Proceedings of the Combustion Institute 40 (2024) 105304

Contents lists available at ScienceDirect

Proceedings of the Combustion Institute

journal homepage: www.elsevier.com/locate/proci

Ensemble predictions of laser ignition with a hybrid stochastic physics-embedded deep-learning framework

Wai Tong Chung [a,*], Charlélie Laurent [b], Donatella Passiatore [b], Matthias Ihme [a,c]

[a] Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA
[b] Center for Turbulence Research, Stanford, CA 94305, USA
[c] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

ARTICLE INFO

Keywords:
Laser ignition
Machine learning
Stochastic modeling
Ignition probability map
Rocket combustion

ABSTRACT

When investigating stochastic ignition phenomena, high-fidelity simulations and experiments for statistical characterization can be a laborious endeavor. Machine learning (ML), in particular deep-learning, offers an accurate and cost-effective approach for reduced-order modeling of forced ignition. In this work, we investigate the potential and limitations of ML in modeling ignition, especially in the presence of limited data and multiple ignition modes. To this end, we introduce a hybrid stochastic physics-embedded deep-learning framework that combines sparse experimental data from Schlieren measurements with inert large-eddy simulations for predicting the spatio-temporal evolution of ignition kernel location and morphology. This model combines a stochastic differential equation for modeling kernel dynamics and a deep-learning model for representing kernel morphology. We evaluate this modeling approach for laser ignition of a gaseous methane/oxygen (CH_4/O_2) model rocket combustor, where morphological effects of the ignition kernel significantly influence ignition behavior. Results demonstrate that this model can reasonably capture behavior associated with the three dominant ignition modes, namely direct, indirect, and failed ignition, along with statistics associated with kernel growth and position, at lower computational costs than high-fidelity reacting simulations. In addition, we demonstrate that this modeling framework can be employed for generating spatially resolved ignition probability maps by incorporating physics to represent kernel interaction with the turbulent jet. We note that limitations in accuracy can be observed when predicting with vastly out-of-distribution data. Nevertheless, these results demonstrate that this physics-embedded ML approach can statistically characterize forced ignition in a more cost-effective manner than reacting high-fidelity simulations, as long as sufficiently representative data is available.

1. Introduction

Laser-induced ignition of propulsion systems offers numerous benefits, including multiple ignition positions, accurate controllability, high kernel deposition energy, and the ability to ignite leaner mixtures [1]. During ignition, stochasticity arises from the interaction of the ignition kernel with the turbulent fuel/oxidizer mixture, and from variations in kernel deposition energy, which can affect the evolution of the ignition kernel [2]. Thus, the characterization of statistical behavior within ignition phenomena is essential for robust deployment of these systems.

Ensembles of experimental measurements are typically used to investigate stochastic ignition under a variety of conditions. For example, Ahmed et al. [3] performed ensemble tests on a CH_4/air counterflow flame, and found that convective and local strain effects were significant in influencing ignition probability in this fundamental configuration. Cordier et al. [4] employed a premixed CH_4/air confined swirl combustor configuration to show the importance of turbulent flow properties in influencing kernel trajectory and ignition probability. Several studies [5–7] have focused on the effects of fuel chemistry and mixture composition in determining ignition probability in experimental configurations that represent gas-turbine combustors. Strelau et al. [8] investigated laser ignition in a non-premixed gaseous CH_4/O_2 model rocket combustor and found three distinct modes of successful ignition: (i) direct ignition, where the hot plasma is deposited in a reactive fuel/oxidizer mixture and rapidly transitions into a sustained flame, (ii) indirect ignition, where the hot plasma is deposited outside the reactive mixture and transitions to a flame after the asymmetric plasma kernel slowly interacts with the turbulent jet, and (iii) failed ignition.

High-fidelity simulations can provide detailed insights about ignition phenomena. For instance, Lacaze et al. [9] performed large-eddy

* Corresponding author.
E-mail address: wtchung@stanford.edu (W.T. Chung).

https://doi.org/10.1016/j.proci.2024.105304
Received 4 December 2023; Accepted 23 May 2024
Available online 29 June 2024

simulations (LES) of a model rocket combustor [10] to examine ignition processes as well as flame propagation, anchoring, and stabilization mechanisms. Wang et al. [11] performed a direct numerical simulation (DNS) of a laser kernel in quiescent air to demonstrate the importance of initial kernel morphology on influencing plasma-ejection that leads to the aforementioned indirect ignition. Ensemble studies with high-fidelity simulations can incur large computational costs, especially when treating combustion with finite-rate chemistry [12,13]. Thus, many ensemble simulations have employed cost-efficient combustion models involving global [14] or tabulated [15] chemistry.

Reduced physics-based representation of ignition phenomena can also present a cost-effective alternative for evaluating the fundamental and statistical behavior of forced ignition phenomena. Richardson and Mastorakos [16] performed 1D counterflow flame calculations (of varying strain rates, ignition kernel temperature, and ignition kernel location) to show that successful ignition relies on both the underlying flow conditions and fuel/oxidizer mixture composition. A few reduced-order models have combined (i) probabilistic arguments with (ii) flowfields obtained from high-fidelity inert simulations. Neophytou et al. [17] estimated ignition probability with an ensemble-based model that predicts volume fraction of burned gases using spark location, initial kernel size, as well as time-averaged turbulent velocity and mixture fraction flowfields that were extracted from inert LES. Esclapez et al. [18] employed flowfields from inert LES to develop a stochastic differential equation-based (SDE) model that directly transports probability density functions (PDFs) to generate an ignition map for a swirl combustor. While these approaches have demonstrated their predictive capabilities, they do not consider phenomena related to kernel morphology and plasma-ejection found in laser ignition systems.

