

# A data-driven kinematic model of a ducted premixed flame

Hans Yu<sup>a</sup>, Matthew P. Juniper<sup>a</sup>, Luca Magri<sup>a,b,\*</sup>

<sup>a</sup>*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK*

<sup>b</sup>*Institute for Advanced Study, Technical University of Munich, Lichtenbergstrasse 2a, 85748 Garching, Germany (visiting fellowship)*

---

## Abstract

Reduced-order models of flame dynamics can be used to predict and mitigate the emergence of thermoacoustic oscillations in the design of gas turbine and rocket engines. This process is hindered by the fact that these models, although often qualitatively correct, are not usually quantitatively accurate. As automated experiments and numerical simulations produce ever-increasing quantities of data, the question arises as to how this data can be assimilated into physics-informed reduced-order models in order to render these models quantitatively accurate. In this study, we develop and test a physics-based reduced-order model of a ducted premixed flame in which the model parameters are learned from high speed videos of the flame. The experimental data is assimilated into a level-set solver using an ensemble Kalman filter. This leads to an optimally calibrated reduced-order model with quantified uncertainties, which accurately reproduces elaborate nonlinear features such as cusp formation and pinch-off. The reduced-order model continues to match the experiments after assimilation has been switched off. Further, the parameters of the model, which are extracted automatically, are shown to match the first order behavior expected on physical grounds. This study shows how reduced-order models can be updated rapidly whenever new experimental or numerical data becomes available, without the data itself having to be stored.

*Keywords:* premixed combustion, thermoacoustics, reduced-order modeling, data assimilation, uncertainty quantification

---

---

\*Corresponding author  
Email address: [lm547@cam.ac.uk](mailto:lm547@cam.ac.uk) (Luca Magri)

## 1. Introduction

Thermoacoustic oscillations are a persistent challenge in the design of jet and rocket engines [1, 2]. On the one hand the *qualitative* mechanism is well-understood: acoustic perturbations cause heat release rate perturbations at the flame through a variety of mechanisms [1]. These heat release rate perturbations, if sufficiently in phase with the acoustic pressure, amplify the acoustic perturbations, closing the feedback loop [3]. On the other hand, the *quantitative* prediction of thermoacoustic dynamics in real engines is difficult and computationally demanding, despite advances in large-eddy simulations (LES) of reacting flows [4]. The quantitative prediction is difficult because thermoacoustic systems are extremely sensitive to small changes to the system [5, 6].

Recent advances in data-driven methods and machine learning introduce new approaches to the development of predictive methods [7, 8]. Most data-driven approaches identify systems and extract models using projections or kernels to map from high-fidelity data to reduced-order models [9]. Examples include proper orthogonal decomposition, dynamic mode decomposition, active subspaces, resolvent analysis and variational autoencoders. In this study, a different approach is developed and tested; this is a data-driven, physics-informed reduced-order model. We begin with the  $G$ -equation, which is a reduced-order model of the flame surface [10]. The surface propagates into the fresh gas, whose velocity is prescribed. The propagation speed is determined by fuel composition and local curvature. Unlike projection- and kernel-based methods, which are fully data-driven, we choose the physically meaningful quantities *a priori*, and define them as parameters and variables in the proposed reduced-order model. Subsequently, this qualitative model is made quantitatively predictive by augmenting it with data. Data assimilation based on the ensemble Kalman filter is performed to achieve a statistically optimal combination of theoretical, computational and experimental knowledge [e.g., 11, 12]. The ensemble Kalman filter is suitable for the treatment of nonlinear dynamics because it is a stochastic technique. By using a Bayesian framework, the objective of this work is to propose an adaptive reduced-order model that predicts the nonlinear flame dynamics with its uncertainties. We collect experimental data, and develop a reduced-order model of the base flow and the flame response (Sections 2, 3). Using the ensemble Kalman filter, we assimilate the experimental data into the reduced-order model for optimal calibration, and quantify the uncertainties in the flame response and the model parameters (Section 4). In Section 5, we summarize the conclusions of this study.

## 2. Experiment

A schematic view of the experimental set-up is given in Fig. 1. The core of the experiment consists of a Bunsen burner inside a tube. The Bunsen burner consists of a straight metal pipe with an inner diameter of 10 mm. A tube with an integrated glass window for optical access acts as a cylindrical enclosure to shield the flame from gusts. Experiments are carried out with a pre-

mixed methane-ethene-air mixture with laminar flow rates set to 0.9, 0.25 and 8 normal liters per minute. The composition of the mixture is controlled with BRONKHORST EL-FLOW mass flow controllers (MFC). A loudspeaker, driven by an amplified sinusoidal signal, is mounted upstream of the Bunsen burner for acoustic forcing. The flame dynamics are recorded with a PHANTOM V4.2 CMOS camera with a glass lens at a resolution of  $1280 \times 800$  pixels and a frame rate of 2800 frames/s. The resolution is sufficiently fine to resolve the flame surface (Fig. 2a).

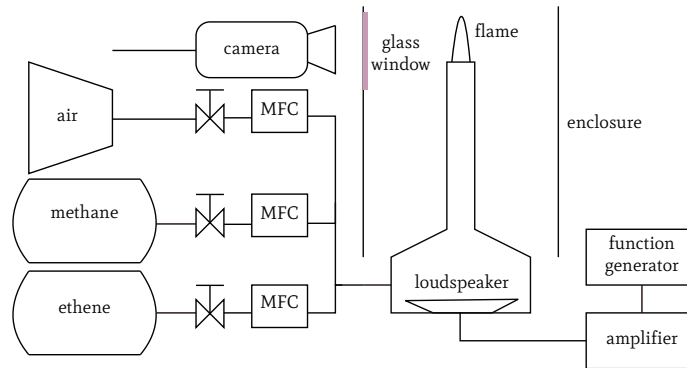


Figure 1: Experimental set-up.

