

# DATA ASSIMILATION AND OPTIMAL CALIBRATION IN NONLINEAR MODELS OF FLAME DYNAMICS

**Hans Yu\***

Department of Engineering  
University of Cambridge  
Cambridge, CB21PZ  
United Kingdom  
Email: hy313@cam.ac.uk

**Thomas Jaravel**

**Matthias Ihme**  
Center for Turbulence Research  
Stanford University  
Stanford, California, 94305

**Matthew P. Juniper**

Department of Engineering  
University of Cambridge  
Cambridge, CB21PZ  
United Kingdom

**Luca Magri**

Department of Engineering  
University of Cambridge  
Cambridge, CB21PZ  
United Kingdom

Visiting Fellow at Institute for Advanced Study  
Technical University of Munich  
85748 Garching  
Germany

## ABSTRACT

We propose an on-the-fly statistical learning method to take a qualitative reduced-order model of the dynamics of a premixed flame and make it quantitatively accurate. This physics-informed data-driven method is based on the statistically optimal combination of (i) a reduced-order model of the dynamics of a premixed flame with a level-set method, (ii) high-quality data, which can be provided by experiments and/or high-fidelity simulations, and (iii) assimilation of the data into the reduced-order model to improve the prediction of the dynamics of the premixed flame. The reduced-order model learns the state and the parameters of the premixed flame on the fly with the ensemble Kalman filter, which is a Bayesian filter used, for example, in weather forecasting. The proposed method and algorithm are applied to two test

cases with relevance to reacting flows and instabilities. First, the capabilities of the framework are demonstrated in a twin experiment, where the assimilated data is produced from the same model as that used in prediction. Second, the assimilated data is extracted from a high-fidelity reacting-flow direct numerical simulation (DNS), which provides the reference solution. The results are analyzed by using Bayesian statistics, which robustly provide the level of confidence in the calculations from the reduced-order model. The versatile method we propose enables the optimal calibration of computationally inexpensive reduced-order models in real time when experimental data becomes available, for example, from gas-turbine sensors.

---

\*Address all correspondence to this author.

## INTRODUCTION

Thermoacoustic instabilities are a persistent challenge in the design of jet and rocket engines. Velocity and pressure oscillations inside the combustion chamber interact with the flame and cause unsteady heat release. If moments of higher heat release coincide with moments of higher pressure (and lower heat release with lower pressure), acoustic oscillations can grow [1]. This can lead to large-amplitude oscillations causing structural damage in the jet or rocket engine [2, 3].

A self-excited thermoacoustic oscillation may display a variety of complex dynamics due to the nonlinear response of the heat source. For an electric heater or a diffusion flame, the nature of the nonlinearity is relatively simple: after a linear growth phase, the acoustic oscillations saturate and form a limit cycle. For a premixed flame, the nature of the nonlinearity becomes more complicated because of the formation of cusps and pinched-off fuel-air pockets. In premixed flames, which are the subject of this study, the thermoacoustic system displays rich nonlinear dynamics, which can even become chaotic through period-doubling or Ruelle-Takens-Newhouse bifurcations [4, 5]. Rich nonlinear dynamics have also been observed in more realistic settings, e.g. a gas-turbine model combustion chamber [6].

The rich nonlinear dynamics of premixed flames is one of the reasons for which their reduced-order modeling is challenging. In addition to nonlinearity, the time-accurate prediction of the flame dynamics is difficult because of

- Aleatoric uncertainties. Under realistic conditions, the flame dynamics are subject to stochastic noise, which cannot be exactly replicated in a simulation. In the worst-case scenario, the stochastic dynamics of the flame may trigger thermoacoustic instabilities before the limit of linear stability is reached [7];
- Epistemic uncertainties (also known as model uncertainties). The flame dynamics are affected by three subsystems: the hydrodynamics, the chemistry and the acoustics. A simulation typically relies on some modeling assumptions, e.g. simplified governing equations, a (relatively) low spatial resolution or a reduced chemical mechanism. The result is a model that may lack relevant degrees of freedom or have inaccurate parameters [8], which need to be calibrated;
- Extreme sensitivity. Both the linear and the nonlinear dynamics of the flame are highly sensitive to uncertain parameters such as boundary conditions and the operating regime, in particular in thermoacoustic systems [9–11].

Most of the reduced-order models available are able to capture the qualitative flame dynamics, but they often fail to be quantitatively accurate [12]. In this paper, we propose a method to make qualitative reduced-order models quantitatively accurate. To achieve this, we augment reduced-order models with data from numerical experiments in the form of high-fidelity simulations using methods based on the theory of stochastic processes,

namely data assimilation and parameter estimation based on a Bayesian approach. Data assimilation gives an optimal estimate of the true state of a system given experimental observations. Parameter estimation optimally calibrates the model on the fly by using the data to find a maximum-likelihood set of parameters for the model. While data assimilation has its root in oceanography and meteorology [13–15], it has in recent years been applied to various other areas in fluid mechanics, including turbulent near-wall flow [16], generation of inflow and initial conditions [17], mean flow reconstruction [18], turbulent flow around aircraft [19], viscous flow around a cylinder [20], optimal sensor placement [21], unsteady separated flow aerodynamics [22] as well as extinction and reignition dynamics in turbulent combustion [23]. The theory of stochastic processes [24] offers a probabilistic, data-driven approach that accurately quantifies the uncertainties in a model and improves its predictions.

This paper is organized as follows: First, the Hamilton-Jacobi equation (the physical model for the flame dynamics) and the ensemble Kalman filter (the statistical model for data assimilation) are introduced. Second, results for two test cases are discussed: a first test case, where the assimilated data is produced from the same model as that used in prediction; and a second test case, where the assimilated data is extracted from a high-fidelity simulation. Third, conclusions and future work are outlined.

## LEVEL-SET METHODS AND DATA ASSIMILATION

A ducted premixed flame subject to upstream velocity forcing is investigated. In the past, its dynamics have been successfully characterized using dynamical system techniques and continuation analysis on a reduced-order model [5, 25]. It has been shown that the kinematics of the flame surface are the major source of nonlinearity in a ducted premixed flame [26].

A physics-informed, data-driven framework for the time-accurate calculation of a ducted premixed flame using the reduced-order model requires two components: (i) a computational method to predict the motion of the flame surface, and (ii) a statistical algorithm to find the optimal estimate from a model prediction and experimental observations. The two components are presented in the following subsections.

### Hamilton-Jacobi equation

The kinematic description of a surface is given by its laws of motion [27]:

$$\frac{d\mathbf{x}}{dt} = \mathbf{u} - s_L \mathbf{n}, \quad (1)$$

where  $\mathbf{x}$  is the position of one point on the surface, and  $\mathbf{n}$  is the normal vector at this point.  $\mathbf{u}$  is the velocity field of the underlying medium, and  $s_L$  is the speed of the surface relative to

the underlying medium. In general,  $s_L$  depends on geometrical properties of the surface, e.g. on the local curvature via the proportionality constant  $\mathcal{L}$ . The laws of motion are equivalent to the following Hamilton-Jacobi equation [28]:

$$\frac{\partial G}{\partial t} + (\mathbf{u} \cdot \mathbf{n} - s_L) = 0, \quad (2)$$

$$\text{subject to } G(\mathbf{x}(t), t) - G(\mathbf{x}(0), 0) = 0, \quad (3)$$

$$\text{and } \sqrt{\nabla G \cdot \nabla G} - 1 = 0. \quad (4)$$

In combustion, Eqn. (2) is also known as the  $G$ -equation [27]. The normal vector  $\mathbf{n}$  depends on the computation of partial derivatives in space, which makes Eqn. (2) a partial differential equation (PDE). Equations (3) and (4) are constraints, which make the solution to the  $G$ -equation unique away from the surface. For the upstream velocity forcing of the premixed flame, the radial and axial components of the underlying velocity field are given by [29]

$$u_r = \frac{2\pi f K r e_a}{U} \cos\left(2\pi f \left(t - \frac{Kz}{U}\right)\right), \quad (5)$$

$$u_z = U + e_a \sin\left(2\pi f \left(t - \frac{Kz}{U}\right)\right), \quad (6)$$

where the base-flow speed is denoted by  $U$ , the frequency by  $f$  and the forcing amplitude by  $e_a$ . Eqns. (5) and (6) are related to each other through the continuity equation. The non-dimensional parameter  $K$  is defined as the ratio between the base-flow speed and the perturbation phase speed. Hence, the perturbation phase speed is given by  $U/K$  and the wave number by  $2\pi f K/U$ . The solution  $G(\mathbf{x}, t)$  to the Hamilton-Jacobi equation is the so-called generating function [30]. The solution  $\mathbf{x}(t)$  to Eqn. (3) gives the location of the surface at every time  $t$ .

