

A deep learning-based model reduction (DeePMR) method for simplifying chemical kinetics

Zhiwei Wang^{a,b}, Yaoyu Zhang^{a,b,c}, Yiguang Ju^d, Weinan E^{e,f},

Zhi-Qin John Xu^{a,b,*}, Tianhan Zhang^{f,*}

^a*Institute of Natural Sciences, School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, 200240, China*

^b*MOE-LSC and Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai, 200240, China*

^c*Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, 200240, China*

^d*Department of Mechanical and Aerospace Engineering, Princeton University, NJ, 08540, US*

^e*School of Mathematical Sciences, Peking University, Beijing, 100871, China;*

^f*AI for Science Institute, Beijing, 100080, China*

**Corresponding authors: xuzhiqin@sjtu.edu.cn, tianhanz@princeton.edu*

Abstract

A deep learning-based model reduction (DeePMR) method for simplifying chemical kinetics is proposed and validated using high-temperature auto-ignitions, perfectly stirred reactors (PSR), and one-dimensional freely propagating flames of n-heptane/air mixtures. The mechanism reduction is modeled as an optimization problem on Boolean space, where a Boolean vector, each entry corresponding to a species, represents a reduced mechanism. The optimization goal is to minimize the reduced mechanism size given the error tolerance of a group of pre-selected benchmark quantities. The key idea of the DeePMR is to employ a deep neural network (DNN) to formulate the objective function in the optimization problem. In order to explore high dimensional Boolean space efficiently, an iterative DNN-assisted data sampling and DNN training procedure are implemented. The results show that DNN-assistance improves sampling efficiency significantly, selecting only 10^5 samples out of 10^{34} possible samples for DNN to achieve sufficient accuracy. The results demonstrate the capability of the DNN to recognize key species and reasonably predict reduced mechanism performance. The well-trained DNN guarantees the optimal reduced mechanism by solving an inverse optimization problem. By comparing ignition delay times, laminar flame speeds, temperatures in PSRs, the resulting skeletal mechanism has fewer species (45 species) but the same level of accuracy as the skeletal mechanism (56 species) obtained by the Path Flux Analysis (PFA) method. In addition, the skeletal mechanism can be further reduced to 28 species if only considering atmospheric, near-stoichiometric conditions (equivalence ratio between 0.6 and 1.2). The DeePMR provides an innovative way to perform model reduction and demonstrates the great potential of data-driven methods in the combustion area.

Keywords: model reduction; machine learning; deep neural network; chemical kinetics

1. Introduction

The development of detailed chemistry mechanisms of hydrocarbon fuels paves the way to realistic simulations of practical combustors [1]. However, due to chemistry stiffness, the simulation of large-size detailed mechanisms become forbiddingly expensive, especially for very large-scale simulation. Figure 1a shows a species-reaction graph of the n-heptane mechanism with 116 species and 830 reactions, revealing an intrinsic complexity of the chemistry mechanism [2]. It is a highly non-trivial task to remove part of the chemistry mechanism but maintain its overall accuracy for simulations in various reactors. Therefore, there is a substantial need to develop a systematic way to simplify chemistry mechanisms.

Previous researchers have proposed various model reduction methods. The first type of reduction method is by sensitivity analysis [3, 4]. The second type is time scale analysis, including quasi-steady-state approximations (QSSA) [5], partial equilibrium assumptions (PEA) [6], computational singularity perturbation (CSP) [7–9], rate-controlled constrained-equilibrium (RCCE) [10], and lifetime analysis [11].

The third type is graph-based methods, including directed relation graph (DRG) [12], directed relation graph with error propagation (DRGEP) [13], path flux analysis (PFA) [14], and directed relation graph with error propagation and sensitivity analysis (DRGEP-SA) [15]. The methods focus on maintaining the fluxes in the reaction network associated with a set of targeted species when removing species. The major difference among the DRG, DRGEP, and PFA methods is the interaction coefficient definition. First, DRG, DRGEP, and PFA use the absolute, net, summation of production and consumption rates to define the interaction coefficient for the directly related species, respectively. In addition, DRG assumes that error does not decay along the graph-search path and adopts the path with maximum error. In contrast, DRGEP assumes that the error induced by indirectly related species decays geometrically. Wu et al. [16] showed that DRG might overestimate the errors in the starting species, and DRGEP tends to underestimate reduction errors since the geometric error decay may overly predict decay speed. The PFA method combines the features of DRG and DRGEP, and proposes a new scheme that considers error decay but only up to the second order. Besides the error estimation method, an important topic for graph-based methods is choosing the target set of species. Chen and Chen [17] showed that with or without H radical in the target set significantly impacts flame speed prediction accuracy. Curtis et al. [18] proposed the relative importance index (RII) method to determine the target species set automatically. In summary, the interaction coefficient definition and the set of target species determine the path flux error triggered by the removal of species.