In order to perform data assimilation (Section 4), we extract the coordinates of the pixels associated with the flame surface, which is identified by its natural luminescence in the visible range in the experimental images (Fig. 2a). A number of edge and ridge detection algorithms are readily available from SCIKIT-IMAGE [13], all of which give comparable results. In Fig. 2b, the result of applying the Sobel filter, which is used in this paper, is shown.

### 3. Reduced-order model

The evolution of a premixed flame is modeled here by the kinematics of a surface. The position of the flame surface is captured by the zero-level set of a continuous scalar variable  $G$ . It is governed by the  $G$ -equation [14]:

$$\frac{\partial G}{\partial t} + (\mathbf{u} - s_L \mathbf{n}) \cdot \nabla G = 0 \quad , \quad (1)$$

where  $\mathbf{u}$  is the underlying flow field,  $s_L$  is the laminar flame speed, and  $\mathbf{n}$  is the unit normal vector. The underlying flow field  $\mathbf{u}$  is the superposition of a base flow  $\mathbf{U}$  (Section 3.1) and a velocity perturbation  $\mathbf{u}'$  (Section 3.2). The laminar flame speed is

$$s_L = s_L^0 (1 - \kappa \mathcal{L}) \quad , \quad (2)$$

where  $s_L^0$  denotes the adiabatic flame speed. The Markstein length  $\mathcal{L}$  makes the flame speed a function of the local curvature  $\kappa$ . Normal vector  $\mathbf{n}$  and curvature  $\kappa$

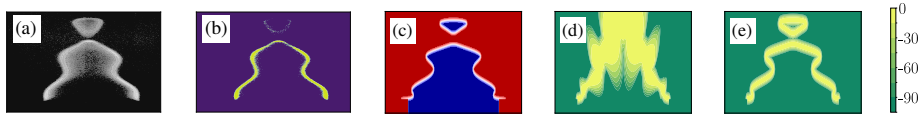


Figure 2: Experimental, computational and statistical visualisations of premixed flame at reduced aspect ratios. (a): Experimental image of acoustic forcing at 200 Hz. Perturbations of the flame surface form at the base and travel to the tip. If the amplitude of the perturbation is large enough, a fuel-air pocket pinches off. (b): Experimental image after postprocessing. Pixels associated with the flame surface are colored yellow. (c): Snapshot of  $G$ -equation simulation for  $K = 0.5$  and  $\varepsilon = 0.36$ . The zero-level set (white) separates the burnt (red) from the unburnt (blue) region. (d): Snapshot of log-normalized likelihood (Eq. (21)) for initial guess of  $K = 0.5$  and  $\varepsilon = 0.2$  with 10% standard deviation in each. Highly likely positions of the flame surface are shown in yellow; less likely positions in green. (e): Snapshot of log-normalized likelihood after combined state and parameter estimation. The model parameters  $K$  and  $\varepsilon$  are optimally calibrated such that they reproduce the propagation of perturbations along the flame surface as well as the formation of pinched-off fuel-air pockets. The spread of the high-likelihood locations (yellow) resembles the precision of the edge detection (b). The computational and statistical results of panels (c,d,e) are thoroughly explained in Sections 3 and 4.

are given in terms of  $G$ :

$$\mathbf{n} = \frac{\nabla G}{\|\nabla G\|} \quad , \quad \kappa = -\nabla \cdot \mathbf{n} \quad . \quad (3)$$

Fig. 3 shows a schematic of this reduced-order model.

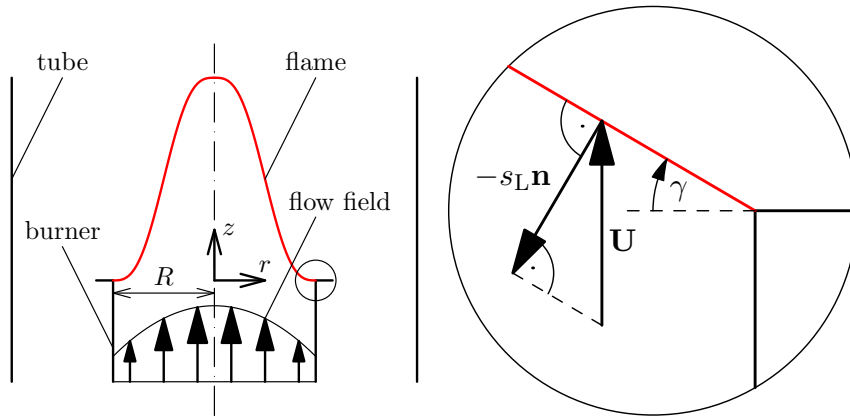


Figure 3: Reduced-order model of the ducted premixed flame (left). In the absence of velocity perturbations, there is a kinematic balance between the base flow and the flame surface, e.g. at the burner lip (right).

The  $G$ -equation is numerically solved using a narrow-band level-set method with distance reinitialization. The computational domain is discretized using a fifth-order, weighted essentially non-oscillatory (WENO) scheme in space and a third-order, total variation diminishing (TVD) Runge-Kutta scheme in time.

At the burner lip, a rotating boundary condition is used [15]. For details on the  $G$ -equation solver, the reader is referred to [16].

