

AFRL-AFOSR-VA-TR-2022-0390

Modeling of non-equilibruim hypersonic air flows by means of the Multi-group Maximum Entropy method

Panesi, Marco
UNIVERSITY OF ILLINOIS
506 S WRIGHT ST
URBANA, IL, 61801
USA

07/15/2022
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
20220715	Final	START DATE 20180715	END DATE 20210714

4. TITLE AND SUBTITLE
Modeling of non-equilibruim hypersonic air flows by means of the Multi-group Maximum Entropy method

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
	FA9550-18-1-0388	61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Marco Panesi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF ILLINOIS
506 S WRIGHT ST
URBANA, IL 61801
USA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	AFRL/AFOSR RTA1	AFRL-AFOSR-VA-TR-2022-0390

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The development of hypersonic weapons, frequently noted as a game-changing capability in military superiority, critically relies on the availability of computationally efficient and accurate models for the description of non-equilibrium flows. This project has constructed a reduced-order model for modeling of non-equilibrium phenomena in high-speed airflows. This work is at the interface between computational chemistry and computational fluid dynamics and aims at developing new models based on microscopic theory and applying them to macroscopic scales. The most physically consistent description of non-equilibrium flows relies on the solution of the master equations for each internal level of the gas particles. However, such a detailed description is impractically expensive. The proposed approach leverages the maximum entropy principle, subject to a series of moment constraints, to reconstruct the logarithm of the distribution function expressed as a power series in internal energy. To improve the accuracy of the method, the internal energy levels are lumped in multiple groups, leveraging an adaptive kinetic-based grouping strategy to select the 'best' arrangement of the states within each group.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	168

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
SARAH POPKIN	000-0000

FINAL REPORT
AFOSR Grant: # FA9550-15-1-0132

MODELING OF NON-EQUILIBRUIM HYPERSONIC AIR FLOWS BY MEANS OF THE MULTI-GROUP MAXIMUM ENTROPY





by the
University of Illinois at Urbana-Champaign
The Center for Hypersonics and Entry Systems Studies (CHESS)
Aerospace Engineering (AE)

Marco Panesi (AE)

Proposed Period of Performance: 05/01/2022 – 04/30/2027

Technical POC: *Name:* **Marco Panesi**
Mailing Address: 314 Talbot Lab, 104 S. Wright Street, Urbana, IL, 61801.
Telephone: (217) 333-2187
Email: mpanesi@illinois.edu

Table of Contents

Table of Contents ..... **I**

Abstract ..... **II**

I. Section 2: Technical Report **1**

A. Objectives 1

B. Key Accomplishments 1

C. Impact 2

Abstract

The development of hypersonic weapons, frequently noted as a game-changing capability in military superiority, critically relies on the availability of computationally efficient and accurate models for the description of non-equilibrium flows. This project has constructed a reduced order model for modeling of non-equilibrium phenomena in high-speed air flows. This work is at the interface between computational chemistry and computational fluid dynamics and aims at developing new models based on microscopic theory and applying them to macroscopic scales. The most physically consistent description of non-equilibrium flows relies on the solution of the master equations for each internal level of the gas particles. However, such a detailed description is impractically expensive. The proposed approach leverages the maximum entropy principle, subject to a series of moment constraints, to reconstruct the logarithm of the distribution function expressed as a power series in internal energy. To improve the accuracy of the method, the internal energy levels are lumped in multiple groups, leveraging an adaptive kinetic-based grouping strategy to select the 'best' arrangement of the states within each group. This approach seeks for an optimal reduced order representation of the distribution function by grouping together individual energy states that are likely to quickly equilibrate with each other, thus maximizing the accuracy of the reconstruction, and drastically reducing the computational cost. This new class of ab-initio based models has the potential to revolutionize engineering and science by enabling truly predictive simulation of nonequilibrium chemistry phenomena, over a wide spectrum of flow conditions. The new model has been implemented in a AFOSR solvers and used to investigate the hypersonic flows.

Approved for Public Release

I. Section 2: Technical Report

A. Objectives

This project aims to construct an adaptive strategy for the construction of reduced order non-equilibrium models using the Multi-group Maximum Entropy method. The method will make use of Quasi-Classical scattering calculations and clustering algorithms to determine the best grouping strategy for a preselected chemical system, without the usual reliance on the a-priori knowledge and/or assumptions on the relaxation dynamics and/or timescales. The specific objectives of the proposed work are summarized below.

1. Compute the State-to-State Dynamics of the Internal States in a Strong Non-equilibrium Environment: The analysis of the non-equilibrium energy transfer and chemical relaxation processes, using a rovibrational state-to-state description, will provide invaluable information on the dynamics of each internal state without the usual reliance on case-specific empiricism.

2. Construction of the Multi-group Maximum Entropy method for air: The method relies on the reconstruction of the internal distribution function by using the maximum entropy principle. The internal energy space will be subdivided into groups and the method of weighted residuals applied to the master equations to obtain macroscopic equations. The system under consideration will include: N2-O2, O2-O2, (N2-N2) and the associated sub-systems.

3. Application to AFOSR CFD solvers and comparison against the experimental data: The adaptive multi-group maximum entropy method will be applied to the Air Force flow codes (i.e., US3D). The US3D code will be used for this purpose.

B. Key Accomplishments

The main accomplishments are discussed in depth in several publications. A brief discussion of the contribution of each publication is discussed hereafter:

1. *SM Jo, M. Panesi and JG Kim "Prediction of shock standoff distance with modified rotational relaxation time of air mixture" Phys. Rev. Fluids, Physics of Fluids 33, 047102 (2021).* This paper proposes an in-depth analysis of the rotational relaxation of N2 molecules. The analysis relies on the use of QCT to derive state specific rates to be used in the CFD. **Merit.** Diatom-Diatom molecules are extremely challenging owing to the overwhelmingly large, virtually infinite. This work proposes an approach to the model reduction that is based on the rotational energy of the distribution function. There is only another work of this kind of analysis and was carried out by the PI and collaborators see 5) & 6).
2. *S. Venturi, M. Sharma; B. Lopez, M. Panesi "A Data-Inspired and Physics-Driven Model Reduction for Dissociation: Application to O2+O System", Journal of Physical Chemistry A, 124, 8359-837 (2020).* This work proposes a new strategy for the modeling of dissociation in diatom-atom collisions. It clearly demonstrates the

invalidity of conventional approaches based on the assumption of rotational equilibrium; it also defines a new model in which the transition probability is function on the distance from the centrifugal barrier. **Merit.** This work clearly demonstrates the invalidity of about 80% of the non-equilibrium models published over the last couple of decades.

3. *R. Macdonald, E. Torres, T. Schwerzentruber, and M. Panesi "State-to-state and direct molecular simulation study of energy transfer and dissociation of nitrogen mixtures" Journal of Physical Chemistry A, 124, 35, 6986–7000 (2020).* This is an in-dept comparison of the Direct-Molecular Simulation and the Master equation approach. The agreement confirms the validity of the DMS model, that was never assessed before.
4. *S. Venturi, R.L. Jaffe, and M. Panesi "A Bayesian Machine Learning Approach to the Quantification of Uncertainties on Ab Initio Potential Energy Surfaces" Journal of Physical Chemistry A, 124, 25, 5129–5146 (2020).* This work represents the first application of the Bayesian Neural Network for the quantification of Uncertainty associated with the Ab-initio based models. **Merit.** This is one of the first works that addresses the problem of quantifying uncertainty associated with quantum mechanical based PES.
5. *R.L. Macdonald, R.L. Jaffe, D.W. Schwenke, M. Panesi "Construction of a Maximum Entropy Quasi-Classical Trajectory Method. Part 1: Theory and Application to N2-N2 System" J. Chemical Physics, 148, 054309 (2018).* This work presents a new methodology for the construction of reduced order model for diatom-diatom processes. All the conventional approaches rely on the assumption of rotational equilibrium which is invalid in the case of hypersonic applications. **Merit.** This work clearly demonstrates that the importance of vibration in conventional chemistry models is due mainly to the assumption of rotational equilibrium. When this assumption is removed, vibration and rotation contribute to dissociation in equal proportion.
6. *R.L. Macdonald, M. Grover, T. Schwartzentruber, M. Panesi "Construction of a Maximum Entropy Quasi-Classical Trajectory Method. Part 2: Comparison against the Direct Molecular Simulation method" J. Chemical Physics, 148, 054310 (2018).* This work is the second part of the two parts paper and validates the model discussed above 5) using Direct Molecular Simulation.
7. *M. Sharma, Y. Liu, M. Panesi "Characterization of Chemical Non-Equilibrium in Hypersonic Flows using Quadratic Reconstruction Operator", Physical Review E, 101, 013307, (2020).* This work presents and high-order method for the reconstruction of the population using the reduced order model constructed by the PI. **Merit.** Using quadratic reconstruction, it is possible to further reduce the computational cost in the modeling of chemically reacting gases.

C. Impact

This work has constructed a new reduced order method for the modeling non-equilibrium flows in Computational Fluid Dynamics. It has been demonstrated to allow for many orders of magnitude speedup while keeping the accuracy of the more costly models, that are often applicable to 0 and 1D simulations. Other research groups are adopting the method in US and abroad and more importantly the conventional approaches can be recast as a sub-class of the new model created.

MACHINE LEARNING AND UNCERTAINTY QUANTIFICATION FRAMEWORK FOR PREDICTIVE *AB INITIO* HYPERSONICS

BY

SIMONE VENTURI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Aerospace Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2021

Urbana, Illinois

Doctoral Committee:

Professor Marco Panesi, Chair and Director of Research
Professor J. Craig Dutton
Professor Kelly Stephani
Dr. Raymond Bemish, Air Force Research Laboratory
Dr. Richard L. Jaffe, NASA Ames Research Center

Abstract

Hypersonics represents one of the most challenging applications for predictive science. Due to the multi-scale and multi-physics characteristics, high-Mach phenomena are generally complex from both the computational and the experimental perspectives. Nevertheless, the related simulations typically require high accuracy, as their outcomes inform design and decision-making processes in safety-critical applications. *Ab initio* approaches aim to improve the predictive accuracy by making the calculations free from empiricism. In order to achieve this goal, these methodologies move the computational resolution down to the interatomic level by relying on first-principles quantum physics. As side effects, the increase in model complexity also results in: i) more physics that could be potentially misrepresented and ii) dramatic inflation of the computational cost. This thesis leverages machine learning (ML), uncertainty quantification (UQ), data science, and reduced order models (ROMs) for tackling these downsides and improving the predictive capabilities of *ab initio* Hypersonics.

The first part of the manuscript focuses on formulating and testing a systematic approach to the reliability assessment of ML-based models based on their non-deterministic extensions. In particular, it introduces a novel methodology for the quantification of uncertainties associated with potential energy surfaces (PESs) computed from first-principles quantum mechanical calculations. The methodology relies on Bayesian inference and ML techniques to construct a stochastic PES and to express the inadequacies associated with the *ab initio* data points and their fit. The resulting stochastic surface is efficiently forward propagated via quasi-classical trajectory (QCT) and master equation calculations by combining high fidelity calculations and reduced order modeling. In this way, the PES contribution to the uncertainty on predefined quantities of interest (QoIs) is explicitly determined. This study is done at both microscopic (e.g., rovibrational-specific rate coefficients) and macroscopic (e.g., thermal and chemical relaxation properties) levels. A correlation analysis is

finally applied to identify the PES regions that require further refinement, based on their effects on the QoI reliability. The methodology is applied to the study of singlet ($1^1A'$) and quintet ($2^5A'$) PESs describing the interaction between O_2 molecules and O atoms in their ground electronic state. The investigation of the singlet surface reveals a negligible uncertainty on the kinetic properties and relaxation times, which are found to be in excellent agreement with the ones previously published in the literature. On the other hand, the methodology demonstrated significant uncertainty on the quintet surface due to inaccuracies in the description of the exchange barrier and the repulsive wall. When forward propagated, this uncertainty is responsible for the variability of one order of magnitude in the vibrational relaxation time and of factor four in the exchange reaction rate coefficient, both at 2 500 K.

The second part of this thesis presents a data-informed and physics-driven coarse-graining strategy aimed to reduce the computational cost of *ab initio* simulations. At first, an in-depth discussion of the physics governing the non-equilibrium dissociation of O_2 molecules colliding with O atoms is proposed. A rovibrationally-resolved database for all of the elementary collisional processes is constructed by including all nine adiabatic electronic states of O_3 in the QCT calculations. A detailed analysis of the *ab initio* data set reveals that, for a rovibrational level, the probability of dissociating is mostly dictated by its deficit in internal energy compared to the centrifugal barrier. Due to the assumption of rotational equilibrium, the conventional vibrational-specific calculations fail to characterize such a dependence, and the new ROM strategy is proposed based on this observation. By relying on a hybrid technique made of rovibrationally-resolved excitation coupled to coarse-grained dissociation, the novel approach is compared to the vibrational-specific model and the direct solution of the rovibrational state-to-state master equation. Simulations are performed in a zero-dimensional isothermal and isochoric chemical reactor for a wide range of temperatures (1 500 - 20 000 K). The study shows that the main contribution to the model inadequacy of vibrational-specific approaches originates from the incapability of characterizing dissociation, rather than the energy transfers. Even when constructed with only twenty groups and only 20% of the original computational cost, the new reduced order model outperforms the vibrational-specific one in predicting all of the QoIs related to dissociation kinetics. At the highest temperature, the accuracy in the mole fraction is improved by 2 000%.

Acknowledgments

I would like to express my gratitude to my advisor, Prof. Marco Panesi, for his mentorship during these five years of Phd. Through his guidance and example, he taught me that mastering both detailed and big-picture analysis is a key requirement for excelling in research. His curiosity and open-mindedness to new ideas and topics have been constant and inspirational, and they profoundly helped me with untangling in such a multi-disciplinary field.

I would like to thank my committee members, Prof. Panesi, Prof. Dutton, Prof. Stephani, Dr. Bemish, and Dr. Jaffe, for their appraisal of this research and valuable feedback.

I am also very thankful to Dr. Michael Barnhardt, Dr. Richard L. Jaffe, Dr. David Hash, and Dr. David Schwenke for having made possible my two internships at NASA Ames. In particular, I would like to thank Dr. Jaffe for his precious mentorship during those two summers and for having always been available to answer my numerous questions about quantum chemistry. Special thanks to Dr. Schwenke also for sharing his VVTC QCT code.

I am indebted to the postdoctoral researchers of the NEQRAD group, Dr. Bruno Lopez and Dr. Alessandro Munafò, the former for the help with creating the CoarseAIR QCT code and the latter for providing access to his PLATO library and KONIG code. The work presented in this thesis would not have been possible without these computational tools.

I would like to thank the other NEQRAD group researchers, Dr. Andrea Alberti, Chiara Amato, Berkan Bolkan, Hugo Chiang, Sanjeev Kumar, Vegnesh Jayaraman, Dr. Sung Min Jo, Dr. Robyn Macdonald, Vishnu Oruganti, Alberto Racca, Przemyslaw Rostkowski, Dr. Amal Sahai, Maitreyee Sharma, Dr. Pablo Solano-Lõpez, Daniel Steinberg, and Dr. Joao Vargas for the continuous brainstorming and constructive feedback. In particular, I want to express my gratitude to Andrea. For his passion, work ethic, and moral integrity, he has been a continuous source of inspiration and a model to follow. I am thankful to Amal for the useful discussions about reduced order modeling

and nonequilibrium kinetics.

I am very grateful for the people I was lucky enough to meet in Urbana-Champaign. In particular, the friendship with Dr. Giovanni Antonio, Dr. Tommaso Biancalani, Dr. Elena Bichi, Giuseppe Bonaiti, Prof. Francesco Cellarosi, Dr. Filippo Ciabatti, Carlo Di Giulio, Giusy Falcone, Dr. Giovanni Fiore, Regina Hinders, Gavin Kumar Ananda, Francesca Nardecchia, Dr. Luca Patriarca, Eleonora Po, Dr. Marcello Rubessa, Patty Rubessa, Francesca Salerno, Dr. Mario Vailati Riboni, and Mariagabriella Stuardi deeply, even if indirectly, contributed to this thesis.

I am also extremely thankful to my lifelong friends for having constantly supported me in the past five years. Even more, for doing that from distance. I would like to mention: Giacomo Alaimo, Paolo Amato, Filippo Ballo, Dr. Andrea Battistoni, Dr. Carlo Bianchi, Luca Biondini, Nicola Careri, Tommaso Costantini, Giancarlo Della Volpe, Alessio Foconi, Gianluca Giampaolo, Giulio Giovannetti, Leonardo Illuminati, Mattia Lanzani, Michele Loperfido, Raffaele Magrini Alunno, Paolo Mancinelli, Giulia Minucci, Maria Novelli, Dr. Pietro Novelli, Cristiana Petrignani, Alessio Savi, and Dr. Giorgio Vallone.

I would like to express my gratitude to my girlfriend, Jillian, for her patience, motivation, love, and with her to all the Van Kampen family.

Finally, my most profound gratitude goes to my parents, Elisabetta and Valter, to my brother, Lorenzo, and to all my relatives back in Italy. This year, COVID-19 made it impossible to see each other in person and to celebrate the conclusion of this doctorate together. This thesis is dedicated to all of you.

This work was supported by the Air Force Office of Scientific Research No. FA9550-18-1-0388 with Program Office Dr. Ivett Leyva. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the AFOSR or the U.S. government. The large computations for this work were carried out on the NASA Pleiades supercomputer at NASA Ames Research Center.

Table of Contents

List of Tables . . . . . **ix**

List of Figures . . . . . **x**

List of Abbreviations . . . . . **xx**

List of Symbols . . . . . **xxii**

Chapter 1 Introduction **1**
1.1 Motivation 1
1.2 Literature Review 3
1.2.1 *Ab Initio* Potential Energy Surfaces and Associated Uncertainties 3
1.2.2 Reduced Order Modeling for Nonequilibrium Kinetics 7
1.2.3 The O_2+O Chemical System 10
1.3 Objective of the Thesis 11
1.4 Related Scientific Contributions 14

I Bayesian Machine Learning Approach to the Quantification of Uncertainties on *Ab Initio* Potential Energy Surfaces (PESs) **18**

Chapter 2 The Stochastic PES: Construction, Propagation, and Refinement . **19**
2.1 Permutation Invariant Polynomials Neural Networks for PESs 22
2.2 The Steps of the SPES Approach 25
2.2.1 Step 1: Bayesian Calibration 25
2.2.2 Step 2: Forward Propagation 28
2.2.3 Step 3: Surface Refinement 29
2.3 Physical Modeling 30
2.3.1 *Ab Initio* Construction of the PESs 31
2.3.2 Rovibrational-Specific QCT Calculations 33
2.3.3 Rovibrational State-to-State Master Equation 34
2.3.4 Reduced-Order Modeling: CG-QCT and CGM 35
2.3.5 Convergence of QCT and CG-QCT Calculations 36

Chapter 3 Applications to the O_2+O System **42**
3.1 Verification Cases for the PIP-BNN 42
3.2 Errors and Uncertainties of the $1^1A'$ and $2^5A'$ Surfaces 48
3.3 Forward Propagation of the SPESs: Heat Bath Simulations at 10 000 K 53

3.4 Temperature-Dependent Kinetic Parameters and Comparisons with Experiments . . 58
3.5 Sensitivity Analysis of the $2^5A'$ PES 60

II Data-Inspired and Physics-Driven Model Reduction for Dissociation 63

Chapter 4 A Novel Grouping Strategy for the Dissociation Mechanisms . . . 64
4.1 O_2 Diatomic Potential and O_2+O Potential Energy Surfaces 64
4.2 QCT Calculations and *Ab Initio* Kinetic Database 66
4.3 Isothermal and Isochoric Chemical Reactor Model 67
4.3.1 Analysis of Dissociation in a 0-D Heat Bath: Pure Dissociation 68
4.3.2 Analysis of Dissociation in a 0-D Heat Bath: Excitation and Dissociation . . 68
4.4 Model Reduction for Dissociation Mechanisms 68
4.4.1 Reduced Order Approach to Master Equation 69
4.4.2 Hybrid Approach to ME: RVS StS Excitation Coupled to CG Dissociation . . 71
4.5 Equivalence of Coarse-Grained and Grouped-Reconstructed Formulations 72

Chapter 5 Applications to the O_2+O System 75
5.1 Analysis of the State-Specific Dissociation Rates 75
5.2 Analysis of Dissociation in a 0-D Heat Bath: Pure Dissociation 79
5.3 Analysis of Dissociation in a 0-D Heat Bath: Excitation and Dissociation 83
5.4 Comparison with Literature 89

III Conclusions 92

Chapter 6 Summary and Future Work 93
6.1 Summary 93
6.1.1 Bayesian ML Approach to the UQ of *Ab Initio* PESs 93
6.1.2 Data-Inspired and Physics-Driven Model Reduction for Dissociation 95
6.2 Future Work 96

Appendix A Part I, Additional Results 98
A.1 PESs 2D Cuts 98
A.2 PESs 3D Cuts 100
A.2.1 Results for $1^1A'$ PESs 100
A.2.2 Results for $2^5A'$ PESs 101
A.3 RVS StS Rate Coefficients 105
A.4 Gas Quantities of Interest at T = 10 000 K 106
A.5 Gas Quantities of Interest at T = 2 500 K 108
A.6 Gas Quantities as Functions of Temperature 110
A.7 Gas QoI Sensitivities to the $2^5A'$ PES Uncertainties 113
A.7.1 Vibrational Relaxation Times 113
A.7.2 Vibrational Relaxation Times Excluding Exchange Processes 114
A.7.3 Rotational Relaxation Times 115

Appendix B Part II, Further Results 116
B.1 Dimensionality Reduction of RVS Dissociation Rate Coefficients 116
B.1.1 RVS Dissociation Rate Coefficients Reconstructed from the VS Model 116
B.1.2 RVS Dissociation Rates Coefficients Reconstructed from the CB Model . . . 118
B.1.3 Depletion Delays 120
B.2 Amount of O_2 Dissociation Taking Place at QSS 121
B.2.1 Heat Bath at $T_{Tran} = 5\,000$ K 121
B.2.2 Heat Bath at $T_{Tran} = 20\,000$ K 122
B.3 Application to Other Chemical Systems 123
B.3.1 Application to O_2+C System 123
B.3.2 Application to CO+O System 124
B.3.3 Application to N_2+N System 125
B.3.4 Application to N_2+O System 126
B.3.5 Application to NO+N System 127

References 128

List of Tables

2.1 Prior Distributions of Parameters and Hyper-parameters: 27
2.2 Test case 1: gas QoIs at $T_{Tran} = 2\,500$ K from StS Calculations 38
2.3 Test case 2: gas QoIs at $T_{Tran} = 10\,000$ K from StS Calculations 39
2.4 Test case 3: gas QoIs at $T_{Tran} = 2\,500$ K from CGM200 40
2.5 Test case 4: gas QoIs at $T_{Tran} = 10\,000$ K from CGM200 41

5.1 Parameters Values for Eq. 5.1 78

List of Figures

1.1 Schematics of the *ab initio* computational framework for nonequlibrium flows. 3
1.2 Evolution of O_2 mole fraction normalized by its initial condition. Blue dash-dotted line: DMS simulation by Grover *et al.* [44] Black continuous line: QCT and master equation, both rovibrationally state-to-state, by Venturi *et al.* [160] Red dotted line: QCT and master equation, both vibrationally state-to-state. To be consistent with Grover *et al.*'s calculations, no recombination and no dissociation correction factor [110] have been considered. 9
1.3 Schematics of the proposed approach to predictive Hypersonics. (Sources of the subfigures: NASA Ames and NEQRAD research group) 12
1.4 Covers of the Journal of Physical Chemistry A, Volume 124, Issue 44, Machine Learning in Physical Chemistry Virtual Special Issue (a), and Volume 124, Issue 41 (b). 16

2.1 Schematic of the probabilistic model generating the Stochastic Potential Energy Surface (SPES). 20
2.2 The steps of the SPES approach. Black arrows represent flows of information content (a): Bayesian calibration, (i.e, solution of the inverse problem). The optimal PDFs characterizing the overall set of parameters and the hyperparameters of the stochastic potential energy surface (SPES) are learnt through Bayesian inference from the *ab initio* data points. (b): Forward propagation of the uncertainties. The posterior distribution is sampled, an ensemble of PESs is obtained, and the surfaces are employed in computational simulations with the aim of predicting QoIs. (c): Surface refinement based on sensitivity analysis. The QoIs can be correlated to the variability of the SPES at a fine grid of atomic configurations, in order to highlight the surface areas where uncertainties are crucial. 21
2.3 Structure of a L-Hidden Layers PIP-NN; the bold letters in the colored boxes correspond to the arrays of parameters characterizing the network. 24
2.4 Spatial configurations and related *ab initio* energies for the $1^1A'$ (a) and $2^5A'$ (b) data points from Varga *et al.* [150]. 32
2.5 Test case 1: evolution of O_2 mole fraction in a StS heat bath at constant T_{Tran} = 2 500 K. The rates for the ME calculations are computed from 100 000 (green solid line) and 20 000 (red dashed line) trajectories per level. 37
2.6 Test case 2: evolution of O_2 mole fraction in a StS heat bath at constant T_{Tran} = 10 000 K. The rates for the ME calculations are computed from 100 000 (green solid line) and 20 000 (red dashed line) trajectories per level. 38

2.7 Test case 3: evolution of O_2 mole fraction in a heat bath at constant T_{Tran} = 2 500 K simulated by means of CGM200. The results are obtained through bin rates computed by grouping the StS ones (green solid line) or directly by CG-QCT calculations (red dashed line). 39
2.8 Test case 4: evolution of O_2 mole fraction in a heat bath at constant T_{Tran} = 10 000 K simulated by means of CGM200. The results are obtained through bin rates computed by grouping the StS ones (green solid line) or directly by CG-QCT calculations (red dashed line). 40

3.1 Means (red dots) and three-sigma confidence intervals (red lines) for the potential energies computed at the data points using 50 PIP-BNN samples for verification case 1 (a), verification case 2 (b), and verification case 3 (c). 44
3.2 RMSEs from the mean of the 50 PIP-BNN samples at the data points of $2^5A'$ for the 3 verification cases (blue for case 1, red for case 2 and green for case 3). Based on their potential energies, the data points have been divided in 12 groups, with upper bounds given by $V_{Max} = \{2.0;\ 4.0;\ 6.0;\ 8.0;\ 10.0;\ 15.0;\ 20.0;\ 25.0;\ 30.0;\ 50.0;\ 100.0;\ 1\,000.0\}$ eV 45
3.3 Posterior distributions of the likelihood function's standard deviation for the 3 verification cases (blue for case 1, red for case 2 and green for case 3). 45
3.4 Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at T_{Tran} = 10 000 K. 46
3.5 Evolution of O_2 dissociation (lower black part) and exchange (upper red part) averaged rates, as predicted by ME (solid lines) and CGM200 (dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (dotted lines), and by CGM200 using the 50 PIP-BNN samples (narrow lines). The histograms on the left report the dissociation (lower Subfig.) and exchange (upper Subfig.) QSS rates predicted through such samples, and they compare them to the outcomes of CGM200 using Varga *et al.*'s fit. The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at T_{Tran} = 10 000 K. 46
3.6 Evolution of O_2 averaged vibrational (green lines) and rotational (red lines) energy, as predicted by ME (solid lines) and CGM200 (dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (dotted lines), and by CGM200 using the 50 PIP-BNN MAP samples (narrow lines). The sub-figures on the top report the histogram of the vibrational and rotational relaxation times predicted through such samples, and they compare them to the outcomes of CGM200 using Varga *et al.*'s fit. The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at T_{Tran} = 10 000 K. 47
3.7 Means and three-sigma confidence intervals of the $1^1A'$ potential energies computed at the data points using 50 PIP-BNN samples (a, red dots) and 3000 PIP-BNN samples (b, red dots), compared to the energies resulting from Varga et al.'s fits (green dots). 49

3.8 Means and three-sigma confidence intervals of the $2^5A'$ potential energies computed at the data points using 50 PIP-BNN samples (a, red dots) and 3000 PIP-BNN samples (b, red dots), compared to the energies resulting from Varga et al.'s fits (green dots). 49
3.9 RMSEs at the $1^1A'$ (a) and $2^5A'$ (b) data points. In green, the errors produced by Varga *et al.*'s fits; in red, the ones generated by the means of 50 PIP-BNN samples. Based on their potential energies, the data points have been divided into 11 groups, with upper bounds given by $\boldsymbol{V_{Max}} = \{2.0;\ 4.0;\ 6.0;\ 8.0;\ 10.0;\ 15.0;\ 20.0;\ 25.0;\ 30.0;\ 50.0;\ 100.0\}$ eV. The minimum of the diatomic potential is here assumed as reference energy. The blue squares represent the number of *ab initio* data points contained in each energy group. 50
3.10 $1^1A'$ (lower part of the figure) and $2^5A'$ (upper part) PESs as a function of r_{A-B} for $\measuredangle O_AO_BO_C = 116.75°$ and $r_{B-C} = 2.282\ a_0$. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fits at the data geometrical configurations, bold lines to PIP-BNN MAP estimates, and the remaining narrow lines to the 50 samples per PIP-BNN. 51
3.11 3D views of the $2^5A'$ PESs at $\measuredangle O_AO_BO_C = 50°$ for $r_1 > r_3$, compared to the correspondent *ab initio* data points (gray dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. 52
3.12 3D views of the $1^1A'$ PESs at $\measuredangle O_AO_BO_C = 90°$ for $r_1 > r_3$, compared to the correspondent *ab initio* data points (gray dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. 53
3.13 Rovibrational state specific dissociation rate coefficients for the $2^5A'$ interactions at $T_{Tran} = 10\,000$ K. Green dots are obtained using Varga *et al.*'s fit, while the red ones are computed through the PIP-BNN MAP surface. 54
3.14 Evolution of O_2 mole fraction and their histogram at t = 0.7 μs, as predicted: by RVS StS ME (green solid line), and by CGM200 (overlapping green dashed line), both starting from Varga's fit; by RVS StS ME from PIP-BNN MAP (red dotted line); by CGM200 from the 50 PIP-BNN samples (gray narrow lines). The simulations have been performed at $T_{Tran} = 10\,000$ K with the rate coefficients from the $1^1A'$ PESs. . 55
3.15 (a): Evolution of O_2 mole fractions and their histogram at t = 3 μs. (b): Histogram of the vibrational relaxation times. Green lines are obtained by starting from Varga *et al.*'s fit; in particular: green solid lines are solutions of RVS StS ME; green dashed line are solutions of CGM200. Red dotted lines are solutions of RVS StS ME from PIP-BNN MAP; gray narrow lines are solutions of CGM200 from the 50 PIP-BNN samples. The simulations have been performed at $T_{Tran} = 10\,000$ K with the rates from the $2^5A'$ PESs. 56

3.16 Evolution of O_2 mole fraction, as predicted by ME (the green solid line) and CGM200 (the overlapping green dashed line), both starting from Varga *et al.*'s fit, and by: (a) CGM200 from the 50 PIP-BNN samples (grey narrow lines), with the two curves at extremes of the distributions being highlighted in blue and black; (b) CGM200 from the 50 PIP-BNN samples (grey narrow lines), with the blue and black curves in (a) being recomputed via ME calculations (blue and black dotted lines). The simulations have been performed at $T_{Tran} = 10\,000$ K with the rates from the $2^5A'$ PESs. 57
3.17 Vibrational relaxation times. The values for the $1^1A'$ and $2^5A'$ PESs are in black and in red, respectively. Of these, the solid lines are the values obtained using Varga *et al.*'s fits, the dots the values computed through the PIP-BNN MAP surfaces, and the vertical lines the three-sigma intervals corresponding to the 50 PIP-BNN samples. The green lines represent the values obtained by using all 9 PESs from Varga *et al.* and by either including (solid line) or excluding (dotted line) the exchange reactions. Pink dashed line and symbols identify the experiments by Ibraguimova *et al.* [53] and Breen *et al.* [18], respectively. 59
3.18 Dissociation (a) and exchange reaction (b) rate coefficients at QSS. Black and red colors corresponds to results from the $1^1A'$ and $2^5A'$ PESs, respectively. In particular, the solid lines report the values obtained using Varga *et al.*'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. The green line represents the values obtained by using all 9 PESs from Varga *et al.* Pink symbols correspond to Shatalov's experiments [4, 138]. 60
3.19 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples at $\measuredangle O_AO_BO_C = 50°$ for $r_1 > r_3$. The surface has been colored based on the absolute value of Pearson correlation coefficients [130] between the PES and the global exchange reaction rate coefficient at the equilibrium distribution of $T_{Tran} = 10\,000$ K (a), and between the PES and the vibrational relaxation time at $T_{Tran} = 10\,000$ K (b). The gray dots identify the locations suggested by the PIP-BNN for performing refinement *ab initio* calculations. The shades of blue in the PES depth are results of rendering effects and do not correspond to different values of Pearson coefficients. 62

4.1 Two-dimensional and three-dimensional views of the O_2 effective diatomic potential as a function of the distance between the oxygen atoms and of the rotational quantum number. In the right figure, the rovibrational levels are represented as dots at their inner and outer turning points, and they are colored based on their vibrational quantum numbers. 65
4.2 Schematics of the procedure for obtaining the data sets of RVS StS rate coefficients employed in the hybrid approach to the master equation. The final database combines the unmodified *ab initio* excitation rates and the dissociation rates encoded trough model reductions and decoded trough reconstructing averages. 71

5.1 O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy (*i.e.*, left side of the 3D diatomic potential in Figure 4.1(b)). The levels are colored based on their *ab initio* RVS dissociation rate coefficients, obtained through QCT calculations at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier. 76

5.2 RVS dissociation rate coefficients at T = 10 000 K. Black dots: rates from the *ab initio* data-set. Red dots: rates from the vibrationally-averaged data-set. Blue dots: rates from the CB-averaged data-set, based on 45 groups. 77
5.3 Scatter plot reporting the fitting error for the group-specific dissociation rate coefficients at T = 5 000 K (red dots), T = 10 000 K (green dots), and T = 15 000 K (blue dots). The 45 groups are constructed based on the energy-distance from the centrifugal barrier. The oblique external lines represent factor-two fitting errors (*i.e.*, absolute errors equal to 100%). 79
5.4 Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 7 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets. 80
5.5 Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 30 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets. 80
5.6 Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 100 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets. 81
5.7 Evolution of O_2 mole fraction in the 0-D simulation in which only dissociative collisions are allowed. The conditions of the isothermal and isochoric heat bath are T^0_{Int} = 10 000 K, $n^0_d = 3.33 \times 10^{23}$ m^{-3}, $[O_2]^0 = 95\%$, and T_{Tran} = 10 000 K. Black and blue lines overlap. Red and pink lines overlap. Black unbroken line: dissociation is computed based on the *ab initio* data set of rate coefficients. Red dash-dotted line: dissociation is computed based on the vibrationally-averaged data set. Blue dashed line: dissociation is computed based on the 45-groups CB-averaged data set. Pink unbroken line: from the fully-VS model. 83
5.8 (a): Evolution of O_2 mole fraction in a heat bath at T = 10 000 K. Black and blue lines overlap. Black unbroken line: fully-StS calculations. Red dotted line: RVS StS excitation coupled to VS dissociation. Blue dashed line: RVS StS excitation coupled to CB dissociation with 45 groups. Pink unbroken line: fully-VS calculations. Crosses: time instants at which $[O_2] = 50\%$. (b): Vibrational distribution functions at the time instants at which $[O_2] = 50\%$. Black dots: fully-StS calculations. Red dots: RVS StS excitation coupled to VS dissociation. 84
5.9 Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K, computed by means of fully-StS calculations. The time-snapshot is taken during QSS, at the instant at which $[O_2] = 50\%$ (*i.e.*, black/blue crosses in Figure 5.8(a)). Levels characterized by the same vibrational quantum numbers are connected through segments of equal colors. 85

5.10 Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K. The time-snapshots are taken during QSS, at the instants at which $[O_2]$ = 50% (*i.e.*, crosses in Figure 5.8(a)). Black dots: fully-StS calculations. Red dots: RVS StS excitation coupled to VS dissociation. Blue dots: RVS StS excitation coupled to CB dissociation with 45 groups. 86
5.11 Delays of O_2 depletion compared to the fully-StS simulations. Red dotted line: RVS StS excitation coupled to VS dissociation. Pink unbroken line: excitation and dissociation are both computed vibrationally-specific. All the other lines are obtained by computing excitation trough RVS StS and dissociation through the CB model with 45 (blue dash-dotted line), 20 (green dash-dotted line), 15 (purple dash-dotted line) and 10 (yellow dash-dotted line) groups. 87
5.12 Evolution of the energy rate constants for dissociation-vibrational (unbroken lines), dissociation-rotational (dotted lines), and dissociation-internal (dashed lines) energy transfer in a heat bath at T_{Tran} = 10 000 K. Black lines: from fully-StS calculations. Red lines: RVS StS excitation coupled to VS dissociation. Yellow lines: RVS StS excitation coupled to CB dissociation with 10 groups. 88
5.13 Energy rate constants for dissociation-vibrational (unbroken lines), dissociation-rotational (dotted lines), and dissociation-internal (dashed lines) energy transfer at QSS. Black lines: fully-StS calculations. Red lines: RVS StS excitation coupled to VS dissociation. Pink lines: fully-VS calculations. Yellow lines: StS excitation coupled to CB dissociation with 10 groups. 89
5.14 QSS dissociation rate coefficients for the O_2 molecule. Small pink circles: Shatalov experiments [139]. Green dotted line: Andrienko *et al.* [4]. Blue hexagram symbols: DMS simulations by Grover *et al.* [44]. Blue dashed line: present work, from fully-StS calculations with 16/3 correction factor applied in post-processing. Black continuous lines: present work, from fully-StS calculations. Yellow dash-dotted lines: present work, RVS StS excitation coupled to CB dissociation with 10 groups. Red dotted lines: present work, from RVS StS excitation coupled to VS dissociation. Pink unbroken lines: present work, from fully-VS calculations. 90
5.15 Percentages of O_2 depletion taking place before (black lines), during (green lines), and after (red lines) QSS, as functions of heat bath temperature. Three different scenarios are analyzed, all with T^0_{Int} = 300 K and $[O_2]^0$ = 95%. Dashed lines: $n^0_d = 3.33 \times 10^{24}\ m^{-3}$ (p^0 = 13806.49 Pa). Unbroken lines: $n^0_d = 3.33 \times 10^{23}\ m^{-3}$ (p^0 = 1380.65 Pa). Dotted lines: $n^0_d = 3.33 \times 10^{22}\ m^{-3}$ (p^0 = 138.07 Pa). 91

A.1 $1^1A'$ (red lower parts of the Fig.s) and $2^5A'$ (black upper parts) PESs at four different sets of geometries. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fit at the data geometries, bold lines to PIP-BNN MAP, and the remaining narrow lines to the 50 PIP-BNN samples. (a): $\measuredangle O_A O_B O_C = 60°$ and $r_1 = 2.646a_0$. (b): $\measuredangle O_A O_B O_C = 70°$ and $r_1 = 2.646a_0$. 98
A.2 $1^1A'$ (red lower parts of the Fig.s) and $2^5A'$ (black upper parts) PESs at four different sets of geometries. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fit at the data geometries, bold lines to PIP-BNN MAP, and the remaining narrow lines to the 50 PIP-BNN samples. (a): $\measuredangle O_A O_B O_C = 110°$ and $r_1 = 2.268a_0$. (b): $\measuredangle O_A O_B O_C = 170°$ and $r_1 = 2.268a_0$. 99

A.3 3D views of the $1^1A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 100°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 100°$, PIP-BNN. . 100
A.4 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 60°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 60°$, PIP-BNN. . . 101
A.5 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 65°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 65°$, PIP-BNN. . . 102
A.6 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 70°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 70°$, PIP-BNN. . . 102
A.7 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 80°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 80°$, PIP-BNN. . . 103
A.8 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 90°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 90°$, PIP-BNN. . . 103
A.9 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the ab initio data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 100°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 100°$, PIP-BNN. . 104
A.10 Rovibrational state specific exchange rate coefficients for the $2^5A'$ interactions at $T_{Tran} = 10\,000$ K, computed as $k_i^E = \sum_{j=1}^{N_{RV}} k_{ij}^E$, where i corresponds to the initial level, j to the final one, and N_{RV} to the number of O_2 ro-vibrational levels. Green dots are obtained using Varga *et al.*'s fit, while the red ones are computed through the PIP-BNN MAP surface. 105

A.11 Evolution of O_2 dissociation global rates, as predicted by ME (green solid lines) and CGM200 (green dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted lines), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig.s report the histograms of the respective QSS rates. The simulations have been performed considering only $2^5A'$ interactions at $T_{Tran} =$ 10 000 K. The dissociation global rate is computed as $\bar{k^D} = \sum_i \sum_j k_i^D n_i(t)/n_{O_2}(t)$; i identifies the O_2 initial state, in case of ME calculations, or the initial group in case of CGM; k_i^D corresponds to the dissociation rate associated with the i-th state/group, and n_i represents the population of the i-th state/group. 106
A.12 Evolution of O_2 exchange global rates, as predicted by ME (green solid lines) and CGM200 (green dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted lines), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig.s report the histograms of the respective QSS rates. The simulations have been performed considering only $2^5A'$ interactions at $T_{Tran} =$ 10 000 K. The exchange global rate is computed as $\bar{k^E} = \sum_i \sum_j k_{ij}^E n_i(t)/n_{O_2}(t)$; i identifies the O_2 initial state, in case of ME calculations, or the initial group in case of CGM, and j the final one; k_{ij}^E corresponds to the exchange rate from the i-th state/group to the j-th, and n_i represents the population of the i-th state/group. . . 107
A.13 Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig. on the right reports the histogram of such samples at $t = 2s$. The PIP-BNN adopted is the one constructed for $1^1A'$, and the simulations are performed at $T_{Tran} = 2\,500$ K. 108
A.14 Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig. on the right reports the histogram of such samples at $t = 10s$. The PIP-BNN adopted is the one constructed for $2^5A'$, and the simulations are performed at $T_{Tran} = 2\,500$ K. 109
A.15 Rotational relaxation times for $1^1A'$ (in black) and $2^5A'$ (in red) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. 110
A.16 Equilibrium and QSS dissociation rates for the $1^1A'$ (in black and green) and $2^5A'$ (in red and blue) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. At $T_{Tran} = 2\,500$ K no QSS is present for any of the two PESs. In the top-right square, the main Fig. is zoomed at $T_{Tran} = 10\,000$ K and $T_{Tran} = 7\,500$ K. 111
A.17 Equilibrium and QSS exchange rates for the $1^1A'$ (in black and green) and $2^5A'$ (in red and blue) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. At $T_{Tran} = 2\,500$ K no QSS is present for any of the two PESs. 112

A.18 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the vibrational relaxation time at T_{Tran} = 10 000 K. (a): $\measuredangle O_AO_BO_C = 50°$. (b): $\measuredangle O_AO_BO_C = 110°$. (c): $\measuredangle O_AO_BO_C = 175°$. 113
A.19 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the vibrational relaxation time at T_{Tran} = 10 000 K, obtained by artificially excluding the exchange processes. (a): $\measuredangle O_AO_BO_C = 50°$. (b): $\measuredangle O_AO_BO_C = 110°$. (c): $\measuredangle O_AO_BO_C = 175°$. 114
A.20 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the rotational relaxation time at T_{Tran} = 2 500 K. (a): $\measuredangle O_AO_BO_C = 50°$. (b): $\measuredangle O_AO_BO_C = 110°$. (c): $\measuredangle O_AO_BO_C = 175°$. 115

B.1 O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy. The levels are colored based on the RVS dissociation rate coefficients from the vibrationally-averaged data set at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier. 116
B.2 Scatter plot representing the error generated by the VS dimensionality reduction of the RVS dissociation rate coefficients at T = 10 000 K. 117
B.3 O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy. The levels are colored based on the RVS dissociation rate coefficients from the CB-averaged data set at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier. 118
B.4 Scatter plot representing the error generated by the CB dimensionality reduction of the RVS dissociation rate coefficients at T = 10 000 K. 119
B.5 Delays of O_2 depletion compared to the fully-StS simulations. The conditions of the isothermal and isochoric heat bath are T^0_{Int} = 300 K, $n^0_d = 3.33 \times 10^{23}$ m^{-3}, $[O_2]^0$ = 95%, and T_{Tran} = 10 000 K. Red dotted line: RVS StS excitation coupled to VS dissociation. Pink unbroken line: excitation and dissociation are both computed vibrationally-specific. All the other lines are obtained by computing excitation trough RVS StS and dissociation through CB model, with 45 (blue dash-dotted line), 20 (green dash-dotted line), 15 (purple dash-dotted line) and 10 (yellow dash-dotted line) groups. For these last four calculations, fitted dissociation rate coefficients are used. 120
B.6 Evolution of O_2 mole fraction (blue dash-dotted line) and global dissociation rate coefficient (black unbroken line) in a heat bath at T_{Tran} = 5 000 K and $n^0_d = 3.33 \times 10^{24}$ m^{-3} (P^0 = 13806.49 Pa). The simulation is performed fully-StS. The two vertical lines represent the time instants at which the quasi-steady state is assumed starting and terminating. 121
B.7 Evolution of O_2 mole fraction (blue dash-dotted line) and global dissociation rate coefficient (black unbroken line) in a heat bath at T_{Tran} = 20 000 K and $n^0_d = 3.33 \times 10^{23}$ m^{-3} (P^0 = 1380.65 Pa). The simulation is performed fully-StS. The two vertical lines represent the time instants at which the quasi-steady state is assumed starting and terminating. 122

B.8 Rovibrational-specific (RVS) dissociation rate coefficients at T_{Tran} = 5 000 K (a), T_{Tran} = 10 000 K (b), and T_{Tran} = 20 000 K (c) for the O2+C system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 49 groups for the O_2 molecule. 123
B.9 Rovibrational-specific (RVS) dissociation rate coefficients at T_{Tran} = 5 000 K (a), T_{Tran} = 10 000 K (b), and T_{Tran} = 20 000 K (c) for the CO+O system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 83 groups for the CO molecule. 124
B.10 Rovibrational-specific (RVS) dissociation rate coefficients at T_{Tran} = 5 000 K (a), T_{Tran} = 10 000 K (b), and T_{Tran} = 20 000 K (c) for the N_2+N system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 61 groups for the N_2 molecule. 125
B.11 Rovibrational-specific (RVS) dissociation rate coefficients at T_{Tran} = 5 000 K (a), T_{Tran} = 10 000 K (b), and T_{Tran} = 20 000 K (c) for the N_2+O system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 54 groups for the N_2 molecule. 126
B.12 Rovibrational-specific (RVS) dissociation rate coefficients at T_{Tran} = 5 000 K (a), T_{Tran} = 10 000 K (b), and T_{Tran} = 20 000 K (c) for the NO+N system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 46 groups for the NO molecule. 127

List of Abbreviations

ADVI	Automatic Differentiation Variational Inference algorithm
AFOSR	U.S. AirForce Office of Scientific Research
BNN	Bayesian Neural Network
CB	Centrifugal-Barrier-Based
CFD	Computational Fluid Dynamics
CG	Coarse-Grained
CG-QCT	Coarse-Grained Quasi-Classical Trajectory method
CGM	Coarse-Grained Model
DMS	Direct Molecular Simulation method
DSEC	Dynamical Scaled External Correlation
GP	Gaussian Processes
IMLS	Interpolating Moving Least-Squares
LHS	Latin Hyper-Cube Sampling Method
LHS50	50 samples obtained via LHS
MAP	Maximum a Posteriori Probability
MC	Monte Carlo method
ME	Master Equation
ML	Machine Learning
MLE	Maximum Likelihood Estimation
MT	Multi-Temperature Models
NASA	National Aeronautics and Space Administration
NN	artificial Neural Networks

ODE	Ordinary Differential Equation
PDF	Probability Density Function
PES	Potential Energy Surface
PIP	Permutation Invariant Polynomials
PIP-BNN	Permutation Invariant Polynomials Bayesian Neural Network
PIP-NN	Permutation Invariant Polynomials Neural Network
QCT	Quasi-Classical Trajectory method
QoI	Quantity of Interest
QSS	Quasi-Steady-State
ROM	Reduced Order Model
RKHS	Reproducing Kernel Hilbert Space
RMSE	Root-Mean-Square Error
RVS	RoVibrational-Specific
SPES	Stochastic Potential Energy Surface
StS	State-to-State
TPS	Thermal Protection System
UQ	Uncertainty Quantification
VS	Vibrational Specific
VDF	Vibrational Distribution Function
WKB	Wentzel–Kramers–Brillouin approximation
XMS-CASPT2	Multi-State Complete Active SPace second-order Perturbation Theory

List of Symbols

$\boldsymbol{\alpha}$	Vector of the ML-based surrogate model parameters
$\boldsymbol{\beta}$	Hyperparameters controlling the PDFs of $\boldsymbol{\alpha}$
$\boldsymbol{\gamma}$	Hyperparameters controlling the representation of the SPES error
ϵ_i	Internal energy of the i-th rovibrational level
ϵ_i^D	Energy deficit of the i-th rovibrational level from the J-dependent barrier
ϵ_V	SPES error
ϵ_{V_i}	SPES error at the i-th Data Point
ϵ_V^{log}	Logarithm of ϵ_{V_i}
η	Free parameter in the CB grouping strategy
$\boldsymbol{\theta}$	Overall set of SPES' parameters and hyperparameters
λ	First tunable parameter for generating the PIP intermediate variables
$\xi_i(T)$	Contribution of the i-th level to the internal partition function at temperature T
$\rho_{x,y}$	Pearson correlation coefficient between the quantities x and y
σ_{b_i}	Standard deviation of the prior distributions corresponding to $\boldsymbol{b}_i$
σ_{Like}	Standard deviation of the likelihood function
σ_{W_i}	Standard deviation of the prior distributions corresponding to $\boldsymbol{W}_i$
τ_M	Relaxation characteristic time of the molecule internal mode M
τ_i^D	Dissociation characteristic time of the i-th level
$\phi(...)$	Function converting the relative distances to the PIP intermediate variables
b_i^k	Bias parameter corresponding to the k-th neuron of the (i)-th layer
C^{DM}	Constants for the dissociation-energy coupling of mode M
$\boldsymbol{D}$	Set of *ab initio* data points

e_M	Energy of the internal mode M
E_P^D	Group-specific averaged distance from the centrifugal barrier
$\hat{\boldsymbol{E}}^D$	Energy grid in the CB grouping strategy
$f(T)$	Boltzmann distribution at the temperature T
$\mathcal{F}_i$	Activation function of the neurons composing the i-th layer
g_i	Degeneracy of the i-th rovibrational level
g^e	Degeneracy of the molecular electronic ground state
$\boldsymbol{G}$	PIP symmetrized polynomial vector
$\hbar$	Reduced Planck constant
Half-$\mathcal{N}(\mu, \sigma^2)$	Half-normal distribution of mean μ and standard deviation σ
J	Rotational quantum number
k_B	Boltzmann's constant
k_i^D	Dissociation reaction rate coefficient for the state i
k_i^R	Recombination reaction rate coefficient for the state i
$\bar{k}_i^D$	Dissociation reaction rate coefficient for the state i, reconstructed after being grouped
$\bar{k}_i^R$	Recombination reaction rate coefficient for the state i, reconstructed after being grouped
k_{ij}	Excitation reaction rate coefficient between states i and j
k_{ij}^E	Exchange reaction rate coefficient between states i and j
k_{ij}^I	Inelastic reaction rate coefficient between states i and j
$K_i^{Eq}(T)$	Equilibrium constant for the level i at temperature T
K_P^D	Dissociation reaction rate coefficient for the group P
K_P^R	Recombination reaction rate coefficient for the group P
$\mathcal{I}$	Set of rovibrational levels contained in the molecule
$\mathcal{I}_P$	Set of rovibrational levels contained in the p-th group
L	Number of NN hidden layers
m	Molecular reduced mass
$\mathcal{M}$	SPES model form
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution of mean μ and standard deviation σ
n_X	Number density of chemical species X

N_D	Number of data points
N_i	Number of neurons in the i-th hidden layer
N_{LHS}	Number of Latin Hypercube samples
$\hat{O}$	Symmetrization operator of the PIP Layer
$\boldsymbol{P}$	Vector of PIP intermediate variables
p	Pressure
$Q_X(T)$	Internal partition function of the species X
$Q_X^t(T)$	Translational partition function of the species X
$\boldsymbol{r}$	Vector of atom distances
r_e	Second tunable parameter for generating the PIP intermediate variables
t	Time
T_{Int}	Internal temperature
T_{Tran}	Translational temperature
$\mathcal{U}(a,b)$	Uniform distribution between the extremes a and b
v	Vibrational quantum number
V	Potential energy predicted by the SPES
$\hat{V}$	Potential energy predicted by the ML-based PES
V^0	Potential energy reference value
V^{Diat}	Diatomic contribution to the potential energy
V^{MB}	Multi-body contribution to the potential energy
$W_i^{j,k}$	Weight parameter connecting the j-th neuron of the (i-1)-th layer and the k-th neuron of the (i)-th layer
x^* and x_{Eq}	Equilibrium value of the quantity x
$\boldsymbol{x_{LHS}}$	Latin Hypercube sampled values for the quantity x
x_{MAP}	Maximum a posteriori estimate of the quantity x
x_{QSS}	QSS value for the quantity x
$\bar{x}$	Averaged value for the quantity x
$\boldsymbol{y}$	Vector of outputs from the computational model

Chapter 1

Introduction

1.1 Motivation

A vehicle traveling at hypersonic speed undergoes severe convective and radiative heat fluxes, the root causes of which can be detected in the air-/space-craft's kinetic energy being partially converted into the internal modes of the surrounding flow. The shock wave standing in front of the vehicle enhances the strength of the collisions between the atoms and the molecules composing the gas mixture; consequently, the populated high-energy quantum states of the particles strongly contribute towards changing the gas chemical composition and emitting radiation. The heat fluxes are functions of the gas mixture's properties (*e.g.*, the composition of the planetary atmosphere, flight altitude, and possible ablated materials) and of the vehicle's characteristics (*e.g.*, shape, dimensions, and configuration) [1, 118]. Because of the coupling effects implicated by radiation and ablation, the multidisciplinary nature of the problem, and the multiple scales affected, the predictions of the heat fluxes and the related material responses are the results of extremely demanding and challenging simulations, from both the computational and the experimental points of view [41]. Nevertheless, a reliable characterization of the thermal stresses experienced by the hypersonic vehicle is required in order to design an appropriate thermal protection system (TPS) [38]. A TPS is a set of components that shields the air-/space-craft, or some of its regions, from extreme heating. In most hypersonic flight applications, such as atmospheric (re-)entry, the TPS is a single-point-of-failure (*i.e.*, the mission accomplishment vitally depends on the system's success). The system is then safety-critical and, at the same time, hard to be modeled in all its environment. These two characteristics enforce the adoption of generous margins of safety during the design process, intending to account for possible uncertainties and errors in predicting the quantities of interest (QoIs). When an improvement in modeling the problem translates into an uncertainty reduction

on the heat flux computations, the previous margin of safety becomes over-conservative and can be relaxed, with remarkable benefit for the mission effectiveness and/or costs.