There are several consequences to solving the Hamilton-Jacobi equation (Eqns. (2)-(4)) instead of the laws of motion (Eqn. (1)). Firstly, the laws of motion require a parameterization of the surface [31]. The quality of the parameterization quickly deteriorates when the metric on the surface or its topology changes. The Hamilton-Jacobi equation avoids these issues by embedding the surface as a level set of the generating function, which is defined over the entire domain. Secondly,

the Hamilton-Jacobi equation needs to be in theory solved over the entire domain even though surfaces are one-codimensional (i.e. surfaces have one less dimension than the volume they inhabit) [32]. This is more computationally expensive but has the advantage that surfaces can be identified by their generating functions as vectors in a state space. The choice of generating functions is highly relevant as they have to live within a state space suitable for data assimilation. This will be further outlined in the next subsection.

The  $G$ -equation is numerically solved using a computationally inexpensive narrow-band level-set method [33]: The computational domain is discretized using a weighted essentially non-oscillatory (WENO) scheme and a total-variation diminishing (TVD) version of the Runge-Kutta scheme [34–36]. This gives third-order accuracy in time and up to fifth-order accuracy in space. Cusps and pinched-off fuel-air pockets are reliably captured. At the base of the flame, a rotating boundary condition is used [37]. The  $G$ -equation solver has been verified and validated in a number of studies [5, 25, 38–40]. Distance reinitialization in the narrow band is performed using a continuation method [41]. The generating function is reconstructed from the solution to the  $G$ -equation using a fast-marching method [42].

### Ensemble Kalman filter

The Kalman filter provides a statistically optimal estimate  $\psi^a$  (‘analysis’) of the true state  $\psi$  of a system from a model prediction  $\psi^f$  (‘forecast’) and experimental observations  $\mathbf{d}$  [43, 44]. The model prediction is mapped from the state space to the observation space through a linear measurement operator  $\mathbf{M}$ . In practice,  $\mathbf{M}$  is given by the transfer functions of the experimental apparatus and sensors. The prediction uncertainty and the experimental errors are represented by covariance matrices  $\mathbf{C}_{\psi\psi}^f$  and  $\mathbf{C}_{\epsilon\epsilon}$  respectively:

$$\psi^f \sim \mathcal{N}\left(\psi, \mathbf{C}_{\psi\psi}^f\right) \quad , \quad \mathbf{d} \sim \mathcal{N}\left(\mathbf{M}\psi, \mathbf{C}_{\epsilon\epsilon}\right), \quad (7)$$

where  $\mathcal{N}$  denotes the (multivariate) normal distribution with mean and covariance matrix. The statistically optimal estimate  $\psi^a$  of the true state and its uncertainty  $\mathbf{C}_{\psi\psi}^a$  are given by [24]

$$\psi^a = \psi^f + \left(\mathbf{M}\mathbf{C}_{\psi\psi}^f\right)^T \left[\mathbf{C}_{\epsilon\epsilon} + \mathbf{M}\mathbf{C}_{\psi\psi}^f\mathbf{M}^T\right]^{-1} (\mathbf{d} - \mathbf{M}\psi^f), \quad (8)$$

$$\mathbf{C}_{\psi\psi}^a = \mathbf{C}_{\psi\psi}^f - \left(\mathbf{M}\mathbf{C}_{\psi\psi}^f\right)^T \left[\mathbf{C}_{\epsilon\epsilon} + \mathbf{M}\mathbf{C}_{\psi\psi}^f\mathbf{M}^T\right]^{-1} \left(\mathbf{M}\mathbf{C}_{\psi\psi}^f\right). \quad (9)$$

Intuitively speaking, Eqn. (8) corrects the model prediction  $\psi^f$  according to the statistical distance between the experimental observation  $\mathbf{d}$  and the model prediction  $\mathbf{M}\psi^f$ , which is projected

from the state space to the observation space. The statistical distance is weighted according to the prediction and experimental covariance matrices  $\mathbf{C}_{\psi\psi}^f$  and  $\mathbf{C}_{\varepsilon\varepsilon}$ .

For the ducted premixed flame under investigation, the state vectors  $\psi^f$  and  $\psi^a$  are the generating functions  $G(\mathbf{x}, t)$  from the previous subsection. After discretization, the state vectors have  $\mathcal{O}(10^5)$  entries and the covariance matrices, accordingly,  $\mathcal{O}(10^{10})$  entries. The computation and manipulation of the covariance matrices make the Kalman filter computationally infeasible. To bypass the explicit formation of covariance matrices, the prediction uncertainty may alternatively be approximated by an ensemble of state vectors  $\psi_i^f$ , where  $i = 1, 2, \dots, N$  with  $N$  as the ensemble size. The statistically optimal estimates  $\psi^a$  and  $\mathbf{C}_{\psi\psi}^a$  from the ensemble Kalman filter are given by [15]

$$\psi_i^a = \psi_i^f + \left( \mathbf{M} \mathbf{C}_{\psi\psi}^f \mathbf{M}^T \right)^{-1} \left[ \mathbf{C}_{\varepsilon\varepsilon} + \mathbf{M} \mathbf{C}_{\psi\psi}^f \mathbf{M}^T \right]^{-1} \left( \mathbf{d} - \mathbf{M} \psi_i^f \right), \quad (10)$$

$$\bar{\psi} = \frac{1}{N} \sum_{i=1}^N \psi_i \quad , \quad \mathbf{C}_{\psi\psi} = \frac{1}{N-1} \sum_{i=1}^N (\psi_i - \bar{\psi})(\psi_i - \bar{\psi})^T. \quad (11)$$

In Fig. 1, a schematic of this level-set data assimilation framework is shown.

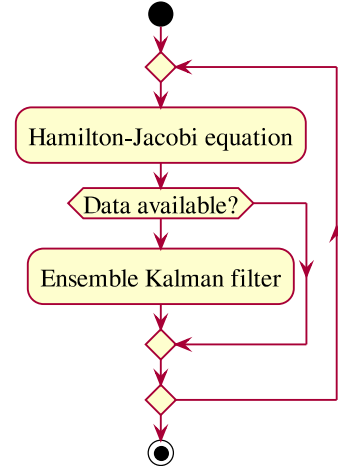
For combined state and parameter estimation, the state vectors are augmented by appending the parameters of interest to the discretized generating functions [45]:

$$\tilde{\psi} = \begin{pmatrix} \psi \\ \theta \end{pmatrix}, \quad (12)$$

where  $\tilde{\psi}$  denotes the augmented state vector and  $\theta$  the parameters. Combined state and parameter estimation works by applying the ensemble Kalman filter (Eqns. (10) and (11)) to  $\tilde{\psi}^f$  instead of  $\psi^f$ .