### 3.1. Base flow

For a time-independent base flow  $\mathbf{u} = \mathbf{U}$ , the  $G$ -equation (Eq. 1) becomes

$$\mathbf{U} \cdot \mathbf{n} - s_L^0(1 - \kappa\mathcal{L}) = 0 \quad . \quad (4)$$

For an axisymmetric burner, we parametrize the zero-level set, i.e. the flame surface, by setting its height  $h$  above the burner outlet as a function of the radius  $r$ :

$$G(r, z = h(r)) = 0 \quad , \quad 0 \leq r \leq R \quad . \quad (5)$$

The normal vector  $\mathbf{n}$  and the curvature  $\kappa$  (Eq. 3) in terms of  $h$  are

$$\mathbf{n} = \frac{1}{[1 + h'(r)^2]^{0.5}} \begin{pmatrix} -h'(r) \\ 1 \end{pmatrix} \quad , \quad (6)$$

$$\kappa = \frac{h''(r)}{[1 + h'(r)^2]^{1.5}} + \frac{h'(r)}{r[1 + h'(r)^2]^{0.5}} \quad . \quad (7)$$

Substituting Eqs. 6 and 7 into Eq. 4 gives a one-dimensional boundary value problem with  $h'(0) = 0$  and  $h(R) = 0$ .

For simplicity, we assume that the base flow  $\mathbf{U}$  is only a function of the radius  $r$ , but not of the height  $h$ . While the conditions at the burner outlet are theoretically known to be those of a Poiseuille-type pipe flow for a given mass flow rate, we introduce two additional parameters to account for the imperfections of this reduced-order model. Firstly, the velocity profile deviates from that of a Poiseuille flow towards a uniform flow the further we move away from the burner outlet [17]. Secondly, we observe that, as the base-flow speed vanishes near the burner wall due to the no-slip condition, the flame speed also decreases due to heat loss to the burner wall [18], which is not properly modeled by the constant adiabatic flame speed  $s_L^0$  (Eq. (2)). Hence, we introduce two additional parameters to the reduced-order model of the base flow: The aspect ratio  $\beta$  gives the ratio between the flame length and the burner radius in a uniform flow without curvature effects:

$$\left(\frac{\bar{U}}{s_L^0}\right)^2 = \beta^2 + 1 \quad , \quad (8)$$

where  $\bar{U}$  denotes the mean flow speed. The shape parameter  $\alpha$  linearly determines the velocity profile such that the mass flow rate is conserved:

$$\frac{U}{\bar{U}} = 1 + \alpha \left[1 - 2\left(\frac{r}{R}\right)^2\right] \quad , \quad (9)$$

where  $\alpha = 0$  corresponds to a uniform flow, and  $\alpha = 1$  corresponds to a Poiseuille flow. In summary, the base-flow model has three parameters  $\theta_{\text{bf}}$  ('base flow'):

the shape parameter  $\alpha$ , the aspect ratio  $\beta$  and the Markstein length  $\mathcal{L}$ . Note that  $\beta$  replaces the parameter  $s_L^0/\bar{U}$  due to Eq. (8).

For a given set of parameters, we solve the boundary value problem (Eqs. (4)-(7)) by using a finite-difference method and a Newton-Raphson solver [19]. In iteration step  $k$ , the residual  $\mathbf{R}^k = \mathbf{R}(\mathbf{h}^k)$  is computed by evaluating the left-hand side of Eq. (4) at every grid point. The discretization of the height,  $h$ , is represented by  $\mathbf{h}$ . The Jacobian  $\mathbf{J}^k = \mathbf{J}(\mathbf{h}^k)$  is computed by applying the chain rule to differentiate  $\mathbf{R}^k$  with respect to  $\mathbf{h}^k$ . Updates are performed by repeatedly solving

$$\mathbf{J}^k \Delta \mathbf{h}^k = \mathbf{R}^k \quad \Longrightarrow \quad \mathbf{h}^{k+1} = \mathbf{h}^k - \Delta \mathbf{h}^k \quad . \quad (10)$$

To infer the values of the parameters, we embed the boundary value problem into a least-squares problem with a cost functional  $\mathfrak{E}$ :

$$\mathfrak{E} = \sum_{m=1}^M [z_m - L_m(\mathbf{h}(\boldsymbol{\theta}_{\text{bf}}))]^2 \quad , \quad (11)$$

where  $L_m$  is a suitable linear interpolation operator for the  $m$ -th measurement  $(r_m, z_m)$  of the flame surface. This optimization problem is solved by line search [20]. The sensitivity of the cost functional  $\mathfrak{E}$  to the parameters  $\boldsymbol{\theta}_{\text{bf}}$  is calculated using the adjoint variable  $\boldsymbol{\lambda}$  [21]:

$$\boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \mathbf{h}} = \frac{\partial \mathfrak{E}}{\partial \mathbf{h}} \quad \Longrightarrow \quad \frac{d\mathfrak{E}}{d\boldsymbol{\theta}_{\text{bf}}} = \boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\theta}_{\text{bf}}} \quad . \quad (12)$$

In Fig. 4, the results are shown for  $\alpha = 0.84$ ,  $\beta = 15.1$  and  $\mathcal{L} = 3$  mm. The base-flow model agrees with the experiment. Furthermore, the base-flow model is confirmed using CANTERA simulations [22], which provide  $\beta \approx 15.8$ , based on the calculated adiabatic flame speed as well as the mass flow rate set in the experiment, and a flame thickness of 1.2 mm, the latter on the same order of magnitude as the inferred Markstein length, in agreement with the theory [23, 24]. Finally,  $\alpha = 0.84$  indicates a velocity profile close to Poiseuille flow as expected. Therefore, our model with three parameters covers a variety of base flows on physical grounds. Using a more complex base-flow model requires no conceptual changes to the data assimilation framework.