Generally speaking, it is the key assumption of graph-based methods that flux error evaluation can

represent the overall performance of the reduced mechanisms. However, the validity of the current assumption heavily relies on the specific interaction coefficient definition. It is unclear when the assumption fails. For example, Sun et al. [14] pointed out that DRGEP cannot identify the importance of NO_x catalytic effect on ignition enhancement due to using net reaction rates in error estimation.

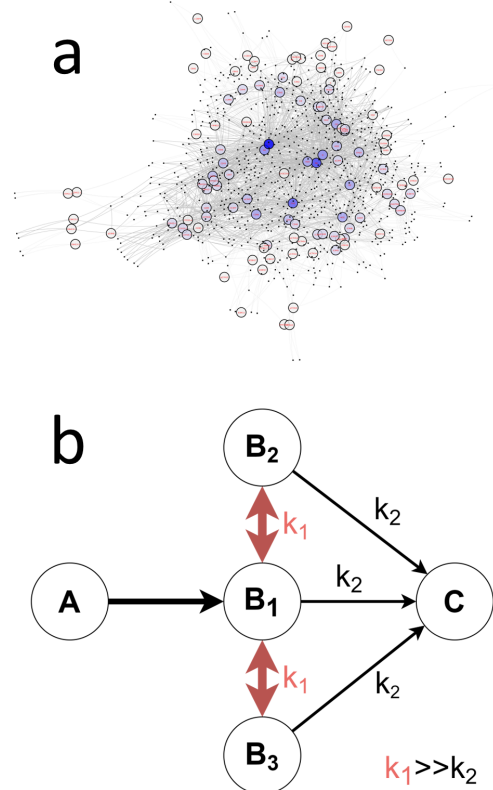


Fig. 1: (a): Species-reaction graph for detailed n-heptane mechanism (116 species, 830 reactions). Circles represent species, and black dots represent reactions. The circles are dyed by the vortex degree; (b): An illustration for a reaction pathway.

Figure 1b shows another example where three graph-based methods fail to perform the reduction. Figure 1b represents a simplified mechanism, where A , B_i , C are the reactant, intermediate species, and product, respectively. The inter-conversion rates among B_1 , B_2 , and B_3 are faster than other reaction rates. Consequently, B_i has similar concentrations and contributions to the target species C . Based on error estimation methods mentioned above, there is little difference among the impacts of B_i 's on the flux to the target species C . As a result, B_i will either be removed or retained altogether. In contrast, it is readily seen that removing B_2 and B_3 will not affect the reactant's consumption rate and product's creation rate, which is an acceptable reduction for the original

mechanism, and graph-based methods do not come up with such a solution. Due to the extreme complexity of the detailed mechanism shown in Figure 1a, it is challenging for graph-based methods to evaluate to what extent the path flux can accurately indicate the overall performance of the reduced mechanisms and where the assumption might fail.

On the other hand, there is a recent spike of machine learning in combustion kinetic studies. For generating detailed chemistry mechanism, Zeng et al. [19] adopted a novel DeePMD method [20] to build a detailed methane combustion mechanism using *ab initio* molecular dynamics (AIMD). In terms of model reduction, Si [21] and Kelly [22] used machine learning algorithms to optimize the parameters in the chemistry mechanism.

This paper develops a data-driven method based on the deep neural network (DNN) to reduce species and reactions in detailed mechanisms. First, we will introduce the basic idea of constructing an optimization problem for model reduction. The mathematical formulation is straightforward, while the challenge is to design a DNN to represent the objective function. The training-sampling iteration, the DNN-assisted data sampling strategy, and the inverse optimization problem will be introduced in detail. Then, the iterative training process will be discussed, including the statistical features of the sampled data and the DNN prediction performance. Afterward, the resulting reduced mechanism (45 species) is validated by comparing the ignition delay times, laminar flame speeds, and extinction curves in PSRs. Moreover, a smaller-size reduced mechanism is achieved by narrowing the range of benchmark quantities, demonstrating the flexibility of the current method. Finally, conclusions are drawn.

2. Methodology

In this section, we would describe our proposed deep learning-based model reduction (DeePMR) method to find optimal reduced mechanisms, followed by the implementation details in next section.

2.1. Optimization framework for searching optimal reduced mechanisms

Suppose we have a detailed mechanism with N_s species and simplify it by removing unnecessary species. First, to represent a reduced mechanism, we employ an N_s -dimensional Boolean vector $\mathbf{x} \in \{0, 1\}^{N_s}$, each entry indicating the status of a species: true or one means remaining, false or zero removing, similar to [23]. For convenience, we use float numbers 0 and 1 instead of True or False to implement our algorithm. Then, we assume the performance of the reduced mechanism can be indicated by a function $\text{Err}(\mathbf{x})$. For example, function $\text{Err}(\mathbf{x}) : \{0, 1\}^{N_s} \rightarrow \mathbb{R}^{N_t}$ can be the errors of the reduced mechanism \mathbf{x} in N_t benchmark quantities, such as the ignition delay times and laminar flame speeds under various initial conditions, temperature in PSR with various resi-

dence times. The ultimate goal is to find such an optimal solution of \mathbf{x} that the error of the reduced mechanism \mathbf{x} is within our tolerance, and the size of the reduced mechanism \mathbf{x} , in other words, the sparsity of the vector \mathbf{x} , is as small as possible. The reduction procedure can be considered as a typical optimization problem:

$$\min F(\mathbf{x}) := \sum_{i=1}^{N_t} \text{Err}(\mathbf{x})_i + \beta \|\mathbf{x}\|_0. \quad (1)$$

where $\text{Err}(\mathbf{x})_i$ is the error of the i -th benchmark quantity, β is the weight of the sparsity.

