Early experiences in transfer learning between quiescent and reacting combustion chambers

By T. Zahtila, E. Saetta, N. Kozak, D. Rossinelli, G. Iaccarino AND G. Geraci†

We present a two-stage transfer learning framework for modeling spatiotemporal fields in highly dynamic, turbulent reacting flow problems. Of particular interest and challenge is the multiphysics nature of the problem consisting of a laser-induced spark with rich dynamics, high-speed coflowing fuel-oxidizer intake stream and turbulent combustion. We aim to accelerate training of a convolutional autoencoder (cAE) and specifically investigate the scenario where a simplified data set retains a subset of the physics in the target problem but some components are missing entirely. Accordingly, a cAE is first trained on a simplification to the target problem of interest, that of a computationally inexpensive quiescent spark chamber: that is, laser-deposited energy kernel evolution in a chamber of a homogeneous O_2 concentration. Subsequently, the cAE is augmented to efficiently compress the full target problem: reacting rocket combustors via latent-space fusion. This approach significantly reduces training cost, enables latent feature re-use, and improves generalization across domains with differing physics such as ignition, turbulence and flame propagation.

1. Introduction

The essential components of the PSAAP-III rocket combustor, as studied in Passiatore *et al.* (2024), consist of a coflowing fuel-oxidizer jet into a combustion chamber, ignited by a laser-deposited spark, and subsequent evolution of a turbulent flame. The dynamics of the spark alone, particularly the initial thermal expansion phase, unfolds on a timescale much faster than the timescales of the concurrently operating fuel-oxidizer intake jet (Wang *et al.* 2020). The dynamics of the laser spark are crucial in migrating hot kernels of gas to regions of favorable chemical composition and ignition success (Brouzet *et al.* 2026). When the focus is restricted to spark kernel evolution alone, the computational domain can be significantly reduced (Brouzet *et al.* 2026). Therefore, in this study, we consider two data sets: (i) laser spark evolution in a quiescent, nonreacting chamber with homogeneous concentration of O_2, and (ii) laser spark evolution in a fully reacting rocket combustor, with the latter representing our target application. We emphasize that simulations of case (i) are cost effective, whereas the target application case (ii) requires costly computations.

Transfer learning (TL) is a deep learning approach in which a model trained for one task is retained as the starting point for a second related task, thereby leveraging reuse of network elements. For example, in an adjacent application of plume tracking, the first task involved learning a Gaussian dispersion model, which was then fine-tuned in the second task of a limited number of Navier-Stokes computations (Shi *et al.* 2023). TL provides a natural strategy to connect the two laser-induced spark scenarios by leveraging latent feature representations: first learning the evolving dynamics of a laser spark in a

† Sandia National Laboratories

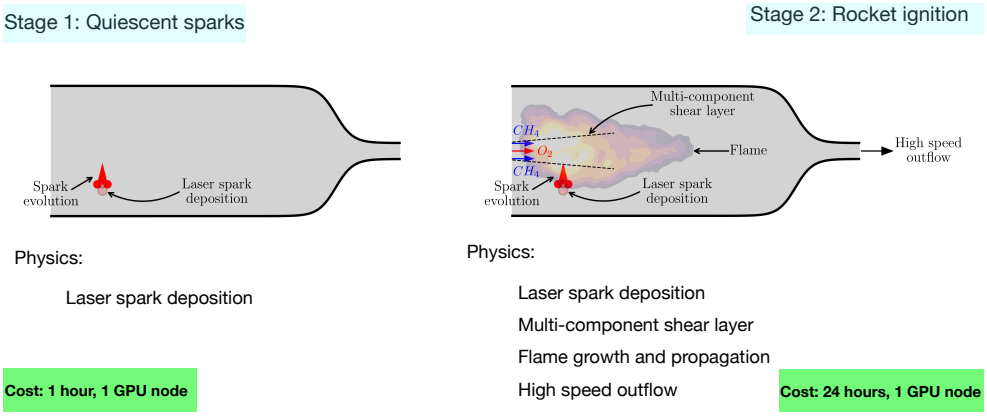

FIGURE 1: Comparison of the physics present in the two data sets: (i) laser spark evolution in a quiescent, nonreacting chamber, and (ii) laser spark evolution in a fully reacting rocket combustor.

quiescent chamber, then re-using and adapting the latent space to a useful representation for the reacting combustor scenario. Unlike conventional supervised learning (Li *et al.* 2020; Zahtila *et al.* 2024), which is tied to domain-specific mappings, TL allows re-use of learned encoders and decoders across related but distinct data sets, promising reduced training cost and improved generalization (Pan & Yang 2009; Weiss *et al.* 2016). The ultimate aim of the present work is an efficient construction of the input-output map via latent variables, where the subcomponent physics of the laser-induced spark is retained from the earlier problem. Presently, however, we only focus on the construction of a meaningful latent variable space that reflects the physical processes.

Recent work with autoencoders (AEs) and both linear and nonlinear multi-fidelity manifolds (Zeng *et al.* 2023; Saetta *et al.* 2024; Zanoni *et al.* 2024; Zeng *et al.* 2025) has been shown to effectively bridge variable-fidelity simulations while maintaining predictive accuracy. Multi-fidelity learning achieves this by jointly combining data sets of varying fidelity but of the same target physics (e.g., RANS, LES, DNS) into a hierarchy. In contrast, TL adapts knowledge from a source domain (e.g., nonreacting spark) to a target domain (e.g., reacting combustor). This paradigm is broadly applicable in fluid mechanics; for example, knowledge gained in flow past an isolated cylinder (Kravchenko & Moin 2000) naturally informs the more complex case of flow past confined cylinders (Lu *et al.* 2023*a*,*b*). The overall reliance on high-fidelity computational fluid dynamics (CFD) for trustworthy and accurate data remains a fundamental bottleneck in the field: A single reacting LES of the PSAAP combustor can demand on the order of $\mathcal{O}(10^5)$ core-hours (Zahtila *et al.* 2023). Traditional high-fidelity workflows therefore advance computations tied to specific and narrow scenarios and potentially for only one-time usage. TL offers a means to mitigate this inefficiency by creating transferable latent representations and re-use of data that took significant computing cost to generate.