### 3.2. Flame response

For the forcing of the premixed flame, the radial and axial components of the velocity perturbation  $\mathbf{u}'$  are [25]

$$\frac{u'_r}{\bar{U}} = -\frac{\varepsilon \pi f K r}{\bar{U}} \cos \left( 2\pi f \left( \frac{Kz}{\bar{U}} - t \right) \right) \quad , \quad (13)$$

$$\frac{u'_z}{\bar{U}} = \varepsilon \sin \left( 2\pi f \left( \frac{Kz}{\bar{U}} - t \right) \right) \quad , \quad (14)$$

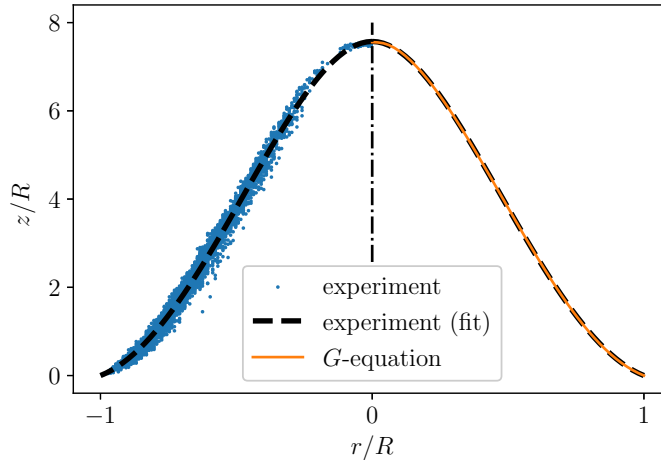


Figure 4: Edge detection (blue dots, left) and least-squares fit from base-flow model (orange line, right). For comparison, the fourth-order polynomial fit  $z/R = a_4(r/R)^4 + a_3(r/R)^3 + a_2(r/R)^2 - (a_4 + a_3 + a_2)$  respecting the boundary conditions is shown (black dashed line, both left and right). The base-flow model reproduces the flame surface observed in the experimental images.

where  $u'_r$  and  $u'_z$  satisfy the continuity equation. The forcing has a frequency  $f$  and an amplitude  $\varepsilon$ . The non-dimensional parameter  $K$  is the ratio between the mean flow speed and the phase speed at which perturbations on the flame surface travel in the axial direction. The model parameters  $\theta_{\text{fr}}$  (‘flame response’) are  $K$  and  $\varepsilon$ . In Fig. 2c, the result from a simulation based on our reduced-order model with parameters chosen to qualitatively match Fig. 2a,b is shown.

In preparation for data assimilation (Section 4), the  $G$ -equation has to be synchronized with the experimental images. This is not straightforward because we control the phase of the velocity perturbations in the  $G$ -equation (Eqs. (13), (14)), whereas the experimental images only depict the flame surface without any velocity information. Therefore, we derive an analytical relationship between the velocity perturbations and the motion of the flame surface at the base of the flame. As shown in Fig. 3 (right), the relationship between the base flow (Eq. (9)) and the flame surface (Eq. (8)) at the burner lip ( $r = R, z = 0$ ), neglecting curvature effects, is

$$\cos(\gamma) = \frac{s_L}{U} \Big|_R \approx \frac{s_L^0}{U(R)} \approx \frac{1}{\beta(1-\alpha)} \quad , \quad (15)$$

where the last approximation is justified for  $\beta = 15.1 \gg 1$  as observed in the experiment. The normal vector  $\mathbf{n}$  is

$$n_r \approx 1 \quad , \quad n_z = \frac{1}{\beta(1-\alpha)} \quad . \quad (16)$$

We consider a small-amplitude perturbation  $\gamma'$  around the angle  $\gamma$  as a result

of the corresponding velocity perturbation  $\mathbf{u}'$ :

$$\mathbf{u}' \cdot \mathbf{n} \approx -\varepsilon \pi f K R \sin(2\pi f t + \Delta\varphi) \quad , \quad (17)$$

where it is assumed that  $(\pi f K R \beta (1 - \alpha) / \bar{U})^2 \gg 1$  and  $\tan(\Delta\varphi) = \pi f K R \beta (1 - \alpha) / \bar{U}$ , which is justified by the inferred values for the model parameters  $\theta_{\text{bf}}$  and the judicious choice of frequencies  $f$  in Section 4. By observing the motion of the flame surface near the burner lip,  $\Delta\varphi$  is calibrated to synchronize the  $G$ -equation with the experimental images.

Before turning to the data-driven estimation of  $K$  and  $\varepsilon$  in the next section, we summarize the *a-priori* insights about the model parameters. Under the assumption that the velocity perturbation felt at the base of the flame, i.e.  $\hat{\gamma}'$ , only depends on the volume of the loudspeaker, it follows from Eq. (17) that the amplitude,  $\varepsilon$ , is inversely proportional to the frequency,  $f$ . Consequently, Eq. (17) implies a low-pass filter for the flame response [26]. The assumption that  $K$  only depends on the base flow, not the forcing frequency, is expected to be valid at small amplitudes, as shown by linear stability analysis [27]. The behavior at larger amplitudes can be investigated with the approach in this paper. While  $K \approx 1$  is reasonable for open flames in quiescent environments [28, 29], we additionally have to take into account the entrainment due to the buoyancy-driven flow surrounding the burner as well as the confinement due to the enclosing tube. Hence we anticipate a frequency-independent phase speed for the velocity perturbation with  $K < 1$ .

#### 4. Data assimilation

The Kalman filter provides a statistically optimal estimate  $\psi^a$  (‘analysis’) of the unknown state  $\psi$  of a system from a model prediction  $\psi^f$  (‘forecast’) and experimental observations  $\mathbf{d}$  [11]. The model prediction is mapped from its state space to the observation space through a measurement operator  $\mathbf{M}$ . The prediction uncertainties and the experimental errors are represented by covariance matrices  $\mathbf{C}_{\psi\psi}^f$  and  $\mathbf{C}_{\varepsilon\varepsilon}$ , respectively.