For today's Hypersonics and many other scientific communities (*e.g.*, plasma physics and material science), the predictive science paradigm consists of avoiding empiricism by integrating first-principles quantum chemistry into computational models. The resulting methodologies are defined as *ab initio* approaches [22, 28, 36, 60, 114, 120]. One example is the *PES-to-rate coefficients* strategy [60], which is becoming the method of choice for first-principles Hypersonics. This approach, which will be described in detail starting from the next section, can be briefly summarized as follows. In order to simulate a single collision between some of the particles composing the gas mixture surrounding the vehicle, a large number of geometrical arrangements are first selected for the colliding atoms. The solutions of the electronic Schrödinger equation are computed at such spatial points, and the so obtained energies are then fitted to an analytical expression: the potential energy surface (PES). A PES represents the energy of atomic interactions at arbitrary geometries, and its gradients can be used as source terms in the Hamiltonian equations for computing the trajectories of the colliding particles, based on the quasi-classical trajectory method (QCT) [66]. By simulating an ensemble of collisions starting from different initial conditions (*i.e.*, target-projectile orientations) and repeating the process for a set of molecules' initial quantum states and target-projectile relative velocities, it is then possible to obtain the state-to-state (StS) cross-sections. Finally, by integrating each cross-section over a Maxwellian distribution of translational energies, StS rate coefficients can be computed. These quantities represent the rates at which the molecule quantum states become populated or depleted due to collisional processes. By employing them in computational fluid dynamics (CFD) calculations, the nonequilibrium kinetic processes can be simulated, and the QoIs for the TPS design (*e.g.*, heat fluxes on the vehicle's surface) can be finally predicted. Theoretically, the introduction of fully physics-based *ab initio* methodologies should result in predictive capabilities that have never been achieved in the past. In practical terms, the increase in the model complexity has two main side effects.

Firstly, more capacity also translates into more physics that could be potentially misrepresented, in analogy with the idea of overfitting in statistics. Only when a systematic and consistent approach to quantifying errors and uncertainties affecting these augmented models is completed, will

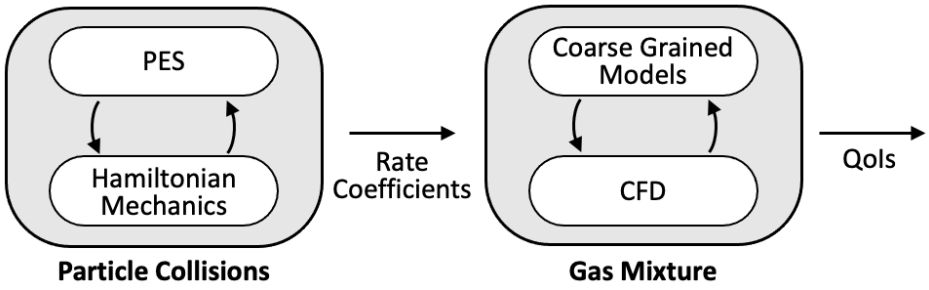


Figure 1.1: Schematics of the *ab initio* computational framework for nonequlibrium flows.

an actual relaxation of the aforementioned margins of safety for the TPS design be fully justified. This fact motivates the effort that the hypersonic community has been addressing for the last 10 years into representing and quantifying the uncertainties related to CFD [15, 51], radiation [64, 100, 101, 115, 152], material response and ablation [127]. Simultaneously, less has been done to characterize the uncertainties on modeling the rovibrational kinetic processes.

Secondly, improving the microscopic resolution on the nonequilibrium flow around the hypersonic vehicle increases the computational cost of already dramatically expensive simulations. Resolving each of the rovibrational states in the gas mixture is computationally impracticable at the moment. As a consequence, the energy states need to be clustered before performing CFD simulations, and the nonequilibrium properties are computed at the group of states level based on the so-called coarse-grained (CG) models [48, 85]. The resulting *ab initio* computational framework is represented in Fig. 1.1. In order to preserve the potential accuracy gained from the first-principles calculations of the atomic interactions, the rovibrational states must be grouped consistently with the nonequilibrium kinetics.

1.2 Literature Review

1.2.1 *Ab Initio* Potential Energy Surfaces and Associated Uncertainties

The *ab initio* methodologies mentioned above rely on the construction of the potential energy surfaces (PESs), followed by scattering calculations based on quasi-classical dynamics [60, 66, 134] or quantum mechanics [61, 163, 164, 167]. To this aim, the energies for a large number of geometrical configurations of the atoms included in the chemical system are computed by solving the

electronic Schrödinger equation and fitted to an analytical expression: the PES [60]. In recent years, the availability of ever-increasing computational resources has enabled the construction of many high-fidelity *ab-initio*-based surfaces for a large number of chemical systems. Examples include the set of surfaces for O_2+O [150], N_2+N [39, 58], NNO [33], NOO [132], COO [136], and CNO [72]. In QCT calculations [60, 66], the PESs are used for producing StS rate coefficients by simulating an ensemble of collisions starting from different initial conditions (*i.e.*, target-projectile orientations, initial quantum states, etc.). The availability of *ab initio* rate coefficients for elementary collisional processes has enabled the development of rovibrational-specific (RVS) StS models [4, 5, 69, 70, 82, 114]. These approaches allow for the detailed study of kinetic processes under strong nonequilibrium conditions, for which the populations of the internal energy levels significantly depart from the equilibrium Boltzmann distribution. Many research areas, such as combustion, atmospheric chemistry, astrochemistry, and hypersonics, rely on *ab initio* rate coefficients for high-resolution simulations. In some applications, the outcomes of these calculations are employed for decision-making or design processes, as in the case of TPSs for (re)-entering capsules.

Despite its importance, very little work has been done so far to assess the reliability of the overall *ab initio* methodology (*i.e.*, from PES to macroscopic quantities). Recently, Jaffe *et al.* made a first attempt at validating the PESs for N_2+N and N_2+N_2 from NASA Ames [57, 58] and the surface for N_2+N_2 from the University of Minnesota [120] by comparing thermal and quasi-steady-state (QSS) [81] dissociation rates to shock-tube experiments [56]. The good agreement between the *ab initio* calculations and the experimental data justifies the need for the use of high-fidelity first-principles approaches. However, while the results from these two previous studies suggest a not negligible sensitivity to the accuracy of the PES, a more systematic framework is necessary to quantify the uncertainty associated with the *ab-initio*-based methodology and its potential impact on the nonequilibrium kinetics.

The *ab initio* approach, briefly outlined above, bears multiple sources of uncertainties:

- *Solution of the electronic Schrödinger equation*: the values of the interaction potential at the selected atomic spatial configuration are not known exactly, due to unavoidable truncation errors in the series expansions, or selection of the atomic orbital basis set, and/or simplifying assumptions in the mathematical formulation (*e.g.*, approximation of the correlation

energies);

- *PES fitting/interpolation technique*: in order to be applied to the simulation of atomic collisions, the potential energy needs to be formulated as a continuous surface, differentiable with respect to the atomic spatial coordinates. The procedure of constructing the PES starting from the *ab initio* data points generally varies from one research group to another. Traditionally, it is either based on a combination of compact support analytical functions [58, 136], or on ad hoc fitting methods (*e.g.*, permutation invariant polynomials (PIPs) [17]), or on ad hoc interpolation techniques (*e.g.*, reproducing kernel Hilbert space (RKHS) [49, 50, 146] and interpolating moving least-squares (IMLS) [95, 125]). In the past twenty years, the construction of *ab initio* PESs has benefited from the use of machine learning (ML), allowing for the automation of the fitting procedure. Artificial neural networks (NNs) [11, 13], Gaussian processes (GPs) [9, 30, 47, 71], and many other ML techniques successfully approached the problem as regression analysis. For a detailed description of these strategies, the interested reader can refer to the review articles [19, 133].

- *Determination of the cross-sections by means of QCT calculations*: to predict the values of the RVS StS cross-sections, the QCT method neglects quantum effects, such as zero-point energy conservation and tunneling [79]. Moreover, the Monte Carlo (MC) integration performed over the space of target-projectile initial conditions generates an error inversely proportional to the square root of the number of trajectories simulated.

- *Fitting of rate coefficients*: StS rate coefficients are obtained over a range of translational temperatures. An analytical function (*e.g.*, Arrhenius form) is then used for extrapolating their values at a generic temperature.

Of the possible sources of uncertainty presented in the list above, the ones due to the MC integration and the fitting of rate coefficients can be easily quantified [135]. On the contrary, the representation of uncertainties generated by the construction of PESs from *ab initio* data constitutes the focus of the present work.

For PESs generated via Gaussian process regression, the uncertainty bound intrinsically accompanies the prediction of the potential energy at a generic atomic configuration [74, 126]. Recently,

this distinctive attribute of GPs has been exploited for constructing PESs based on active learning approaches [46, 147]. One of the most successful examples is the work by Vargas-Hernández *et al.* [151], which reconstructed the H_3 and OH_3 PESs from the observed exchange reaction probabilities of H_2 and HO, as a solution of an inverse quantum scattering problem. Unfortunately, GPs have the drawback of being significantly slow to be evaluated, and their cost escalates with the number of training points [65, 74, 124, 126] rendering them impractical to use in many applications. An example is given by the simulation of nonequilibrium hypersonic flows. These calculations generally rely on quasi-classical dynamics [114, 134], which require the evaluation of the PES gradients. For a GP-based PES, the cost of evaluating derivatives is $\mathcal{O}(nd)$, where n is the number of training points, and d is the surface dimensionality [35, 160]. Moreover, at high temperatures, all the rovibrational energy levels of the molecules become significantly populated. This means that tens of millions of reaction probabilities need to be computed [4, 114], requiring billions of trajectory calculations.

Neural networks constitute an efficient alternative to GPs [65, 83]. However, in their classical formulation, they are deterministic tools and do not quantify the uncertainty in the prediction. When constructed based on an NN, the training process of the PES relies on the minimization of a *loss function* (*e.g.*, root-mean-square error, RMSE) for a subset of the *ab initio* data points. This procedure corresponds to a maximum likelihood estimation (MLE) of the parameters, which is just a single point prediction of the optimal values for weights and biases [109]. Following this approach to ML, the error associated with the fit can only be estimated at the data points (*i.e.*, only where the *ab initio* data energy is known), and cannot be characterized at a generic spatial configuration. In addition, QCT calculations require the use of spatial gradients of the PESs, and a simple comparison of the predicted potential energies with the *ab initio* data points is inadequate to assess the reliability of the surface.

Contextualizing the discussion by Kennedy and O'Hagan [67], the sources of uncertainties on an NN-based PES can be distinguished in:

- *Error in the data set*, which is the error in the solution of the electronic Schrödinger equation;
- *Parameter uncertainty.* The algorithm adopted for training the surrogate PES estimates parameter values that are only approximations of their optima. Moreover, these estimates

are generally sensitive to hyper-parameters (*e.g.*, learning rate), and initialization (*e.g.*, initial values of weights and biases) [80];

- *Model inadequacy* (also called *structural inadequacy*), which is the error due to the deficiencies or limitations of the model structure (*e.g.*, symmetry functions, number of layers, number of neurons per layer, activation functions, etc.). The *universal approximation property* of NNs states that "Multilayer feedforward networks with as few as a single hidden layer and an appropriately smooth hidden layer activation function are capable of an arbitrarily accurate approximation to an arbitrary function and its derivatives." [52]; in practice, the number of neurons and the approximation capabilities are constrained, on the one hand, by the risk of over-fitting due to the limited training data available, and, on the other hand, by the cost of the PES evaluations during molecular simulations.

Li *et al.* [80] already identified the need for robust and systematic uncertainty quantification (UQ) frameworks to prove the reliability of NN-based potentials in atomistic simulations. These authors, however, focused their studies on the inadequacies affecting the PES training procedure (*i.e.*, inaccuracies of NN hyper-parameters, such as learning rate and initial weights), rather than on characterizing the parameter and model uncertainties proper of the neural network itself. Recent attempts have been made in the direction of representing the variability of NN predictions. Some examples are the negative of the squared difference surface (*i.e.*, the difference between two independent NNs) by Lin *et al.* [83] and the standard deviation of a NN committee (*i.e.*, the standard deviation of the predictions from multiple NNs initialized using different weights) by Zhang *et al.* [166] However, these techniques compare differently-trained NNs only between each other, not to the *ab initio* data. As a result, in the best-case scenario, they only estimate the parameter contribution to the uncertainty. As stated by Zhang *et al.*: "There may be situations in which the physics is poorly described by a model, yet the corresponding ensemble of predictions has small variance" [166].

1.2.2 Reduced Order Modeling for Nonequilibrium Kinetics

Over the past decades, the availability of unprecedented computational resources has allowed the use of quantum-chemistry databases for constructing the physical models adopted to describe the

nonequilibrium chemical relaxation processes that take place in hypersonic flow regimes. Often, however, these efforts are limited to the calculation of some empirical parameters of outdated legacy nonequilibrium models, many times constructed on overly simplified, and sometimes even invalid, assumptions. Even if computed from first principles and using the most sophisticated PESs, these models inherit an intrinsic bias that can introduce large inaccuracies in the predictions.

In the vast majority of the nonequilibrium models, the rotational and vibrational energy modes are decoupled, and the population of the rotational levels is assumed to be in equilibrium at a temperature (*i.e.*, rotational temperature) close to the translational one [117]. Moreover, effective rate coefficients are derived by assuming a dependence on an empirical temperature [117]. Traditionally considered the method of choice, the VS model [21, 26] falls in this category as it relies on the assumption of rotational equilibrium. Several *ab initio* databases of rate coefficients have been computed based on this approach [2, 3, 12, 24, 37], and some of them have been used for further modeling [45, 113, 140]. Even if able to correctly account for deviation from equilibrium on the vibrational population, the VS model provides an overly simplified description of the rotational relaxation, which compromises its accuracy. As an example, Figure 1.2 shows the comparison of the composition profiles obtained studying the relaxation of oxygen molecules in an isothermal chemical reactor with three models consistently derived using the same *ab initio* PESs [150]: the VS model, the RVS StS model [4, 68, 82, 114, 116], and the direct molecular simulation (DMS) model [44, 148]. This figure unequivocally shows that the VS strategy is unable to reproduce the results of the more advanced approaches. Similar conclusions were observed for nitrogen-based systems, N_2-N [93, 94, 116] and N_2-N_2 [89, 91].

In RVS StS and DMS models, the transitions between the internal levels of the molecules are explicitly considered, and no assumption concerning the population of the states is required (*i.e.*, no internal temperature is needed). This allows for the kinetics of the rovibrational levels to be correctly described, thus justifying the results shown above. However, for the RVS StS and DMS models, intensive requirements in terms of computational power have not allowed for efficient implementation in multidimensional codes, and most calculations found in the literature have been limited to 0-D and 1-D configurations. To overcome the limitations of the VS models while retaining computational efficiency, a new class of nonequilibrium reduced order models (ROMs), referred

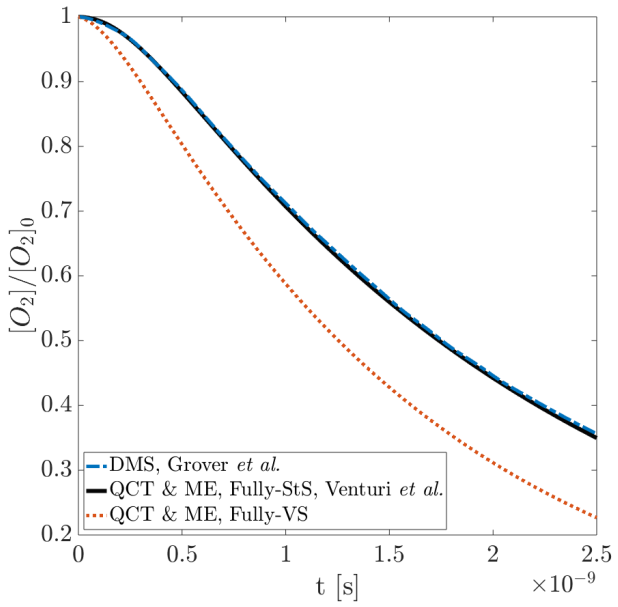


Figure 1.2: Evolution of O_2 mole fraction normalized by its initial condition. Blue dash-dotted line: DMS simulation by Grover *et al.* [44] Black continuous line: QCT and master equation, both rovibrationally state-to-state, by Venturi *et al.* [160] Red dotted line: QCT and master equation, both vibrationally state-to-state. To be consistent with Grover *et al.*'s calculations, no recombination and no dissociation correction factor [110] have been considered.

to as coarse-grained (CG) models, have been developed over the years [48, 85, 93, 104, 106, 129]. In the CG approach, the energy levels are clustered in macroscopic groups (also referred to as bins), and the populations within each group are prescribed by Boltzmann distributions [85]. Based on this definition, the VS model can be considered as a particular case of the CG approach. In 1987, Haug *et al.* studied dissociation of para-H_2 due to Ar collisions at 4 500 K [48] and compared the vibrational-specific model, rotational-specific model, and a reduced order approach that coupled rotation and vibration. They concluded that "inclusion of rovibrational coupling in a lumping model is more important than merely including a larger number of lumped states" [48]. More recently, CG models were solely based on an energy-based binning strategy [94, 104], underestimating the importance of a physics-based construction of the groups. Recently, however, a number of authors have demonstrated that the accuracy of the method is very sensitive to the way internal levels are clustered [102, 129]. Sahai *et al.* [129] leveraged machine learning techniques to construct an automated kinetics-based grouping method. Throughout the thermo-chemical relaxation, the approach demonstrated significant improvement in reproducing the main features of the nonequilibrium dis-

tribution function obtained from rovibrational StS models. However, since dissociation processes were not included in the clustering algorithm, the method required some ad-hoc fixes. This fact is not surprising since the coarse-grained approach implicitly assumes that the excitation processes among states within the same group are much faster than the dissociation reactions [88].

1.2.3 The O_2+O Chemical System

The O_2-O chemical system is fundamental for the scientific community in a number of practical applications involving reacting flows, ranging from hypersonic flight (*e.g.*, Earth re-entry) to atmospheric chemistry (*e.g.*, ozone formation and depletion [8, 96, 121]). In 2019, Venturi *et al.* [160] compared the results of isothermal heat bath excitation using QCT-based RVS StS rates to the direct molecular simulations (DMS) performed in the same year by Grover *et al.* [44]. Both studies relied on the PESs from Varga *et al.* [150], and, despite the different methodologies adopted, the resulting QoI values (*i.e.*, thermal and QSS dissociation rates, vibrational time constants, and mole fractions) are in excellent agreement over the entire range of translational temperatures. However, these predictions do not agree with the ones by Andrienko *et al.* [4], obtained in 2016 using RVS StS QCT and heat bath simulations based on the surface from Varandas *et al.* [149]. This fact suggests a non-negligible sensitivity of the gas QoIs to some of the O_2+O PES regions, consistent with what was previously found by Kulakhmetov and coworkers [76]. These last authors compared the results obtained using the surface by Varandas *et al.* to the ones generated through a simplified potential. While the thermal dissociation rate coefficients were found in good agreement for temperatures up to 10 000 K, Kulakhmetov *et al.* showed a strong sensitivity of the vibrational and rotational excitation dynamics on the topological details of the PES.

The nine PESs by Varga et al. [150] are employed in this work. The set of surfaces characterizes the dynamics O_2+O collisions in high-temperature environments encountered in hypersonic flows. In their paper, Varga et al. write [150]: "We make special efforts to make the surfaces realistic up to very high energy as required to treat dissociation of O_2 in high-energy collisions with ground-state O. The surfaces should also be useful for studying vibrational and rotational energy transfer in spin-conserving O_2+O collisions." In 2019, Grover *et al.* [44] showed that all the nine PESs of the O_2+O system are relevant for the characterization of the nonequlibrium kinetics in

high-temperature ranges. Previous studies (*e.g.*, Dawes *et al.* [31, 32] and Powell *et al.* [121]) only focused on the set of singlet O_3 PESs, which has only a small statistical weight (1/9), or on the surfaces relevant to ozone photodissociation (*e.g.* Grebenshchikov *et al.* [43]). Recently, in their comparison to further validate Varga et al.'s PESs, Geistfeld and Schwartzentruber [40] found good agreement between their QCT calculations and the experimental data by Lahankar *et al.* [77].

1.3 Objective of the Thesis

This thesis presents a novel approach to *ab initio* Hypersonics that leverages machine learning, uncertainty quantification, Data Science, and reduced order models for improving the predictive capabilities of the computational simulations. In this framework, schematized in Fig. 1.3, ML algorithms are designated as efficient and automatized techniques for constructing surrogate models (*e.g.*, ML-based PESs) and assisting the development of reduced-order models (*e.g.*, as for the unsupervised methods by Sahai *et al.* [129] for clustering rovibrational levels). Surrogate models are here intended as continuous and differentiable hyper-surfaces obtained from a discrete set of data points to decrease the complexity of the computational pipeline. As mentioned above, in Hypersonics, the simulation outcomes are used for design and decision-making processes in safety-critical applications. For this reason, UQ acts on the ultimate QoI by extending its information content from a single-value prediction to a probability distribution. In order to achieve this, the ML-based surrogate is augmented to a probabilistic (*i.e.*, stochastic) model, and Bayesian inference is used for solving the inverse problem starting from the data points (*i.e.*, training phase, also referred to as calibration). When the data set size is relatively small (*e.g.*, the *ab initio* data sets for atom configurations and interaction energies), this extension to stochastic is also an effective strategy for reducing the risk of overfitting, as will be motivated in the following chapter. Moreover, by correlating the uncertainty on the QoI to the one on the stochastic surrogate, it is possible to highlight the most important regions of the distribution of hyper-surfaces. If needed, new locations can be sampled from these portions, and they can be successively labeled with further experiments or reliable simulations to produce additional data points. These can be used for retraining the surrogate model and improving its accuracy. Finally, Data Science techniques are here employed for identifying confounding variables and essential patterns in large data sets (*e.g.*, the database

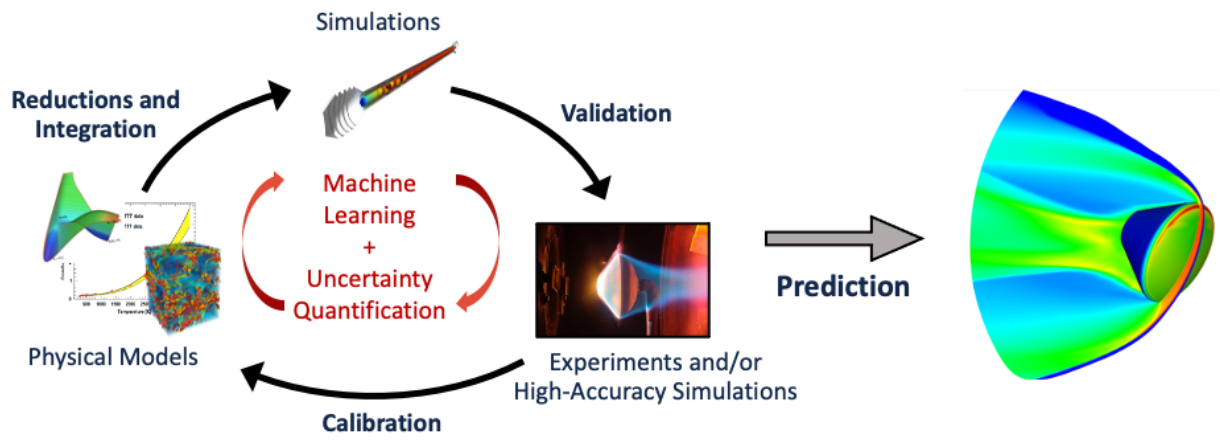


Figure 1.3: Schematics of the proposed approach to predictive Hypersonics. (Sources of the subfigures: NASA Ames and NEQRAD research group)

of *ab initio* rate coefficients) with the aim of gaining insights on the important physics. In turn, this allows improving the model accuracy without increasing its computational cost or reducing the model complexity without affecting its reliability.

The thesis is divided into three parts. The first part follows a systematic approach to the uncertainty quantification of ML-based PESs, as it relies on the idea by Neal [109] of augmenting a generic NN to a Bayesian neural network (BNN). The application of BNNs to the construction of PESs was proposed by Venturi *et al.* [159], and it generates surfaces that are stochastic by nature (*i.e.*, a stochastic PES, abbreviated as SPES). As for the GPs, the non-deterministic character is the expression of the *ab initio* and fitting uncertainties affecting the surface. BNNs provide most of the advantages of GPs to the PES construction, such as generating uncertainty bounds and lacking overfitting. Compared to the latter, however, they are significantly more efficient when employed in molecular simulations. It is worth mentioning that GPs can be interpreted as BNNs with an infinite number of hidden units and normally distributed parameters [74]. The use of a Bayesian framework [100, 101, 127, 152] allows one to gain physical insight on the PES by studying the propagation of uncertainty from micro-scale to the macro-scale QoIs. The availability of such information content has two main advantages. Firstly, it enables an assessment of the PES predictive capabilities. Secondly, it permits one to identify the surface features that mostly contribute to the uncertainties on state-to-state rate coefficients and their derived variables, such

as QSS rates, relaxation time-constants, and energy transfer coefficients. As a consequence, the surface can be refined by computing additional *ab initio* energies in its most critical regions. The novel approach is here tested for O_2+O interactions but can be easily applied to chemical systems with more than three atoms and used for validating ML-based PESs in a multitude of fields.

The second part of the manuscript employs Data Science techniques for discovering patterns in the dissociation mechanisms of diatomic molecules. These insights are then used for complementing the adaptive clustering model proposed by Sahai *et al.* [129] The tools adopted in this investigation include rovibrational QCT calculations [23, 135] and the solution of the master equation (ME) [68, 69, 114]. The combined use of these techniques allowed for the derivation of a novel CG strategy relying on grouping the rovibrational levels based on their energy-deficit from the centrifugal barrier. The role of this quantity has been widely explored in the past [7, 25, 27, 55, 87, 92, 97, 122, 162]. Only recently, however, it has been suggested to use such energy-deficit as the cornerstone of a grouping strategy [159]. The new model is validated against the ME's direct solution and further bench-marked against the conventional VS model, thus exposing the invalidity of the assumptions underlying this last approach.

Finally, Part III summarizes key conclusions and presents future work.

The thesis is organized as follows:

- Chapter 2 outlines the steps to follow for generating the SPES, forward propagating its uncertainty, and refining its important regions. It starts by presenting the Permutation Invariant Polynomials Neural Network as an example of ML-based techniques for the PES construction, and it then describes the stochastic extension for these deterministic approaches. In the following section, the chapter introduces the O_2+O chemical system on which the novel framework will be tested and details the calculations performed for predicting the gas QoIs, which start from the first principles of quantum chemistry.

- The study of energy transfer and dissociation processes for a wide range of conditions for the O_2-O system in an ideal chemical reactor is studied in Chapter 3. The first part discusses the results of the SPES calibration process by providing comparisons with the *ab initio* data points and the original Varga PESs [150]. The remainder of the chapter addresses the forward propagation of the uncertainty to predefined quantities of interest. To this end, extensive

QCT calculations are performed to generate the reaction rate parameters to be used in the 0D chemical reactor solver. Finally, the sensitivities of the QoIs to the accuracy of the PESs are analyzed.

- Chapter 4 details the quasi-classical trajectory calculations, the resulting kinetic databases, and a hybrid methodology applied to the study of dissociation. A brief discussion of the O_2 diatomic potential is followed by a description of the methods used to investigate the non-equilibrium relaxation in a 0-D chemical reactor. The necessary equations, already extensively discussed in literature [114], are briefly outlined. Reduced-order techniques are discussed next, including both the conventional vibrational-specific approach and the proposed centrifugal-barrier-based strategy. The section is concluded by presenting a novel hybrid methodology, which combines a rovibrational-specific treatment of excitation processes and a physics-based model reduction of the dissociation process.

- The hybrid methodology introduced above is used in Chapter 5 to study the performance of the reduced-order models by comparing the data sets of rate coefficients (*ab initio* and reconstructed) and the results of 0-D thermochemical simulations. The observable QoIs for the latter analysis are the composition profiles and the rovibrational distributions at different times. Finally, for validation purposes, the macroscopic quasi-steady-state (QSS) [81] rate coefficients are compared to those in the literature.

- Chapter 6.1 summarizes the key conclusions in Sec.s 6.1.2 - 6.1.1 and presents future work in Sec. 6.2.

1.4 Related Scientific Contributions

The theoretical findings and the results discussed in this thesis have been published in peer-reviewed journals and presented at international conferences. Additional articles are under preparation.

- Journal publications:

  1. **S. Venturi**, M. P. Sharma, B. Lopez, and M. Panesi, *Data-Inspired and Physics-Driven Model Reduction for Dissociation: Application to the O_2+O System*, J. Phys. Chem. A

2020, 124, 41, 8359–8372 [153]

2. **S. Venturi**, R. L. Jaffe, and M. Panesi, *Bayesian Machine Learning Approach to the Quantification of Uncertainties on Ab Initio Potential Energy Surfaces*, J. Phys. Chem. A 2020, 124, 25, 5129–5146 [158]

3. R. L. Jaffe, M. Grover, **S. Venturi**, D. W. Schwenke, P. Valentini, T. E. Schwartzentruber, and M. Panesi, *Comparison of Potential Energy Surface and Computed Rate Coefficients for N_2 Dissociation*, J. of Thermophysics and Heat Transfer, 32, 4, 869-881 [56]

4. **S. Venturi**, A. Sahai, and M. Panesi, *Diatomic Potential-Based Strategy for Clustering Rovibrational Levels: Application to Earth and Mars Atmospheres*, In Preparation

5. **S. Venturi**, A. Munafò, and M. Panesi, *Dynamics of $CO(^1\Sigma^+)+O(^3P)$ Collisions: Rovibrational QCT Calculations and their Validation*, In Preparation

- Journal covers:

1. J. Phys. Chem. A 2020, Volume 124, Issue 44, Machine Learning in Physical Chemistry Virtual Special Issue (Figure 1.4(a))

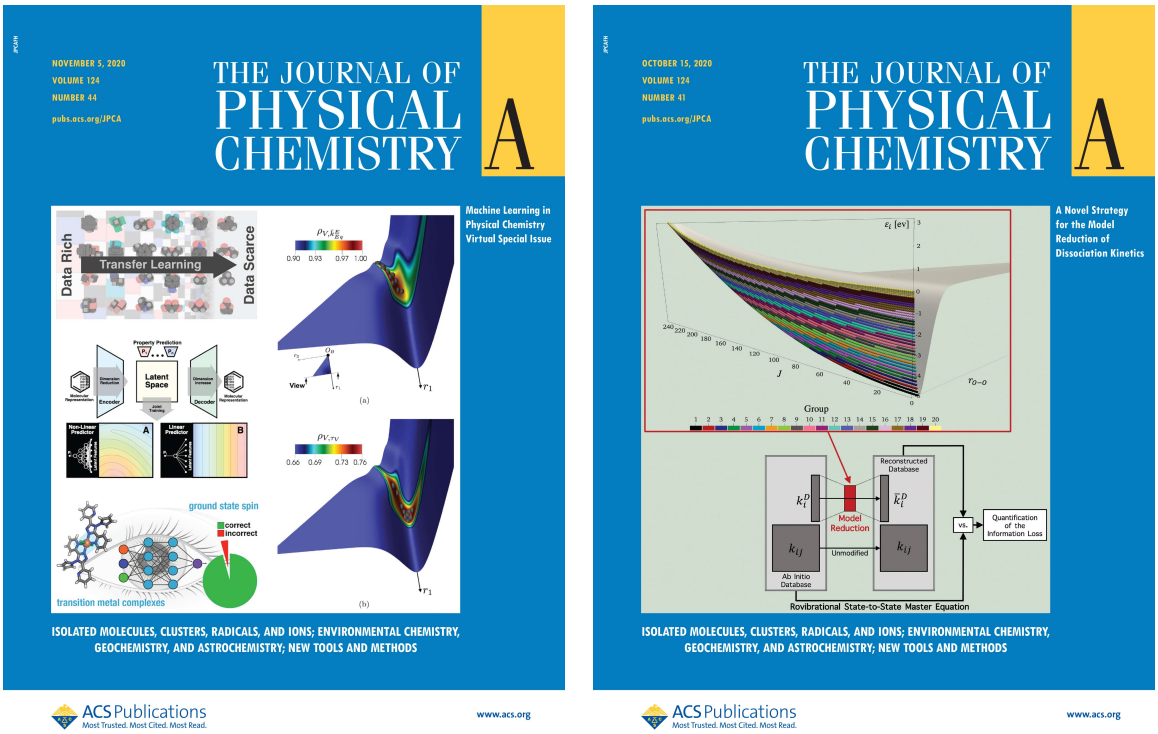
2. J. Phys. Chem. A 2020, Volume 124, Issue 41 (Figure 1.4(b))

- Conference proceedings:

1. A. Munafò, **S. Venturi**, M. P. Sharma, and M. Panesi, *Reduced-Order Modeling for Nonequilibrium Air Flows*, AIAA Scitech 2020 Forum, AIAA 2020-1226 [106]

2. **S. Venturi**, M. P. Sharma, A. Racca, and M. Panesi, *Effects of Ab Initio Potential Energy Surfaces on O_2-O Nonequilibrium Kinetics*, AIAA Aviation 2019 Forum, AIAA 2019-3358 [160]

3. **S. Venturi**, A. Munafò, and M. Panesi, *Modeling the Contribution of Rotation to Ro-Vibrational Nonequilibrium Kinetics*, 8-th International Workshop on Radiation of High Temperature Gases for Space Missions, 18-20 March 2019, Madrid, Spain [154]



(a) (b)

Figure 1.4: Covers of the Journal of Physical Chemistry A, Volume 124, Issue 44, Machine Learning in Physical Chemistry Virtual Special Issue (a), and Volume 124, Issue 41 (b).

4. **S. Venturi**, A. Racca, and M. Panesi, *Quantifying the Uncertainty on Ab Initio Rate Coefficients by Means of Bayesian Machine Learning*, Uncertainty Quantification and Optimization (UQOP) Conference, 25-29 March 2019, Paris, France [156]

5. **S. Venturi**, M. P. Sharma, and M. Panesi, *A Machine Learning Framework for the Quantification of the Uncertainties Associated with Ab Initio Based Modeling of Nonequilibrium Flows*, AIAA Scitech 2019 Forum, AIAA 2019-0788 [159]

6. M. P. Sharma, **S. Venturi**, and M. Panesi, *Application of Ab Initio Based Grouped Rates for Modeling Nonequilibrium Flow Physics*, AIAA Scitech 2019 Forum, AIAA 2019-0792 [123]

7. **S. Venturi**, and M. Panesi, *Investigating CO Dissociation by Means of Coarse Grained Ab Initio Rate Constants*, 2018 AIAA Aerospace Sciences Meeting, AIAA 2018-1232 [155]

8. A. Munafò, **S. Venturi**, R. L. Macdonald, and M. Panesi, *State-to-State and Reduced-*

Order Models for Recombination and Energy Transfer in Aerothermal Environments, 54th AIAA Aerospace Sciences Meeting, AIAA 2016-0505 [105]

9. R. L. Jaffe, D. W. Schwenke, M. Grover, P. Valentini, T. E. Schwartzentruber, **S. Venturi**, and M. Panesi, *Comparison of Quantum Mechanical and Empirical Potential Energy Surfaces and Computed Rate Coefficients for N_2 Dissociation*, 54th AIAA Aerospace Sciences Meeting, AIAA 2016-0503 [59]

- Computer programs and toolboxes:

1. **S. Venturi** and B. Lopez, *CoarseAIR: Fortran 2008 Code for StS and Coarse-Grained Quasi-Classical Trajectory Method*, https://github.com/simoneventuri/CoarseAIR [153]
2. **S. Venturi**, *MCoarseAIR: MATLAB Code for Postprocessing CoarseAIR and Heat Bath Results*, https://github.com/simoneventuri/MCoarseAIR [158]
3. **S. Venturi**, *Python Code for Postprocessing the Ab Initio Databases of Rate Coefficients and CoarseAIR Results*, https://github.com/simoneventuri/PyCoarseAIR [158]
4. **S. Venturi**, *Spebus: Python+MATLAB Toolbox for Machine Learning-Based Potential Energy Surfaces (PESs)*, https://github.com/simoneventuri/Spebus [153]

- Databases:

1. **S. Venturi**, M. P. Sharma, B. Lopez, A. Munafò, and M. Panesi, *The CHESS Database of Ab Initio Rate Coefficients for Hypersonic Applications*, April 2020 [157]

Part I

Bayesian Machine Learning Approach to the Quantification of Uncertainties on *Ab Initio* Potential Energy Surfaces (PESs)

Chapter 2

The Stochastic PES: Construction, Propagation, and Refinement

A broad objective of predictive science is to develop models for evaluating the response of a physical system to changes in the external conditions and for quantifying the uncertainties connected to such estimates, given the availability of a limited set of reliable data [111, 112]. Following this idea, a PES employed in predictive simulations should not only fit the *ab initio* data points but should also inform about the induced uncertainties. To fulfill this gap, the proposed methodology provides a stochastic representation of the potential energy surface by augmenting a generic preexisting ML-based surrogate model (*e.g.*, neural network, permutation invariant polynomials, etc.) in two ways: by treating its parameters, $\boldsymbol{\alpha}$, as random variables, characterized by their probability density functions (PDFs); by introducing a random noise, ϵ_V, as a function of a set of hyperparameters, $\boldsymbol{\gamma}$, to account for the inadequacies due to the *ab initio* calculations and the fitting model form [142, 143, 152]. In this way, the PES is formulated as a probabilistic model (Figure 2.1), and the potential energy V can be finally predicted for a generic set of atomic distances, $\boldsymbol{r}$, as:

$$V = \hat{V}(\boldsymbol{r}, \boldsymbol{\alpha})\epsilon_V(\boldsymbol{\gamma}), \tag{2.1}$$

where $\hat{V}$ represents the output of the original ML fit. The formulation of the problem in Eq. 2.1 is reminiscent of the "potential morphing approach" introduced by Meuwly and Hutson [99] and earlier by Bowman and Gazdy [16]. However, there is a fundamental difference in the modeling purposes, manifested through the stochastic character of $\epsilon_V(\boldsymbol{\gamma})$: while the "morphing" correction (*e.g.*, the multiplicative term in Eq. 3 from Meuwly and Hutson [99]) has the objective of improving the PES (*i.e.*, "to give an optimal fit to the reference data" [99]), $\epsilon_V(\boldsymbol{\gamma})$ is here used to describe the model inadequacy (*i.e.*, to represent the unknown unknowns).

As a result of both parametric and modeling uncertainty, instead of providing a single model out-

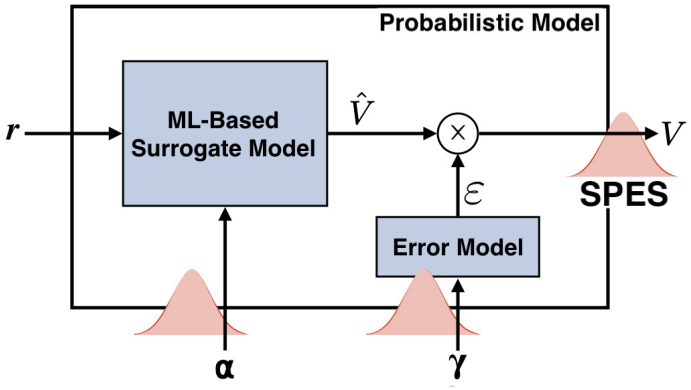


Figure 2.1: Schematic of the probabilistic model generating the Stochastic Potential Energy Surface (SPES).

put to an input, the probabilistic model generates a distribution of potential energies, hereafter referred to as Stochastic Potential Energy Surfaces (SPESs). The proposed UQ framework relies on the SPES in three distinct stages, as indicated in Figure 2.2: Bayesian calibration, forward propagation, and surface refinement.

Step one is the **Bayesian calibration**, in which the SPES is constructed via Bayesian inference [143] as a solution of an *inverse problem*. The process begins with the selection of *prior distributions* for the overall set of parameters and hyperparameters, $\boldsymbol{\theta} = \boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{\gamma}$, where $\boldsymbol{\beta}$ represents a vector of hyperparameters used for constructing $\boldsymbol{\alpha}$'s PDFs. The priors represent the pre-calibration knowledge available for the values of the coefficients. By using *ab initio* data and invoking Bayes' theorem, the prior PDFs are updated to the *posterior distributions*, as sketched in Figure 2.2(a). This results in an optimal stochastic surface (*i.e.*, a family of potential energy surfaces) shaped around the data points: the SPES.

Step two is the **forward propagation** of the SPES to the QoIs. The calibrated stochastic PES is sampled, and the resulting ensemble of deterministic surfaces is forward propagated to a set of QoIs (Figure 2.2(b)). In order to do so, additional physical models have to be included in the methodology, depending on the choice of the quantities of interest. The applications analyzed in the present part follow the work of Panesi *et al.* [114, 116], where the QoIs include a number of physical variables resulting from the analysis of kinetic processes in nonequilibrium gas mixtures,

such as phenomenological rate parameters, compositions, characteristic times, and more. For this reason, the forward propagation step requires the solution of the scattering problem, to obtain a complete set of state-specific rate coefficients, followed by the solution of the master equations, to describe the state-specific kinetics of the gas in a chemical reactor.

Step three is the surface **refinement**, which relies on Bayesian-informed sensitivity analysis. Global sensitivity analysis techniques [54] are employed to quantify the influence that the SPES' features have on the QoIs' variance (Figure 2.2(c)). In this way, the surface can be refined by introducing additional *ab initio* calculations, so that the uncertainties on the predicted QoIs are reduced.

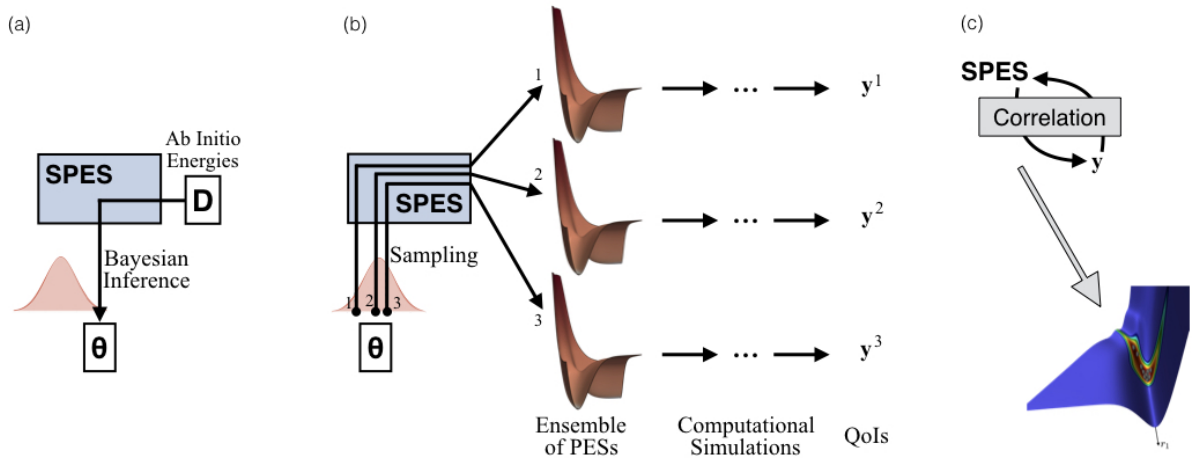


Figure 2.2: The steps of the SPES approach. Black arrows represent flows of information content (a): Bayesian calibration, (i.e, solution of the inverse problem). The optimal PDFs characterizing the overall set of parameters and the hyperparameters of the stochastic potential energy surface (SPES) are learnt through Bayesian inference from the *ab initio* data points. (b): Forward propagation of the uncertainties. The posterior distribution is sampled, an ensemble of PESs is obtained, and the surfaces are employed in computational simulations with the aim of predicting QoIs. (c): Surface refinement based on sensitivity analysis. The QoIs can be correlated to the variability of the SPES at a fine grid of atomic configurations, in order to highlight the surface areas where uncertainties are crucial.

The remainder of this chapter provides a more detailed discussion of individual components of the proposed methodology. In particular, Sec. 2.1 presents the Permutation Invariant Polynomials Neural Network as an example of the state-of-the-art algorithms for constructing ML-based PESs. The generalized Bayesian framework for extending the deterministic PES to stochastic objects is described in Sec. 2.2. Finally, the physical models including potential energy surfaces, scattering

calculations, and the master equation solution of the 0D chemical reactor are discussed in Sec. 2.3.