Earlier PSAAP-III efforts explored linear approaches to multi-fidelity. For instance, interpolative decomposition (ID) methods developed in Cutforth *et al.* (2025) constructed low-rank subspaces to capture cluster membership for successful versus unsuccessful ignitions, giving rise to stochastic multimodal ID. These approaches are effective when the quantity of interest (QoI) admits a low-rank representation, as with scalar time series such as pressure traces. However, linear methods do not readily admit low-rank struc-

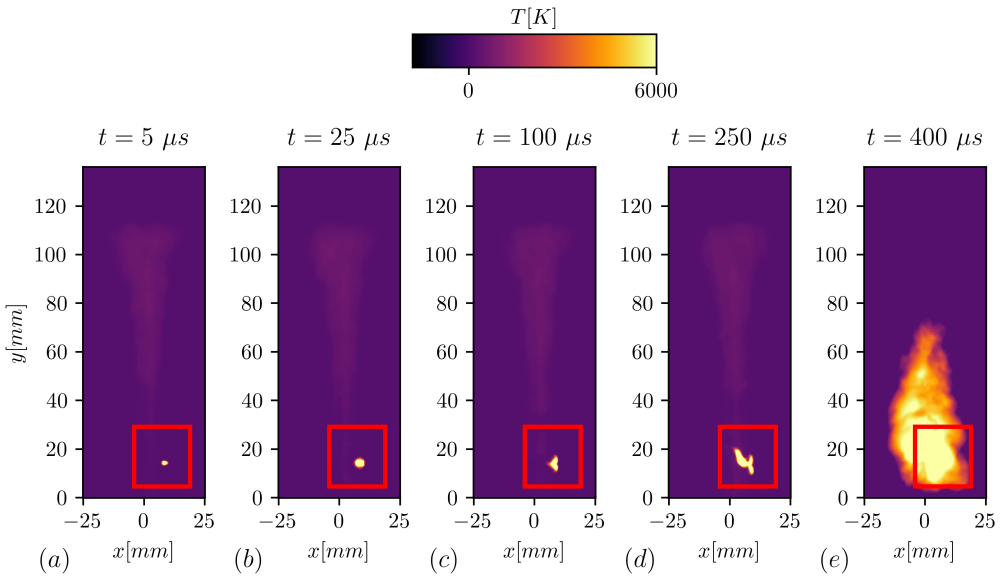

FIGURE 2: Evolution of the reacting spark kernel at representative times ($t = 5$–400 μs). Red rectangles denote the extracted subdomain used in subsequent analyses.

tures when applied to more complex quantities (Lee *et al.* 2025), and rank explosion is accentuated for spatiotemporal quantities such as the field quantities considered in the present work. Therefore, in general, linear methods are not a natural candidate for developing TL strategies.

The underlying goal of the present work is retention of learned physics compressed representation in a second more complex scenario. We develop a TL pipeline for combustor data sets that integrates nonlinear cAEs and seeks to leverage the common information among the disparate data sets. The computational cost and relevant physics present in each data set are shown in Figure 1. The key difficulty is that kernel evolution associated with laser spark deposition, for which a common physics exists across data sets, is only a subset of the physics that occur in the reacting combustor which includes high-speed flows, shocks, ignition events and flame. Isolating the laser spark kernel physics is therefore a key challenge. In the rest of this brief, Section 2 describes the PSAAP combustor data sets and preprocessing, Section 3 details the TL methodology, Section 4 presents results for transfer between reacting and nonreacting cases, and Section 5 concludes with future directions for multitask learning in high-fidelity combustion simulations.

2. Data sets and basis for transfer learning

The data sets considered in this study originate from large-eddy simulations (LESs) of the PSAAP-III combustor. Two scenarios are examined: (i) a quiescent laser spark evolution in a nonreacting chamber, and (ii) a reacting rocket combustor in which ignition and flame development occur. For each individual realization of the input uncertainty vector $\boldsymbol{\xi}$ [which is not discussed here but contains variabilities in laser spark operating conditions as well as aleatoric and model-form uncertainties; see Cutforth *et al.* (2025) for details] data sets are derived from identical initial spark deposition parameter sam-