## RESULTS

In the following, axisymmetric ducted premixed flames are considered. In axial and radial coordinates, a uniform  $400 \times 400$  Cartesian grid is used. Example solutions of the  $G$ -equation for the ducted premixed flame are shown in Fig. 2. The flame is attached to the burner lip, while the perturbations are convected from the base of the flame to the tip. If the perturbations are sufficiently large, a fuel-air pocket pinches off. In our reduced-order model, the perturbations are mainly governed by two non-dimensional parameters [29]: (i) the parameter  $K$ , which governs the perturbation phase speed, and (ii)  $e_a$ , which governs the amplitude of the response of the flame surface to acoustic excitation. Both parameters enter the  $G$ -equation via the underlying velocity



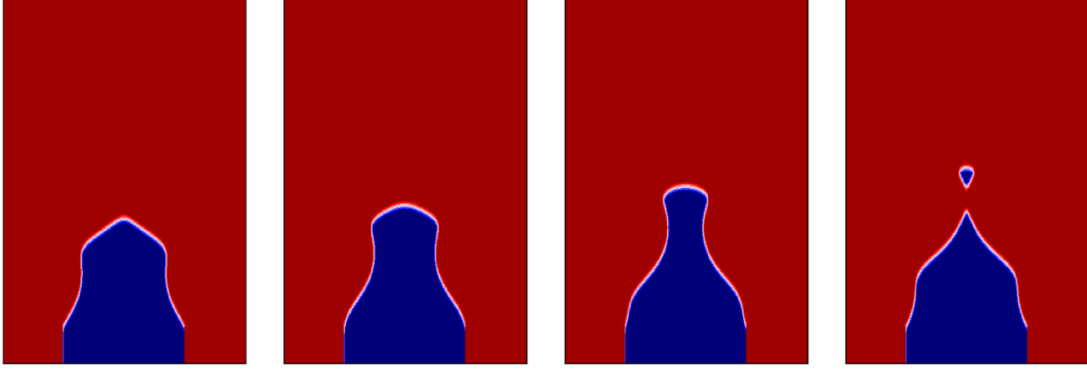
**FIGURE 1:** Activity diagram for level-set data assimilation framework. At every timestep, the flame surface is advanced using our Hamilton-Jacobi equation. When experimental observations become available, data is assimilated by the ensemble Kalman filter.

field  $\mathbf{u}$ . Neither parameter is accurately known a priori, which is a major source of uncertainty.

In the following subsections, two types of problems are addressed: The forward problem quantifies the propagation from the input (the uncertainty in the parameters) to the output (the uncertainty in the location of the flame surface). In the inverse problem, data assimilation is used to identify the location of the flame surface, and reduce the uncertainty in the location and the parameters. The distinct notions of forward and inverse problems are well established in the statistical learning community [46]. In the first subsection, the forward problem is solved by treating the Hamilton-Jacobi equation as a stochastic process. In the second subsection, the inverse problem is solved by using synthetic data from the same  $G$ -equation solver. In the third subsection, the inverse problem is solved by using data from a direct numerical simulation of the premixed flame, which provides the reference solution. Parameter estimation for the inverse problems is performed over both  $K$  and  $e_a$ .

### Stochastic flame dynamics

A stochastic process describes how the probability distribution over a state evolves in time. In the ensemble Kalman filter, the probability distributions over the discretized generating functions are evolved in time using a Markov chain Monte Carlo method, which makes the algorithm well suited for parallelization. The marginal probability distribution of the  $k$ -th entry in



**FIGURE 2:** Solutions of  $G$ -equation over one period of harmonic forcing of velocity at flame base (Tab. 1, middle column). The fuel-air mixture leaves the burner at the bottom of each frame. The infinitely thin flame surface separates the burnt (red) from the unburnt (blue) gas.

**TABLE 1:** Parameters for  $G$ -equation simulations of twin experiment and assimilation of data from high-fidelity simulation.

Parameter	Twin experiment	DNS assimilation
$U$ [m/s]	1.0	1.41
$s_L$ [m/s]	0.164	0.329
$\mathcal{L}$ [m]	0.02	0.02
$e_a$ [m/s]	0.1	0.5
$K$ [-]	1.0	0.214

the state vector reads

$$\psi[k] \sim \mathcal{N}(\bar{\psi}[k], \mathbf{C}_{\psi\psi}[k, k]). \quad (13)$$

The mean  $\bar{\psi}[k]$  and the variance  $\mathbf{C}_{\psi\psi}[k, k]$  are computed from Eqn. (11). Hence the likelihood for the flame surface to be found at the location  $k$  is explicitly given by

$$p[k] = \frac{1}{\sqrt{2\pi\mathbf{C}_{\psi\psi}[k, k]}} \exp\left(-\frac{\bar{\psi}[k]^2}{2\mathbf{C}_{\psi\psi}[k, k]}\right). \quad (14)$$

The normalization condition at every location  $k$  is given by  $\int_{-\infty}^{\infty} p[k] d\bar{\psi}[k] = 1$ . To compare likelihoods at different locations, it is instructive to normalize by the maximum of the probability  $p_0$  at each location  $k$ . Similar to Gaussian processes, it is instructive to look at logarithms for visualization [47]:

$$\log\left(\frac{p[k]}{p_0[k]}\right) = -\frac{\bar{\psi}[k]^2}{2\mathbf{C}_{\psi\psi}[k, k]}. \quad (15)$$

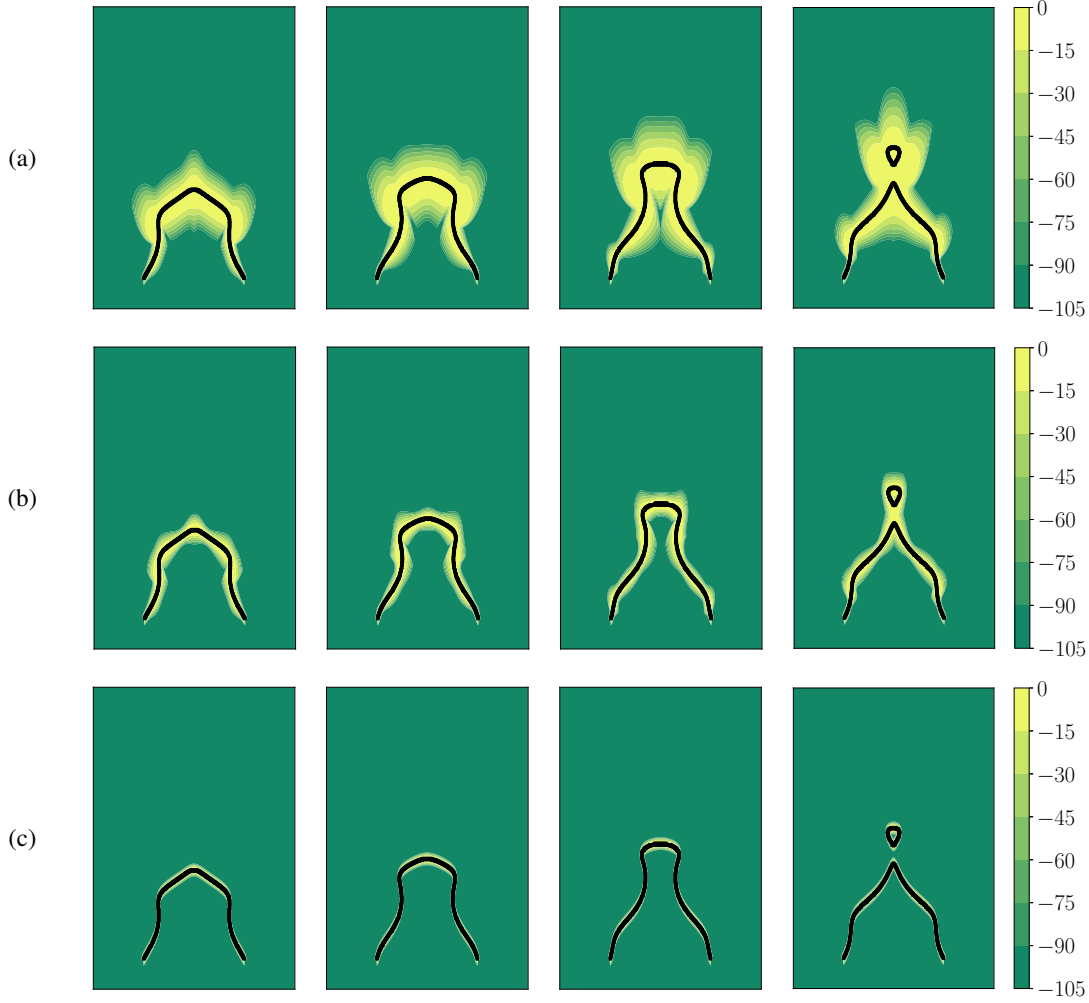
**TABLE 2:** Base-flow (top) and forcing (bottom) parameters for high-fidelity reacting-flow simulation at burner and coflow inlet.

