·Advanced Interdisciplinary Science·

Excited state reaction kinetics regression based on sequence-to-sequence learning*

Bai Tianzi[1,2]， Huai Ying[1,3]， Liu Tingting[1]， Jia Shuqin[1,3]， Duo Liping[1,3]

（1. *Key Laboratory of Chemical Lasers*, *Dalian Institute of Chemical Physics*, *Chinese Academy of Sciences*, *Dalian* 116023, *China*;

2. *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*;

3. *State Key Laboratory of Molecular Reaction Dynamics*, *Dalian Institute of Chemical Physics*,

Chinese Academy of Sciences, *Dalian* 116023, *China*）

Abstract: [**Background**] The reaction kinetics in lasers often involves a lot of excited state species. The mutual effects and numerical stiffness arising from the excited state species pose significant challenges in numerical simulations of lasers. The development of artificial intelligence has made neural networks (NNs) a promising approach to address the computational intensity and instability in excited state reaction kinetics (ESRK). [**Purpose**] However, the complexity of ESRK poses challenges for NN training. These reactions involve numerous species and mutual effects, resulting in a high-dimensional variable space. This demands that the NN possess the capability to establish complex mapping relationships. Moreover, the significant change in state before and after the reaction leads to a broad variable space coverage, which amplifies the demand for NN's accuracy. [**Methods**] To address the aforementioned challenges, this study introduced successful sequence-to-sequence learning from large language learning into ESRK to enhance prediction accuracy in complex, high-dimensional regression. Additionally, a statistical regularization method was proposed to improve the diversity of the outputs. NNs with different architectures were trained using randomly sampled data, and their capabilities were compared and analyzed. [**Results**] The proposed method is validated using a vibrational reaction mechanism for hydrogen fluoride, which involves 16 species and 137 reactions. The results demonstrate that the sequential model achieves lower training loss and relative error during training. Furthermore, experiments with different hyperparameters reveal that variation in the random seed can significantly impact model performance. [**Conclusions**] In this work, the introduction of the sequential model successfully reduced the parameter count of the conventional wide model without compromising accuracy. However, due to the intrinsic complexity of ESRK, there remains considerable room for improvement in NN-based regression tasks for this domain.

Key words: excited state, reaction kinetics, sequence-to-sequence learning, complexity

CLC number: TN248.5 **Document code:** A **doi**: 10.11884/HPLPB202638.250298

Chemical lasers represent a typical class of high-power lasers[1]. Numerical simulations of such systems provide valuable guidance for device optimization and design, thereby reducing research and development costs. The reaction kinetics in the optical cavity usually includes various excited states, which significantly affect key performance parameters such as power density and spectral line distribution, making it a critical process in laser numerical simulation[2]. However, the complex interactions between excited state species and the numerical stiffness caused by disparate reaction rates pose considerable challenges to the calculation of excited state reaction kinetics (ESRK). For instance, each additional species in the system requires solving an additional species transport equation, leading to increased computational costs and difficulty in simulation convergence[3].

In recent years, benefiting from the abundance of data and the reduction of hardware costs, The adoption of the neural

* **Received date:** 2025-09-01； **Revised date:** 2025-12-16

Foundation item: supported by National Key R&D Program of China (2024YFB4006600); Research Foundation (232-CXCY-A01-09-05-01); Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0970204)

Biography: Bai Tianzi, btzan@dicp.ac.cn。

Corresponding author: Huai Ying, huaiying@dicp.ac.cn。

networks (NNs) has gained traction in the research of computational fluid dynamics (CFD) with chemical reactions[4], particularly due to the NN's capacity to enhance the computational efficiency without reduction of the governing equations[5]. However, for engineering practice, the complexity of relationships[6] and the shortage of data[7] restrict the application of NNs.

For the NN driven by data, the curse of dimensionality presents exponentially growing parameter space related to the number of species[8]. This issue is exacerbated by the species mutual effects that resist conventional dimensionality reduction techniques such as convolution and pooling operations[9]. The results from CFD simulations are usually employed as databases for the NN training, while this approach turns out to be prohibitive computational costs[10]. Moreover, the trained data-driven models exhibit confined predictive accuracy within the training parameter space, demonstrating limited generalization capability when designing new configurations[11].

Meanwhile, the species mutual effects in ESRK lead to the kinetics being sensitive to the changes in the flow state as the inputs of the NN and the outputs often span multiple orders of magnitude[12]. Although there have been some attempts to introduce normalization methods to enable the NN to learn the reaction kinetics, the processes of normalization and denormalization will result in certain information loss[13].

To address these challenges, a statistical NN framework is proposed to integrate the data distribution characteristics into the training process, thereby enhancing data utilization efficiency. The existing approaches like variance controlled NNs have demonstrated improved performance by incorporating the output variance regularization into loss functions for multi-output problems[14]. However, they have not exploited the statistical correlations between data points, which the statistical NN framework takes into account[15]. At the same time, our training process indicates that the statistical regularization induces oscillatory gradient behavior and aggravates the impact of stochasticity on NN's performance.