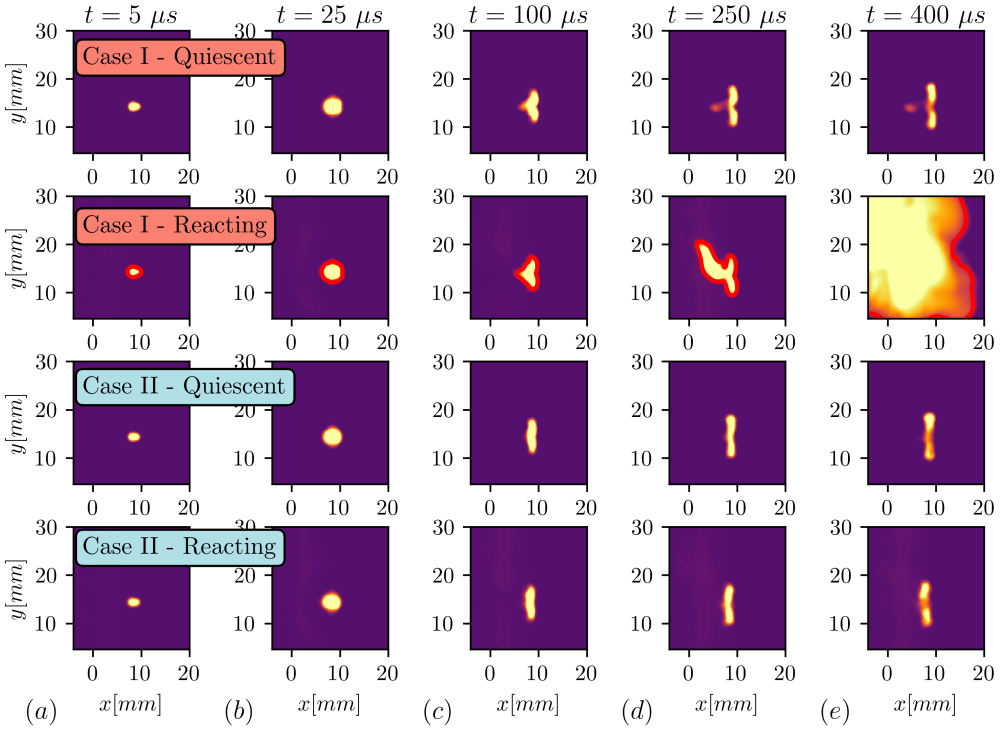

FIGURE 3: Comparison of kernel evolution between quiescent (first and third row) and reacting (second and fourth row) chambers. Each column corresponds to a temporal snapshot, illustrating the similarity of early kernel dynamics and the divergence that emerges once combustion begins. Threshold contour of the laser spark temperature field at $T = 1000$ K is shown in the second row.

pling, allowing for direct comparison of kernel morphology across domains. Each data set contains $N_c = 300$ realizations, with $N_t = 20$ temporal snapshots per simulation, yielding a spatiotemporal ensemble of images with size $\mathbf{I} \in \mathbb{R}^{592\times240}$ at each time step. The ensemble structure and sampling allows us to assess both variability across realizations and systematic differences between the quiescent and reacting chamber scenarios.

Figure 2 shows the temporal evolution in the target problem for a representative case of the reacting rocket combustor. At the earliest stage ($t = 5\ \mu$s), the deposited energy appears as a compact, localized hotspot near the bottom of the chamber. By $t = 25$–$50\ \mu$s, the spark kernel expands radially and the introduced baroclinic torque drives its dynamical evolution. At the critical ignition event time ($t = 250\ \mu$s), the kernel grows in size as the characteristics of the growing flame appear, and finally at $t = 400\ \mu$s, a developed flame can be observed. The red rectangles highlight the cropped region used for further pre-processing analysis, which isolates the spark while removing unaffected portions of the chamber.

To reduce a high-dimensional CFD field into the utilized 2D image, the raw simulation outputs are converted to an integrated line-of-sight quantity. The primary development of this path-tracing procedure is detailed in Rossinelli *et al.* (2025) and is numerical emulation of the classical experimental Schlieren methodology, for which light is distorted

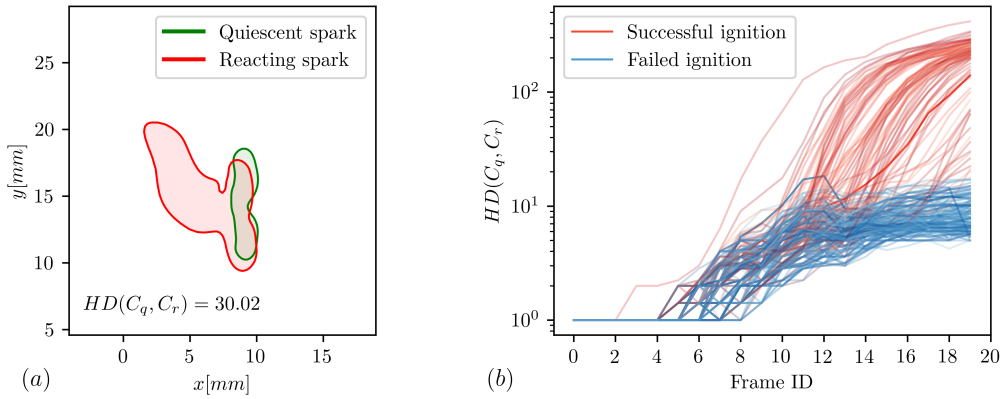

FIGURE 4: Quantitative comparison of kernel morphology using the Hausdorff distance. (a) Single instance contour overlay for quiescent (green) and reacting rocket combustor (red) sparks at $t = 250$ μs, with distance $HD(C_q, C_r)$ indicated. (b) Trajectories of Hausdorff distance over time for all realizations (transparent curves) split by failed or successful ignition.

due to variation in the density field refractive medium (Strelau *et al.* 2024). In the present work, we instead employ a thermal emission modality, simulating the propagation of uncollimated light from the 3D volume to a finite-aperture sensor.