2.1 Permutation Invariant Polynomials Neural Networks for PESs

Between all the machine learning techniques adopted for fitting PESs, artificial neural networks are the most widely used. What really differentiates the multiple approaches to NN-PESs is the way in which they enforce the possible symmetries between the particle configurations [17]. In fact, the PES must guarantee the prediction invariance to translation, to rotation, and to permutation of identical atoms in the system (*e.g.*, for O_2+O, $\hat{V}(r_1, r_2, r_3)$ must be identical to $\hat{V}(r_2, r_1, r_3)$). The easiest solution for dealing with the first two of such constraints is to describe the atomic interactions through relative distances, instead of Cartesian coordinates. The third requirement is more difficult to fulfill, and the most common strategies for addressing this problem rely on feature engineering: instead of directly providing the atom distances as NN input, these are first symmetrized through an additional ad hoc layer. For high dimensional PESs, this transformation can depend on cutoff functions and angular terms [11], and can become particularly complicated. For systems composed of 3 or 4 atoms, the symmetry functions can be relatively simple: as suggested by Jiang and Guo [62], the PES can be constructed through a hybrid approach, which is generated by connecting in series permutation invariant polynomials and the neural network. As a first step of this PIP-NN technique, the atomic relative distances, $\mathbf{r}$, are transformed to intermediate variables, $\mathbf{P} = \phi(\mathbf{r})$, generally through exponential functions, Morse functions, Gaussians, or combinations of them. A symmetrized polynomial vector, $\mathbf{G}$, is then computed as:

$$G_i = \hat{S}_i \Big[P_1^{l_1} P_2^{l_2} P_3^{l_3} \Big] \tag{2.2}$$

where l_j are exponents, and $\hat{\mathbf{S}}$ is the symmetrization operator. In the present work, the Morse function is used for transforming r_i in P_i:

$$P_i = \exp(-\lambda(r_i - r_e)), \qquad i = 1, 3 \tag{2.3}$$

where λ and r_e are tunable parameters; the symmetrized polynomial vector is constructed from the intermediate variables as:

$$\begin{cases} G_1 = P_1P_2 + P_2P_3 + P_1P_3, \\ G_2 = P_1P_2P_3, \\ G_3 = P_1^2P_2 + P_1P_2^2 + P_3^2P_2 + P_3P_2^2 + P_1^2P_3 + P_1P_3^2, \\ G_4 = P_1^3P_2 + P_1P_2^3 + P_3^3P_2 + P_3P_2^3 + P_1^3P_3 + P_1P_3^3, \\ G_5 = P_1^2P_2P_3 + P_1P_2^2P_3 + P_1P_2P_3^2, \\ G_6 = P_1^2P_2^2 + P_3^2P_2^2 + P_1^2P_3^2. \end{cases} \tag{2.4}$$

Compared to the **G** vector suggested by Jiang and Guo, the component $P_1 + P_2 + P_3$ has been removed; the many-body contribution to the PES, indeed, has to be exactly zero when one of the three atoms is at infinite distance from the others, and this can be guaranteed only when no unconnected terms [120] are given as input to the NN. Furthermore, the new **G** has 6 dimensions, compared to the 3 of the original paper, and a polynomial degree of 4, compared to 3. The author tried various constructions of the symmetrized polynomial vector and found that the increase in **G** complexity proposed here significantly improves the accuracy of the PIP-NN.

As a subsequent step of the approach, **G** is fed to the neural network, and it flows through its layers as a series of weighted linear combinations alternated to non-linear functions. Indeed, the output of the k-th neuron in the i-th hidden layer, y_i^k, can be computed as:

$$\begin{cases} z_i^k = \sum_{j=1}^{N_{i-1}} W_i^{j,k} q_{i-1}^j + b_i^k, \\ y_i^k = \mathcal{F}_i\Big(z_i^k\Big), \qquad\qquad i = 1, L \quad \text{and} \quad k = 1, N_i \end{cases} \tag{2.5}$$

where L quantifies the number of hidden layers, and N_i the number of neurons present in the i-th layer. Corresponding to this i-th level, there are 2 arrays of parameters: a matrix of weights, $\mathbf{W_i}$, and a vector of biases, $\mathbf{b_i}$. Moreover, $q_i^j = G_j$ when $i = 0$, and $q_i^j = y_i^j$ otherwise. Hyperbolic tangent and sigmoid function have been tested by the author as neuron activations, $\mathcal{F}_i(...)$, with $tanh(...)$ showing to perform better in terms of training and accuracy.

$\hat{V}$ can be finally computed as the output of the network:

$$\hat{V} = \sum_{j=1}^{N_L} W_{L+1}^{j} y_L^j + b_{L+1}, \tag{2.6}$$

and a schematic of the overall PIP-NN architecture is reported in Figure 2.3.

Venturi *et al.* [160] tested a PIP-NN for the computation of O_2+O RVS StS rate coefficients

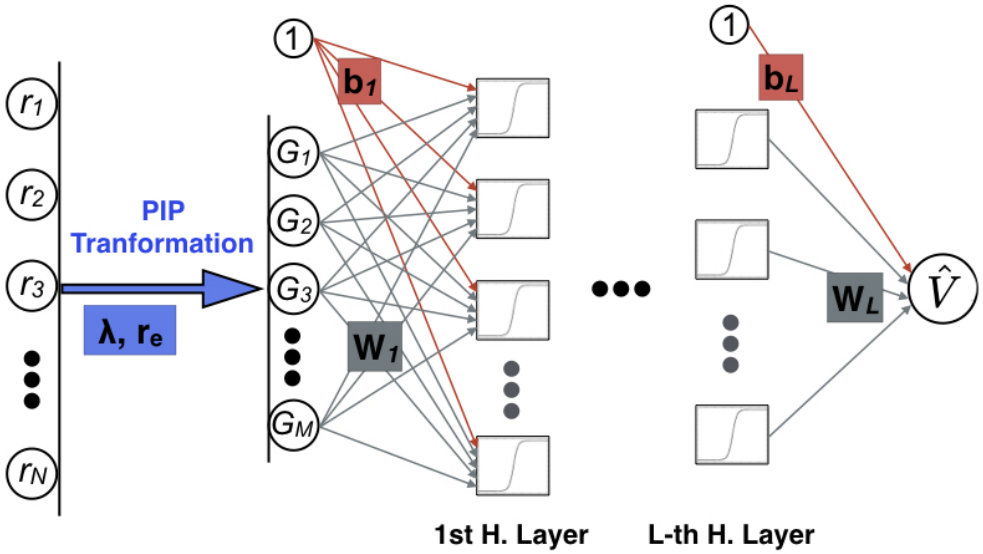


Figure 2.3: Structure of a L-Hidden Layers PIP-NN; the bold letters in the colored boxes correspond to the arrays of parameters characterizing the network.

based on the $2^5A'$ PES, and for the consecutive heat bath simulation at constant translational temperature of 10 000 K. Two hidden layers with respectively 20 and 10 neurons have been used for constructing the PES, with a total of 2 calibrated parameters for the PIP part, and 361 for the remaining NN. The resulting surface presented a significant reduction of the *ab initio* points fitting error compared to the original PIP-Fit by Varga *et al.*; the RMSE has been decreased over all the energy range, especially by more than 50% for the surface regions between 0 and 4 eV. This improvement in the PES accuracy also translated to non-negligible differences between the gas QoIs (particularly mole fractions and vibrational excitation timescales) predicted through the original PES from Varga *et al.* and through the ML-constructed surface. In the present work, the PIP-NN structure is implemented through the Python libraries Theano [144] and Lasagne [34].

2.2 The Steps of the SPES Approach

2.2.1 Step 1: Bayesian Calibration

In the context of the Bayesian framework, the learning process relies on updating the prior distributions, $p(\boldsymbol{\theta}|\mathcal{M})$, via the Likelihood function, $p(\mathbf{D}|\boldsymbol{\theta},\mathcal{M})$, constructed using the *ab initio* data points, $\mathbf{D}$. This task relies on Bayes' theorem [141, 143]:

$$p(\boldsymbol{\theta}|\mathbf{D},\mathcal{M}) \propto p(\mathbf{D}|\boldsymbol{\theta},\mathcal{M})p(\boldsymbol{\theta}|\mathcal{M}), \tag{2.7}$$

where $p(\boldsymbol{\theta}|\mathbf{D},\mathcal{M})$ refers to the posterior distribution and $\mathcal{M}$ indicates the model form. The permutation invariant polynomials neural network (PIP-NN) [62, 63] is selected to generate an analytical representation of the PES, due to its computational efficiency. The modifications made to the original approach have been described in the previous section. Starting from such a fitting technique, the SPES is constructed by augmenting the NN to Bayesian neural networks [109]. In the remaining part of this manuscript, this extension will be called permutation invariant polynomials Bayesian neural network (PIP-BNN), and its functional form will be represented by:

$$\mathcal{M} = \Big\{\phi(...), \hat{O}(...), L, \big(N_i, \mathcal{F}_i(...)\big)_{i=1:L}\Big\}, \tag{2.8}$$

where $\phi(...)$ identifies the function adopted for converting the relative distances between the atoms to some intermediate variables, $\hat{O}(...)$ expresses the symmetrization operator selected for transforming the intermediate variables in the symmetrized polynomial vector, L quantifies the number of NN hidden layers, and, finally, N_i and $\mathcal{F}_i(...)$ respectively represent the number of neurons and non-linear functions chosen for the i-th layer, as they have been introduced in Sec. 2.1. The overall set of parameters and hyperparameters will be identified by:

$$\boldsymbol{\theta} = \boldsymbol{\alpha} \cup \boldsymbol{\beta} \cup \boldsymbol{\gamma} = \Big\{\lambda, r_e, \big(\mathbf{W_i}, \mathbf{b_i}\big)_{i=1:L}\Big\} \cup \Big\{\big(\sigma_{W_i}\big)_{i=1:L}, \big(\sigma_{b_i}\big)_{i=1:L-1}\Big\} \cup \Big\{\sigma_{Like}\Big\}, \tag{2.9}$$

where λ and r_e characterize the transformation to intermediate variables, $\mathbf{W_i}$ and $\mathbf{b_i}$ respectively indicate weights and biases of the NN, σ_{W_i} and σ_{b_i} represent the standard deviations of the pa-

rameter prior distributions, and finally, σ_{Like} expresses the standard deviations of the likelihood. It should be noticed that the parameter b_3 corresponds to the reference energy of the surface, V^0. The prior distribution is constructed assuming that the $\boldsymbol{\alpha}$ parameters are random variables independent from each other, and that the same holds for the hyperparameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. We can then write:

$$p(\boldsymbol{\theta}|\mathcal{M}) = \prod_{i}^{N_\alpha} \prod_{j}^{N_\beta} p(\alpha_i|\beta_j, \mathcal{M}) p(\beta_j|\mathcal{M}) \prod_{k}^{N_\gamma} p(\gamma_k|\mathcal{M}), \tag{2.10}$$

where N_i represents the size of the correspondent (hyper-)parameter vector, $\boldsymbol{i}$, with $\boldsymbol{i} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$, and the $p(\theta_i|\mathcal{M})$ are reported in Table 2.1, with $\mathcal{U}(a, b)$, $\mathcal{N}(\mu, \sigma^2)$, and Half-$\mathcal{N}(\mu, \sigma^2)$ identifying uniform, normal and half-normal distributions, respectively. For the parameters of the symmetry function (λ and r_e) and for the PES reference energy ($b_3 = V^0$), uninformative priors are selected and only maximum and minimum values are imposed for such quantities. The prior distributions of the NN weights and biases are represented through zero-mean Gaussians, the standard deviations of which (σ_{W_i} and σ_{b_i}) are treated as new hyper-parameters. Such choices rely on the fact that even weakly informative priors can robustly contribute to regularization [14] by promoting models with as little complexity as possible.

The multiplicative stochastic model in Eq. 2.1 has been chosen for two reasons. Firstly, a large fraction of the particle trajectories visit only the low energy regions of the potential surface; as a consequence, these portions should be most accurately reproduced. Secondly, the quantities that drive the collision dynamics are the interaction forces, computed as the PES gradients with respect to the relative atom-atom distances. The derivative operator filters out any additive quantities that are not functions of the spatial configuration, preventing them from propagating any further. The error in the PES prediction at the i-th data location is formulated as $\epsilon_{V_i} = \hat{V}_i / D_i$ or, in its logarithmic form, as:

$$\epsilon_{V_i}^{log} = \log(\epsilon_{V_i}) = \log(\hat{V}_i) - \log(D_i); \tag{2.11}$$

$\hat{V}_i$ represents the sum of the reference value ($\hat{V}^0$) and the many-body contribution ($\hat{V}_i^{MB}$). In this work, in fact, the overall PES is obtained based on the expansion:

$$V(r_1, r_2, r_3) = \sum_{i=1}^{3} V^{Diat}(r_i) + V^0 + V^{MB}(r_1, r_2, r_3), \tag{2.12}$$

where r_i indicates the interatomic distance and the diatomic potential (V^{Diat}) is not varied from its original formulation. The reasoning behind these choices will be detailed in Sec. 2.3.1. The stochastic error is described through the likelihood function as a zero-mean Gaussian:

$$p(\mathbf{D}|\boldsymbol{\theta}, \mathcal{M}) = \frac{1}{\sqrt{(2\pi\sigma_{Like}^2)^{N_D}}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N_D}\left(\frac{\epsilon_{V_i}^{log}}{\sigma_{Like}}\right)^2\right], \tag{2.13}$$

where N_D represents the number of data points. The assumption underlying the characterization of the likelihood function through a scalar hyper-parameter, σ_{Like}, is that the prediction errors affecting the different points of the surface are fully correlated to each other.

The number of parameters involved in the construction of the PIP-BNN is relatively small when compared to other NN-derived algorithms. Even for such relatively low dimensionality, however, the computation of the posterior distribution through sampling methods becomes impracticable.

Table 2.1: Prior Distributions of Parameters and Hyper-parameters:

α_i	$p(\alpha_i\|\beta_j, \mathcal{M})$
λ $[1/a_0]$	$\sim \mathcal{U}(0,1)$
r_e $[a_0]$	$\sim \mathcal{U}(1,4)$
$\left(\mathbf{W}_i\right)_{i=1:3}$	$\sim \mathcal{N}(0, \sigma_{W_j}^2)$, with $j = i$
$\left(\mathbf{b}_i\right)_{i=1:2}$	$\sim \mathcal{N}(0, \sigma_{b_j}^2)$, with $j = i$
b_3	$\sim \mathcal{U}(-5,5)$

β_j	$p(\beta_j\|\mathcal{M})$
$\left(\sigma_{W_j}\right)_{j=1:3}$	$\sim$ Half-$\mathcal{N}(0, 5^2)$
$\left(\sigma_{b_j}\right)_{j=1:2}$	$\sim$ Half-$\mathcal{N}(0, 5^2)$

γ	$p(\gamma\|\mathcal{M})$
σ_{Like}	$\sim$ Half-$\mathcal{N}(0, 1^2)$

For this reason, we approximate it by means of variational Bayesian learning [10], through the automatic differentiation variational inference method (ADVI) by Kucukelbir *et al.* [75]. The procedures for the SPES calibration are performed by means of the in-house code Spebus, which uses the Theano [144] and PyMC3 [131] libraries.

2.2.2 Step 2: Forward Propagation

The output of the calibration process consists of a continuous distribution of potential energy surfaces, representative of the uncertainty associated with the PES. To assess the impact of the uncertainties on the QoIs, the SPES must be propagated through the physical models, which are generally deterministic in nature. Through the analysis pursued in the present part, we aim to verify two qualities of the PES: how well does the PES reproduce the *ab initio* data points, and how uncertain are the predictions that it generates. For the former objective, a suitable quantity to analyze derives from the maximum a posteriori probability estimate (MAP), which is the mode of the posterior distribution:

$$\boldsymbol{\theta}_{MAP} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, p(\boldsymbol{\theta}|\mathbf{D}, \mathcal{M}). \tag{2.14}$$

Given a generic atomic configuration $\boldsymbol{r}$ and following the schematics in Figure 2.1, the point in the parameter space $\boldsymbol{\alpha}_{MAP}$ can be used for predicting $\hat{V}_{MAP}$, which represents the many-body component of the potential energy shifted by the reference value. At this point, Eq. 2.1 should be applied for taking into account the structural inadequacy. Due to the fact that we imposed a multiplicative error and a normal-distributed likelihood (Eq. 2.13), the noise should be computed from [112]:

$$\epsilon_V = \exp\left(\epsilon_V^{\log}\right) = \exp\Big(\ \mathcal{N}(0, \sigma_{Like})\ \Big). \tag{2.15}$$

However, in order to obtain a single MAP sample from the SPES, we select the most probable value for the noise, which corresponds to $\epsilon_V = 1$, and we finally add the diatomic contribution:

$$V_{MAP} = \sum_{i=1}^{3} V^{Diat}(r_i) + \mathcal{P}(\boldsymbol{r}, \boldsymbol{\alpha} = \boldsymbol{\alpha}_{MAP}, \epsilon_V = 1), \tag{2.16}$$

where $\mathcal{P}$ represents the PIP-BNN operator. Given the set of *ab initio* data points and the prior distributions, the V_{MAP} is the stochastic surface sample that corresponds to the highest value of

parameter posterior probability.

On the other hand, in order to assess the variability of the SPES at each atomic configuration, it is possible to generate a finite and discrete representation of the posterior distribution by taking advantage of Latin hyper-cube sampling (LHS) [98]. In this way, the highly dimensional parameter space can be efficiently explored, and a small set of samples can be produced ($\boldsymbol{\theta}^j_{LHS}$ for $j = 1, N_{LHS}$). When this collection is given to the neural network, it generates an ensemble of surfaces, which needs to be multiplied by the random noise. In order to maintain the deterministic character gained through the LHS, $\epsilon_V^{\log}$ is randomly sampled from the normal distributions with standard deviations $\sigma^j_{Like_{LHS}}$. When the diatomic contribution is added, we obtain:

$$V^j_{LHS} = \sum_{i=1}^{3} V^{Diat}(r_i) + \mathcal{P}(\boldsymbol{r}, \boldsymbol{\alpha} = \boldsymbol{\alpha}^j_{LHS}, \epsilon_V = \epsilon^j_{V_{LHS}}), \quad j = 1, N_{LHS}. \tag{2.17}$$

Through this procedure, the MAP and LHS surfaces can be forward propagated through the physical model (see Sec. 2.3) without requiring any modification to its computational implementation. In this way, the effect of the PES unreliability on the predictive capabilities of microscopic and macroscopic QoIs, y^j, can be assessed. The accuracy required on the prediction (*i.e.*, the extent of the uncertainty bound required on y^j) is highly application-dependent, and it is also dictated by the impact that the QoI has on the overall design/decision-making process.

2.2.3 Step 3: Surface Refinement

One of the advantages of producing an ensemble of PESs and independently forward propagating multiple samples is that a one-to-one relationship can be constructed between each sampled surface, V^j_{LHS}, and the value of the gas QoI, y^j, computed from it. In this way, the Pearson coefficient [130] correlating these two variables can be computed at each atomic configuration (r_1, r_2, r_3) as:

$$\rho_{V,y}(r_1, r_2, r_3) = \frac{\sum_{j=1}^{N_{LHS}} (V^j_{LHS} - \bar{V}_{LHS})(y^j - \bar{y})}{\sqrt{\sum_{j=1}^{N_{LHS}} (V^j_{LHS} - \bar{V}_{LHS})^2 \sum_{j=1}^{N_{LHS}} (y^j - \bar{y})^2}}, \tag{2.18}$$

where V^j_{LHS} and $\bar{V}_{LHS}$ are both functions of (r_1, r_2, r_3), and the over-barred letters correspond to the mean values. Eq. 2.18 assumes linearity between the V and y variables. A study of the

$\rho_{V,y}$ magnitudes on a grid of (r_1, r_2, r_3) configurations allows the identification of the PES regions that have the largest effect on the gas QoI. Further steps, when required, are the selection of a set of novel configurations from these high sensitivity portions, and the *ab initio* calculations of the potential energies at these new coordinates. The resulting points can then be added to the training and validation data in order to reduce the variability and/or increase the accuracy of the PES in such problematic regions.

It is important to notice that the PES refinement as proposed here is physics-based and highly QoI specific. Through this approach, indeed, the target variable y for which the improving reliability of the prediction is desirable does not have to represent a gas quantity, as in the cases analyzed in the present work. Depending on the application, y can be located further upstream or downstream in the sequence of computational simulations.

2.3 Physical Modeling

This section outlines the sequence of physical models that allow the prediction of the gas QoIs starting from first-principles quantum mechanics:

- ***Ab Initio* Construction of the PESs:** As mentioned in the introduction, the foundation of the *ab initio* methodology is the accurate modeling of the quantum forces acting between the atoms; the resulting potential energies are here represented through the SPES, which is tested for the O_2+O chemical system introduced in Subsec. 2.3.1.

- **Rovibrational-Specific QCT Calculations:** The PES gradients are then employed as forcing terms in the Hamiltonian equations. Based on the QCT method described in Subsec. 2.3.2, trajectories of many individual atom-molecule collisions are simulated, and the RVS cross-sections and rate coefficients are computed.

- **Rovibrational-Specific State-to-State Master Equation:** Through the ME, the StS rate coefficients are used for investigating the kinetics of nonequilibrium processes taking place in isochoric heat baths (Subsec. 2.3.3).

- **Reduced-Order Modeling:** A fully RVS StS concatenation of QCT and ME becomes computationally impractical when it comes to forward propagating an entire ensemble of

PESs. A reduced order model sucha as the one described in Subsec. 2.3.4 becomes crucial as an alternative to the rovibrational quantum resolution.

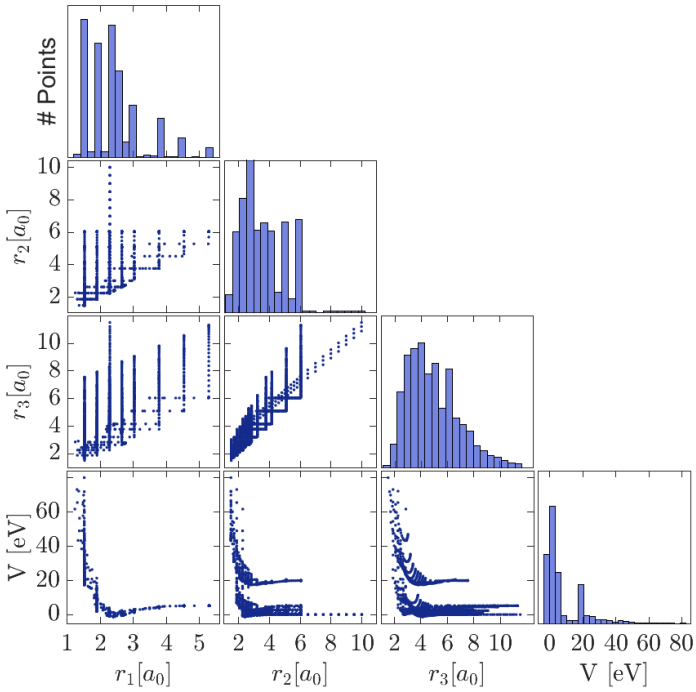
The following parts of the section will individually describe the aforementioned models.

2.3.1 *Ab Initio* Construction of the PESs

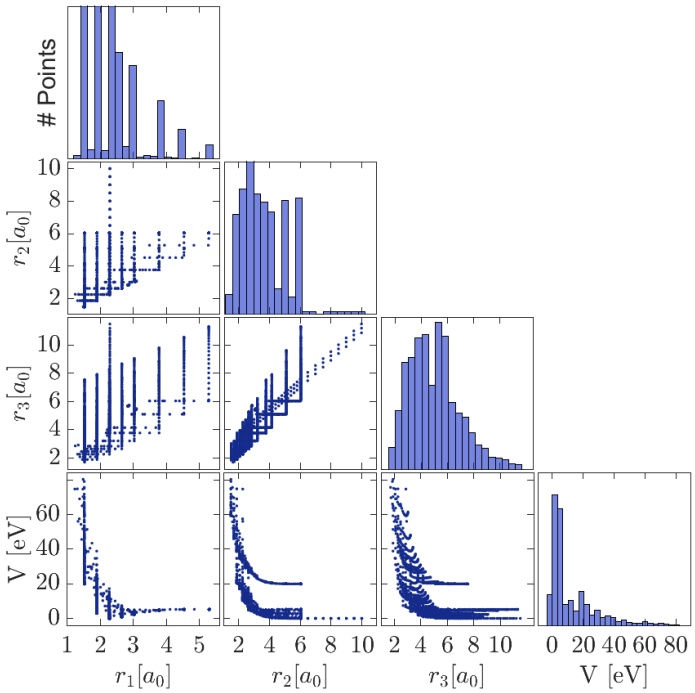
In the present study, we focus on the interaction of O_2 molecules and O atoms, which can take place on singlet, triplet, and quintet PESs, due to the two unpaired electrons of both molecular and atomic oxygen in their electronic ground-states. Moreover, the threefold spatial degeneracy of the oxygen atom makes the collision likely to occur on surfaces of different multiplicities: 3 singlets (with symmetries $1^1A'$, $2^1A'$, and $1^1A''$), with statistical weights 1/27 each, 3 triplets ($1^3A'$, $2^3A'$, and $1^3A''$), with statistical weights 3/27 each, and 3 quintets ($1^5A'$, $2^5A'$, and $1^5A''$), with statistical weights 5/27 each. Recently, Varga *et al.* [150] computed all 9 adiabatic PESs. These surfaces are hereafter referred to as either Varga's PESs or Varga's fits. There is also an older PES for O_2+O interactions by Varandas *et al.* [149], only for the $1^1A'$ surface.

In the present work, the $1^1A'$ and $2^5A'$ potential surfaces are adopted as test cases for the UQ framework presented in the previous sections. For these two surfaces, Varga *et al.*'s database contains 1686 and 1617 spatial configurations respectively, which correspond to potential energies over a range of 80 eV. The data points are reported in Figure 2.4. In particular, we selected the $1^1A'$ surface because it contains the ground-state of the ozone molecule, $O_3(^1A_1)$, which is crucial in atmospheric chemistry. The $2^5A'$ PEs was chosen because it has much higher statistical weight as well as the highest fitting error [150].

In Varga *et al.*'s work, the resulting set of 9 O_2+O PESs were obtained as fits to the *ab initio* data through the many-body expansion reported in Eq. 2.11. The many-body part, V^{MB}, and the pairwise contribution represented by the sum of the three diatomic potentials, V^{Diat}, were fit separately. V^{MB} has been computed at a discrete set of geometric configurations through DSEC-corrected XMS-CASPT2 calculations, and has been then fitted with permutationally invariant polynomials in mixed exponential-Gaussians [165]. The separation of 2-body and 3-body terms in Eq. (2.12) is used also in the present work for separating the contributions to the PES and for isolating the multi-atom interactions. The diatomic potential strongly influences the nonequilibrium



(a)



(b)

Figure 2.4: Spatial configurations and related *ab initio* energies for the $1^1A'$ (a) and $2^5A'$ (b) data points from Varga *et al.* [150].

kinetics, especially for the dissociation rate coefficients. From an uncertainty perspective, however, the representation inadequacy of V^{Diat} can be considered negligible compared to the inaccuracies generated by the modeling of the three-body contribution. This is because the *ab initio* computation of the diatomic interactions are easier to be performed, due to the lower dimensionality of the problem and the fact that the validation phase can generally rely on accurate experimental data. For this reason, the diatomic potential adopted by Varga *et al.* is assumed to be highly accurate, is not varied from its original formula, and the parameters are not changed. All the uncertainty analysis will be focused on the three-body contribution, which is constructed using our new fitting strategy, starting with *ab initio* data generated by Varga *et al.*

2.3.2 Rovibrational-Specific QCT Calculations

All the atoms and molecules in this work are considered to be in their electronic ground-state, and only rovibrational and chemical processes are taken into account. The rotationless ($J = 0$) component of the O_2 diatomic potential from Varga *et al.* is populated by 45 vibrational levels and a total of 6115 bound and quasi-bound rovibrational levels are computed using the WKB [135, 145] semiclassical approximation. In the manuscript, each of them is identified either with the index $O_2(i)$, where $i = 1, 6115$, or the 2-tuple $O_2(v, J)$, which represent the vibrational and rotational quantum numbers. Based on a semiclassical approximation [6, 161], all the 6115 levels are included by accounting for the rotational degeneracy as $g_i = (2J + 1)/2$. As a consequence, the collisions between O_2 molecules and oxygen atoms result in 75 million possible outcomes, which can be divided into:

- Inelastic (non-reactive) processes, whereby the colliding atom does not break the molecule bond:

$$O_2(i) + O_C \underset{k^I_{ji}}{\overset{k^I_{ij}}{\rightleftharpoons}} O_2(j) + O_C, \quad j > i; \tag{2.19}$$

- Exchange reactions, when one of the two atoms composing the molecule ends being replaced by the free particle:

$$O_2(i) + O_C \underset{k^E_{ji}}{\overset{k^E_{ij}}{\rightleftharpoons}} O_2(j) + O_l, \quad j > i, \quad l = A \text{ or } B; \tag{2.20}$$

- Dissociation reactions:

$$\mathrm{O}_2(i) + \mathrm{O}_C \underset{k_i^R}{\overset{k_i^D}{\rightleftharpoons}} \mathrm{O}_A + \mathrm{O}_B + \mathrm{O}_C. \tag{2.21}$$

In Eqs. (2.19), (2.20) and (2.21) the letters A, B and C are used for distinguishing the three atoms taking part in the collision, while k_{ij}^I, k_{ij}^E, k_i^D and k_i^R are adopted for identifying inelastic, exchange, dissociation and recombination rate coefficients. The MAP solutions of the PIP-BNN corresponding to the surfaces $1^1A'$ and $2^5A'$ are employed in this work for simulating the interactions between the oxygen atoms, and for computing the exothermic (*i.e.*, $\epsilon_i > \epsilon_j$, where ϵ_i and ϵ_j are the rovibrational energies of the states i and j, respectively) and dissociation rate coefficients. This operation is performed through QCT calculations, the details of which can be found in the papers by Chaban *et al.* [23], Schwenke *et al.* [135] and Jaffe *et al.* [60], and in the Ph.D. thesis by Macdonald [88]. In the present work, 20 000 collisions per initial O_2 rovibrational state are simulated for each of the translational temperatures in the following list: $T_{Tran} = \{2\,500, 5\,000, 7\,500, 10\,000\}$ K. The number of trajectories is chosen in order for the QCT methodology to result in small sampling error, as shown in Subsec.2.3.5, thereby adding a negligible contribution to the uncertainties in the PES. The backward (endothermic and recombination) rate coefficients are obtained by exploiting microscopic-reversibility [114]. The rate coefficient calculations are conducted through the new computer code CoarseAIR, an in-house Fortran 2008 modernization of the original VVTC code developed at NASA Ames by D. Schwenke [135]. The new object-oriented version implements the coarse-grained quasi-classical (CG-QCT) method [90], which is discussed later in this section, and the capability of dealing with ML-based PESs (PIP-NN, PIP-BNN and Gaussian Processes).

2.3.3 Rovibrational State-to-State Master Equation

In this work, we analyze the time evolution of a gas mixture composed of atomic and molecular species in an isochoric reactor. The gas is initially in thermal equilibrium at $T_{Tran}^0 = T_{Int}^0 = 1\,400$ K, pressure $p = 9\,940$ Pa, and molar composition of 5% of O and 95% of O_2. At the start of the simulation, the translational temperature is artificially raised to a higher T_{Tran} and then kept constant. The sudden heating causes the excitation of the O_2 rovibrational levels, which were initially distributed following a Boltzmann function at T_{Int}^0, and the eventual dissociation of the oxygen molecules. The resulting kinetics of each of the ro-vibational levels are described by the

RVS StS master equation (ME):

$$\begin{cases} \frac{d}{dt}n_i = \sum_{j\neq i}(k_{ji}n_j - k_{ij}n_i)n_O + k_i^R n_O^3 - k_i^D n_i n_O \\ \\ \frac{d}{dt}n_O = \sum_i (k_i^D n_i n_O - k_i^R n_O^3), \end{cases} \tag{2.22}$$

where n_O and n_i respectively represent the populations of the oxygen atoms and of the oxygen molecule in the i-th rovibrational state, and where the excitation rate coefficient, $k_{ij} = k_{ij}^I + k_{ij}^E$, does not distinguish the exchange reaction from the inelastic mechanism. The RVS StS ME represents a system of 6115+1 coupled ordinary differential equations (ODEs).

Through the e-folding method [119], it is finally possible to reconstruct the characteristic times for vibration and rotation relaxation (τ_V and τ_R) of the Landau-Teller model [78]:

$$\frac{de_{Mode}(t)}{dt} = \frac{e_{Mode}(t) - e^*_{Mode}(T_{Tran})}{\tau_{Mode}}, \tag{2.23}$$

where $e_{Mode} = e_V$ (i.e, vibrational energy) is here distinguished from $e_{Mode} = e_R$ (*i.e.*, rotational energy) in a vibration-prioritized framework [114]. The heat bath simulations through the RVS StS ME and the coarse-grained model (CGM) [85, 86], which is discussed later in this section, are performed employing the KONIG code and the PLATO thermochemical library, both by A. Munafò [102, 103, 107, 108].

2.3.4 Reduced-Order Modeling: CG-QCT and CGM

In order to obtain the full set of converged RVS StS rate coefficients for the MAP surface, the number of trajectories required is approximately 50×10^8. Consequently, forward propagating the PES uncertainties by means of computing the RVS StS rate coefficients from the ensemble of surfaces in Eq. (2.17) would demand N_{LHS}-times more collisions. In order to avoid an intractable computational effort, for the QCT calculations, we reduce the resolution to groups of states, based on the CG-QCT method suggested by Macdonald *et al.* [90] and extensively discussed in Macdonald's Ph.D thesis [88]. In this approach, instead of constraining the molecule to start from a specified rovibrational energy, at each collision the initial level $O_2(i)$ is randomly picked from a group of

states p, also called bin; this selection takes place through a Boltzmann-weighted sampling. The outcomes of this procedure are the rate coefficients for each group,

$$k_{pq} = \sum_{i \in p} \sum_{j \in q} k_{ij} f_p^i(T_{Tran}), \tag{2.24}$$

where $f_p^i(T_{Tran})$ represents the Boltzmann distribution at the translational temperature. The rate coefficients can be used for simulating nonequilibrium kinetics in the heat bath by means of the efficient CGM developed by Liu *et al.* [85, 86]. In this approach, the reduced model is constructed with the ansatz that rovibrational states that are connected by preferential transition pathways are more likely to be found in local equilibrium with each other. For this reason, they are kept in the same bin. The specific grouping strategy adopted in this work is the adaptive grouping method proposed by Sahai *et al.* [129]. Details with regards to this technique can be found in Sahai's Ph.D. thesis [128], while the generic reduced-order approach, only briefly mentioned in this section, will be the main topic of the second part of this manuscript.

2.3.5 Convergence of QCT and CG-QCT Calculations

In this subsection, the convergences of the QCT and CG-QCT calculations are presented, with the objective of showing that a further increase in the number of computed trajectories would not significantly change the predicted values of the gas QoIs.

Four test cases are analyzed, all of them involving the simulations of gas mixtures composed of atomic and molecular oxygen undergoing thermal baths in an isochoric reactor. The interactions between O_2 and O are assumed to take place only on the first singlet surface, $1^1A'$. We preferred $1^1A'$ to the other 8 O_2+O available PESs due to the relatively large number of resulting processes that its low exchange barrier makes likely to happen.

The gas is considered in initial thermal equilibrium at $T^0_{Tran} = T^0_{Int} = 1\,400$ K, pressure $p = 9\,940$ Pa, and mole fraction composition of 5% of O and 95% of O_2. At the first instants of the simulations, the translational temperature is artificially raised to higher T_{Tran} ($T_{Tran} = 2\,500$ K for the first and the third test cases, and $T_{Tran} = 10\,000$ K for the second and the fourth ones), and then kept constant for all the following times. The heat baths are simulated by means of the ME (test

cases 1 and 2) and 200 bin CGM (test cases 3 and 4).

In the first two test cases, the results obtained by using StS rates computed through 20 000 trajectories per level are compared to the ones produced by StS rates generated through 100 000 trajectories per level (Figure 2.5 with Table 2.2 for $T_{Tran} = 2\,500$ K and Figure 2.6 with Table 2.3 for $T_{Tran} = 10\,000$ K). In the remaining two cases, instead, the results of CGM obtained by using bin rates grouped from the StS ones are compared to the outcomes of CGM with bin rates directly computed through CG-QCT, with 50 000 trajectories per group (Figure 2.7 with Table 2.4 for $T_{Tran} = 2\,500$ K and Figure 2.8 with Table 2.5 for $T_{Tran} = 10\,000$ K). For both the highest and lowest temperatures, the figures and the table prove that any increase in the number of trajectories simulated in the study chapter would not produce any significant improvement in the accuracy on the main QoIs. Moreover, the model reduction performed before the trajectory calculations based on the CG-QCT strategy does not affect the reliability of the group rate coefficients and their resulting heat bath quantities.

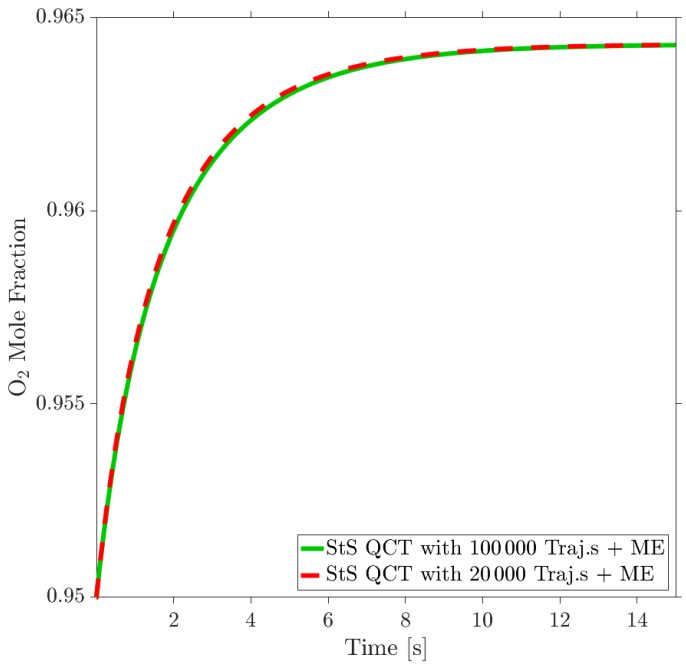


Figure 2.5: Test case 1: evolution of O_2 mole fraction in a StS heat bath at constant $T_{Tran} = 2\,500$ K. The rates for the ME calculations are computed from 100 000 (green solid line) and 20 000 (red dashed line) trajectories per level.

Table 2.2: Test case 1: gas QoIs at $T_{Tran} = 2\,500$ K from StS Calculations

QoI	20 000 Traj.s	100 000 Traj.s
$p\tau_V$ $[atm * s]$	2.308e-09	2.315e-09
$p\tau_R$ $[atm * s]$	2.116e-09	2.136e-09
k^D_{QSS} $[cm^3/s]$	-	-
k^D_{Eq} $[cm^3/s]$	3.077e-19	2.995e-19
k^E_{QSS} $[cm^3/s]$	-	-
k^E_{Eq} $[cm^3/s]$	1.257e-10	1.266e-10

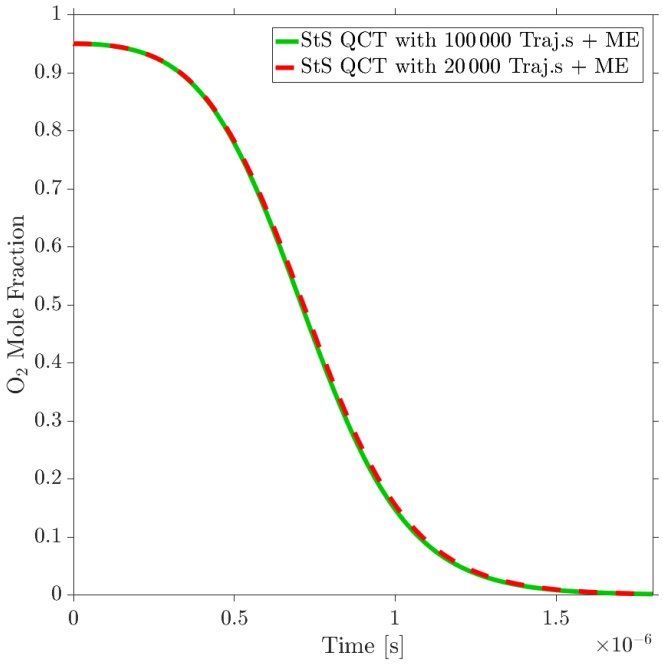


Figure 2.6: Test case 2: evolution of O_2 mole fraction in a StS heat bath at constant $T_{Tran} =$ 10 000 K. The rates for the ME calculations are computed from 100 000 (green solid line) and 20 000 (red dashed line) trajectories per level.

Table 2.3: Test case 2: gas QoIs at $T_{Tran} = 10\,000$ K from StS Calculations

QoI	20 000 Traj.s	100 000 Traj.s
$p\tau_V$ $[atm * s]$	7.208e-09	7.205e-09
$p\tau_R$ $[atm * s]$	5.454e-09	5.459e-09
k^D_{QSS} $[cm^3/s]$	5.932e-12	6.011e-12
k^D_{Eq} $[cm^3/s]$	9.570e-12	9.654e-12
k^E_{QSS} $[cm^3/s]$	3.128e-10	3.116e-10
k^E_{Eq} $[cm^3/s]$	3.201e-10	3.190e-10

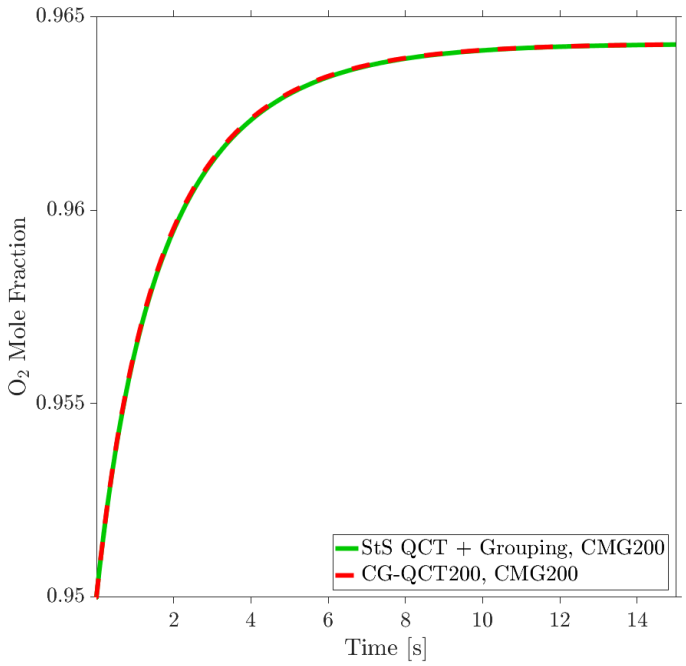


Figure 2.7: Test case 3: evolution of O_2 mole fraction in a heat bath at constant $T_{Tran} = 2\,500$ K simulated by means of CGM200. The results are obtained through bin rates computed by grouping the StS ones (green solid line) or directly by CG-QCT calculations (red dashed line).

Table 2.4: Test case 3: gas QoIs at $T_{Tran} = 2\,500$ K from CGM200

QoI	CG-QCT200	QCT + 200 Bins Grouping
$p\tau_V$ $[atm * s]$	2.039e-09	2.035e-09
$p\tau_R$ $[atm * s]$	1.543e-09	1.540e-09
k^D_{QSS} $[cm^3/s]$	-	-
k^D_{Eq} $[cm^3/s]$	3.009-19	2.977e-19
k^E_{QSS} $[cm^3/s]$	-	-
k^E_{Eq} $[cm^3/s]$	1.275e-10	1.266e-10

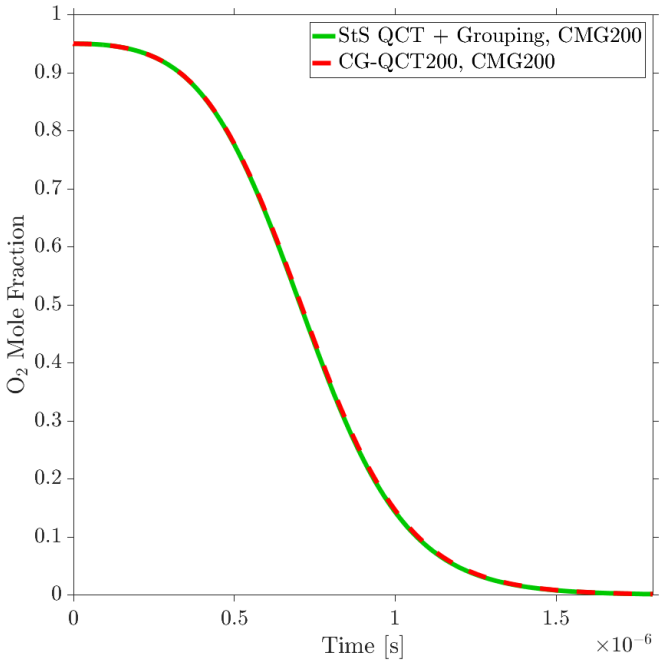


Figure 2.8: Test case 4: evolution of O_2 mole fraction in a heat bath at constant $T_{Tran} = 10\,000$ K simulated by means of CGM200. The results are obtained through bin rates computed by grouping the StS ones (green solid line) or directly by CG-QCT calculations (red dashed line).

Table 2.5: Test case 4: gas QoIs at $T_{Tran} = 10\,000$ K from CGM200

QoI	CG-QCT200	QCT + 200 Bins Grouping
$p\tau_V$ $[atm * s]$	6.860e-09	6.884e-09
$p\tau_R$ $[atm * s]$	4.937e-09	4.917e-09
k^D_{QSS} $[cm^3/s]$	6.020e-12	6.054e-12
k^D_{Eq} $[cm^3/s]$	9.651e-12	9.654e-12
k^E_{QSS} $[cm^3/s]$	3.140e-10	3.118e-10
k^E_{Eq} $[cm^3/s]$	3.213e-10	3.190e-10

Chapter 3

Applications to the O_2+O System

In this chapter, the methodology discussed earlier is used to study energy transfer and dissociation processes for a wide range of conditions for the O_2-O system in an ideal chemical reactor. The first part of the section discusses the results of the calibration process by providing comparisons with the *ab initio* data points and the original Varga's PESs [150]. The remainder of the section addresses the forward propagation of the uncertainty to predefined quantities of interest. To this end, extensive QCT calculations are performed to generate the reaction rate parameters to be used in the 0D chemical reactor solver. The section ends discussing the sensitivities of the quantities of interest to the accuracy of the potential energy surfaces.

3.1 Verification Cases for the PIP-BNN

The first step in the SPES methodology is the calibration of the PES using *ab initio* data. To gain confidence in the validity of the methodology, we have first performed extensive calibration and forward propagation using manufactured data obtained by sampling the Varga's PES [150]. Three different cases are proposed, and in each of them a stochastic PES is generated based on fabricated data. In particular, the geometric arrangements for the data points are selected as follows:

- Validation Case 1: The 1617 atomic configurations for the $2^5A'$ surface reported in Varga *et al.*'s paper, plus 8383 geometries for which r_1, r_3 and α are sampled through LHS with $1.5a_0 < r_1 < 10.0a_0$, $1.5a_0 < r_3 < 10.0a_0$, and $35^\circ < \alpha < 175^\circ$ ($r_1 = \overline{O_AO_B}$, $r_3 = \overline{O_BO_C}$ and $\alpha = \angle O_AO_BO_C$);

- Validation Case 2: The 1617 atomic configurations for the $2^5A'$ surface reported in Varga *et al.*'s paper, plus 3383 geometries sampled through the same strategy as in Test Case 1;

- Validation Case 3: The 1617 atomic configurations for the $2^5A'$ surface reported in Varga *et al.*'s paper.

In order to generate labeled data, the corresponding potential energies are obtained by evaluating Varga *et al.*'s fit for $2^5A'$ at such geometric arrangements. The results from the three test cases are compared to each other, in order to show that the error in reconstructing Varga *et al.*'s PES monotonically reduces with increasing the number of data points (Figures 3.1, 3.2, 3.3).

The PIP-BNN stochastic PES generated through verification case 3 is then adopted for QCT and CG-QCT calculations at $T_{Tran} = 10\,000$ K; the resultant StS and group rates are used for simulating a heat bath, with the goal of showing that the gas QoIs predicted by the so constructed PIP-BNN are in good agreement with the ones generated with Varga *et al.*'s fit. The gas is considered in initial thermal equilibrium at $T^0_{Tran} = T^0_{Int} = 1\,400$ K, pressure $P = 9\,940$ Pa, and mole fraction composition of 5% of O and 95% of O_2. At the first instants of the simulations, the translational temperature is artificially raised to $T_{Tran} = 10\,000$ K, and then kept constant for all the following times. i) Mole fractions (Figure 3.4), ii) QSS and equilibrium dissociation rate coefficients (Figure 3.5), and iii) vibrational and rotational relaxation times (Figure 3.6) have been analyzed as quantities of interest. For all these QoIs, the predictions made through the SPES are in relatively good agreement with the ones obtained by relying on the original PES. These results, other than verifying the correct implementation of the SPES methodology, prove that data point locations and NN structure complexity are sufficient to the Bayesian neural network for reproducing the fitting technique by Varga *et al.*

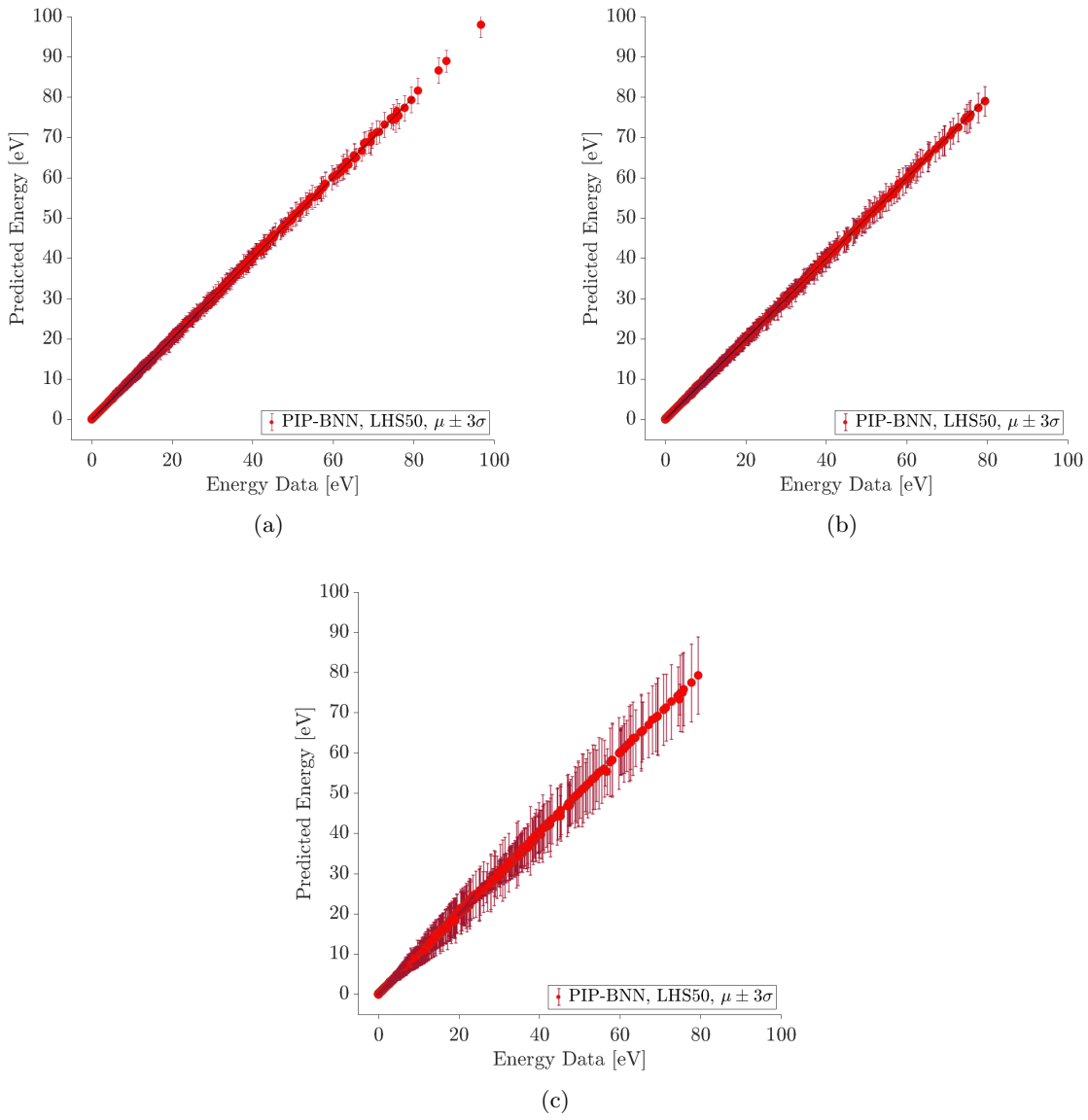


Figure 3.1: Means (red dots) and three-sigma confidence intervals (red lines) for the potential energies computed at the data points using 50 PIP-BNN samples for verification case 1 (a), verification case 2 (b), and verification case 3 (c).