The application of the Kalman filter to the proposed reduced-order model is challenging for at least two reasons: Firstly, the  $G$ -equation is highly nonlinear, which is exemplified by the occurrence of cusps and pinched-off fuel-air pockets. This complicates the treatment of the time-dependent covariance matrix  $\mathbf{C}_{\psi\psi}^f$ . Secondly, the proposed reduced-order model has  $\mathcal{O}(10^5)$  degrees of freedom after discretization, which makes the computation and inversion of covariance matrices computationally intractable. To make the analysis statistically and computationally feasible, we instead generate an ensemble of  $N$  model predictions  $\psi_i^f$  with  $i = 1, \dots, N$ . This variation of the Kalman filter, the ensemble Kalman filter, gives for  $\psi_i^a$  and its statistics [11, 30]:

$$\psi_i^a = \psi_i^f + \left( \mathbf{M} \mathbf{C}_{\psi\psi}^f \right)^T \left[ \mathbf{C}_{\varepsilon\varepsilon} + \mathbf{M} \mathbf{C}_{\psi\psi}^f \mathbf{M}^T \right]^{-1} \left( \mathbf{d} - \mathbf{M} \psi_i^f \right) \quad , \quad (18)$$



$$\bar{\psi} = \frac{1}{N} \sum_{i=1}^N \psi_i \quad , \quad \mathbf{C}_{\psi\psi} = \frac{1}{N-1} \sum_{i=1}^N (\psi_i - \bar{\psi}) (\psi_i - \bar{\psi})^T \quad . \quad (19)$$

Finally, the question remains as to whether the (ensemble) Kalman filter preserves the kinematic nature of our reduced-order model. The choice of a suitable state space, i.e. the construction of  $\psi$  from  $G$ , is crucial and not straightforward. Hence, we adopt the level-set data assimilation framework developed by Yu et al. [30], which is based on the Hamilton-Jacobi formalism.

In state estimation, the ensemble  $\psi_i$  represents one realization from a probability distribution in  $\psi$  with mean  $\bar{\psi}$  and covariance matrix  $\mathbf{C}_{\psi\psi}$ . As such, the marginal probability distribution in the  $k$ -th entry of  $\psi$  is given by the mean  $\bar{\psi}[k]$  and the variance  $\mathbf{C}_{\psi\psi}[k, k]$ . Consequently, the likelihood of finding the flame surface at the location corresponding to the  $k$ -th entry of  $\psi$ , regardless of the position of the flame surface elsewhere, is [30]

$$p[k] = \frac{1}{\sqrt{2\pi\mathbf{C}_{\psi\psi}[k, k]}} \exp\left(-\frac{\bar{\psi}[k]^2}{2\mathbf{C}_{\psi\psi}[k, k]}\right) \quad . \quad (20)$$

Alternatively, the log-normalized likelihood is given by

$$\log\left(\frac{p[k]}{p_0[k]}\right) = -\frac{\bar{\psi}[k]^2}{2\mathbf{C}_{\psi\psi}[k, k]} \quad , \quad (21)$$

where  $\log(p/p_0) = 0$  identifies the most likely position of the flame surface. In Fig. 2d, the log-normalized likelihood is visualized for an educated guess of  $K \approx 0.5$  and  $\varepsilon \approx 0.2$  before any data assimilation. The position of the flame surface becomes highly uncertain for just a modest amount of standard deviation in the model parameters.

For combined state and parameter estimation, an augmented  $\tilde{\psi}$  is obtained by appending the model parameters  $\boldsymbol{\theta}_{\text{fr}}$ , i.e.  $K$  and  $\varepsilon$ , to  $\psi$ , and applying the ensemble Kalman filter to  $\tilde{\psi}$  [30]. In Fig. 2e, the results are shown for the same initial guess of  $K$  and  $\varepsilon$  as in Fig. 2d. In comparison, the identification of high-likelihood positions of the flame surface has significantly improved after combined state and parameter estimation. In Fig. 5, the joint probability distribution in  $K$  and  $\varepsilon$  is visualized. It is computed by marginalizing  $\psi$  from the probability distribution in  $\tilde{\psi}$ . Although the means have the same order of magnitude, the standard deviation in  $K$  is three times smaller. The parameter  $K$  is easier to infer because the pinch-off timing strongly depends on  $K$  and is captured accurately by the proposed data assimilation method. Moreover,  $K$  and  $\varepsilon$  are only weakly correlated, which confirms their distinct roles within the proposed reduced-order model.

To assess the proposed reduced-order model over a range of operating conditions, combined state and parameter estimation is performed for multiple frequencies ( $f = 200 - 400$  Hz). Due to the low-pass nature of premixed flames, low frequency forcing leads to large fuel-air pockets, which form cusps and pinch

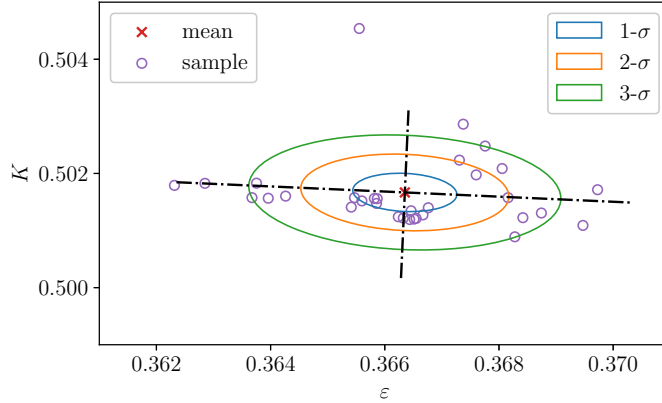


Figure 5: Sample of  $K$  and  $\varepsilon$  and reconstructed probability distribution after combined state and parameter estimation for  $f = 200$  Hz. The 1-, 2- and 3- $\sigma$  (blue/orange/green, respectively) confidence regions correspond to 39, 86 and 99% probabilities of sampling a given set of parameters, respectively.