Beyond the challenges of data acquisition, in typical ESRK systems such as chemical lasers and combustion, there are high-dimensional complex correlations resulting from the interactions among the species[16]. Such high-dimensional correlations are difficult to be separated in a simple manner[17]. Moreover, characteristic processes such as combustion and relaxation correspond to different locations within the parameter space[18]. This results in the data of ESRK being non-uniformly distributed across the parameter space, with aggregation occurring in regions associated with the reaction pathways. The combined effects of high dimensionality and localized aggregation create a paradox where the parameter space exhibits vast dimensions while containing only limited regions of meaningful data[19]. This phenomenon imposes stringent demands on the complex learning capacity of NNs[20]. To mitigate this difficulty, we adapt sequence-to-sequence learning from natural language processing, which improves the complexity of the NN and separates the learning modules of each output partly[21]. This architecture enables the establishment of sophisticated associative mappings while maintaining a compact parameter budget[22].

This article is organized as follows. Section 1 delineates the proposed methodology, with its implementation outcomes examined in Section 2 through a chemical mechanism involving 16 species and 137 reactions. The summary and prospect are articulated in Section 3.

1 Methodology

1.1 Reaction kinetics

Reaction kinetics characterizes the spatiotemporal evolution of species composition in reaction systems. The species transport governing equation can be formulated as

$$\frac{\partial}{\partial t}(\rho Y_i) + \nabla \cdot (\rho Y_i \boldsymbol{u}) = -\nabla \cdot \boldsymbol{J}_i + \dot{\omega}_i \tag{1}$$

where Y_i denotes the mass fraction of the i-th species, ρ represents the gas mixture density, $\boldsymbol{u}$ signifies the velocity vector, $\dot{\omega}_i$ expresses the chemical source term governed by Arrhenius kinetics, and $\boldsymbol{J}_i$ is the diffusion flux of the i-th species. This partial differential equation describes the coupled processes of fluid dynamics, mass transport, and chemical reactions.

The reaction processes occurring in chemical lasers exemplify a prototypical ESRK. Illustrated by the reactions that occurred in Hydrogen Fluoride (HF) chemical laser[23]:

Vibration-Translation (V-T) Collision Relaxation

$$\mathrm{HF}(\upsilon)+M\rightarrow\mathrm{HF}(\upsilon-1)+M \tag{2}$$

Vibration-Vibration (V-V) Energy Exchange

$$\mathrm{HF}(\upsilon)+\mathrm{HF}(\upsilon')\rightarrow\mathrm{HF}(\upsilon+1)+\mathrm{HF}(\upsilon'-1) \tag{3}$$

where M denotes collision partners mediating vibration-to-translation energy dissipation and υ represents the vibration level of excited HF.

The preceding examples reveal pronounced mutual effects resulting in prohibitive computational demands, motivating the present investigation to introduce the NN to decrease the computational complexity. Notably, in the species transport governing equation Eq. (1), the chemical source term is not directly related to the convection and diffusion processes, but is exclusively determined by the instantaneous flow state (temperature, pressure and species composition, T, p, Y_i)

$$\dot{\omega}_i = MW_i\sum_{j=1}^{L}\left(\upsilon''_{ij}-\upsilon'_{ij}\right)q_j \tag{4}$$

$$q_j = k_{fj}\prod_{i=1}^{N}[C_i]^{\nu'_{ij}} - k_{rj}\prod_{i=1}^{N}[C_i]^{\nu''_{ij}} \tag{5}$$

$$C_i = \frac{n_i}{V} = \frac{p_i}{RT} = \frac{PY_i\overline{MW}/MW_i}{RT} \tag{6}$$

$$k_j = AT^b\exp\left(-\frac{E_a}{RT}\right) \tag{7}$$

where $\overline{MW}$ represents the molecular weight of the mixture, MW_i represents the molecular weight of i-th species, ν is the stoichiometric coefficient, q_j denotes the molar rate of j-th reaction, k_j represents the rate constant of j-th reaction, C_i expresses the molar concentration of i-th species, n_i is the amount of substance of i-th species, V signifies the volume, p_i denotes the partial pressure of i-th species, R is the universal gas constant, A presents the pre-exponential factor of the Arrhenius formula, b is the temperature exponent and E_a denotes the activation energy.

This decoupling characteristic inspires the proposed methodology. The chemical source terms corresponding to flow states are predicted by the NN to decrease the computational complexity.