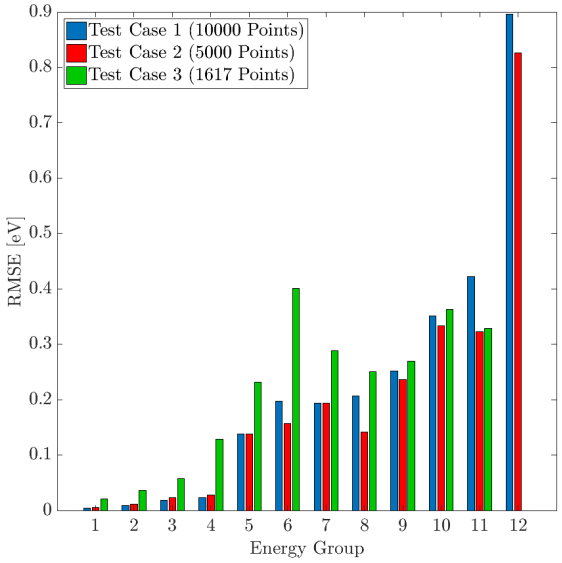


Figure 3.2: RMSEs from the mean of the 50 PIP-BNN samples at the data points of $2^5 A'$ for the 3 verification cases (blue for case 1, red for case 2 and green for case 3). Based on their potential energies, the data points have been divided in 12 groups, with upper bounds given by $V_{Max} = \{2.0;\ 4.0;\ 6.0;\ 8.0;\ 10.0;\ 15.0;\ 20.0;\ 25.0;\ 30.0;\ 50.0;$
$100.0;\ 1\,000.0\}$ eV

.

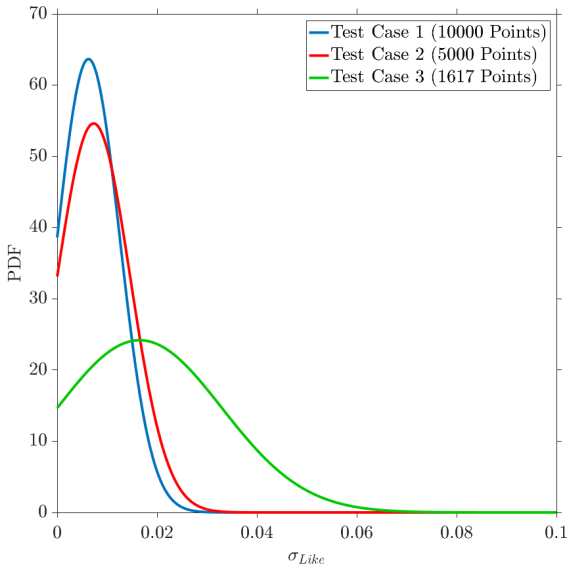


Figure 3.3: Posterior distributions of the likelihood function's standard deviation for the 3 verification cases (blue for case 1, red for case 2 and green for case 3).

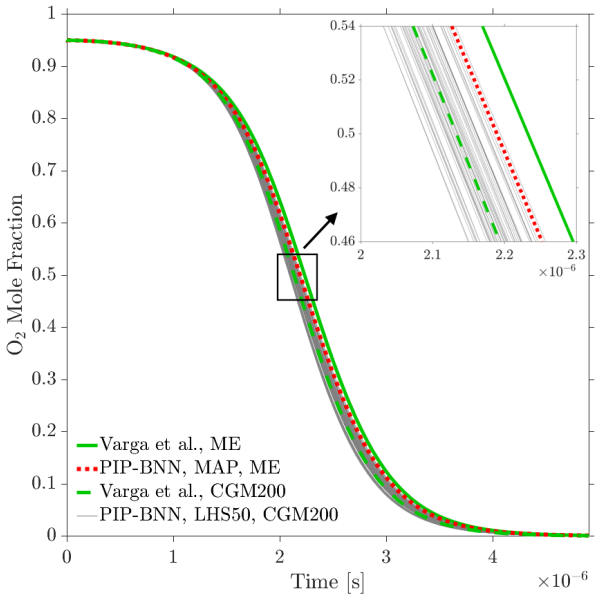


Figure 3.4: Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at $T_{Tran} = 10\,000$ K.

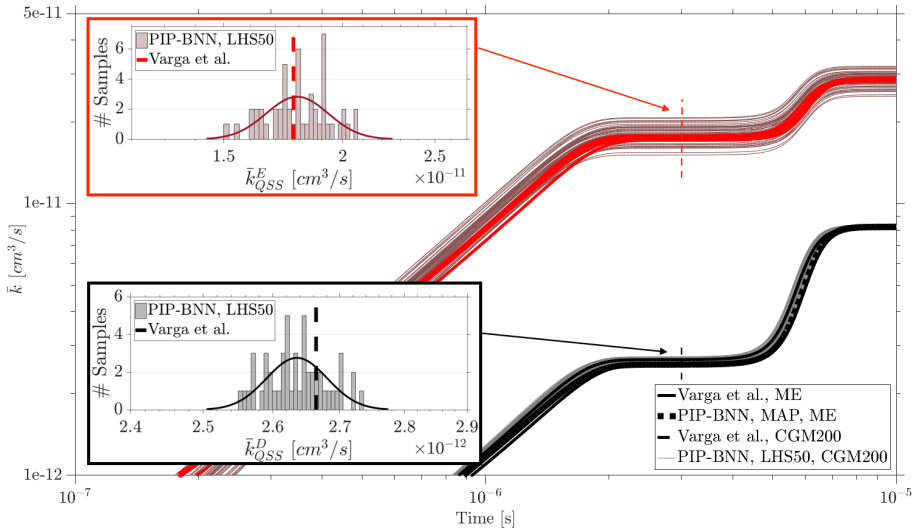


Figure 3.5: Evolution of O_2 dissociation (lower black part) and exchange (upper red part) averaged rates, as predicted by ME (solid lines) and CGM200 (dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (dotted lines), and by CGM200 using the 50 PIP-BNN samples (narrow lines). The histograms on the left report the dissociation (lower Subfig.) and exchange (upper Subfig.) QSS rates predicted through such samples, and they compare them to the outcomes of CGM200 using Varga *et al.*'s fit. The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at $T_{Tran} = 10\,000$ K.

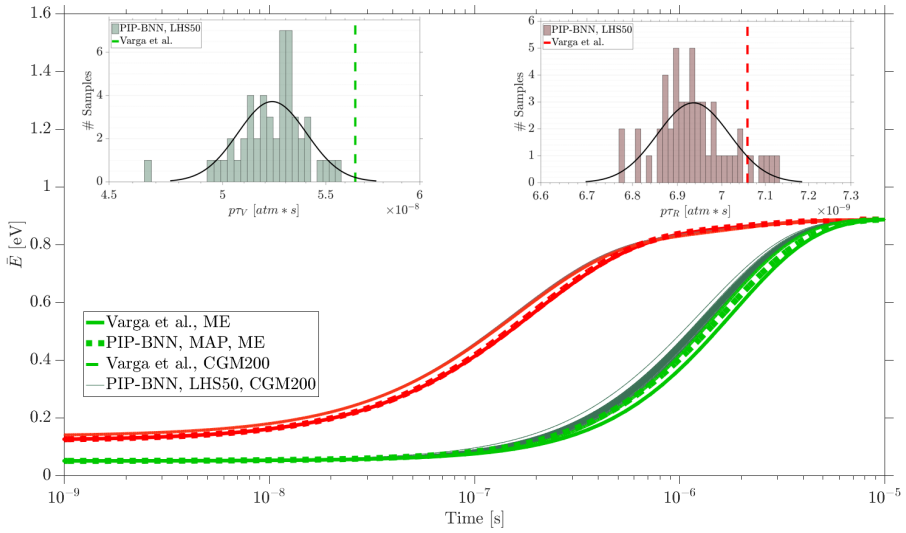


Figure 3.6: Evolution of O_2 averaged vibrational (green lines) and rotational (red lines) energy, as predicted by ME (solid lines) and CGM200 (dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (dotted lines), and by CGM200 using the 50 PIP-BNN MAP samples (narrow lines). The sub-figures on the top report the histogram of the vibrational and rotational relaxation times predicted through such samples, and they compare them to the outcomes of CGM200 using Varga *et al.*'s fit. The PIP-BNN adopted is the one constructed for test case 3, and the simulations are performed at $T_{Tran} = 10\,000$ K.

3.2 Errors and Uncertainties of the $1^1A'$ and $2^5A'$ Surfaces

The SPES methodology was then applied to the construction of a stochastic representation of the $1^1A'$ and $2^5A'$ surfaces by using the actual *ab initio* data points. Figure 3.7 and Figure 3.8 respectively compare the errors of the Varga $1^1A'$ and $2^5A'$ PESs to the related calibrated SPES. Despite the low number of surfaces produced from the PES posteriors, the confidence intervals do not appear to be suffering from undersampling. In fact, if 3 000 samples are generated rather than 50, the expected values and the standard deviations of the potential energies predicted at the data points do not appreciably change. The small number of samples required for representing the posterior distribution is credited to LHS, which ensures that the resulting family of surfaces resembles the real variability of the stochastic PES.

The confidence intervals for the high energy data points are associated with standard deviations up to 2.5 eV (about 3% of the energy values) for the $1^1A'$ PES, and up to 3.5 eV (about 4%) for the $2^5A'$ PES. Through the PIP-BNN, the construction of the PES regions characterized by lower energies is not independent of the shaping of the upper portions. As a single hyper-parameter σ_{Like} is adopted for characterizing the likelihood function over the entire potential domain, any inaccuracy at low energies influences the uncertainty on the higher values and vice versa. The choice of multiplicative error imposed through Eq. 2.11, however, ensures the asymmetry of this relation by guaranteeing higher accuracy in the regions more important for the collision simulations, and it results in narrower error bars at lower energies.

In order to further compare the fitting capabilities of the deterministic approach by Varga *et al.* to the stochastic PIP-BNN, Figure 3.9 compares the errors generated by the two methodologies. Based on the values of the potential energies, the data is divided into 11 groups, and the RMSEs are computed for each of them. The RMSEs produced by averaging LHS50 are consistently lower than the ones obtained with Varga's fits for both quintet and singlet surfaces. Concerning the singlet PES, the groups' RMSEs are negatively correlated to the number of data points populating the respective ranges. The PIP-BNN grants remarkable fitting improvement in the regions characterized by potential energies higher than 10 eV, with the errors being reduced by more than 50% with respect to the original fit. For the quintet PES, comparable refinements are localized in the

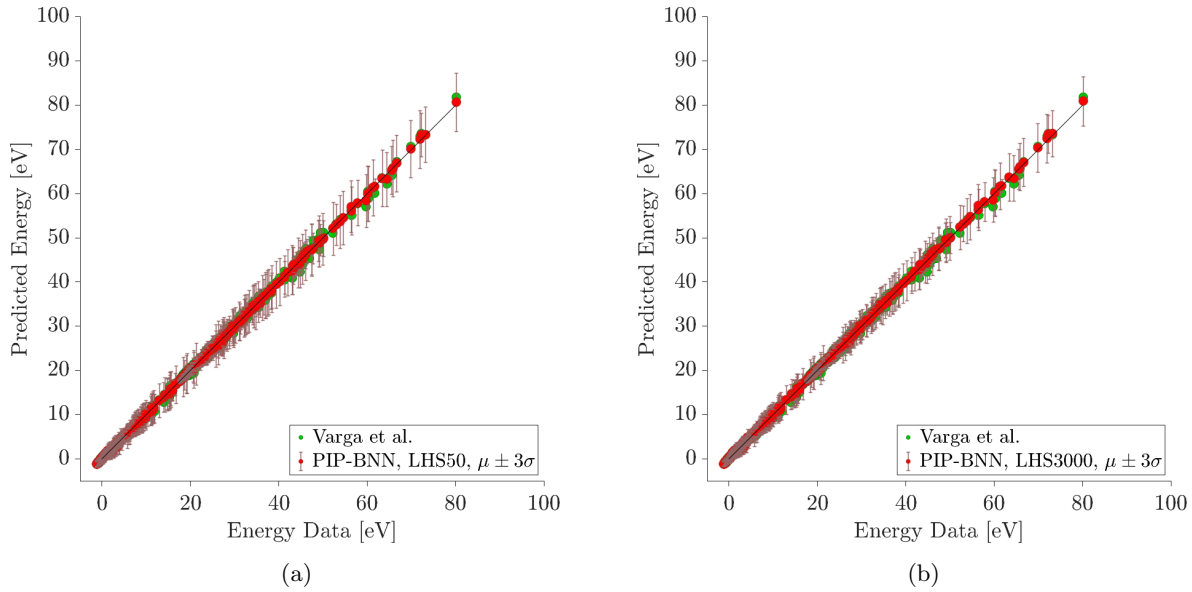


Figure 3.7: Means and three-sigma confidence intervals of the $1^1A'$ potential energies computed at the data points using 50 PIP-BNN samples (a, red dots) and 3000 PIP-BNN samples (b, red dots), compared to the energies resulting from Varga et al.'s fits (green dots).

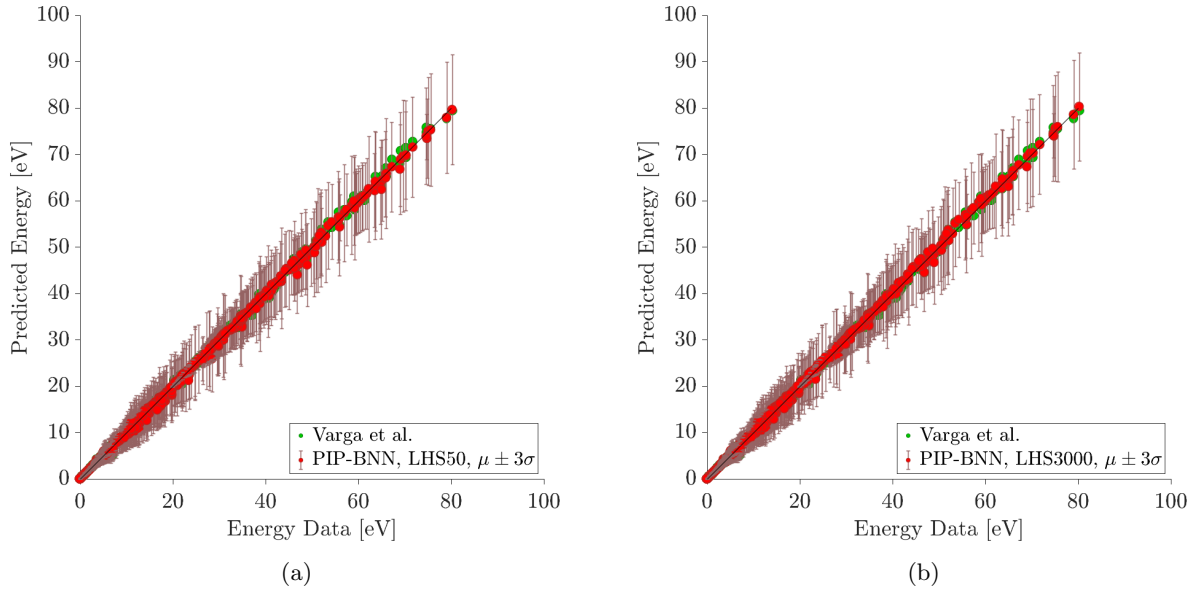


Figure 3.8: Means and three-sigma confidence intervals of the $2^5A'$ potential energies computed at the data points using 50 PIP-BNN samples (a, red dots) and 3000 PIP-BNN samples (b, red dots), compared to the energies resulting from Varga et al.'s fits (green dots).

lower energy regions of the surface. For energies below 5 eV, both the Varga and PIP-BNN fits perform considerably better for the quintet surface than the singlet, even if significantly more data

points are available in this energy range for the latter. This is a consequence of the complexity of the $1^1A'$ surface, which includes a deep well characterizing the ozone molecule.

This characteristic is more evident in Figure 3.10, which presents the two-dimensional cuts of

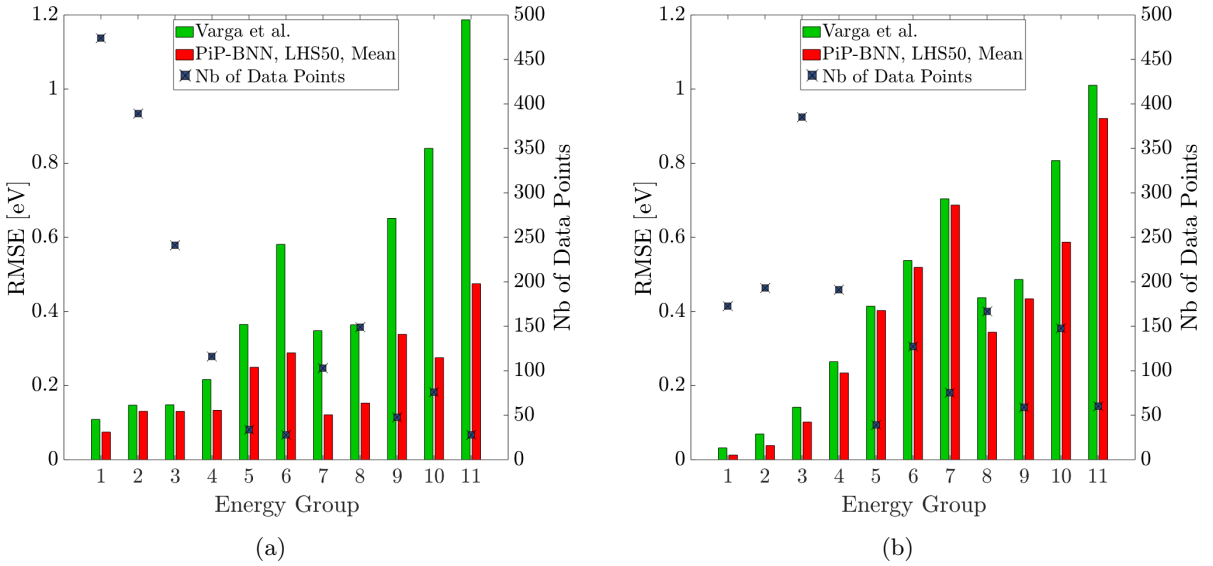


Figure 3.9: RMSEs at the $1^1A'$ (a) and $2^5A'$ (b) data points. In green, the errors produced by Varga *et al.*'s fits; in red, the ones generated by the means of 50 PIP-BNN samples. Based on their potential energies, the data points have been divided into 11 groups, with upper bounds given by $\boldsymbol{V_{Max}} = \{2.0;\ 4.0;\ 6.0;\ 8.0;\ 10.0;\ 15.0;\ 20.0;\ 25.0;\ 30.0;\ 50.0;\ 100.0\}$ eV. The minimum of the diatomic potential is here assumed as reference energy. The blue squares represent the number of *ab initio* data points contained in each energy group.

the various surfaces as functions of the distance between the O_A and O_B atoms (r_{A-B}) for an arrangement of the projection plane chosen by Varga *et al.* for cutting the PESs close to the minimum-energy path of ozone dissociating in O_2+O [150]. As the LHS50 lines attest, the uncertainty at low energy values computed by the PIP-BNN is reduced as r_{A-B} increases. As r_{A-B} tends to infinity, indeed, the O_2+O interaction energy tends to zero. At smaller atomic distances, the PES variability of the quintet interactions is associated with the slope of the repulsive wall, where Varga's fit and the PIP-BNN predictions are in fairly good agreement. In the case of the singlet interactions, instead, the uncertainty mostly affects the location and the value of the energy at the local minimum, and Varga's fit markedly overestimates the slope of the potential for O_A moving towards O_B.

For the case of 3 like atoms, 3D plots can be extremely helpful in illustrating the overall deficiencies

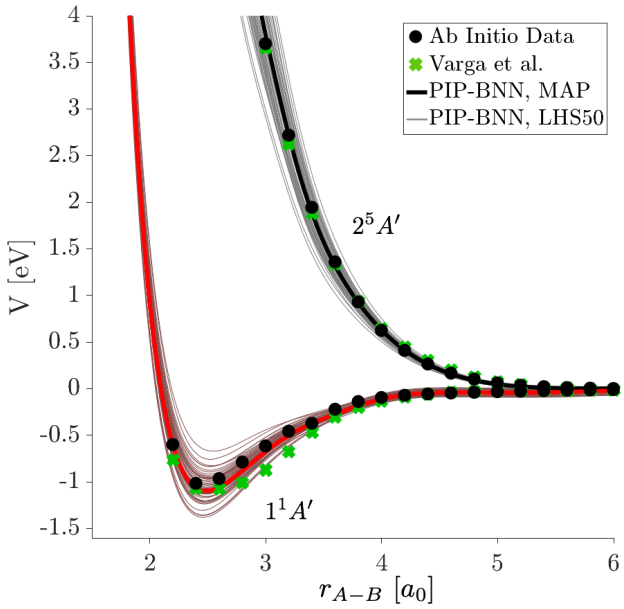
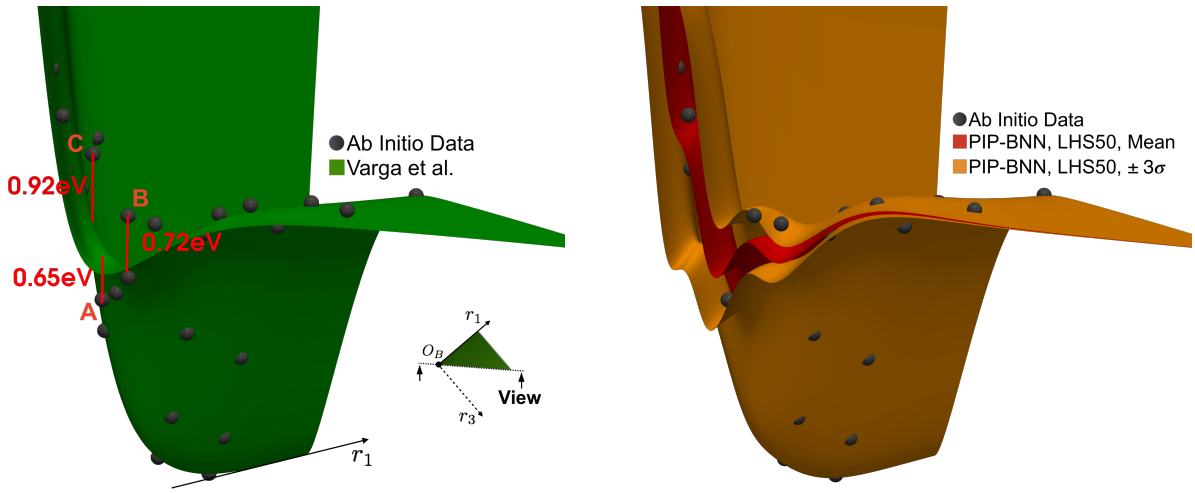


Figure 3.10: $1^1A'$ (lower part of the figure) and $2^5A'$ (upper part) PESs as a function of r_{A-B} for $\measuredangle O_AO_BO_C = 116.75°$ and $r_{B-C} = 2.282\ a_0$. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fits at the data geometrical configurations, bold lines to PIP-BNN MAP estimates, and the remaining narrow lines to the 50 samples per PIP-BNN.

of the fitting techniques. For example, Figure 3.11 compares the $2^5A'$ PESs at $\measuredangle O_AO_BO_C = 50°$ to the related *ab initio* points, identified by gray dots. The PESs and the data points have been cut through the plane perpendicular to the $r_1 - r_3$ quadrant and parallel to the energy axis, and only their $r_1 > r_3$ region is shown. The energy values range between 0 and 8 eV, and they plateau at the dissociation energy of the rovibrational ground-state (*i.e.*, about 5.2 eV). The figure highlights two regions of the $2^5A'$ interactions where Varga's fit fails to reproduce the *ab initio* energies: the exchange barrier, with data points A and B being of from Varga's PES by -0.65 eV and +0.72 eV respectively, and the repulsive wall, with the vertical distance between the point C and the fit measuring 0.92 eV. With regards to the first region, the Varga surface is unable to capture the local maximum occurring in the region around $r_1 = r_3 = 4\ a_0$, and the fit cuts through points A and B. Focusing on the repulsive wall, the original fit too weakly protrudes for replicating point C, characterized by an energy value of 5.56 eV. These two shortcomings are a common occurrence for $\measuredangle O_AO_BO_C$ between 50° and 100° (see Figures A.4, A.5, A.6, A.7, A.8, and A.9 of Appendix A). In their paper, Varga *et al.* mentioned the deficiency of their fit in reproducing

the sharp changes of the *ab initio* points between 45° and 65°, which they attributed to multiple avoided crossings. The SPES methodology allows for significant improvement in the description of the main features of the data points: the new exchange barrier presents a hump, which propagates as a wavefront starting at $\measuredangle O_A O_B O_C \sim 50°$ from the plateau region of the dissociation energy to the wall at $\measuredangle O_A O_B O_C \sim 100°$. Furthermore, for small r_1, the repulsive portions of the LHS50 surfaces present complicated details due to the C data point, rather than smoothly approaching infinity as in the case of Varga's fit. Thanks to these refinements, at points A, B, and C, the mean of the LHS50 surfaces shows fitting errors of -0.34, 0.33, and 0.5 eV, respectively. However, the uncertainties associated with the PIP-BNN predictions are extensive in both the exchange barrier and repulsive wall regions, especially in the aforementioned $\measuredangle O_A O_B O_C$ range. This uncertainty is mostly a consequence of the lack of *ab initio* data where the sharp changes in the quintet interactions take place.

For the $1^1A'$ PES, the PIP-BNN uncertainties are localized near the ozone equilibrium geometry

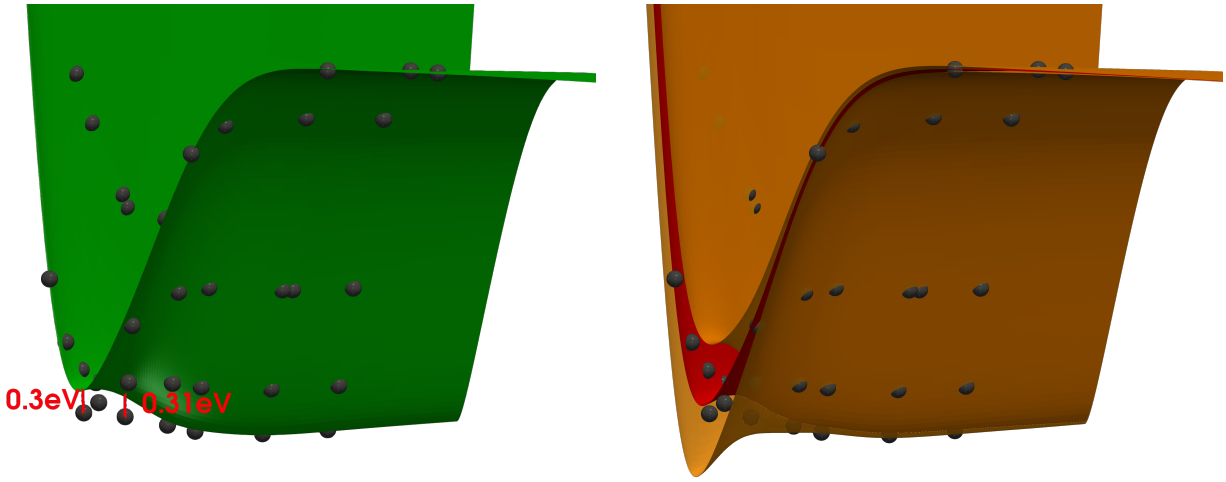


(a) (b)

Figure 3.11: 3D views of the $2^5A'$ PESs at $\measuredangle O_A O_B O_C = 50°$ for $r_1 > r_3$, compared to the correspondent *ab initio* data points (gray dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples.

and exchange region, mainly due to the inability of the fitting to reproduce a few *ab initio* points for $\measuredangle O_A O_B O_C$ between 60° and 100°. The mean-surface from the PIP-BNN samples is associated with a reduction of the fitting error only near the exchange barrier for $\measuredangle O_A O_B O_C$ between 140°

and 160°.



(a) (b)

Figure 3.12: 3D views of the $1^1A'$ PESs at $\measuredangle O_A O_B O_C = 90°$ for $r_1 > r_3$, compared to the correspondent *ab initio* data points (gray dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples.

3.3 Forward Propagation of the SPESs: Heat Bath Simulations at 10 000 K

In this section, the posterior PDFs obtained through the SPES approach are used to assess the uncertainty of the kinetic properties in an isothermal chemical reactor at 10 000 K. A combination of RVS StS and reduced-order model simulations are presented hereafter. Given their computational costs, RVS StS simulations are restricted to the Varga and MAP PESs for the singlet and quintet surfaces. For these cases, RVS QCT calculations (Sec. 2.3.2) are used to compute the RVS StS rate coefficients, which are then employed in the solution of the ME (Sec. 2.3.2). Examples of dissociation RVS StS rate coefficients from the calculations for the $2^5A'$ PES at $T_{Tran} = 10\,000$ K are reported in Figure 3.13. On the other hand, the CG-QCT and CGM methods (Sec. 2.3.4) are applied in series for forward propagating the SPES samples based on 200 distinct groups of energy levels. For each sample and for each group, 50 000 trajectories are run at a fixed translational

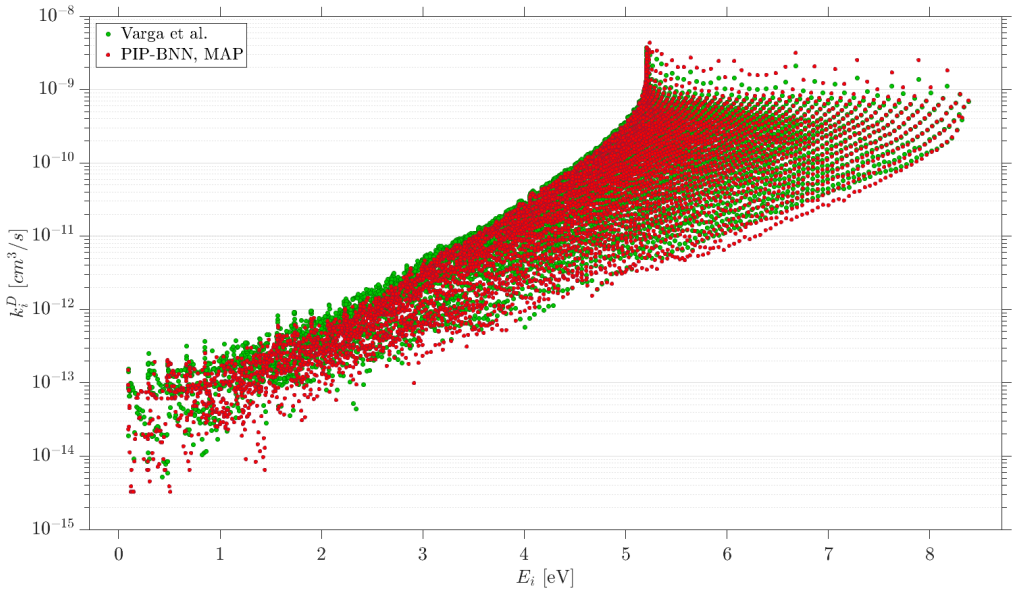


Figure 3.13: Rovibrational state specific dissociation rate coefficients for the $2^5A'$ interactions at $T_{Tran} = 10\,000$ K. Green dots are obtained using Varga *et al.*'s fit, while the red ones are computed through the PIP-BNN MAP surface.

temperature. It is important to mention that the computational cost of the CG-QCT model is one order of magnitude smaller than the RVS StS model.

The results of RVS StS ME and CGM calculations using the rate coefficients for the singlet surface are shown in Figure 3.14. The comparison of the composition profiles obtained through MAP and Varga's fit demonstrates that the modeling error introduced by the coarse-grained procedure is small when compared to the uncertainty associated with the PES, represented by the spread of the gray curves. The largest value of the O_2 mole fraction uncertainty is about $\pm 6\%$ with a 95% degree of confidence in the midst of the dissociation process (t = 0.7 μs). The results obtained with Varga's fit at T = 10 000 K are contained within the confidence interval, and are located in a high-density probability region together with the outcomes of the RVS StS calculations from the MAP surface. Similar observations are made for the singlet PES at other translational temperatures. The predicted exchange and dissociation global rate coefficients, (obtained as averages of the state- or group-specific rate coefficients weighted by the time-dependent distribution function, $\bar{k}^E(t) = \sum_i \sum_j k_{ij}^E f^i(t)$ and $\bar{k}^D(t) = \sum_i k_i^D f^i(t)$ respectively) are also characterized by small uncertainty and good agreement with Varga's PES. The consistency between the kinetic quantities

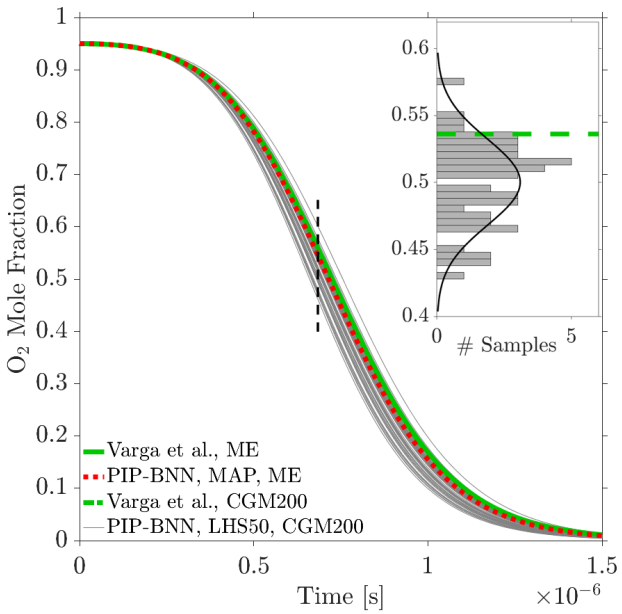


Figure 3.14: Evolution of O_2 mole fraction and their histogram at t = 0.7 μs, as predicted: by RVS StS ME (green solid line), and by CGM200 (overlapping green dashed line), both starting from Varga's fit; by RVS StS ME from PIP-BNN MAP (red dotted line); by CGM200 from the 50 PIP-BNN samples (gray narrow lines). The simulations have been performed at T_{Tran} = 10 000 K with the rate coefficients from the $1^1A'$ PESs.

computed through the original and the MAP surfaces is a consequence of the minor fitting improvements and small variability generated by the PIP-BNN at the $1^1A'$ low energy regions. In this sense, the study of this singlet surface further validates the legitimacy of the SPES methodology and its implementation.

On the other hand, the heat bath calculations for the $2^5A'$ PES show large discrepancies between the results obtained with the PIP-BNN surfaces and the ones from the original Varga's fit. The differences are noticeable at T_{Tran} = 10 000 K and are even larger at lower temperatures, as presented in the next section. For this quintet PES, the mole fraction estimates given by PIP-BNN have twice the variability of the $1^1A'$ equivalent case, and the predictions from the original fit lie outside the 3σ confidence interval (Figure 3.15(a)). Furthermore, based on what is computed by means of RVS StS ME starting from the MAP PES, the oxygen dissociation is slower by 25% when compared to results given by Varga's fit.

To assess the influence of the CGM procedure on the overall uncertainty, we performed RVS StS ME calculations for the PESs responsible for the fastest and slowest dissociation. The analysis,

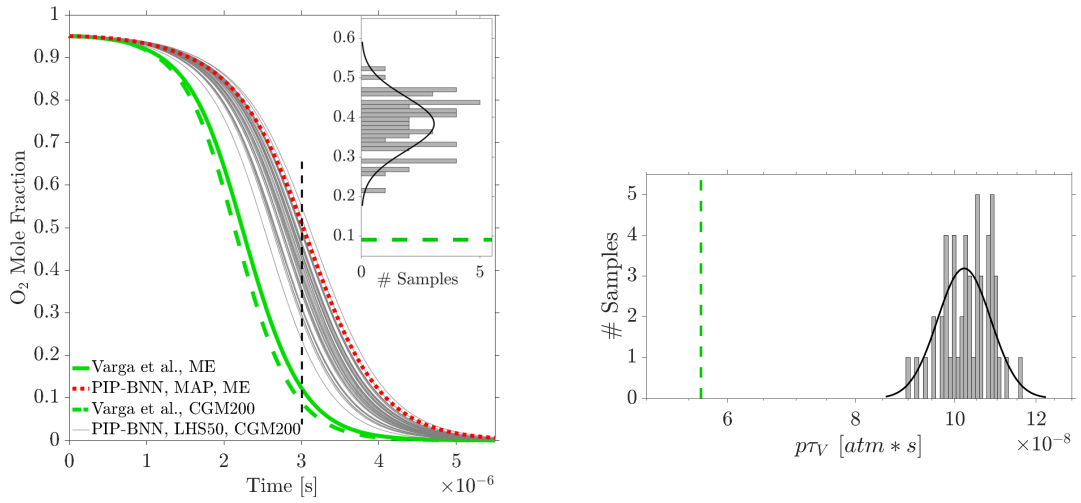


Figure 3.15: (a): Evolution of O_2 mole fractions and their histogram at t = 3 μs. (b): Histogram of the vibrational relaxation times. Green lines are obtained by starting from Varga *et al.*'s fit; in particular: green solid lines are solutions of RVS StS ME; green dashed line are solutions of CGM200. Red dotted lines are solutions of RVS StS ME from PIP-BNN MAP; gray narrow lines are solutions of CGM200 from the 50 PIP-BNN samples. The simulations have been performed at T_{Tran} = 10 000 K with the rates from the $2^5A'$ PESs.

explained in Figure 3.16, confirms the trends obtained with the reduced-order model. The PIP-BNN results are also in good agreement with our previous work [160], which reported the value of 34.5% for the O_2 mole fraction at t = 3 μs. In that paper, the heat bath was simulated through a ME using the RVS StS rate coefficients from an ad hoc PIP-NN PES, while the surface was constructed by using the same set of $2^5A'$ *ab initio* energies adopted here, and the neural network structure was identical to the present one. At that time, however, the parameters were treated deterministically. The small discrepancies between the prediction from the PIP-NN and the one from the MAP are justified by the different regularization adopted by the PIP-BNN under the form of the prior distribution in Eq. 2.10.

The mole fraction variability resulting from the PIP-BNN construction of the $2^5A'$ surfaces necessitates a deeper analysis. For this purpose, we study the thermal relaxation by excluding the dissociation processes in the ME simulations. The relaxation time constants for the vibrational energy resulting from the reduced-order method are shown in Figure 3.15(b). With a confidence level of 99.7%, the vibrational excitation predicted through the original fit takes place between

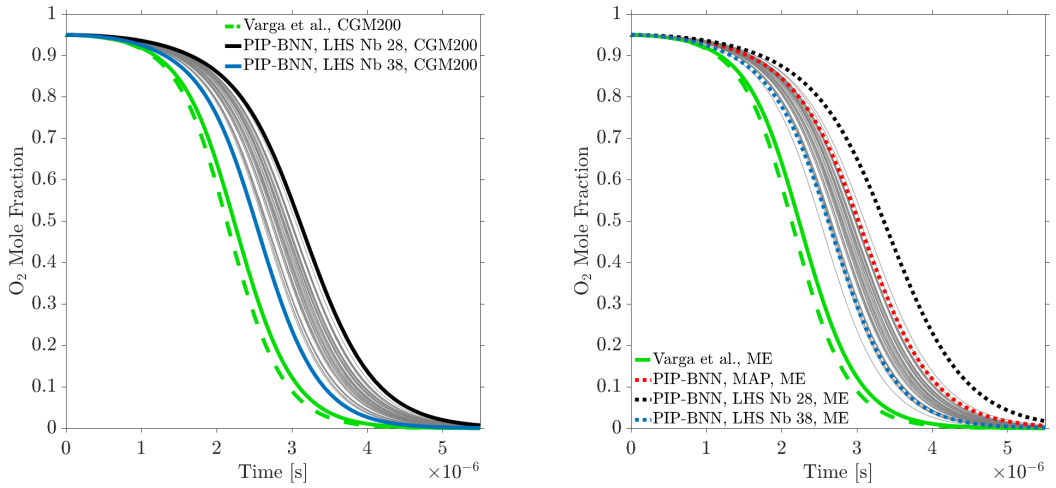


Figure 3.16: Evolution of O_2 mole fraction, as predicted by ME (the green solid line) and CGM200 (the overlapping green dashed line), both starting from Varga *et al.*'s fit, and by: (a) CGM200 from the 50 PIP-BNN samples (grey narrow lines), with the two curves at extremes of the distributions being highlighted in blue and black; (b) CGM200 from the 50 PIP-BNN samples (grey narrow lines), with the blue and black curves in (a) being recomputed via ME calculations (blue and black dotted lines). The simulations have been performed at $T_{Tran} = 10\,000$ K with the rates from the $2^5A'$ PESs.

1.5 to 2.2 times faster than that estimated through the PIP-BNN. The uncertainty affecting this prediction is a direct consequence of the PES variability in the exchange barrier and repulsive wall regions.

The same inadequacies also affect the exchange reaction and dissociation global rate coefficients. However, while for the former quantity the variability after thermalization is comparable to the one during the quasi-steady state, the latter shows uncertainty only at QSS. During that phase, in fact, the 95% confidence interval of the dissociation global rate is bounded by ±9% around the mean and does not contain the prediction from the original fit. On the contrary, its equilibrium value is almost in perfect agreement with the prediction computed through Varga's fit, as it is most sensitive to the diatomic contribution to the PES, which has not been varied from its original formulation. Further details and figures for the global rate coefficients can be found in Subsec.s A.4 and A.5 of the appendix.

3.4 Temperature-Dependent Kinetic Parameters and Comparisons with Experiments

We now analyze the impact that the discrepancies in the $1^1A'$ and $2^5A'$ PESs have on the gas characteristic quantities as a function of translational temperature. These QoIs not only allow insight into the propagation dynamics of the *ab initio* inaccuracies to the macroscopic level, as we saw in the previous sections, but they are also key parameters in modeling nonequilibrium kinetics in a CFD framework [42, 118]. The hybrid nature of the results presented should be noted: we artificially impose the confidence intervals computed from the 50 samples to be centered at the values obtained from the MAP PESs, which are then treated as best estimates. This procedure is followed in order to take full advantage of the high-resolution calculations, while still exploiting the information content gained from the reduced-order calculations with regards to the uncertainties affecting the PESs. The quantification of the uncertainties connected to the remaining 7 PESs for the O_2+O electronic ground-state lies outside the scope of the present work. However, in order to compare with the experiments, we also report the results for the full O_2+O system computed based on all 9 PESs from Varga *et al.* by leveraging RVS StS QCT and RVS StS ME calculations. Figure 3.17 presents the vibrational relaxation times obtained for the $1^1A'$ (in black) and $2^5A'$ (in red) interactions from heat baths at constant $T_{Tran} = \{2\,500, 5\,000, 7\,500, 10\,000\}$ K. For the $1^1A'$ interactions, the graph shows very good agreement between the values resulting from the PIP-BNN and the ones obtained through Varga's fit, which, in turn, almost perfectly match the DMS predictions (not shown) by Grover *et al.* [44]; the uncertainty characterizing these points is inversely proportional to temperature and relatively small, even at 2 500 K. However, for the $2^5A'$ PESs, the calculations from Varga's surface always underestimate the vibrational relaxation rate, and the SPES predicts uncertainties in excess of one order of magnitude at 2,500 K. The simulations for the full O_2+O system based on Varga's fits predict vibrational relaxation times almost constant with temperature. While, at low T_{Tran}, these results are included in the uncertainty bound of the old experiments by Breen *et al.* [18], the trend is in contrast with the more recent trials by Ibraguimova *et al.* [53]. The difference with the latter decreases from a factor four at 2 500 K to a factor two at high temperatures. This discrepancy is consistent with the inaccuracies in the characterization of

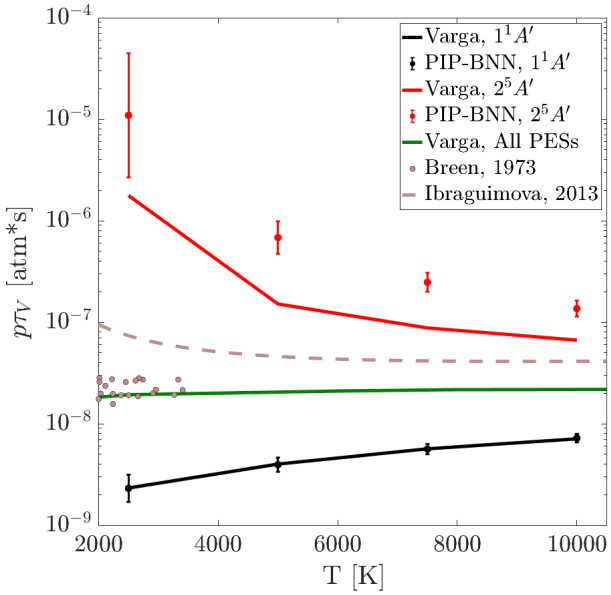


Figure 3.17: Vibrational relaxation times. The values for the $1^1A'$ and $2^5A'$ PESs are in black and in red, respectively. Of these, the solid lines are the values obtained using Varga *et al.*'s fits, the dots the values computed through the PIP-BNN MAP surfaces, and the vertical lines the three-sigma intervals corresponding to the 50 PIP-BNN samples. The green lines represent the values obtained by using all 9 PESs from Varga *et al.* and by either including (solid line) or excluding (dotted line) the exchange reactions. Pink dashed line and symbols identify the experiments by Ibraguimova *et al.* [53] and Breen *et al.* [18], respectively.

$2^5A'$ interactions revealed by the PIP-BNN approach.

The dissociation and exchange reaction rate coefficients at QSS are reported in Figure 3.18 for the $1^1A'$ and $2^5A'$ interactions at $T_{Tran} = \{5\,000, 7\,500, 10\,000\}$ K. At $T_{Tran} = 2\,500$ K, for the scenario analyzed, recombination takes place for a short amount of time before the start of dissociation, and no quasi-steady state is observed for either the singlet or the quintet PESs [160]. In Figure 3.18(b), all the dissociation rate coefficients resulting from simulations are corrected by a factor 16/3, in order to account for the effect of the excited electronic states [110]. The $1^1A'$ predictions from Varga's fit and the ones from the $1^1A'$ PIP-BNN surfaces are in good agreement, but the results involving the $2^5A'$ PESs show noticeable inconsistencies. At QSS, the dissociation rate coefficient estimated by Varga lies outside the three-sigma intervals from the PIP-BNN. However, the discrepancy is relatively small, especially if compared to the uncertainty bounds of the experiments by Shatalov [138] for equilibrium dissociation rate coefficients, corrected for QSS by Andriendko *et al.* [4]. No experiments are at the moment available for the exchange reaction rate coefficient. This

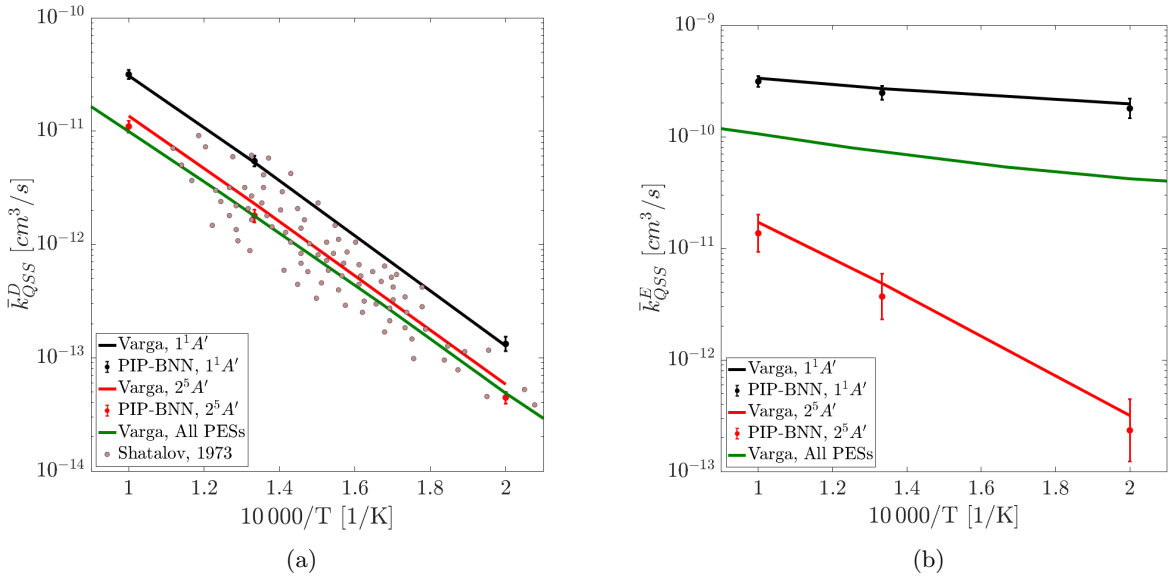


Figure 3.18: Dissociation (a) and exchange reaction (b) rate coefficients at QSS. Black and red colors corresponds to results from the $1^1A'$ and $2^5A'$ PESs, respectively. In particular, the solid lines report the values obtained using Varga *et al.*'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. The green line represents the values obtained by using all 9 PESs from Varga *et al.* Pink symbols correspond to Shatalov's experiments [4, 138].

quantity is found significantly higher for the $1^1A'$ PES than for $2^5A'$ one and nearly independent of the translational temperature, due to the relatively small energy barrier. While almost no variability is propagated from the former PES and good agreement is found between PIP-BNN and Varga's fit, the results from the $2^5A'$ interactions show large uncertainty during QSS as well as at equilibrium (see Figure A.17 in Appendix A).