2.2. Deep neural network modeling performance function

Once the objective function $F(\mathbf{x})$ is determined, the optimal solution can be achieved by optimization. However, the major difficulty is parametrizing the function $\text{Err}(\mathbf{x})$. The function $\text{Err}(\mathbf{x})$ needs to reflect the overall performance of a reduced mechanism. The corresponding benchmark quantity can include the ignition delay time, equilibrium temperature, laminar flame speed, and temperature in PSRs.

The current work proposes an end-to-end DNN to build up the relationship between the reduced mechanism and its accuracy in predicting benchmark quantities. The input is an N_s -dimensional Boolean vector \mathbf{x} , meaning a reduced mechanism, and output $\mathbf{u}_\theta(\mathbf{x})$ is an N_t -dimensional real vector. Each element of the output vector $\mathbf{u}_\theta(\mathbf{x})$ represents the prediction of the benchmark quantities of the reduced model. The structure is shown in Figure 2a. An L -layer neural network is,

$$\begin{aligned} \mathbf{u}_\theta(\mathbf{x}) = & \mathbf{W}^{[L-1]} \sigma \circ (\mathbf{W}^{[L-2]} \sigma \circ (\dots (\\ & \mathbf{W}^{[1]} \sigma \circ (\mathbf{W}^{[0]} \mathbf{x} + \mathbf{b}^{[0]}) + \mathbf{b}^{[1]}) \dots) \\ & + \mathbf{b}^{[L-2]}) + \mathbf{b}^{[L-1]}, \end{aligned}$$

where $\mathbf{W}^{[l]} \in \mathbb{R}^{m_{l+1} \times m_l}$, $\mathbf{b}^{[l]} \in \mathbb{R}^{m_{l+1}}$, $m_0 = N_s$, $m_L = N_t$, "o" means entry-wise operation, σ is ReLU. We denote the set of parameters by θ . Due to the intrinsic non-linearity and high-dimension of the performance function $\text{Err}(\mathbf{x})$, we use the DNN of three hidden layers, with 2000, 1000, 1000 nodes, respectively. In addition, each type of benchmark quantity requires a DNN. For example, three DNNs are trained to predict ignition delay time, temperature, and flame speed. The loss function is Mean Square Error (MSE), and the optimizer is Stochastic Gradient Descent (SGD) with batch size 128.

The well-trained DNN serves as the objective function in the inverse optimization problem. Here, the DNN training process is considered as the 'forward' optimization problem, which optimizes the parameters in the DNN. The 'inverse' means finding the optimal input \mathbf{x} given the objective function, which is in the form of a neural network. SGD is performed to find an optimal solution \mathbf{x} of the objective function

$F(\mathbf{x})$. It is worth noting that each element x_i of the original optimal solution \mathbf{x} can be any real number, but it is preferred that x_i is close to zero or one so that \mathbf{x} can be easily mapped to a reduced mechanism. Consequently, an additional Sigmoid layer is added in front of the frozen DNN. The structure of the performance function is shown in Figure 2b.

2.3. DNN-assisted sampling method

The data sampling strategy plays an essential role in algorithm implementation. The dataset size needs to be large enough to train the DNN accurately. However, it is seen that the exhaustive search of the entire input space is not feasible. For the detailed mechanism of 116 species, all possible reduced mechanisms are around 10^{34} . If labeling each sample costs one millisecond, it will take more than billions of years. As a result, an iterative DNN-assisted sampling method is proposed in this work. The procedure is illustrated in Figure 2c. In the beginning, a few samples are generated randomly with the sparsity k_0 , which means removing k_0 species randomly from the detailed mechanism. The random deleting species surely cause significant errors in some reduced mechanisms. Consequently, part of the sampled reduced mechanisms are labeled as bad reduction ones. More specifically, errors of the benchmark quantities of sampled reduced mechanisms are calculated by Cantera. The problematic reduction leads to large errors in the label. The reduced mechanisms and their errors in benchmark quantities are assembled into the dataset. The DNN training and optimization are performed accordingly. Then, new samples with higher sparsity are generated and merged into the existing dataset.

The next step is the key for reduced mechanism sampling: using the under-training DNN to evaluate new samples instead of randomly deleting species. Since the sample sparsity grows incrementally, DNN is expected to have the ability to predict the newly sampled mechanisms fast and accurately. DNN helps select only the promising reduced mechanisms to be labeled, discussed in detail in the next section.

