DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES S

doi:10.3934/dcdss.2021138

AN OUT-OF-DISTRIBUTION-AWARE AUTOENCODER MODEL FOR REDUCED CHEMICAL KINETICS

Pei Zhang, Siyan Liu, Dan Lu and Ramanan Sankaran

Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak Ridge, TN 37830, USA

Guannan Zhang*

Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN 37830, USA

Abstract. While detailed chemical kinetic models have been successful in representing rates of chemical reactions in continuum scale computational fluid dynamics (CFD) simulations, applying the models in simulations for engineering device conditions is computationally prohibitive. To reduce the cost, data-driven methods, e.g., autoencoders, have been used to construct reduced chemical kinetic models for CFD simulations. Despite their success, data-driven methods rely heavily on training data sets and can be unreliable when used in out-of-distribution (OOD) regions (i.e., when extrapolating outside of the training set). In this paper, we present an enhanced autoencoder model for combustion chemical kinetics with uncertainty quantification to enable the detection of model usage in OOD regions, and thereby creating an OOD-aware autoencoder model that contributes to more robust CFD simulations of reacting flows. We first demonstrate the effectiveness of the method in OOD detection in two well-known datasets, MNIST and Fashion-MNIST, in comparison with the deep ensemble method, and then present the OOD-aware autoencoder for reduced chemistry model in syngas combustion.

1. **Introduction.** Accelerating simulations has been a critical research topic motivated by the need for faster device designs and real-time control applications. For simulations of chemically reacting flows, which are crucial in power, propulsion and energy conversion systems, one primary computational bottleneck comes from the modeling of chemical reactions. To model the rates of chemical reactions, which occur at atomic and molecular scales, accurately for computational fluid dynamics (CFD) simulations at continuum scales, detailed kinetic models consisting of a large

2020 *Mathematics Subject Classification.* 37M05, 37N99, 35K57, 68T07.

Key words and phrases. Chemical kinetics, dimension reduction, neural network, uncertainty quantification, out of distribution.

* Corresponding author: Guannan Zhang.

number of species and elementary reactions are needed. The large size and the complexity involved greatly increase the computational cost due to the need to solve a large set of highly coupled differential equation system. The wide range of chemical time scales in a multicomponent system adds to numerical stiffness and further increases the computational cost. Therefore, incorporation of such detailed chemical kinetic models in CFD simulations presents a huge challenge for the computational power and time. Reduced chemical kinetic models can accelerate CFD predictions of reacting flows and hence enable faster designs. The method in this work can also be used to develop reduced order models of systems useful for real-time control in engineering applications.

Data-driven approaches, such as principal component analysis (PCA) [4,5,9,10,15–17,21,27,28] and deep neural networks (DNNs) [18,19,30,31], have been applied for reducing chemical kinetic models for combustion. The reduction approaches fall into two major categories. One is to identify the critical species and elementary reactions via sensitivity analysis [5,28] and eliminate the less important ones, under certain target conditions. These works are based on the assumption that only a subset of species and elementary reactions are needed to achieve the engineering level accuracy. Esposito and Chelliah [5] applied PCA with sensitivity of kinetic parameters for ethylene-air combustion under three conditions, i.e., ignition, flame propagation and extinction, and reduced the dimension from 111 species and 784 reactions to fewer than 40 species. In the other category, instead of eliminating species and reactions, the reduction is achieved by defining a smaller set of new compressed variables as a function of the original thermo-chemical variables, e.g., a truncated set of principal components (PCs) in PCA. The governing equations of the PCs are then transported as part of the CFD to solve the chemical state. This category of work is inspired by the observation that chemical states are mostly distributed along low dimensional manifolds (LDMs) in the high dimensional thermo-chemical state space. Various techniques, e.g., PCA and autoencoder (AE) neural networks, have been used to identify these manifolds. Sutherland and Parente [27] conducted PCA analysis for CO/H_2 combustion dataset from a high-fidelity three-dimensional (3-D) direct numerical simulation (DNS) of turbulent flames and identified a linear approximation of the LDMs with two PCs for the original dataset with dimension of 12. Mirgolbabaei et al. [19] and Zhang et al. [31] found nonlinear approximations of the LDMs by training AE NNs using datasets collected for H_2 auto-ignition in turbulent jet from one-dimensional turbulence (ODT) simulation, and for CO/H_2 combustion from a large ensemble of 0-D stirred reactors under various parametric conditions, respectively.

Although great success has been achieved in the data-driven work, where reduced models are derived from a training dataset consisting of thermo-chemical states collected under certain conditions, potential risks arise when applying the reduced models in new CFD predictions, especially when the combustion conditions are far away from (or in out-of-distribution (OOD) regions of) the training dataset. The solution accuracy can be significantly deteriorated. Even worse, one can get nonphysical solutions that violate physical laws, e.g., mass conservation, and false extinction/ignition prediction in combustion engine simulations, which eventually lead to poor design choices. Data-driven reduced chemistry models that can not only provide us fast and accurate predictions but also allow the interpretation of confidence and detection of OOD samples are needed. In this work, we present a

reduced chemical kinetic model with OOD-detection capability leveraging AE NN in conjunction with an uncertainty quantification (UQ) method.

Various approaches for UQ of NN models have been developed, including fully Bayesian NNs [14], assumption-based variational inference [6, 8], and empirical ensemble approaches [1, 11, 23]. Whilst being shown successful, these methods require either high computational demands, strong assumptions or large memory costs to store the ensemble of models. As an attractive alternate, prediction interval (PI) methods [20, 22, 24, 26] directly communicate uncertainty, by providing a lower and upper bound for a NN output such that the value of the prediction falls between the bounds for some target percentage (e.g., 95%) of the unseen data, and hence present more understandable information for decision making [22, 26]. Nix and Weigend [20] proposed a maximum likelihood framework for building PIs with two NNs, where one predicts the value and the other predicts the variance. This method assumes Gaussian distribution for model errors and may fail in capturing bounds for asymmetric distributions. The quality-driven (QD) approach [22] requires no distributional assumption by defining a sophisticated loss function, but it is unable to generate point estimates and has a fragile training process. Built on QD, the prediction intervals with specific value prediction (PIVEN) method in [26] adds the capability to calculate point estimates and the PI method in [24] further improves the training stability of QD by integrating a penalty function to the loss.

While showing promising results, the recently developed PI methods [22, 24, 26] rely on sophisticated loss functions to obtain a well-calibrated PI. This inadvertently introduces highly sensitive hyperparameters which require delicate tuning [24] in training process to achieve the desired performance and make these methods less practical and less robust when deployed. Recently, we developed the prediction interval based on three neural networks (PI3NN)—a novel method for calculating PIs [29]. Different from the other methods [22, 24, 26] that rely on sophisticated loss functions to obtain a well-calibrated PI, the PI3NN method uses the standard loss functions, e.g., mean squared error (MSE), for training, enabling simple and robust training. At the same time, it possesses the desirable properties as the state-of-the-art PI methods have—such as requiring no distributional assumption and producing tight PI bounds, and in addition, it is noninstrusive and shows the capability to capture domain shift and reasonably quantify larger uncertainty on OOD samples in benchmark regression tasks.

In this work, we combine the PI3NN method with AE NNs and present an enhanced AE reduced model, AE-PI3NN, for chemical kinetics with OOD-detection capability. In section 2, we present the problem setting with combustion background and 0-D stirred chemical reactor configurations. Section 3 summarizes the methodology details. We assess the AE-PI3NN performance in detecting OOD samples in comparison with a state-of-the-art deep ensemble (DE) method [11] and present the results in section 4, where 4.1 shows a case with two well-known image datasets: MNIST and Fashion-MNIST, and 4.2 presents the reduced modeling results of syngas combustion chemistry. Section 5 summaries the conclusions of the work.