3.5 Sensitivity Analysis of the $2^5A'$ PES

Sensitivity analysis, described in Sec. 2.2.3, can be used to study the regions of the $2^5A'$ PES that mostly contribute to the uncertainties of the quantities of interest, y. Such regions can significantly vary in relation to the quantity selected when performing the sensitivity study. In order to choose the PES locations to be refined, we first compute the Pearson coefficient [130] $\rho_{V,y}$ for a grid of 362 400 configurations including atomic distances between 1.5 and 8.0 a_0 and angles in the range 35°-175°. After sorting the grid points based on their $\rho_{V,y}$, we choose the nodes with the top 1%

values of $\rho_{V,y}$ as candidates for further selection based on importance sampling. Every time a new candidate point is obtained, we also make sure that the minimum distance to the nodes already accepted is larger than 0.2 a_0 before adding it to the final list.

In Figure 3.19 we present the results for $\bar{k}^E_{Eq}$, on the left, and for τ_V, on the right, which are the two variables that appeared to be affected most by the PES variability. The 3D views show the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples at $\measuredangle O_A O_B O_C = 50°$. The surfaces are identical to the red one in Figure 3.11, but they are colored based on the magnitude of $\rho_{V,y}$.

As expected, the exchange reaction rate at equilibrium is most sensitive to the transition state location on the PES. This is in accord with the physical intuition that the exchange reaction is governed by the effortlessness with which one of the atoms composing the molecule can traverse the energy barrier, a necessary condition for it to break one bond and form another. What is most interesting is that the structure of the valley between the exchange barrier and the repulsive wall has more relevance than the shape of the channel dividing the bump from the plateau, which is more important for dissociation.

The resulting vibrational relaxation times are highly correlated to both the exchange barrier and the repulsive wall regions. The highest sensitivity is found around the area where the many-body term in the potential contributes the most. In particular, the lowest part of the repulsive wall strongly impacts the non-reactive collisional dynamics. Computing more *ab initio* data points in these regions should reduce the uncertainty associated with the $2^5A'$ PES.

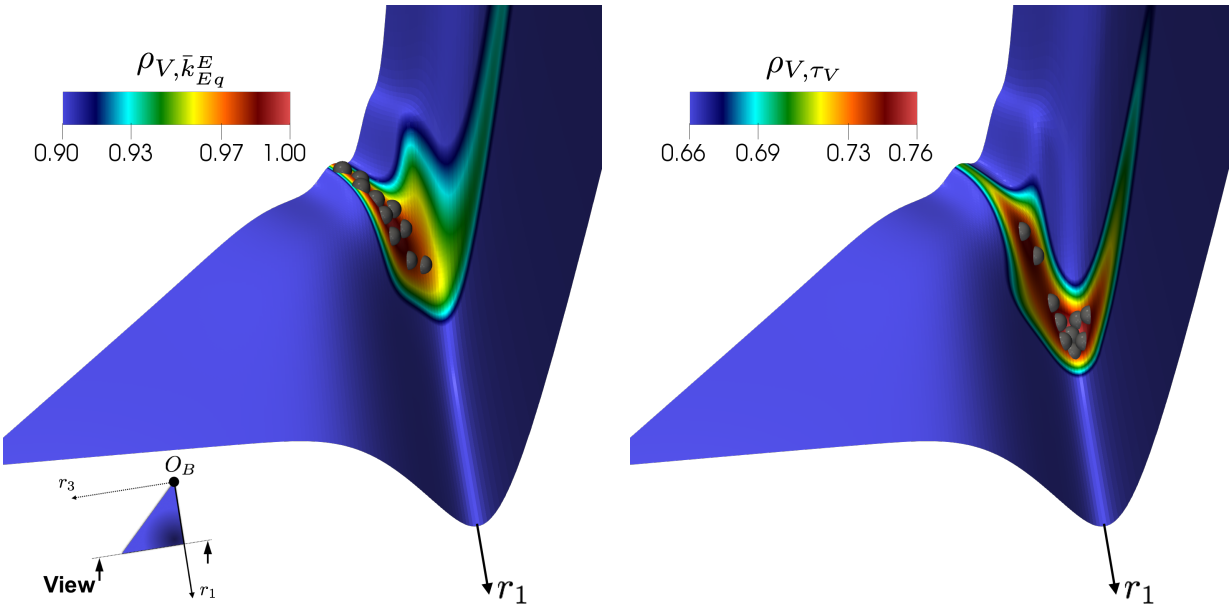


Figure 3.19: 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples at $\measuredangle O_A O_B O_C = 50°$ for $r_1 > r_3$. The surface has been colored based on the absolute value of Pearson correlation coefficients [130] between the PES and the global exchange reaction rate coefficient at the equilibrium distribution of $T_{Tran} = 10\,000$ K (a), and between the PES and the vibrational relaxation time at $T_{Tran} = 10\,000$ K (b). The gray dots identify the locations suggested by the PIP-BNN for performing refinement *ab initio* calculations. The shades of blue in the PES depth are results of rendering effects and do not correspond to different values of Pearson coefficients.

Part II

Data-Inspired and Physics-Driven Model Reduction for Dissociation

Chapter 4

A Novel Grouping Strategy for the Dissociation Mechanisms

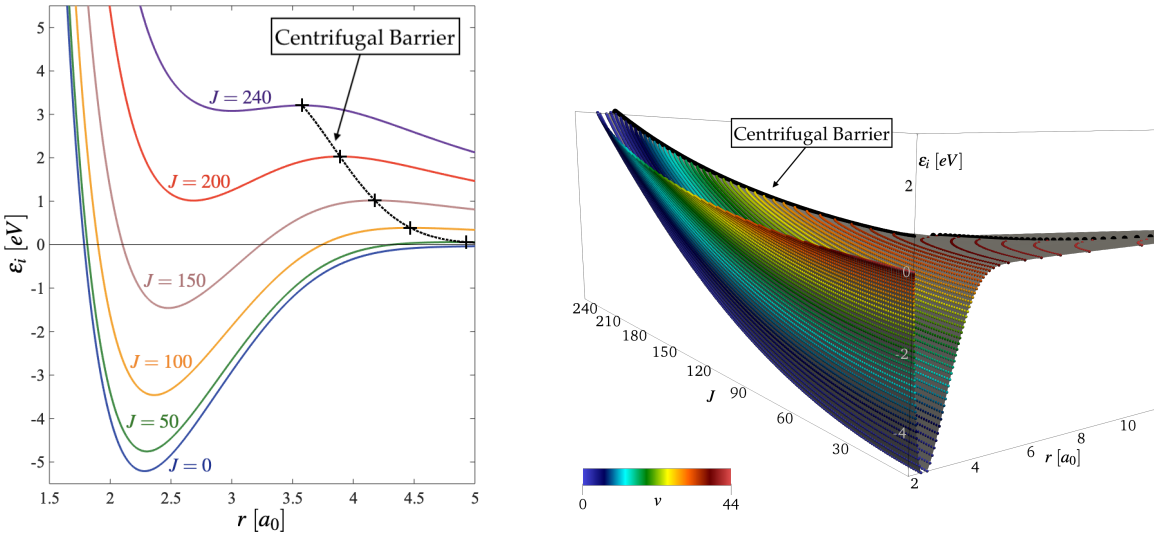
This chapter details the quasi-classical trajectory calculations, the kinetic databases, and the hybrid methodology applied to the study of dissociation.

This chapter is structured as follows: a brief discussion of the O_2 diatomic potential is followed by a description of the methodology used to investigate the nonequilibrium relaxation in a 0-D chemical reactor. The necessary equations, already extensively discussed in the literature [114], are briefly outlined. The reduced order techniques adopted in this work are discussed next, including both the conventional vibrational-specific approach and the proposed centrifugal-barrier-based strategy. The section is concluded by presenting a novel hybrid methodology, which combines a rovibrational-specific treatment of excitation processes and a physics-based model reduction of the dissociation process.

4.1 O_2 Diatomic Potential and O_2+O Potential Energy Surfaces

In the calculations performed in this section, all the nine PESs from Varga et al. [150], already discussed in Sec. 2.3.1, are considered, and statistical weights of 1/27, 3/27, and 5/27 are adopted for the singlet, triplet, and quintet based on the high-temperature limit [150]. As already mentioned, starting from the *ab initio* data points, the PESs are constructed through a two-step fitting procedure based on a multi-body expansion represented in Eq. 2.12. The V^{Diat} is obtained by fitting an accurate diatomic potential, and V^{MB} is constructed by using permutationally invariant polynomials in mixed exponential-Gaussians [165]. The effective diatomic potential is obtained by adding the rotational contribution:

$$V_J^{Diat}(r, J) = V^{Diat}(r) + \frac{J(J+1)\hbar^2}{2mr^2}, \tag{4.1}$$



(a) (b)

Figure 4.1: Two-dimensional and three-dimensional views of the O_2 effective diatomic potential as a function of the distance between the oxygen atoms and of the rotational quantum number. In the right figure, the rovibrational levels are represented as dots at their inner and outer turning points, and they are colored based on their vibrational quantum numbers.

where J indicates the rotational quantum number, $\hbar$ the reduced Planck constant, and m the molecular reduced mass [60]. Figure 4.1(a) shows V_J^{Diat} as a function of the inter-atomic bond length and rotational quantum number. The main effect of molecular rotation is the formation of a centrifugal barrier (black dotted line) [55, 122]: a rotating molecule will dissociate if its internal energy exceeds the barrier value, referred to as $V_{J_{Max}}^{Diat}(J) = \text{argmax}_r \left[V_{J_{Max}}^{Diat}(r, J)\right]$.

The 6 115 rovibrational levels described in Sec. 2.3.2, which include both bound and quasi-bound states, are used also in this study. We indicate the set of rovibrational levels, sorted according to increasing energy, with the symbol $\mathcal{I}$. Again, each state is identified either through an i index or through the 2-tuple $O_2(v, J)$, representing its vibrational and rotational quantum numbers. In the remaining part of the manuscript, the internal energy of the rovibrational level is indicated as ϵ_i, and the dissociation energy of the rovibrational ground-state is chosen as its reference value. As a result, ϵ_1 = -5.113 eV, while ϵ_{6115} = 3.174 eV. The energy deficit from the J-dependent barrier is indicated as:

$$\epsilon_i^D = V_{J_{Max}}^{Diat}(J) - \epsilon_i. \tag{4.2}$$

In Figure 4.1(b), the levels are marked with symbols at their inner and outer turning points and colored based on their vibrational quantum numbers. The three-dimensional visualization proved essential to understanding the behavior of RVS dissociation rate coefficients, as later discussed in chapter 5.

4.2 QCT Calculations and *Ab Initio* Kinetic Database

Also in this part, all the atoms and molecules are considered at their electronic ground-state, and only rovibrational excitation and chemical processes are taken into account. Due to the large number of rovibrational levels involved, the collisions between O_2 molecules and O atoms lead to about 75 million possible outcomes, which have been already described in Sec. 2.3.2, with the rate coefficients for the different collisional processes identified by k^I_{ij}, k^E_{ij}, k^D_i and k^R_i for inelastic, exchange reaction, dissociation and recombination rates respectively. The forward rates (exothermic and dissociation) are computed via QCT calculations, as discussed in detail by Jaffe *et al.* [60] The backward rates are retrieved from the forward ones based on micro-reversibility (*i.e.*, detailed balance) [114]. In particular, the recombination rate coefficients are computed as:

$$k^R_i = k^D_i \frac{1}{K^{Eq}_i(T)} = k^D_i \frac{Q_{O_2}(T)\ Q^t_{O_2}(T)}{[Q_O\ Q^t_O(T)]^2}, \quad \forall\ i \in \mathcal{I}, \tag{4.3}$$

where the equilibrium constant, $K^{Eq}_i(T)$, depends on the partition functions of O (translational, $Q^t_O(T)$, and internal, Q_O) and O_2 (translational, $Q^t_{O_2}(T)$, and internal, $Q_{O_2}(T)$). In particular, $Q_O = 9$ and $Q_{O_2}(T)$ is given by the expression:

$$Q_{O_2}(T) = \sum_{i\in\mathcal{I}} \xi_i(T) = \sum_{i\in\mathcal{I}} g^e\ g_i\ \exp\Big(-\frac{\epsilon_i}{k_B T}\Big), \tag{4.4}$$

with $\xi_i(T)$ being the i-th level contribution, k_B the Boltzmann's constant, and $g^e = 3$, the degeneracy of the O_2 electronic ground-state.

In this part, 50 000 collisions per initial O_2 rovibrational state are simulated for each of the translational temperatures in the following list: $T_{Tran} = \big\{1\,500,\ 2\,500,\ 5\,000,\ 6\,000,\ 8\,000,\ 10\,000,\ 12\,000,\ 14\,000,\ 15\,000,\ 20\,000\big\}$ K. The trajectory calculations are performed with CoarseAIR [153], the

in-house Fortran 2008 modernization of the original VVTC code developed at NASA Ames by D. Schwenke [135]. The resulting dissociation rate coefficients are corrected by a factor 16/3 [110]. The 9 PESs employed in this work only characterize the interaction of the electronic ground-state of O_2 with the O atom. The oxygen molecule, however, presents low-lying electronic excited states, and the dissociation can take place from one of these levels relatively easily. As a consequence, the multi-surface correction factor is included in this work with the aim of accounting for the effect of the excited electronic states. An in-depth discussion about the reasoning behind the 16/3 coefficient can be found in Esposito *et al.* [37] and in Andrienko *et al.* [4]. The overall set of *ab initio* rovibrational-resolved rate coefficients form the first database obtained using all nine PESs for the ground-state of O_2 colliding with O.

4.3 Isothermal and Isochoric Chemical Reactor Model

In this part, we study the nonequilibrium relaxation of oxygen in an isothermal chemical reactor [4, 69, 114]. In particular, two 0-D cases are analyzed: a purely-dissociative case and a more exhaustive one where both dissociation and excitation are taken into account. In all the simulations presented, unless explicitly indicated otherwise, the full rovibrational kinetics are studied. Three rovibrational databases are constructed using the excitation rate coefficients (Eqs. 2.19-2.20) obtained by QCT scattering calculations and the dissociation rate coefficients from three different data sets:

1. The *ab initio data set*, as obtained from the QCT calculations;
2. The *vibrationally-averaged data set*, reconstructed from the vibrational-specific (VS) reduced order model;
3. The *CB-averaged data set*, reconstructed from the novel centrifugal-barrier-based (CB) reduced order model.

Hereafter, the calculations performed by employing the data set (a) will be referred to as fully-StS. In all the scenarios analyzed, molecules are allowed to interact via collisional processes with atoms, but not with each other. The heat bath simulations are performed with the KONIG code and the PLATO thermochemical library [102, 103, 107, 108]. Before moving to the description of the reduced order strategies, the two heat bath scenarios are briefly discussed below.

4.3.1 Analysis of Dissociation in a 0-D Heat Bath: Pure Dissociation

The numerical simulations presented are carried out in a 0-D isochoric and isothermal chemical reactor, where density and temperature are set as initial conditions. At the beginning of the simulation, the O_2 molecules are distributed according to a Maxwell-Boltzmann function at the translational temperature, $T^0_{Int} = T_{Tran}$, and their concentration is out of equilibrium. The most accurate analysis of the resulting dynamics is performed by integrating in time the RVS master equation:

$$\begin{cases} \dfrac{dn_i}{dt} = -k_i^D n_i n_O + k_i^R n_O^3, \quad \forall\ i \in \mathcal{I}, \\ \dfrac{dn_O}{dt} = \displaystyle\sum_{i\in\mathcal{I}} k_i^D n_i n_O - \sum_{i\in\mathcal{I}} k_i^R n_O^3, \end{cases} \tag{4.5}$$

where n_i and n_O represent the number densities of the i-th rovibrational level and atomic oxygen, respectively.

4.3.2 Analysis of Dissociation in a 0-D Heat Bath: Excitation and Dissociation

In this case, both dissociation-recombination and energy transfer processes are considered. A Boltzmann distribution at a low internal temperature is prescribed as an initial condition for the O_2 molecules (*i.e.*, 'cold' oxygen mixture). At the beginning of the simulation, the translational temperature is instantaneously artificially raised to a higher value, $T_{Tran} \gg T^0_{Int}$, and thereafter kept constant. The inclusion of rovibrational excitation of the O_2 molecules requires modification to the system of equations presented above. The resulting RVS master equation for this analysis has been already presented as Eq. 2.22.

4.4 Model Reduction for Dissociation Mechanisms

This section summarizes the details of the methodology used to construct the *vibrationally-averaged* and the *CB-averaged* dissociation data sets used for solving the RVS master equation. In particular, the goal of this part is to put forward the differences and similarities between the proposed CG approaches and VS model published in the literature. The hybrid technique used in our analysis is presented by starting from the methodology of model reduction discussed in a number of

publications [85, 102, 129].

4.4.1 Reduced Order Approach to Master Equation

The general model-reduction approach for the nonequilibrium kinetics used in the remaining part of this thesis has been extensively discussed by Liu *et al.* [85] and briefly mentioned in Sec.2.3.4. However, the two distinct grouping strategies adopted in this part are different from the one by Sahai *et al.* [129] introduced above. The goal of the present study is the analysis and coarse-graining of the dissociation mechanisms, rather than the excitation ones. For this reason, the two following grouping strategies are the well known VS model [21, 26] and the new CB model. The procedure for simulating the nonequilibrium kinetics is the same for both strategies and is briefly summarized below for a purely-dissociative case:

1. *Grouping the Rovibrational Levels*: The levels are first separated in $\mathcal{N}_{\mathcal{G}}$ groups (also referred to as bins). The set of states contained in the P-th bin is indicated by $\mathcal{I}_P$.

2. *Boltzmann-Averaging the RVS Rate Coefficients*: The group-specific dissociation rate coefficient, K_P^D, is obtained from the RVS ones as a weighted average based on the Boltzmann distribution function over $\mathcal{I}_P$:

$$K_P^D(T_{Tran}, T_P) = \sum_{i \in \mathcal{I}_P} k_i^D(T_{Tran}) f_P^i(T_P), \tag{4.6}$$

   where the the Boltzmann distribution is defined as:

$$f_P^i(T_P) = \frac{\xi_i(T_P)}{\sum_{i \in \mathcal{I}_P} \xi_i(T_P)} = \frac{\xi_i(T_P)}{Q_P(T_P)}, \quad \forall i \in \mathcal{I}_P, \tag{4.7}$$

   where T_P and $Q_P(T_P)$ indicate the internal temperature and partition function of the P-th group, respectively.

3. *Solving the CG Master Equation*: The population of the P-th group, n_P, is obtained by solving the coarse-grained master equation, derived by computing the moments of the corresponding

RVS StS equation. For a purely-dissociative case, this is given by:

$$\begin{cases} \dfrac{dn_P}{dt} = -K_P^D n_P n_O + K_P^R n_O^3, \quad \forall\ \mathcal{I}_\mathcal{P} \in \mathcal{I}, \\ \dfrac{dn_O}{dt} = \displaystyle\sum_{\mathcal{I}_\mathcal{P} \in \mathcal{I}} K_P^D n_P n_O - \sum_{\mathcal{I}_\mathcal{P} \in \mathcal{I}} K_P^R n_O^3. \end{cases} \tag{4.8}$$

The distinction between the VS and CB models is limited to step 1 in the procedure, that lumps the rovibrational states into groups. Thus, different low dimensional representations of the distribution function are used for the two strategies:

- *VS Grouping:* All the levels characterized by the same vibrational quantum number are grouped together:

$$\mathcal{I}_P = \big\{(v, J) \mid v = (P-1)\big\}, \quad \forall\ \mathcal{I}_P \subset \mathcal{I}. \tag{4.9}$$

- *CB Grouping:* The levels are clustered based on the energy-deficit from the centrifugal barrier. The strategy requires the construction of a grid over ϵ_i^D:

$$\hat{E}_G^D = \left[1 - \frac{G}{N_G}\right]^\eta \epsilon_1^D, \quad \forall\ G \in \big\{0, ..., N_G\big\}, \tag{4.10}$$

where $\epsilon_1^D = \text{argmax}_i\left(\epsilon_i^D\right)$, and η is a free parameter. Each level is then assigned to the group that respects the following constraint:

$$\hat{E}_{P-1}^D \leq \epsilon_i^D < \hat{E}_P^D. \tag{4.11}$$

The value of $\eta = 1$ chosen in this work yields a uniform grid. Differently, $\eta > 1$ would correspond to a grid refined in the proximity of the centrifugal barrier.

In order for the comparison with the VS approach to be consistent, $T_P = T_{Tran}$ (*i.e.*, the group temperature is fixed and equal to the translation one) is also imposed for the CB model. It should be noticed, however, that this does not correspond to constraining $T_{Rot} = T_{Tran}$.

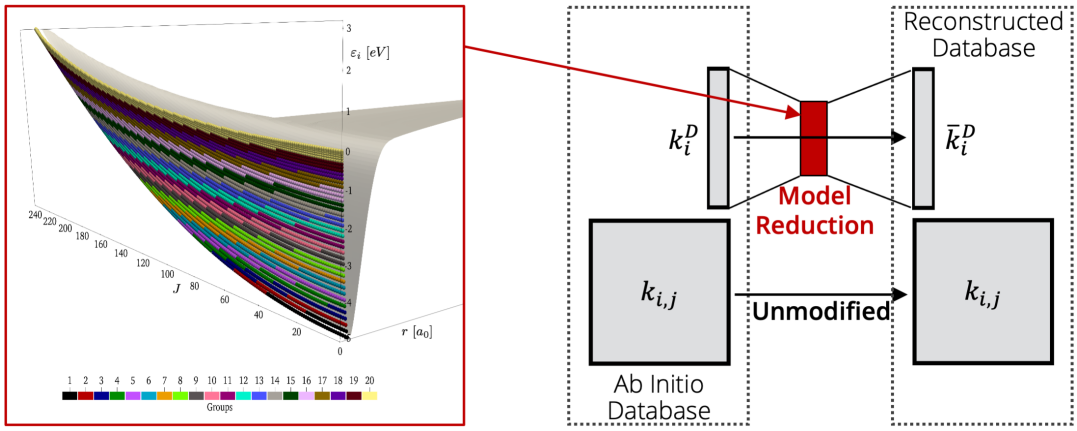


Figure 4.2: Schematics of the procedure for obtaining the data sets of RVS StS rate coefficients employed in the hybrid approach to the master equation. The final database combines the unmodified *ab initio* excitation rates and the dissociation rates encoded trough model reductions and decoded trough reconstructing averages.

Finally, the group-specific averaged distance from the centrifugal barrier is defined as:

$$E_P^D = \sum_{i \in \mathcal{I}_P} \epsilon_i^D f_P^i(T). \tag{4.12}$$

4.4.2 Hybrid Approach to ME: RVS StS Excitation Coupled to CG Dissociation

In order to quantify the impact that the model reduction for the dissociation kinetics has on the overall nonequilibrium relaxation, no additional inadequacy must be introduced while modeling the energy transfers. For this reason, while still coarse-graining the dissociation kinetics, we employ an exact treatment of the excitation processes by using their correspondent (unmodified) *ab initio* RVS StS rate coefficients. The hybrid procedure that we follow, briefly schematized in Fig. 4.2, is outlined hereafter:

1. *Grouping the Rovibrational Levels*: As discussed in the previous section.

2. *Boltzmann-Averaging the RVS Rate Coefficients*: As discussed in the previous section.

3. *Reconstruction of the Rovibrational Rate Coefficients*: The RVS rate coefficients are reconstructed from the group-specific values computed in the previous step.

4. *Solution of the RVS StS Master Equation.*

The first two steps have been already discussed in the previous section, and are basically unchanged for the dissociation processes. On the other hand, the reconstruction at step 3 needs to be consistent with the Boltzmann-averaging performed in step 2. For this reason, the reconstructed (rovibrational-specific) dissociation rate coefficients are assigned as follows:

$$\bar{k}_i^D = K_P^D, \qquad \forall\, i \in \mathcal{I}_P,\ \forall\, \mathcal{I}_P \subset \mathcal{I}. \tag{4.13}$$

Finally, instead of solving the coarse-grained master equations (*e.g.*, Eq. 4.8), we solve the RVS formulations (Eq.s 2.22 and 4.5). In the following, this approach will be referred to as grouped-reconstructed. It is important to notice that, for the purely-dissociative heat bath presented in Sec. 4.3.1, the coarse-grained formulation and the grouped-reconstructed one are mathematically equivalent. as shown in the next section.

4.5 Equivalence of Coarse-Grained and Grouped-Reconstructed Formulations

This section focuses on the purely-dissociative heat bath scenario presented in Sec. 4.3.1. Its objective is proving that the solution of the ro-vibrational-specific master equation (*i.e.*, Eq. 4.5) under the constraint of Eq. 4.13 is mathematically equivalent to the solution of the coarse-grained master equation (*i.e.*, Eq. 4.8).

We start with selecting the P-th bin and multiplying the equation characterizing its population (first equation in the system of Eq. 4.8) by the Boltzmann distribution function for its levels, f_P^i. By relying on Eq. 4.7, we can write:

$$\frac{dn_P \frac{\xi_i}{Q_P}}{dt} = -K_P^D n_P n_O \frac{\xi_i}{Q_P} + K_P^R n_O^3 \frac{\xi_i}{Q_P}, \quad \forall i \in \mathcal{I}_P. \tag{4.14}$$

We should recall that a state is in local equilibrium with all the other levels in the same group. Therefore:

$$\frac{n_i}{n_p} = f_P^i = \frac{\xi_i}{Q_P}, \quad \forall i \in \mathcal{I}_P. \tag{4.15}$$

By substitution, it is then possible to transform n_P in n_i:

$$\frac{dn_i}{dt} = -K_P^D n_i n_O + K_P^R \frac{\xi_i}{Q_P} n_O^3, \quad \forall i \in \mathcal{I}_P. \tag{4.16}$$

By rewriting the group-specific recombination rate coefficient in terms of the dissociation one (similarly to in Eq. 4.3 for the state-specific recombination rate), Eq. 4.16 becomes:

$$\frac{dn_i}{dt} = -K_P^D n_i n_O + K_P^D \frac{\xi_i Q_{O_2}^t}{[g_O Q_O^t]^2} n_O^3, \quad \forall i \in \mathcal{I}_P. \tag{4.17}$$

If Eq. 4.13 is now employed, we obtain:

$$\frac{dn_i}{dt} = -\bar{k}_i^D n_i n_O + \bar{k}_i^D \frac{\xi_i Q_{O_2}^t}{[g_O Q_O^t]^2} n_O^3, \quad \forall i \in \mathcal{I}_P. \tag{4.18}$$

By invoking micro-reversibility, once again, the resulting equation for the evolution of the levels in the P-th group is:

$$\frac{dn_i}{dt} = -\bar{k}_i^D n_i n_O + \bar{k}_i^R n_O^3, \quad \forall i \in \mathcal{I}_P. \tag{4.19}$$

When the same process is repeated for all the $\mathcal{I}_P \subset \mathcal{I}$, it is found that the first of the euations in the system of Eq. 4.5 exactly corresponds to the first of equations in the system of Eq. 4.8.

By comparing Eq. 4.16 and Eq. 4.19, it should also be noticed that:

$$\bar{k}_i^R = K_P^R \frac{\xi_i}{Q_P}, \quad \forall i \in \mathcal{I}_P. \tag{4.20}$$

By rewriting this last equation as $K_P^R = \sum_{i \in \mathcal{I}_P} \bar{k}_i^R$ and by using the constraint $n_P = \sum_{i \in \mathcal{I}_P} n_i$, the equivalence between the second equations in the systems of Eq. 4.5 and Eq. 4.8 can be proved in

few steps:

$$\begin{aligned}\frac{dn_O}{dt} &= -\sum_{\mathcal{I}_P\subset\mathcal{I}} K_P^D n_P n_O + \sum_{\mathcal{I}_P\subset\mathcal{I}} K_P^R n_O^3 = \\ &= -\sum_{\mathcal{I}_P\subset\mathcal{I}} K_P^D \sum_{i\in\mathcal{I}_P} n_i n_O + \sum_{\mathcal{I}_P\subset\mathcal{I}} \sum_{i\in\mathcal{I}_P} \bar{k}_i^R n_O^3 = \\ &= -\sum_{i\in\mathcal{I}} \bar{k}_i^D n_i n_O + \sum_{i\in\mathcal{I}} \bar{k}_i^R n_O^3.\end{aligned} \tag{4.21}$$

Chapter 5

Applications to the O_2+O System

The hybrid methodology discussed in the previous chapter is now adopted to study the performance of the reduced order models by comparing the data sets of rate coefficients (*ab initio* and reconstructed) and the results of 0-D thermochemical simulations. The observable QoIs for the latter analysis are the composition profiles and the rovibrational distributions at different times. Finally, for validation purposes, the macroscopic QSS [81] rate coefficients are compared to the ones in the literature.

5.1 Analysis of the State-Specific Dissociation Rates

The analysis and comparison of the RVS dissociation rate coefficients obtained via the dimensionality reduction process constitute necessary prerequisites for the understanding of the nonequilibrium kinetics, discussed in later sections. In Figure 5.1, the rovibrational levels are plotted as functions of their rotational quantum numbers and internal energies, and they are colored based on the magnitude of their dissociation rate coefficients. The black curves indicate the states characterized by the same energy deficit from the J-dependent centrifugal barrier, $\epsilon_i^D(J)$. These curves help to visualize the strong dependency of the k_i^D coefficients on $\epsilon_i^D(J)$: the larger the distance from the barrier, the lower the likelihood of dissociation. The translational temperature, T_{Tran}, controls the energy available during the collision and acts as a scaling factor of the overall trend. The comparison of Figure 5.1 and Figure 4.1(b) demonstrates that energy levels with the same vibrational quantum number can have drastically different dissociation rate coefficients, especially for small vs. As an example, at T = 10 000 K, $\bar{k}^D_{(0,0)}$ is four orders of magnitude lower than $\bar{k}^D_{(0,240)}$.

An effective reduced order model should minimize the information lost during the grouping step (*i.e.*, during the encoding phase $k_i^D \rightarrow K_P^D$), so that the reconstruction of the state-specific

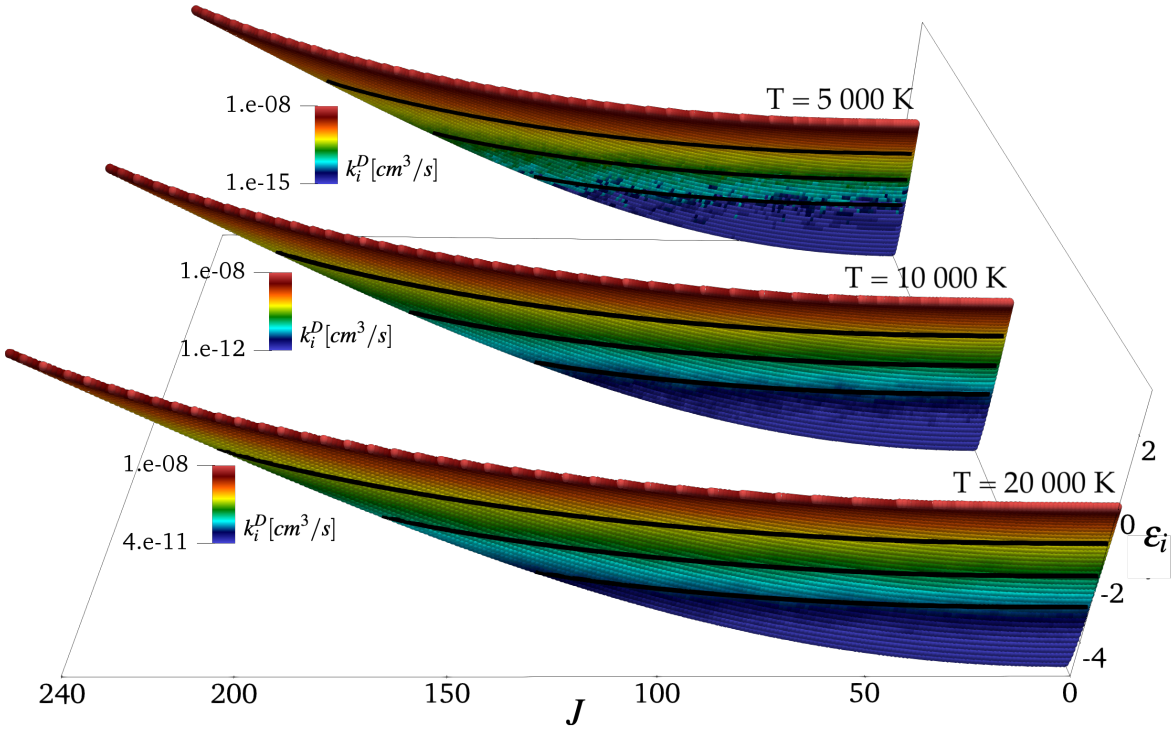


Figure 5.1: O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy (*i.e.*, left side of the 3D diatomic potential in Figure 4.1(b)). The levels are colored based on their *ab initio* RVS dissociation rate coefficients, obtained through QCT calculations at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier.

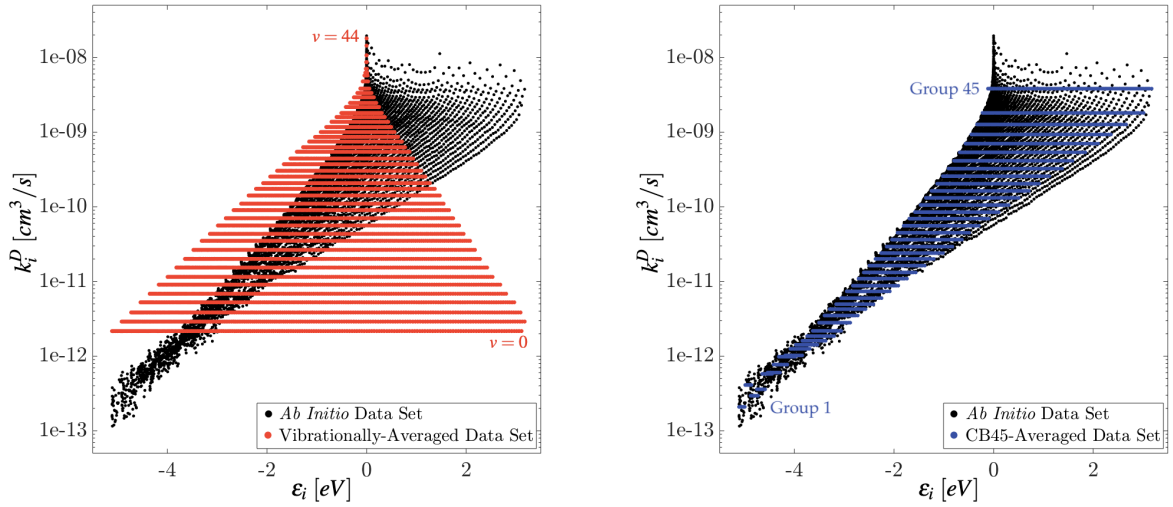


Figure 5.2: RVS dissociation rate coefficients at T = 10 000 K. Black dots: rates from the *ab initio* data-set. Red dots: rates from the vibrationally-averaged data-set. Blue dots: rates from the CB-averaged data-set, based on 45 groups.

rates (*i.e.*, the decoding phase $K_P^D \rightarrow \bar{k}_i^D$) would produce $\bar{k}_i^D$ very similar to k_i^D. In practice, the dimensionality reduction is not perfect, and its accuracy can be assessed by a generic error $e = \sum_{i \in \mathcal{I}} |\bar{k}_i^D - k_i^D|$. The more the reconstructed rate coefficients are similar to their *ab initio* correspondents, the better the model is in reducing the dissociation kinetics.

To better illustrate this important point, Figure 5.2 shows comparisons between the RVS rate coefficients obtained from QCT calculations (in black) and the ones resulting from the dimensionality reductions detailed in Sec. 4.4 (in red and blue). To perform a fair comparison, since the vibrational-specific reduction uses the 45 v quantum numbers as separate groups, we have used 45 bins in the CB model as well. As a result of the Boltzmann-averaging process, the VS rate coefficients form horizontal strands when they are associated with states characterized by the same vibrational quantum number (see Eq. 4.13). For this reason, the rates for the low-lying and high-lying energy levels are strongly mispredicted (*e.g.*, in the case of $k^D_{(0,0)}$, the rate is overestimated by more than one order of magnitude, and for $k^D_{(0,240)}$ it is underestimated by more than three orders of magnitude). On the contrary, the novel strategy provides rate coefficients that are in excellent agreement with the *ab initio* ones. The maximum relative error between the reconstructed and

original rates is reduced to 5.6, compared to 1250.0 of the previous method.

An additional advantage of the CB approach to model reduction is the possibility of fitting the group dissociation rate coefficients with remarkable precision:

$$K_P^D = Q_P^E C_1 \exp\Big(- \frac{C_2 E_P^D + C_3}{k_B T}\Big), \quad \forall\, \mathcal{I}_P \subset \mathcal{I}, \tag{5.1}$$

where $Q_P^E = \sum_{i \in \mathcal{I}_P} \exp\left(-\frac{\epsilon_i^0}{k_B T}\right)$ and $\epsilon_i^0 = \epsilon_i - \epsilon_1$. C_1, C_2, and C_3 identify three tunable parameters, the values of which are reported in Table 5.1. The resulting fitting accuracy is showed in 5.7.

Table 5.1: Parameters Values for Eq. 5.1

	C_1 $[cm^3/s]$	C_2 $[-]$	C_3 $[eV]$
45 Groups	1.231e-11	2.352	-5.035
20 Groups	5.238e-12	2.362	-5.099
15 Groups	3.724e-12	2.364	-5.157
10 Groups	2.141e-12	2.347	-5.298

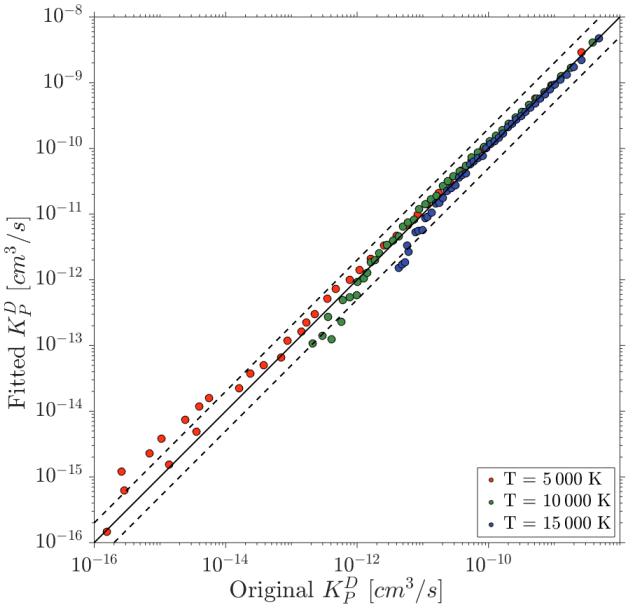
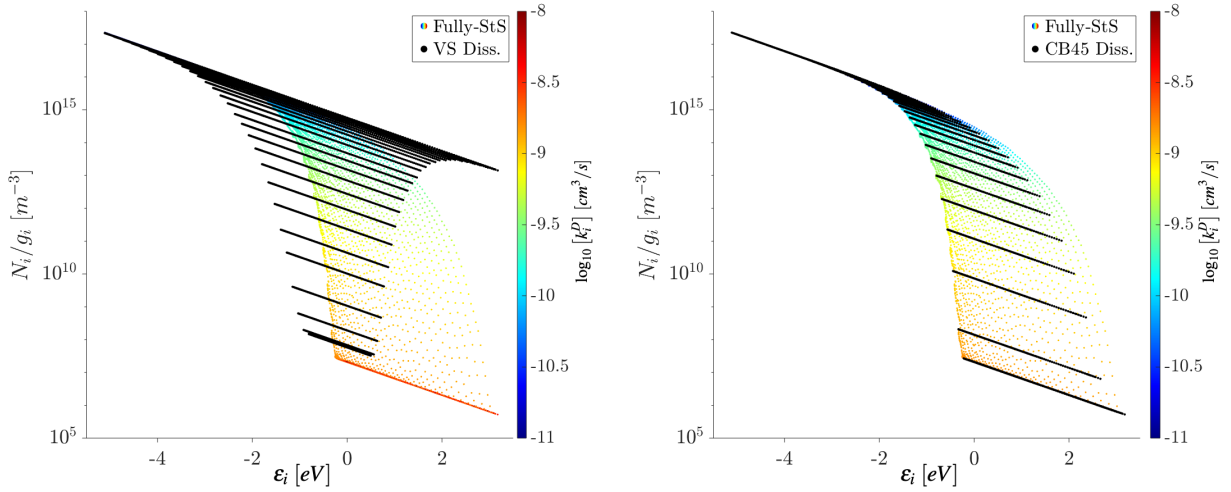


Figure 5.3: Scatter plot reporting the fitting error for the group-specific dissociation rate coefficients at T = 5 000 K (red dots), T = 10 000 K (green dots), and T = 15 000 K (blue dots). The 45 groups are constructed based on the energy-distance from the centrifugal barrier. The oblique external lines represent factor-two fitting errors (*i.e.*, absolute errors equal to 100%).

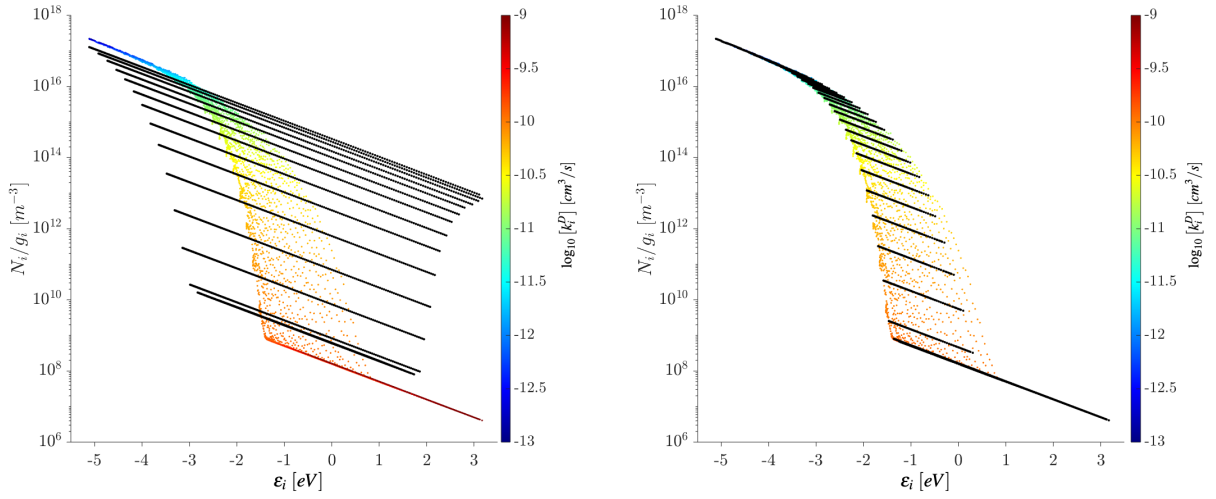
5.2 Analysis of Dissociation in a 0-D Heat Bath: Pure Dissociation

The section proceeds by analyzing the differences between the VS and CB reduced order models in the characterization of the O_2+O dissociation kinetics. To this aim, we simulate the nonequilibrium processes taking place in an isochoric and isothermal chemical reactor. In this analysis, we eliminate the mechanisms leading to the excitation of the molecule's internal degrees of freedom. Only dissociation is included. At the beginning of the numerical simulation, the gas number density is $3.33 \times 10^{23}\ m^{-3}$ and the mixture is made up of 95% O_2 and 5% of O. Moreover, the molecules are initially populated according to a Maxwell-Boltzmann distribution at $T_{Tran} = T^0_{Int} = 10\,000$ K. The translational temperature is kept constant throughout the chemical relaxation. The study of the kinetics is carried out by solving the master equation for the entire rovibrational energy ladder. Three different sets of rovibrational dissociation rate coefficients are used in this comparison: i) the *ab initio* data-set; ii) the vibrationally-averaged data-set; iii) the CB-averaged data-set. In the last



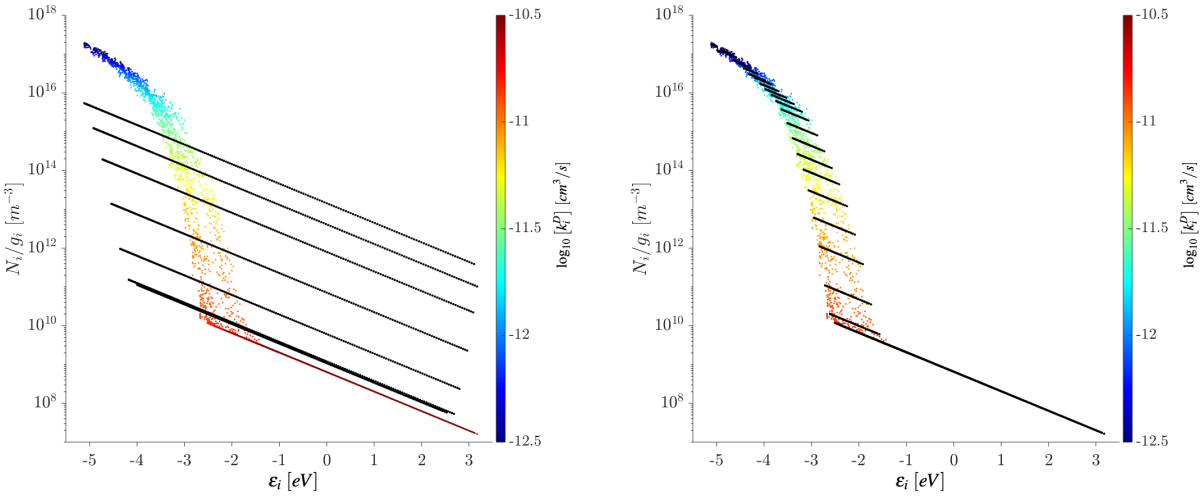
(a) (b)

Figure 5.4: Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 7 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets.



(a) (b)

Figure 5.5: Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 30 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets.



(a) (b)

Figure 5.6: Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K in which only dissociative collisions are allowed. The time-snapshots are taken at t = 100 μs. Colored dots: dissociation is computed based on the *ab initio* data set of rate coefficients, and levels are colored based on their correspondent rates. Black dots: results based on the vibrationally-averaged (a) and on the 45-groups CB-averaged (b) data sets.

two cases, the reconstructed coefficients are obtained by assigning the value of the group-specific rate to all the states within the bin, which is consistent with Eq. 4.13 and Figure 5.2.

Figures 5.4, 5.5, and 5.6 report the rovibrational distribution function for the three models. The symbols in color represent the population at t = 7 μs, t = 30 μs, and t = 100 μs, respectively, for the case employing the *ab initio* k_i^D. The outcomes of the averaged models are presented in black and are overlaid on the reference solution. In all the simulations, the levels showing the same values of dissociation rate coefficients are in equilibrium with each other at $T = T_{Tran}$. This behavior can be mathematically explained by starting from Eq. 4.5. Assuming that n_O is known at each time instant, the equation has the analytical solution:

$$n_i(t) = \frac{1}{K_i^{Eq}(T)} n_O^2(t) + C \exp\Big(-k_i^D n_O^2(t) t\Big), \tag{5.2}$$

where n_O indicates the initial number of O atoms, and C_1 is given by:

$$C = \left(\frac{1 - n_O(0)}{n_O(0)} \frac{g_i}{Q_{O_2}(T)} \exp\Big(-\frac{\epsilon_i}{k_B T}\Big) - \frac{1}{K_i^{Eq}(T)} \right) n_O(0). \tag{5.3}$$

Eq. 5.2 describes a first-order relaxation with time constant:

$$\tau_i^D = \frac{1}{k_i^D n_O^2(t)}. \tag{5.4}$$

At the beginning of the simulation, a Boltzmann distribution is imposed for all the rovibrational levels. From the combination of this assumption and Eq. 5.4, it follows that, if $k_i^D = k_j^D$, then the local equilibrium between the states i and j is preserved for the entire dissociation process.

It can be noticed that states characterized by internal energies close to the centrifugal barrier (*i.e.*, the ones with the highest k_i^D) are the ones to be depleted first, even if they lie at the vibrational ground-state. For this reason, the CB approach to model reduction significantly outperforms the VS one in reproducing the features of the *ab initio* mechanisms, even for these simplified heat bath conditions. It is worth highlighting that the inadequacy of the VS model in predicting the dissociation kinetics is not limited to the quasi-bound levels. The consequence on the description of macroscopic quantities such as the O_2 mole fraction can be seen in Figure 5.7. It is worth highlighting that this improvement in accuracy is obtained by solving only 45 equations compared to the original 6 116 of the fully-StS model, with a reduction of the computational cost by more than two orders of magnitude.