In the pre-processing stage, a rectangular subdomain was extracted, corresponding to the initial region of spark kernel formation. This region was chosen to increase sensitivity of the analysis to early kernel growth. Figure 3 compares the temporal progression of spark kernels in the quiescent (first and third row) and reacting (second and fourth row) chambers for two representative cases: one exhibiting successful ignition and one resulting in failure. In the early frames ($t = 5$–50 μs), the two cases are nearly indistinguishable, both exhibiting symmetric radial growth governed primarily by diffusion and pressure relaxation (Wang *et al.* 2020). When the ignition trial is successful the divergence becomes obvious and pronounced at later times ($t = 250$–400 μs). The quiescent kernels remain smooth and symmetric, while the reacting counterparts develop large-scale flame structures associated with turbulent combustion. However, when the ignition trial is a failure, even at later times, the images remain highly correlated. Evidently, a variance exists within the data for the potential for transferable physics of early kernel growth, conditional on the bifurcating outcomes of ignition success and failure. As a proxy for similarity between individual realizations at a given time of the quiescent and reacting laser sparks, we quantify morphological differences between kernels through the Hausdorff distance, which has been previously applied in the analysis of interfacial multiphase flows (Cutforth & Mirjalili 2025), a worst-case measure analogous to the L^∞ norm. For two contours Γ_X and Γ_Y, it is defined as

$$HD(\Gamma_X, \Gamma_Y) = \max\left\{ \sup_{x\in\Gamma_X} \inf_{y\in\Gamma_Y} \|x-y\|, \sup_{y\in\Gamma_Y} \inf_{x\in\Gamma_X} \|x-y\| \right\}, \tag{2.1}$$

which measures the maximum distance of a point in one set to the nearest point in the other. As this metric requires an input of closed contours rather than an image, we threshold the laser spark temperature field at $T = 1000$ K; the metric was not

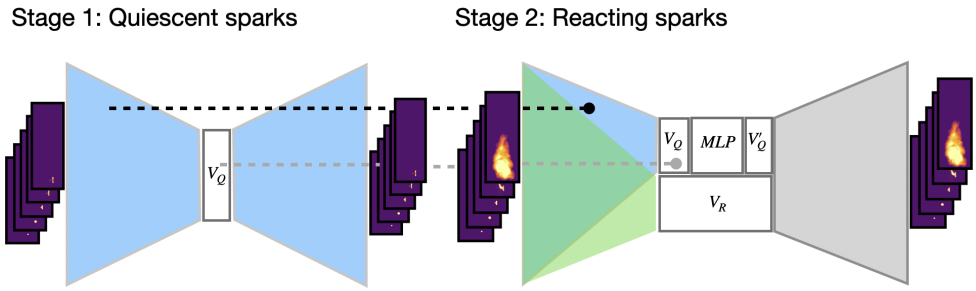

FIGURE 5: Schematic of the proposed two-stage architecture. A cAE is first trained on quiescent spark image sequences, yielding 4D latent vectors V_Q. A nonlinear transformation $V'_Q = \mathrm{MLP}(V_Q)$ adapts this embedding for re-use in the reacting decoder, which fuses it with an 8D vector V_R produced by a second encoder.

sensitive to choice of threshold, and example contours are shown in Figure 3. We note that the choice of Hausdorff distance is a preliminary choice and in future work will investigate alternatives. Figure 4(a) provides an illustrative realization of the Hausdorff distance overlaying quiescent and reacting kernel contours Γ_X and Γ_Y at $t = 250$ μs, which represents the bifurcating instance of ignition in the latter data. A clear degree of mismatch between the contours is visible, with the reacting kernel evidently growing with the flame region. Figure 4(b) presents ensemble trajectories of the Hausdorff distance $HD(\Gamma_X, \Gamma_Y)$ over time. The temporal growth of $HD(C_q, C_r)$ captures the increasing divergence between quiescent and reacting kernel morphology. At early times, the low typical value of $HD(\Gamma_X, \Gamma_Y)$ indicates strong similarity between fields. At later times, two distinct ensemble branches emerge: non-igniting cases with $HD(\Gamma_X, \Gamma_Y) \approx \mathcal{O}(10^1)$ and igniting cases with $HD(\Gamma_X, \Gamma_Y) \approx \mathcal{O}(10^2)$.

3. Proposed architecture

Our architecture comprises two sequential stages: a cAE trained on quiescent sparks constituting stage-1, followed by a stage-2 second cAE trained on the target reacting rocket combustor data set, which retains and re-uses the encoder and latent space structure from stage-1 (see Figure 5).

Stage 1: Quiescent spark compression. We first train a cAE on infrared (IR) image sequences $x_Q(t_j)$ obtained from the quiescent sparks nonreacting LES computations. The encoder e_Q maps each frame to a latent vector $V_Q(t_j) \in \mathbb{R}^4$ in

$$V_Q(t_j) = e_Q(x_Q(t_j)). \tag{3.1}$$

The decoder d_Q reconstructs the input as $\hat{x}_Q(t_j) = d_Q(V_Q(t_j))$, minimizing the mean squared error across time steps. The resulting latent trajectories $V_Q(t)$ compactly encode the early spark dynamics governed by diffusion, thermal expansion and baroclinic torque. Given the exploratory nature of the present work, the latent space dimension N_l is chosen based on preceding experience (Zahtila *et al.* 2024) for this problem data, and sensitivity tests will be conducted at a later stage. Briefly, the quiescent encoder consists of a set of convolutional layers with max pooling, residual blocks and two fully connected layers.

The decoder mirrors this structure in reverse. Training is performed with batch size $N_b = 200$ and learning rate $\mathrm{lr} = 10^{-3}$.