2. **Problem description.**

Combustion of complex hydrocarbon fuels is a multi-scale multi-physics process. To simulate the evolution of chemical state and flow, we need to solve a set of governing equations consisting of mass continuity, Navier-Stokes momentum, species and energy equations. The governing equations are

supplied with constitutive relations that approximate atomic processes at continuum scales. For combustion of a mixture of ideal gases, a few commonly used constitutive relations with simplified assumptions, for governing equations of species and energy, are summarized below. For more details we refer the readers to [3].

- Species diffusion can be driven by various gradients of, e.g., species concentration, temperature and pressure. We only consider the diffusion due to species concentration, i.e., Fickian diffusion. The diffusion coefficients are approximated with a simplified mixture-averaged model [2].
- Heat flux generally consists of contributions from heat conduction, heat transport due to mass diffusion, thermal diffusion (also known as Dufour effect) and thermal radiation. In the work, we only consider contributions from heat conduction and mass diffusion.

Based on the above assumptions, the governing equations for thermo-chemical state variables, i.e., species mass fractions Y_k $(k = 1, \ldots, n_s)$ and total enthalpy h, are

$$\begin{aligned}
&\frac{\partial \rho Y_k}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} Y_k) = \nabla \cdot \left[\rho D_k \left(\nabla Y_k + Y_k \nabla \ln \overline{M}\right)\right] + M_k \dot{\omega}_k, \\
&\frac{\partial \rho h}{\partial t} + \nabla \cdot (\rho \boldsymbol{u} h) = \frac{Dp}{Dt} + \nabla \cdot \left(\lambda \nabla T + \sum_{k=1}^{n_s} h_k \rho D_k \left(\nabla Y_k + Y_k \nabla \ln \overline{M}\right) \right) + \boldsymbol{\tau} : \nabla \boldsymbol{u},
\end{aligned} \tag{1}$$

where n_s is the number of species, ρ is density, p is pressure, $\boldsymbol{u}$ is velocity vector, $\boldsymbol{\tau}$ is the stress tensor, λ is the thermal conductivity, D_k, M_k and $\dot{\omega}_k$ are the mixture-averaged diffusion coefficient, molecular weight and net molar production rate of the kth species, respectively, and $\overline{M} = \left(\sum_{k=1}^{n_s} Y_k / M_k\right)^{-1}$ is the mixture mean molecular weight of the n_s species. The rates of chemical reactions are computed using continuum kinetic models. For chemical kinetic models consisting of n_e elementary reactions, the net production rate of kth species is,

$$\dot{\omega}_k = \sum_{i=1}^{n_e} \nu_{k,i} \gamma_i,$$

where $\nu_{k,i}\gamma_i$ is the contribution from the ith elementary reaction with γ_i being the reaction rate and $\nu_{k,i}$ being the stoichiometric coefficient(zero if the kth species is not in the ith reaction). For an elementary reaction, e.g.,

$$\mathrm{H} + \mathrm{O}_2 \rightleftharpoons \mathrm{O} + \mathrm{OH},$$

the reaction rate is $\gamma_i = k_f[\mathrm{H}][\mathrm{O}_2] - k_r[O][OH]$, where $[\cdot]$ is species concentration. The forward reaction rate constant k_f in the modified Arrhenius form is $k_f = AT^n \exp(-E_a/RT)$ and the reverse reaction rate constant is calculated as $k_r = k_f / K_{\mathrm{eq}}$ with K_{eq} being the equilibrium constant. In the reaction rate constant, T is temperature and R is the universal gas constant. The parameters (A, n, E_a) are constants provided by chemical kinetic models.

Equations in (1) involving multiple physics, e.g., turbulence, mixing and chemical reactions, are challenging to solve, especially when detailed kinetic models are used to describe the reaction kinetics. Detailed kinetic models consisting of a large number of chemical species and elementary reactions are needed for accurately representing combustion kinetics, e.g., AramcoMech 2.0 [12] with $n_s = 493$ species and $n_e = 2716$ elementary reactions for C0–C4 combustion. The large size means a large set of equations to solve and inevitably strong stiffness. For example, the time

scales associated with chemical reactions span over several orders of magnitude, from picoseconds to seconds, for n-heptane combustion [13]. Solving such large stiff partial differential equation systems is computationally challenging if not prohibitive for CFD simulations in engineering device conditions. To reduce the cost, we develop a data-driven reduced chemistry model in this work.

We collect datasets of thermo-chemical state variables from a large ensemble of inexpensive 0-D stirred reactor, called perfectly stirred reactor(PSR) [25], calculations. PSR is a canonical 0-D configuration with inlets and outlets and has been widely used in studies of chemical kinetics. The fuel and oxidizer mixture stream enters, resides and reacts inside the reactor for certain time, represented by a time scale called residence time scale τ_r, before existing the reactor. By controlling the residence time scale τ_r, various thermo-chemical states inside the reactor can be obtained from equilibrium (infinite residence time) to unreacted (zero residence time) states.

Let $\boldsymbol{x} = [x_1, x_2, \ldots, x_{n_x}] = [Y_1, \ldots, Y_{n_s}, T \text{ or } h]$ being the thermo-chemical state vector, consisting of the mass fractions of n_s species and temperature/enthalpy, the governing equations of a PSR reactor are

$$\begin{aligned} \frac{dY_k}{dt} &= \frac{Y_{k,\text{in}} - Y_k}{\tau_r} + M_k \dot{\omega}_k / \rho, \\ \frac{dT}{dt} &= \frac{\sum_{k=1}^{n_s} (h_{k,\text{in}} - h_k) Y_{k,\text{in}} / c_p}{\tau_r} - \sum_{k=1}^{n_s} M_k \dot{\omega}_k h_k / (\rho c_p), \end{aligned} \tag{2}$$

where h_k is the specific enthalpy of the kth species, c_p is the mixture-averaged specific heat, variables with subscript "in" represent the inflow state, and $n_x = n_s + 1$ is the dimension of the thermo-chemical state space. A steady state solution of the PSR reactor can be obtained when the time derivative is approaching zero. For a combustible mixture inflow $\boldsymbol{x}_{\text{in}}$, a steady state solution equal to the equilibrium state of the inflow mixture is obtained when τ_r is infinite, while an extinguished/nonreacting state solution is obtained when the time scale τ_r is too small for the mixture to react. By varying the inflow state $\boldsymbol{x}_{\text{in}}$ and the residence time scale τ_r, we can populate a thermo-chemical database with different combustion sample states from the steady PSR reactor.

In a previous work [30], a database consisting of $N_{\text{samp}} = 1.63$ million samples, $\mathcal{D}_{\text{PSR}} = \{\boldsymbol{x}_i\}_{i=1}^{N_{\text{samp}}}$, has been collected for the training and testing of AE NN models. The same dataset is used here in this work. The dataset considers the syngas CO/H_2 combustion case in [7], where the chemical mechanism with $n_s = 11$ species is used to describe the syngas oxidization. The thermo-chemical state vector $\boldsymbol{x}_i$ is composed of 12 components, i.e., temperature T and mass fractions of the 11 chemical species Y_k $(k = 1, \ldots, n_s)$. To have a fair representation of all feature variables, which vary by several orders of magnitude, in the model, the data is scaled to $[-1, 1]$ via $\boldsymbol{x}^* = 2(\boldsymbol{x} - \boldsymbol{x}_{\min})/(\boldsymbol{x}_{\max} - \boldsymbol{x}_{\min}) - 1$, where $\boldsymbol{x}_{\min}$ and $\boldsymbol{x}_{\max}$ are the minimum and maximum values of samples in the dataset $\mathcal{D}_{\text{PSR}}$. In the following, we continue to use the notation $\boldsymbol{x}$ for the scaled data $\boldsymbol{x}^*$ for simplicity. The dataset $\mathcal{D}_{\text{PSR}}$ is split into multiple parts: 70% for training, *i.e.,* $\mathcal{D}_{\text{train}} = \{\boldsymbol{x}_i\}_{i=1}^{N_{\text{train}}}$ with the number of training samples $N_{\text{train}} \approx 1.14$ million, and 20% for testing $\mathcal{D}_{\text{test}} = \{\boldsymbol{x}_i\}_{i=1}^{N_{\text{test}}}$ with N_{test} being the number of testing samples.