Machine learning (ML) methods offer opportunities for reduced-order modeling within a wide range of combustion phenomena [19] – in both fundamental [20,21] and system-level [22,23] problems. In a forced ignition study, Sforzo and Seitzman [24] developed a support vector machine-based approach for predicting ignition probability within partially-stirred reactor simulations. Popov et al. [25] showed that convolutional neural networks (CNNs) outperformed other reduced-order approaches in ignition boundary prediction of a jet-in-crossflow configuration. The authors noted that this deep learning approach can extract spatial information effectively, but suffers from poor out-of-distribution (OOD) predictions if developed without considering essential physics.

The objective of this work is to identify the utility and limitations of ML methods for spatially predicting and statistically characterizing ignition phenomena, especially in the presence of multiple ignition modes and limited data. To this end, this work introduces a reduced-order physics-embedded SDE-ML framework for modeling ignition. We evaluate this approach for modeling ignition within the gaseous CH_4/O_2 model rocket combustor [8] detailed in Section 2, while the methods employed within this work are introduced in Section 3. We discuss results from this work in Section 4, before providing concluding remarks in Section 5.

2. Configuration

In this work, we model ignition within the gaseous CH_4/O_2 model rocket combustor by Strelau et al. [8]. This optically accessible configuration was designed specifically for statistically characterizing laser ignition phenomena of gaseous mixtures within rockets. The experimental configuration consists of a shear co-axial injector where the oxidizer flows through a central axisymmetric jet with diameter d_o = 3.57 mm, while fuel is injected through an annulus with inner and outer diameters $d_{f,i}$ = 5.33 mm and $d_{f,o}$ = 6.35 mm. The cylindrical combustion chamber has a total length of 111 mm and a diameter of d_{ch} = 50.8 mm. The combustor operates at a global oxidizer-to-fuel ratio of approximately $\phi = 3$, with mass flow rates of oxidizer $\dot{m}_o$ = 6.58 g/s and fuel $\dot{m}_f$ = 2.11 g/s that correspond to sonic and subsonic conditions,

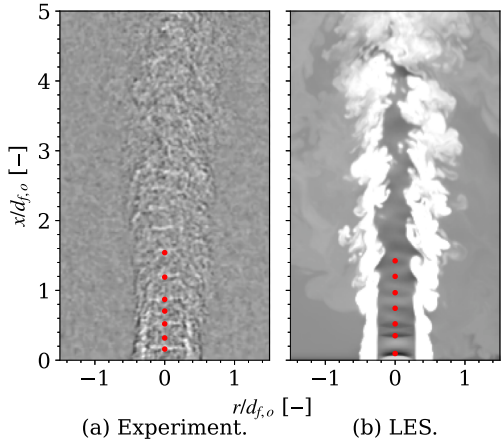


Fig. 1. Instantaneous (a) experimental Schlieren measurements [8] with (b) density fields from the present LES. Shock train locations shown in red are quantified in Table 1.

respectively. The temperature of the oxidizer and the fuel supplied at the injector inlet are T_o = 242 K and T_f = 282 K, respectively. Prior to ignition, the reactants are injected to pressurize the chamber until reaching a quasi-steady-state with a nominal operating pressure of 1.4 bar. The laser is then deployed near this non-premixed mixture via an Innolas Spitlight Standard 600-10 laser in single-shot mode, which generated a 532 nm laser pulse with laser energy of $E_{laser} = 22 \pm 7$ mJ.

In this work, we generate datasets for developing and testing the SDE-ML model. In particular, we process Schlieren measurements of N_{exp} = 153 laser ignition tests, across 21 laser deposition locations in the jet centerplane. The measurements are recorded at an acquisition rate of 500 kHz. Approximately five ignition tests were performed in laser locations where ignition probability $P_{ig} = 0$ and $P_{ig} = 1$, resulting in an estimated uncertainty of ±0.2. For laser locations where ignition probability $0 < P_{ig} < 1$, approximately ten ignition tests were performed, resulting in an estimated uncertainty of ±0.1.

Previous studies [17,18] have noted that reduced-order models for forced ignition phenomena can be constructed with inputs from inert simulations to represent flows during the early stages of ignition — where dilational effects from heat release are not yet relevant. A non-reacting LES is calculated with a high-order finite-difference compressible flow solver [26]. For spatial discretization, a sixth-order skew-symmetric scheme is employed with a targeted essentially non-oscillatory (TENO) scheme [27] for shock-capturing, while time advancement is performed via a third-order strong-stability-preserving Runge–Kutta (SSP-RK3) method [28]. LES closure is provided via the Smagorinsky model [29], with turbulent Prandtl and Schmidt numbers of 0.7. Multi-component species transport [30] is employed to represent molecular mixing within this configuration. This LES was advanced with a timestep size/ of 16 ns, corresponding to an approximate acoustic CFL = 0.1, on 96 NVIDIA V100 GPUs with a wall-clock-time of 0.6 s per timestep.

The LES domain consists of a single-block curvilinear mesh with 221M (960 × 480 × 480) mesh points, which is stretched in all three directions. The finest mesh spacing is at the injector exit, where the mesh is locally uniform with cell size of 88 μm, was chosen by examining the sensitivity of LES predictions of shock train locations across different mesh resolutions. These shock train locations, which have previously been used to validate jet topology of rocket combustors [4], were determined by observing the presence of peaks in pixel intensity and density magnitudes in the centerline of time-averaged measurements and LES, respectively. Fig. 1 compares instantaneous experimental Schlieren and simulated density fields, with reasonable agreement seen between the visible shock train locations in the predictions. These shock train locations are quantified in Table 1, showing a maximum

Table 1
Comparison of shock train locations from LES against experimental measurements.