Parameter	Burner inlet	Coflow inlet
Diameter [mm]	12	25
Mass-flow rate [g/s]	0.199	0.784
Temperature [K]	300	300
Mass fraction methane [-]	0.0237	-
Mass fraction ethylene [-]	0.0673	-
Mass fraction air [-]	0.909	1.0
Forcing amplitude [m/s]	0.141	-
Forcing frequency [Hz]	200	-

In Fig. 3(a), the logarithm of the normalized likelihood is shown. The zero-level set gives the maximum-likelihood location of the flame surface. The more negative the value at a location is, the less likely the flame surface is to be found there. As the perturbation is convected from the base of the flame to the tip, the high-likelihood region for the location of the flame surface grows. The high-likelihood region is largest when fuel-air pockets pinch off, which represents maximal uncertainty. Physically speaking, this means that the accurate detection of pinch-off events is the most challenging task.

### Twin experiment

In a twin experiment, both the model predictions and the experimental observations are generated by solving the  $G$ -equation. This allows us to validate our data assimilation framework. More



**FIGURE 3:** Snapshots of the logarithm of the normalized likelihood (Eqn. (15)) over one period for (a) the forward problem and the inverse problems with either (b) state estimation or (c) combined state and parameter estimation respectively. In the twin experiment, the experimental observations are extracted from a  $G$ -equation simulation (black line). High-likelihood (yellow) and low-likelihood (green) regions are shown. Each snapshot from left to right is distanced by a quarter of the forcing period.

specifically, we expect the maximum-likelihood solution to collapse onto the flame surface when the parameters are optimally calibrated. Moreover, a twin experiment can give us physical insight into where parameter uncertainties most severely affect the model predictions. For the model predictions, an ensemble of twenty  $G$ -equations is solved. This number is a compromise between stochastic convergence and available computational resources. The same initial condition is used for the whole ensemble, but a different set of parameters  $K$  and  $e_a$  is chosen for each simulation. They are sampled from two independent normal distributions with 20% standard deviation respectively. For the

generation of synthetic experimental observations, a separate set of parameters  $K$  and  $e_a$  is chosen. The measurements are taken at all grid points where the absolute value of the discretized generating function is less than one grid size. This gives a cloud of grid points inside a narrow band around the zero-level set of the generating function. Corresponding to the zero-level set, the vector of measurements  $\mathbf{d}$  is zero-valued. The measurement matrix  $\mathbf{M}$  is a restriction operator, which maps the state space to the observation space and indicates the locations of the grid points inside the narrow band. The covariance matrix  $\mathbf{C}_{\varepsilon\varepsilon}$ , which represents the experimental errors, is a diagonal matrix with  $\sigma_\varepsilon^2$  on

its diagonal. The simulation parameters are given in Tab. 1.

In a first twin experiment, the ensemble Kalman filter is used to perform state estimation. In Fig. 3(b), the logarithm of the normalized likelihood is shown. A qualitative comparison to Fig. 3(a) shows that the high-likelihood region for the location of the flame surface is significantly reduced. A quantitative measure for the uncertainty in the location of the flame surface is given by the root mean square (RMS) error, which is defined as the square-root of the trace of the covariance matrix:

$$\text{RMS error} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\psi_i - \bar{\psi})^T (\psi_i - \bar{\psi})}. \quad (16)$$

In Fig. 4, the RMS error is plotted over time. At timestep 0, the error is zero because the same initial condition is used for the whole ensemble. In the forward problem, the RMS error grows until it reaches a high-uncertainty plateau. Spikes in the high-uncertainty plateau coincide with the moments when fuel-air pockets are pinched off. With state estimation, the uncertainty is regularly capped, and the predictions of the location of the flame surface significantly improve.

In a second twin experiment, the ensemble Kalman filter is augmented to perform combined state and parameter estimation, which is explained at the end of the previous section on the ensemble Kalman filter. In Fig. 4, the RMS error is plotted over time. The initial behavior is similar to that of the twin experiment with state estimation. After the first data assimilation at timestep 1,000, the uncertainty remains at a relatively constant level. At this point, the combined state and parameter estimation has optimally calibrated the parameters. Consequently, the subsequent uncertainty does not grow and the parameters do not improve. When the first fuel-air pocket is pinched off, at timestep 4,000, the uncertainty grows rapidly until the combined state and parameter estimation updates the state and the parameters again. The update step coincides with the moment of pinch-off, which is also a moment of high uncertainty. Thus the parameters can be found very accurately, which leads to a low-uncertainty plateau. High uncertainties at later pinch-off events are suppressed without much change in the parameters.

In a parameter study, the experimental error  $\sigma_\varepsilon^2$  is varied to investigate its effect on data assimilation. In Fig. 5, the RMS error is plotted over time. It becomes evident that the experimental error indeed poses an epistemic bound on how low the uncertainty can be reduced by combined state and parameter estimation. Hence, uncertainty in the optimally calibrated parameters remains unless the experimental error vanishes. Furthermore, moments of pinch-off regularly induce significant uncertainty regardless of the experimental error. In each case, combined state and parameter estimation intervenes, and caps the RMS error before it grows out of bounds. The results from data assimilation confirm the extreme sensitivity observed elsewhere

that undermines the a-priori calibration of models in favor of optimal, on-the-fly calibration [11]. In Fig. 6, the residuals of  $K$  and  $e_a$  are plotted over time. While  $e_a$  remains virtually constant after 10,000 timesteps,  $K$  significantly varies in sync with the pinch-off cycles. In terms of an inverse problem, the uncertainty in the state (Fig. 5) has propagated upstream to the parameters and mainly to  $K$ . The results from combined state and parameter estimation confirm the physical significance of  $K$  as a model parameter [29].

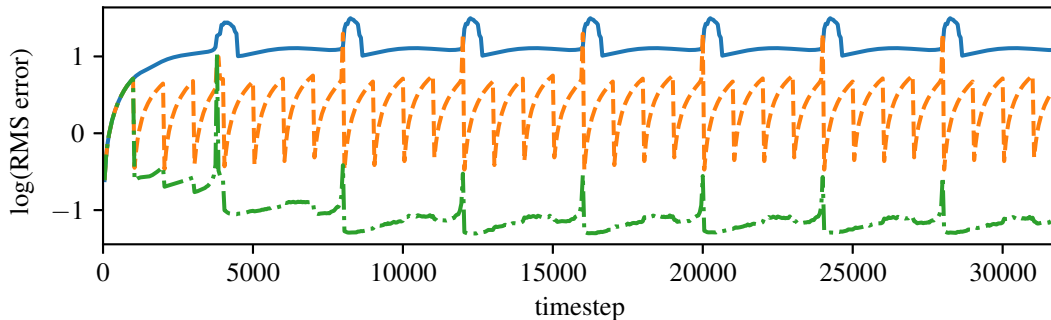
### Assimilation of data from high-fidelity simulation

After the twin experiment, the assimilated data is extracted from a DNS of a ducted premixed flame. The simulation is performed with the finite-volume solver CharLES<sup>X</sup> (Cascade Technologies), solving the compressible reacting Navier-Stokes equations. For convective fluxes, a hybrid scheme is applied, consisting of a fourth-order central scheme and a second-order essentially non-oscillatory scheme [48], which minimizes the numerical dissipation [49].

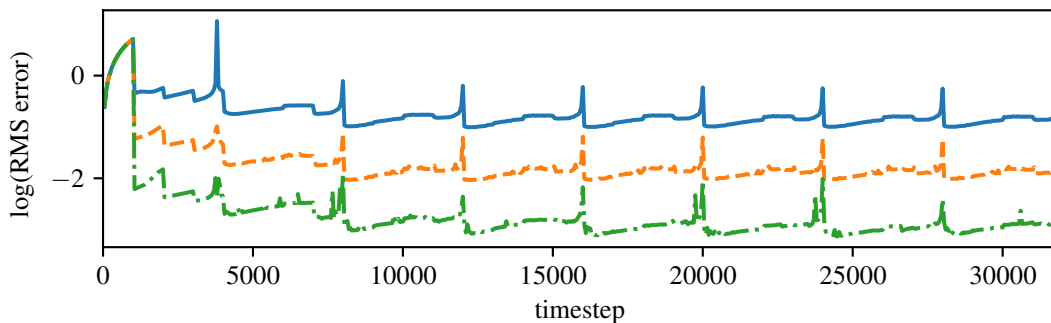
The computational domain is three-dimensional (Fig. 7). Because of the axisymmetry of the configuration, only a quarter (90° angular sector) of the full geometry with two symmetry boundary planes is considered. The domain is 0.12 m long, and extends two diameters upstream into the burner to allow for potential flash back of the flame. It also includes a coflow of fresh air. This models the flow of fresh air that would arise in a vertical duct due to buoyancy effects. The stretched grid comprises 500,000 hexahedral elements with characteristic element sizes ranging from 0.15 mm in the flame region to 3 mm at the outlet.