3. Methodology.

3.1. Autoencoder for reduced chemistry.

For a high dimensional dataset, AE NN learns nonlinear mappings between the original high dimensional space and the reduced low dimensional space, *i.e.*, $\mathscr{F} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ and $\mathscr{G} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_x}$, where $n_z \ll n_x$, from data. More specifically, it learns simultaneously two groups of functions, *i.e.*, encoding functions $\boldsymbol{z} = \boldsymbol{f}_{\boldsymbol{w}_e}(\boldsymbol{x}) = \{z_i = f_{\boldsymbol{w}_e i}(\boldsymbol{x}) : i = 1, \ldots, n_z\}$ and decoding functions $\hat{\boldsymbol{x}} = \boldsymbol{g}_{\boldsymbol{w}_d}(\boldsymbol{z}) = \{\hat{x}_j = g_{\boldsymbol{w}_d j}(\boldsymbol{z}) : j = 1, \ldots, n_x\}$, parameterized by vectors $\boldsymbol{w}_e$ and $\boldsymbol{w}_d$, respectively, from a given training dataset, $\mathcal{D}_{\text{train}} = \{\boldsymbol{x}_i\}_{i=1}^{N_{\text{train}}}$, by minimizing the difference between $\mathcal{D}_{\text{train}}$ and the corresponding reconstructed dataset, $\widehat{\mathcal{D}} = \{\boldsymbol{g}_{\boldsymbol{w}_d}\left(\boldsymbol{f}_{\boldsymbol{w}_e}(\boldsymbol{x}_i)\right)\}_{i=1}^{N_{\text{train}}}$. The training of AE can be done with a standard MSE loss function. In the AE reduced chemistry work [30, 31], in addition to the reconstruction MSE loss, the loss function considers also a penalty loss ε_{ele} due to the violation of elemental mass conservation, as in

$$L_{\text{AE}} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \|\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i\|_2^2 + \varepsilon_{\text{ele}}. \tag{3}$$

It has been shown that the AE NN can reduce the mechanism size for CO/H_2 combustion from 12 to 2 with little loss of accuracy [30, 31].

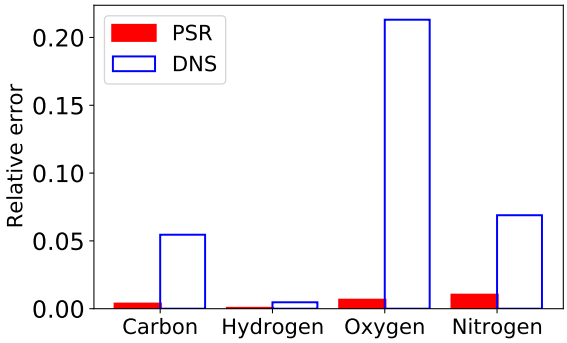

FIGURE 1. Elemental mass conservation errors. An elemental error of 20% is expected to cause large errors in temperature prediction thus false extinction/ignition events in engine simulations. With OOD-aware model, non-physical solutions with conservation laws violated can be detected before leading to poor engine designs.

3.2. OOD-aware autoencoder.

Data-driven modeling approaches like AE NN, although flexible and powerful, can be very dangerous to use in practice. One potential risk is model extrapolability. A trained data-driven model can always provide users predictions even for conditions that are very different from the training set. While this may be also true for some physical models, they usually come with explicit physical assumptions and well-known suitable application scenarios as well as clear explainability and interpretability. Using a model without knowing the applicable regions in CFD can dangerously lead to very poor even non-physical predictions. For example, Figure 1 shows the elemental mass conservation errors of two test sets, *i.e.*, PSR and DNS, from the reduced chemistry model. The maximum error in PSR test set, which has the same distribution as the training set, is less than

1%, while it shoots up to 20% in the DNS set [7], which has a different distribution than the training set. Such a large elemental error in reduced chemistry models is expected to cause significant errors in energy/temperature prediction, thus leading to false ignition or extinction solutions in engine simulations. An OOD-aware AE model for reduced chemistry that is able to detect the potential danger is thus needed for robust CFD predictions of combustion systems. In this work, we assess a recently developed uncertainty estimation methods, PI3NN [29], in capturing OOD samples for AE models, in comparison with the state-of-the-art DE method [11].

3.2.1. *The DE method.* Lakshminarayanan et al. [11] proposed a scalable deep ensemble method for predictive uncertainty quantification and assessed the method performance in predicting uncertainty for OOD samples. The method assumes a Gaussian data distribution and uses M NNs with random-initialized parameters to get an ensemble prediction of the mean and variance. Each NN outputs the predicted mean $\mu_{\boldsymbol{\theta}_m}(\boldsymbol{x})$ and variance $\sigma^2_{\boldsymbol{\theta}_m}(\boldsymbol{x})$ $(m = 1, \ldots, M)$, trained with the negative log-likelihood (NLL) loss function,

$$L_{\text{DE}} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \left[\frac{\log \sigma_i^2}{2} + \frac{(y_i - \mu_i)^2}{2\sigma_i^2} \right], \tag{4}$$

where N_{train} is the number of samples in the training dataset $\mathcal{D}_{\text{train}} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N_{\text{train}}}$. The final predicted mean and variances are,

$$\begin{aligned} \mu_{DE}(\boldsymbol{x}) &= \frac{1}{M} \sum_{m=1}^{M} \mu_{\boldsymbol{\theta}_m}(\boldsymbol{x}) \\ \sigma^2_{DE}(\boldsymbol{x}) &= \frac{1}{M} \sum_{m=1}^{M} \left(\sigma^2_{\boldsymbol{\theta}_m}(\boldsymbol{x}) + \mu^2_{\boldsymbol{\theta}_m}(\boldsymbol{x}) \right) - \mu^2_{DE}(\boldsymbol{x}). \end{aligned} \tag{5}$$

The method is shown to be able to capture the uncertainty difference between in-distribution (ID) and OOD samples in the classification problem with MNIST and NotMNIST images in [11].

3.2.2. *The PI3NN method.* The PI3NN method was proposed recently for quantifying uncertainty of regression models [29]. For a regression task, the method can provide a lower bound and an upper bound of the PI such that predictions fall between the bounds at a target percentile γ. It requires no data-distribution assumption, introduces no sensitive hyper-parameters and importantly, introduces no accuracy loss for the regression models.

