Compact Representation of a Multi-dimensional Combustion Manifold Using Deep Neural Networks

UNIVERSITY OF WATERLOO

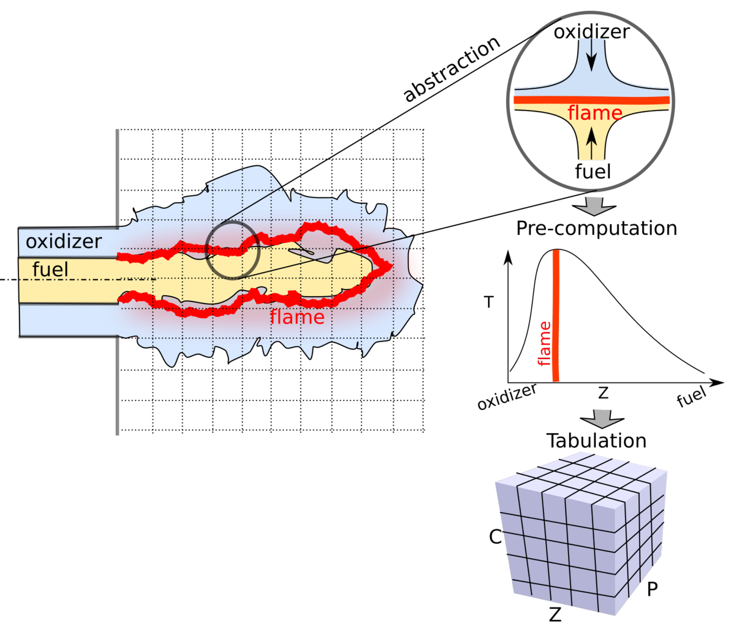
Sushrut Bhalla, Matthew Yao, Jean-Pierre Hickey, Mark Crowley

University of Waterloo

UNIVERSITY OF WATERLOO

Introduction

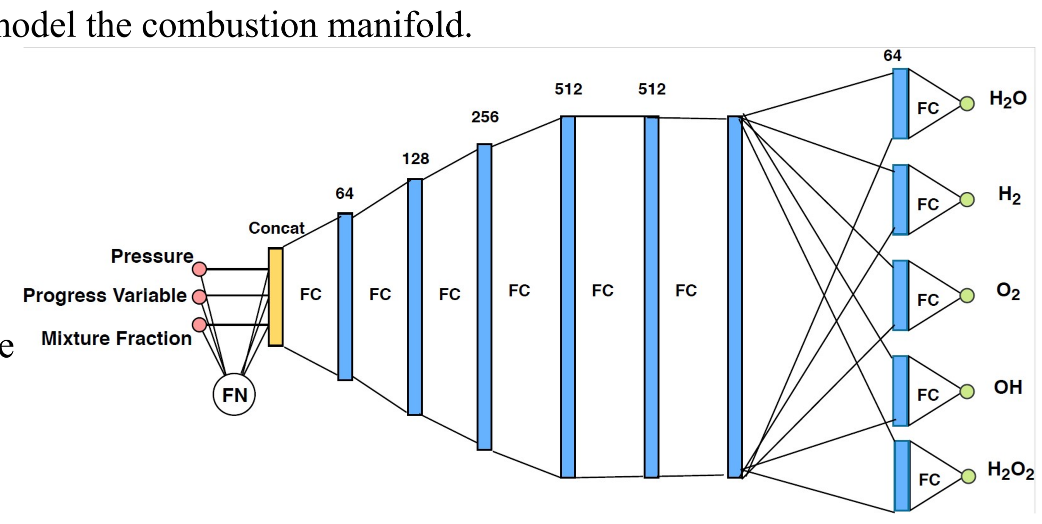
- The computational challenges in turbulent combustion simulations stem from the physical complexities and multi-scale nature of the problem which make it intractable to compute scale-resolving simulations.
- For most engineering applications, the large scale separation between the flame (typically sub-millimeter scale) and the characteristic turbulent flow (typically centimeter or meter scale) allows us to evoke simplifying assumptions--such as done for the flamelet model--to pre-compute all the chemical reactions and map them to a low-order manifold.
- The resulting manifold is then tabulated and looked-up at run-time. As the physical complexity of combustion simulations increases (including radiation, soot formation, pressure variations etc.) the dimensionality of the resulting manifold grows which impedes an efficient tabulation and look-up.
- A simplified illustration of the flamelet model for non-premixed combustion is shown in Figure on the right.