3. Results and discussion

3.1. Sampled data distribution

The current work proposes a DNN-filter data sampling strategy. The goal is to select promising reduced mechanism samples for DNN training. Therefore, the key question is how to define 'promising' for the selection. First of all, the promising candidates can ignite. Moreover, the reduced mechanism prediction difference compared with the detailed mechanism should be relatively small. For example, the ignition delay time of the detailed mechanism and the reduced mechanism are $\tau_{detailed}$, $\tau_{reduced}$, respectively. The difference should be within two orders of magnitude, i.e., $0.01 < \tau_{reduced}/\tau_{detailed} < 100$.

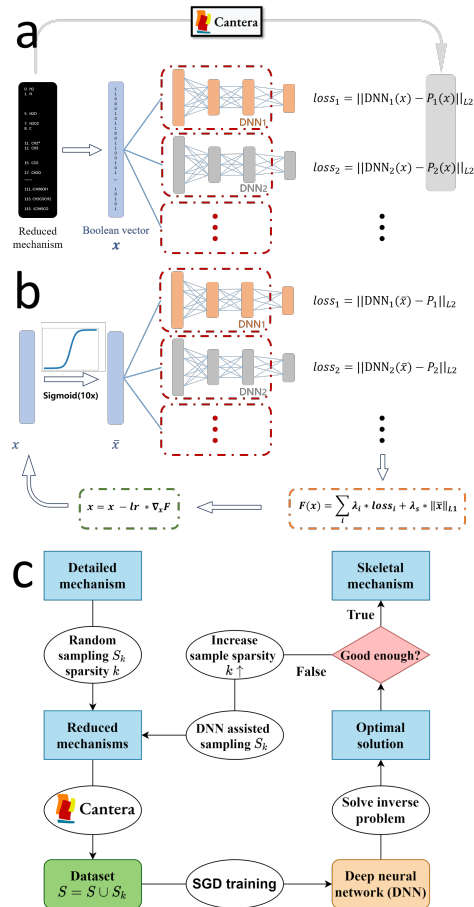


Fig. 2: (a): illustration for the data labeling and DNN structure; (b): illustration for the inverse optimization and the DNN structure with an additional Sigmoid layer; (c): the flow chart for the sampling-training iteration.

In the beginning, 8192 reduced mechanisms are randomly generated by deleting eight species. The benchmark quantities, such as ignition delay times, equilibrium temperatures, are calculated to label the sampled mechanisms. The resulting dataset is adopted to train the DNN.

In the training-sampling iteration, the sparsity of the sampled mechanism keeps increasing. Higher sparsity indicates a higher possibility that the randomly selected reduced mechanism deteriorate. Consequently, the under-training DNN is utilized to label the new reduced mechanisms. The number of newly sampled reduced mechanisms is about 16384 for each iteration. In the current work, the sparsity increases by two for each iterative step. The choice of the sparsity step size is mainly due to the under-training DNN prediction ability. Because a too large step size may lead to a large difference between the datasets of two consecutive iterations, it will potentially cause a large

DNN prediction error.

Figure 3a shows sampled data distribution and how it evolves with the iteration. The retaining ratio for the i -th species is defined as the number of reduced mechanisms containing the i -th species divided by the number of all sampled mechanisms. The dotted line is the averaged retaining ratio. At the beginning of the iteration, at sparsity $k_0 = 8$ (the blue histogram), all species have a similar retaining ratio due to random sampling. As the training and sampling proceed (the green and red histograms), the retaining ratio significantly differs for different species. The DNN tends to keep important species; thus, these species have higher chances to be selected and higher retaining ratios, such as n-heptane, O_2 , CO_2 , and H_2O .

Figure 3b shows the sampled mechanism distribution of the benchmark quantity of the ignition delay time. The x-axis indexes 30 different initial conditions. The y-axis is the ratio between the ignition delay times of the reduced mechanisms and the detailed mechanism. The number of sampled mechanisms dyes the color. It shows that most of the sampled mechanisms (red and green parts) are located near the centerline $y = 1$. In other words, most of the reduced mechanisms have ignition delay times close to the correct one. It is also important to notice the wide error distribution along the y-direction. It implies that a relatively small part of the reduced mechanisms in the dataset have ignition delay times significantly different from the result of the detailed mechanism, which helps the DNN learn the wrong patterns to avoid similar problematic reductions. This current sample-training iteration proves to be highly efficient. In total, only around 10^5 samples out of 10^{34} all possible samples are selected.

3.2. DNN training and prediction performance

At the end of the sampling-training iteration, all labeled reduced mechanisms are shuffled and assembled into a comprehensive dataset. 80% of the dataset is used for the DNN training, while the other is the test dataset. The batch size is 128. The initial learning rate is set to 0.01, which decays by 3% \sim 10% every ten epochs. The training is considered saturated when the test error increases, so-called early stopping.