Shock train locations normalized by jet diameter $d_{f,o}$ [−].							
Exp.	0.16	0.32	0.52	0.70	0.87	1.19	1.54
LES	0.10	0.35	0.52	0.74	0.97	1.20	1.42
Error	0.06	0.03	0.00	0.04	0.10	0.01	0.12

discrepancy of $0.12d_{f,o}$. We note that one of the key challenges presented by this complex high-speed multi-physics flow configuration is the development of a data-driven model that is sufficiently robust to the limited information provided by experimental measurements, as well as discrepancies between simulation and measured data.

3. Methods

3.1. SDE-ML framework

In this section, we describe the 2D SDE-ML approach for utilizing (i) both 3D LES and 2D experimental data for modeling kernel morphology via the deep learning approach, and (ii) LES statistics for informing SDE-based kernel transport, as summarized by Fig. 2. Ensemble calculations can be performed with this approach to evaluate statistical behavior of laser ignition in the present configuration.

In this work, we represent space- and time-dependent kernel morphology $Y^n(\mathbf{x}, t^n)$ as a binary flowfield of ignited and non-ignited segments. Thus, the spatio-temporal evolution of kernel morphology for timestep n is expressed as:

$$Y^n(\mathbf{x}, t^n) = \begin{cases} 1 & \text{ignited cell at } t^n\,, \\ 0 & \text{non-ignited cell at } t^n\,. \end{cases} \tag{1}$$

To obtain this binary flowfield, we process Schlieren measurements of the present configuration. We note that the proposed SDE-ML method is agnostic to the source of kernel morphology data, and could be extended to employ data from high-fidelity reacting LES. However, ML models typically perform more accurately with increasing data diversity and volume [31]. Thus, we employ already available data from the 153 experimental tests to demonstrate the present framework. In future work, this data scarcity could be addressed by extending the present framework with multi-fidelity training [32] that combines small samples of high-fidelity reacting LES with large samples of more affordable and lower-fidelity simulation data.

The use of conventional image processing techniques, involving spatio-temporal filters and edge detection algorithms, can be difficult for this specific configuration due to (i) differences in timescales encountered between direct and indirect ignition, and (ii) measured structures from the jet obstructing ignition kernel structure. Thus, we employ an open-source video segmentation tool [33] for extracting a total of 19,765 frames of kernel morphology segments Y^n (partially shown in orange in Fig. 2). Each ignition test case is segmented starting from time after laser deposition $\tau = t - t_{laser} = 4\,\mu s$ to avoid imaging artifacts that arise from the pressure wave that forms immediately after laser deposition. Each of the frames of the kernel segments is manually inspected for quality prior to ML training.

An ML model f_{ML} is employed for autoregressively generating next-timestep predictions $\hat{Y}^{n+1}$ of kernel morphology via segmentation, which classifies flowfield segments as ignited and non-ignited via:

$$\hat{Y}^{n+1}(\mathbf{x}, t^{n+1}) = f_{\mathrm{ML}}(\hat{Y}^n, \overline{\mathbf{u}}, \mathbf{u}', \overline{Z}, Z')\,, \tag{2}$$

with an initial condition $\hat{Y}^0 \equiv Y^0$. The ML inputs consist of stacked channels of (i) mean $\bar{\cdot}$ and fluctuating $\cdot'$ components of velocity $\mathbf{u} = [u_x, u_r]^\intercal$ and mixture fraction Z extracted from temporal- and azimuthal-averaged flowfields over 320 μs (representative of the time taken after laser deposition to transition to a flame), as well as (ii) the kernel segments from the previous timestep. This choice of ML inputs is motivated by previous work involving rule-based and analytic reduced-order ignition models [17,18], as well as fundamental understanding of the relationship between successful ignition with the underlying flow and mixture conditions [16]. These seven inputs are stacked together with axial, transverse, and channel dimensions of $N_x \times N_r \times N_c = 250\times160\times7$ that can be fed into a convolutional neural network (CNN).

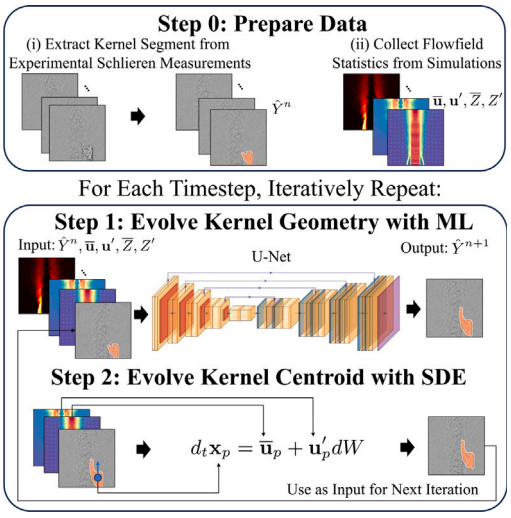


Fig. 2. SDE-ML framework for modeling stochastic ignition.

To predict the evolution of kernel morphology, we employ a U-Net [34] model (shown in Fig. 2) – a well-established deep-learning architecture suited for segmentation tasks. This CNN uses a small moving window, known as a filter, that performs a mathematical operation (typically convolution and pooling operations [19]) on a spatial neighborhood of values (in the first CNN layer containing $[\hat{Y}^n, \overline{\mathbf{u}}, \mathbf{u}', \overline{Z}, Z']^\intercal$), thereby enabling higher prediction accuracy in spatial problems compared to other ML algorithms. A key aspect of this model design involves the use of skip connections that connect processed features from the later hidden layers with unprocessed features from the earlier layers — which helps in preserving fine-grained details during segmentation.

When generating ML predictions, CNN filters at the input layer supplement the kernel segment with surrounding flowfield information from the inert LES statistics. Even though the LES flowfield remains the same across data samples, the flowfields vary from the perspective of the convolutional filter and provide local mixture information for the CNN window containing an ignited kernel segment. If these channels containing local mixture information were not present, the CNN would not have sufficient information to propagate or extinguish the ignition kernel, since there would be no inputs that inform the CNN on the proximity of the ignition kernel to the reactive/non-reactive mixtures.