1.2 Sequence-to-sequence learning

A fundamental challenge in ESRK regression lies in its high-dimension. The dimensions of flow states and chemical source terms, as the NN's inputs and outputs, are positively correlated with the number of species in the reaction mechanism. Moreover, the chemical kinetic equations Eq. (4)—Eq. (7) demonstrate that the chemical source term depends on all participating species, preventing complexity reduction through variable decoupling. This multivariate dependence imposes stringent demands on the capacity of the NN.

To address this challenge, we adapt sequence-to-sequence learning from recent breakthroughs in natural language processing[21]. Nature language is usually regarded as a high-dimensional system. Considering its dimension, the sequence-to-sequence learning first processes the input text through an encoder to generate a hidden context victor. Then, it decodes this vector through a loop of decoders to generate the output tokens. This paradigm captures the correlations within inputs while enabling relatively independent regression of the multiple outputs. The proposed NN structure implementing this paradigm is illustrated in Fig.1.

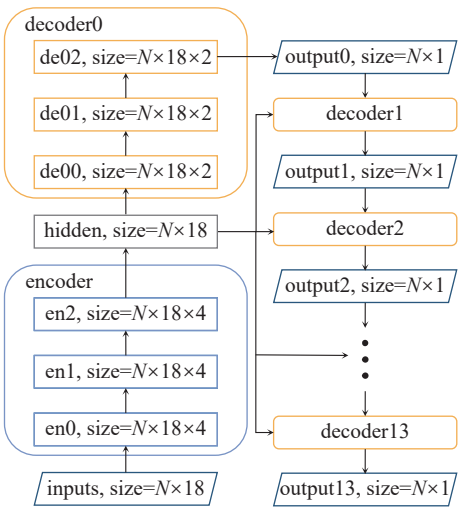

Fig. 1 Structure of proposed neural network

As illustrated in Fig.1, the input flow states are firstly mapped by an encoder module to a hidden vector. Each serial decoder then subsequently generates its assigned chemical source term using both the hidden vector and, when available, output from the prior

decoder as inputs. In contrast to conventional structures that directly predicts all source terms through a single wide layer, this sequence-to-sequence structure achieves two synergistic benefits: parameter reduction through serial decoders with fewer interactions, and partial decoupling of outputs — maintaining essential mutual effects correlations while allowing computationally tractable independence. These features improve the NN's capability to resolve the high-dimensional regression task.

1.3 Statistical neural network

During the process of training the NN to achieve the regression of chemical source terms, our observations reveal two critical characteristics of ESRK: the parameter space exhibits a high dimensionality which is proportional to the number of species, and the meaningful data are concentrated along the reaction pathway. These characteristics result in sparse data distribution in the parameter space, where the chemical source terms (outputs of the NN) demonstrate sensitivity to numerical variations in flow states (inputs of the NN).

The challenges of sparsity and sensitivity undermine the establishment of the regression relationships by the NN. Typical manifestations include the NN either collapsing into uniform predictions across all inputs or exhibiting numerical oscillations during the training.

To address these challenges in the regression of ESRK, we propose a statistical NN framework with a statistical regularization in the loss function

$$L_{\mathrm{s}} = L + C\Big/\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{8}$$

where the L denotes the original loss function, which in this study is set as the mean average error (MAE); C is a constant whose effect is evaluated in the Section 2.3.

In the NN training process, an epoch refers to one complete iteration through the entire training database. A batch segmentation strategy is customarily implemented, whereby the database is partitioned into multiple subsets for sequential training processing. Each subset is called a batch, and all batches together complete one epoch of the NN training. This framework introduces the standard deviation of each output in a batch to the loss function as a regularizer. That is, from multiple outputs, extracting one specific output and calculating the standard deviation of it in a batch. The regularization is implemented by the weighted average reciprocal of the standard variances of all outputs. The application outcomes and analyses of the proposed framework are presented in Section 2.

2 Results

2.1 Setup of application

To evaluate the proposed methodology's efficacy in ESRK regression, the vibrational mechanism of HF chemical lasers with 16 species and 137 reactions is selected as the demonstration system[23]. The NNs are trained with the open-source deep learning framework TensorFlow[24] on an NVIDIA Tesla V100s. The flow states as the NN's inputs are generated via uniform random sampling within predefined parameter space and their ranges are specified in Table 1. The chemical source terms are calculated through an open-source chemical kinetics software, Cantera[25]. To align the generated data with realistic reaction pathways, all reaction source terms are constrained below 1.0×10^{6} kg/(m^3·s) through explicit thresholding during database generation.