Stage 2: Reacting spark compression. For reacting sequences $x_R(t_j)$, we train a second autoencoder while retaining from stage-1 e_Q and keeping it fixed (frozen). A second encoder e_R maps $x_R(t_j)$ to an eight-dimensional latent vector $V_R(t_j) \in \mathbb{R}^8$. Simultaneously, the quiescent latent vector $V_Q(t_j)$ is passed through a fully connected multilayer perceptron (MLP) to yield a transformed vector $V'_Q(t_j) \in \mathbb{R}^4$. These are concatenated and decoded by d_R to reconstruct the reacting image

$$\hat{x}_R(t_j) = d_R\left([V'_Q(t_j), V_R(t_j)]\right). \tag{3.2}$$

The reacting decoder thus receives a twelve-dimensional latent space vector, with four components from the quiescent kernel encoder and eight from the reacting encoder. The MLP serves as a latent-space filter: It retains structure from V_Q when this aids reconstruction of the evolving kernel (e.g., during early times) but learns to suppress it once the reacting physics (ignition, turbulence, flame anchoring) diverge from the quiescent case. This selective re-use improves generalization and enables the decoder to adaptively balance prior and task-specific information, thereby recalling the quiescent spark information when useful, and forgetting when the information is not useful. The total loss is required to separate out the individual physics contributions and is therefore defined as

$$\mathcal{L} = \mathcal{L}_{\mathrm{recon}} + \lambda_{\mathrm{id,eff}} L_{\mathrm{id}} + \lambda_{\mathrm{reg,eff}} L_{\mathrm{reg}}, \tag{3.3}$$

where the adaptive weights $\lambda_{\mathrm{id,eff}}$ and $\lambda_{\mathrm{reg,eff}}$ control the relative importance of quiescent spark expression and suppression as the kernel evolves. These weights are modulated by the normalized hot-region fraction area of each sample A_{norm}, which is a proxy for the Hausdorff distance, using sigmoidal functions

$$\lambda_{\mathrm{id,eff}} = \lambda_{\mathrm{id}}\left[1 - \sigma\big(\alpha_{\mathrm{id}}(A_{\mathrm{norm}} - m_{\mathrm{id}})\big)\right], \tag{3.4}$$

$$\lambda_{\mathrm{reg,eff}} = \lambda_{\mathrm{reg}}\,\sigma\big(\alpha_{\mathrm{reg}}(A_{\mathrm{norm}} - m_{\mathrm{reg}})\big), \tag{3.5}$$

where $\sigma(\cdot)$ denotes the logistic sigmoid, A_{norm} is the normalized hot-region area and α and m are steepness and midpoint parameters respectively and $\lambda_{id}, \lambda_{reg}$ are hyperparameters. Rather than using a Hausdorff distance, which is nondifferentiable and therefore unsuitable for gradient-based optimization, the hot-region fraction was employed as a differentiable proxy for front extent, computed as the fraction of bright image pixels within a specified intensity range. The adaptive weights as a function of A_{norm} are shown in Figure 6, with A_{norm} computed as a simple linear re-scaling between the 5^{th} and 90^{th} percentiles of the underlying hot-region distribution. Lastly,

$$\mathcal{L}_{\mathrm{recon}} = \|\hat{x} - x\|_2^2, \tag{3.6}$$

$$L_{\mathrm{id}} = \|V'_Q - V_Q\|_2^2, \tag{3.7}$$

$$L_{\mathrm{reg}} = \|V'_Q\|_2^2, \tag{3.8}$$

where x and $\hat{x}$ denote the input and reconstructed images, respectively. The adaptive formulation allows the model to gradually shift emphasis from latent identity preservation to regularization as reacting dynamics become dominant.

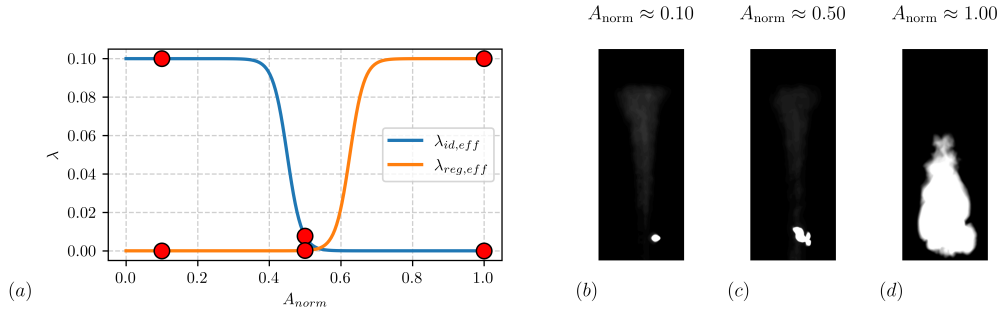

FIGURE 6: (a) Sigmoid-based loss weights $\lambda_{id,eff}$ and $\lambda_{reg,eff}$ vs normalized area A_{norm}. Masked hot-kernel region according to temperature level set $T = 1000$ K, with three representative marked A_{norm} values shown in (b–d).

Metric	$t = 5\,\mu$s	$t = 25\,\mu$s	$t = 100\,\mu$s	$t = 250\,\mu$s	$t = 400\,\mu$s
Std dev	3.99	6.54	6.62	12.32	33.37
Shannon entropy	1.039	1.084	1.095	1.182	1.820
FFT energy	15.94	42.77	43.82	151.7	1113

TABLE 1: Temporal evolution of standard deviation, Shannon entropy and spectral energy image metrics, based on the representative image sequence in Figure 2.

4. Results

We first present the stage-2 latent space vector to assess the ability of the proposed architecture to separate physical processes as embeddings in the latent space, shown in Figure 7. As it illustrates this separation more clearly, only the latent space corresponding to successfully igniting cases are under consideration, as it encompasses the full spectrum of physics to be captured. Also plotted for reference is the nominal ignition timescale, as it represents the onset of significant divergence between the stage-1 data and the target problem. As a reminder of the individual components, $V_Q = [V_1, ...V_4]^T$ and $V'_Q = MLP(V_Q)$ and $V_R = [V_5, ..., V_{12}]^T$; note that given the common encoder architecture, the value of V'_Q matching the value of V_Q physically indicates retention of the quiescent spark dynamics. We suggest that the stage-1 quiescent encoder latent space vector (V_Q) beyond this time grows to spurious values, most evident in Figure 7(a,c) as the trained stage-1 encoder fails. The mapped encoder MLP-mapped V'_Q therefore retains (V_Q) before this timescale and gradually decays to zero as ignition proceeds, which is satisfactory given the early stages of successful ignition do appear as the combination of a deposited spark and early flame growth [see Figure 2(d)] before the flame eventually overcomes the entire space.