The computational domain ($N_x \times N_r = 250 \times 160$) for this reduced-order model is discretized with a uniform mesh size of 180 μm in both axial and radial directions. This grid size was selected by considering limitations in computational memory during multi-GPU training of the ML model, and is similar in size with the inert LES grid described in Section 2. A convergence study examining the effects of varying grid sizes is presented in the supplementary material.

To model stochastic variations of different ignition trajectories, we treat the ignition kernel centroid as a Lagrangian particle with position $\mathbf{x}_p$ that is advanced via the following SDE:

$$d_t\mathbf{x}_p = \overline{\mathbf{u}}_p + \mathbf{u}'_p dW\,, \tag{3a}$$

$$\mathbf{u}_p = \int \hat{Y}\mathbf{u}\, d\hat{Y} \,/ \int \hat{Y}\, d\hat{Y}\,, \tag{3b}$$

$$\mathbf{x}_p = \int \hat{Y}\mathbf{x}\, d\hat{Y} \,/ \int \hat{Y}\, d\hat{Y}\,, \tag{3c}$$

where a Gaussian distribution $\mathcal{N}(0,\sigma)$ introduces turbulent fluctuations via a Wiener process dW [35]. The Gaussian standard deviation $\sigma = 0.125$ was determined with a hyperparameter search on a validation set discussed in Section 3.2. Both mean and fluctuating velocity components of the particle are approximated by averaging the velocity field that intersects with the area of kernel segments. Eq. (3a) is advanced via a forward Euler scheme with a timestep of $\Delta t_{\mathrm{SDE}} = \Delta t_{\mathrm{ML}} = 8\,\mu s$, which enables the kernel centroid to be transported across the different cells in the SDE-ML grid, while ensuring that the number of Schlieren samples are sufficiently large ($\mathcal{O}(10^4)$) for training a deep-learning model. During position updates, the ML channel with the kernel segment $\hat{Y}^n$ is updated with the new kernel $\hat{Y}^{n+1}$, while preserving the other channels (since flowfield statistics are assumed to be time-independent), before repeating the ML and SDE steps for the next timestep iteration. We note that prior to SDE advancement, all ignited cells are spatially translated by $\Delta \mathbf{x}_{\mathrm{ML}} = \mathbf{x}_p(Y^{n+1}) - \mathbf{x}_p(Y^n)$ so that kernel transport is governed solely by the SDE without advective effects from the ML model. A wall-clock-time of 0.3 s with one V100 GPU generates a single ignition trajectory with 128 timesteps that compounds to a model duration of 1 ms.

3.2. ML setup

The present U-Net contains 82 layers, corresponding to approximately 31M trainable parameters, with the initial weights set via He initialization [36]. During training, the Adam optimizer minimizes the cell-wise cross-entropy loss with an initial learning rate of `1e-4` and batch size of 32. These hyperparameters were selected to match default U-Net settings [34]. Training this U-Net via distributed data parallelism and mixed precision on `PyTorch Lightning 1.6.5` [37] on four V100 GPUs requires approximately 1 h wall-clock-time.

The processed dataset, with the input–output pair described by Eq. (2), is split to consider 18 in-distribution and three OOD laser deposition locations. The three OOD locations are selected to each represent direct/indirect/failed ignition phenomena that can be used to evaluate the SDE-ML model's behavior when extrapolating beyond seen laser deposition locations. 80% of the in-distribution data is used for training, while the remaining 20% are split evenly for validating and unit testing the ML component of the present framework via cell accuracy (defined as the percentage of cells that are classified correctly). Without the SDE component, teacher forced (where the ML input consists only of the ground-truth samples Y^n) next-frame predictions from this model provides a cell accuracy of 99.7% in the test set. Without the SDE component, we note that in the autoregressive (where the ML outputs $\hat{Y}^n$ are iteratively used as the next timestep's inputs, as described by Eq. (2)) next-frame predictions, small errors from previous iterations accumulate (a well-known ML property [38]) — leading to approximately 90% test cell accuracy after 30 ML timesteps (with an in-prediction duration of 240 μs).

4. Results

Here, we report results from ensemble kernel trajectory predictions made by the present SDE-ML modeling framework on the present configuration. In this work, each sample of initial kernel morphology condition Y^0 is used to generate an ensemble with 100 stochastic variations, *i.e.*, approximately 100-fold larger than the experimental ensemble.

Fig. 3 shows instantaneous predictions from the SDE-ML model (in red) that qualitatively captures kernel behavior seen in experimental measurements (in grayscale) for (a) direct, (b) indirect, and (c) failed ignition phenomena. At time after laser deposition $\tau = 4\,\mu s$, a small ignition kernel is observed at the three different laser deposition locations, with a pressure wave that is surrounding the kernel. In the direct ignition seen in Fig. 3a, the kernel deposited in the central jet rapidly transitions into a sustained flame within $\tau = 28\,\mu s$ and continues to propagate, which is also seen with the SDE-ML model. In Fig. 3b,

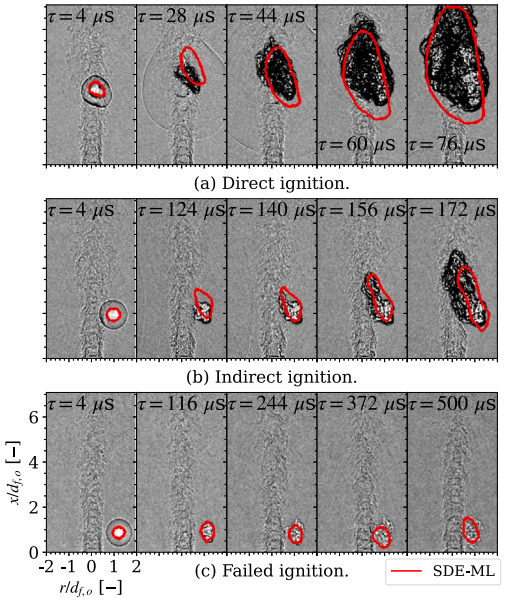


Fig. 3. Comparisons of ignition kernel predictions from the SDE-ML model against experimental measurements of direct/indirect/failed ignition for time after laser deposition τ.