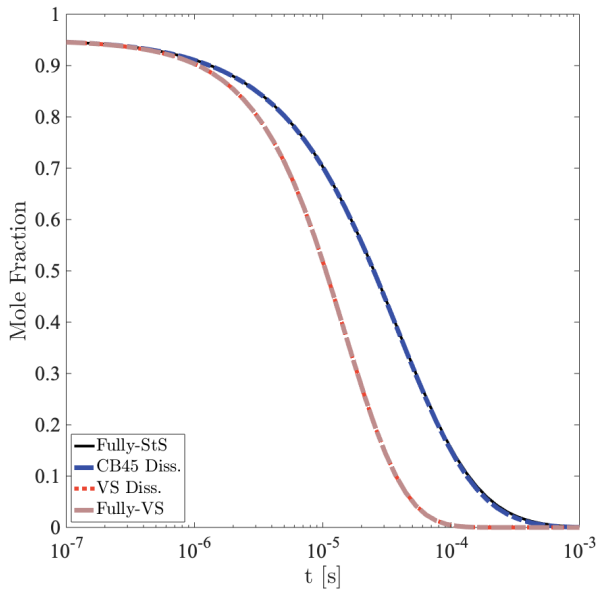


Figure 5.7: Evolution of O_2 mole fraction in the 0-D simulation in which only dissociative collisions are allowed. The conditions of the isothermal and isochoric heat bath are $T^0_{Int} = 10\,000$ K, $n^0_d = 3.33 \times 10^{23}$ m^{-3}, $[O_2]^0 = 95\%$, and $T_{Tran} = 10\,000$ K. Black and blue lines overlap. Red and pink lines overlap. Black unbroken line: dissociation is computed based on the *ab initio* data set of rate coefficients. Red dash-dotted line: dissociation is computed based on the vibrationally-averaged data set. Blue dashed line: dissociation is computed based on the 45-groups CB-averaged data set. Pink unbroken line: from the fully-VS model.

5.3 Analysis of Dissociation in a 0-D Heat Bath: Excitation and Dissociation

In this section, the analysis of the nonequilibrium relaxation is extended to include the collision processes leading to internal energy exchange. While the translational temperature of the heat bath is still fixed at $T_{Tran} = 10\,000$ K, the relaxation starts with 'cold' molecules. Thus, we initialize the internal population of O_2 using a Boltzmann distribution at $T^0_{Int} = 300$ K. The thermal and chemical nonequilibrium kinetics are studied by solving the RVS master equation in Eq. 2.22. The main goal of this analysis is to investigate how the internal excitation of the molecules affects the dissociation mechanisms observed so far. In this work, we are not interested in studying the physics of the energy exchanges, nor in reducing their dimensionality. Therefore, the values of the excitation rate coefficients, k_{ij}, remain unchanged. On the contrary, the three sets of k^D_i are separately employed. It is important to mention that, while we are still solving Eq. 2.22 for each

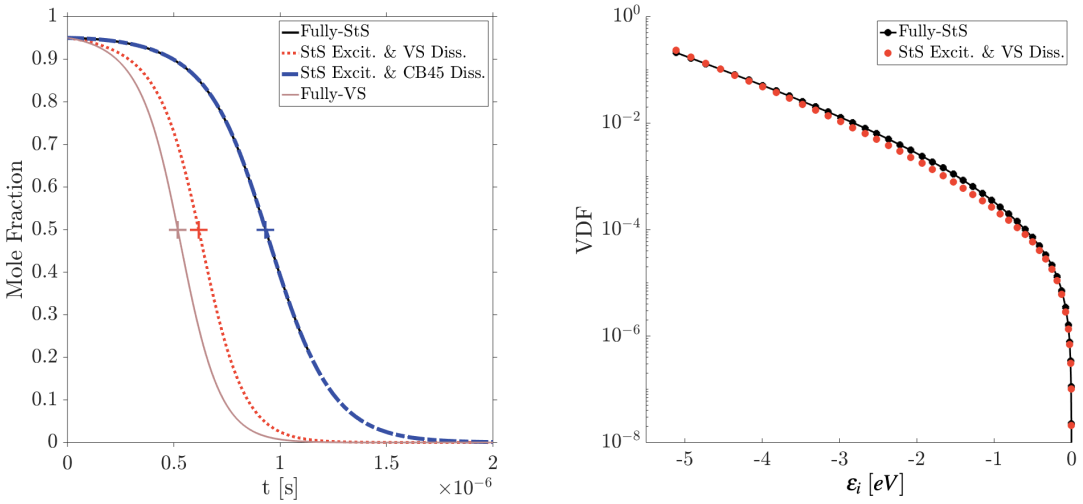


Figure 5.8: (a): Evolution of O_2 mole fraction in a heat bath at T = 10 000 K. Black and blue lines overlap. Black unbroken line: fully-StS calculations. Red dotted line: RVS StS excitation coupled to VS dissociation. Blue dashed line: RVS StS excitation coupled to CB dissociation with 45 groups. Pink unbroken line: fully-VS calculations. Crosses: time instants at which $[O_2]$ = 50%. (b): Vibrational distribution functions at the time instants at which $[O_2]$ = 50%. Black dots: fully-StS calculations. Red dots: RVS StS excitation coupled to VS dissociation.

rovibrational level, the adoption of the reconstructed rate coefficients is equivalent to simulating the dissociation processes as if they were grouped according to the two strategies outlined above (see B).

The time evolution of the resulting O_2 mole fractions is presented in Figure 5.8(a). The dissociation predicted by the VS model is much too rapid, about 33% faster than the one predicted by the fully-StS kinetics. On the contrary, the CB model introduces negligible inaccuracies in the prediction of the chemical composition. For the sake of completeness, we also report the result of the fully-VS model [21, 26], which groups also the energy exchange processes based on vibrational quantum numbers. The comparisons of the pink and red curves with the black one demonstrate that the main shortcoming in the VS model is its inability to predict dissociation correctly. This result is consistent with what was found by Macdonald *et al.* for the N_2+N_2 system [90]. The analysis of the vibrational distribution function (VDF), reported in Figure 5.8(b), shows excellent agreement between the predictions provided by the VS model and the fully-STS result. Clearly, the ability to capture the VDF is not sufficient to correctly predict the dissociation. A countless

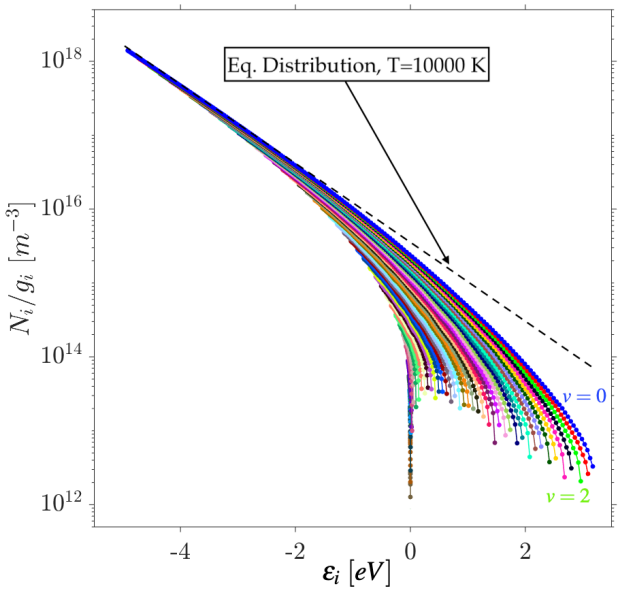


Figure 5.9: Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K, computed by means of fully-StS calculations. The time-snapshot is taken during QSS, at the instant at which $[O_2] = 50\%$ (*i.e.*, black/blue crosses in Figure 5.8(a)). Levels characterized by the same vibrational quantum numbers are connected through segments of equal colors.

number of models focus uniquely on the modeling of the vibrational nonequilibrium, completely disregarding the nonequilibrium in the rotational degrees of freedom. The result presented constitutes a clear proof that exclusively relying on the VDF to construct nonequilibrium models is a mistake, and rotation must be correctly accounted for [89].

The fundamental assumption of rotational equilibrium (*i.e.*, the rotational levels follow a Boltzmann distribution) adopted by the VS model is unable to describe the internal distributions of O_2 molecules during dissociation. Figure 5.9 shows the rovibrational populations computed by the fully-StS simulation in the midst of QSS, when about half of the molecules have dissociated. In the figure, the levels have been colored as functions of their vibrational quantum numbers. The rotational population distribution of the vibrational states exhibits a strong deviation from the equilibrium one. To correctly account for its curvature, higher-order moments must be used [137]. The RVS distribution at QSS shares many similarities with the one observed in Figures 5.4, 5.5, and 5.6, where only dissociation processes were included. The establishment of QSS results from temporary 'equilibrium' between the incoming flux of molecules, excited from the highly populated states, and the outgoing flux, caused by the direct dissociation events. Noticeable exceptions are

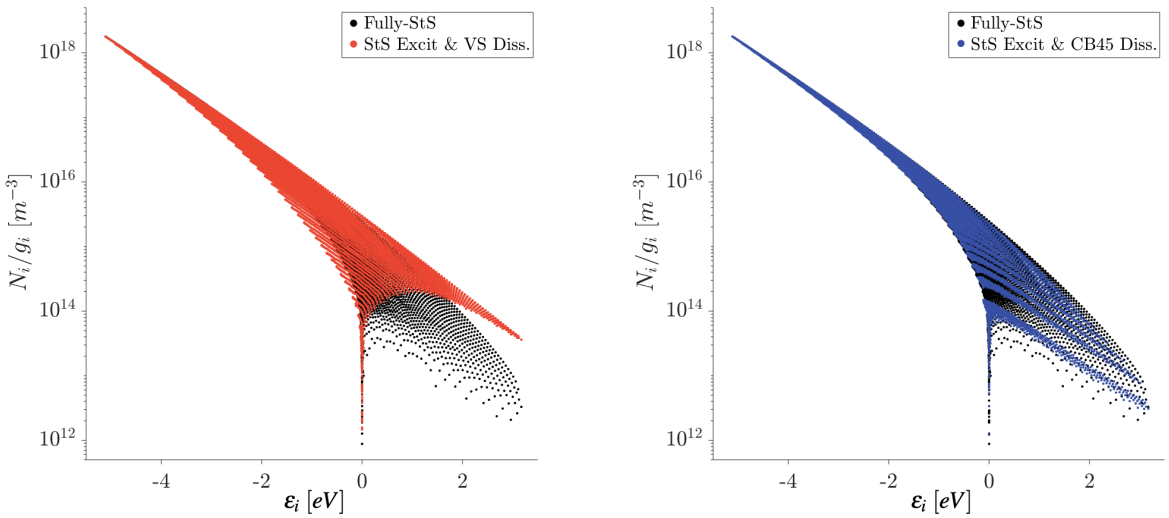


Figure 5.10: Normalized population distribution of the rovibrational levels for a heat bath at T = 10 000 K. The time-snapshots are taken during QSS, at the instants at which $[O_2] = 50\%$ (*i.e.*, crosses in Figure 5.8(a)). Black dots: fully-StS calculations. Red dots: RVS StS excitation coupled to VS dissociation. Blue dots: RVS StS excitation coupled to CB dissociation with 45 groups.

the few states characterized by high v and internal energy close to 0 eV. The larger the vibrational quantum number, the smaller the energy-spacing between the levels. Therefore, while jumps between high-v states are very likely, the exchanges with levels close to the ground-state are improbable. For this reason, high-v states tend to be isolated in a single cluster and to be mostly governed by dissociation, given their proximity to the centrifugal barrier [29].

The use of a Maxwell-Boltzmann distribution to describe all the rotational states characterized by the same v leads to a severe over-prediction of the populations of the quasi-bound levels (Figure 5.10(a)). The thermo-chemical relaxation is significantly accelerated, as a consequence. On the contrary, grouping the levels based on their energy-deficit from the centrifugal barrier permits the imposition of the condition of local equilibrium only to the states showing similar k_i^D, thus allowing for the QSS distribution to be correctly captured (Figure 5.10).

The heat bath simulations are then repeated for a range of translational temperatures between 1 500 and 20 000 K. Figure 5.11 reports the relative differences between the times required by the models to achieve the value of 50% in O_2 mole fraction. Different numbers of groups are tested for the CB approach: 10, 15, 20, and 45. In general, the novel strategy outperforms the VS model

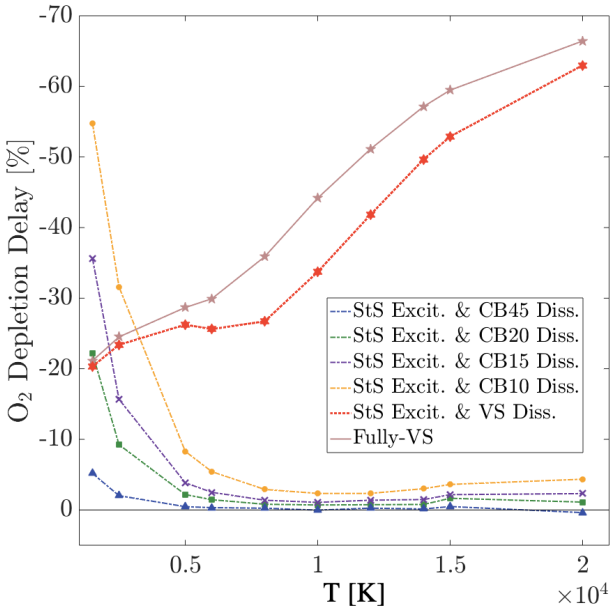


Figure 5.11: Delays of O_2 depletion compared to the fully-StS simulations. Red dotted line: RVS StS excitation coupled to VS dissociation. Pink unbroken line: excitation and dissociation are both computed vibrationally-specific. All the other lines are obtained by computing excitation trough RVS StS and dissociation through the CB model with 45 (blue dash-dotted line), 20 (green dash-dotted line), 15 (purple dash-dotted line) and 10 (yellow dash-dotted line) groups.

over the entire range of conditions and for all the numbers of groups. Given the same extent of coarsening, with the number of bins fixed to 45, the CB model performs five times better than the VS one at 1 500 K and sixty times better at 20 000 K. For T_{Tran} larger than 5 000 K, even with just 10 groups, the error of the CB approach is at least three times smaller than the one characterizing the VS model. With 20 bins, the accuracy in the prediction of the mole fraction at $T_{Tran} =$ 20 000K is improved by 2000%. At very low temperatures, where molecular dissociation is barely occurring, the fidelity of the CB reduction deteriorates. This behavior is explicable by analyzing Figure 5.1, once again. When T_{Tran} is relatively small, the levels close to the centrifugal barrier are characterized by dissociation rate coefficients spanning many orders of magnitude. When a small number of groups is used, these states end up clustered in the same bin, and their rates are averaged into a single value. Therefore, a larger number of groups is necessary to guarantee the same accuracy observed at high temperatures. This number could be decreased further by selecting $\eta > 1$ in Eq. 4.11 and by consequently clustering the proximity of the centrifugal barrier in more bins.

In order to gather more insight into the dissociation dynamics, we analyze the dissociation-vibrational, dissociation-rotational, and dissociation-internal energy transfer. While Figure 5.12 shows the evolution of these energy rate constants, C^{DV}, C^{DR}, and C^{DI}, respectively, in a heat bath at $T_{Tran} = 10\,000$ K, Figure 5.13 presents their QSS values as function of the translational temperature. For temperatures higher than 18 000 K, the vibrational mode is not anymore the preferential channel for energy removal. It should be noticed that the same phenomenon was already inferred for the N_2+N system, but expected only at $T_{Tran} > 50\,000$ K [114]. For the O_2+O system, the contribution of rotation is proven to be not negligible even at relatively low temperatures, and inadequacies in its characterization necessarily result in a poor estimate of the overall amount of energy stored in the molecules during their depletion. Once again, the CB order model outperforms the VS one even when constructed with significantly fewer groups.

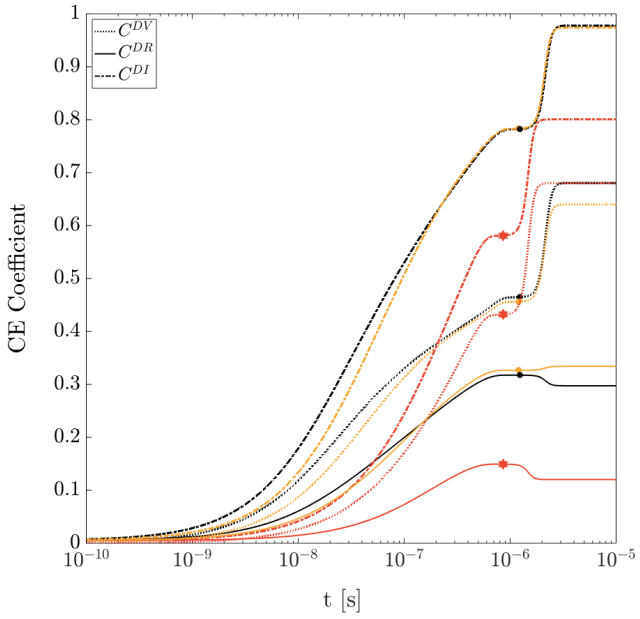


Figure 5.12: Evolution of the energy rate constants for dissociation-vibrational (unbroken lines), dissociation-rotational (dotted lines), and dissociation-internal (dashed lines) energy transfer in a heat bath at $T_{Tran} = 10\,000$ K. Black lines: from fully-StS calculations. Red lines: RVS StS excitation coupled to VS dissociation. Yellow lines: RVS StS excitation coupled to CB dissociation with 10 groups.

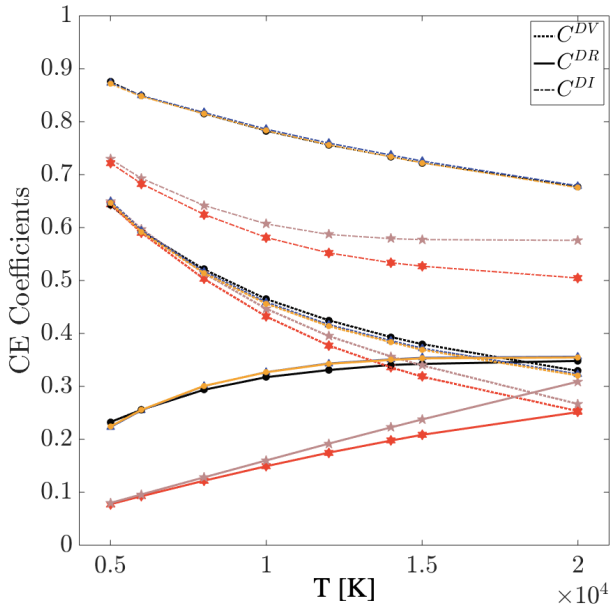
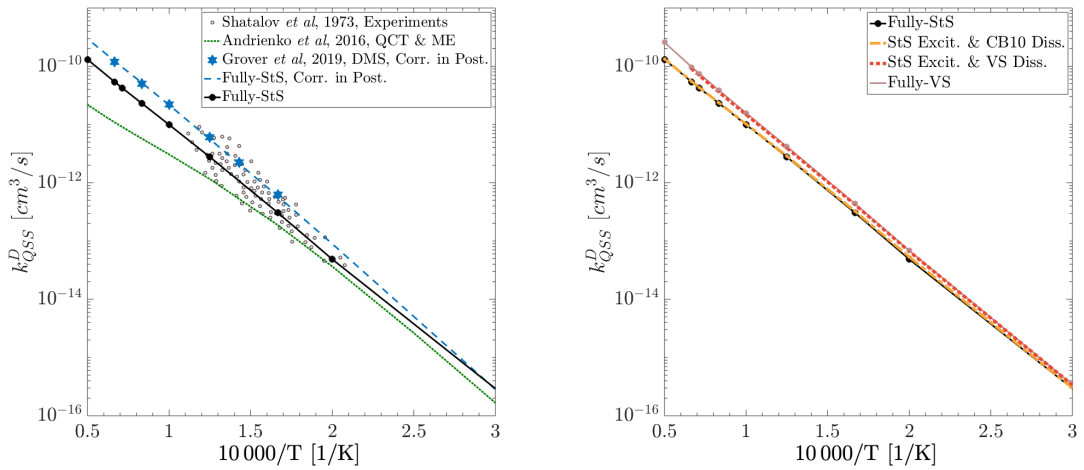


Figure 5.13: Energy rate constants for dissociation-vibrational (unbroken lines), dissociation-rotational (dotted lines), and dissociation-internal (dashed lines) energy transfer at QSS. Black lines: fully-StS calculations. Red lines: RVS StS excitation coupled to VS dissociation. Pink lines: fully-VS calculations. Yellow lines: StS excitation coupled to CB dissociation with 10 groups.

5.4 Comparison with Literature

Before closing the results section, comparisons against available experimental data and other theoretical calculations are provided. Figure 5.14(a) compares the QSS dissociation rates obtained through the fully-StS simulations to the ones already in the literature. The novel set of rate coefficients is in good agreement with the experiments by Shatalov *et al.* [139], and it differs by a factor up to five from the RVS calculations by Andrienko *et al.* [4], which used the $^1A'$ PES by Varandas *et al.* [149]. The discrepancies with the work by Grover *et al.* [44] are generated by inconsistent use of the correction factor 16/3 to account for the contribution of the electronic levels. Since it does not rely on rate coefficients, the DMS approach permits to apply such factor only in the post-processing phase. Given the non-linearities in Eq. 2.22, such *an posteriori* remedy is not equivalent to rectifying the dissociation probabilities at run time, and it leads to the error shown in the figure. In order to compare against Grover *et al.*[44], we have repeated the calculations and have applied the correction factor in the same way as them (as we have done for obtaining Figure 1.2).



(a) (b)

Figure 5.14: QSS dissociation rate coefficients for the O_2 molecule. Small pink circles: Shatalov experiments [139]. Green dotted line: Andrienko *et al.* [4]. Blue hexagram symbols: DMS simulations by Grover *et al.* [44]. Blue dashed line: present work, from fully-StS calculations with 16/3 correction factor applied in post-processing. Black continuous lines: present work, from fully-StS calculations. Yellow dash-dotted lines: present work, RVS StS excitation coupled to CB dissociation with 10 groups. Red dotted lines: present work, from RVS StS excitation coupled to VS dissociation. Pink unbroken lines: present work, from fully-VS calculations.

The QSS rate coefficients predicted by the CB model result are accurate even when the description of the dissociation processes relies on only ten distinct $\bar{k}_i^D$ (Figure 5.14(b)). Differently, the VS model computes fully-reliable rate coefficients only up to 4 000 K. At higher temperatures, the VS values gradually diverge from the rovibrational-StS ones, and the error reaches a factor of two at 20 000 K. Nevertheless, such disagreement is not as remarkable as the factor ten found in previous studies [4].

As a final remark, Figure 5.15 shows the percentage of dissociation taking place during QSS, as predicted by the fully-StS study relying on the *ab initio* database. Three different initial number densities have been tested: 3.33×10^{22}, 3.33×10^{23}, and 3.33×10^{24} m^{-3}. The heat bath simulations at 1 500 and 2 500 K resulted in QSS dissociation rate coefficients being higher than their correspondent equilibrium values, as a consequence of the recombination processes. For this reason, in Figure 5.15, as well as in Fig.B.7, we did not report temperatures lower than 5 000 K. After reaching a maximum between 5 000 and 10 000 K, depending on the initial number density imposed, such a quantity rapidly drops as the temperature increases. Already at 14 000 K, more than half of

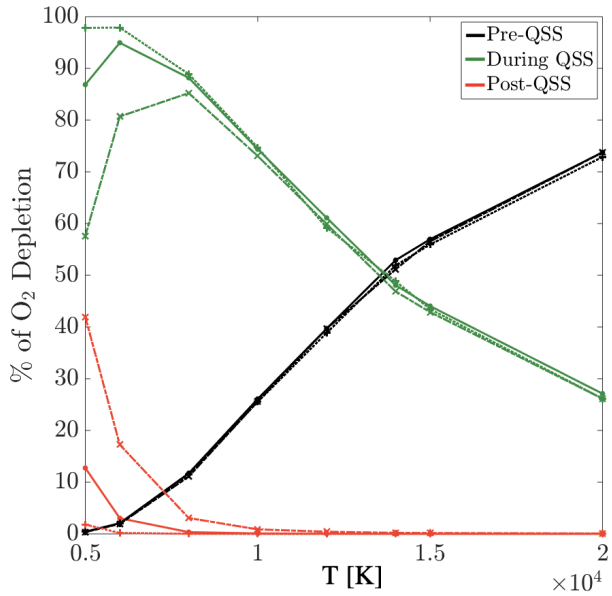


Figure 5.15: Percentages of O_2 depletion taking place before (black lines), during (green lines), and after (red lines) QSS, as functions of heat bath temperature. Three different scenarios are analyzed, all with $T^0_{Int} = 300$ K and $[O_2]^0 = 95\%$. Dashed lines: $n^0_d = 3.33 \times 10^{24}\ m^{-3}$ ($p^0 = 13806.49$ Pa). Unbroken lines: $n^0_d = 3.33 \times 10^{23}\ m^{-3}$ ($p^0 = 1380.65$ Pa). Dotted lines: $n^0_d = 3.33 \times 10^{22}\ m^{-3}$ ($p^0 = 138.07$ Pa).

O_2 depletion occurs before the rovibrational distribution starts to be temporarily unvaried. From this figure, we can conclude that analyzing the O_2+O system properties only at a single point-in-time snapshot might at best result in a partial understanding of the dissociation dynamics. The correctness of the QSS characterization is a necessary, but not sufficient, condition for validating the predictive capabilities of reduced order models.

Part III

Conclusions

Chapter 6

Summary and Future Work

6.1 Summary

The proposed work leveraged machine learning, uncertainty quantification, Data Science, and reduced order models for improving the predictive capabilities of *ab initio* Hypersonics. The first part of the manuscript presented an example of how Bayesian machine learning can be employed for constructing surrogate models starting from a limited set of data points. The non-deterministic properties of the resulting hyper-surfaces can be exploited for: i) quantifying the effects that the related uncertainties have on QoIs, and for ii) proposing possible corrections. The results from this first part are summarized in Sec. 6.1.1. The rest of the work analyzed large data sets obtained from expensive and accurate calculations for gathering additional insights on the dominant physics. The discovered patterns have been successively used for coarse-graining the dynamics and reducing the complexity of the computational simulations. The information lost due to reduced order modeling has been finally quantified via an encoding-decoding approach. Sec. 6.1.2 summarizes the findings from this second part.

6.1.1 Bayesian ML Approach to the UQ of *Ab Initio* PESs

This part proposed a new Bayesian framework that relies on a non-deterministic extension of the potential energy surface, referred to as *stochastic PES* (SPES). The approach comprises three stages: i) stochastic calibration of PES parameters and hyperparameters by using *ab initio* data points; ii) forward propagation of the SPES through a sequence of simulations involving QCT and ME calculations; iii) and sensitivity analysis of predetermined quantities of interest to the uncertainties of the surface. The methodology has been applied to the O_2+O system for a range of conditions of interest in hypersonic aerothermodynamics, using the *ab initio* database provided

by Varga *et al.* [150]. Of the nine PESs for the electronic ground-state asymptote, only the $1^1A'$ and $2^5A'$ surfaces have been analyzed. It has been found that:

- Despite the uncertainties in the representation of the ozone region, in an isothermal chemical reactor study at 10,000 K the inaccuracies on the singlet PES ($1^1A'$) result in only $\pm 6\%$ uncertainty in the computed composition profile. The uncertainties of the vibrational and rotational relaxation times are also negligible, as well as the ones for dissociation and exchange reaction rate coefficients.

- The quintet PES ($2^5A'$) manifests significantly larger uncertainty, as the composition profile in the chemical reactor has a variability as large as $\pm 20\%$ at 10,000 K. At that temperature, the actual rate of O_2 dissociation can be up to 6 times larger than the value predicted using Varga *et al.*'s fit.

- For the quintet PES, the vibrational relaxation time is characterized by uncertainty as large as 1 order of magnitude at low temperatures, and its actual value can be up to ten times larger than that predicted through Varga *et al.*'s fit.

- For the quintet surface, the uncertainties in the QoIs are traceable to the inadequacies in the characterization of the repulsive wall and energy barrier. More *ab initio* data points are needed in these PES regions.

Compared to the deterministic machine learning techniques, the proposed methodology significantly increases the information content gained from the *ab initio* based data. The discussed procedure allows for the quantification of the uncertainties on the quantities of interest (QoIs) as a result of inaccuracies in the PES, and more importantly, it allows to identify the areas that mostly contribute to such unreliability. Any additional refinement of the surface, then, can be circumscribed to such locations, allowing a significant reduction of the computationally expensive *ab initio* calculations. This procedure facilitates sensitivity studies of the uncertainty in *ab initio* PESs and can lead to effective validation of the *ab initio*-QCT rate coefficients used in complex chemical models.

6.1.2 Data-Inspired and Physics-Driven Model Reduction for Dissociation

By leveraging QCT calculations, this work has constructed the first rovibrational state-to-state database for the O_2-O interactions described by Varga *et al.*'s PESs [150]. The *ab initio* data-set of rate coefficients has been used to consistently construct two coarse-grained (CG) models: the vibrational-specific (VS) model and the new CB approach, which groups the rovibrational levels based on their energy-deficits from the centrifugal barrier. The two strategies have been tested in a hybrid framework, which combines an exact treatment of excitation and a coarse-grained dissociation. In a temperature range of hypersonic interest (1 500 - 20 000 K), comparisons against the fully-StS model in isothermal and isochoric chemical reactors have revealed that:

- An accurate description of the vibrational nonequilibrium is necessary but not sufficient for accurately characterizing the dissociation mechanisms. Indeed, even if the VS model was able to reproduce the vibrational distribution function from the fully-StS model, it was unable to describe the composition profiles correctly.

- The VS model under-predicts the time scales of the overall molecular dissociation by up to 65% at 20,000 K. The error is primarily due to the VS inadequacy in characterizing the dissociation dynamics, rather than the energy exchange processes.

- The predictions based on the newly developed CB model are in excellent agreement with results obtained via the direct solution of the rovibrational-specific state-to-state master equation. Using only 20 groups, the CB model outperforms the VS one in characterizing the dissociation mechanisms. Compared to the latter approach, the accuracy in the prediction of mole fraction is improved up to a factor 20 (*i.e.*, 2000%) at the highest temperature analyzed.

It is worth emphasizing that, based on what is presented in this work, vibrational-specific databases of rate coefficients are not able to accurately characterize the dissociation, even if computed from the first principles of quantum chemistry and by relying on the most sophisticated PESs.
The CB grouping approach only depends on the diatomic potential, and it does not require any *a priori* knowledge of dissociation probabilities. If the diatomic potential remains unchanged, the same groups can be used for modeling interactions taking place on different PESs, or even with

different collision partners. Consequently, this novel strategy is particularly suitable for CG-QCT simulations, in which the group-specific rates are directly obtained from scattering calculations [90].

6.2 Future Work

The framework presented in this thesis can be employed in different applications and extended in multiple ways.

Adding an additional step to the forward propagation would significantly improve the SPES approach presented above. The PDFs of QSS dissociation coefficients and vibrational relaxation times, which are consequences of PES uncertainties, should be inputted to multi-temperature models [41, 118] in CFD simulations. In this way, the effects of the PES inaccuracies can be directly quantified on the ultimate QoIs (*e.g.*, TPS heat fluxes).

The SPES methodology could also be enhanced by replacing the deterministic description of the diatomic contributions in Eq. 2.12 with stochastic formulations. For example, this could be achieved by representing the pairwise terms via the eventempered Gaussian fitting function of Bytautas *et al.* [20] and by treating the related coefficients as probability distribution functions. Such an extension would generate non-negligible complications; above all, the necessity of computing a separate list of rovibrational levels for each diatomic curve sampled from the pairwise potential distribution. At the same time, this improvement would also allow a sensitivity analysis of the macroscopic QoIs (*e.g.*, QSS dissociation rates or vibrational relaxation times) to the diatomic potential features.

New first-principles calculations should be performed for the $2^5A'$ O_2+O interactions at the high-sensitivity locations proposed in Sec. 3.5. The augmented set of data points should be fitted and successively compared to the original $2^5A'$ PES by Varga *et al.*

The Spebus toolbox developed by the author of this thesis is currently being used at the NEQRAD group, University of Illinois at Urbana-Champaign, for constructing PIP-NN-based PESs of the CNH system. Soon, the toolbox will be extended for being able to handle chemical systems composed of more than three atoms.

Regarding the dissociation mechanisms analysis, work is currently underway in performing equivalent studies on different chemical systems and on recombining mixtures. In the past months, the

CB grouping strategy has been applied to the dissociation of O2+C, CO+O, N_2+N, N_2+O, and NO+N. For these triatomic chemical systems, the author of this thesis computed the RVS rate coefficients by employing CoarseAIR [153] in QCT simulations based on the PESs by Schwenke *et al.* [136] for O2+C and CO+O, the PES by Jaffe *et al.* [56, 59] for N2+N, and the PESs from Lin *et al.* [84] for N_2+O and NO+N. The data sets or rate coefficients have been collected in *The CHESS Database of Ab Initio Rate Coefficients for Hypersonic Applications* [157]. Sec. B.3 of Appendix B reports the comparison between the CB-averaged data-sets and the original *ab initio* dissociation rate coefficients.

Finally, it should be highlighted that the CB model for dissociation mechanisms can be easily integrated with the adaptive binning for energy exchange processes [129], thus removing one of the limitations that has affected the accuracy of the CG models. The approach can also be employed for selecting the subset of rovibrational levels to conduct accurate *ab initio* calculations and generate unbiased training data for neural network representations [73].

Appendix A

Part I, Additional Results

This part of the appendix reports additional results from the application of the PIP-BNN approach to the $1^1A'$ and $2^5A'$ potential energy surfaces of the O_2+O System.

A.1 PESs 2D Cuts

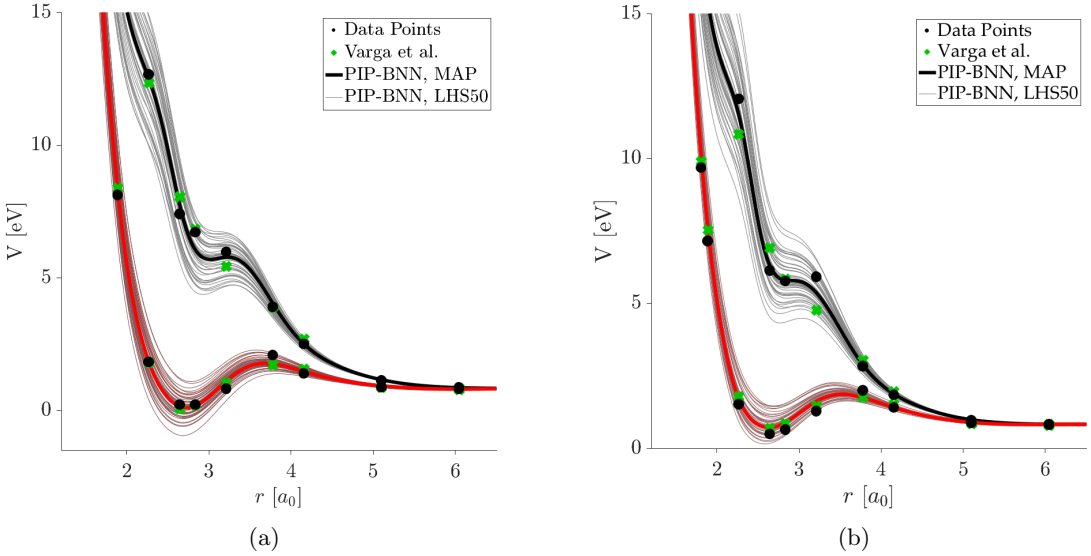


Figure A.1: $1^1A'$ (red lower parts of the Fig.s) and $2^5A'$ (black upper parts) PESs at four different sets of geometries. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fit at the data geometries, bold lines to PIP-BNN MAP, and the remaining narrow lines to the 50 PIP-BNN samples. (a): $\measuredangle O_A O_B O_C = 60°$ and $r_1 = 2.646 a_0$. (b): $\measuredangle O_A O_B O_C = 70°$ and $r_1 = 2.646 a_0$.

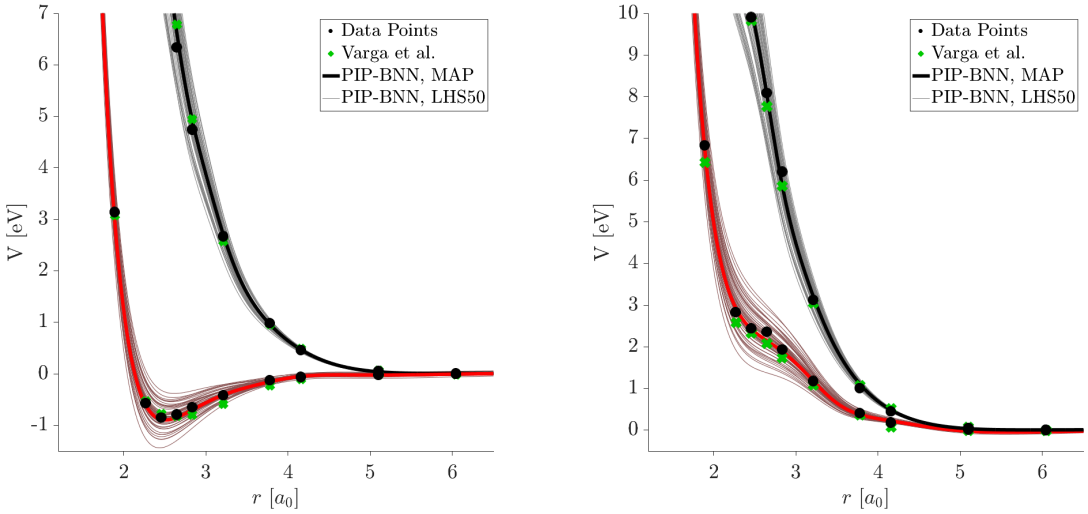
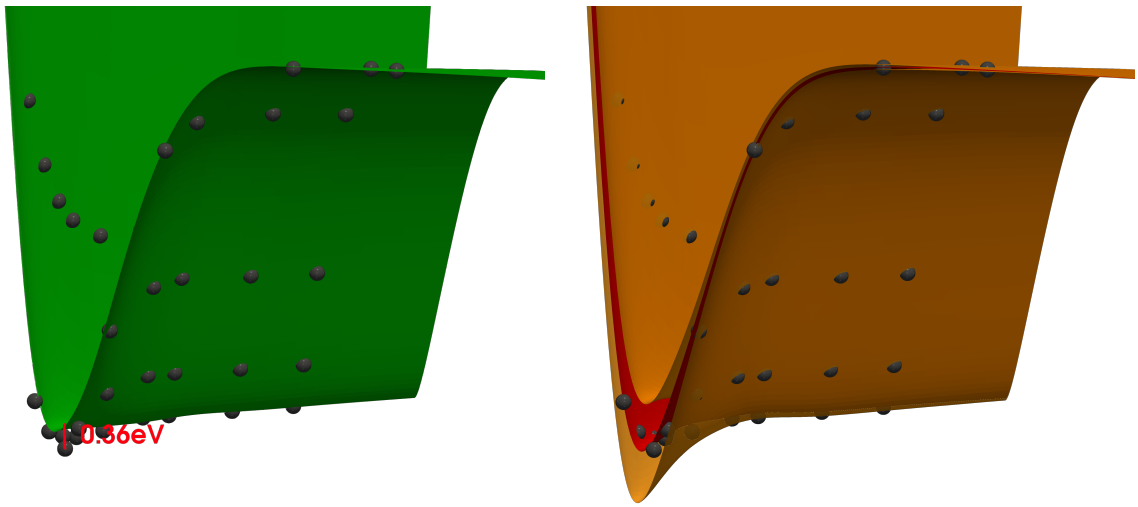


Figure A.2: $1^1A'$ (red lower parts of the Fig.s) and $2^5A'$ (black upper parts) PESs at four different sets of geometries. Black dots correspond to the *ab initio* data points, green crosses to Varga *et al.*'s fit at the data geometries, bold lines to PIP-BNN MAP, and the remaining narrow lines to the 50 PIP-BNN samples. (a): $\measuredangle O_A O_B O_C = 110°$ and $r_1 = 2.268 a_0$. (b): $\measuredangle O_A O_B O_C = 170°$ and $r_1 = 2.268 a_0$.

A.2 PESs 3D Cuts

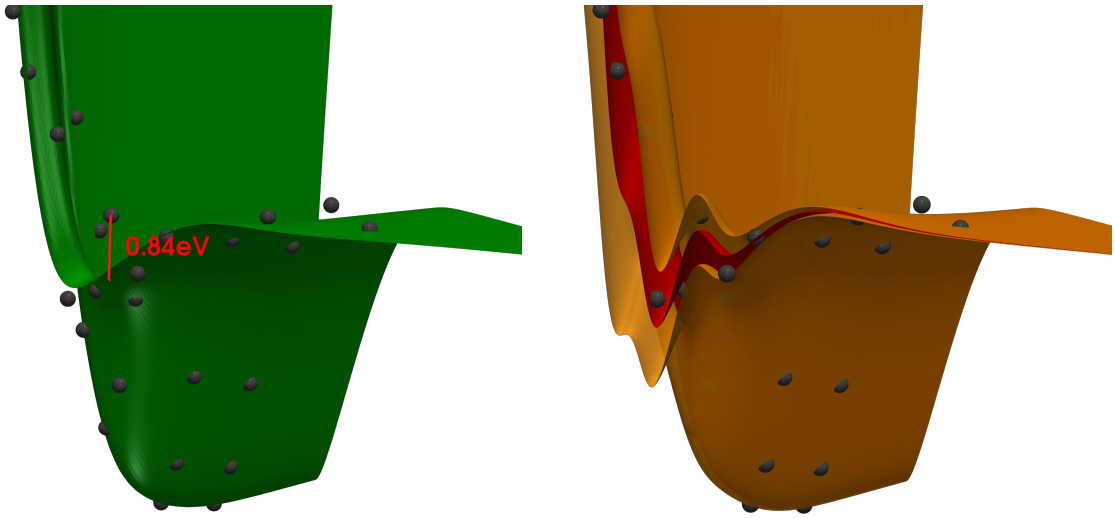
A.2.1 Results for $1^1A'$ PESs



(a) (b)

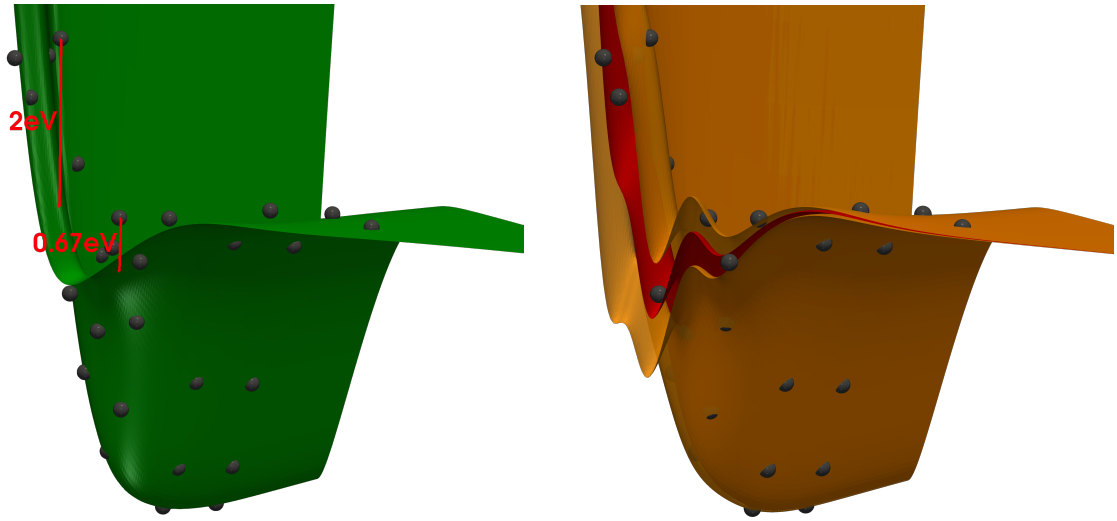
Figure A.3: 3D views of the $1^1A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_A O_B O_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red one identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_A O_B O_C = 100°$, Varga *et al.*. (b):$\measuredangle O_A O_B O_C = 100°$, PIP-BNN.

A.2.2 Results for $2^5A'$ PESs



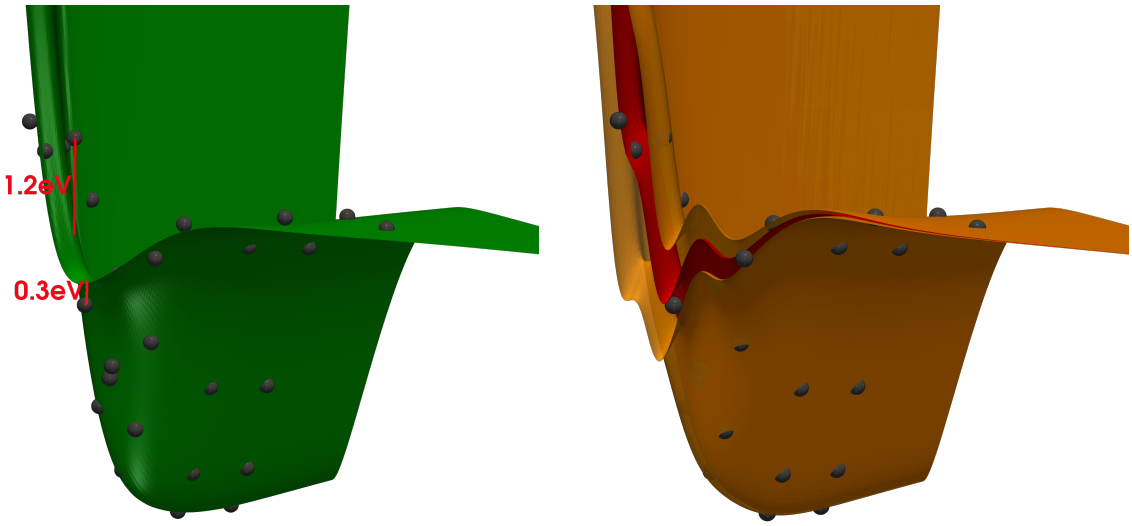
(a) (b)

Figure A.4: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 60°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 60°$, PIP-BNN.



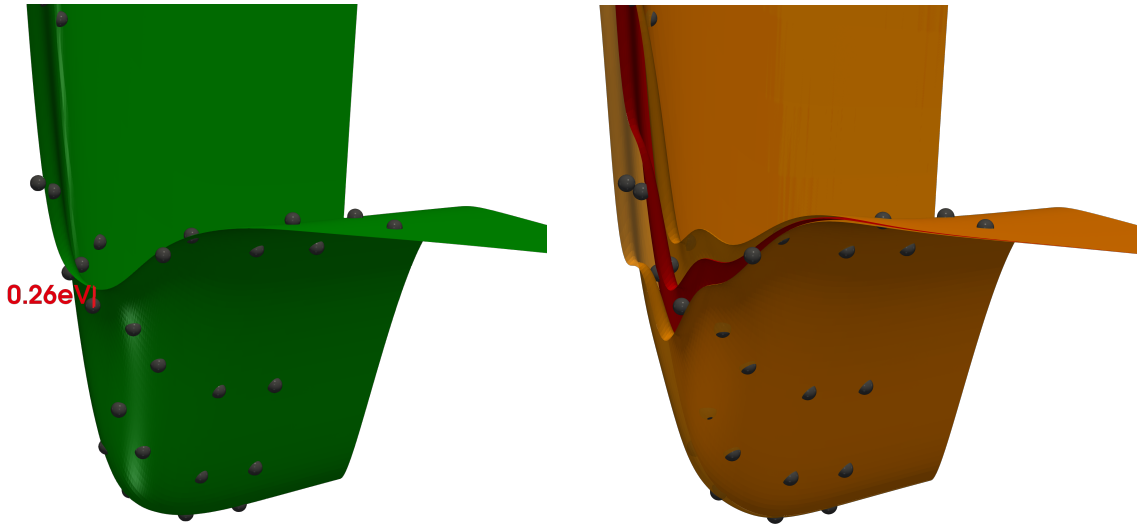
(a) (b)

Figure A.5: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 65°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 65°$, PIP-BNN.



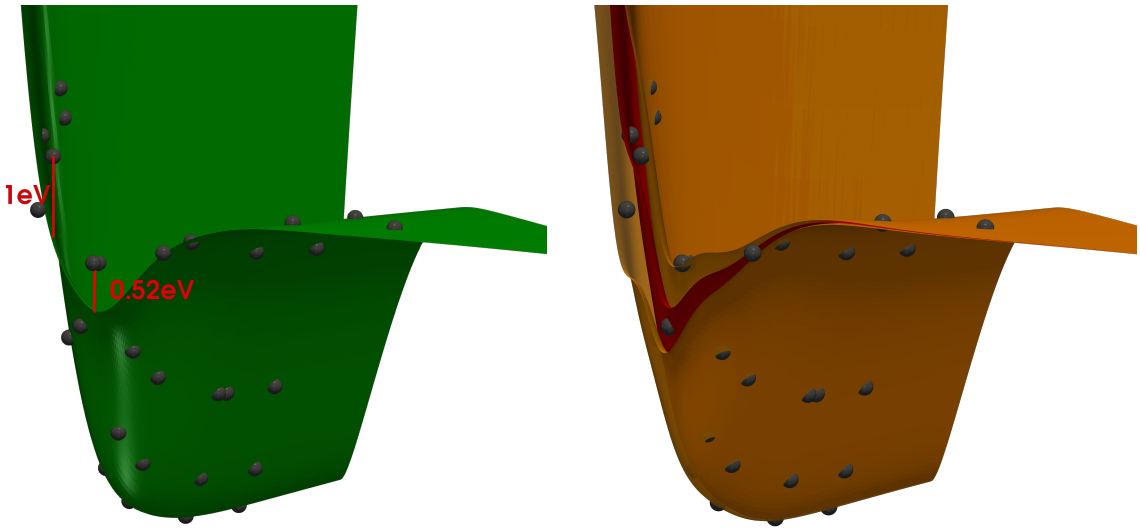
(a) (b)

Figure A.6: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 70°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 70°$, PIP-BNN.