Figure 4a shows the DNN prediction for a randomly selected test sample. The x-axis is the ignition delay time of the detailed mechanism calculated by Cantera. The y-axis is the ignition delay time by the DNN prediction (green circle) and Cantera (red dot) for a reduced mechanism. The selected mechanism from the test dataset shows a large deviation compared with the detailed mechanism. However, this deviation is predicted by the DNN accurately. A typical DNN training process is shown in the inset of Figure 4a. Figure 4b shows the DNN prediction error distribution on test samples. The axes are the same as Figure 3b: the x-axis is the initial condition, the y-axis is the ignition delay time error between the reduced and detailed mechanisms. The DNN prediction error determines the color. Generally speaking,

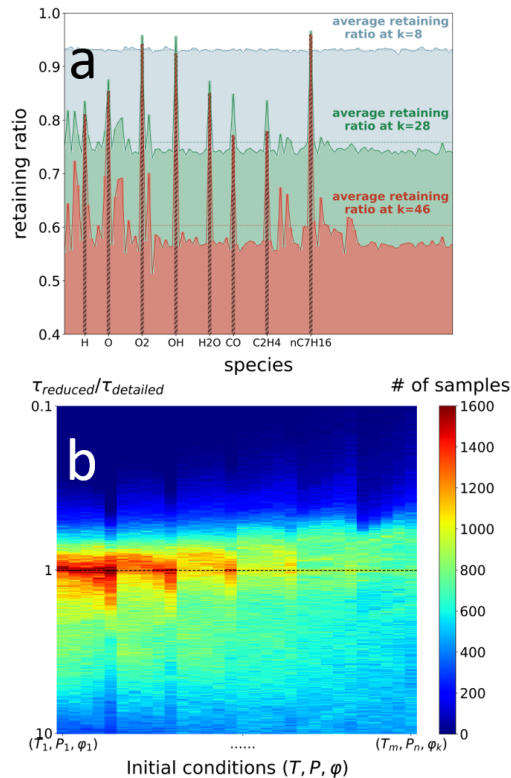


Fig. 3: (a): histogram of the retaining ratio for selected reduced mechanisms with different sparsities, the x-axis is the species index, the y-axis is the retaining ratio of species; (b): sampled data distribution. The x-axis represents the initial condition. The y-axis is the ignition delay time ratio between the reduced and detailed mechanisms.

the DNN achieves reasonable accuracy in predicting ignition delay times of reduced mechanisms in different initial conditions. Most of the prediction error is rather small, consistent with the observation in Figure 4a. Since the samples shown in Figure 4b are excluded from the DNN training, the good performance on the unseen data demonstrates the reliability of the DNN, which is the prerequisite to obtain a desirable reduced mechanism from the optimization problem.

3.3. Inverse optimization problem

Optimal reduced mechanisms are achieved by solving inverse optimization problems on the DNN input. Given the DNN function, the optimization goal is to locate the minimum point by SGD. The first step is to generate a random vector, denoted as \mathbf{x}_0 . Since we represent a reduced mechanism by a Boolean vector consisting of zero and one, we expect that the inverse problem can produce results similar to Boolean vector. Therefore, an entry-wise Sigmoid transformation $\bar{\mathbf{x}} = \text{Sigmoid}(10\mathbf{x})$ is adopted so that the each entry of the solution obtained by the inverse optimization

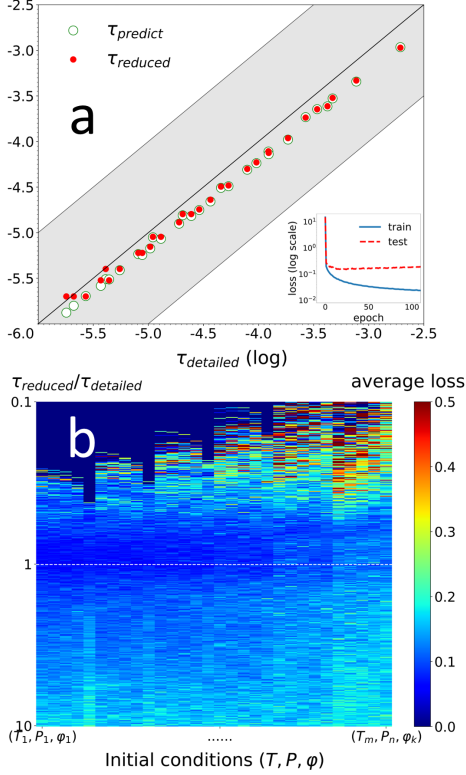


Fig. 4: (a) the comparison between DNN prediction and Cantera results of a reduced mechanism; inset: a typical DNN training process; (b) DNN prediction error on the test dataset. The x- and y-axis are the same with Figure 3b: initial conditions and ratio between reduced and detailed mechanism.

tion problem is close to zero or one.