- In this work we present a novel approach to train deep neural networks to model the high-dimensional combustion manifold.
- We approximate the combustion manifold using a neural network function approximator and use it to predict the temperature and composition of the reaction.
- We present a novel training procedure which is developed to generate a smooth prediction curves for temperature over the course of a reaction.
- We then evaluate our work against the current approach of tabulation with linear interpolation in combustion simulations.
- We also provide an ablation study of our training procedure in the context of overfitting in our model.

Methods

- A deep neural network is used to model the combustion manifold.
- Four different neural networks are designed for predicting the Species, Heat Release, Temperature and Source Term.
- Temperature (T) and Source Term (W) are modelled using a separate neural network because of the large scale difference in their output range.



Over Sampling Hard Examples

- Hard examples are sample points where our model error exceeds 75th percentile error mark between all training data points.
- To improve accuracy on hard-examples, we perform over-sampling of hard data-points.
- The training batch consists of 50% hard data-points and 50% easy data-points.

Importance Weights Error and Gradient Clipping

- We incorporate the weighting of the loss function differently for hard and easy examples as it leads to better model predictions. We use the following loss for hard examples:

$$L_{imp}(d)_{d\sim D} = \alpha * MSE(T(d), \hat{T}(d))$$

- We also incorporate gradient clipping to clip large gradients in the backpropagation phase of neural network training to reduce any damage to our trained model caused by anomalies in our data.

Regularization

- To reduce the chance of oscillation and over-fitting, we combine *L1* and *L2* regularizers with an *L1L2* regularized loss function.

$$L(d)_{batch} = MSE(d)_{batch} + \lambda_{l1}\Sigma|W_{NN}| + \lambda_{l2}\Sigma|(W_{NN})^2|$$

- We arrive at values of $\lambda_{l1} = 0.00015$ and $\lambda_{l2} = 0.000125$. through cross validation of the temperature curve smoothness.
- Additional experiments were performed with the current state of the art regularization techniques, e.g, Batch Normalization, Layer Noramalization and Dropout.

Ensemble Model

- An ensemble model helps reduce the overfitting in the final predictions and improve accuracy by cherry picking the individual models.
- Ensemble model helps reduce the regularization performed to each individual neural network while reducing overfitting.
- We create an averaging ensemble of 5 trained deep neural networks. The output of only 4 models is used to generate the prediction. The 4 models are chosen based on least variance criteria.
- Individual neural networks are trained using different initializations, learning rates and initial seed values.

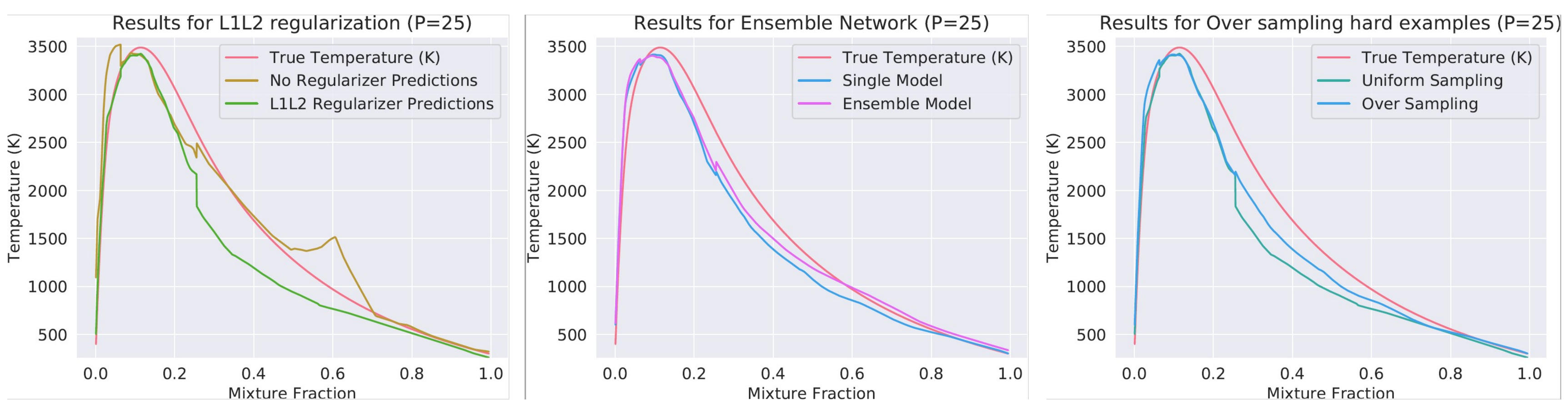
Results

- The following table shows the accuracy and the loss for each prediction variable.
- The equation to consider a prediction accurate is:

$$\hat{T}(d) \epsilon \{T(d) - ET,\ T(d) + ET\}$$

	H	O2	O	OH	H2	H2O	HO2	H2O2	HR	T(K)	W
Acc	21.44	12.56	37.43	43.34	25.41	14.59	41.42	36.05	81.18	54.60	58.14

- The following figures show an ablation comparison of each approach discussed in the Methods section.
- L1L2 regularizer reduces over-fitting in the trained model.
- Using an ensemble model helps reduce the error in the trained model.
- The over-sampling of hard examples during training reduces the model's error on areas with large variation.