we see that the kernel deposited outside the jet does not ignite the fuel/oxidizer mixture until approximately $\tau = 140\,\mu s$. After this time, the hot plasma ejected from the asymmetric kernel interacts with the central jet, which causes a transition to a sustained flame with an oblong morphology. This indirect ignition phenomenon highlights the influence of kernel morphology in laser ignition, which is reasonably captured by the SDE-ML model. In Fig. 3c, we observe that the SDE-ML can also accurately capture the absence of a sustained flame up until a long duration of $\tau = 500\,\mu s$.

We examine the statistical behavior of SDE-ML ensemble predictions by comparing the temporal evolution of mean kernel centroid positions against corresponding measurements in Fig. 4. These positions are overlaid on top of mean velocity magnitudes and unit vectors. The SDE-ML model predicts trajectories that qualitatively agree with the measurements. Specifically, it can be seen that the kernels are transported (i) downstream by the jet, and (ii) radially due to entrainment. Near the central jet core ($r < 0.5 d_{f,o}$), mean velocity trajectory predictions agree with the experiment in time and position. However, we note that outside this range, quantitative discrepancies are observed in the SDE-ML predictions. On average, the ignition kernels predicted by the SDE-ML model in these locations are not transported as deeply (radial differences up to $0.4 d_{f,o}$) into the jet as the averaged experimental measurements, indicating an under-prediction of kernel centroid radial velocity by the SDE-ML model. In addition, the kernels in these locations require up to 50% more time to reach the same axial location, when compared to experimental measurements. Sources of deviation in the kernel trajectory in this reduced-order model result from the potential discrepancies between LES input and real-world velocity. LES improvements (including turbulence modeling, boundary condition treatment, and grid refinement) would result in flowfields that transport the modeled kernel in a more similar trajectory to the experiments. In addition, another source of discrepancy is the 2D treatment of this 3D configuration, which captures the trajectory of the kernel in a projected 2D plane. As such, any azimuthal velocity in the 3D configuration will cause an apparent motion of the kernel towards and away from the jet, which is not accounted for in the present modeling approach. Another potential limitation in the current 2D approach is the absence of laser deposition angles within the modeling framework. However, laser angle was not varied in the corresponding experimental configuration [8],

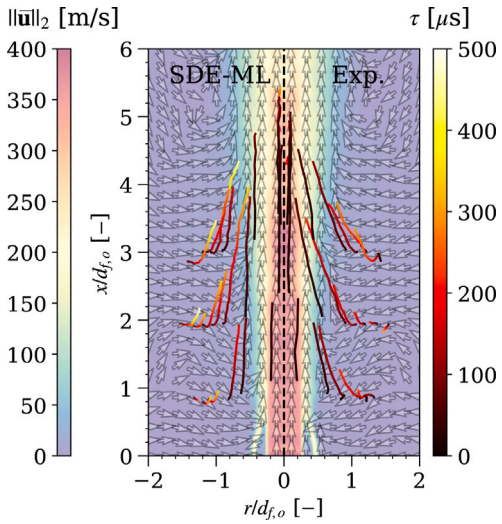


Fig. 4. Mean kernel position trajectory from ensemble SDE-ML predictions against ensemble experimental measurements across time after laser deposition τ. LES mean velocity magnitude (with translucent unit vectors) is also shown.

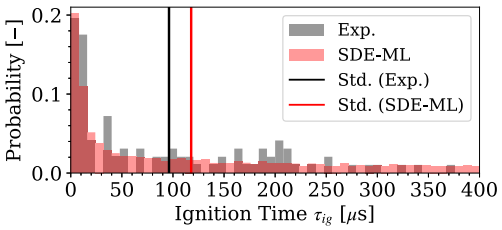


Fig. 5. Comparison of normalized ignition time τ_{ig} distributions from SDE-ML predictions against measurements.

so examining the full extent of 3D ignition is beyond the scope of the present work, and should be investigated in future work if high quality data relating 2D flowfields with 3D effects becomes available.

Nevertheless, this velocity treatment is still sufficient for capturing the probability of ignition delay τ_{ig} from the ensemble experiments shown in Fig. 5. In another reduced-order modeling study [17], successful ignition was determined by monitoring the ignition kernel growth rate. Here, ignition time is defined as the time taken for the kernel to exceed a growth rate of 0.2 mm^2/s in both SDE-ML predictions and experimental measurements. Estimated probabilities of the ignition time from ensemble predictions and measurements show good agreement in the distribution peak (~20%) at approximately $\tau_{ig} \approx 10\,\mu$s, which correspond to the large proportion of direct ignition cases within the present ensemble. After $\tau_{ig} \approx 10\,\mu$s, long-tailed ignition time distributions are seen for both predictions and measurements, which are a result of the delayed indirect ignition that is influenced by stochastic variations of kernel interactions with the turbulent fuel/oxidizer mixture. Ensemble SDE-ML predictions of τ_{ig} possess a standard deviation of 123 µs, which is within 18% difference with measured standard deviation of 88 µs.