At the main inlet, a parabolic velocity profile is imposed in space along with temperature and species composition corresponding to the ethylene-methane-air mixture given in Tab. 2. For pulsed cases, a sinusoidal velocity amplitude is imposed in time. The coflow speed is prescribed based on a one-dimensional analysis, inspired by the experimental configuration of a vertical Rijke tube [50]. The choice of the coflow speed does not affect the flame dynamics in a qualitative manner. Lateral and injection walls are modeled with isothermal no-slip boundary conditions with an imposed temperature  $T = 300\text{ K}$  to correctly reproduce the flame stabilization. At the outlet of the domain, a constant pressure boundary is used to impose atmospheric pressure conditions, along with a sponge layer region to suppress reflected acoustic waves.

The reduced chemical mechanism includes 15 species and 5 quasi-steady state species, which correctly predicts laminar flame speed and auto-ignition delay for the target operating conditions [51]. In the solver, the chemical source terms are evaluated using finite-rate kinetics with a semi-implicit Rosenbrock-Krylov scheme [52, 53]. From one-dimensional laminar computations, it was verified a priori that the grid resolution in the flame region (0.15 mm) is sufficient to correctly resolve the ther-



**FIGURE 4:** Root mean square (RMS) error over time for the forward problem (blue line) and the inverse problems with either state estimation (orange dash) or combined state and parameter estimation (green dash-dot). One forcing period corresponds to 4,000 timesteps. Data is available every 1,000 timesteps.



**FIGURE 5:** Root mean square (RMS) error over time for combined state and parameter estimation with different experimental errors ( $\sigma_\varepsilon^2 = 10^{-4}$  blue line,  $\sigma_\varepsilon^2 = 10^{-6}$  orange dash,  $\sigma_\varepsilon^2 = 10^{-8}$  green dash-dot).

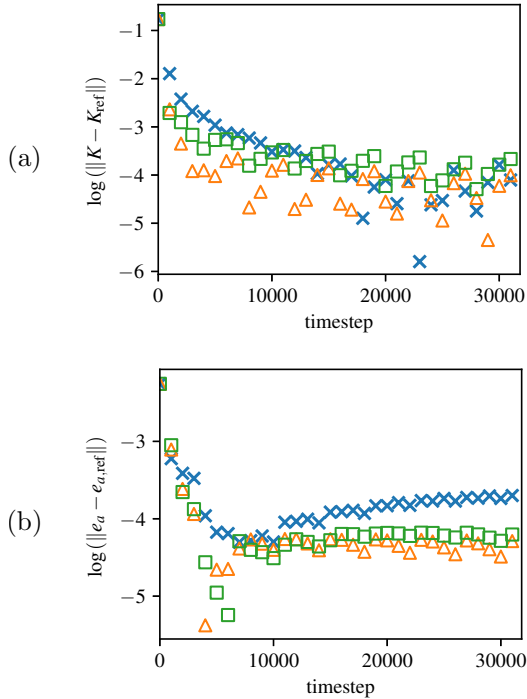
mal thickness of the flame (0.4 mm).

In Fig. 8, snapshots of the DNS are shown. The flame is attached to the burner lip on the edge of flashing back. The DNS displays the same qualitative behavior as the  $G$ -equation as perturbations are observed to travel from the base to the tip of the flame. For the given simulation set-up, no pinched-off fuel-air pockets are observed in the DNS.

For the model predictions, an ensemble of twenty-four  $G$ -equations is solved. The same initial condition is used for the whole ensemble, but a different set of parameters  $K$  and  $e_a$  is chosen for each simulation. They are sampled from two independent normal distributions with 20% standard deviation respectively. The covariance matrix  $C_{\varepsilon\varepsilon}$  is again a diagonal matrix, i.e., the experimental errors are assumed to be statistically independent. The simulation parameters are given in Tab. 1. In Fig. 9, the logarithm of the normalized likelihood is shown for the forward problem and the inverse problems, the latter involving either state estimation or combined state and parameter estimation.

As in Fig. 3, data assimilation significantly improves the ability of the reduced-order  $G$ -equation model to capture the motion of the flame surface. In a qualitative comparison, it can be seen that the high-likelihood region from state estimation (Fig. 9(b)) is a better envelope to the flame surface observed in the DNS than the unfiltered stochastic process (Fig. 9(a)). Combined state and parameter estimation (Fig. 9(c)) provides an even narrower envelope, which is testament to the reduced uncertainty. This assessment is quantitatively confirmed by the RMS error shown in Fig. 10. While state estimation does not appreciably reduce the RMS error, it provides a dynamic, truly a-posteriori estimate of the uncertainty that is rigorously based on Bayesian statistics. On the contrary, the RMS error of the unfiltered stochastic process only gives an a-priori estimate of the uncertainty based on an educated initial guess for the possible ranges of values for the parameters. In Fig. 11, the values of  $K$  and  $e_a$  with their error bars are plotted over time for combined state and parameter estimation. After a burn-in phase of approximately 4,000 timesteps, the





**FIGURE 6:** Residuals of (a)  $K$  and (b)  $e_a$  over time for combined state and parameter estimation with different experimental errors ( $\sigma_\varepsilon^2 = 10^{-4}$  blue cross,  $\sigma_\varepsilon^2 = 10^{-6}$  orange triangle,  $\sigma_\varepsilon^2 = 10^{-8}$  green square).

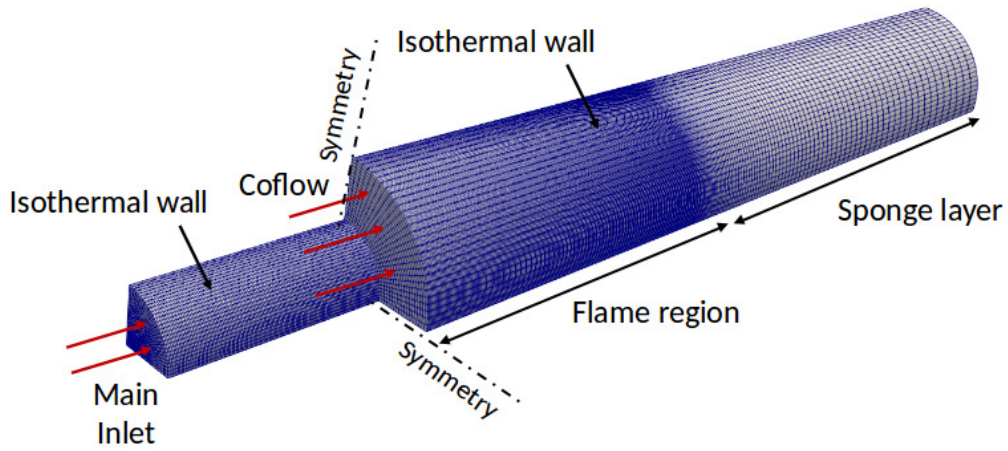
parameters are optimally calibrated to values which are in agreement with the experimental observations. This is reflected in Fig. 10 by the low-uncertainty plateau, which is only interrupted by the moments when perturbations reach the tip of the flame and neutralize. The outlier around timestep 3,500, where combined state and parameter estimation performs worse in terms of RMS error than the unfiltered stochastic process and state estimation, is explained by the formation of a pinched-off fuel-air pocket that is not observed in the DNS.