Table 1 Ranges of flow states as the neural network's inputs

name	T / K	P / pa	$Y_{\mathrm{H_2}}$	$Y_{\mathrm{H_2(1)}}$	Y_{H}	$Y_{\mathrm{F_2}}$	Y_{F}	Y_{DF}	$Y_{\mathrm{HF(0)}}$
lower limit	3.0×10^{1}	1.0×10^{2}	0	0	0	0	0	0	0
upper limit	1.0×10^{3}	3.5×10^{3}	5.0×10^{-1}	5.0×10^{-3}	1.2×10^{-2}	3.0×10^{-3}	3.0×10^{-1}	5.0×10^{-1}	2.0×10^{-1}
name	$Y_{\mathrm{HF(1)}}$	$Y_{\mathrm{HF(2)}}$	$Y_{\mathrm{HF(3)}}$	$Y_{\mathrm{HF(4)}}$	$Y_{\mathrm{HF(5)}}$	$Y_{\mathrm{HF(6)}}$	$Y_{\mathrm{HF(7)}}$	Y_{He}	$Y_{\mathrm{N_2}}$
lower limit	0	0	0	0	0	0	0	0	0
upper limit	7.5×10^{-2}	1.5×10^{-1}	4.0×10^{-2}	3.0×10^{-4}	2.0×10^{-4}	1.5×10^{-4}	1.0×10^{-4}	9.0×10^{-1}	2.0×10^{-1}

The training process proceeds as follows. First, 300 000 flow states are uniformly randomly sampled and the corresponding chemical source terms are calculated, forming a database with 20% data reserved for testing. Then, the remaining training data is batched (size=20 000) for 200 epochs of NN training, the validation data is composed of 20% training data and the performance of trained NN is evaluated by the reserved test data. The parameters of NN are updated by back-propagating gradients through the Adam optimizer during every batch. Finally, the mean absolute relative error (MARE), Eq. (9) and the mean squared relative error (MSRE), Eq. (10) are recorded

$$\mathrm{MARE} = \sum\nolimits_N \left| \left(x - x_{\mathrm{pre}} \right) / x \right| \tag{9}$$

$$\mathrm{MSRE} = \sqrt{\sum\nolimits_N \left(x - x_{\mathrm{pre}} \right)^2 \Big/ \sum\nolimits_N (x)^2} \tag{10}$$

where the subscript pre means the prediction of the NN.

The training process mentioned above constitutes one iteration. Upon completing one iteration, new sampling is performed for subsequent iterations until reaching the preset count. The random seed is fixed to guarantee that all the NNs are trained with the same database.

2.2 Effect of proposed algorithm

The proposed methodology is subsequently evaluated on an ESRK regression task. Following the setup in Section 2.1, both sequential and wide NNs were trained. The wide NN architecture connects the serial decoders in parallel to constitute a wide layer, where all source terms are predicted simultaneously rather than sequentially. The effects of different loss functions, the traditional MAE loss and MAE with statistical regularization, were also evaluated. The NN employing statistical regularization is named with a postfix “Std”. The configurations of different NNs are listed in Table 2, where we can see that the sequential models have only 12.34% of the total parameters compared to wide models. The training losses for different NNs are shown in Fig.2.

Table 2 The configurations of different neural networks

name	structure	loss	total parameters
Seq	serial decoders	MAE	57 631
SeqStd	serial decoders	MAE+C	57 631
Wide	parallel decoders	MAE	467 167
WideStd	parallel decoders	MAE+C	467 167

The NNs' MSRE of test data are shown in Fig.3. The error is calculated by Eq. (10) where the x_{pre} is the chemical source terms predicted by the NN and the x is calculated by reaction kinetics Eq. (4)—Eq. (7). For the sequential models (solid lines), their error curves evolve similarly with their loss trends, while the wide models (dashed lines) remain near 100% relative error. This indicates that the sequential models have learned a more accurate relationship between the flow states and chemical source terms, achieving better generalization to unseen test data compared to the wide models. Meanwhile, NNs with statistical regularization are denoted by triangular markers and the NNs with traditional MAE loss function are denoted by circular markers. Analysis reveals that the improvement of statistical regularization in the sequential model is marginal but for the wide model, the regularization could provide about 20% reduction of the test errors. This phenomenon originates from the conventional NNs' scarce capacity to accommodate widely distributed data, often collapsing into local minima that produce homogeneous outputs across inputs. The introduced statistical regularization mitigates this collapse by introducing the penalty of the output similarity, thereby enhancing prediction accuracy.

Subsequent evaluation has been conducted using 5 000 random sampling data. Fig.4(a) displays the relative errors of different NNs on data sampling from whole parameter space, where all models demonstrate unsatisfactory accuracy while the sequential models maintain lower relative errors compared to wide models. Fig.4(b) presents the results of data sampling from the parameter space within the constrained (0.6, 0.65) range of flow states. As can be seen, in this situation the MARE of all NNs decreased to the value close to the MSRE, which demonstrates the competence of NNs in the regression task. In the