The reacting components V_R, however, show some temporal evolution at early times, which is not satisfactory, and future work will be aimed at orthogonalizing the activation of these components with respect to V_Q. These components are intended to initially capture the addition of a coflowing jet, which is relatively uniform in the early times, and thereafter turbulent combustion. In Figure 7(i,l), we do see an increase in variance after the ignition timescale corresponding to the capture of flame dynamics. A better treatment for the early-time behavior of V_R is currently the focus of ongoing work.

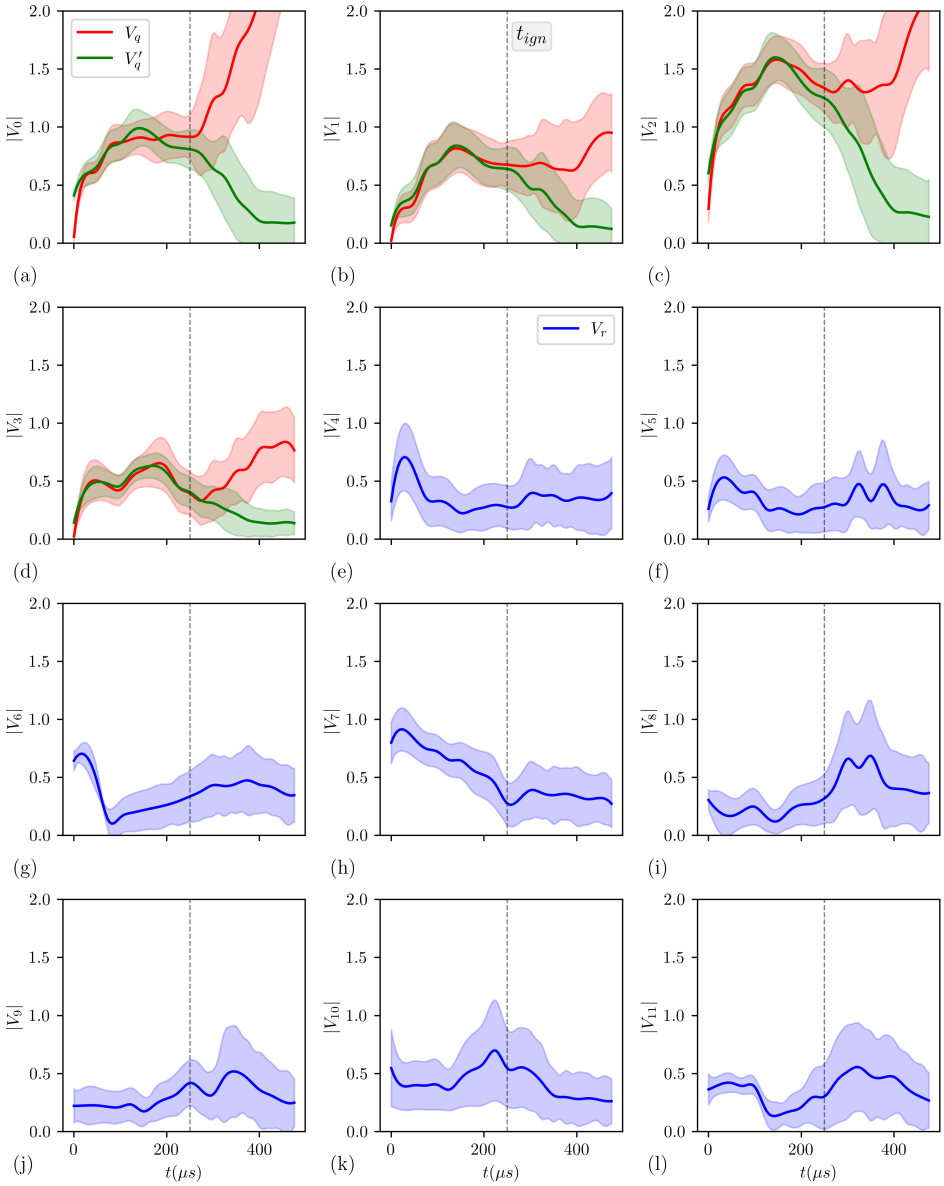

FIGURE 7: Evolution of latent-space vector magnitude for the quiescent encoder output V_Q and its transformed representation V_Q' in panels (a)–(d), and the reacting encoder output V_R in panels (e)–(l), for the successfully ignited reacting cases. Solid lines indicate mean value and shaded regions indicate ensemble variance across test cases. The vertical dashed lines mark the nominal ignition timescale t_{ign}.

In Figure 8 we compare the MLP-mapped V_Q' for the cases where ignition was successful or failed. In the latter case, the correlation with the stage-1 quiescent spark quantities remains high. The MLP-mapping satisfies this requirement at a qualitative level, but we