Suppose the benchmark quantities are P_i 's, the objective function $F(\mathbf{x})$ can be written as:

$$F(\mathbf{x}) = \sum_i \lambda_i (P_i - \text{DNN}_i(\text{Sigmoid}(10\mathbf{x})))^2 + \lambda_s \|\text{Sigmoid}(10\mathbf{x})\|_{L1} \quad (2)$$

where λ_i is the weight for the benchmark quantities, λ_s is the weight for the sparsity, DNN_i is the i -th output of the DNN. In experiments, we use the same value for λ 's of the benchmark quantities of the same type, such as the 100 for ignition delay times, 10 for temperatures, 0.1 for sparsity weight. Note that $L0$ -norm is difficult to optimize, it is a conventional approach to use $L1$ -norm in optimization. The input vector is updated as:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - lr * \nabla_{\mathbf{x}_k} F(\mathbf{x}_k),$$

where lr is the learning rate.

3.4. Reduced mechanisms and validation

To demonstrate the validity of the DeePMR, ignition delay time, laminar flame speed, and extinction curves in PSRs are compared using detailed and reduced mechanisms. PFA and the DeePMR generate reduced mechanisms with species number 56 (PFA) and 45 (DNN), respectively. All the calculations are performed by the open-source Python package Cantera. The benchmark tests include different initial conditions from lean to rich mixtures at low and high pressures. The laminar flame speed, ignition delay time, and extinction curves comparison are shown for initial pressure $P = 1\text{atm}$ and $P = 10\text{atm}$ in Figures 5a, 5b, and 5c, respectively. In most benchmark tests, reduced mechanisms from the DNN and the PFA produce equally accurate results, the averaged error of ignition delay time, laminar flame speed, and PSR temperature is within 15%.

The DNN method can reduce the detailed mechanism more aggressively if the targeted scenarios are near stoichiometric conditions. Figure 6 shows the reduced mechanism with 28 species, half of the size of the PFA reduced mechanism. The only difference in performing DeePMR is to remove non-stoichiometric benchmark quantities from the DNN data labeling. The reduced mechanism shows a reasonable agreement with the detailed mechanism at $P = 1\text{atm}$ and equivalence ratio between 0.6 and 1.2.

Figure 7 shows the CPU time cost comparison using different mechanisms in perfectly stirred reactor, ideal gas reactor and laminar flame simulations. The results demonstrate that the reduced mechanism significantly accelerates various simulations. The reduced mechanism by DNN with 45 species have an averaged 3x speedup factor compared with the detailed mechanism, and 30% efficiency improvement compared with the reduced mechanism with 56 species from PFA method. The reduced mechanism with 28 species save additional 50% CPU time compared with the 45 species mechanism.

4. Conclusions

The current paper proposes a deep learning-based model reduction (DeePMR) method for simplifying combustion chemistry mechanisms. The key idea is to use an end-to-end DNN to predict a reduced mechanism accuracy. The input is the reduced mechanism, and the output is the selected benchmark quantities. The benchmark quantities can be ignition delay time, equilibrium temperature, laminar flame speed, depending on the scenario for the reduced mechanism.

The sampling is the major challenge due to the extremely large number of possible reductions. Consequently, an iterative procedure of sampling and DNN training is adopted. The iteration takes advantage of the DNN prediction to filter reduced mechanisms. Only promising candidates are retained and labeled. The sampling-training iteration generates a representative dataset efficiently.

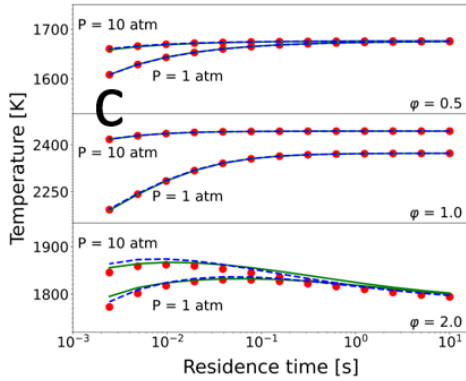
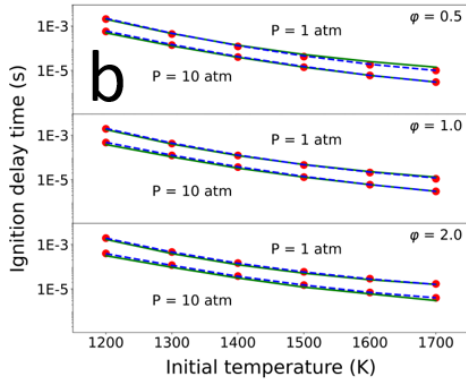
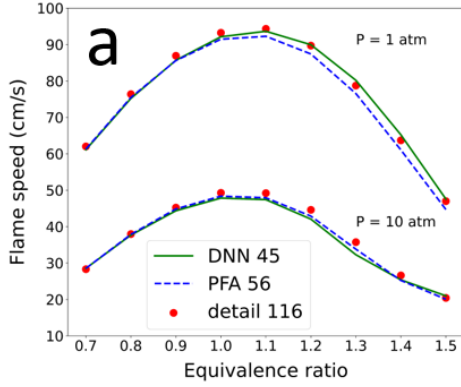