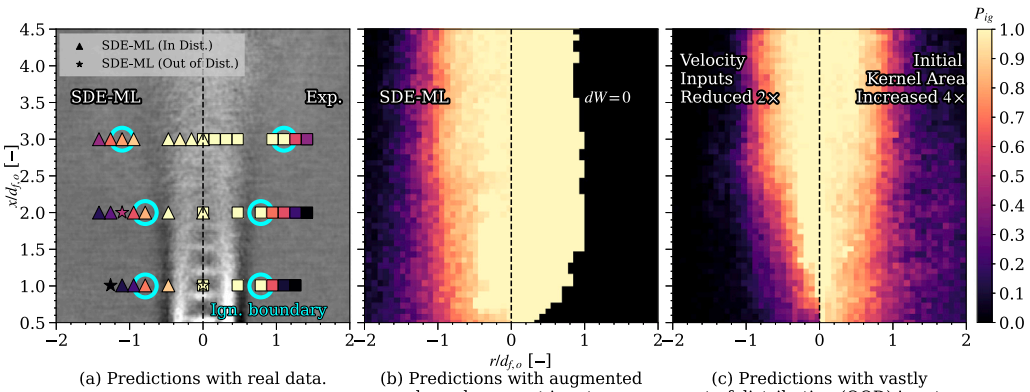
In this work, we map the probability of successful ignition with real data and three augmented datasets. Real data consists of the segmented Schlieren data (stacked with inert LES statistics), as described in Section 2. Since experimental ignition tests were only performed for 21 different laser deposition locations, an augmented dataset was generated with the purpose of estimating a spatially resolved ignition probability map. This augmented kernel dataset was created by spatially translating all ignited cells of a single frame of a kernel segment $Y^0(\tau = 4\,\mu\text{s},\ x = 2d_{f,o},\ r = 1.1d_{f,o})$ by $\Delta\mathbf{x}_{aug} = \mathbf{x}_{aug} - \mathbf{x}_p^{\text{ML}}(Y^0)$. This generates 1740 augmented initial kernel morphologies Y^0_{aug} that mimic kernels deposited with laser energy $E_{laser} = 26\,\text{mJ}$ at locations $3\,\text{mm} \le x_{aug} < 30$ mm and $r_{aug} < 14$ mm. These locations were chosen to avoid potential prediction issues that can arise from the interaction between the kernel and the domain boundaries, which is substantiated since laser deposition does not typically occur near the combustor walls. Effects from choosing different initial ignition kernels among the 153 tests are discussed in the supplementary material. The other two datasets are created to evaluate SDE-ML performance when tested with vastly OOD input conditions (to provide insights on the limitations of the present SDE-ML model) by (i) reducing the LES velocity in the augmented kernel dataset by two-fold and (ii) increasing the initial kernel segment area in the augmented kernel dataset by four-fold, respectively.

Fig. 6a compares ignition probability maps predicted with real data against experimental measurements in the 21 laser deposition locations. This spatially sparse ignition probability map is shown on top of ensemble-averaged experimental Schlieren measurements. Cyan circles are used to highlight the ignition boundary measured from the experiments, which is defined here as the radial distance before ignition probability decreases below unity. At the ignition boundaries, the SDE-ML model predicts ignition probabilities of 0.72, 0.84, and 0.8 at axial locations $x = d_{f,o}$, $2d_{f,o}$, and $3d_{f,o}$, respectively. Of the three ignition probabilities, only $P_{ig}(x = d_{f,o})$ exceeds the measured uncertainty of ±0.2. This under-prediction of ignition probability is likely caused by the under-prediction of mean radial velocity magnitude observed in Fig. 4, which reduces the likelihood of the hot kernel interacting with the cold reactants near the central jet.

The ignition probabilities at three OOD laser deposition locations that lead to direct ($x = d_{f,o}$, $r = 0$), indirect ($x = 2d_{f,o}$, $r = 1.1d_{f,o}$), and failed ($x = d_{f,o}$, $r = 1.3d_{f,o}$) ignition are also evaluated. At these locations, the SDE-ML model correctly predicts unity and zero probability for the direct and failed ignition cases, respectively. For the indirect ignition case, a predicted ignition probability of $P_{ig}^{pred} = 0.5$ is observed, which is within the uncertainty range of $P_{ig}^{exp} = 0.6 \pm 0.1$. Thus, the SDE-ML model can predict the ignition probability behavior reasonably well at slightly OOD deposition locations.

The spatially resolved ignition probability map predicted with the augmented kernel dataset is shown in Fig. 6b. Here, we compare SDE-ML predictions with and without the stochastic component ($dW = 0$). Both approaches are seen to generate a spatially coherent ignition probability map, even though the training data does not extend beyond sparse laser deposition locations. However, we note that without the stochastic component representing the effects of turbulent diffusion [35], the ML model can only generate deterministic predictions, resulting in binary probability map on the right side of Fig. 6b.

Fig. 6c compares two fully resolved ignition probability maps predicted with the two vastly OOD datasets. On the left, when reducing the velocity inputs by two-fold, a smaller region with unity ignition probability is observed above $x > d_{f,o}$, which is a result of lower mean radial velocities. Below $x < d_{f,o}$, the model predicts only failed ignition when $dW = 0$ indicating that the predicted ignition kernel loses hot plasma significantly before being transported to a reactive fuel/oxidizer mixture. On the right, when increasing the initial ignition kernel segment area by four-fold, narrower regions of zero and unity ignition probability are observed, when $r > 1.4d_{f,o}$ and $r \le 0.6d_{f,o}$, respectively. The narrower regions on the zero probability are likely due to increased interaction of the kernel area with the underlying velocity flowfield (as described by Eq. (3b)), which subjects the ignition kernel to more significant advection towards the reactive fuel/oxidizer core that in greater proportions of successful ignition in those laser deposition locations. In the training data, large ignition kernels near the jet shear layer are prevalent in the failed ignition cases, where the kernels dissipate and cool over time (as seen beyond $\tau = 244\,\mu$s in Fig. 3). As such, it is likely that the narrower regions on the unity probability arise from the relationship between large ignition kernels with lower ignition temperature that has been learned by the SDE-ML



(a) Predictions with real data. (b) Predictions with augmented kernel segment inputs. (c) Predictions with vastly out-of-distribution (OOD) inputs.

