

Campagne 2022 Contrats Doctoraux Instituts/Initiatives

Proposition de Projet de Recherche Doctoral (PRD)

Appel à projet ISCD-Institut des Sciences du calcul des Données 2022

Intitulé du Projet de Recherche Doctoral : Swarms-of-trajectories methods for complex chemical reaction pathways in atmospheric aerosols

Directeur de Thèse porteur du projet (titulaire d'une HDR) :

NOM : **LAAGE** Prénom : **Damien**
Titre : Directeur de Recherche ou
e-mail : damien.laage@sorbonne-universite.fr
Adresse professionnelle : Département de chimie, Ecole Normale Supérieure, rue
(site, adresse, bât., bureau) Lhomond, Paris

Unité de Recherche :

Intitulé : PASTEUR
Code (ex. UMR xxxx) : UMR8640

ED388-ChimiePhysiqueChimieAnalytiq

Ecole Doctorale de rattachement de l'équipe & d'inscription du doctorant :

Doctorants actuellement encadrés par le directeur de thèse (préciser le nombre de doctorants, leur année de 1ère inscription et la qualité d'encadrement) : 3 doctorants (1^e inscription en 2019, 2020 et 2021, tous 100%)

Co-encadrant :

NOM : **SPEZIA** Prénom : **Riccardo**
Titre : Directeur de Recherche ou HDR
e-mail : riccardo.spezia@sorbonne-universite.fr

Unité de Recherche :

Intitulé : LCT
Code (ex. UMR xxxx) : UMR7616

ED388-ChimiePhysiqueChimieAnalytique ParisCer
Ou si ED non Alliance SU :

Ecole Doctorale de rattachement :

Doctorants actuellement encadrés par le co-directeur de thèse (préciser le nombre de doctorants, leur année de 1ère inscription et la qualité d'encadrement) : 2 doctorants (1^e inscription en 2020 et 2021, à 100 % et 40 %, respectivement)

Cotutelle internationale : Non Oui, précisez Pays et Université :

Description du projet de recherche doctoral (en français ou en anglais)

Context and objectives

Our project aims at developing a comprehensive and versatile simulation methodology to determine chemical reaction pathways from swarms of trajectories. Its key features will include neural network(NN)-based reactive force fields, a

transition-path sampling determination of reaction pathways at finite temperature, an explicit treatment of nuclear quantum effects (NQE) via ring polymer molecular dynamics (RPMD), and the inclusion of electronically non-adiabatic dynamics. These simulation techniques will allow elucidating a series of key questions regarding chemical reactivity in aqueous aerosols, of critical importance in atmospheric chemistry.

The 2021 report of the United Nations' International Panel on Climate Change identified the role played by aerosols in the energy exchanges of the atmosphere as the largest uncertainty for the modeling of global climate ¹. Chemical reactivity at the surface of these aerosols is now established to be dramatically different from those in the bulk and in the gas phase, and it is imperative to gain a better understanding of this interfacial reactivity to quantitatively model the impact of emissions on climate change.

Some computational studies have started to address these systems but face several limitations. We describe these challenges and how they will be solved in our project:

- Electronic structure rearrangements: these have so far been described with ab initio quantum chemistry or DFT-based molecular dynamics (MD), but their large computational cost does not allow for the extensive sampling of reaction pathways that is required. We will take advantage of recent developments in machine learning to train NN-based reactive force fields that combine a DFT-level quality with a computational cost only slightly larger than that of force-field MD.
- Sampling of reaction pathways: many atmospheric reactions, including the reaction of the Criegee intermediate to be studied here, follow a pathway that involves a complex combination of coordinates; multiple pathways can be in competition and their relative importance may depend, e.g., on the temperature conditions. We will therefore use a swarm-of-trajectories approach and sample the different pathways in an agnostic way by using transition-path-sampling.
- Nuclear quantum effects: it has been shown that important atmospheric reactions involve hydrogen atom transfers via tunneling. We will therefore describe NQE (including both tunneling and zero-point energy effects) via a RPMD approach.
- Photochemical reactions: many atmospheric reactions involve a photoexcitation to an electronically excited state, and the reaction's products sensitively depend on the branching ratio during the excited-state relaxation dynamics. We will describe non-adiabatic transitions between electronic states via a surface-hopping approach. This method involves swarms of trajectories propagated with a stochastic description of transitions. It has recently been combined with RPMD to include NQE.

Selected atmospheric chemistry questions

Our computational approach will be used to elucidate the following key questions for atmospheric chemistry.

Aerosol nucleation and growth. The growth of aerosols at the early stages following gas-liquid nucleation is key to understand the stability and size distribution of these particles, but is poorly described by classical nucleation theory, especially in the presence of acidic defects ². Recent experimental work ³ has identified the key role that nitric acid plays in the anomalously large growth rates of airborne particles above polluted macro-cities. The models usually employed to rationalize such results are based on macroscopic thermodynamic parameters, and lack molecular-scale information on the growth mechanism. A challenge for the simulation of aerosol growth is the need to describe the progressive shift in acid dissociation equilibrium when the aerosol size increases. We will combine NN-based reactive MD simulations with enhanced sampling algorithms (e.g. aggregation-volume bias Monte Carlo) to obtain aerosol growth free energy profiles for the paradigm case of water/nitric acid

aerosols. The calculation of free energy profiles for increasing HNO₃ concentrations will provide an improved understanding of the impact of pollution on aerosol growth and rationalize the abundant experimental results. Furthermore, we will extend our approach to examine the kinetics of aerosol growth: classical nucleation theory is often formulated in terms of collision/desorption event kinetics, and we will propagate a swarm of non-equilibrium trajectories generated from a distribution of initial conditions to study the successful uptake of a gas-phase molecule by the aerosol upon collision.

Reactivity of Criegee intermediates. Criegee intermediates are highly reactive carbonyl oxides that play a key role in atmospheric ozonolysis of unsaturated hydrocarbons. Their reactivity is radically different at the surface of aqueous aerosols than in the gas phase. Their reaction with water and with nitric acid has been shown to involve multiple possible reaction pathways⁴, and hydrogen transfer via tunneling has been suggested⁵ to be important. We will use a transition-path sampling approach with RPMD and a NN reactive potential to determine the dominant reaction pathway and study how this pathway is affected by temperature.

Photodissociation of nitric acid. A recent study in the Laage group explained the experimental observation that nitric acid is less dissociated at the aqueous aerosol interface than in the bulk. This has important consequences for the photodissociation of nitric acid in the atmosphere. We will determine and contrast the photodissociation mechanisms and dynamics of undissociated nitric acid and of nitrate ion adsorbed on aerosols⁶. These mechanisms have remained elusive and have been suggested to involve intersystem crossing between excited singlet and triplet states⁷. We will therefore extend the machine learning force field approach to train separate NN in the ground, excited singlet and triplet states, and we will include non-adiabatic transitions via a surface-hopping scheme. Swarms of trajectories initiated from ground state equilibrium conditions will be propagated to obtain the branching ratios and establish the photodissociation mechanism.

Computational methodology

While some of the simulation features needed for our project are already available in codes that we have successfully used, there is presently no consistent computational framework that includes all the required features. We will therefore combine pre-existing but independent codes, and implement the missing features. The existing building blocks are:

- DeePMD-kit⁸ to train NN-based reactive force fields on reference DFT calculations performed with CP2K
- i-PI⁹ to perform RPMD
- OpenPathSampling¹⁰ library for transition path sampling simulations
- Newton-X¹¹ library for non-adiabatic surface-hopping dynamics
- VENUS¹² to generate ensembles of trajectories for collisions (the Spezia group is the code main support team)
- GraphTheory package to analyze trajectories ensembles developed by the Spezia group¹³

One main outcome of the project will be to build up an environment which will be able to handle and couple with these tools which will be useful also for future projects in the broad area of simulating chemical reactions.

Relevance to ISCD

A common feature to the different parts of our project is the need to run swarms of trajectories: transition path sampling for reaction pathways, surface-hopping for non-adiabatic transitions, and non-equilibrium for collisions. All the trajectories will be propagated using NN potential force fields. This computationally intensive approach

will generate a substantial amount of data that will be analyzed to obtain key physical chemical information. The federative context of ISCD therefore provides an ideal environment for the development of these novel techniques. The applicability of the proposed approach thus extends much beyond the specific questions addressed here and will be valuable for future applications in the ISCD context and for national and international grant applications.

PhD advisors

Damien Laage is the coordinator of the Theoretical Chemistry group at PASTEUR. He will provide his expertise on the following aspects:

- NN-based reactive force fields: the Laage group has developed reactive force fields using deep machine learning with DeePMD for systems ranging from proton diffusion to acid dissociation and peptide bond formation
- impact of solvation on chemical reactions: the Laage group has a recognized expertise in the molecular description of solvation's impact on chemical reactions, including at aqueous interfaces and in complex environments ¹⁴
- ring-polymer molecular dynamics (RPMD): the Laage group has already applied RPMD run with i-PI to study NQE in liquid water ¹⁵, and has implemented its combination with machine-learning force fields
- transition path sampling: the Laage group has recently combined the OpenPathSampling python library with NN-based molecular dynamics to run swarm of trajectories and determine reaction pathways

Riccardo Spezia is the leader of the "Dynamics Simulations: Structure and Reactivity" group at LCT. He will provide his expertise in:

- non-adiabatic surface-hopping dynamics: RS pioneered ¹⁶ the study of solvation effects on non-adiabatic surface hopping dynamics and showed the solvent's impact on excited state dynamics via its friction on the reaction coordinate and its distortion of the potential energy surface. These two aspects will be critical here to determine the effect of the aqueous aerosol interface environment on reactions.
- collisional reactions: RS has a considerable expertise ¹⁷ in reactive collisional dynamics based on ensembles of out-of-equilibrium trajectories with a distribution of initial collisions. This will be key to understand the growth of aerosols.
- data analysis: RS has developed ¹³ a graph theory approach to obtain relevant structural and dynamical information in an automatic fashion from large sets of trajectories. This will be used to analyze all our swarms of trajectories.

Student profile

This project will require a student with a strong background in physical chemistry and/or molecular simulations. Prior experience with MD, machine-learning based force fields, and/or non-adiabatic dynamics will be a plus.

References

(publications by the PhD advisors highlighted in blue)

- (1) Masson-Delmotte, V. et al. *IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: 2021;
- (2) Seinfeld, J. H.; Pandis, S. N. *Atmospheric Chemistry and Physics*; John Wiley & Sons: 2016; pp 1152.
- (3) Wang, M. et al. *Nature* **2020**, *581*, 184-189.
- (4) Zhu, C. et al. *J Am Chem Soc* **2016**, *138*, 11164-11169.
- (5) Long, B.; Bao, J. L.; Truhlar, D. G. *J Am Chem Soc* **2016**,
- (6) Anglada, J. M.; Martins-Costa, M. T. C.; Francisco, J. S.; Ruiz-López, M. F. *J Am Chem Soc* **2021**, *143*, 453-462.
- (7) Svoboda, O.; Slaviček, P. *J Phys Chem Lett* **2014**, *5*, 1958-1962.
- (8) Wang, H.; Zhang, L.; Han, J.; Weinan, E. *Comput Phys Commun* **2018**, *228*, 178-184.
- (9) Ceriotti, M.; More, J.; Manolopoulos, D. E. *Comp Phys Comm* **2014**, *185*, 1019-1026.
- (10) Swenson, D. W. H.; Prinz, J. H.; Noe, F.; Chodera, J. D.; Bolhuis, P. G. *J Chem Theory Comput* **2019**, *15*, 813-836.
- (11) Barbatti, M. et al. *WIREs Computational Molecular Science* **2014**, *4*, 26-33.
- (12) <https://www.depts.ttu.edu/chemistry/Venus/index.php>
- (13) Perez-Mellor, A. F.; Spezia, R. *J Chem Phys* **2021**, *155*, 124103.
- (14) Maffucci, I.; Laage, D.; Sterpone, F.; Stirmemann, G. *Chemistry* **2020**, *26*, 10045-10056.
- (15) Wilkins, D. M.; Manolopoulos, D. E.; Pipolo, S.; Laage, D.; Hynes, J. T. *J Phys Chem Lett* **2017**, *8*, 2602-2607.
- (16) Malhado, J. P.; Spezia, R.; Hynes, J. T. *J Phys Chem A* **2011**, *115*, 3720-3735.
- (17) Macaluso, V.; Homayoon, Z.; Spezia, R.; Hase, W. L. *Phys Chem Chem Phys* **2018**, *20*, 19744-19749.