Fig. 5: Reduced mechanism (45 species) performance compared with the PFA result and the detailed mechanism. (a): laminar flame speed comparison at $P=1$ and 10 atm, equivalence ratio $\phi = 0.7 \sim 1.5$, initial temperature 500K ; (b): ignition delay time at $P=1$ and 10 atm, equivalence ratio $\phi = 0.5, 1.0, 2.0$, temperature $T = 1200 \sim 1700$ K; (c): temperature under different residence time in PSRs at $P=1$ and 10 atm, equivalence ratio $\phi = 0.5, 1.0, 2.0$, initial temperature $T = 500\text{K}$, residence time $t = 10^{-3} \sim 10$ s.

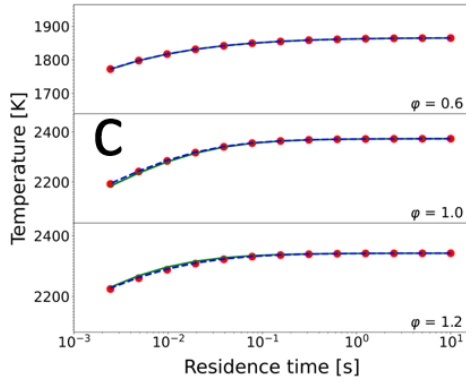
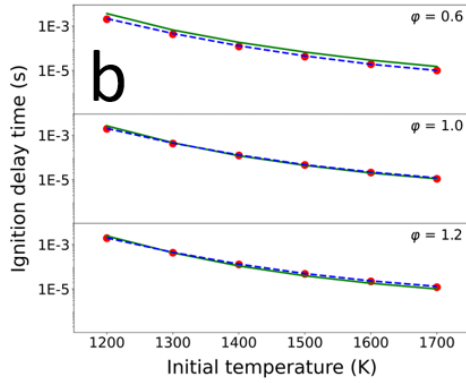
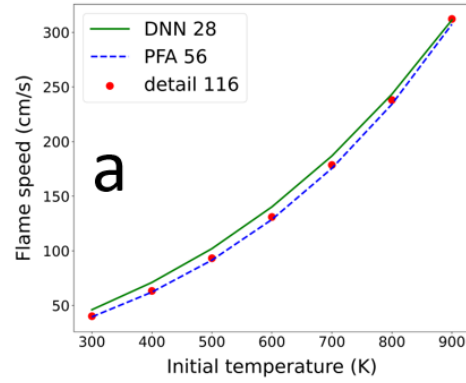


Fig. 6: Reduced mechanism (28 species) performance compared with the PFA result and the detailed mechanism. (a): laminar flame speed comparison at $P=1$ atm, equivalence ratio $\phi = 1.0$, initial temperature $T = 300 \sim 900$ K; (b): ignition delay time comparison at $P=1$ atm, equivalence ratio $\phi = 0.6, 1.0, 1.2$, initial temperature $T = 1200 \sim 1700$ K; (c): temperature comparison under different residence time in PSRs at $P=1$, equivalence ratio $\phi = 0.6, 1.0, 1.2$, initial temperature $T = 500\text{K}$, residence time $t = 10^{-3} \sim 10$ s.

The final reduced mechanism is obtained by solving an inverse optimization problem. The optimiza-

tion target is to simultaneously minimize the reduced mechanism's error and size, where the well-trained

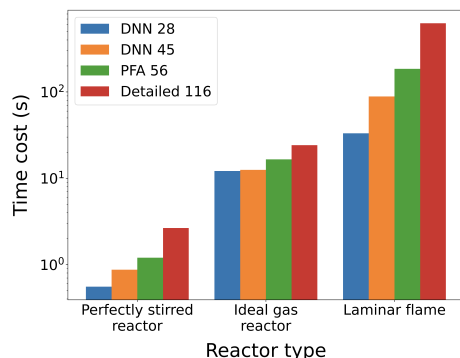


Fig. 7: Computation cost comparison among different mechanisms.

DNN provides the objective function.

The results show that the DNN reduced mechanisms have a smaller size and comparable accuracy than the path flux analysis method, which demonstrates the great potential for machine-learning methods to perform model reduction for combustion kinetics.

Acknowledgments

This work is supported by the Shanghai Sailing Program, the Natural Science Foundation of Shanghai Grant No. 20ZR1429000 (Z. X.), the National Natural Science Foundation of China Grant No. 62002221 (Z. X.), the National Natural Science Foundation of China Grant No. 12101402 (Y. Z.), Shanghai Municipal of Science and Technology Project Grant No. 20JC1419500 (Y.Z.), Shanghai Municipal of Science and Technology Major Project No. 2021SHZDZX0102 (Z. X., Y.Z.), and the HPC of School of Mathematical Sciences (Z. X., Y.Z.) and the Student Innovation Center (Z. X., Y.Z.) at Shanghai Jiao Tong University, and AI for Science Institute, Beijing.