For a regression NN, $y = f_{\boldsymbol{w}}(\boldsymbol{x})$, with the parameter vector $\boldsymbol{w}$ learned from a training set $\mathcal{D}_{\text{train}} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N_{\text{train}}}$, the PI3NN method introduces two additional stand-alone NNs, denoted by $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $v_{\boldsymbol{\xi}}(\boldsymbol{x})$, to quantify the upper and the lower bounds of the PI, respectively. The two NNs, $u_{\boldsymbol{\theta}}(\boldsymbol{x})$ and $v_{\boldsymbol{\xi}}(\boldsymbol{x})$, are trained separately with regular loss functions such as MSE, by using the datasets

$$\mathcal{D}_{\text{upper}} = \left\{ (\boldsymbol{x}_i, y_i - f_{\boldsymbol{\omega}}(\boldsymbol{x}_i)) \;\middle|\; y_i \geq f_{\boldsymbol{\omega}}(\boldsymbol{x}_i), i = 1, \ldots, N_{\text{train}} \right\}$$

and

$$\mathcal{D}_{\text{lower}} = \left\{ (\boldsymbol{x}_i, f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) - y_i) \;\middle|\; y_i < f_{\boldsymbol{\omega}}(\boldsymbol{x}_i), i = 1, \ldots, N_{\text{train}} \right\}.$$

The upper and lower bounds of the PI for a target percentile γ are then defined as $[f_{\boldsymbol{\omega}} - \beta_r v_{\boldsymbol{\xi}}, f_{\boldsymbol{\omega}} + \alpha_r u_{\boldsymbol{\theta}}]$ with β_r and α_r being the roots of the following equations,

$$Q_{upper}(\alpha) = \sum_{(\boldsymbol{x}_i, y_i) \in \mathcal{D}_{\text{upper}}} \mathbb{1}_{y_i \geq f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) + \alpha u_{\boldsymbol{\theta}}(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) - \frac{N_{\text{train}}(1-\gamma)}{2} = 0,$$

$$Q_{lower}(\beta) = \sum_{(\boldsymbol{x}_i, y_i) \in \mathcal{D}_{lower}} \mathbb{1}_{y_i \leq f_{\boldsymbol{\omega}}(\boldsymbol{x}_i) - \beta v_{\boldsymbol{\xi}}(\boldsymbol{x}_i)}(\boldsymbol{x}_i, y_i) - \frac{N_{\text{train}}(1-\gamma)}{2} = 0,$$

where $\mathbb{1}(\cdot)$ is the indicator function.

In Figure 2, we demonstrate the PI3NN method on a 1-D toy regression dataset with additive asymmetric non-Gaussian data noise, and compare it with the DE method. The training samples are drawn from $y = x^3 + \varepsilon$ with $x \in [-4, 4]$. The asymmetric noise ε is defined as,

$$\varepsilon = \begin{cases} 10\zeta, & \text{if } \zeta \geq 0 \\ 2\zeta, & \text{otherwise} \end{cases}, \text{ with } \zeta \sim \mathcal{N}(0, 1).$$

For such asymmetric noise, the 95% PI produced by PI3NN captures about 95% of training data with *tight* bounds in $[-4, 4]$, while DE produces an unnecessarily wide lower bound in $[-4, 4]$ due to its Gaussian assumption on noise distribution. The PI3NN method works well with the asymmetric noise since it handles the upper and lower bounds of the PI separately with two NNs. The DE method can only predict symmetric PIs since it assumes Gaussian data distribution with mean μ_{DE} and variance σ^2_{DE} as in Eq. (5). Both PI3NN and DE produce reasonably wide PIs in the OOD region $[-7, -4] \cup [4, 7]$.

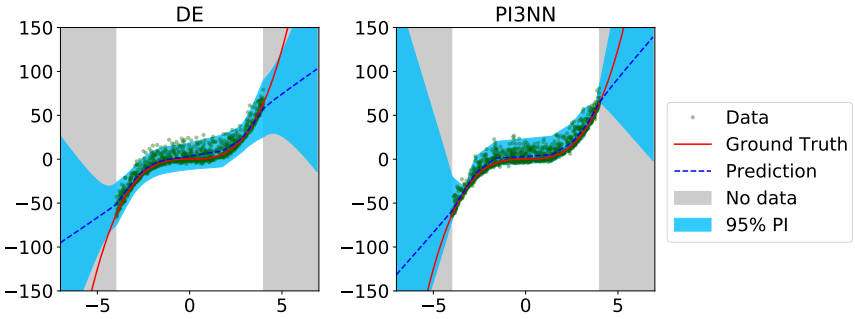

FIGURE 2. Estimation of 95% PI prediction in toy regression task $y = x^3 + \varepsilon$ with asymmetric noise ε. The 95% PI produced by PI3NN captures about 95% of training data with *tight* bounds within training region, while DE produces an unnecessarily wide lower bound. Both methods produce reasonably wide PIs in the OOD region.

4. **Experiments.** In this work, we combine the two methods, i.e., DE and PI3NN, with AE and check the OOD-detection performance in two experiments. Figures 3 (a) and (b) show the NN architecture of AE with DE and PI3NN, respectively. In the first experiment, the NNs are trained and tested on two well-known image datasets: MNIST and Fashion-MNIST. In the second experiment, two combustion datasets: PSR and DNS, are used for the assessment.

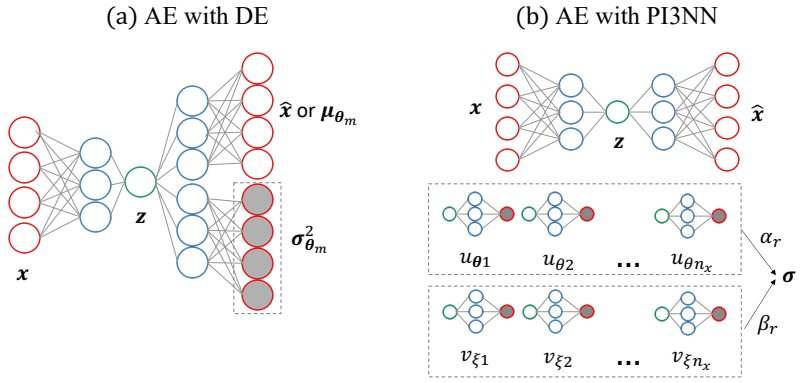

FIGURE 3. Architecture diagrams of AE with DE and PI3NN. Neurons are represented by the circles with different edge colors for different types of layers, i.e., red for input and output layers, green for bottleneck layer, and blue for other hidden layers. The red circles filled with gray are outputs for variance/standard deviation variables. The diagrams shown here are simplified for illustration only. The actual numbers of layers and neurons in each layer vary in different experiments.

4.1. **MNIST VS. Fashion-MNIST.** In the image experiment, the NNs are trained on MNIST training set with 60,000 digit images, and tested on MNIST and Fashion-MNIST test sets, which are unseen in training and referred to as the ID and OOD test sets, respectively. The ID test set has 10,000 digit images and the OOD has 10,000 fashion and clothing images.

The standalone AE has five dense layers: input layer, hidden layer with 392 neurons and ReLU activation function, bottleneck layer with 64 neurons and ReLU activation function, hidden layer with 392 neurons and ReLU activation function, and output layer with sigmoid activation function. When combining AE with DE method, as shown in Figure 3(a), we add one more branch, consisting of a hidden layer with 784 neurons with ReLu activation function and an output layer with $(\cdot)^2$ followed to guarantee nonnegative outputs, to the bottleneck layer $\boldsymbol{z}$, for variance $\boldsymbol{\sigma}^2_{\boldsymbol{\theta}_m}$ prediction. When the PI3NN method is applied for uncertainty estimation, a class of additional NNs are used for variance prediction as shown in Figure 3(b), without any modifications to AE. The architecture of variance NNs in PI3NN can be flexible, ranging from one NN for one output feature, i.e., one pixel, to one NN for all output features, i.e., 784 pixels. In the work, we divided the 784 features into 7 groups and trained a NN for each with 112 output features.