Finally, we build on our understanding of the state and parameter uncertainties to attempt model uncertainty quantification. Unlike optimization-based approaches, Bayesian statistics do not focus on the minimization of one cost functional but revolve around probability distributions. As such, they can be analyzed using histograms. In Fig. 12, two histograms, which are representative of the twin experiment and the assimilation of DNS data respectively, are shown. In the twin experiment (Fig. 12(a)), most observation points are located in the high-likelihood region. The likelihood of finding the flame surface in the low-likelihood region decays exponentially. When assimilating DNS data (Fig. 12(b)), the histogram has a heavy tail. While the likelihood of finding the flame surface in the high-

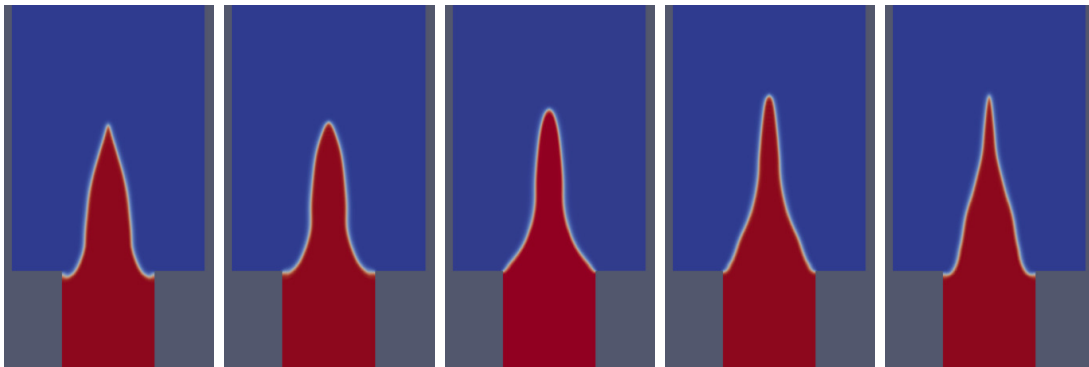
likelihood region decays fast, the likelihood of finding the flame surface in the low-likelihood region does not vanish, reminiscent of a power-law distribution. The deviation from an exponentially decaying distribution, a piece of information detected by our Bayesian approach but absent from an optimization-based approach, signifies the presence of model errors [54]. Such a deviation informs us of the inherent limitations of a reduced-order model, even when optimally calibrated. Model errors can only be decreased by adding complexity, i.e., more physics, to the reduced-order model assumptions. Further analysis reveals that experimental observations located in low-likelihood regions are predominantly found at the base of the flame, in particular when the flame is on the edge of flashing back. It is well known that the dynamics at the base of a premixed flame are crucial to the flame dynamics [55]. Although the motion of the flame surface is accurately captured towards the tip of the flame, level-set methods, which do not obey conservation laws, cannot model heat transfer between the flame and the burner lip or the momentum transfer in the wake region behind the burner lip. This shows how model uncertainty, which is traditionally the type of uncertainty that is the most difficult to quantify, can be detected by Bayesian statistics to inspire more predictive models.

## CONCLUSIONS

We propose an on-the-fly statistical learning method, based on data assimilation with the ensemble Kalman filter, to improve the prediction of the dynamics of premixed flames. In the first part of the paper, the capabilities of the framework are demonstrated in a twin experiment, where the assimilated data is produced from the same model as that used in prediction. This guarantees that the assimilated data is consistent with the predictions, which is particularly useful to validate the algorithm and identify the regions that are most affected by parameter uncertainties. The uncertainties of the calculations are estimated by using Bayesian statistics. From a calibration point of view, it is shown that the major source of uncertainty originates from the model parameter  $K$ , which essentially governs the speed at which perturbations travel on the flame surface. From a dynamical point of view, it is shown that the pinch-off events are extremely sensitive to both parameters and initial conditions. In the second part of the paper, data is extracted from a high-fidelity reacting-flow simulation and assimilated into the reduced-order flame model. State estimation as well as combined state and parameter estimation provide significant improvements in the model predictions. While the motion of the flame surface is well captured towards the tip of the flame, the attachment of the flame to the burner lip is not. Bayesian statistics reveal the reason to be neither state nor parameter uncertainties but deficiencies in the model. While the traveling of perturbations along the flame surface is accurately modeled, the reduced-order  $G$ -equation model does not include the physical mechanisms necessary to model the flame



**FIGURE 7:** Computational domain, computational grid and boundary conditions of direct numerical simulation (DNS).



**FIGURE 8:** Ethylene mass fractions in direct numerical simulation (DNS) over one period. The fuel-air mixture leaves the burner at the bottom. The flame surface separates the burnt (blue) from the unburnt (red) gas.

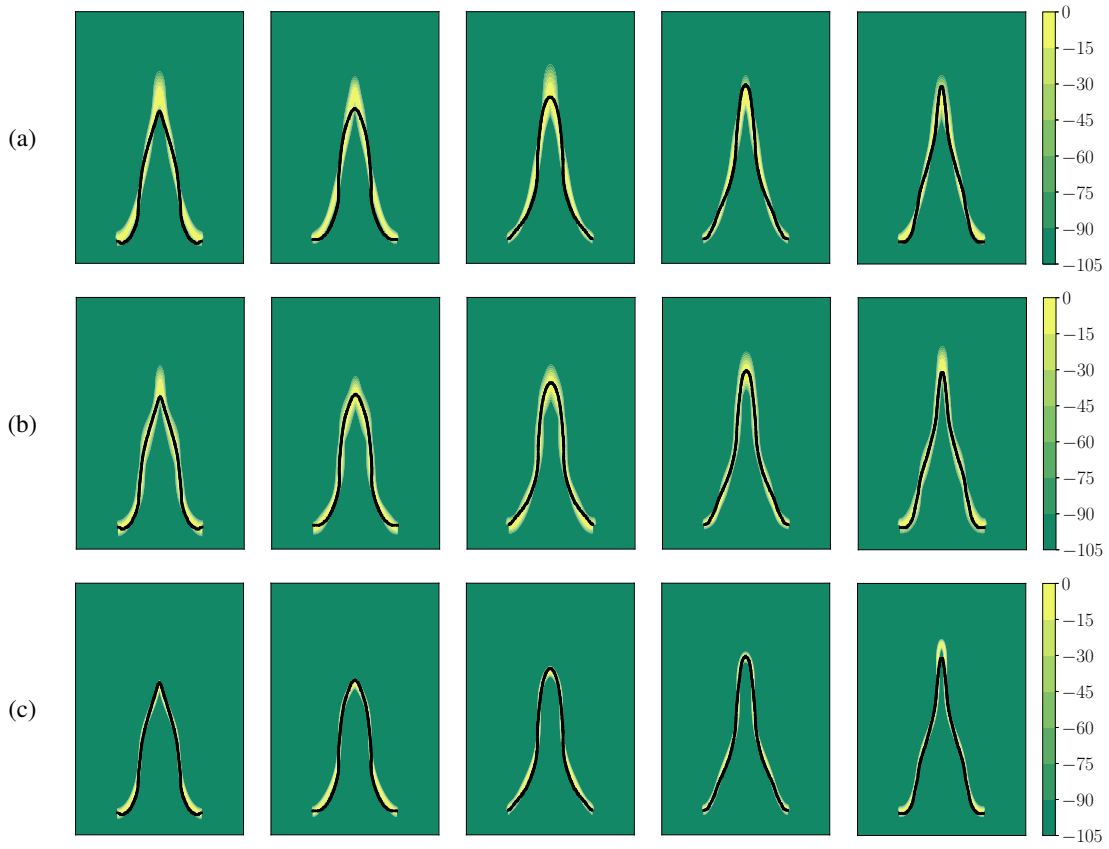
attachment, i.e. heat loss to the burner and the shear layer in the wake of the burner. This physical insight into the significance of the pinch-off events and the flame attachment will inform both the modeling assumptions as well as the design of experiments towards truly accurate prediction, for example in turbulent flows [23, 56].