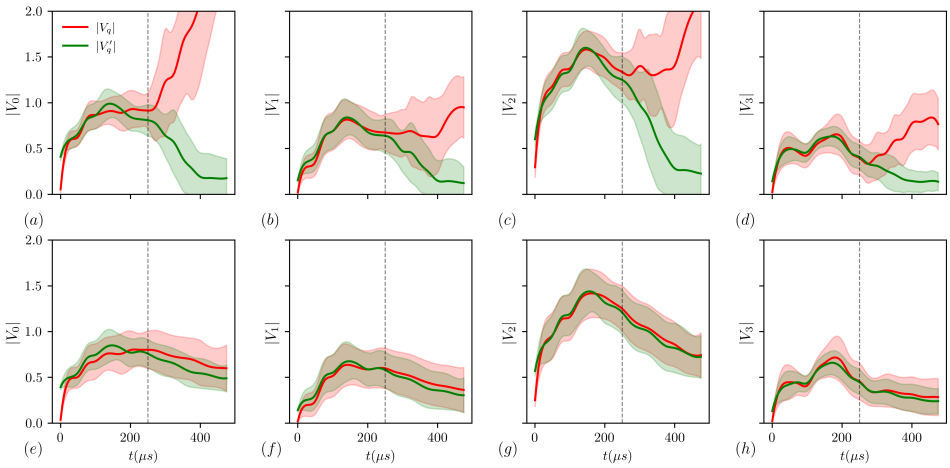

FIGURE 8: Comparison of the stage-1 encoder latent space V_Q and the MLP-mapped V_Q' for panels (a)–(d) ignition success and panels (e)–(f) failed ignition.

note the difficulty of the optimization landscape for the loss function presented in Section 3 and to this we attribute the small but not negligible disagreement.

With the major stated aim of the proposed architecture being the reduced amount of training data for the target problem of reacting sparks, we here document the validation performance on an unseen set of reacting sparks. We plot the validation loss as a function of the number of available cases shown to the architecture in Figure 9. The image complexity increases markedly over time for reacting cases, as there is considerably more spatial complexity once the combustion process starts. This was quantified in Table 1 which reports the standard deviation, Shannon entropy and total spectral energy, which in turn increases nearly two orders of magnitude across the $t = 5$–$400\,\mu$s sequence. Note the spectral energy E_{FFT} satisfies Parseval's identity, such that $E_{\mathrm{FFT}} = \sigma^2 \cdot N$, where σ^2 is the image variance and N the number of pixels. This is pertinent, as reconstruction errors increase at later times when spatial complexity and spectral energy are highest. For this reason, we investigate two data folds, (i) the first being to consider all available data, while (ii) the second is considering just the restricted initial 250 μs interval, that is, the first 10 images of each ignition trial sequence. The second data fold therefore largely ignores the combustion process sequence and focuses on adapting the modified laser ejecta dynamics that occur between the quiescent combustor and in the reacting counterpart.

We will first restrict discussion to the validation performance on the target problem where the full data sequence (i.e., no temporal reduction) is available for training, shown in Figure 9(a). For the standalone quiescent AE, it is unsurprising that performance only improved slowly beyond $N_c > 50$, as additional quiescent cases add little correlation between the quiescent sparks and reacting images. For the standalone reacting sparks AE, good performance is achieved with $N_c > 50$. The dual-stage AE did not yield a measurable performance here, attributable to the order of magnitude larger spectral content in the reacting images. However, when we consider the second fold, which is the restricted initial 250 μs interval, a substantial performance gain was achieved by the dual-stage AE compared to the standalone high-fidelity AE when the number of available reacting

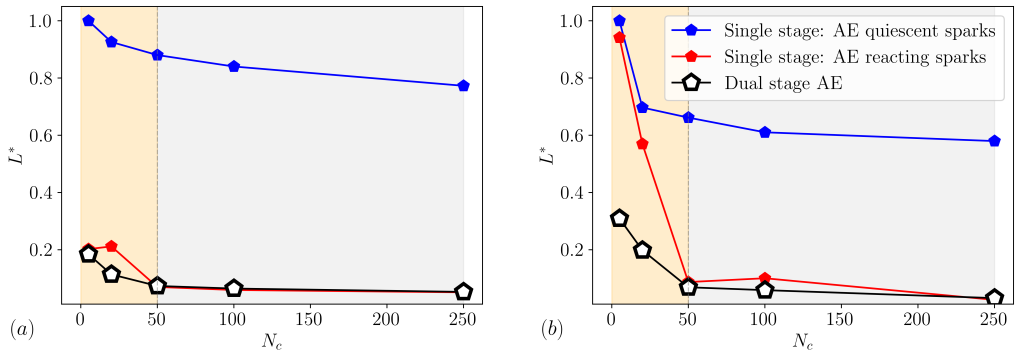

FIGURE 9: Validation loss for the reacting spark data set, comparing an AE trained on the standalone quiescent sparks, AE trained on the standalone reacting sparks and the current proposed dual-stage AE architecture, where N_c denotes the number of reacting spark cases available to the training (except in the quiescent case). (a) Shows the full ignition sequence, and (b) the restricted initial 250 μs interval that is, the first 10 images of each ignition trial sequence.

rocket combustor trials is less than $N_c \leq 50$. While not yet performed, it is important to contextualize the magnitude of this loss by inspecting the dual-stage reconstruction to appraise quality for constituent features in the image. This demonstrates the potential benefits for dual-stage training for the present class of problems, where the first stage problem contains a subset of physics in the full problem and is a fairer comparison of the ability of learned physics to be transferred across domains.

5. Future outlook

A key component that has been overlooked in the present work is transition of training from stage-1 to stage-2. Currently, this is a discontinuous jump and certainly does not reflect best practices in transfer learning. Whereas most transfer-learning (TL) methods fine-tune a network pretrained on another task, our approach extends the architecture sequentially to the more complex problem and selectively recalls physics when beneficial. The training of neural networks and the theory of their training is the subject of seminal work on information bottlenecks (Shwartz-Ziv & Tishby 2017). Broadly, in that work, the characteristics in the mutual information plane across layers of the architecture are established to ensure that sufficient data and training is present for compression and that irrelevant information is discarded, without overcompression where generalization is lost. In future work, we will investigate transitions from stage-1 to stage-2 that conserve a continual compression of information without sharp losses as the data stream is changed. A further connection with the machine learning literature that requires investigation is the analogy between the present work and the factual-recall paradigm (Geva *et al.* 2021) where the reuse of the latent space is analogous to memory retrieval.