Fig. 6. SDE-ML predictions of ignition probability P_{ig} maps. Ensemble-averaged experimental Schlieren measurements (along with measured ignition boundaries in cyan) are shown in (a), while SDE-ML predictions without the stochastic component ($dW = 0$) are shown in (b) and (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

model, which results in greater proportions of failed ignition in those laser deposition locations.

Whether these predictions with OOD augmented datasets are physically plausible is uncertain within this work, given the absence of measurements under these velocity conditions. We note that all ML-based models cannot predict reliably with vastly OOD inputs [19]. ML-tailored uncertainty quantification techniques could be employed with physics-based approaches to partially address this issue by providing further interpretability on model confidence in the presence of unfamiliar inputs, and can potentially be employed for guiding additional experimental measurements [39]. However, the integration of these methods are an active scientific pursuit that are currently beyond the scope of this work.

5. Conclusions

This work introduces a reduced-order physics-embedded SDE-ML modeling framework that employs both experimental and simulation data for statistically characterizing ignition behavior within an experimental gaseous CH_4/O_2 model rocket combustor configuration.

This SDE-ML framework is shown to qualitatively capture the evolution of ignition kernels in laser ignition. In particular, the SDE-ML model is able to capture the influence of asymmetric kernel morphology on stochastic indirect ignition. In addition, ensemble ignition time predictions showed reasonable statistical agreement with ensemble experimental measurements. Ensemble predictions of mean kernel trajectory and ignition probability demonstrate qualitative agreement with the ensemble measurements. Quantitative differences arise from limitations in the treatment of kernel velocity (potential discrepancies in the LES input and 2D treatment of this model). This results in errors in ignition probability at ignition boundary and kernel position of up to 0.28 and $0.4d_{f,o}$, respectively.

To demonstrate that this SDE-ML framework can develop reasonable predictions when tested on unseen situations, we showed that the model reasonably captures ignition probabilities associated with direct, indirect, and failed ignition modes at slightly OOD laser deposition locations. The versatility of the SDE-ML model is further evaluated with augmented datasets that mimic inputs for generating a spatially resolved ignition probability map under different velocity conditions. The SDE-ML model can generate a spatially coherent ignition probability map, with only a training dataset consisting of ignition kernels deposited in sparse laser deposition locations. However, we note that providing the model with vastly OOD inputs potentially results in spurious phenomena. Extrapolation with OOD inputs is an active research topic within all ML techniques that may be addressed with future investigations involving physics-informed ML and uncertainty quantification [19]. Other potential directions for future work involve extending the present framework towards 3D approaches, and further inclusion of multi-fidelity ML training [32].

These results demonstrate the potential and limitations in combining stochastic modeling approaches with data-driven methods for leveraging sparse experimental measurements and non-reacting simulation data in capturing different ignition modes within a complex configuration with lower costs than reacting high-fidelity simulations.

Novelty and Significance Statement

The novelty of this research is the introduction of a hybrid stochastic physics-embedded deep-learning modeling framework for statistically characterizing stochastic ignition phenomena. We show that the presence of different ignition modes within laser ignition phenomena, which arises from ignition kernel morphology, can be captured by this approach in a cost-efficient manner with only sparse experimental measurements and non-reacting simulations. This research is significant for statistical characterization of ignition within combustion systems.

CRediT authorship contribution statement

Wai Tong Chung: Designed/performed research, Analyzed data, Writing – original draft. **Charlélie Laurent:** Designed research, Writing – original draft. **Donatella Passiatore:** Generated simulation data, Writing – original draft. **Matthias Ihme:** Designed research, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge financial support and computing resources from the U.S. Department of Energy, National Nuclear Security Administration under award No. DE-NA0003968. We thank the members of the PSAAP-III Center for numerous fruitful discussions, and especially Ryan Strelau and Carson Slabaugh for providing their experimental data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.proci.2024.105304. Links to code, model weights, and processed data are shared at https://github.com/IhmeGroup/sde-ml-ign. Supplementary results on mesh converges and model sensitivity to the choice of initial ignition kernel morphology is also provided.

References

[1] S.A. O'Briant, S.B. Gupta, S.S. Vasu, Review: Laser ignition for aerospace propulsion, Propuls. Power Res. 5 (1) (2016) 1–21.

[2] E. Mastorakos, Forced ignition of turbulent spray flames, Proc. Combust. Inst. 36 (2) (2017) 2367–2383.

[3] S.F. Ahmed, R. Balachandran, E. Mastorakos, Measurements of ignition probability in turbulent non-premixed counterflow flames, Proc. Combust. Inst. 31 (1) (2007) 1507–1513.

[4] M. Cordier, A. Vandel, G. Cabot, B. Renou, A.M. Boukhalfa, Laser-induced spark ignition of premixed confined swirled flames, Combust. Sci. Tech. 185 (3) (2013) 379–407.

[5] E. Bach, J. Kariuki, J.R. Dawson, E. Mastorakos, H.-J. Bauer, Spark ignition of single bluff-body premixed flames and annular combustors, 2013, AIAA Paper 2013-1182.

[6] K. Prieur, D. Durox, J. Beaunier, T. Schuller, S. Candel, Ignition dynamics in an annular combustor for liquid spray and premixed gaseous injection, Proc. Combust. Inst. 36 (3) (2017) 3717–3724.

[7] B. Sforzo, H. Dao, S. Wei, J. Seitzman, Liquid fuel composition effects on forced, nonpremixed ignition, J. Eng. Gas Turbine Power 139 (3) (2017) 031509.

[8] R. Strelau, M. Frederick, W.C. Senior, R. Gejji, C.D. Slabaugh, Modes of laser spark ignition of a model rocket combustor, 2023, AIAA Paper 2023-2377.