All the NNs are implemented with Tensorflow 2.4.0. For the training of NNs in AE with DE, the specification of learning rate l_r is found to be tricky with the NLL loss function in Eq. (4). The reconstruction part, $(y_i - \mu_i)^2$, prefers a learning rate around 10^{-3}, while a much smaller learning rate around 10^{-6} is needed when σ_i^2 included. For faster training, we first trained the NNs with the reconstruction MSE loss for 800 epochs at $l_r = 10^{-3}$ to get a well-reconstructed image output, and then changed the loss function to the NLL loss function in Eq. (4) and trained

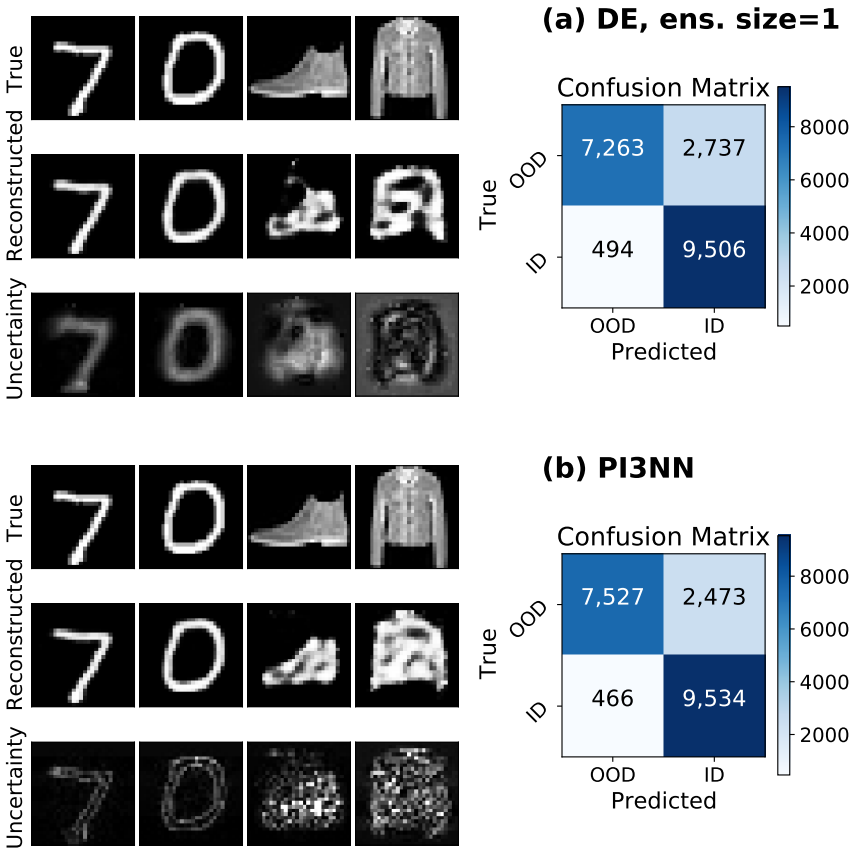

FIGURE 4. Comparison of OOD detection accuracy in MNIST (ID) and Fashion-MNIST (OOD) test sets between DE and PI3NN methods. Both methods capture the difference between ID and OOD with larger uncertainties for OOD samples. The confusion matrices show that the majority of test samples are correctly detected as OOD (true positive) or ID (true negative).

the NNs for another 20,000 epochs at $l_r = 10^{-6}$. The training of the variance NNs in AE with PI3NN was faster with 500 epochs at $lr = 10^{-4}$. Training of NNs in the image experiment was conducted on a 2019 MacBook Pro with a single CPU. It took around 8 h in total to train the 7 variance NNs in AE with PI3NN. The training of one NN in AE with DE took around 13 h.

Figure 4 shows the comparison results of DE with a single run ($M = 1$) and PI3NN. A good uncertainty estimation method should predict small uncertainties for samples from the same distribution with training set, i.e., ID, and large uncertainties for samples outside of training distribution, i.e., OOD. Both DE and PI3NN succeed in predicting larger uncertainties, shown as more and brighter white pixels in Figure 4, for samples from Fashion-MNIST/OOD than samples from MNIST/ID. To have a quantitative comparison, we first define an index for OOD detection as the L^2-norm of uncertainty vector of ith image, $\sigma_{l_2 i} = \|\boldsymbol{\sigma}_i\|_2$, and then specify a threshold value σ_{ood} such that any images with $\sigma_{l_2 i} > \sigma_{ood}$ are detected as OOD

samples. The subplots in the right hand side of Figure 4 show the OOD detection accuracy for DE and PI3NN, where σ_{ood} is defined as the 95% percentile of $\{\sigma_{l_2 i}\}_{i=1}^{N_{\text{train}}}$ of the training set. Clearly, both methods show reasonably good detection accuracy with 95% samples from ID test set and 75% (or 74% for DE) samples from OOD test set correctly classified. By varying the threshold value σ_{ood}, the receiver operating characteristic (ROC) curves can be obtained, as shown in Figure 5, where DE with ensemble size of two ($M = 2$) is also shown. In this experiment, DE with $M = 2$ shows an excellent OOD detection accuracy followed by PI3NN and both are better than DE with $M = 1$.

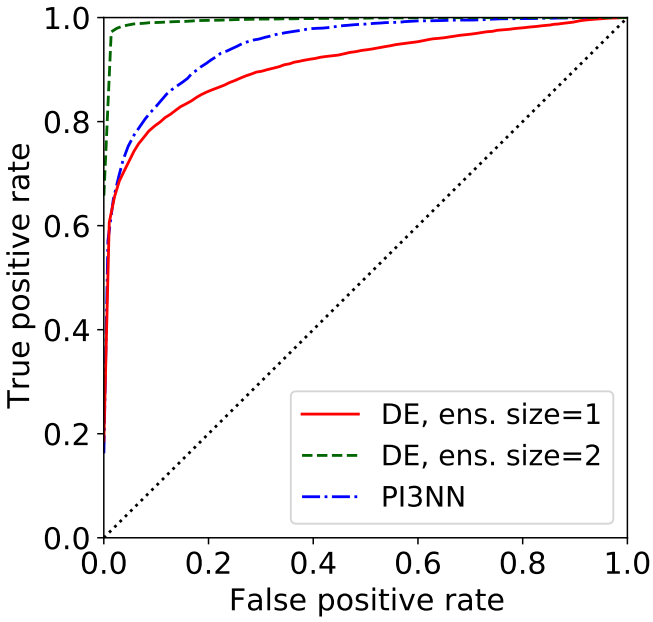

FIGURE 5. Receiver operating characteristic curves of DE and PI3NN in OOD detection. All methods can capture the OOD samples well in the image experiment. Though DE with ensemble size of 2 shows better accuracy than PI3NN, it requires two runs while PI3NN only needs a single run. DE with a single run shows lower OOD detection accuracy than PI3NN.

One observed difference in Figure 4 between the two methods in the MNIST/ID test set is that PI3NN predicts high uncertainties only at the edges of digits where the reconstruction errors are high, while DE predicts high uncertainties across the body area of digits. The correlation coefficients between uncertainties and reconstruction errors of all ID samples are -0.18 and 0.40 for DE and PI3NN, respectively. The better alignment between uncertainties and errors of samples suggests that PI3NN is a better error indicator. In addition, PI3NN is nonintrusive to AE, while the accuracy of AE deteriorates in DE due to the multiple training tasks in the NLL loss (Eq. (4)). The reconstruction MSE error in AE when combined with DE was found to increase from 0.0025 to 0.0037 for the MNIST test set and from 0.0806 to 0.0937 for the Fashion-MNIST test set, while the reconstruction accuracy was not affected when with PI3NN.