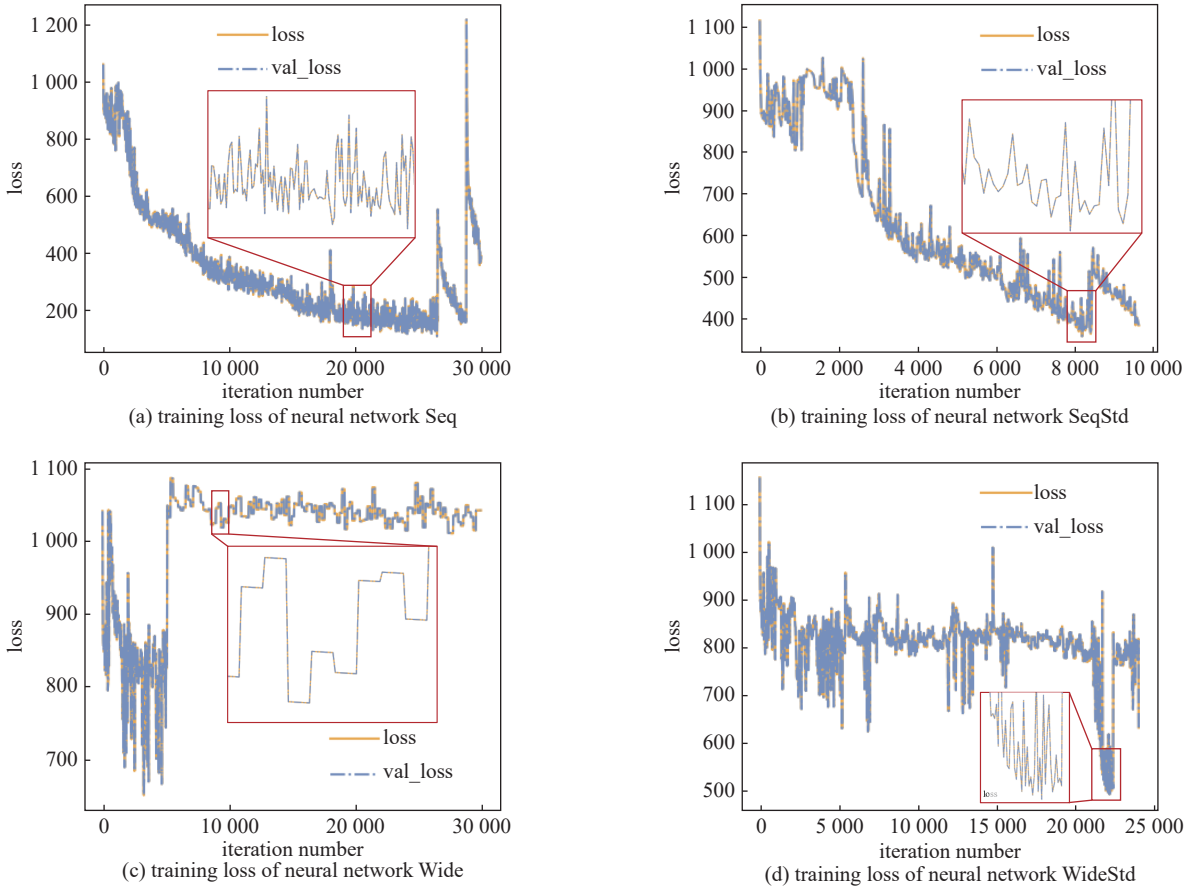

Fig. 2 Training losses for different neural networks

meantime, the relative errors of sequential models are around 20%, indicating that the NNs can achieve effective prediction within this range. However, for the wide models, the errors of NN with traditional loss, Wide, have decreased while the errors of NN with statistical regularization, WideStd, increased. Given identical training data ensured through fixed random seeds, this outcome shows differences in the regression capacity of NNs. While sequential models demonstrate localized accuracy enhancement by successfully establishing regression mappings within subspace domains, they exhibit limitations in global prediction due to the complexity arising from species mutual effects in ESRK.

