"Reduced-order modeling of turbulent reacting flows using data-driven approaches"

Candidate: Dr. Kamila Zdybał Université Libre de Bruxelles kamilazdybal@gmail.com +48 604 453 455	**Supervisor:** Professor Alessandro Parente Université Libre de Bruxelles **Co-supervisor:** Professor James C. Sutherland The University of Utah

Brief summary of the work

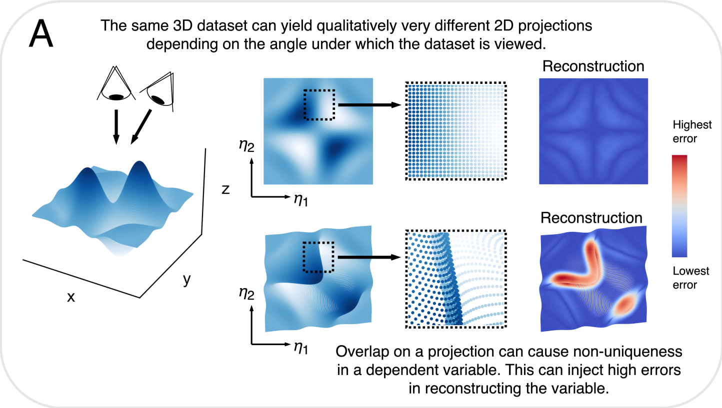
Simulating turbulent flames is a computationally challenging task. This remains true even with the current advances in numerical algorithms and high-performance computing. Combustion of even simple fuels, like methane, is described by many thermo-chemical state variables, including temperature, pressure and tens or even hundreds of chemical compounds that form in a burning flame. Each such variable adds an extra model dimension that needs to be accounted and solved for in a simulation. This can increase the computational time of a simulation to the point that accurate simulations become unaffordable. Dimensionality reduction has thus been explored in the literature to alleviate the high computational cost of reacting flow simulations. However, there remain many challenges and unanswered questions that dimensionality reduction techniques pose. One of the main challenges is: How can we represent the high-dimensional reacting system well in lower dimensions? Projecting the thermo-chemical state onto new low-dimensional coordinates can introduce non-uniqueness in representing modeled quantities (see **Fig. A**). This can significantly hamper reduced-order model performance. The present research project addresses the outstanding questions related to dimensionality reduction of combustion datasets. We demonstrate how ill-behaved projections can explain the model mispredictions reported in the literature thus far. We then provide novel strategies and insights on assessing and improving the quality of low-dimensional projections. The techniques developed in this thesis advance the state-of-the-art in model reduction and provide tools that can significantly improve the accuracy of numerical simulations of turbulent flames. This progress is much needed to support the rapidly developing combustion technologies. Accurate numerical simulations will be an indispensable tool to guarantee environmentally safer and more efficient combustion across industries. This research project also provides novel software that can be used by the next generation of students, researchers and engineers, working in reduced-order modeling.

Scientific summary

The main goal of my research is to address the outstanding questions in reduced-order modeling of turbulent combustion, such as: How do poor data projections affect modeling? How can we quantify and improve data projections? What benefits can we expect in model reduction once the projection is well-behaved? The challenge of data-driven dimensionality reduction is illustratively visualized in **Fig. A**, where an overlap in projection topology yields high errors in reconstructing a dependent variable.

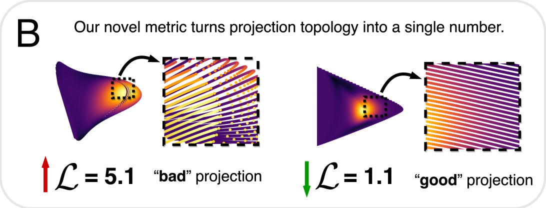
During my doctorate, I have developed a novel quantitative metric to characterize the quality of a data projection [3]. The metric pays attention to two aspects that affect modeling in particular: non-uniqueness and large gradients. A larger value indicates a more problematic projection that can include overlapping observations, twists, large gradients, or large curvatures in a projection topology. A smaller value indicates an improved projection (see **Fig. B**). Such metric was not available in the research community thus far. Many *ad hoc* and heuristic guidelines in dimensionality reduction had to be followed in turbulent combustion research, and in engineering in general. The metric now provides researchers with a way to automate many decisions that impact how a projection looks like. Those include: How to scale the data? Which independent variables to select? Which dimensionality reduction technique to use? How low can the projection dimensionality be without losing modeling accuracy? For example, for two experimentally measured sets of turbulent flames, **F1** and **F2**, our metric correctly identifies that flames **F2** are more difficult to parameterize and would require a higher projection dimensionality to achieve the same parameterization quality as flames **F1** (see **Fig. C**). This agrees with the physical understanding of the flames, where **F2** are known to contain stronger effects of turbulence acting to partially extinguish the flame. **F2** thus contain many more possible thermo-chemical states of the flame than **F1**, and our quantitative metric is capable of capturing this phenomenon. The metric proposed can find applications in virtually any domain of science and engineering. In my research, I propose various interdisciplinary applications of the metric to reacting flows, atmospheric physics, plasma flows, and even biomedical sciences. In the original publication, published in the Nature Portfolio, we argue that further improvements in parameterization quality can be achieved in many areas of research if the low-dimensional parameter space is thoroughly explored and then assessed using the proposed metric.