The image experiment with MNIST and Fashion-MNIST is designed for two purposes. First, it is of interest to examine if PI3NN and DE can detect OOD samples in image dimension reduction problems, and MNIST and Fashion-MNIST are two popular image datasets in OOD studies. MNIST has been used as ID set in a classification task [11], where the DE method is shown to be able to capture larger uncertainties for OOD samples from NotMNIST set consisting of alphabets. Second, it serves as a transition from the 1D toy regression to the complex combustion problem in section 4.2 with increasing complexity. PI3NN and DE have been shown capable of detecting OOD samples in both the 1D toy regression and the image experiment. In the next, we look at the combustion problem.

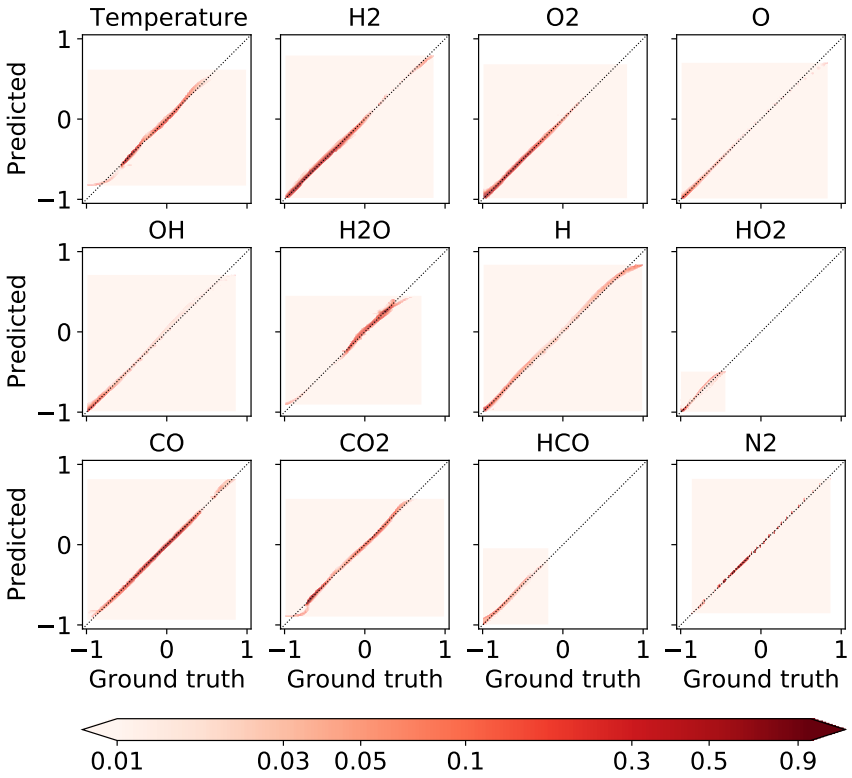

FIGURE 6. Normalized joint histogram of predicted values by AE (with $n_z = 2$) and true values for 12 thermo-chemical state variables, i.e., temperature and mass fractions of 11 species, in PSR test set. The dark red area along the diagonal line shows that AE can reduce the state dimension of syngas CO/H_2 combustion from $n_x = 12$ to $n_z = 2$ without much loss of accuracy.

4.2. **Syngas combustion chemistry.** In the reduced chemistry experiment, the training set consists of around 1.14 million combustion/thermo-chemical states from steady 0-D PSR reactors (Eq. (2)). The ID test set is composed of around 0.34 combustion states also from steady 0-D PSR but at different parametric conditions with the training set. The OOD test set, with around 1.0 million combustion states,

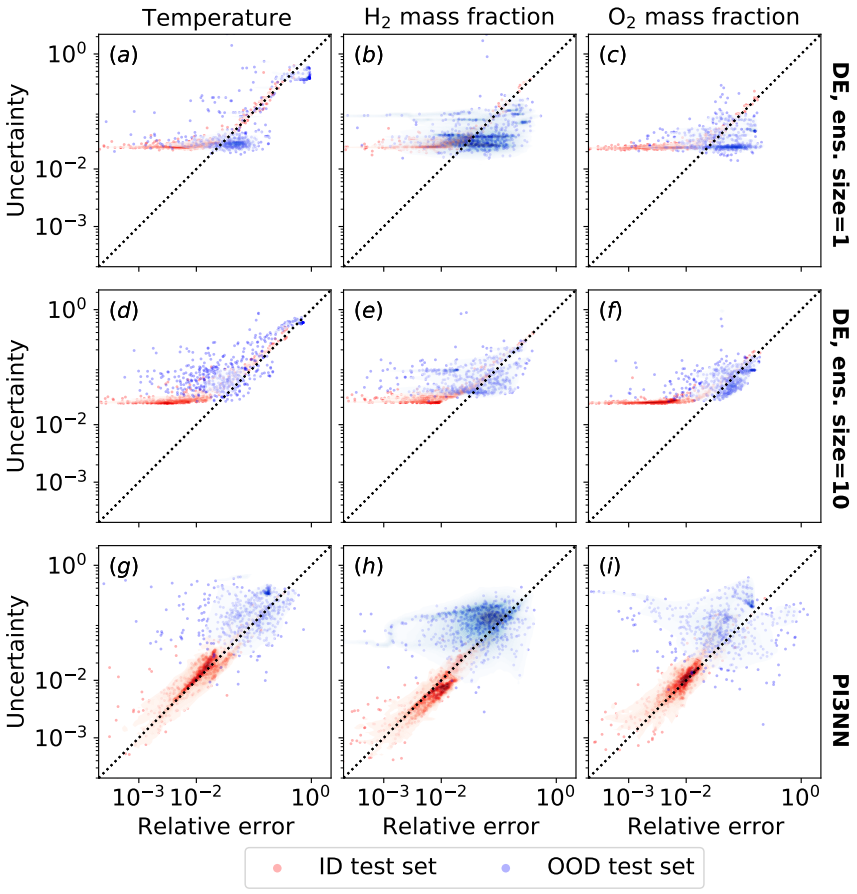

FIGURE 7. Predictive uncertainty (the width of PI) vs. predictive error of the ID (red) and OOD (blue) test sets from the PI3NN and DE methods. Scatter points are samples randomly selected from the two test sets. The filled contours show the normalized joint histogram for samples in ID/OOD test sets, with light red/blue colors representing fewer samples and dark red/blue representing more samples. DE with a single run ($(a) - (c)$) fails to capture the difference in uncertainty for ID and OOD samples. DE with 10 runs ($(d) - (f)$) shows improved but still limited separation of OOD samples from ID samples and it fails to produce the uncertainty-error correlation. PI3NN ($(g) - (i)$) shows a strong correlation between the uncertainty and the error and clearly demonstrates that OOD and ID have different uncertainty magnitudes.

is taken from the 3-D DNS work in [7], where the full set of governing equations for turbulent combustion including Eq. (1) are solved by using a high-fidelity solver [3]. Both ID and OOD test sets are not seen during training.