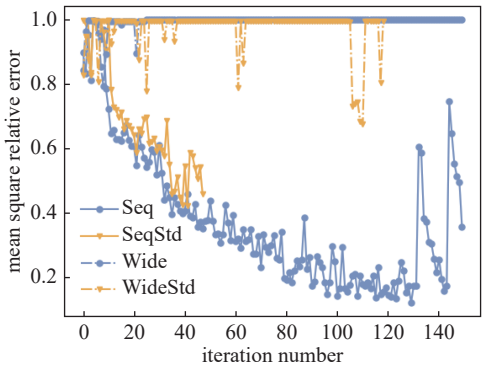

Fig. 3 Mean square relative errors on test data of different neural networks

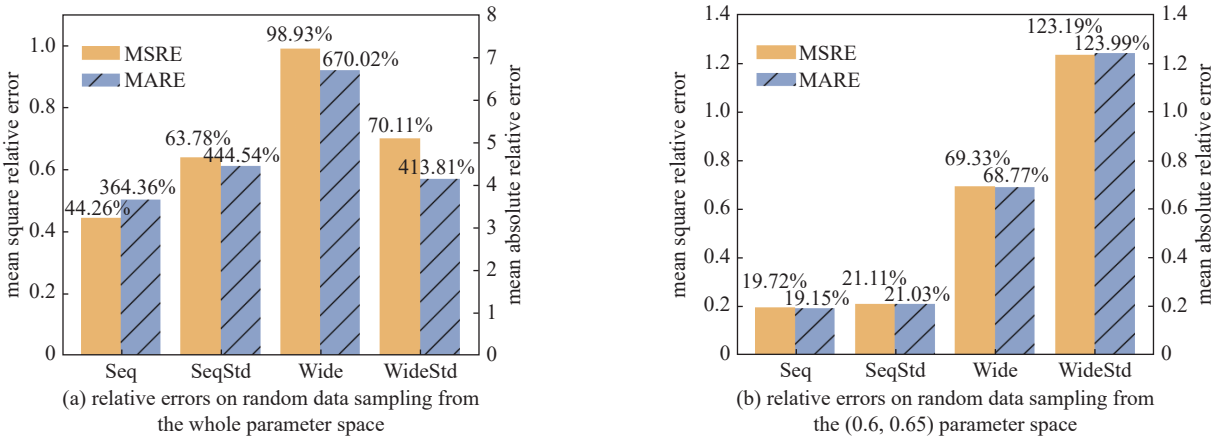

Fig. 4 Relative errors on random sampling data of different neural networks

2.3 Influence of hyper-parameter

The results discussed in Section 2.2 were obtained with the constant C=500 in Eq. (8). The influences of C values are further investigated in this section. The relative errors of NNs with different structures on the data random sampling from the whole parameter space are shown in Fig.5.

The Fig.5 reveals different stability characteristics of the sequential models and wide models: the sequential models (solid lines) demonstrate relatively small perturbation in MARE and MSRE across C variations, whereas the wide models (dashed lines) exhibit pronounced fluctuations. The quantitative evaluation shows the sequential models' MSRE variability is 63.54% of the wide models' amplitude, with MARE variability further reduced to 10.02%, verifying the numerical stability of sequential models.

To account for the impacts of stochasticity, particularly in training data distribution and NN initialization, additional validation trials under different random seeds were conducted. Fig.6 illustrates test error trajectories during training, Fig.7 presents relative errors on resampled data and Fig.8 details performance variations with regularization constant adjustments.

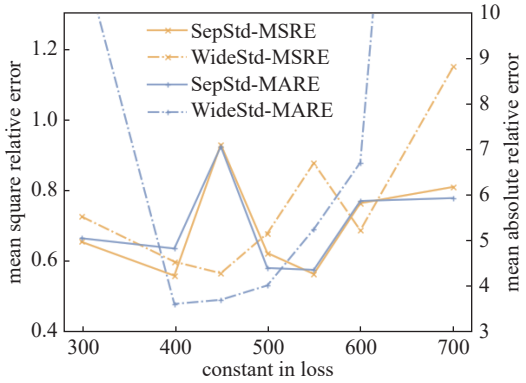

Fig. 5 Relative errors of neural networks with different constant in regularization

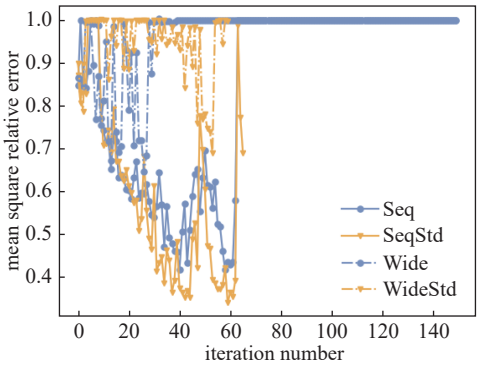

Fig. 6 Mean square relative errors on test data of neural networks trained with another random seed