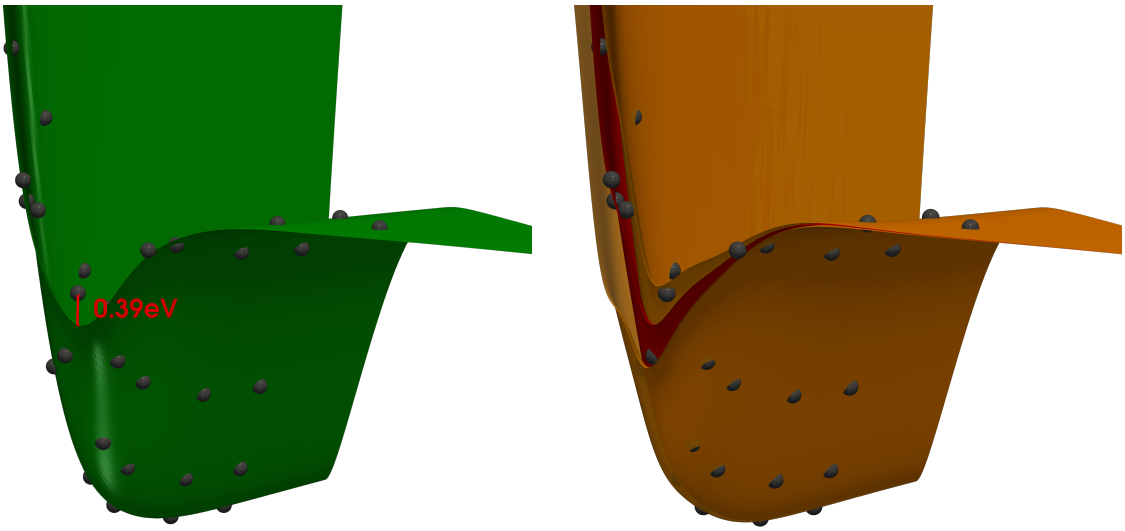
(a) (b)

Figure A.7: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 80°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 80°$, PIP-BNN.



(a) (b)

Figure A.8: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the *ab initio* data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 90°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 90°$, PIP-BNN.



(a) (b)

Figure A.9: 3D views of the $2^5A'$ PESs at $r_1 > r_3$ and multiple $\measuredangle O_AO_BO_C$, compared to the ab initio data points (black dots). The green surface represents Varga *et al.*'s fit, the red ones identifies the mean of the 50 PIP-BNN samples, while the orange ones correspond to the bounds of the three-sigma confidence intervals generated by such samples. (a): $\measuredangle O_AO_BO_C = 100°$, Varga *et al.*. (b):$\measuredangle O_AO_BO_C = 100°$, PIP-BNN.

A.3 RVS StS Rate Coefficients

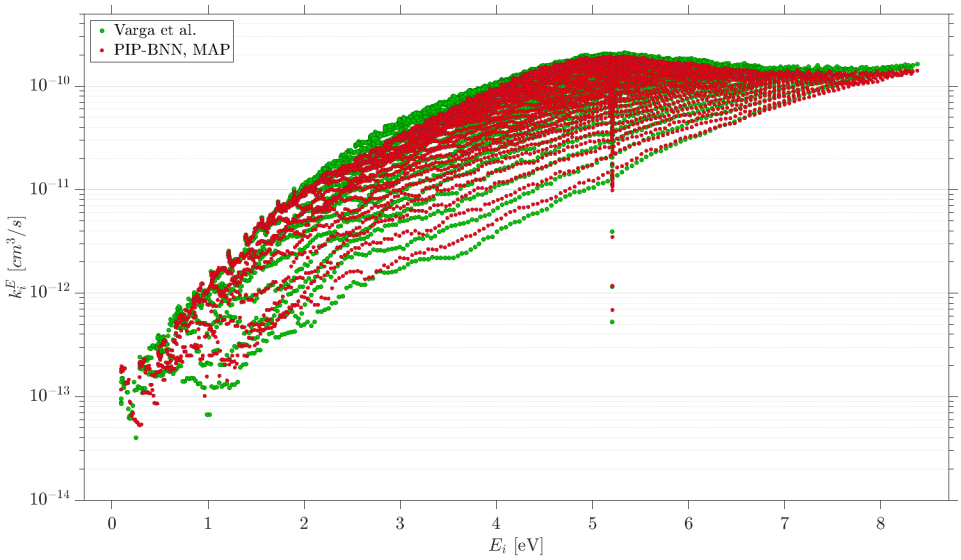


Figure A.10: Rovibrational state specific exchange rate coefficients for the $2^5A'$ interactions at $T_{Tran} = 10\,000$ K, computed as $k_i^E = \sum_{j=1}^{N_{RV}} k_{ij}^E$, where i corresponds to the initial level, j to the final one, and N_{RV} to the number of O_2 ro-vibrational levels. Green dots are obtained using Varga *et al.*'s fit, while the red ones are computed through the PIP-BNN MAP surface.

A.4 Gas Quantities of Interest at T = 10 000 K

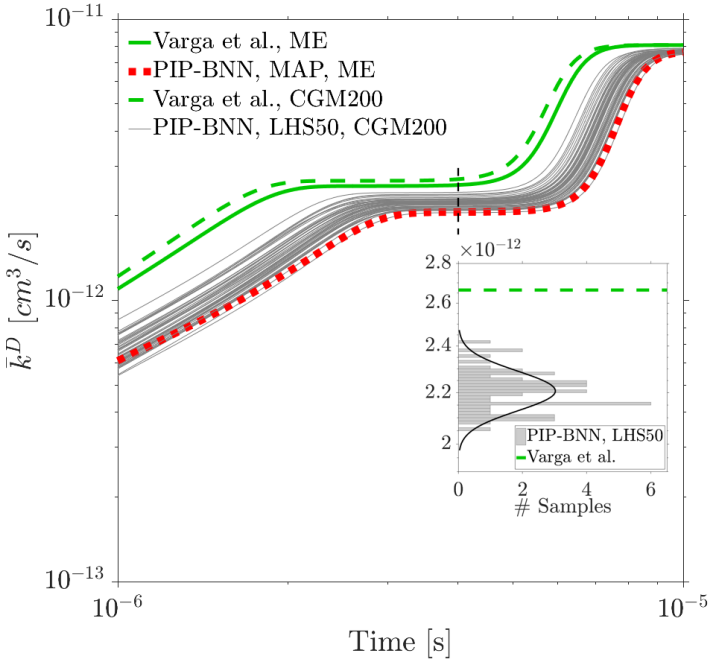


Figure A.11: Evolution of O_2 dissociation global rates, as predicted by ME (green solid lines) and CGM200 (green dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted lines), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig.s report the histograms of the respective QSS rates. The simulations have been performed considering only $2^5A'$ interactions at $T_{Tran} = 10\,000$ K. The dissociation global rate is computed as $\bar{k^D} = \sum_i \sum_j k_i^D n_i(t)/n_{O_2}(t)$; i identifies the O_2 initial state, in case of ME calculations, or the initial group in case of CGM; k_i^D corresponds to the dissociation rate associated with the i-th state/group, and n_i represents the population of the i-th state/group.

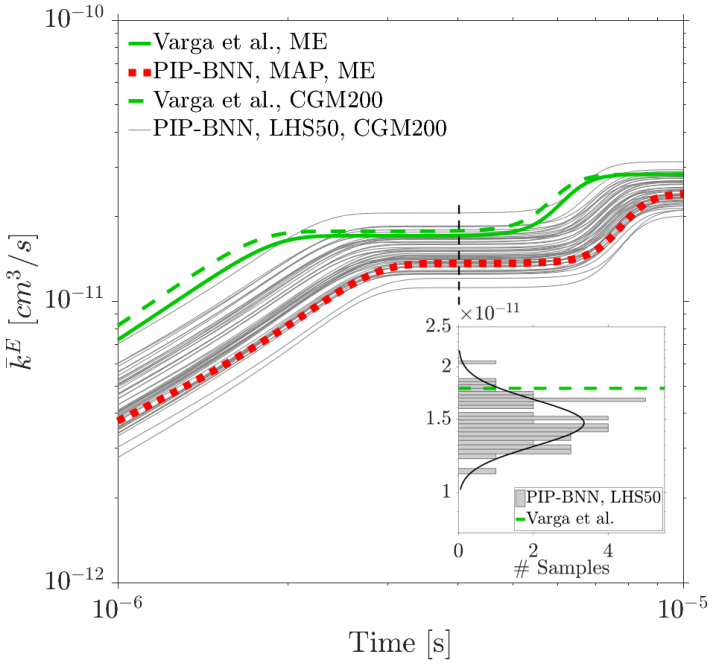


Figure A.12: Evolution of O_2 exchange global rates, as predicted by ME (green solid lines) and CGM200 (green dashed lines) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted lines), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig.s report the histograms of the respective QSS rates. The simulations have been performed considering only $2^5A'$ interactions at $T_{Tran} = 10\,000$ K. The exchange global rate is computed as $\bar{k^E} = \sum_i \sum_j k^E_{ij} n_i(t)/n_{O_2}(t)$; i identifies the O_2 initial state, in case of ME calculations, or the initial group in case of CGM, and j the final one; k^E_{ij} corresponds to the exchange rate from the i-th state/group to the j-th, and n_i represents the population of the i-th state/group.

A.5 Gas Quantities of Interest at T = 2 500 K

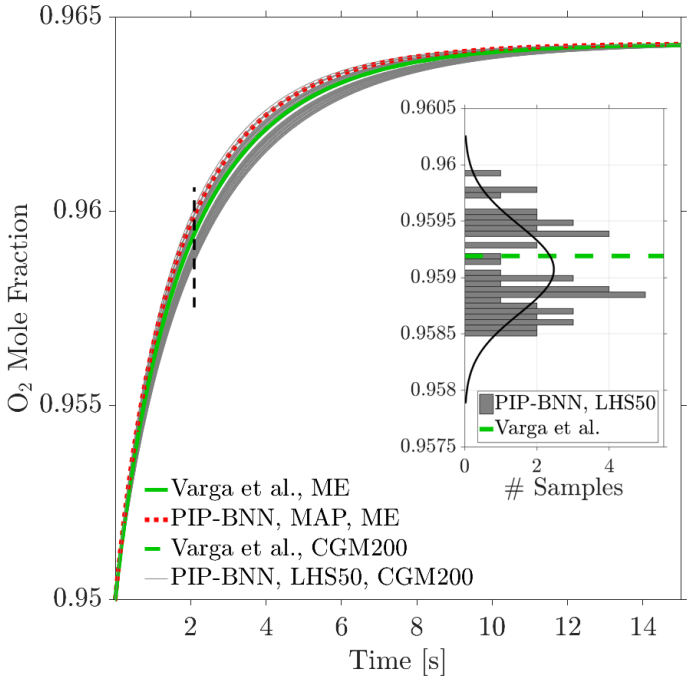


Figure A.13: Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig. on the right reports the histogram of such samples at $t = 2s$. The PIP-BNN adopted is the one constructed for $1^1A'$, and the simulations are performed at $T_{Tran} = 2\,500$ K.

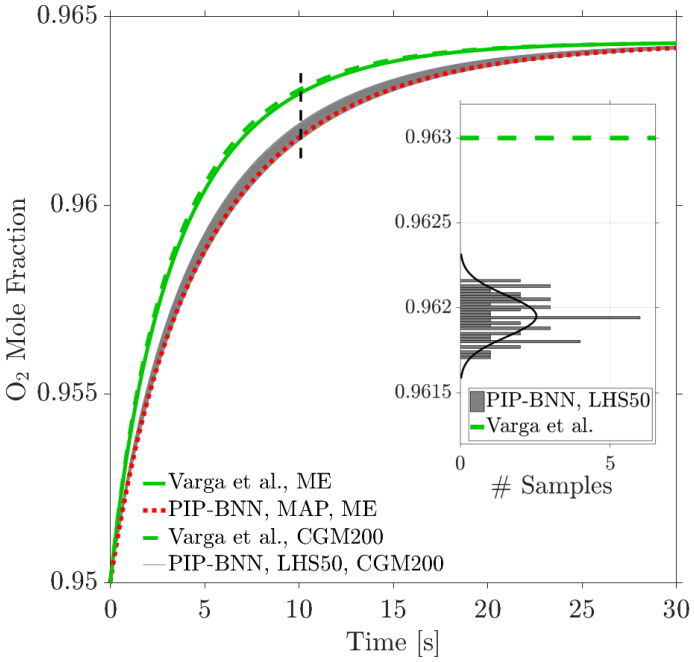


Figure A.14: Evolution of O_2 mole fraction, as predicted by ME (green solid line) and CGM200 (green dashed line) using Varga *et al.*'s fit, by ME using the PIP-BNN MAP surface (red dotted line), and by CGM200 using the 50 PIP-BNN samples (grey narrow lines). The Subfig. on the right reports the histogram of such samples at $t = 10s$. The PIP-BNN adopted is the one constructed for $2^5 A'$, and the simulations are performed at $T_{Tran} = 2\,500$ K.

A.6 Gas Quantities as Functions of Temperature

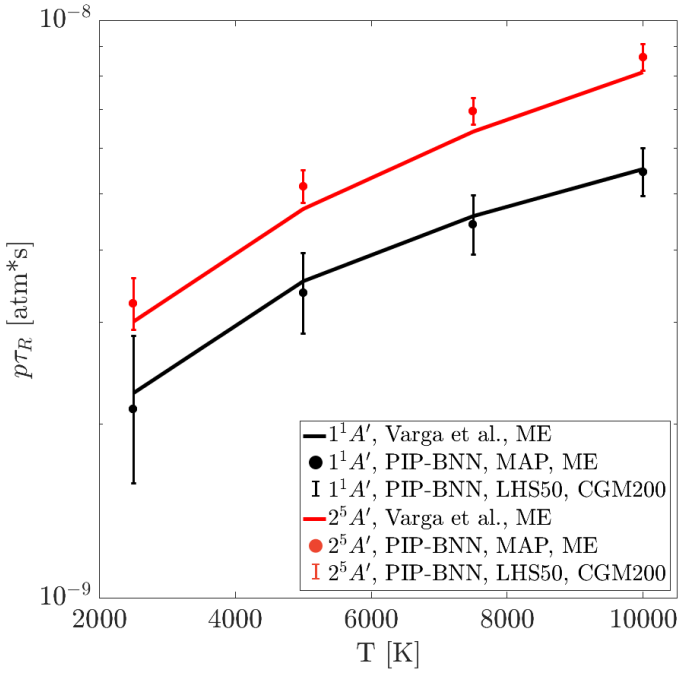


Figure A.15: Rotational relaxation times for $1^1A'$ (in black) and $2^5A'$ (in red) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples.

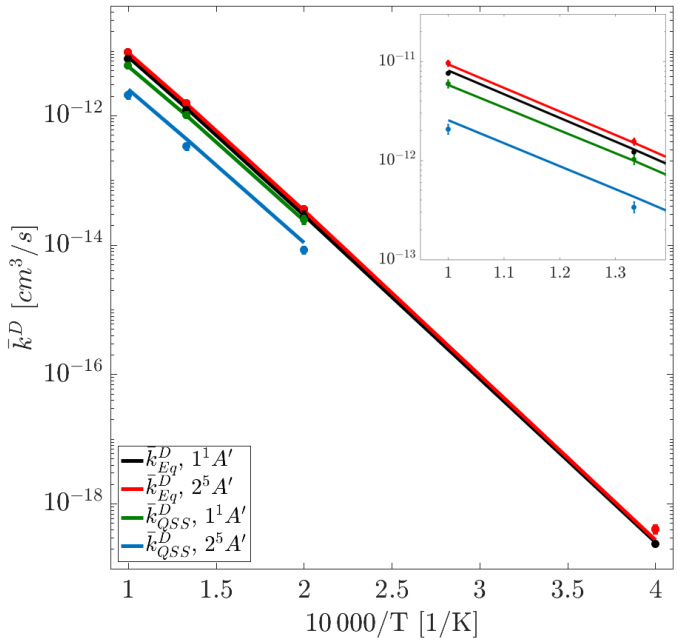


Figure A.16: Equilibrium and QSS dissociation rates for the $1^1A'$ (in black and green) and $2^5A'$ (in red and blue) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. At $T_{Tran} = 2\,500$ K no QSS is present for any of the two PESs. In the top-right square, the main Fig. is zoomed at $T_{Tran} = 10\,000$ K and $T_{Tran} = 7\,500$ K.

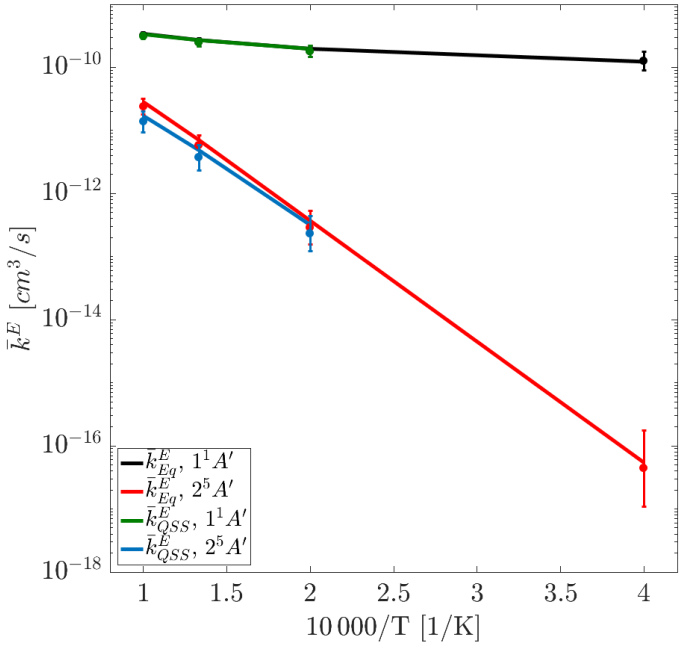


Figure A.17: Equilibrium and QSS exchange rates for the $1^1A'$ (in black and green) and $2^5A'$ (in red and blue) PESs. The solid lines report the values obtained using Varga et al.'s fits, the dots the ones computed through the PIP-BNN MAP surfaces, while the vertical lines identify the three-sigma intervals corresponding to the 50 PIP-BNN samples. At $T_{Tran} = 2\,500$ K no QSS is present for any of the two PESs.

A.7 Gas QoI Sensitivities to the $2^5A'$ PES Uncertainties

A.7.1 Vibrational Relaxation Times

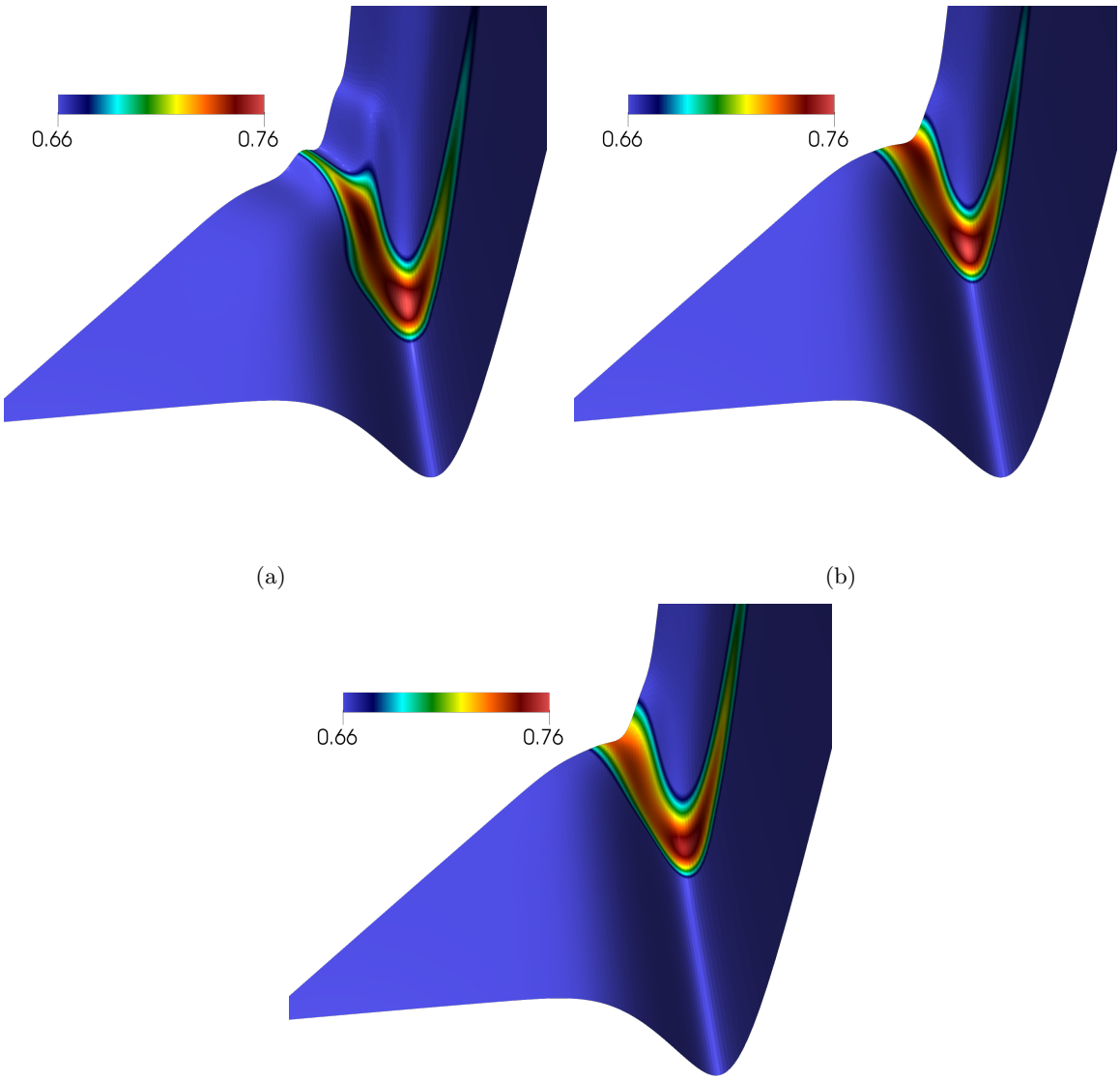
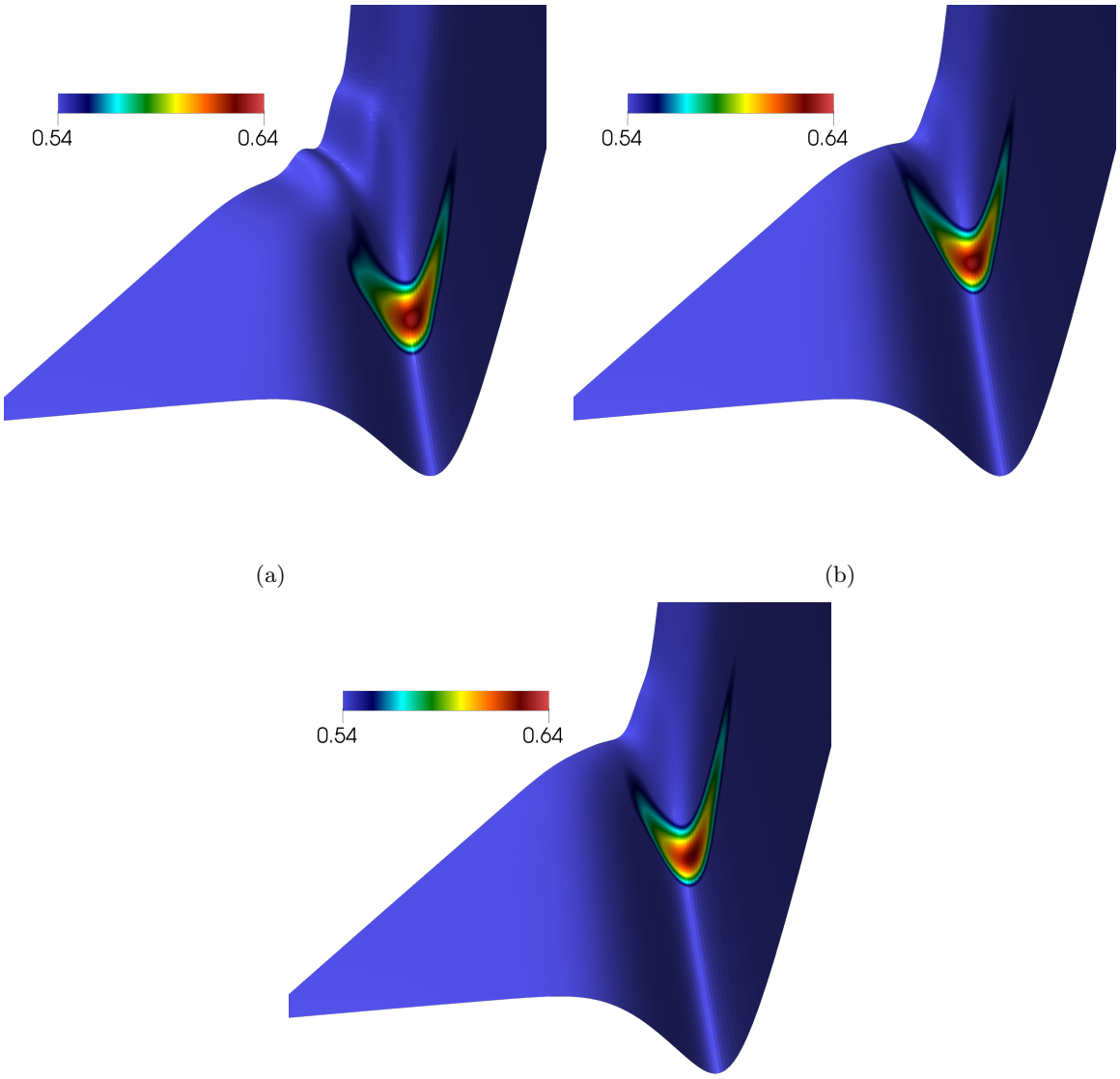


Figure A.18: 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the vibrational relaxation time at $T_{Tran} = 10\,000$ K. (a): $\measuredangle O_AO_BO_C = 50°$. (b): $\measuredangle O_AO_BO_C = 110°$. (c): $\measuredangle O_AO_BO_C = 175°$.

A.7.2 Vibrational Relaxation Times Excluding Exchange Processes

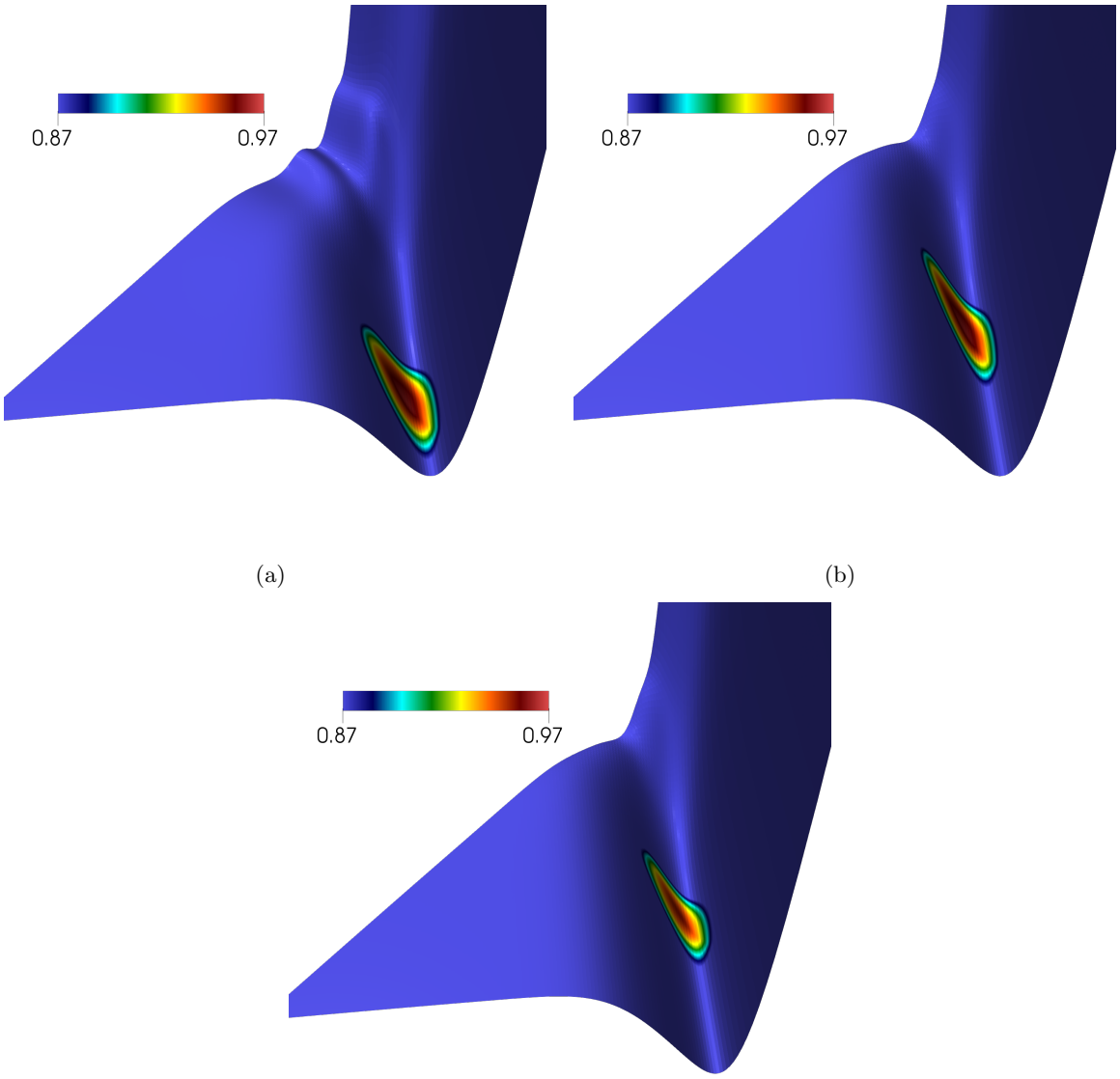


(a) (b)

(c)

Figure A.19: 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the vibrational relaxation time at $T_{Tran} = 10\,000$ K, obtained by artificially excluding the exchange processes. (a): $\measuredangle O_A O_B O_C = 50°$. (b): $\measuredangle O_A O_B O_C = 110°$. (c): $\measuredangle O_A O_B O_C = 175°$.

A.7.3 Rotational Relaxation Times



(a) (b)

(c)

Figure A.20: 3D views of the $2^5A'$ PES resulting from the average of the 50 PIP-BNN samples for $r_1 > r_3$. The surfaces have been colored based on the absolute values of Pearson correlation coefficients between the PES and the rotational relaxation time at $T_{Tran} = 2\,500$ K. (a): $\measuredangle O_A O_B O_C = 50°$. (b): $\measuredangle O_A O_B O_C = 110°$. (c): $\measuredangle O_A O_B O_C = 175°$.

Appendix B

Part II, Further Results

This part of the appendix reports additional results from the application of reduced order models to the dissociation kinetics of O_2+O and other chemical systems.

B.1 Dimensionality Reduction of RVS Dissociation Rate Coefficients

B.1.1 RVS Dissociation Rate Coefficients Reconstructed from the VS Model

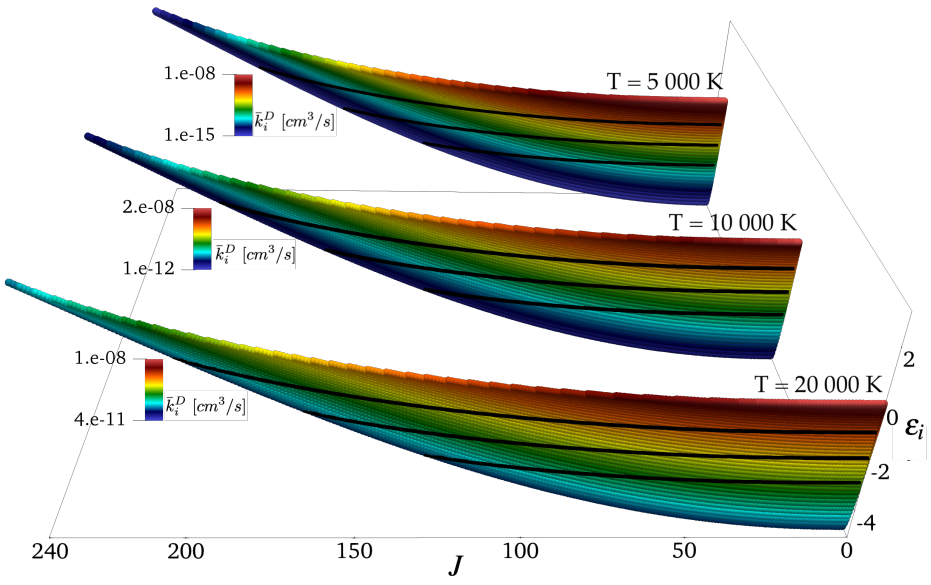


Figure B.1: O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy. The levels are colored based on the RVS dissociation rate coefficients from the vibrationally-averaged data set at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier.

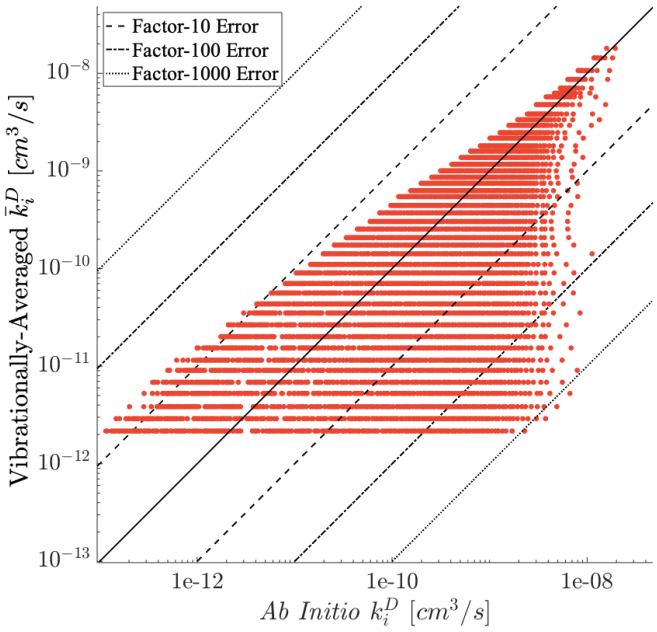


Figure B.2: Scatter plot representing the error generated by the VS dimensionality reduction of the RVS dissociation rate coefficients at T = 10 000 K.

B.1.2 RVS Dissociation Rates Coefficients Reconstructed from the CB Model

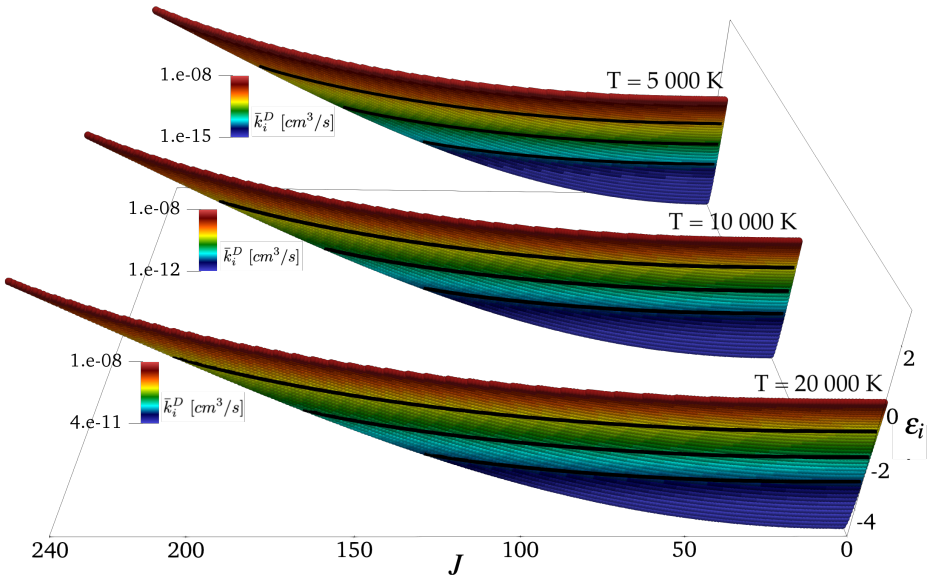


Figure B.3: O_2 rovibrational levels plotted as functions of rotational quantum number and internal energy. The levels are colored based on the RVS dissociation rate coefficients from the CB-averaged data set at T = 5 000, 10 000, and 20 000 K. Black lines represent isolines from the centrifugal barrier.

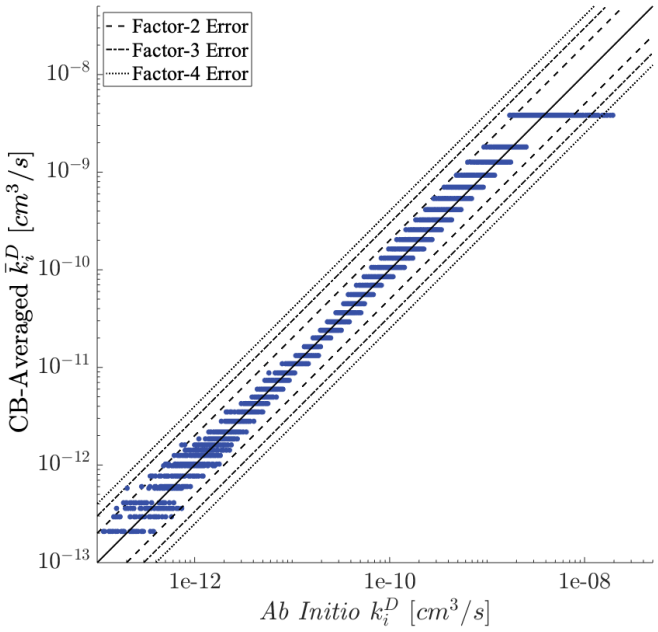


Figure B.4: Scatter plot representing the error generated by the CB dimensionality reduction of the RVS dissociation rate coefficients at T = 10 000 K.

B.1.3 Depletion Delays

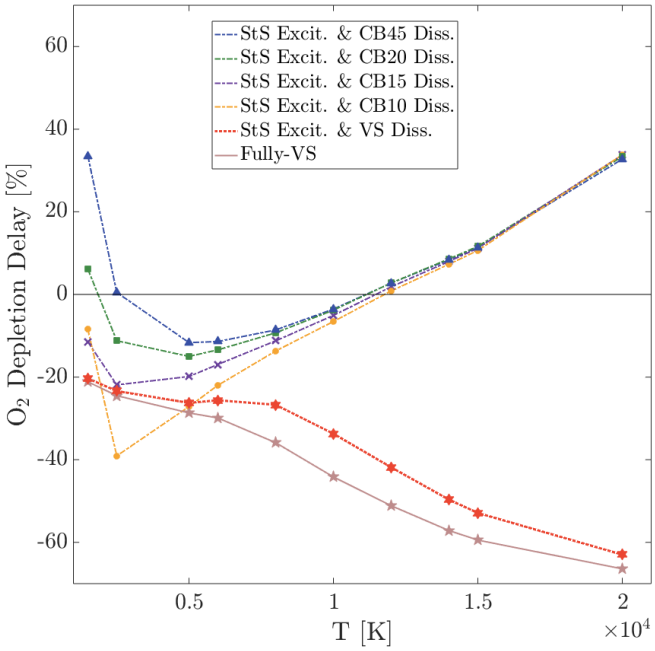


Figure B.5: Delays of O_2 depletion compared to the fully-StS simulations. The conditions of the isothermal and isochoric heat bath are $T^0_{Int} = 300$ K, $n^0_d = 3.33 \times 10^{23}$ m^{-3}, $[O_2]^0 = 95\%$, and $T_{Tran} = 10\,000$ K. Red dotted line: RVS StS excitation coupled to VS dissociation. Pink unbroken line: excitation and dissociation are both computed vibrationally-specific. All the other lines are obtained by computing excitation trough RVS StS and dissociation through CB model, with 45 (blue dash-dotted line), 20 (green dash-dotted line), 15 (purple dash-dotted line) and 10 (yellow dash-dotted line) groups. For these last four calculations, fitted dissociation rate coefficients are used.

B.2 Amount of O_2 Dissociation Taking Place at QSS

B.2.1 Heat Bath at $T_{Tran} = 5\,000$ K

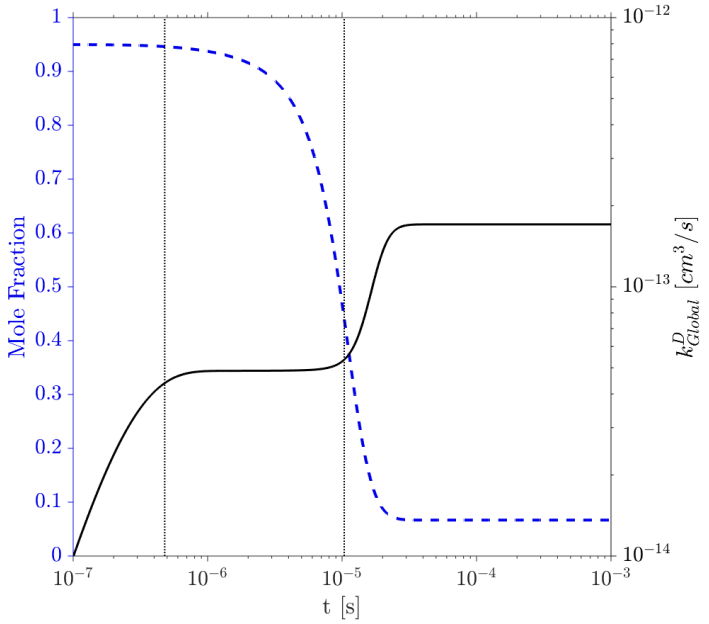


Figure B.6: Evolution of O_2 mole fraction (blue dash-dotted line) and global dissociation rate coefficient (black unbroken line) in a heat bath at $T_{Tran} = 5\,000$ K and $n_d^0 = 3.33 \times 10^{24}$ m^{-3} (P^0 = 13806.49 Pa). The simulation is performed fully-StS. The two vertical lines represent the time instants at which the quasi-steady state is assumed starting and terminating.

B.2.2 Heat Bath at T_{Tran} = 20 000 K

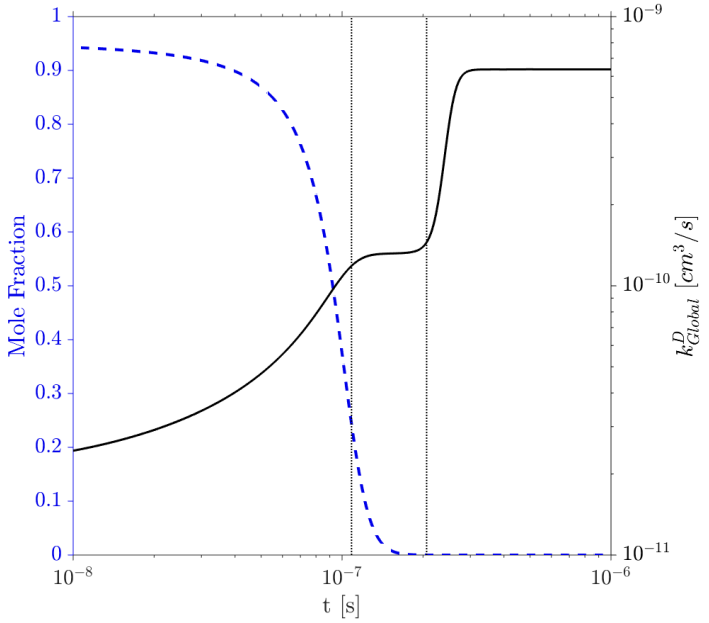


Figure B.7: Evolution of O_2 mole fraction (blue dash-dotted line) and global dissociation rate coefficient (black unbroken line) in a heat bath at T_{Tran} = 20 000 K and $n_d^0 = 3.33 \times 10^{23}$ m^{-3} (P^0 = 1380.65 Pa). The simulation is performed fully-StS. The two vertical lines represent the time instants at which the quasi-steady state is assumed starting and terminating.

B.3 Application to Other Chemical Systems

B.3.1 Application to O_2+C System

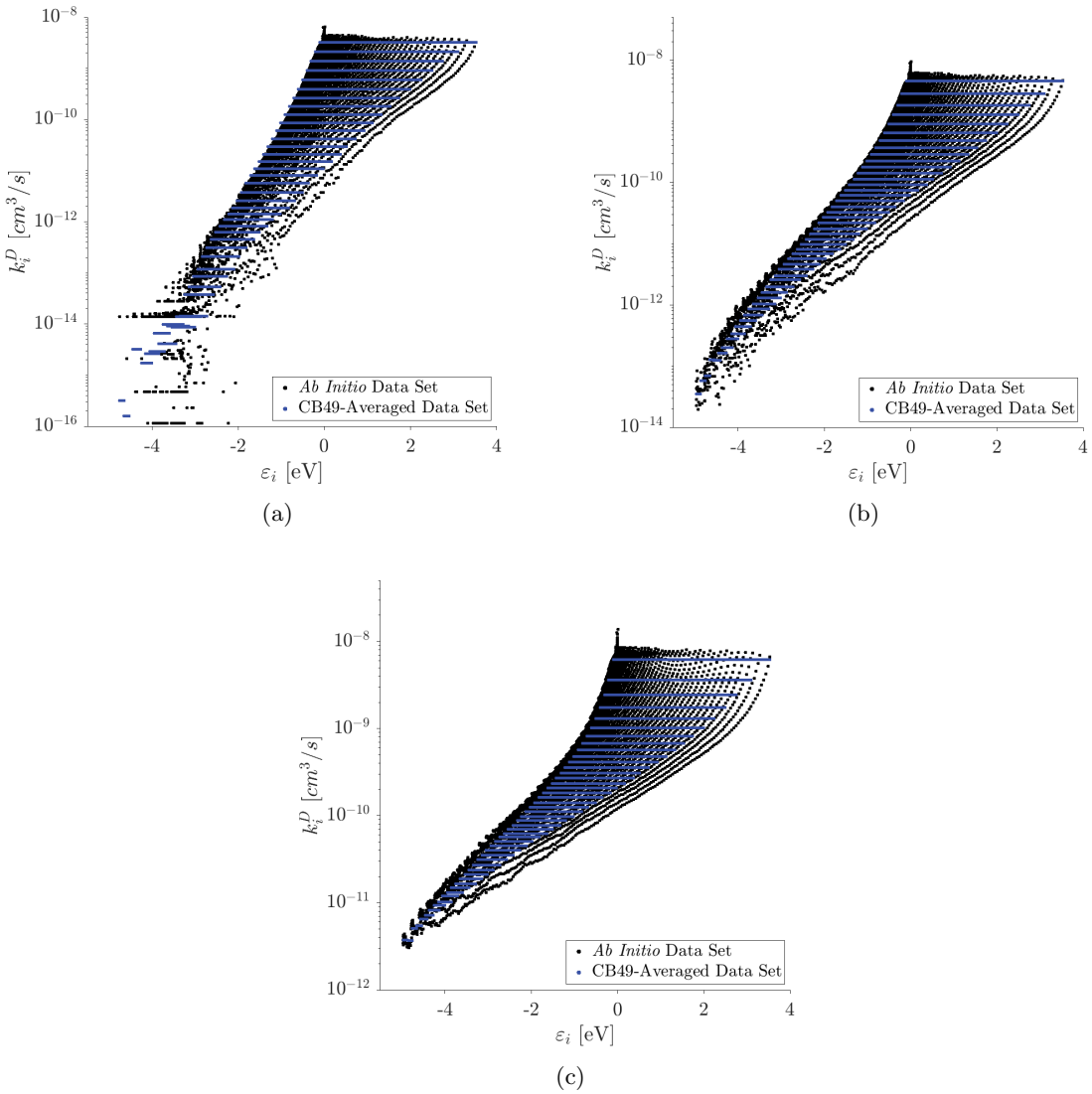


Figure B.8: Rovibrational-specific (RVS) dissociation rate coefficients at $T_{Tran} = 5\,000$ K (a), $T_{Tran} = 10\,000$ K (b), and $T_{Tran} = 20\,000$ K (c) for the O2+C system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 49 groups for the O_2 molecule.

B.3.2 Application to CO+O System

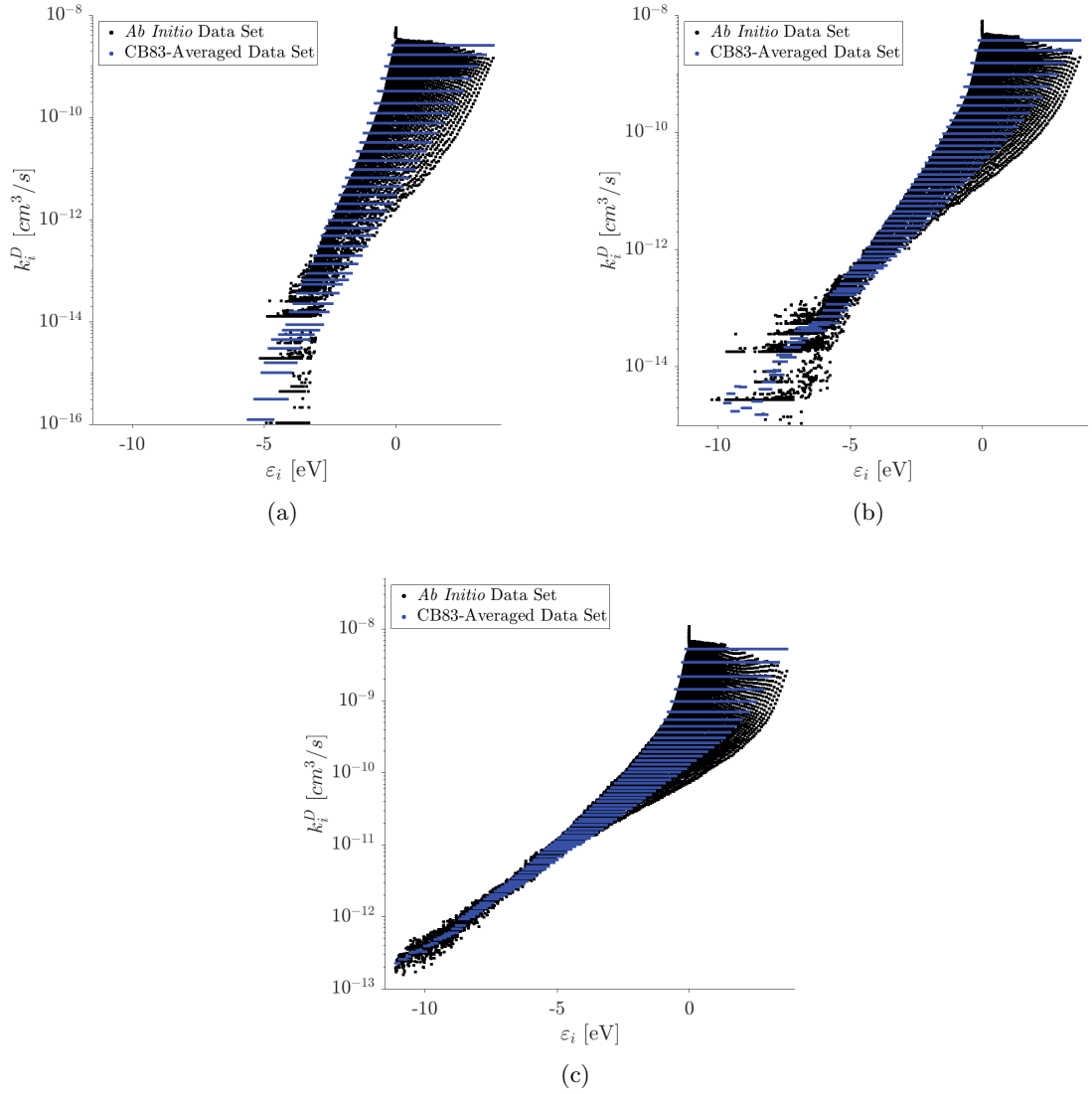


Figure B.9: Rovibrational-specific (RVS) dissociation rate coefficients at $T_{Tran} = 5\,000$ K (a), $T_{Tran} = 10\,000$ K (b), and $T_{Tran} = 20\,000$ K (c) for the CO+O system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 83 groups for the CO molecule.