off, while high frequency forcing leads to small perturbations which travel downstream without cusps. The range of forcing frequencies is chosen in order to avoid flame blow-off at lower frequencies as well as vanishing flame response at higher frequencies. An ensemble of 32  $G$ -equation simulations with the same initial condition is considered each time. The model parameters  $K$  and  $\varepsilon$  are sampled from independent normal distributions based around educated guesses of their values with 10% standard deviation in each. At first, the  $G$ -equation simulations are solved without data assimilation to make sure that the dynamics are consistent with the parameters and are free of transient effects. After 10 periods of forcing, the postprocessed experimental images ( $f = 2800$  Hz) are assimilated for 5 periods. The covariance matrix  $\mathbf{C}_{\varepsilon\varepsilon}$ , which represents the experimental errors, is a diagonal matrix with  $\sigma_\varepsilon^2$  on its diagonal. The choice of  $\sigma_\varepsilon = 1$  mm is based on the thickness of the detected edges from postprocessing (Fig. 2b). In general, the precision of the experimental data does not affect the accuracy of the ensemble in the long run, but it reduces the uncertainty in the ensemble overall [30]. Finally, the  $G$ -equation simulations are solved for another 5 periods to observe the performance with optimally calibrated model parameters and without any data assimilation. The overall computational time is less than 40 minutes for 20 periods on a cluster node with 1 processor per  $G$ -equation simulation. This includes the very frequent output of solution files in order to study time series and statistics for this paper. This is marginally longer than a single simulation of the  $G$ -equation as the ensemble Kalman filter is parallel by design.

In Fig. 6, the root-mean-square (RMS) error,

$$\text{RMS error} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\psi_i - \bar{\psi})^T (\psi_i - \bar{\psi})} \quad , \quad (22)$$

is plotted over time for combined state and parameter estimation at 200, 300 and 400 Hz. Within 5 periods of forcing, the dynamics for the sampled sets of parameters fully form. Between 10 and 15 periods, the ensemble Kalman filter gradually improves the estimates by up to two orders of magnitude. After 15 periods, the dynamics adapt to the optimally calibrated model parameters with low uncertainty. The remaining spurious oscillations are the result of noise in the experimental data. In Fig. 7, the optimally calibrated model parameters and their confidence intervals are shown for  $f = 200 - 400$  Hz. In agreement with theory, the model parameter  $K$  remains nearly constant while the model parameter  $\varepsilon$  decreases in inverse proportion to the frequency  $f$  (Section 3.2). The state of the system strongly varies with the operating conditions because of nonlinear effects. On the other hand, the model and its parametrization either follow a certain scaling or are constant. (This is because the proposed data-driven reduced-order model is physics-based, as opposed to traditional machine learning algorithms that are physics-blind.) This is clearly the case for the model parameters  $K$  and  $\varepsilon$  in Fig. 7. Therefore, we expect the reduced-order model to interpolate well despite a limited amount of data. The confidence intervals at the different frequencies are of comparable height. In agreement with Fig. 5,  $K$  is significantly more certain than  $\varepsilon$ .

In Fig. 8, the outcome of combined state and parameter estimation is visualized for forcing at 200, 300 and 400 Hz, respectively. The optimally calibrated reduced-order model accurately captures the perturbations traveling along the flame surface as well as the fuel-air pockets pinching off. While no individual  $G$ -equation simulation captures the motion of the flame surface completely, the  $G$ -equation simulations as an ensemble form an envelope in which the flame surface is fully contained, thus quantifying the uncertainty in the reduced-order model. While the pinched-off fuel-air pockets are clearly detectable in the experimental images for 200 Hz, the pinched-off fuel-air pockets are smaller in size for higher frequencies due to the low-pass nature of the flame response, and exist for shorter periods of time. Although the light intensity is diminished towards the tip of the flame surface, which complicates edge detection and observation in general, the optimally calibrated reduced-order model correctly infers the precise flame dynamics (300 Hz, left/middle left) that lead to the short-lived fuel-air pocket pinching off (300 Hz, middle right). As the perturbations traveling along the flame surface decrease in magnitude, so does the signal-to-noise ratio. Combined with the reduced relative frame rate at higher frequencies, the experimental images are ambiguous as to whether a fuel-air pocket pinches off, or the tip of the flame surface only retracts so rapidly because of the high local curvature (400 Hz, middle right). This ambiguity is reflected in the ensemble of  $G$ -equation simulations, especially in the elevated uncertainty towards the end

of the assimilation window, where some exhibit pinched-off fuel-air pockets with lifespans below the frame rate while others do not (Fig. 6).

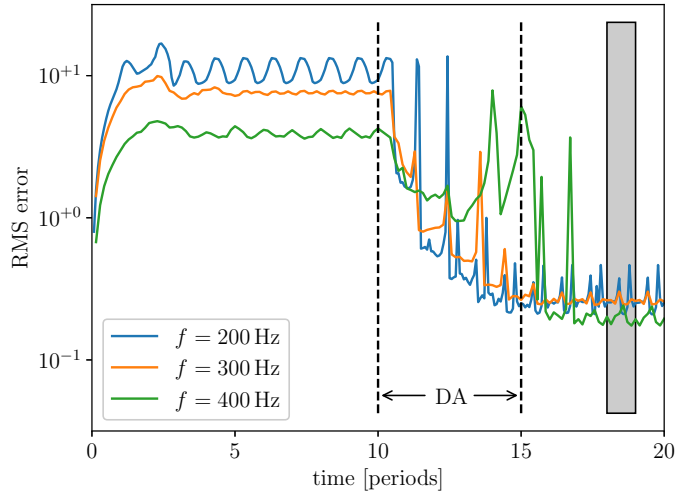


Figure 6: Root-mean-square (RMS) error for forcing at 200, 300 and 400 Hz (blue/orange/green, respectively). Data assimilation (DA) takes place between 10 and 15 periods. The grey window is representative of the period depicted in Fig. 8.

## 5. Conclusions

We develop and test a data-driven reduced-order model of a ducted premixed flame. This reduced-order model is nonlinear, adaptive and based on physical principles. For the first time, experimental data is rigorously assimilated into the model. This is a significant advancement compared to the assimilation of synthetic or simulation data [31]. Firstly, even a direct numerical simulation introduces assumptions into its underlying physical model, e.g. regarding the validity of reduced chemical mechanisms, ignored conjugate heat transfer or artificial flame thickening. Secondly, the experiment produces pinched-off fuel-air pockets unobserved in the available simulation data, which poses a particular challenge to the level-set method. Thirdly, the experimental data requires more sophisticated data processing in order to identify the flame surface and to remove light emission noise.

The model is validated by comparing its behavior with that from experimental data that it cannot observe. The two key aspects of the analysis are the following: Firstly, the  $G$ -equation is a fully nonlinear model. This includes non-smooth features, e.g. cusps on the flame surface, and discontinuities such as topological merging and break-up. Unlike classical, sensitivity-based approaches designed under linear assumptions, this probabilistic, ensemble-based

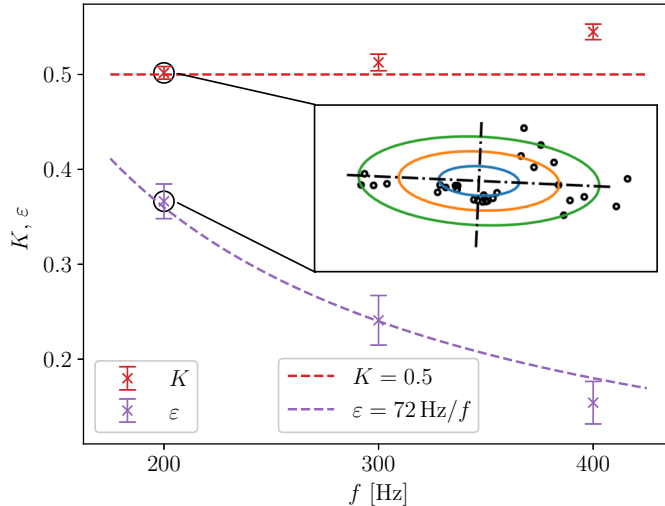


Figure 7: Optimally calibrated estimates and their uncertainties for  $K$  (red) and  $\varepsilon$  (purple). The  $10\text{-}\sigma$  (chosen for clearer visualisation) confidence intervals are computed by marginalizing the corresponding joint probability distributions (Fig. 5). Every joint probability distribution is reconstructed from an ensemble of 32  $G$ -equation simulations, 96 simulations in total. The dashed lines show the behavior estimated from theory (Section 3.2) for  $K$  (constant) and  $\varepsilon$  (inversely proportional to  $f$ ).

approach successfully models the nonlinearities and delivers an optimally calibrated reduced-order model. Secondly, the proposed level-set data assimilation framework is based on Bayesian inference [30]. As such, all estimates are equipped with statistically rigorous uncertainty quantification. This is highly relevant to the design of combustion systems: In thermoacoustics, for example, slight errors in the model deduced from a single burner could have a large impact on predictions in a different configuration, such as inside an annular combustor [5, 6]. The data assimilation framework was developed for level-set methods in general. As such, it is readily generalizable to other flame shapes modeled by the  $G$ -equation. In practice, the reduced-order model is optimally calibrated to laboratory experiments, and adapts on-the-fly when new observations become available during operation.

This study highlights the role that data can play in the future of combustion modeling for thermoacoustics. It is increasingly impractical to store data, particularly as experiments become automated [32] and numerical simulations become more detailed. Rather than store the data itself, the technique in this paper optimally assimilates the data into the parameters of a physics-based model. With this technique, rapid prototyping of combustion systems can feed into rapid calibration of their reduced-order models and then into gradient-based design optimization. While it has been shown, e.g. in the context of ignition and extinction, that large-eddy simulations become quantitatively predictive when

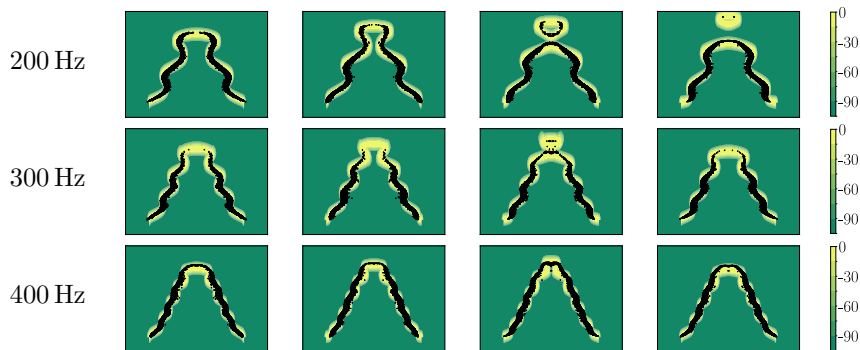


Figure 8: Snapshots of log-normalized likelihood (Eq. (21)) over one forcing period after combined state and parameter estimation for 200, 300 and 400 Hz (top/middle/bottom row, respectively). Highly likely positions of the flame surface are shown in yellow; less likely positions in green. The flame surface as detected from experimental images is included (black dots).

augmented with data [33], the reduced-order modeling of flame dynamics in turbulent flows remains challenging. For these challenging situations, this work opens up new possibilities for the development of reduced-order models that adaptively change any time that data from experiments or simulations becomes available.

### Acknowledgments

The authors would like to thank U. Sengupta for the acquisition of the experimental data. H. Yu is supported by the Cambridge Commonwealth, European & International Trust under a Schlumberger Cambridge International Scholarship. L. Magri gratefully acknowledges support from the Royal Academy of Engineering Research Fellowships scheme 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.