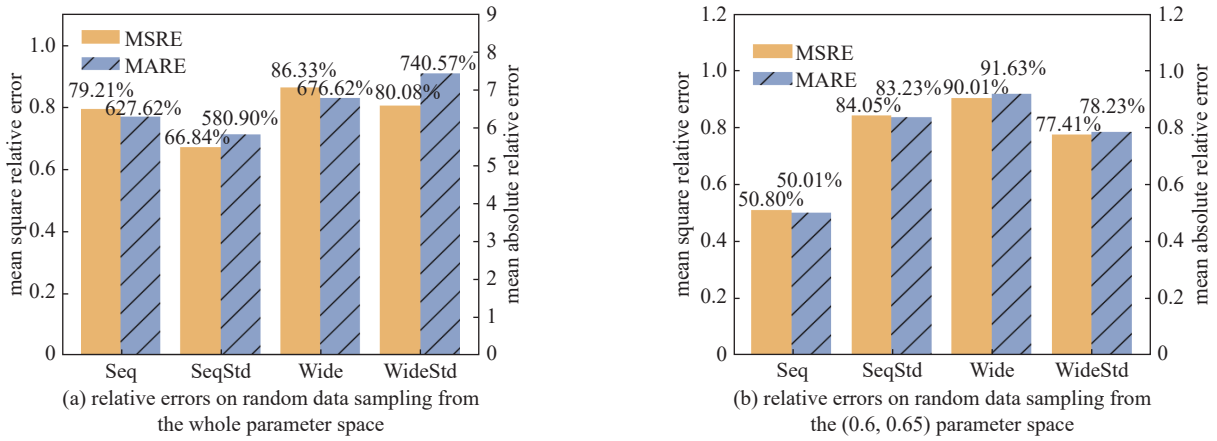

Fig. 7 Relative errors on random sampling data of neural networks trained with another random seed

Comparative analysis of Fig.6-8 against Fig.3-5 reveals that the performance of the sequential model remains slightly better than that of the wide model after changing the random seed. However, divergent characteristics emerge. The model exhibited degraded accuracy within the (0.6, 0.65) parameter space compared to the prior random seed, suggesting that the sparser sampling density in this subspace constrains the NNs' capacity. This further demonstrates the complexity of kinetic regression and its sensitivity to sampling data. Fig.8 reveals altered response patterns to regularization constant variations compared to Fig.5 and this phenomenon is expected.

This behavioral divergence stems from stochasticity: random seeds induce variations in both NN initialization states and training data distributions, thereby modifying the influence of regularization constant. These observations reinforce two aspects of ESRK regression: (1) complexity arising from high-dimensional species mutual effects, and (2) sensitivity to both

algorithmic initialization and database distribution.

3 Conclusion

This study investigates the regression for ESRK, yielding three principal findings.

(1) Complexity and sensitivity: Species mutual effects induce the complexity in the reaction kinetics and the sensitivity of the chemical source terms to the flow states, fundamentally complicating the ESRK regression tasks. These manifest as error amplification when extrapolating beyond training domains.

(2) Proposed algorithm advantages: The proposed sequential model achieves a reducing parameter count by 12.34% compared to the wide model while enhancing prediction accuracy.

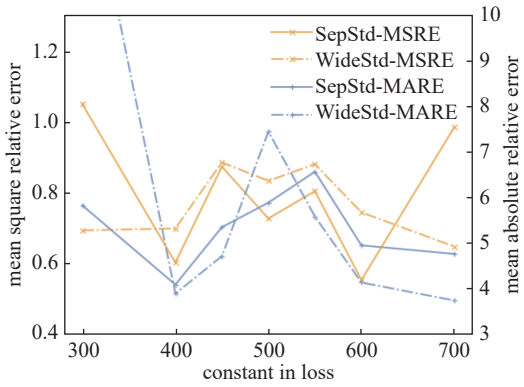

Fig. 8 Relative errors of different constant regularized neural networks trained with another random seed

(3) Stochasticity influence: The high-dimensional parameter space necessitates comprehensive database coverage, yet sampling stochasticity causes non-negligible performance variance.

Future work will integrate reinforcement learning frameworks to decouple the dependency on training data distribution, aiming for high-fidelity predictions across an extended parameter space with reduced data requirements.

References:

[1] Waichman K, Barmashenko B D, Rosenwaks S. Comparing modeling and measurements of the output power in chemical oxygen-iodine lasers: a stringent test of I2 dissociation mechanisms[J]. The Journal of Chemical Physics, 2010, 133: 084301.

[2] Li Hui, Zhao Tianliang, Li Jiaxu, et al. State-to-state chemical kinetic mechanism for HF chemical lasers[J]. Combustion Theory and Modelling, 2020, 24(1): 129-141.

[3] D'Alessio G, Sundaresan S, Mueller M E. Automated and efficient local adaptive regression for principal component-based reduced-order modeling of turbulent reacting flows[J]. Proceedings of the Combustion Institute, 2023, 39(4): 5249-5258.

[4] Kochkov D, Smith J A, Alieva A, et al. Machine learning–accelerated computational fluid dynamics[J]. Proceedings of the National Academy of Sciences of the United States of America, 2021, 118: e2101784118.

[5] Ding Tianjie, Readshaw T, Rigopoulos S, et al. Machine learning tabulation of thermochemistry in turbulent combustion: an approach based on hybrid flamelet/random data and multiple multilayer perceptrons[J]. Combustion and Flame, 2021, 231: 111493.

[6] Ortega A G, Shirin A. Neural network-based descent control for Landers with sloshing and mass variation: a cascade and adaptive PID strategy[J]. Aerospace, 2024, 11: 1009.

[7] Zhang Shihong, Zhang Chi, Wang Bosen. CRK-PINN: a physics-informed neural network for solving combustion reaction kinetics ordinary differential equations[J]. Combustion and Flame, 2024, 269: 113647.

