillations with wave-based acoustics. *J. Fluid. Mech.* **775**, 387–414.
ORCHINI, A., MAGRI, L., SILVA, C.F., MENSAH, G.A. & MOECK, J.P. 2020 Degenerate perturbation theory in thermoacoustics: high-order sensitivities and exceptional points. *J. Fluid Mech.* **903**, A37.
PHAM, D.T. 2001 Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Weath. Rev.* **129** (5), 1194–1207.
PITSCH, H. & DE LAGENESTE, L.D. 2002 Large-eddy simulation of premixed turbulent combustion using a level-set approach. *Proc. Combust. Inst.* **29** (2), 2001–2008.
POINSOT, T. 2017 Prediction and control of combustion instabilities in real engines. *Proc. Combust. Inst.* **36** (1), 1–28.
POLIFKE, W. 2020 Modeling and analysis of premixed flame dynamics by means of distributed time delays. *Prog. Energy Combust. Sci.* **79**, 100845.
POLIFKE, W., PONCET, A., PASCHEREIT, C.O. & DÖBBELING, K. 2001 Reconstruction of acoustic transfer matrices by instationary computational fluid dynamics. *J. Sound Vib.* **245** (3), 483–510.
RACCA, A. & MAGRI, L. 2021 Robust optimization and validation of echo state networks for learning chaotic dynamics. *Neural Netw.* **142**, 252–268.
RAYLEIGH, LORD 1878 The explanation of certain acoustical phenomena. *Nature* **18**, 319–321.
REICH, S. & COTTER, C. 2015 *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press.
RUBIO, P.-B., CHAMOIN, L. & LOUF, F. 2019 Real-time Bayesian data assimilation with data selection, correction of model bias, and on-the-fly uncertainty propagation. *C. R. Méc* **347** (11), 762–779.
RUMELHART, D.E., HINTON, G.E. & WILLIAMS, R.J. 1986 Learning representations by back-propagating errors. *Nature* **323** (6088), 533–536.
SAKOV, P. & OKE, P.R. 2008 Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Weath. Rev.* **136** (3), 1042–1053.
SASAKI, Y. 1955 A fundamental study of the numerical prediction based on the variational principle. *J. Met. Soc. Japan* **33** (6), 262–275.
SCHUERMANS, B. 2003 Modeling and control of thermoacoustic instabilities. EPFL. doi:10.5075/epfl-thesis-2800.
SHAMPINE, L.F. & REICHELT, M.W. 1997 The MATLAB ODE suite. *SIAM J. Sci. Comput.* **18** (1), 1–22.
DA SILVA, A.F.C. & COLONIUS, T. 2018 Ensemble-based state estimator for aerodynamic flows. *AIAA J.* **56** (7), 2568–2578.
DA SILVA, A.F.C. & COLONIUS, T. 2020 Flow state estimation in the presence of discretization errors. *J. Fluid Mech.* **890**, A10.
SILVA, C.F., NICOUD, F., SCHULLER, T., DUROX, D. & CANDEL, S. 2013 Combining a Helmholtz solver with the flame describing function to assess combustion instability in a premixed swirled combustor. *Combust. Flame* **160** (9), 1743–1754.
SUJITH, R.I. & UNNI, V.R. 2020 Complex system approach to investigate and mitigate thermoacoustic instability in turbulent combustors. *Phys. Fluids* **32** (6), 061401.
TARANTOLA, A. 2005 *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

TRAVERSO, T. & MAGRI, L. 2019 Data assimilation in a nonlinear time-delayed dynamical system with Lagrangian optimization. In *Computational Science – ICCS 2019* (ed. J.M.F. Rodrigues, P.J.S. Cardoso, J. Monteiro, R. Lam, V.V. Krzhizhanovskaya, M.H. Lees, J.J. Dongarra & P.M.A. Sloot), vol. 11539. pp. 156–168.

TREFETHEN, L.N. 2000 *Spectral Methods in MATLAB*, vol. 10. SIAM.

XIAO, H., WU, J.-L., WANG, J.-X., SUN, R. & ROY, C.J. 2016 Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier–Stokes simulations: a data-driven, physics-informed Bayesian approach. *J. Comput. Phys.* **324**, 115–136.

YU, H., JUNIPER, M.P. & MAGRI, L. 2019 Combined state and parameter estimation in level-set methods. *J. Comput. Phys.* **399**, 108950.

ZINN, B.T. & LORES, M.E. 1971 Application of the Galerkin method in the solution of non-linear axial combustion instability problems in liquid rockets. *Combust. Sci. Technol.* **4** (1), 269–278.