The AE NN with the same architecture as in [31] is used. It has five dense layers, i.e., input layer, hidden layer in the encoder part with 12 neurons, bottleneck latent layer with 2 neurons, hidden layer in the decoder part with 12 neurons, and output

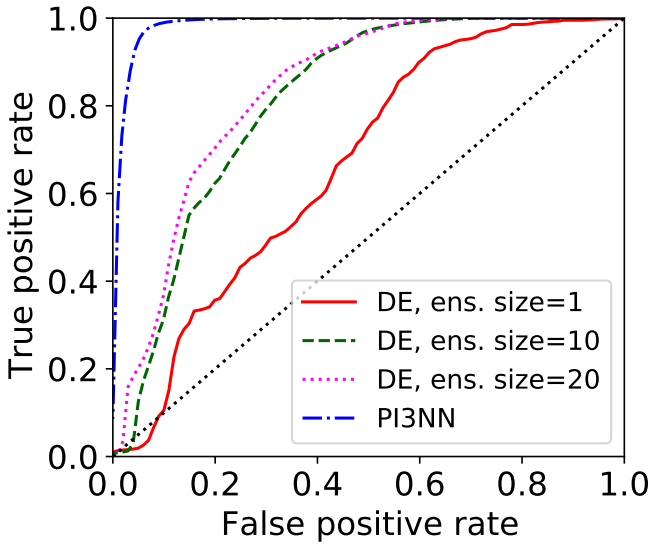

FIGURE 8. Receiver operating characteristic curves of PI3NN and DE ($M = 1, 10, 20$) methods in OOD detection. PI3NN shows the best OOD detection accuracy. DE shows improved detection accuracy with increasing ensemble size (M).

layer. Hyperbolic tangent activation function, Tanh, is used. Also following [31], AE is trained with a loss function consisting of two parts, i.e., the reconstruction MSE loss and a soft penalty term ε_{ele} due to the violation of elemental mass conservation law, as in Eq. (3). Figure 6 shows the normalized joint histogram of AE predicted values and true values for samples in PSR/ID test set. The large sample number distribution along the diagonal dashed line, represented by the dark red color, shows that AE with $n_z = 2$ has an accurate prediction of thermo-chemical state variables in the syngas CO/H_2 combustion case. The entire PSR dataset, $\mathcal{D}_{\text{PSR}}$, is scaled to -1 and 1, before split into the training and test sets, i.e., $\mathcal{D}_{\text{train}}$ and $\mathcal{D}_{\text{test}}$. This explains why, for some state variables, e.g., HCO and HO_2, the samples from the test set do not span the full range of -1 and 1 in Figure 6. HCO and HO_2 are radical species that are produced and consumed in the thin flame region. The radical species have much more complicated structures than the major species, e.g., H_2 and O_2, and hence pose a bigger challenge to accurate prediction. The good agreement in Figure 6 suggests that the AE with $n_z = 2$ can capture all the state variables including the radical species even satisfactorily.

When combining AE with PI3NN, we use two sets of 12 variance NNs for the upper and lower uncertainty profiles of the 12 output features, respectively. Each NN outputs the upper or lower uncertainty of a single feature, where 1 dense hidden layer with 1200 neurons and Tanh activation function is employed. Here, the variance NNs take the thermo-chemical vector $\boldsymbol{x}$ as input with 12 neurons, the same as AE, instead of the bottleneck layer $\boldsymbol{z}$ as in the image experiment. The reason is that $\boldsymbol{z}$ with two neurons is found to be, though representing combustion states satisfactorily as shown in Figure 6, incapable of describing the uncertainties of combustion states well, which are much more complicated than the images expectedly. With

the outputs from the variance NNs, the PI is then calculated following the procedure in section 3.2.2 with a target percentage γ of the training samples covered. Different from the DE method, the variance NNs in PI3NN are trained independently from AE, hence providing a nonintrusive method for uncertainty. In the combustion problem, $\gamma = 0.6826$ is specified to get the reported PIs. The learning rate is specified as 5×10^{-4} for the training of the NNs. All NNs in the combustion experiment were trained by using the Compute and Data Environment for Science (CADES) computing condo in Oak Ridge National Laboratory with a single CPU. The training of two sets of 12 variance NNs with 6000 epochs for the upper and lower PI bounds took about 43 h and 49 h, respectively.

When combining AE with DE, we use an additional NN for feature variances. It is designed in this way, instead of having a branch attached to the bottleneck latent layer $\boldsymbol{z}$ as in the image experiment, to be consistent with PI3NN. The variance NN takes the same input as AE and outputs variances for the 12 features with three dense hidden layers. Every hidden layer has 60 neurons with Tanh activation function. Softplus activation function is used for the output layer to guarantee a nonnegative variance. The two NNs, i.e., AE and variance NN, are trained together with the loss function,

$$L_{DE} = \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \left[\frac{1}{2} \sum_{k=1}^{n_x} \log \sigma_{i,k}^2 + \frac{1}{2} \sum_{k=1}^{n_x} \frac{(x_{i,k} - \mu_{i,k})^2 + \varepsilon_{ele}}{\sigma_{i,k}^2} \right], \tag{6}$$

where N_{train} is the number of samples in the training set, $n_x = 12$ is the number of input/output features, $\mu_{i,k}$ and $\sigma_{i,k}^2$ are the predicted mean and variance of the kth feature of the ith sample, $x_{i,k}$, respectively. The loss function in Eq. (6) is a modified version of the NLL loss in Eq.(4) with output being a vector and elemental mass conservation constraint added. In the training process, a competition between the two terms in the loss function is observed, which leads to an undesirable accuracy deterioration of the AE model, similar to the observation in the image experiment. A decreasing stepwise learning rate with 6000 epochs is specified in the training, where 0.01, 0.005 and 0.001 are used for the first 3000 epochs and 5×10^{-4} is used for the rest. The training of one NN in DE with AE took around 4.5 h, and the total training time for DE with ensemble size of 10 and 20 was around 45 h and 90 h, respectively.

A sound PI method should be able to detect OOD samples by providing a large uncertainty on OOD and a small uncertainty on ID. It should also produce a large predictive uncertainty when the predictive error is high and vice versa, showing a close correspondence between the uncertainty and the error. Figure 7 shows the predictive uncertainty against the relative error for Temperature and mass fractions of H_2 and O_2 of samples from ID and OOD test sets. Three methods, i.e., DE with $M = 1$, DE with $M = 10$ and PI3NN, are compared. Results of PI3NN, in subplots $(g) - (i)$, show a strong correlation between the uncertainty and the error with a large number of samples clustered along the diagonal line. Additionally, PI3NN can clearly separate OOD samples (in blue colors) from ID samples (in red colors) with different uncertainty and error magnitudes. In comparison, DE with $M = 1$, in subplots $(a) - (c)$, cannot distinguish OOD samples from ID. Although DE with ensemble size of ten ($M = 10$), in subplots $(d) - (f)$, shows an improved separation of OOD from ID, it does not produce as good uncertainty-error correlation as PI3NN. We further increase the ensemble size in DE method to $M = 20$, the ROC curve of which is plotted together with the other three methods

in Figure 8. Clearly, PI3NN shows the best OOD detection accuracy with the ROC curve close to the top left corner, representing a high true positive rate and low false positive rate. For a high true positive rate of 90%, the false positive rate is as low as 4% in PI3NN. The DE methods shows improved detection accuracy with ensemble size M increasing. But the accuracy appears to plateau at around $M = 20$. In DE with $M = 20$, the false positive rate is 38%, which is much higher than the case in PI3NN, when the true positive rate is 90%. Furthermore, PI3NN is easy to train, while the training of DE is tedious due to the multiple objectives involved in its loss function.

5. **Conclusions.** In this work, we developed an enhanced AE model for reducing combustion chemical kinetics with OOD-detection capability by combining PI3NN with AE, which will lead to more robust reacting flow simulations. The method performance was first examined in the image experiment with samples from MNIST and Fashion-MNIST, where a state-of-the-art uncertainty approach—DE—was also considered. Both methods were found to capture the OOD samples in the image experiment well. The methods, i.e, AE-PI3NN and AE-DE, were then applied to syngas CO/H_2 combustion. The dimension of combustion states was reduced by AE from 12 to 2 at a good accuracy, implying that the number of differential equations to solve in combustion simulations is 6 times smaller with the reduced chemistry model. The AE-PI3NN showed excellent OOD-detection accuracy, e.g., 90% true positive rate and 4% false positive rate, for syngas combustion with a clear separation of uncertainty magnitudes for OOD samples from ID. Even though the AE-DE showed improved detect accuracy with ensemble size M increasing, it appeared to plateau at around $M = 20$ and produced a false positive rate of 38% for the 90% true positive rate, which is much higher than the 4% in AE-PI3NN. Furthermore, AE-PI3NN is nonintrusive and easy to train, while AE-DE shows an undesirable accuracy deterioration in combustion prediction and the training is sensitive to learning rate due to the multiple tasks in the loss function.