In future work, we are going to apply our level-set data assimilation framework to experiments of the ducted premixed flame in order to optimally calibrate the model parameters and quantify model uncertainties. The assimilated data will include images of the premixed flame taken with a high-speed camera and pressure signals recorded at different locations inside the tube. This will further demonstrate the versatility of the ensemble Kalman filter when applied to reduced-order models as it is non-intrusive and flexible in handling data from a variety of ex-

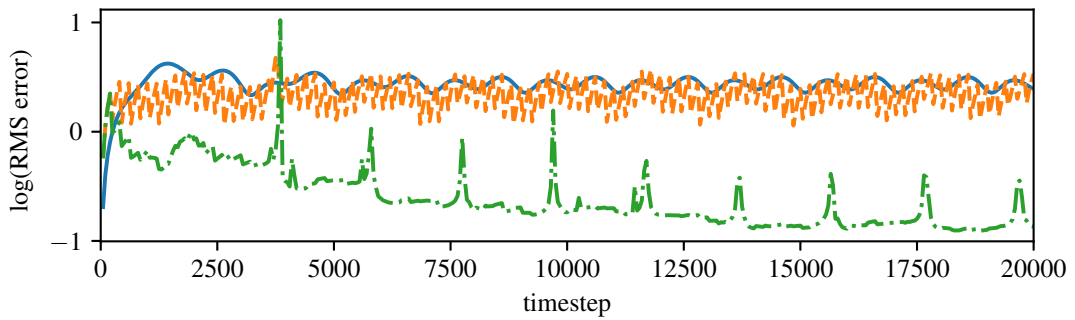
periments and sensors. The statistical methods proposed in this paper open up new possibilities for the optimal calibration of reduced-order models in real time when experimental data becomes available, e.g. from the sensors of a gas turbine. The proposed data assimilation framework was developed for a generic level-set method, therefore, it can be applied to other interface-tracking problems such as multiphase flows.

#### ACKNOWLEDGMENT

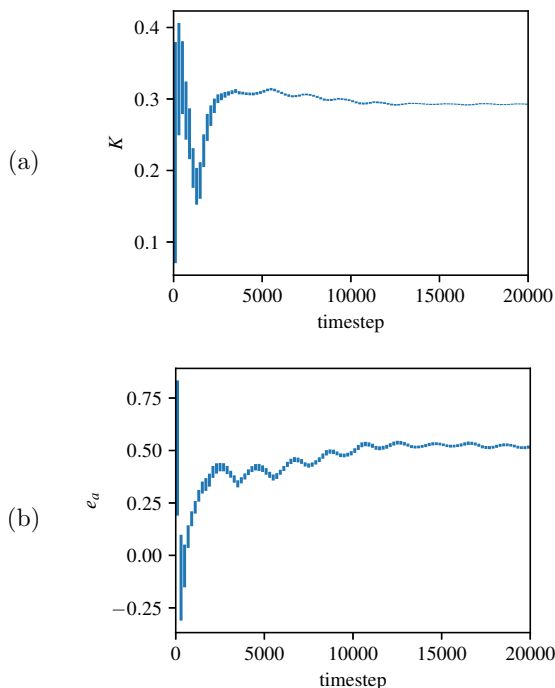
The authors are grateful to the 2018 Center for Turbulence Research Summer Program for having partly sponsored this project. L.M. gratefully acknowledges support of (i) the Royal Academy of Engineering Research Fellowship Scheme and (ii) the visiting fellowship of the Technical University of Munich



**FIGURE 9:** Snapshots of the logarithm of the normalized likelihood (Eqn. (15)) over one period for (a) the forward problem and the inverse problems with either (b) state estimation or (c) combined state and parameter estimation respectively. For the data assimilation, the experimental observations are extracted from a DNS (black line). High-likelihood (yellow) and low-likelihood (green) regions are shown. Each snapshot from left to right is distanced by one fifth of the forcing period.



**FIGURE 10:** Root mean square (RMS) error over time for the forward problem (blue line) and the inverse problems with either state estimation (orange dash) or combined state and parameter estimation (green dash-dot). One forcing period corresponds to 2,000 timesteps. Data is available every 200 timesteps.

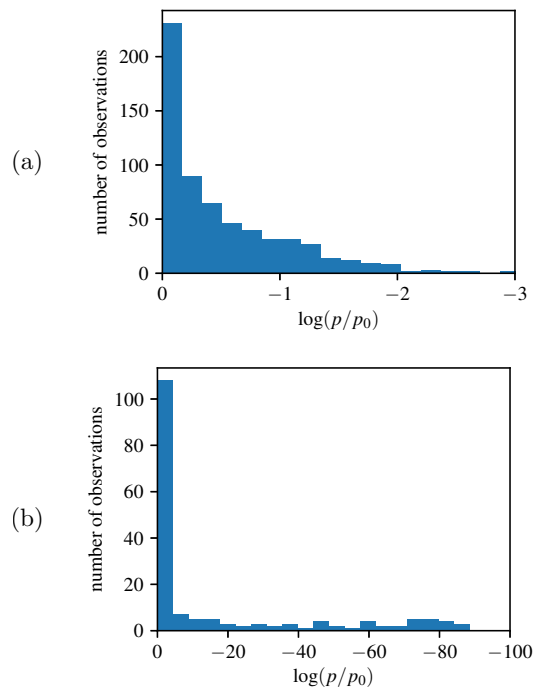


**FIGURE 11:** Three-sigma error bars for (a)  $K$  and (b)  $e_a$  for combined state and parameter estimation. The convergence of the parameters shows that the reduced-order flame model is optimally calibrated.

– Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763. Discussions on data assimilation with Dr. J. W. Labahn are gratefully acknowledged.

## REFERENCES

- [1] Strutt, J. W., 2011. *The Theory of Sound*, Vol. 2. Cambridge University Press.
- [2] Lieuwen, T. C., and Yang, V., 2005. *Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms, and Modeling*. American Institute of Aeronautics and Astronautics, Inc.
- [3] Culick, F. E. C., 2006. *Unsteady motions in combustion chambers for propulsion systems*. RTO AGARDograph AG-AVT-039, North Atlantic Treaty Organization.
- [4] Kabiraj, L., Saurabh, A., Wahi, P., and Sujith, R. I., 2012. “Route to chaos for combustion instability in ducted laminar premixed flames”. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **22**(2), June, p. 023129.
- [5] Kashinath, K., Waugh, I. C., and Juniper, M. P., 2014. “Nonlinear self-excited thermoacoustic oscillations of a



**FIGURE 12:** Histograms of experimental observations over logarithm of normalized likelihood for (a) twin experiment and (b) assimilation of DNS data right. High-likelihood bins are located on the left of each histogram, low-likelihood bins on the right.

ducted premixed flame: bifurcations and routes to chaos”. *Journal of Fluid Mechanics*, **761**, Dec., pp. 399–430.

- [6] Gotoda, H., Nikimoto, H., Miyano, T., and Tachibana, S., 2011. “Dynamic properties of combustion instability in a lean premixed gas-turbine combustor”. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **21**(1), Mar.
- [7] Waugh, I., Geuß, M., and Juniper, M., 2011. “Triggering, bypass transition and the effect of noise on a linearly stable thermoacoustic system”. *Proceedings of the Combustion Institute*, **33**(2), Jan., pp. 2945–2952.
- [8] Kiureghian, A. D., and Ditlevsen, O., 2009. “Aleatory or epistemic? Does it matter?”. *Structural Safety*, **31**(2), Mar., pp. 105–112.
- [9] Magri, L., Bauerheim, M., and Juniper, M. P., 2016. “Stability analysis of thermo-acoustic nonlinear eigenproblems in annular combustors. Part I. Sensitivity”. *Journal of Computational Physics*, **325**, Nov., pp. 395–410.
- [10] Magri, L., Bauerheim, M., Nicoud, F., and Juniper, M. P., 2016. “Stability analysis of thermo-acoustic nonlinear eigenproblems in annular combustors. Part II. Uncertainty quantification”. *Journal of Computational Physics*, **325**, Nov., pp. 411–421.
- [11] Juniper, M. P., and Sujith, R., 2018. “Sensitivity and Non-