[9] G. Lacaze, B. Cuenot, T. Poinsot, M. Oschwald, Large eddy simulation of laser ignition and compressible reacting flow in a rocket-like configuration, Combust. Flame 156 (6) (2009) 1166–1180.

[10] O. Gurliat, V. Schmidt, O. Haidn, M. Oschwald, Ignition of cryogenic H_2/LOX sprays, Aerosp. Sci. Technol. 7 (7) (2003) 517–531.

[11] J.M. Wang, D.A. Buchta, J.B. Freund, Hydrodynamic ejection caused by laser-induced optical breakdown, J. Fluid Mech. 888 (2020) A16.

[12] T. Jaravel, J. Labahn, B. Sforzo, J. Seitzman, M. Ihme, Numerical study of the ignition behavior of a post-discharge kernel in a turbulent stratified crossflow, Proc. Combust. Inst. 37 (4) (2019) 5065–5072.

[13] K. Maeda, T. Teixeira, J.M. Wang, J. Hokanson, C. Melone, M. Di Renzo, S. Jones, J. Urzay, G. Iaccarino, An integrated heterogeneous computing framework for ensemble simulations of laser-induced ignition, 2023, AIAA Paper 2023-3597.

[14] L. Esclapez, E. Riber, B. Cuenot, Ignition probability of a partially premixed burner using LES, Proc. Combust. Inst. 35 (3) (2015) 3133–3141.

[15] Y. Tang, M. Hassanaly, V. Raman, B.A. Sforzo, J. Seitzman, Probabilistic modeling of forced ignition of alternative jet fuels, Proc. Combust. Inst. 38 (2) (2021) 2589–2596.

[16] E.S. Richardson, E. Mastorakos, Numerical investigation of forced ignition in laminar counterflow non-premixed methane-air flames, Combust. Sci. Tech. 179 (1–2) (2007) 21–37.

[17] A. Neophytou, E. Richardson, E. Mastorakos, Spark ignition of turbulent recirculating non-premixed gas and spray flames: A model for predicting ignition probability, Combust. Flame 159 (4) (2012) 1503–1522.

[18] L. Esclapez, F. Collin-Bastiani, E. Riber, B. Cuenot, A statistical model to predict ignition probability, Combust. Flame 225 (2021) 180–195.

[19] M. Ihme, W.T. Chung, A.A. Mishra, Combustion machine learning: Principles, progress and prospects, Prog. Energy Combust. Sci. 91 (2022) 101010.

[20] T. Ding, T. Readshaw, S. Rigopoulos, W. Jones, Machine learning tabulation of thermochemistry in turbulent combustion: An approach based on hybrid flamelet/random data and multiple multilayer perceptrons, Combust. Flame 231 (2021) 111493.

[21] W.T. Chung, A.A. Mishra, M. Ihme, Interpretable data-driven methods for subgrid-scale closure in LES for transcritical LOX/GCH_4 combustion, Combust. Flame 239 (2022) 111758.

[22] H. Ma, Y.-X. Zhang, O.J. Haidn, N. Thuerey, X.-Y. Hu, Supervised learning mixing characteristics of film cooling in a rocket combustor using convolutional neural networks, Acta Astronaut. 175 (2020) 11–18.

[23] W.T. Chung, A.A. Mishra, N. Perakis, M. Ihme, Data-assisted combustion simulations with dynamic submodel assignment using random forests, Combust. Flame 227 (2021) 172–185.

[24] B. Sforzo, J. Seitzman, Modeling ignition probability for stratified flows, J. Propuls. Power 33 (5) (2017) 1294–1304.

[25] P.P. Popov, D.A. Buchta, M.J. Anderson, L. Massa, J. Capecelatro, D.J. Bodony, J.B. Freund, Machine learning-assisted early ignition prediction in a complex flow, Combust. Flame 206 (2019) 451–466.

[26] M. Di Renzo, L. Fu, J. Urzay, HTR solver: An open-source exascale-oriented task-based multi-GPU high-order code for hypersonic aerothermodynamics, Comput. Phys. Comm. 255 (2020) 107262.

[27] L. Fu, X.Y. Hu, N.A. Adams, A family of high-order targeted ENO schemes for compressible-fluid simulations, J. Comput. Phys. 305 (2016) 333–359.

[28] S. Gottlieb, C.-W. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. Soc. Ind. Appl. Math. 43 (1) (2001) 89–112.

[29] J. Smagorinsky, General circulation experiments with the primitive equations: I. The basic experiment, Mon. Weather Rev. 91 (3) (1963) 99–164.

[30] A. Ern, V. Giovangigli, Multicomponent Transport Algorithms, Springer Berlin, Heidelberg, Germany, 1994.

[31] C. Sun, A. Shrivastava, S. Singh, A. Gupta, Revisiting unreasonable effectiveness of data in deep learning era, in: Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 843–852.

[32] S. De, A. Doostan, Neural network training using l1-regularization and bi-fidelity data, J. Comput. Phys. 458 (2022) 111010.

[33] J. Yang, M. Gao, Z. Li, S. Gao, F. Wang, F. Zheng, Track anything: Segment anything meets videos, 2023, arXiv preprint 2304.11968.

[34] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, Proc. Med. Image Comput. Comput.-Assist. Interv. (2015) 234–241.

[35] S.B. Pope, Turbulent Flows, Cambridge University Press, Cambridge, U.K., 2000, pp. 463–557.

[36] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, in: Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 1026–1034.

[37] W. Falcon, The PyTorch Lightning team, PyTorch Lightning, 2022.

[38] P. Sharma, W.T. Chung, B. Akoush, M. Ihme, A review of physics-informed machine learning in fluid mechanics, Energies 16 (5) (2023) 2343.

[39] E. Saetta, R. Tognaccini, G. Iaccarino, Uncertainty quantification in autoencoders predictions: Applications in aerodynamics, J. Comput. Phys. 506 (2024) 112951.