### References

- [1] T. C. Lieuwen, V. Yang (Eds.), *Combustion Instabilities in Gas Turbine Engines: Operational Experience, Fundamental Mechanisms and Modeling*, Progress in Astronautics and Aeronautics, AIAA, 2005.
- [2] F. E. C. Culick, *Unsteady motions in combustion chambers for propulsion systems*, Tech. rep., AGARD (2006).
- [3] J. W. Strutt, The explanation of certain acoustical phenomena, *Nature* 18 (455) (1878) 319–321.

- [4] T. Poinso, Prediction and control of combustion instabilities in real engines, *Proc. Combust. Inst.* 36 (1) (2017) 1–28.
- [5] M. P. Juniper, R. Sujith, Sensitivity and Nonlinearity of Thermoacoustic Oscillations, *Annu. Rev. Fluid Mech.* 50 (1) (2018) 661–689.
- [6] L. Magri, Adjoint Methods as Design Tools in Thermoacoustics, *Appl. Mech. Rev.* 71 (2) (2019) 020801.
- [7] K. Duraisamy, G. Iaccarino, H. Xiao, Turbulence Modeling in the Age of Data, *Annu. Rev. Fluid Mech.* 51 (1) (2019) 357–377.
- [8] S. L. Brunton, B. R. Noack, P. Koumoutsakos, Machine Learning for Fluid Mechanics, *Annu. Rev. Fluid Mech.* 52 (1).
- [9] P. Benner, S. Gugercin, K. Willcox, A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems, *SIAM Review* 57 (4) (2015) 483–531.
- [10] A. P. Dowling, A kinematic model of a ducted flame, *J. Fluid Mech.* 394 (1999) 51–72.
- [11] G. Evensen, *Data Assimilation*, Springer, 2009.
- [12] L. Magri, N. A. K. Doan, *Physics-Informed Data-Driven Prediction of Turbulent Reacting Flows with Lyapunov Analysis and Sequential Data Assimilation*, Springer International Publishing, Cham, 2020, pp. 177–196. doi:10.1007/978-3-030-44718-2\_9. URL [https://doi.org/10.1007/978-3-030-44718-2\\_{\\_}9](https://doi.org/10.1007/978-3-030-44718-2_{_}9)
- [13] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, scikit-image: image processing in python, *PeerJ* 2 (2014) e453.
- [14] N. Peters, *Turbulent Combustion*, Cambridge University Press, 2000.
- [15] I. Waugh, *Methods for Analysis of Nonlinear Thermoacoustic Systems*, Ph.D. thesis, University of Cambridge (Mar. 2013).
- [16] S. Hemchandra, *Dynamics of Turbulent Premixed Flames in Acoustic Fields* Dynamics of Turbulent Premixed Flames in, Ph.D. thesis, Georgia Institute of Technology (Aug. 2009).
- [17] H. Schlichting, K. Gersten, *Grenzschicht-Theorie*, Springer, 2006.
- [18] A. Cuquel, D. Durox, T. Schuller, Impact of flame base dynamics on the non-linear frequency response of conical flames, *Comptes Rendus Mécanique* 341 (1-2) (2013) 171–180.
- [19] U. M. Ascher, R. M. M. Mattheij, R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, SIAM, 1995.

- [20] J. Nocedal, S. J. Wright, Numerical Optimization, Springer, 2006.
- [21] M. D. Gunzburger, Perspectives in Flow Control and Optimization, SIAM, 2002.
- [22] D. G. Goodwin, R. L. Speth, H. K. Moffat, B. W. Weber, Cantera: An Object-oriented Software Toolkit for Chemical Kinetics, Thermodynamics, and Transport Processes (Aug. 2018).
- [23] P. Clavin, F. A. Williams, Effects of molecular diffusion and of thermal expansion on the structure and dynamics of premixed flames in turbulent flows of large scale and low intensity, *J. Fluid Mech.* 116 (1982) 251–282.
- [24] M. Matalon, B. J. Matkowsky, Flames as gasdynamic discontinuities, *J. Fluid Mech.* 124 (1982) 239.
- [25] K. Kashinath, S. Hemchandra, M. P. Juniper, Nonlinear thermoacoustics of ducted premixed flames: The influence of perturbation convection speed, *Combust. Flame* 160 (12) (2013) 2856–2865.
- [26] S. Ducruix, D. Durox, S. Candel, Theoretical and experimental determinations of the transfer function of a laminar premixed flame, *Proc. Combust. Inst.* 28 (1) (2000) 765–773.
- [27] L. Landau, On the theory of slow combustion, in: P. Pelcé (Ed.), *Dynamics of Curved Fronts*, Academic Press, San Diego, 1988, pp. 403 – 411.
- [28] T. Schuller, S. Ducruix, D. Durox, S. Candel, Modeling tools for the prediction of premixed flame transfer functions, *Proc. Combust. Inst.* 29 (1) (2002) 107–113.
- [29] T. Schuller, D. Durox, S. Candel, A unified model for the prediction of laminar flame transfer functions, *Combust. Flame* 134 (1-2) (2003) 21–34.
- [30] H. Yu, M. P. Juniper, L. Magri, Combined state and parameter estimation in level-set methods, *J. Comput. Phys.* 399 (2019) 108950.
- [31] H. Yu, T. Jaravel, M. Ihme, M. Juniper, L. Magri, Data Assimilation and Optimal Calibration in Nonlinear Models of Flame Dynamics, *J. Eng. Gas Turbine Power*.
- [32] G. Rigas, N. P. Jamieson, L. K. B. Li, M. P. Juniper, Experimental sensitivity analysis and control of thermoacoustic systems, *J. Fluid Mech.* 787.
- [33] J. W. Labahn, H. Wu, B. Coriton, J. H. Frank, M. Ihme, Data assimilation using high-speed measurements and LES to examine local extinction events in turbulent flames, *Proc. Combust. Inst.*