B.3.3 Application to N_2+N System

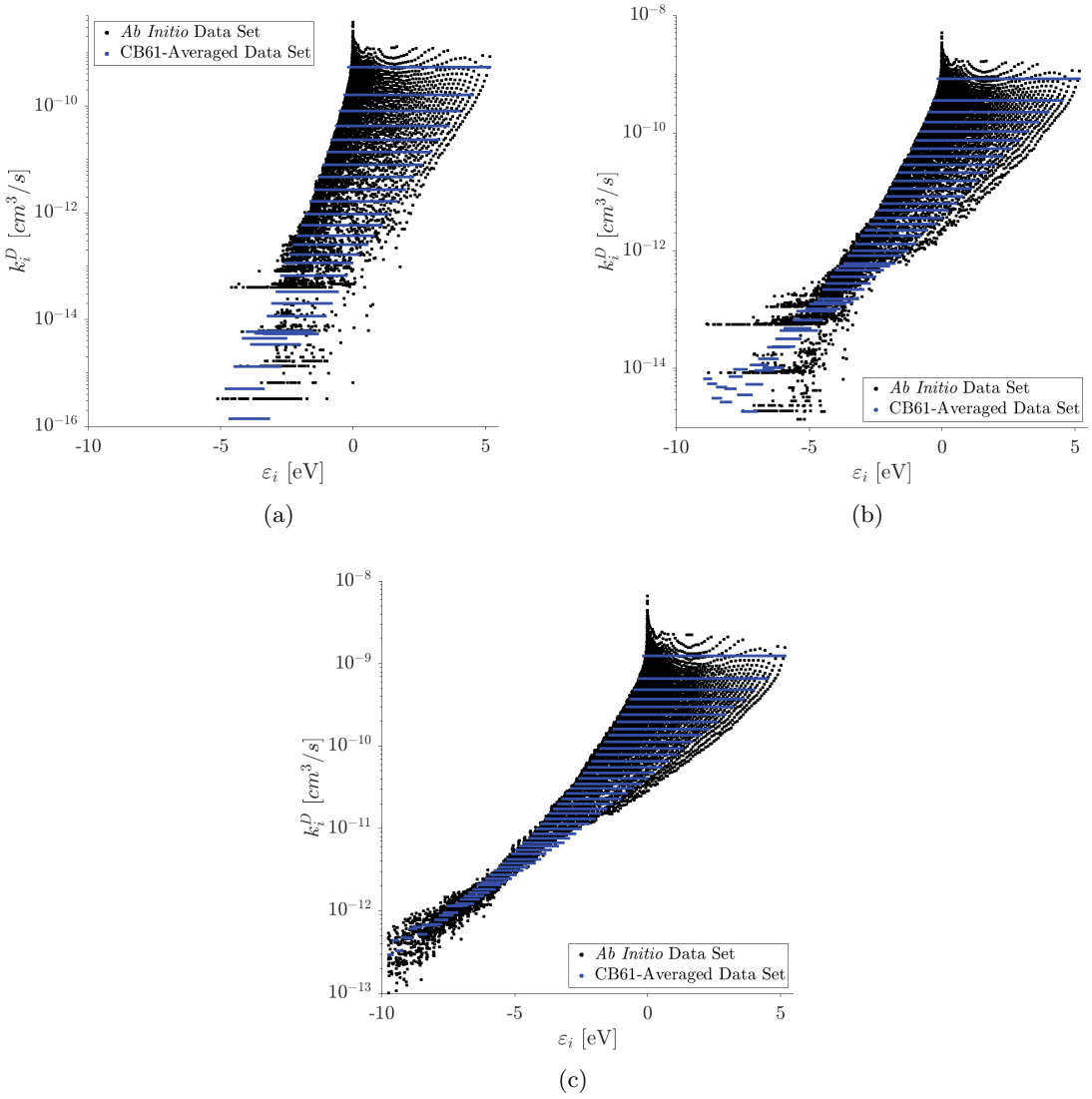


Figure B.10: Rovibrational-specific (RVS) dissociation rate coefficients at $T_{Tran} = 5\,000$ K (a), $T_{Tran} = 10\,000$ K (b), and $T_{Tran} = 20\,000$ K (c) for the N_2+N system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 61 groups for the N_2 molecule.

B.3.4 Application to N_2+O System

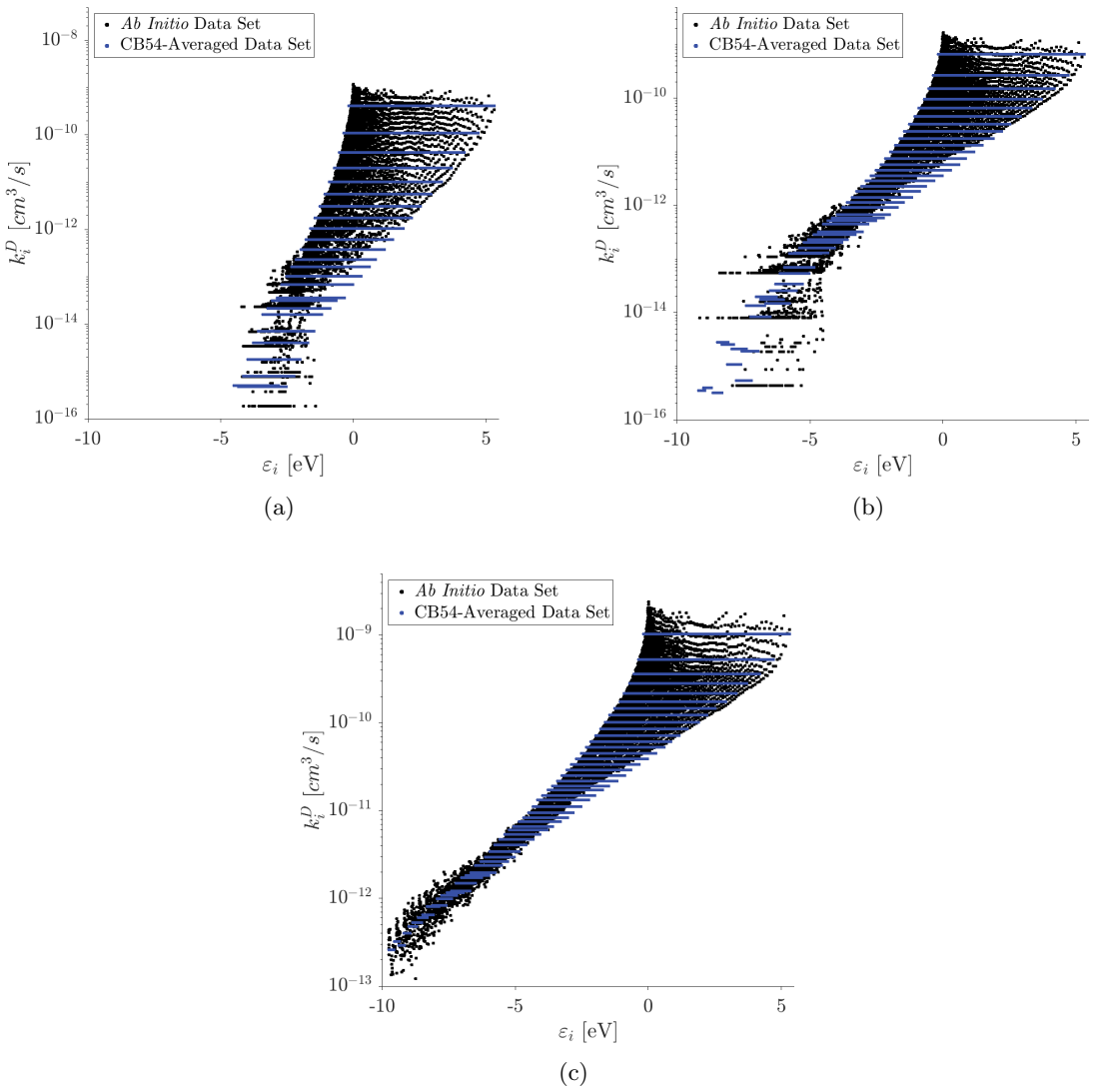


Figure B.11: Rovibrational-specific (RVS) dissociation rate coefficients at $T_{Tran} = 5\,000$ K (a), $T_{Tran} = 10\,000$ K (b), and $T_{Tran} = 20\,000$ K (c) for the N_2+O system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 54 groups for the N_2 molecule.

B.3.5 Application to NO+N System

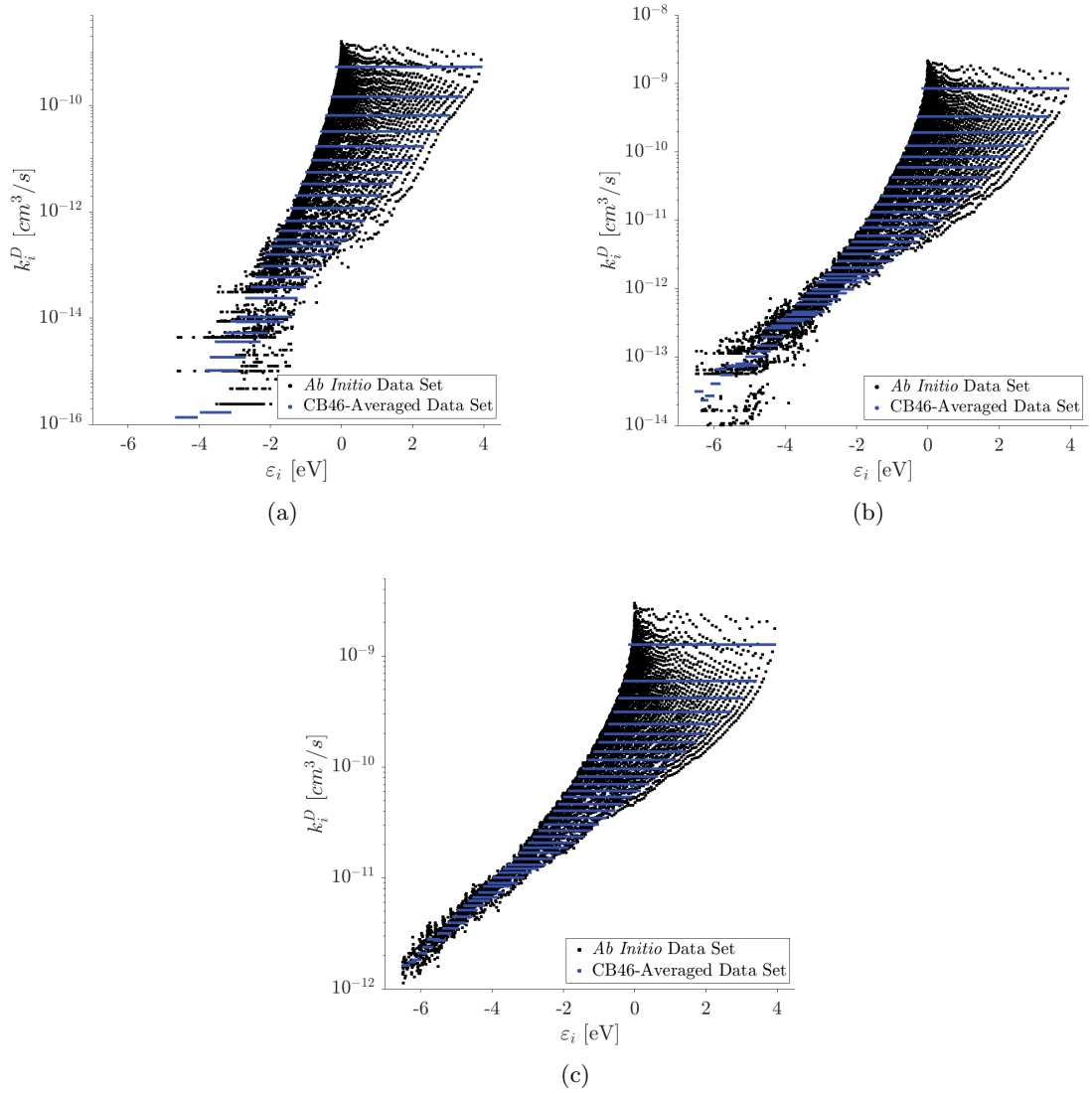


Figure B.12: Rovibrational-specific (RVS) dissociation rate coefficients at $T_{Tran} = 5\,000$ K (a), $T_{Tran} = 10\,000$ K (b), and $T_{Tran} = 20\,000$ K (c) for the NO+N system. Black dots: rates from the *ab initio* data-set. Blue dots: rates from the CB-averaged data-set, based on 46 groups for the NO molecule.

References

[1] J. D. Anderson. *Hypersonic and High Temperature Gasdynamics*. McGraw-Hill, New York, NY, 1989.

[2] D. Andrienko and I. D. Boyd. Investigation of oxygen vibrational relaxation by quasi-classical trajectory method. *Chemical Physics*, 459:1 – 13, 2015.

[3] D. A. Andrienko. The importance of O_3 excited potential energy surfaces in $O_2 - O$ high-temperature kinetics. *The Journal of Chemical Physics*, 152(4):044305, 2020.

[4] D. A. Andrienko and I. D. Boyd. Rovibrational energy transfer and dissociation in $O_2 + O$ collisions. *The Journal of Chemical Physics*, 144(10):104301, 2016.

[5] D. A. Andrienko and I. D. Boyd. Thermal relaxation of molecular oxygen in collisions with nitrogen atoms. *The Journal of Chemical Physics*, 145(1):014309, 2016.

[6] P. Atkins and J. de Paula. *Atkins' Physical Chemistry*. Oxford University Press, Oxford, 2014.

[7] E. E. Aubanel, D. M. Wardlaw, L. Zhu, and W. L. Hase. Role of angular momentum in statistical unimolecular rate theory. *International Reviews in Physical Chemistry*, 10(3):249–286, 1991.

[8] M. Ayouz and D. Babikov. Global permutationally invariant potential energy surface for ozone forming reaction. *The Journal of Chemical Physics*, 138(16):164311, 2013.

[9] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104:136403, Apr 2010.

[10] M. J. Beal. *Variational algorithms for approximate Bayesian inference*. PhD thesis, 2003.

[11] J. Behler and M. Parrinello. Generalized neural-network representation of high-dimensional potential energy surfaces. *Phys. Rev. Lett.*, 98:146401, Apr 2007.

[12] J. D. Bender, P. Valentini, I. Nompelis, Y. Paukku, Z. Varga, D. G. Truhlar, T. Schwartzentruber, and G. V. Candler. An improved potential energy surface and multi-temperature quasiclassical trajectory calculations of $N_2 + N_2$ dissociation reactions. *J. Chem. Phys.*, 143(5), 2015.

[13] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren. Neural network models of potential energy surfaces. *The Journal of Chemical Physics*, 103(10):4129–4137, 1995.

[14] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. Weight uncertainty in neural networks. *Proceedings of the 32nd International Conference on Machine Learning*, 2015.

[15] D. Bose, J. L. Brown, D. K. Prabhu, P. Gnoffo, C. O. Johnston, and B. Hollis. Uncertainty assessment of hypersonic aerothermodynamics prediction capability. *Journal of Spacecraft and Rockets*, 50(1):12–18, 2013.

[16] J. M. Bowman and B. Gazdy. A simple method to adjust potential energy surfaces: Application to hco. *The Journal of Chemical Physics*, 94(1):816–817, 1991.

[17] B. J. Braams and J. M. Bowman. Permutationally invariant potential energy surfaces in high dimensionality. *International Reviews in Physical Chemistry*, 28(4):577–606, 2009.

[18] J. E. Breen, R. B. Quy, and G. P. Glass. Vibrational relaxation of O_2 in the presence of atomic oxygen. *The Journal of Chemical Physics*, 59(1):556–557, 1973.

[19] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547–555, 2018.

[20] L. Bytautas, N. Matsunaga, and K. Ruedenberg. Accurate ab initio potential energy curve of O_2. II. core-valence correlations, relativistic contributions, and vibration-rotation spectrum. *The Journal of Chemical Physics*, 132(7):074307, 2010.

[21] M. Capitelli. *Nonequilibrium Vibrational Kinetics*, volume 39 of *Topics in Current Physics*. Springer, Berlin, 1986.

[22] M. Capitelli, I. Armenise, D. Bruno, M. Cacciatore, R. Celiberto, G. Colonna, O. D. Pascale, P. Diomede, F. Esposito, C. Gorse, and *et al.* Non-equilibrium plasma kinetics: a state-to-state approach. *Plasma Sources Science and Technology*, 16(1):S30–S44, jan 2007.

[23] G. Chaban, R. L. Jaffe, D. W. Schwenke, and W. Huo. Dissociation cross-sections and rate coefficients for nitrogen from acccurate theoretical calculations. AIAA Paper 2008–1209, 2008. 46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV.

[24] R. S. Chaudhry, J. D. Bender, T. E. Schwartzentruber, and G. V. Candler. Quasiclassical trajectory analysis of nitrogen for high-temperature chemical kinetics. *Journal of Thermophysics and Heat Transfer*, 32(4):833–845, 2018.

[25] R. S. Chaudhry and G. V. Candlero. Statistical analyses of quasiclassical trajectory data for air dissociation. AIAA Paper 2019-0789, 2019. AIAA Scitech 2019 Forum.

[26] G. Colonna, I. Armenise, D. Bruno, and M. Capitelli. Reduction of state-to-state kinetic to macroscopic models in hypersonic flows. *J. Thermophys. Heat Transfer*, 20(3):477–486, 2006.

[27] G. Colonna, F. Esposito, and M. Capitelli. The role of rotation in state-to-state vibrational kinetics. AIAA Paper 2006-3423, 2006. 9th AIAA/ASME Joint Thermophysics and Heat Transfer Conference.

[28] G. Colonna, D. Pietanza, and M. Capitelli. Recombination-assisted nitrogen dissociation rates under non-equilibrium conditions. *Journal of Thermophysics and Heat Transfer*, 22(3), July-September 2008.

[29] G. Colonna, L. D. Pietanza, and M. Capitelli. Reduced two-level approach for air kinetics in recombination regime. *AIP Conference Proceedings*, 1333(1):1365–1370, 2011.

[30] J. Cui and R. V. Krems. Gaussian process model for collision dynamics of complex molecules. *Phys Rev Lett.*, 115(7):073202, Aug 2015.

[31] R. Dawes, P. Lolur, A. Li, B. Jiang, and H. Guo. Communication: An accurate global potential energy surface for the ground electronic state of ozone. *The Journal of Chemical Physics*, 139(20):201103, 2013.

[32] R. Dawes, P. Lolur, J. Ma, and H. Guo. Communication: Highly accurate ozone formation potential and implications for kinetics. *The Journal of Chemical Physics*, 135(8):081102, 2011.

[33] O. Denis-Alpizar, R. J. Bemish, and M. Meuwly. Reactive collisions for $NO(^2\Pi) + N(^4S)$ at temperatures relevant to the hypersonic flight regime. *Phys. Chem. Chem. Phys.*, 19:2392–2401, 2017.

[34] S. Dieleman, J. Schlüter, C. Raffel, E. Olson, S. K. Sønderby, D. Nouri, D. Maturana, M. Thoma, E. Battenberg, J. Kelly, and *et al.* Lasagne: first release, Aug. 2015.

[35] D. Eriksson, K. Dong, E. H. Lee, D. Bindel, and A. G. Wilson. Scaling gaussian process regression with derivatives. In *NeurIPS*, pages 6868–6878, 2018.

[36] F. Esposito, I. Armenise, and M. Capitelli. $N - N_2$ state-to-state vibrational relaxation and dissociation rate cefficients based on quasi-classical calculations. *Chem. Phys.*, 331(1):1–8, 2006.

[37] F. Esposito and M. Capitelli. Quasiclassical trajectory calculations of vibrationally specific dissociation cross-sections and rate constants for the reaction $O + O_2(v) \rightarrow 3O$. *Chemical Physics Letters*, 364(1):180 – 187, 2002.

[38] L. S. Fletcher. *Aerodynamic Heating and Thermal Protection Systems*. American Institute of Aeronautics and Astronautics, New York, 1978.

[39] B. R. L. Galvão and A. J. C. Varandas. Accurate double many-body expansion potential energy surface for $N_3(^4A'')$ from correlation scaled ab initio energies with extrapolation to the complete basis set limit. *J. Phys. Chem. A*, 113(52):14424, 2009.

[40] E. C. Geistfeld and T. E. Schwartzentruber. Qct calculations of O_2+O collisions: comparison to molecular beam experiments. Aiaa paper, 2020. AIAA Scitech 2020 Forum, 6-10 January 2020, Orlando, FL.

[41] P. A. Gnoffo. Planetary-entry gas dynamics. *Annu. Rev. Fluid Mech.*, 31:459–494, 1999.

[42] P. A. Gnoffo, R. N. Gupta, and J. L. Shinn. Conservation equations and physical models for hypersonic air flows in thermal and chemical nonequilibrium. NASA Technical Paper 2867, 1989.

[43] S. Y. Grebenshchikov, Z.-W. Qu, H. Zhu, and R. Schinke. New theoretical investigations of the photodissociation of ozone in the hartley, huggins, chappuis, and wulf bands. *Phys. Chem. Chem. Phys.*, 9:2044–2064, 2007.

[44] M. S. Grover, T. E. Schwartzentruber, Z. Varga, and D. G. Truhlar. Vibrational energy transfer and collision-induced dissociation in $O_2 + O$ collisions. *Journal of Thermophysics and Heat Transfer*, 33(3):797–807, 2019.

[45] M. S. Grover, P. Valentini, E. Josyula, and R. S. Chaudhry. Vibrational state-to-state and multiquantum effects for $N_2 + N_2$ interactions at high temperatures for aerothermodynamic applications. AIAA Paper 2020-1227, 2020. AIAA Scitech 2020 Forum.

[46] Y. Guan, S. Yang, and D. H. Zhang. Construction of reactive potential energy surfaces with gaussian process regression: active data selection. *Molecular Physics*, 116(7-8):823–834, 2018.

[47] C. M. Handley, G. I. Hawe, D. B. Kell, and P. L. A. Popelier. Optimal construction of a fast and accurate polarisable water potential based on multipole moments trained by machine learning. *Phys. Chem. Chem. Phys.*, 11:6365–6376, 2009.

[48] K. Haug, D. G. Truhlar, and N. C. Blais. Monte carlo trajectory and master equation simulation of the nonequilibrium dissociation rate coefficient for $Ar + H_2 \leftrightarrow Ar + 2H$ at 4500 k. *The Journal of Chemical Physics*, 86(5):2697–2716, 1987.

[49] T. Ho and H. Rabitz. A general method for constructing multidimensional molecular potential energy surfaces from ab initio calculations. *The Journal of Chemical Physics*, 104(7):2584–2597, 1996.

[50] T. Hollebeek, T.-S. Ho, and H. Rabitz. Constructing multidimensional molecular potential energy surfaces from ab initio data. *Annual Review of Physical Chemistry*, 50(1):537–570, 1999.

[51] B. Hollis and D. Prabhu. Assessment of laminar, convective aeroheating prediction uncertainties for mars-entry vehicles. *Journal of Spacecraft and Rockets*, 50(1):56–68, 2013.

[52] K. Hornik, M. Stinchcombe, and H. White. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3(5):551 – 560, 1990.

[53] L. B. Ibraguimova, A. L. Sergievskaya, V. Y. Levashov, O. P. Shatalov, Y. V. Tunik, and I. E. Zabelinskii. Investigation of oxygen dissociation and vibrational relaxation at temperatures 4000–10 800 k. *The Journal of Chemical Physics*, 139(3):034317, 2013.

[54] B. Iooss and P. Lemaître. *A Review on Global Sensitivity Analysis Methods*, pages 101–122. Springer US, Boston, MA, 2015.

[55] R. L. Jaffe. Rate constants for chemical reactions in high-temperature nonequilibrium air. In *Thermophysical Aspects of Re-Entry Flows*, pages 123–151. American Institute of Aeronautics.

[56] R. L. Jaffe, M. Grover, S. Venturi, D. W. Schwenke, P. Valentini, T. E. Schwartzentruber, and M. Panesi. Comparison of potential energy surface and computed rate coefficients for N_2 dissociation. *Journal of Thermophysics and Heat Transfer*, 32(4):869–881, 2018.

[57] R. L. Jaffe, D. W. Schwenke, and G. Chaban. Vibration-rotation excitation and dissociation in N_2-N_2 collisions from accurate theoretical calculations. AIAA Paper 2010–4517, 2010. 10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Chicago, IL.

[58] R. L. Jaffe, D. W. Schwenke, G. Chaban, and W. Huo. Vibrational and rotational excitation and relaxation of nitrogen from accurate theoretical calculations. AIAA Paper 2008–1208, 2008. 46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV.

[59] R. L. Jaffe, D. W. Schwenke, M. Grover, P. Valentini, T. E. Schwartzentruber, S. Venturi, and M. Panesi. *Comparison of quantum mechanical and empirical potential energy surfaces and computed rate coefficients for N_2 dissociation.* American Institute of Aeronautics and Astronautics, 2018/06/06 2016.

[60] R. L. Jaffe, D. W. Schwenke, and M. Panesi. First principles calculation of heavy particle rate coefficients. In *Hypersonic nonequilibrium flows: fundamentals and recent advances*, pages 103–158. American Institute of Aeronautics, 2015.

[61] L. M. C. Janssen, A. van der Avoird, and G. C. Groenenboom. Quantum reactive scattering of ultracold NH($X\ ^3\Sigma^-$) radicals in a magnetic trap. *Phys. Rev. Lett.*, 110:063201, Feb 2013.

[62] B. Jiang and H. Guo. Permutation invariant polynomial neural network approach to fitting potential energy surfaces. *The Journal of Chemical Physics*, 139(5):054112, 2013.

[63] B. Jiang, J. Li, and H. Guo. Potential energy surfaces from high fidelity fitting of ab initio points: the permutation invariant polynomial - neural network approach. *International Reviews in Physical Chemistry*, 35(3):479–506, 2016.

[64] J. Kalyanaraman, Y. Kawajiri, R. P. Lively, and M. J. Realff. Uncertainty quantification via bayesian inference using sequential monte carlo methods for CO_2 adsorption process. *AIChE J.*, 62, 2016.

[65] A. Kamath, R. A. Vargas-Hernández, R. V. Krems, T. Carrington, and S. Manzhos. Neural networks vs gaussian process regression for representing potential energy surfaces: A comparative study of fit quality and vibrational spectrum accuracy. *The Journal of Chemical Physics*, 148(24):241702, 2018.

[66] M. Karplus, R. N. Porter, and R. D. Sharma. Exchange reactions with activation energy. i. simple barrier potential for (h, h_2). *The Journal of Chemical Physics*, 43(9):3259–3287, 1965.

[67] M. C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[68] J. G. Kim and I. D. Boyd. State-resolved master equation analysis of thermochemical nonequilibrium of nitrogen. *Chem. Phys.*, 415(4):237–246, 2013.

[69] J. G. Kim, O. J. Kwon, and C. Park. Master equation study and nonequilibrium chemical reactions for $H + H_2$ and $He + H_2$. *J. Thermophys. Heat Transfer*, 23(3):443–453, 2009.

[70] J. K. Kim and I. D. Boyd. State-resolved thermochemical nonequilibrium analysis of hydrogen mixture flows. *Phys. Fluids*, 24(8):086102, 2012.

[71] B. Kolb, P. Marshall, B. Zhao, B. Jiang, and H. Guo. Representing global reactive potential energy surfaces using gaussian processes. *The Journal of Physical Chemistry A*, 121(13):2552–2557, 2017. PMID: 28287725.

[72] D. Koner, R. J. Bemish, and M. Meuwly. The $C(^3P) + NO(X^2\Pi) \rightarrow O(^3P) + CN(X^2\Sigma^+), N(^2D)/N(^4S) + CO(X^1\Sigma^+)$ reaction: Rates, branching ratios, and final states from 15 k to 20 000 k. *The Journal of Chemical Physics*, 149(9):094305, 2018.

[73] D. Koner, O. T. Unke, K. Boe, R. J. Bemish, and M. Meuwly. Exhaustive state-to-state cross sections for reactive molecular collisions from importance sampling simulation and a neural network representation. *The Journal of Chemical Physics*, 150(21):211101, 2019.

[74] R. V. Krems. Bayesian machine learning for quantum molecular dynamics. *Phys. Chem. Chem. Phys.*, 21:13392–13410, 2019.

[75] A. Kucukelbir, D. Tran, R. Ranganath, A. Gelman, and D. M. Blei. Automatic differentiation variational inference. *J. Mach. Learn. Res.*, 18(1):430–474, Jan. 2017.

[76] M. Kulakhmetov, M. Gallis, and A. Alexeenko. Effect of $O_2 + O$ ab initio and morse additive pairwise potentials on dissociation and relaxation rates for nonequilibrium flow calculations. *Physics of Fluids*, 27(8):087104, 2015.

[77] S. A. Lahankar, J. Zhang, T. K. Minton, H. Guo, and G. Lendvay. Dynamics of the O-atom exchange reaction $^{16}O(^3P) + {}^{18}O{}^{18}O(^3\Sigma_g^-) \rightarrow {}^{16}O{}^{18}O(^3\Sigma_g^-) + {}^{18}O(^3P)$ at hyperthermal energies. *The Journal of Physical Chemistry A*, 120(27):5348–5359, 2016. PMID: 27043455.

[78] L. Landau and E. Teller. Theory of sound dispersion. *Phys. Z Sowjetunion*, 10(1):34–43, 1936. in German.

[79] Y. Li, Z. Sun, B. Jiang, D. Xie, R. Dawes, and H. Guo. Communication: Rigorous quantum dynamics of $O + O_2$ exchange reactions on an ab initio potential energy surface substantiate the negative temperature dependence of rate coefficients. *The Journal of Chemical Physics*, 141(8):081102, 2014.

[80] Y. Li, P. Wang, and W. Xiao. Uncertainty quantification of atomistic materials simulation with machine learning potentials. AIAA Paper 2018-2166, 2018. 2018 AIAA Non-Deterministic Approaches Conference.

[81] C. Lim and D. G. Truhlar. Study of mixture effects in the nonequilibrium kinetics of homonuclear diatomic dissociation and recombination. *The Journal of Physical Chemistry*, 88(4):778–792, 1984.

[82] C. Lim and D. G. Truhlar. The effect of vibrational-rotational disequilibrium on the rate constant for an atom-transfer reaction. *The Journal of Physical Chemistry*, 90(12):2616–2634, 1986.

[83] Q. Lin, Y. Zhang, B. Zhao, and B. Jiang. Automatically growing global reactive neural network potential energy surfaces: A trajectory-free active learning strategy. *The Journal of Chemical Physics*, 152(15):154104, 2020.

[84] W. Lin, Z. Varga, G. Song, Y. Paukku, and D. G. Truhlar. Global triplet potential energy surfaces for the $N_2(X^1\Sigma) + O(^3P) \leftrightarrow NO(X^2\Pi) + N(^4S)$ reaction. *The Journal of Chemical Physics*, 144(2):024309, 2016.

[85] Y. Liu, M. Panesi, A. Sahai, and M. Vinokur. General multi-group macroscopic modeling for thermo-chemical non-equilibrium gas mixtures. *J. Chem. Phys.*, 142(13):134109, 2015.

[86] Y. Liu, M. Vinokur, M. Panesi, and T. E. Magin. A multi-group maximum entropy model for thermo-chemical nonequilibrium. AIAA Paper 2010–4332, 2010. 10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Chicago, IL.

[87] H. Luo, I. B. Sebastião, A. A. Alexeenko, and S. O. Macheret. Classical impulsive model for dissociation of diatomic molecules in direct simulation monte carlo. *Phys. Rev. Fluids*, 3:113401, Nov 2018.

[88] R. L. Macdonald. *Reduced-order model framework for thermochemical non-equilibrium hypersonic flows*. PhD thesis, University of Illinois at Urbana-Champaign, 2019.

[89] R. L. Macdonald, M. S. Grover, T. E. Schwartzentruber, and M. Panesi. Construction of a coarse-grain quasi-classical trajectory method. II. comparison against the direct molecular simulation method. *The Journal of Chemical Physics*, 148(5):054310, 2018.

[90] R. L. Macdonald, R. L. Jaffe, D. W. Schwenke, and M. Panesi. Construction of a coarse-grain quasi-classical trajectory method. I. theory and application to N_2–N_2 system. *The Journal of Chemical Physics*, 148(5):054309, 2018.

[91] R. L. Macdonald, A. Munafò, and M. Panesi. Rovibrational grouping for $N_2(^1\Sigma_g^+) + N_2(^1\Sigma_g^+)$ energy transfer using state-to-state model. AIAA Paper 2016-4315, 2016. 46th AIAA Thermophysics Conference.

[92] S. O. Macheret, A. A. Fridman, I. V. Adamovich, J. W. Rich, and C. E. Treanor. Mechanism of nonequilibrium dissociation of diatomic molecules. AIAA Paper 1994–1984, 1994. 6th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Colorado Springs, CO.

[93] T. E. Magin, A. Bourdon, R. Jaffe, and D. W. Schwenke. Rovibrational internal energy excitation and dissociation of molecular nitrogen in hypersonic flows. AIAA Paper 2010–4336, 2010. 10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Chicago, IL.

[94] T. E. Magin, M. Panesi, A. Bourdon, R. L. Jaffe, and D. W. Schwenke. Coarse-grain model for internal energy excitation and dissociation of molecular nitrogen. *Chemical Physics*, 398:90 – 95, 2012. Chemical Physics of Low-Temperature Plasmas (in honour of Prof Mario Capitelli).

[95] G. G. Maisuradze, A. Kawano, D. L. Thompson, A. F. Wagner, and M. Minkoff. Interpolating moving least-squares methods for fitting potential energy surfaces: Analysis of an application to a six-dimensional system. *The Journal of Chemical Physics*, 121(21):10329–10338, 2004.

[96] T. K. Mankodi, U. V. Bhandarkar, and B. P. Puranik. Dissociation cross sections for $N_2 + N \rightarrow 3N$ and $O_2 + O \rightarrow 3O$ using the qct method. *The Journal of Chemical Physics*, 146(20):204307, 2017.

[97] R. Marcus. Dissociation and isomerization of vibrationally excited species. III. *The Journal of Chemical Physics*, 43(8):2658–2661, 1965.

[98] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

[99] M. Meuwly and J. M. Hutson. Morphing ab initio potentials: A systematic study of N*e*–HF. *The Journal of Chemical Physics*, 110(17):8338–8347, 1999.

[100] K. Miki, M. Panesi, E. E. Prudencio, and S. Prudhomme. Estimation of the nitrogen ionization reaction rate using electric arc shock tube data and bayesian model analysis. *Physics of Plasmas*, 19, 2012.

[101] K. Miki, M. Panesi, E. E. Prudencio, and S. Prudhomme. Probabilistic models and uncertainty quantification for the ionization reaction rate of atomic nitrogen. *Journal of Computational Physics*, 231(Issue 9):3871–3886, 2012.

[102] A. Munafò, Y. Liu, and M. Panesi. Physical models for dissociation and energy transfer in shock-heated nitrogen flows. *Phys. Fluids*, 27(12):127101, 2015.

[103] A. Munafò, N. N. Mansour, and M. Panesi. A reduced-order NLTE kinetic model for radiating plasmas of outer envelopes of stellar atmospheres. *Astrophys. J.*, 838(2):126, 2017.

[104] A. Munafò, M. Panesi, and T. E. Magin. Boltzmann rovibrational collisional coarse-grained model for internal energy excitation and dissociation in hypersonic flows. *Phys. Rev. E*, 89(2):023001, 2014.

[105] A. Munafò, S. Venturi, R. L. Macdonald, and M. Panesi. State-to-state and reduced-order models for recombination and energy transfer in aerothermal environments. AIAA Paper 2016-0505, 2016. 54th AIAA Aerospace Sciences Meeting.

[106] A. Munafò, S. Venturi, M. S. Priyadarshini, and M. Panesi. Reduced-order modeling for non-equilibrium air flows. AIAA Paper 2020-1226, 2020. AIAA Scitech 2020 Forum.

[107] A. Munafò, A. Alberti, C. Pantano, J. Freund, and M. Panesi. Modeling of laser-induced breakdown phenomena in non-equilibrium plasmas. AIAA Paper 2018-0171, 2018. AIAA Scitech 2018, 8–12 January 2018, Kissimmee, Florida.

[108] A. Munafò, A. Alberti, C. Pantano, J. B. Freund, and M. Panesi. A computational model for nanosecond pulse laser-plasma interactions. *Journal of Computational Physics*, 406:109190, 2020.

[109] R. M. Neal. *Bayesian Learning for Neural Networks*. PhD thesis, 1996.

[110] E. Nikitin, M. Kearsley, and O. U. Press. *Theory of Elementary Atomic and Molecular Processes in Gases*. Clarendon Press, 1974.

[111] W. L. Oberkampf and C. J. Roy. *Verification and Validation in Scientific Computing*. Cambridge University Press, USA, 1st edition, 2010.

[112] T. A. Oliver, G. Terejanu, C. S. Simmons, and R. D. Moser. Validating predictions of unobserved quantities. *Computer Methods in Applied Mechanics and Engineering*, 283:1310 – 1335, 2015.

[113] T.-J. Pan, T. J. Wilson, and K. A. Stephani. Vibrational state-specific model for dissociation and recombination of the $O_2(^3\Sigma_g^-) + O(^3P)$ system in DSMC. *The Journal of Chemical Physics*, 150(7):074305, 2019.

[114] M. Panesi, R. L. Jaffe, D. W. Schwenke, and T. E. Magin. Rovibrational internal energy transfer and dissociation of $N(^4S_u) + N_2(^1\Sigma_g^+)$ sytem in hypersonic flows. *J. Chem. Phys*, 138(4):044312, 2013.

[115] M. Panesi, K. Miki, S. Prudhomme, and A. Brandis. On the assessment of a bayesian validation methodology for data reduction models relevant to shock tube experiments. *Computer Methods in Applied Mechanics and Engineering*, 213-216, 2013.

[116] M. Panesi, A. Munafò, T. E. Magin, and R. L. Jaffe. Study of the non-equilibrium shock heated nitrogen flows using a rovibrational state-to-state method. *Phys. Rev. E*, 90:013009, 2014.

[117] C. Park. Assessment of a two-temperature kinetic model for dissociating and weakly ionizing nitrogen. *Journal of Thermophysics and Heat Transfer*, 2(1):8–16, 2018/06/06 1988.

[118] C. Park. *Nonequilibrium Hypersonic Aerothermodynamics*. Wiley, New York, NY, 1990.

[119] C. Park. Rotational relaxation of N_2 behind a strong shock wave. *Journal of Thermophysics and Heat Transfer*, 18(4):527–533, 2004.

[120] Y. Paukku, K. R. Yang, Z. Varga, and D. G. Truhlar. Global ab initio ground-state potential energy surface of N4. *The Journal of Chemical Physics*, 139(4):044309, 2013.

[121] A. D. Powell, N. S. Dattani, R. F. K. Spada, F. B. C. Machado, H. Lischka, and R. Dawes. Investigation of the ozone formation reaction pathway: Comparisons of full configuration interaction quantum monte carlo and fixed-node diffusion Monte Carlo with contracted and uncontracted MRCI. *The Journal of Chemical Physics*, 147(9):094306, 2017.

[122] H. O. Pritchard. Why atoms recombine more slowly as the temperature goes up. *Accounts of Chemical Research*, 9(3):99–105, 1976.

[123] M. S. Priyadarshini, S. Venturi, A. Munafò, and M. Panesi. Application of ab-initio based grouped rates for modeling non-equilibrium flow physics. AIAA Paper 2019-0792, 2019. AIAA Scitech 2019 Forum.

[124] C. Qu, Q. Yu, B. L. Van Hoozen, J. M. Bowman, and R. A. Vargas-Hernández. Assessing gaussian process regression and permutationally invariant polynomial approaches to represent high-dimensional potential energy surfaces. *Journal of Chemical Theory and Computation*, 14(7):3381–3396, 2018. PMID: 29847723.

[125] E. Quintas-Sánchez and R. Dawes. Autosurf: A freely available program to construct potential energy surfaces. *Journal of Chemical Information and Modeling*, 59(1):262–271, 2019.

[126] C. E. Rasmussen. *Gaussian Processes in Machine Learning*, pages 63–71. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[127] P. Rostkowski, S. Venturi, M. Panesi, A. Omidy, H. Weng, and A. Martin. Calibration and uncertainty quantification of vista ablator material database using bayesian inference. *Journal of Thermophysics and Heat Transfer*, pages 1–14, 2018.

[128] A. Sahai. *Reduced-order modeling of non-Boltzmann thermochemistry and radiation for hypersonic flows*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2019.

[129] A. Sahai, B. Lopez, C. O. Johnston, and M. Panesi. Adaptive coarse graining method for energy transfer and dissociation kinetics of polyatomic species. *J. Chem. Phys.*, 147:054107, 2017.

[130] A. Saltelli, K. Chan, and E. Scott. *Sensitivity Analysis*. Wiley, 2009.

[131] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck. Probabilistic programming in python using PyMC3. *PeerJ Computer Science*, 2:e55, apr 2016.

[132] J. C. San Vicente Veliz, D. Koner, M. Schwilk, R. J. Bemish, and M. Meuwly. The $N(^4S) + O_2(X^3\Sigma_g^-) \leftrightarrow O(^3P) + NO(X^2\Pi)$ reaction: thermal and vibrational relaxation rates for the $^2A'$, $^4A'$ and $^2A''$ states. *Phys. Chem. Chem. Phys.*, 22:3927–3939, 2020.

[133] G. Schmitz, I. H. Godtliebsen, and O. Christiansen. Machine learning for potential energy surfaces: An extensive database and assessment of methods. *The Journal of Chemical Physics*, 150(24):244113, 2019.

[134] T. E. Schwartzentruber, M. S. Grover, and P. Valentini. Direct molecular simulation of nonequilibrium dilute gases. *Journal of Thermophysics and Heat Transfer*, 32(4):892–903, 2018.

[135] D. W. Schwenke. Calculations of rate constants for three-body recombination of H_2 in the presence of H_2. *J. Chem. Phys.*, 89(4):2076–2091, 1988.

[136] D. W. Schwenke, R. L. Jaffe, and G. M. Chaban. Collisional dissociation of CO: ab initio potential energy surfaces and quasiclassical trajectory rate coefficients. NASA Technical Report ARC-E-DAA-TN30024, 2016. NASA Ames Research Center; Moffett Field, CA United States.

[137] M. P. Sharma, Y. Liu, and M. Panesi. Coarse-grained modeling of thermochemical nonequilibrium using the multigroup maximum entropy quadratic formulation. *Phys. Rev. E*, 101:013307, Jan 2020.

[138] O. Shatalov. Molecular dissociation of oxygen in the absence of vibrational equilibrium. *Combust Explos Shock Waves*, 9:610–613, 1973.

[139] O. Shatalov. Molecular dissociation of oxygen in the absence of vibrational equilibrium. *Combustion, Explosion and Shock Waves*, 9(5):610–613, Sep 1973.

[140] N. Singh and T. Schwartzentruber. Consistent kinetic-continuum dissociation model i: kinetic formulation, 2019. arXiv:1912.11025.

[141] D. S. Sivia. *Data Analysis: A Bayesian Tutorial*. Clarendon (Oxford Univ. Press), Oxford, 1996 (ISBN: 0-19-851762-9 or 0-19-851889-7 in paperback).

[142] R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. Society for Industrial and Applied Mathematics, USA, 2013.

[143] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.

[144] Theano Development Team. Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-prints*, abs/1605.02688, May 2016.

[145] D. G. Truhlar and J. T. Muckerman. *Reactive Scattering Cross Sections III: Quasiclassical and Semiclassical Methods*, pages 505–566. Springer US, Boston, MA, 1979.

[146] O. T. Unke, J. C. Castro-Palacio, R. J. Bemish, and M. Meuwly. Collision-induced rotational excitation in $N_2^+(^2\Sigma_g^+, v = 0)-$Ar: Comparison of computations and experiment. *The Journal of Chemical Physics*, 144(22):224307, 2016.

[147] E. Uteva, R. S. Graham, R. D. Wilkinson, and R. J. Wheatley. Active learning in gaussian process interpolation of potential energy surfaces. *The Journal of Chemical Physics*, 149(17):174114, 2018.

[148] P. Valentini, T. E. Schwartzentruber, J. D. Bender, I. Nompelis, and G. V. Candler. Direct molecular simulation of nitrogen dissociation based on an ab initio potential energy surface. *Phys. Fluids*, 27:086102, 2015.

[149] A. J. C. Varandas and A. A. C. C. Pais. A realistic double many-body expansion (dmbe) potential energy surface for ground-state O3 from a multiproperty fit to ab initio calculations, and to experimental spectroscopic, inelastic scattering, and kinetic isotope thermal rate data. *Molecular Physics*, 65(4):843–860, 1988.

[150] Z. Varga, Y. Paukku, and D. G. Truhlar. Potential energy surfaces for $O + O_2$ collisions. *J. Chem. Phys.*, 147:154312/1–17, 2017.

[151] R. A. Vargas-Hernández, Y. Guan, D. H. Zhang, and R. V. Krems. Bayesian optimization for the inverse scattering problem in quantum reaction dynamics. *New Journal of Physics*, 21(2):022001, feb 2019.

[152] S. Venturi. Physics-based stochastic framework for the quantification of uncertainty in non-equilibrium hypersonic flows. Master's thesis, Polytechnic Univ. of Milan, Milan, Italy, 2016. M.S. Thesis.

[153] S. Venturi, R. L. Jaffe, and M. Panesi. Bayesian machine learning approach to the quantification of uncertainties on ab initio potential energy surfaces. *The Journal of Physical Chemistry A*, 124(25):5129–5146, 2020.

[154] S. Venturi, A. Munafò, and M. Panesi. Modeling the contribution of rotation to ro-vibrational non-equilibrium kinetics. In *8-th International Workshop on Radiation of High Temperature Gases for Space Missions, 18-20 March 2019, Madrid, Spain*, 2019.

[155] S. Venturi and M. Panesi. *Investigating* CO *Dissociation by means of Coarse Grained Ab-Initio Rate Constants.* American Institute of Aeronautics and Astronautics, 2018/06/06 2018.

[156] S. Venturi, A. Racca, and M. Panesi. Quantifying the uncertainty on ab-initio rate coefficients by means of bayesian machine learning. In *UQOP : Uncertainty Quantification and Optimization, 18-20 Mar 2019, Paris, France*, 2019.

[157] S. Venturi, M. P. Sharma, B. Lopez, A. Munafo, and M. Panesi. CHESS database of ab initio rate coefficients, April 2020. Database, https://www.chess.aerospace.illinois.edu/databases.

[158] S. Venturi, M. P. Sharma, B. Lopez, and M. Panesi. Data-inspired and physics-driven model reduction for dissociation: Application to the $O_2 + O$ system. *The Journal of Physical Chemistry A*, 124(41):8359–8372, 2020. PMID: 32886505.

[159] S. Venturi, M. P. Sharma, and M. Panesi. A machine learning framework for the quantification of the uncertainties associated with ab-initio based modeling of non-equilibrium flows. AIAA Paper 2019–0788, 2019. AIAA Scitech 2019 Forum, 7-11 January 2019, San Diego, California.

[160] S. Venturi, M. P. Sharma, A. Racca, and M. Panesi. Effects of ab-initio potential energy surfaces on $O_2 + O$ non-equilibrium kinetics. AIAA Paper 2019-3358, 2019. AIAA Aviation 2019 Forum, 17-21 June 2019, Dallas, Texas.

[161] W. J. Vincenti and C. H. Kruger. *Introduction to Physical Gasdynamics*. Wiley, New York, NY, 1965.

[162] E. Waage and B. S. Rabinovitch. Centrifugal effects in reaction rate theory. *Chemical Reviews*, 70(3):377–387, 1970.

[163] D. Wang, W. M. Huo, C. E. Dateo, D. W. Schwenke, and J. R. Stallcop. Reactive resonances in the $N + N_2$ exchange reaction. *Chemical Physics Letters*, 379(1):132 – 138, 2003.

[164] D. Wang, J. R. Stallcop, W. M. Huo, C. E. Dateo, D. W. Schwenke, and H. Partridge. Quantal study of the exchange reaction for $N + N_2$ using an ab initio potential energy surface. *The Journal of Chemical Physics*, 118(5):2186–2189, 2003.

[165] Z. Yang K. R. and, Varga and D. G. Truhlar. *PIPFit Version 2015: A program to perform linear least-squares fitting of potential energy surfaces using permutationally invariant polynomials*. University of Minnesota, Minneapolis,, 2015.

[166] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and W. E. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Materials*, 3:023804, Feb 2019.

[167] B. Zhao and H. Guo. State-to-state quantum reactive scattering in four-atom systems. *WIREs Computational Molecular Science*, 7(3):e1301, 2017.