- linearity of Thermoacoustic Oscillations”. *Annual Review of Fluid Mechanics*, **50**(1), Jan., pp. 661–689.
- [12] Poinso, T., 2017. “Prediction and control of combustion instabilities in real engines”. *Proceedings of the Combustion Institute*, **36**(1), pp. 1–28.
- [13] Bennett, A. F., 1992. *Inverse Methods in Physical Oceanography*. Cambridge University Press, Cambridge.
- [14] Bennett, A. F., 2002. *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, Cambridge.
- [15] Evensen, G., 2009. *Data Assimilation*. Springer Berlin Heidelberg.
- [16] Colburn, C. H., Cessna, J. B., and Bewley, T. R., 2011. “State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter”. *Journal of Fluid Mechanics*, **682**, Sept., pp. 289–303.
- [17] Gronskis, A., Heitz, D., and Mémin, E., 2013. “Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation”. *Journal of Computational Physics*, **242**, June, pp. 480–497.
- [18] Foures, D., Dovetta, N., Sipp, D., and Schmid, P., 2014. “A data-assimilation method for Reynolds-averaged Navier-Stokes-driven mean flow reconstruction”. *Journal of Fluid Mechanics*, **759**, Nov., pp. 404–431.
- [19] Kato, H., Yoshizawa, A., Ueno, G., and Obayashi, S., 2015. “A data assimilation methodology for reconstructing turbulent flows around aircraft”. *Journal of Computational Physics*, **283**, Feb., pp. 559–581.
- [20] Mons, V., Chassaing, J.-C., Gomez, T., and Sagaut, P., 2016. “Reconstruction of unsteady viscous flows using data assimilation schemes”. *Journal of Computational Physics*, **316**, July, pp. 255–280.
- [21] Mons, V., Chassaing, J.-C., and Sagaut, P., 2017. “Optimal sensor placement for variational data assimilation of unsteady flows past a rotationally oscillating cylinder”. *Journal of Fluid Mechanics*, **823**, July, pp. 230–277.
- [22] Darakananda, D., Eldredge, J., da Silva, A., Colonius, T., and Williams, D. R., 2018. “EnKF-based Dynamic Estimation of Separated Flows with a Low-Order Vortex Model”. American Institute of Aeronautics and Astronautics.
- [23] Labahn, J. W., Wu, H., Coriton, B., Frank, J. H., and Ihme, M., 2018. “Data assimilation using high-speed measurements and LES to examine local extinction events in turbulent flames”. *Proceedings of the Combustion Institute*, June.
- [24] Jazwinski, A. H., 2007. *Stochastic Processes and Filtering Theory*. Dover Publications.
- [25] Waugh, I. C., Kashinath, K., and Juniper, M. P., 2014. “Matrix-free continuation of limit cycles and their bifurcations for a ducted premixed flame”. *Journal of Fluid Mechanics*, **759**, Nov., pp. 1–27.
- [26] Dowling, A. P., 1999. “A kinematic model of a ducted flame”. *Journal of Fluid Mechanics*, **394**, Sept., pp. 51–72.
- [27] Peters, N., 2000. *Turbulent Combustion*. Cambridge University Press.
- [28] Yu, H., Jaravel, T., Ihme, M., Juniper, M. P., and Magri, L., 2018. Physics-informed data-driven prediction of premixed flame dynamics with data assimilation. Tech. rep., Center for Turbulence Research.
- [29] Kashinath, K., Hemchandra, S., and Juniper, M. P., 2013. “Nonlinear thermoacoustics of ducted premixed flames: The influence of perturbation convection speed”. *Combustion and Flame*, **160**(12), Dec., pp. 2856–2865.
- [30] Landau, L., and Lifshitz, E., 1969. *Mechanics*, second ed., Vol. 1 of *A Course of Theoretical Physics*. Pergamon Press.
- [31] Osher, S., and Sethian, J. A., 1988. “Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations”. *Journal of Computational Physics*, **79**(1), Nov., pp. 12–49.
- [32] Adalsteinsson, D., and Sethian, J. A., 1995. “A Fast Level Set Method for Propagating Interfaces”. *Journal of Computational Physics*, **118**(2), May, pp. 269–277.
- [33] Peng, D., Merriman, B., Osher, S., Zhao, H., and Kang, M., 1999. “A PDE-Based Fast Local Level Set Method”. *Journal of Computational Physics*, **155**(2), Nov., pp. 410–438.
- [34] Liu, X.-D., Osher, S., and Chan, T., 1994. “Weighted Essentially Non-oscillatory Schemes”. *Journal of Computational Physics*, **115**(1), Nov., pp. 200–212.
- [35] Jiang, G.-S., and Shu, C.-W., 1996. “Efficient Implementation of Weighted ENO Schemes”. *Journal of Computational Physics*, **126**(1), June, pp. 202–228.
- [36] Shu, C.-W., and Osher, S., 1988. “Efficient implementation of essentially non-oscillatory shock-capturing schemes”. *Journal of Computational Physics*, **77**(2), Aug., pp. 439–471.
- [37] Waugh, I., 2013. “Methods for Analysis of Nonlinear Thermoacoustic Systems”. PhD thesis, University of Cambridge, Mar.
- [38] Preetham, Santosh, H., and Lieuwen, T., 2008. “Dynamics of Laminar Premixed Flames Forced by Harmonic Velocity Disturbances”. *Journal of Propulsion and Power*, **24**(6), nov, pp. 1390–1402.
- [39] Orchini, A., Rigas, G., and Juniper, M. P., 2016. “Weakly nonlinear analysis of thermoacoustic bifurcations in the Rijke tube”. *Journal of Fluid Mechanics*, **805**, Oct., pp. 523–550.
- [40] Semlitsch, B., Orchini, A., Dowling, A. P., and Juniper, M. P., 2017. “G-equation modelling of thermoacoustic oscillations of partially premixed flames”. *International Journal of Spray and Combustion Dynamics*, **9**(4), Dec., pp. 260–276.
- [41] Sussman, M., Smereka, P., and Osher, S., 1994. “A Level Set Approach for Computing Solutions to Incompressible

- Two-Phase Flow”. *Journal of Computational Physics*, **114**(1), Sept., pp. 146–159.
- [42] Sethian, J. A., 1996. “A fast marching level set method for monotonically advancing fronts”. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(4), Feb., pp. 1591–1595.
- [43] Kalman, R. E., 1960. “A New Approach to Linear Filtering and Prediction Problems”. *Journal of Basic Engineering*, **82**(1), p. 35.
- [44] Kalman, R. E., and Bucy, R. S., 1961. “New Results in Linear Filtering and Prediction Theory”. *Journal of Basic Engineering*, **83**(1), p. 95.
- [45] Sarkka, S., 2013. *Bayesian Filtering and Smoothing*. Cambridge University Press, Cambridge.
- [46] Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics.
- [47] Rasmussen, C. E., and Williams, C. K. I., 2006. *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press. OCLC: ocm61285753.
- [48] Ma, P. C., Lv, Y., and Ihme, M., 2017. “An entropy-stable hybrid scheme for simulations of transcritical real-fluid flows”. *Journal of Computational Physics*, **340**, July, pp. 330–357.
- [49] Khalighi, Y., Ham, F., Nichols, J., Lele, S., and Moin, P., 2011. “Unstructured Large Eddy Simulation for Prediction of Noise Issued from Turbulent Jets in Various Configurations”. In 17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference), American Institute of Aeronautics and Astronautics.
- [50] Rigas, G., Jamieson, N. P., Li, L. K. B., and Juniper, M. P., 2016. “Experimental sensitivity analysis and control of thermoacoustic systems”. *Journal of Fluid Mechanics*, **787**, Jan.
- [51] Narayanaswamy, K., Blanquart, G., and Pitsch, H., 2010. “A consistent chemical mechanism for oxidation of substituted aromatic species”. *Combustion and Flame*, **157**(10), Oct., pp. 1879–1898.
- [52] Tranquilli, P., and Sandu, A., 2014. “Rosenbrock–Krylov Methods for Large Systems of Differential Equations”. *SIAM Journal on Scientific Computing*, **36**(3), Jan., pp. A1313–A1338.
- [53] Wu, H., Ma, P. C., and Ihme, M., 2017. “Efficient time stepping for reactive turbulent simulations with stiff chemistry”. *arXiv:1712.00953 [physics]*, Dec.
- [54] Hamill, T. M., 2001. “Interpretation of Rank Histograms for Verifying Ensemble Forecasts”. *Monthly Weather Review*, **129**(3), Mar., pp. 550–560.
- [55] Cuquel, A., Durox, D., and Schuller, T., 2013. “Impact of flame base dynamics on the non-linear frequency response of conical flames”. *Comptes Rendus Mecanique*, **341**(1-2), Jan., pp. 171–180.
- [56] Nastac, G., Labahn, J. W., Magri, L., and Ihme, M., 2017. “Lyapunov exponent as a metric for assessing the dynamic content and predictability of large-eddy simulations”. *Physical Review Fluids*, **2**(9), Sept.