[8] Hughes G. On the mean accuracy of statistical pattern recognizers[J]. IEEE Transactions on Information Theory, 1968, 14(1): 55-63.

[9] Han Peilun, Shen Xiaoqian, Shen Boxiong. A simulation study on NOx reduction efficiency in SCR catalysts utilizing a modern C3-CNN algorithm[J]. Fuel, 2024, 363: 130985.

[10] Ihme M, Chung W T, Mishra A A. Combustion machine learning: principles, progress and prospects[J]. Progress in Energy and Combustion Science, 2022, 91: 101010.

[11] Shin J, Hansinger M, Pfitzner M, et al. A priori analysis on deep learning of filtered reaction rate[J]. Flow, Turbulence and Combustion, 2022, 109(2): 383-409.

[12] Zhang Tianhan, Yi Yuxiao, Xu Yifan, et al. A multi-scale sampling method for accurate and robust deep neural network to predict combustion chemical kinetics[J]. Combustion and Flame, 2022, 245: 112319.

[13] Bai Tianzi, Huai Ying, Liu Tingting, et al. Acceleration of the complex reacting flow simulation with a generalizable neural network based on meta-learning[J]. Fuel, 2024, 372: 132173.

[14] Kretzschmar R, Karayiannis N B, Eggimann F. Feedforward neural network models for handling class overlap and class imbalance[J]. International Journal of Neural Systems, 2005, 15(5): 323-338.

[15] Baskerville N P, Granziol D, Keating J P. Appearance of Random Matrix Theory in deep learning[J]. Physica A: Statistical Mechanics and its Applications, 2022, 590: 126742.

[16] Zhang Yu, Du Wenli. Intelligent time-scale operator-splitting integration for chemical reaction systems[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(8): 3366-3376.

［17］ Marcato A, Santos J E, Boccardo G, et al. Prediction of local concentration fields in porous media with chemical reaction using a multi scale convolutional neural network［J］. Chemical Engineering Journal, 2023, 455: 140367.

［18］ Sun Jie, Wang Yiqing, Tian Baolin, et al. DetonationFoam: an open-source solver for simulation of gaseous detonation based on OpenFOAM［J］. Computer Physics Communications, 2023, 292: 108859.

［19］ Salunkhe A, Deighan D, DesJardin P E, et al. Physics informed machine learning for chemistry tabulation［J］. Journal of Computational Science, 2023, 69: 102001.

［20］ Xu Kailai, Darve E. Physics constrained learning for data-driven inverse modeling from sparse observations［J］. Journal of Computational Physics, 2022, 453: 110938.

［21］ Sutskever I, Vinyals O, Le Q V. Sequence to Sequence Learning with Neural Networks［C］//Proceedings of the 28th International Conference on Neural Information Processing Systems. 2014: 3104-3112.

［22］ Liu Bowen, Ramsundar B, Kawthekar P, et al. Retrosynthetic reaction prediction using neural sequence-to-sequence models［J］. ACS Central Science, 2017, 3(10): 1103-1113.

［23］ Manke II G C, Hager G D. A review of recent experiments and calculations relevant to the kinetics of the HF laser［J］. Journal of Physical and Chemical Reference Data, 2001, 30(3): 713-733.

［24］ Abadi M, Barham P, Chen Jianmin, et al. TensorFlow: a system for large-scale machine learning［C］//Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation. 2016: 265-283.

［25］ Goodwin D G, Moffat H K, Speth R L. Cantera: an object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes［R］. Version 2.2. 0, 2015.

基于序列学习的激发态反应动力学回归

白天滋[1,2]， 怀 英[1,3]， 刘婷婷[1]， 贾淑芹[1,3]， 多丽萍[1,3]

（1. 中国科学院 大连化学物理研究所 化学激光重点实验室，辽宁 大连 116023； 2. 中国科学院大学，北京 100049；
3. 中国科学院 大连化学物理研究所 化学反应动力学全国重点实验室，辽宁 大连 116023）

摘 要： 激光器中的反应动力学常包含大量激发态物种。激发态物种之间的相互作用与由此导致的数值刚性是激光器数值模拟的一大挑战。通过神经网络建立激发态反应动力学关系回归可有效降低计算复杂度，为更加准确精细的激光器数值模拟提供可能。但激发态反应动力学的复杂性同样要求神经网络具有较强的回归性能。引入了序列神经网络来在较低参数量的前提下提升网络复杂回归的能力，同时提出了统计网络框架来进一步增加网络输出的多样性。所提出的方法在包含 16 个物种和 137 个反应的氟化氢振动态反应机理中进行了验证。在验证过程中，同时发现了随机性对网络性能的影响。

关 键 词： 激发态；反应动力学；序列学习；复杂性