References

- [1] L. Lu, S. B. Pope, An improved algorithm for in situ adaptive tabulation, *Journal of Computational Physics* 228 (2) (2009) 361–386.
- [2] M. Chaos, A. Kazakov, Z. Zhao, F. L. Dryer, A High-Temperature Chemical Kinetic Model for Primary Reference Fuels (2007).
- [3] H. Rabitz, M. Kramer, D. Dacol, Sensitivity Analysis in Chemical Kinetics, *Annual Review of Physical Chemistry* 34 (1) (1983) 419–461.
- [4] Turanyi, Tamas, Sensitivity analysis of complex kinetic systems. Tools and applications, *Journal of Mathematical Chemistry* 5 (3) (1990) 203–248.
- [5] N. Peters, Numerical and asymptotic analysis of systematically reduced reaction schemes for hydrocarbon flames, in: *Numerical Simulation of Combustion Phenomena*, Springer-Verlag, Berlin/Heidelberg, 1985, pp. 90–109.
- [6] M. D. Smoke, V. Giovangigli, Formulation of the premixed and nonpremixed test problems, in: *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Lecture Notes in Physics, Springer-Verlag, 1991, pp. 1–28.
- [7] S. Lam, D. Goussis, Understanding complex chemical kinetics with computational singular perturbation, *Symposium (International) on Combustion* 22 (1) (1989) 931–941.
- [8] S. H. Lam, Using CSP to Understand Complex Chemical Kinetics, *Combust. Sci. and Tech* 89 (1993) 375–404.
- [9] M. VALORANI, F. CRETA, D. GOUSSIS, J. LEE, H. NAJM, An automatic procedure for the simplification of chemical kinetic mechanisms based on CSP, *Combustion and Flame* 146 (1-2) (2006) 29–51.
- [10] J. C. Keck, Rate-controlled constrained-equilibrium theory of chemical reactions in complex systems, *Progress in Energy and Combustion Science* 16 (2) (1990) 125–154.
- [11] T. Løvås, D. Nilsson, F. Mauss, Automatic reduction procedure for chemical mechanisms applied to premixed methane/air flames, *Proceedings of the Combustion Institute* 28 (2) (2000) 1809–1815.
- [12] T. Lu, C. K. Law, A directed relation graph method for mechanism reduction, *Proceedings of the Combustion Institute* 30 (1) (2005) 1333–1341.
- [13] P. Pepiot-Desjardins, H. PITSCH, An efficient error-propagation-based reduction method for large chemical kinetic mechanisms, *Combustion and Flame* 154 (1-2) (2008) 67–81.
- [14] W. Sun, Z. Chen, X. Gou, Y. Ju, A path flux analysis method for the reduction of detailed chemical kinetic mechanisms, *Combustion and Flame* 157 (7) (2010) 1298–1307.
- [15] K. E. Niemeyer, C. J. Sung, M. P. Raju, Skeletal mechanism generation for surrogate fuels using directed relation graph with error propagation and sensitivity analysis, *Combustion and Flame* 157 (9) (2010) 1760–1770. [arXiv:1607.05079](https://arxiv.org/abs/1607.05079),
- [16] Y. Wu, Y. Liu, T. Lu, A linearized error propagation method for skeletal mechanism reduction, *Combustion and Flame* 211 (2020) 303–311.
- [17] Y. Chen, J. Y. Chen, Towards improved automatic chemical kinetic model reduction regarding ignition delays and flame speeds, *Combustion and Flame* 190 (2018) 293–301.
- [18] N. J. Curtis, K. E. Niemeyer, C. J. Sung, An automated target species selection method for dynamic adaptive chemistry simulations, *Combustion and Flame* 162 (4) (2015) 1358–1374. [arXiv:1804.01591](https://arxiv.org/abs/1804.01591),
- [19] J. Zeng, L. Cao, M. Xu, T. Zhu, J. Z. Zhang, Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation, *Nature Communications* 11 (1) (2020) 1–9.
- [20] L. Zhang, J. Han, R. Car, Deep Potential Molecular Dynamics : A Scalable Model with the Accuracy of Quantum Mechanics, *Physical Review Letters* 120 (14) (2018) 143001.
- [21] J. Si, G. Wang, P. Li, J. Mi, A new skeletal mechanism for simulating MILD combustion optimized using Artificial Neural Network, *Energy* 237 (2021) 121603.
- [22] M. Kelly, S. Dooley, G. Bourque, Toward machine learned highly reduced kinetic models for methane/air combustion, *Proceedings of the ASME Turbo Expo 3A-2021* (2021) 1–12. [arXiv:2103.08377](https://arxiv.org/abs/2103.08377),
- [23] B. Bhattacharjee, D. A. Schwer, P. I. Barton, W. H.

Green, Optimally-reduced kinetic models: Reaction elimination in large-scale kinetic mechanisms, *Combustion and Flame* 135 (3) (2003) 191–208.