Acknowledgments

This investigation was funded by the U.S. Department of Energy's National Nuclear Security Administration (NNSA) via the Stanford PSAAP-III Center for prediction of

laser-induced ignition of rocket engines (Grant Number DE-NA0003968). This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

REFERENCES

Brouzet, D., Rossinelli, D., Zahtila, T., Voci, A., Strelau, R., Warner, A., Gejji, R. M., Slabaugh, C. D. & Iaccarino, G. 2026 Thermal expansion-driven laser ignition in a gas subscale rocket combustor. *Combust. Flame* **284**.

Cutforth, M., Fan, T., Zahtila, T., Doostan, A. & Darve, E. 2025 Bi-fidelity interpolative decomposition for multimodal data. arXiv:2509.12243 [math.NA].

Cutforth, M. & Mirjalili, S. 2025 Convolutional autoencoders for the reconstruction of three-dimensional interfacial multiphase flows. arXiv:2508.04084 [cs.CE].

Geva, M., Schuster, R., Berant, J. & Levy, O. 2021 Transformer feed-forward layers are key-value memories. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 5484–5495. Punta Cana, Dominican Republic.

Kravchenko, A. G. & Moin, P. 2000 Numerical studies of flow over a circular cylinder at $Re_D = 3900$. *Phys. Fluids* **12**, 403–417.

Lee, J., Chan, L., Zahtila, T., Lu, W., Iaccarino, G. & Ooi, A. 2025 Surrogate models for multiregime flow problems. *Phys. Rev. Fluids* **10**, 024703.

Li, Z., Kovachki, N., Azizzadeh-Sheikh, K., Liu, B., Bhattacharya, K., Stuart, A. & Anandkumar, A. 2020 The Fourier neural operator for parametric partial differential equations. arXiv:2010.08895 [cs.LG].

Lu, W., Aljubaili, D., Zahtila, T., Chan, L. & Ooi, A. 2023*a* Asymmetric wakes in flows past circular cylinders confined in channels. *J. Fluid Mech.* **958**, A8.

Lu, W., Nguyen, Q. D., Chan, L., Lei, C. & Ooi, A. 2023*b* Flows past cylinders confined within ducts: effects of the duct width. *Int. J. Heat Fluid Flow* **104**, 109208.

Pan, S. J. & Yang, Q. 2009 A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359.

Passiatore, D., Wang, J. M., Rossinelli, D., Di Renzo, M. & Iaccarino, G. 2024 Computational study of laser-induced modes of ignition in a coflow combustor. *Flow Turbul. Combust.* **113**, 1055–1079.

Rossinelli, D., Li, G., Voci, A., Marshall, L., Hu, J., Brouzet, D., Fan, T., Williams, C., Grace, O., Khanwhale, M. *et al.* 2025 High-fidelity modeling of Schlieren and shadowgraph imaging. *Annual Research Briefs,* Center for Turbulence Research, Stanford University, pp. 89-104.

Saetta, E., Tognaccini, R. & Iaccarino, G. 2024 Autoencoders with sequential training and adaptive resolution. SSRN 4898782.

Shi, J., Xie, W., Li, J., Zhang, X., Huang, X., Usmani, A. S., Khan, F. & Chen, G. 2023 Real-time plume tracking using a transfer learning approach. *Comput. Chem. Eng.* **172**, 108172.

Shwartz-Ziv, R. & Tishby, N. 2017 Opening the black box of deep neural networks via information. arXiv:1703.00810 [cs.LG].

Strelau, R. M., Frederick, M. D., Winter, T. R., Senior, W. C. B., Gejji, R. M. & Slabaugh, C. D. 2024 Laser-induced spark ignition of a gaseous methane–oxygen model rocket combustor. *Combust. Flame* **265**, 113463.

Wang, J. M., Buchta, D. A. & Freund, J. B. 2020 Hydrodynamic ejection caused by laser-induced optical breakdown. *J. Fluid Mech.* **888**, A16.

Weiss, K., Khoshgoftaar, T. M. & Wang, D. 2016 A survey of transfer learning. *J. Big Data* **3**, 9.

Zahtila, T., Passiatore, D. & Iaccarino, G. 2023 Progress on multifidelity simulations of a laser-ignited rocket combustor. *Annual Research Briefs,* Center for Turbulence Research, Stanford University, pp. 57–68.

Zahtila, T., Saetta, E., Brouzet, D., Cutforth, M., Rossinelli, D. & Iaccarino, G. 2024 Neural ordinary differential equation-based representations of unsteady laser-deposited energy kernels. *Annual Research Briefs,* Center for Turbulence Research, Stanford University, pp. 107–117.

Zanoni, A., Geraci, G., Salvador, M., Menon, K., Marsden, A. L. & Schiavazzi, D. E. 2024 Improved multifidelity Monte Carlo estimators based on normalizing flows and dimensionality reduction techniques. *Comput. Methods Appl. Mech. Eng.* **429**, 117119.

Zeng, X., Geraci, G., Eldred, M. S., Jakeman, J. D., Gorodetsky, A. A. & Ghanem, R. 2023 Multifidelity uncertainty quantification with models based on dissimilar parameters. *Comput. Methods Appl. Mech. Eng.* **415**, 116205.

Zeng, X., Geraci, G., Gorodetsky, A. A., Jakeman, J. D. & Ghanem, R. 2025 Boosting efficiency and reducing graph reliance: basis adaptation integration in Bayesian multi-fidelity networks. *Comput. Methods Appl. Mech. Eng.* **436**, 117657.