Acknowledgments. This material is based upon work supported in part by the U.S. Department of Energy, Office of Science, Offices of Advanced Scientific Computing Research, under the contract ERKJ352, and by the AI Initiative at the Oak Ridge National Laboratory (ORNL). ORNL is operated by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

REFERENCES

[1] A. Amini, W. Schwarting, A. Soleimany and D. Rus, Deep evidential regression, In *Advances in Neural Information Processing Systems*, **33** (2020), 14927–14937.

[2] R. B. Bird, W. E. Stewart and E. N. Lightfoot, *Transport Phenomena*, John Wiley and Sons, New York, 1960.

[3] J. H. Chen, A. Choudhary, B. de Supinski, M. DeVries, E. R. Hawkes, S. Klasky, W. K. Liao, K. L. Ma, J. Mellor-Crummey, N. Podhorszki, R. Sankaran, S. Shende and C. S. Yoo, Terascale direct numerical simulations of turbulent combustion using S3D, *Computational Science & Discovery*, **2** (2009), 015001.

[4] A. Coussement, O. Gicquel and A. Parente, MG-local-PCA method for reduced order combustion modeling, *Proceedings of the Combustion Institute*, **34** (2013), 1117–1123.

[5] G. Esposito and H. Chelliah, Skeletal reaction models based on principal component analysis: Application to ethylene-air ignition, propagation, and extinction phenomena, *Combustion and Flame*, **158** (2011), 477–489.

[6] Y. Gal and Z. Ghahramani, Dropout as a Bayesian approximation: Representing model uncertainty in deep learning, In *Proceedings of The 33rd International Conference on Machine Learning*, **48** (2016), 1050–1059.
[7] E. R. Hawkes, R. Sankaran, J. C. Sutherland and J. H. Chen, Scalar mixing in direct numerical simulations of temporally evolving plane jet flames with skeletal CO/H_2 kinetics, *Proceedings of the Combustion Institute*, **31** (2007), 1633–1640.
[8] M. D. Hoffman, D. M. Blei, C. Wang and J. Paisley, Stochastic variational inference, *J. Mach. Learn. Res.*, **14** (2013), 1303–1347, URL http://jmlr.org/papers/v14/hoffman13a.html.
[9] B. J. Isaac, A. Coussement, O. Gicquel, P. J. Smith and A. Parente, Reduced-order PCA models for chemical reacting flows, *Combustion and Flame*, **161** (2014), 2785–2800.
[10] B. J. Isaac, J. N. Thornock, J. Sutherland, P. J. Smith and A. Parente, Advanced regression methods for combustion modelling using principal components, *Combustion and Flame*, **162** (2015), 2592–2601.
[11] B. Lakshminarayanan, A. Pritzel and C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, **17** (2017), 6405–6416.
[12] Y. Li, C.-W. Zhou, K. P. Somers, K. Zhang and H. J. Curran, The oxidation of 2-butene: A high pressure ignition delay, kinetic modeling study and reactivity comparison with isobutene and 1-butene, *Proceedings of the Combustion Institute*, **36** (2017), 403–411.
[13] T. Lu and C. K. Law, Toward accommodating realistic fuel chemistry in large-scale computations, *Progress in Energy and Combustion Science*, **35** (2009), 192–215.
[14] D. J. C. MacKay, A practical Bayesian framework for backpropagation networks, *Neural Comput.*, **4** (1992), 448–472.
[15] M. R. Malik, B. J. Isaac, A. Coussement, P. J. Smith and A. Parente, Principal component analysis coupled with nonlinear regression for chemistry reduction, *Combustion and Flame*, **187** (2018), 30–41.
[16] H. Mirgolbabaei and T. Echekki, A novel principal component analysis-based acceleration scheme for LES–ODT: An a priori study, *Combustion and Flame*, **160** (2013), 898–908.
[17] H. Mirgolbabaei and T. Echekki, Nonlinear reduction of combustion composition space with kernel principal component analysis, *Combustion and Flame*, **161** (2014), 118–126.
[18] H. Mirgolbabaei and T. Echekki, The reconstruction of thermo-chemical scalars in combustion from a reduced set of their principal components, *Combustion and Flame*, **162** (2015), 1650–1652.
[19] H. Mirgolbabaei, T. Echekki and N. Smaoui, A nonlinear principal component analysis approach for turbulent combustion composition space, *International Journal of Hydrogen Energy*, **39** (2014), 4622–4633.
[20] D. A. Nix and A. S. Weigend, Learning local error bars for nonlinear regression, In *Advances in neural information processing systems*, (1995), 489–496.
[21] A. Parente, J. C. Sutherland, L. Tognotti and P. J. Smith, Identification of low-dimensional manifolds in turbulent flames, *Proceedings of the Combustion Institute*, **32** (2009), 1579–1586.
[22] T. Pearce, A. Brintrup, M. Zaki and A. Neely, High-quality prediction intervals for deep learning: A distribution-free, ensembled approach, In *Proceedings of the 35th International Conference on Machine Learning*, **80** (2018), 4075–4084.
[23] T. Pearce, F. Leibfried and A. Brintrup, Uncertainty in neural networks: Approximately Bayesian ensembling, In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, **108** (2020), 234–244.
[24] T. S. Salem, H. Langseth and H. Ramampiaro, Prediction intervals: Split normal mixture from quality-driven deep ensembles, In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, Proceedings of Machine Learning Research, **124** (2020), 1179–1187.
[25] R. Shan and T. Lu, Ignition and extinction in perfectly stirred reactors with detailed chemistry, *Combustion and Flame*, **159** (2012), 2069–2076.
[26] E. Simhayev, G. Katz and L. Rokach, PIVEN: A deep neural network for prediction intervals with specific value prediction, 2021.
[27] J. C. Sutherland and A. Parente, Combustion modeling using principal component analysis, *Proceedings of the Combustion Institute*, **32** (2009), 1563–1570.
[28] S. Vajda, P. Valko and T. Turányi, Principal component analysis of kinetic models, *International Journal of Chemical Kinetics*, **17** (1985), 55–81.

[29] P. Zhang, S. Liu, D. Lu, G. Zhang and R. Sankaran, A prediction interval method for uncertainty quantification of regression models, In *ICLR 2021 SimDL Workshop*, Virtual, 2021.
[30] P. Zhang, R. Sankaran and E. R. Hawkes, A priori examination of reduced chemistry models derived from canonical stirred reactors using three-dimensional direct numerical simulation datasets, In *AIAA Scitech*, (2021), 1784,
[31] P. Zhang, R. Sankaran, M. Stoyanov, D. Lebrun-Grandie and C. E. Finney, Reduced models for chemical kinetics derived from parallel ensemble simulations of stirred reactors, In *AIAA Scitech Forum*, (2020), 0177.

Received August 2021; revised September 2021; early access November 2021.

E-mail address: zhangp1@ornl.gov
E-mail address: lius1@ornl.gov
E-mail address: lud1@ornl.gov
E-mail address: zhangg@ornl.gov
E-mail address: sankaranr@ornl.gov