I have further applied the proposed metric in dimensionality reduction strategies that improve the projection quality (top panel of **Fig. D**). We apply the metric as an objective function to minimize in selecting the appropriate thermo-chemical state vector for model reduction purpose [4]. My goal was to address the gap in the reacting flow simulation literature, where various authors have performed *ad hoc* state vector selections, without giving detailed insights into how selecting or discarding some variables impacts the projection quality. I this work, we offer two quantitative algorithms that allow for significant improvements in parameterization quality. The associated paper is published in the most prestigious venue for combustion researchers, The Proceedings of the Combustion Institute. Following excellent reviews, this paper has received the Distinguished Paper Award from the Combustion Institute.

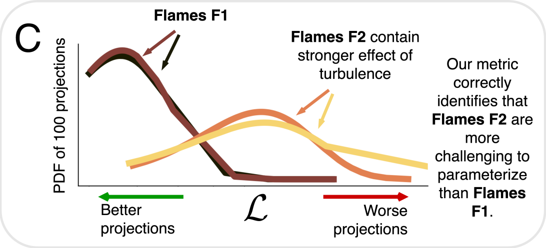


The same 3D dataset can yield qualitatively very different 2D projections depending on the angle under which the dataset is viewed.

Overlap on a projection can cause non-uniqueness in a dependent variable. This can inject high errors in reconstructing the variable.

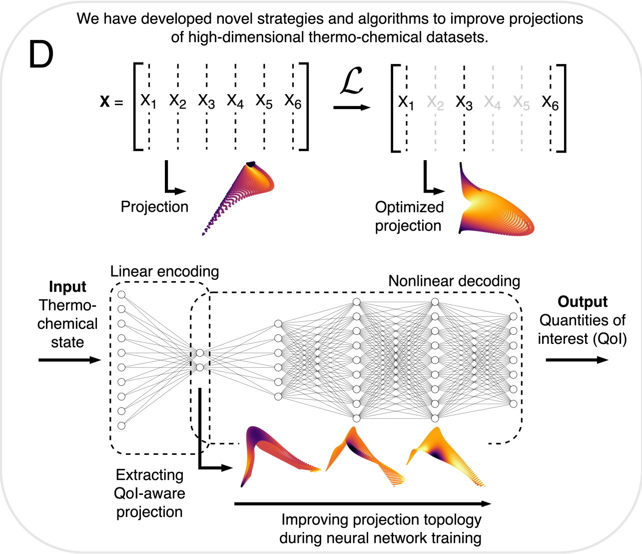


Our novel metric turns projection topology into a single number.



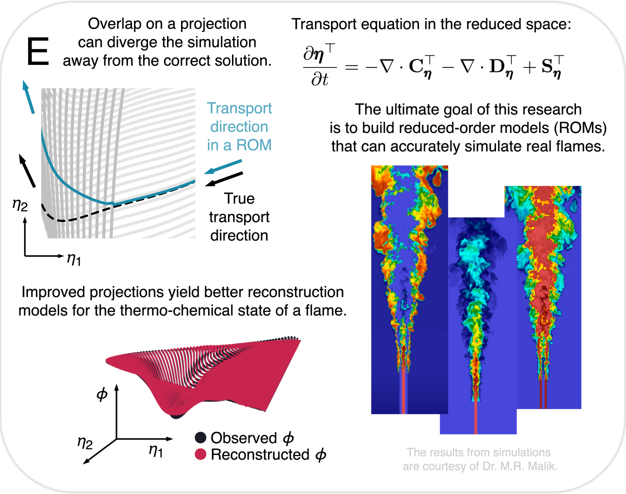
Flames F2 contain stronger effect of turbulence

Our metric correctly identifies that **Flames F2** are more challenging to parameterize than **Flames F1**.



We have developed novel strategies and algorithms to improve projections of high-dimensional thermo-chemical datasets.