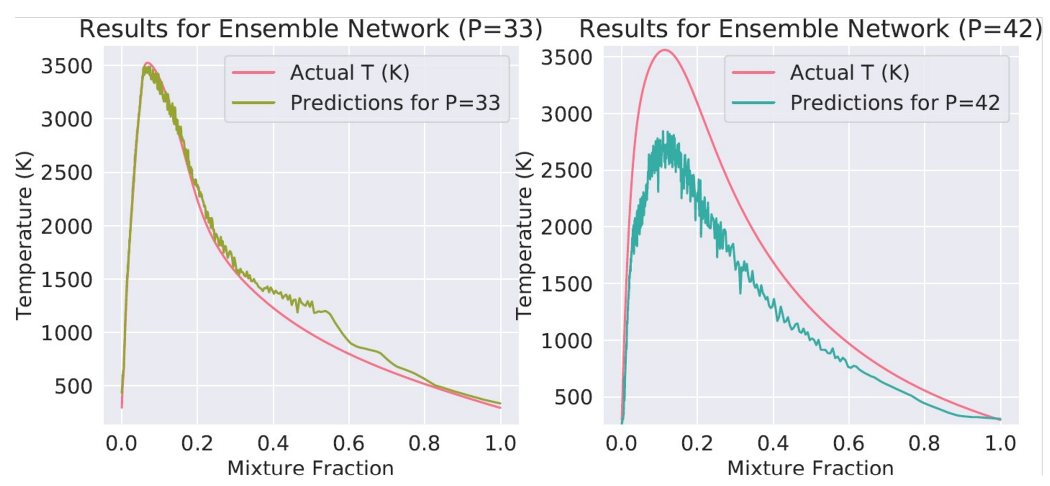
- The following table show the accuracy gain when using *L1L2* regularization.
- The table includes comparisons with popular techniques like Layer normalization and Batch normalization.

	No Regularization	L1L2 Regularization	Layer Normalization	Batch Normalization
Acc	31.44	39.83	23.63	12.28

- The following table show the accuracy comparison when we replace uniform sampling during training with over sampling.
- The use of ensemble model does not improve the model accuracy but provides smoother temperature prediction curves which are needed for the combustion simulator.

	Uniform Sampling	Over Sampling	Ensemble Model
Acc	39.83	48.87	47.73

- The figure on the right shows a qualitative comparison of our model's predictive power when interpolating to testing data.
- The training data included curves at P=30,35,40 and 50Pa. As we see in the figures, our model interpolates well for 5Pa pressure difference.



- The following table presents the memory and the running time analysis of our approach when compared to the current tabular approach used in combustion community.

	Parallel Inference Time (in ms)	Serial Inference Time (in s)	Memory Requirements (in MB)
Tabulation Method	1.2×10^5	10.997	184.64
Deep Neural Network	13.92	55.27	24.158

Conclusions

- Proposed a novel training procedure for modelling high dimensional combustion manifolds using deep neural networks.
- Proposed a novel loss function for regression tasks with examples of varying degree of hardness.
- We also propose a fast over sampling methodology based on the loss of each data point.
- Presented a trained model which achieves sufficient accuracy when compared with tabulated data and runs fast enough to integrate into high dimensional multi-physics simulators of combustion.
- Our model allows for cheap computation of very complex physics, compared to the traditional tabulation methods which can not scale to high dimensions.
- We plan to extend this work to focus on dimensionality reduction to understand core aspects of the combustion manifold.

Bibliography

1. Ma, P.C., Wu, H., Ihme, M., Hickey, J.P.: Nonadiabatic Flamelet Formulation for Predicting Wall Heat Transfer in Rocket Engines. AIAA Journal
2. Ihme, M., Schmitt, C., Pitsch, H.: Optimal artificial neural networks and tabulation methods for chemistry representation in les of a blu-body swirl-stabilized flame. Proceedings of the Combustion Institute
3. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollar, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision.