Machine learning interatomic potentials with accurate long-range interactions for molecular dynamics collision simulations of atmospherically-relevant molecules Ivo Neefjes1 , Jakub Kubecka ˇ 2 , and Jonas Elm1

Overall

Calculating the capture rate requires the long-range potential energy surface (PES) to be accurately defined. This paper uses molecular dynamics simulations (trajectory calculations) to explore the capture rate coefficient for three systems:

H2SO4 + H2SO4; binding energy ~ 66 kJ mol^{-1}

H2SO4 + NH(CH3)2; binding energy ~ 71 kJ mol^{-1}

H2SO4 + $HSO4^{-}$: binding energy ~ 125 kJ mol-1^{-1}

2 Machine learning architectures have been used to explore the PES for these systems based on data from molecular dynamics simulations at 2 levels of theory, semiempirical and DFT.

Comparison is made between the Machine Learning methods and how things need to be improved to calculated accurate capture rates. All the capture rate coefficients are fast at 300 K.

General criticism

This paper is mainly about the method of calculating the capture rate coefficient rather than the importance of the capture rate coefficients under consideration. If I wanted to calculate these capture rate coefficients, would it not be easier to do a high-level PES, much higher level than in this study, and then use capture rate theory to calculate the capture rate coefficient?

An even more important point is that these capture rate coefficients are not the values for atmospheric chemistry. The calculated capture rate coefficient is for infinite pressure and I’m very confident that the systems

explored here are nowhere near the high-pressure limit. The value required for atmospheric chemistry requires a Master Equation calculation using the derived PES for the system. It is also likely, the adducts are too weakly bound for them to not re-dissociate back to reagents at 300 K. Therefore, the equilibrium coefficient for these systems might also be important. You have not provided the atmospherically relevant rate coefficient and the equilibrium coefficient for the system.

The above two points raises the question of if ACP is the right journal for this paper. I do not think it is.

Additionally, can you really explore the long-range potential with such a low level of theory?

Is it easy to calculate the capture rate coefficient over a range of temperatures? Only 300 K is presented.

Overall

For the reasons I outline above I do not think ACP is the right journal for this paper; these reactions are going to be pressure dependent and a long way from the capture rate coefficient at atmospheric pressure. To make this an ACP paper more work is required in order to calculate the atmospherically relevant rate coefficients, including the equilibrium coefficient.

Line-by-Line

25 Most atmospheric aerosol particles form through a gas-to-particle conversion process called new particle formation (NPF),

Are the collisions you are considering in this study important for NPF? I think it is later collisions that are rate determining, not the first step collisions considered in the present system.

further through condensation and coagulation (Zhang et al., 2012). The earliest stages of this formation are inherently dynamic: molecules approach each other under the influence of long-range attractive forces (e.g., van der Waals or electrostatic interactions), then rearrange and relax to accommodate one another while forming a thermodynamically stable cluster.

30 Cluster-forming collisions are quantified by a collision rate coefficient, which represents the frequency of collisions per unit concentration. In practical applications, such as simulations using the Atmospheric Cluster Dynamics Code (ACDC) (McGrath et al., 2012), this coefficient is typically calculated using kinetic gas theory. In this framework, colliding partners are

The capture rate coefficient is appropriate at the high-pressure limit. A lot of associations – including the systems in this paper – are going to be pressure dependent at atmospheric pressure and will be much smaller than the capture limit. A Master Equation calculation is required.

leau, 1977; Knattrup et al., 2025; Kubečka et al., 2025). However, properties derived from such methods can exhibit significant quantitative—and even qualitative—errors. Furthermore, because the computational cost of GFN1-xTB scales cubically with

What sort of errors. Do you just mean rate coefficients?

their computational cost makes even short MD simulations of small systems prohibitively expensive.

Can MD simulations be run long enough to be relevant for more atmospheric reactions? Are you stuck at pico-second processes or shorter? For instance, adding H2O to sulfuric acid dimer is important and the question is how many waters are added (together with further reactions) can MD tackle this problem, or is the timescale to long for MD?

If you have an accurate PES from a much higher-level QM calculation, can this PES be used to calculate an accurate rate coefficient via capture rate theory?

2.1 Collision systems

We investigated three collision systems containing the atmospherically relevant species

The systems considered can be calculated at a much higher level via standard *ab initio* techniques. Would it not be better to do this as well, in order to really assign the accuracy of the present calculations?

Can the T be lowered and still work? Is 100 K possible?

In general, the models achieve excellent accuracy. All MAEs fall below the standard chemical accuracy thresholds of 1 kcal mol^{-1} for energies and 1 kcal mol^{-1} Å^{-1} for forces. Notably, for the neutral H_2SO_4–H_2SO_4 and H_2SO_4–$NH(CH_3)_2$ systems, the errors are exceptionally low, with MAEs below 0.1 kcal mol^{-1} and 0.1 kcal mol^{-1} Å^{-1}. The Δ-PaiNN methodology proves particularly effective. It correctly predicts negligible corrections in the GFN1-xTB sanity check, and consistently

This is not absolute chemical accuracy. What are you comparing against to state errors below 0.1 kCal mol-1?