Improving projection topology during neural network training



Overlap on a projection can diverge the simulation away from the correct solution.

Transport equation in the reduced space:

$$\frac{\partial \boldsymbol{\eta}^\top}{\partial t} = -\nabla \cdot \mathbf{C}_{\boldsymbol{\eta}}^\top - \nabla \cdot \mathbf{D}_{\boldsymbol{\eta}}^\top + \mathbf{S}_{\boldsymbol{\eta}}^\top$$

The ultimate goal of this research is to build reduced-order models (ROMs) that can accurately simulate real flames.

Improved projections yield better reconstruction models for the thermo-chemical state of a flame.

The results from simulations are courtesy of Dr. M.R. Malik.

Another strategy to improve projection quality employs artificial neural networks (bottom panel of **Fig. D**). We introduce a novel modification to encoding-decoding architectures, where we optimize the projection directly for any projection-independent and projection-dependent quantities of interest (QoIs). This strategy naturally promotes improved projections to emerge during neural network training since non-uniqueness and steep gradients increase the mean-squared-error loss function. Thus, projections that do exhibit topological issues are immediately penalized and discarded. We demonstrate that this strategy can become an effective replacement of standalone dimensionality reduction techniques used thus far in the combustion literature, since it provides vast improvements to projection topologies.

The ultimate goal of this research is to improve numerical simulations of real flames (**Fig. E**). Problems reported in the literature thus far can stem from poorly-behaved low-dimensional projections and high errors in models that predict thermo-chemistry *a posteriori*. **Fig. E** shows how overlap on a projection can cause the transport direction to diverge away from the correct solution. This is due to injecting ambiguity in reconstruction models, enacted during and after simulation. This research provides strategies that can be deployed in numerical simulations. We show that significant improvements in predicting thermo-chemistry can be achieved following our proposed guidelines. This can have a broad impact on combustion engineering and combustion industries that require faster tools to help improve combustion as a process. The future goal is to move towards more efficient combustion with lower emissions of harmful pollutants.

Throughout my doctoral research, I have also been actively involved in developing novel software. The primary output of my work in this regard is an open-source Python library, **PCAfold** [5]. It collects all of the novel tools and algorithms that I have developed during my Ph.D., and has already been appreciated by many researchers worldwide in combustion and beyond. I paid particular attention to documenting the software precisely and creating numerous illustrative tutorials that can guide future students in the subject. The documentation is available at: **https://pcafold.readthedocs.io/**. The second library that I have developed, **multipy**, has a didactic purpose and is aimed at helping students learn the subject of multicomponent mass transfer. It can be particularly useful to students performing research in fluid dynamics and multicomponent flows, including combustion. I have also been dedicated to sharing my code for reproducing our research results which has been widely appreciated by the academic community.

Five selected publications & indication of the total research output volume

[1] **K. Zdybał**, G. D'Alessio, A. Attili, A. Coussement, J.C. Sutherland, A. Parente, Local manifold learning and its link to domain-based physics knowledge, Applications in Energy and Combustion Science (14) 100131, 2023.

[2] **K. Zdybał**, G. D'Alessio, G. Aversano, M.R. Malik, A. Coussement, J.C. Sutherland, A. Parente, Advancing reactive flow simulations with data-driven models, Chapter in edited collection: *"Data-Driven Fluid Mechanics: Combining First Principles and Machine Learning"*, Cambridge University Press, 304-329, 2023.

[3] **K. Zdybał**, E. Armstrong, J.C. Sutherland, A. Parente, Cost function for low-dimensional manifold topology assessment, Scientific Reports (12) 14496, 2022.

[4] **K. Zdybał**, J.C. Sutherland, A. Parente, Manifold-informed state vector subset for reduced-order modeling, Proceedings of the Combustion Institute 39(4):5145-5154, 2022.
This paper has received the Distinguished Paper Award from The Combustion Institute.

[5] **K. Zdybał**, E. Armstrong, A. Parente, J.C. Sutherland, PCAfold: Python software to generate, analyze and improve PCA-derived low-dimensional manifolds, SoftwareX (12) 100630, 2020.

My research resulted in total of six peer-reviewed journal articles, five of which I am the lead author of. I have also written two book chapters in edited collections as the lead author (for Cambridge University Press and Springer). I currently have one more paper under review in the journal *Patterns* and one more paper to be submitted in July 2023 to the journal *Combustion and Flame*, both on research coming from my Ph.D. dissertation. I am currently working on a review paper on dimensionality reduction for reacting flows for the journal *Progress in Energy and Combustion Science* as the lead author. This paper will present and critically review many of the insights gained during my doctoral work.