Simulation of ozonolysis reactions by the master equation approach and adaption of the code for use on a supercomputer

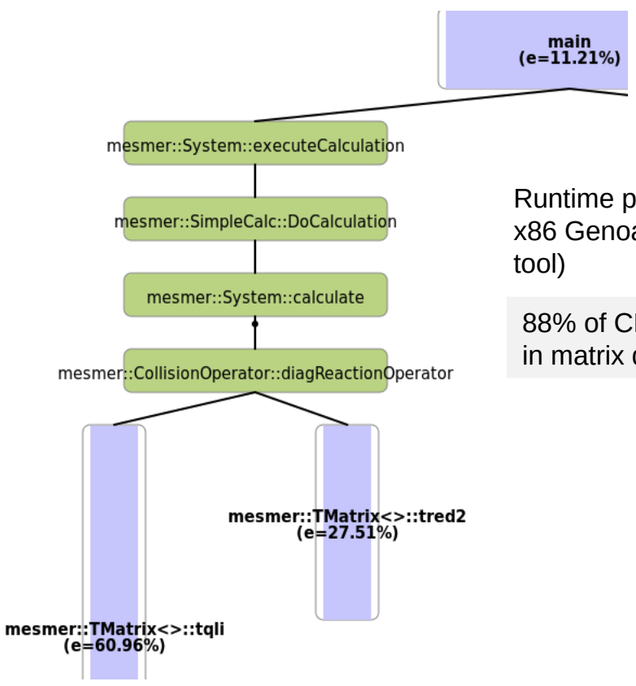
N.Derbel[1,2], A. Alijah[3], S. H. Robertson[4], L. Joly[3], T. Lauvaux[3], P. Bousquet-Mélou[5]
[1] Laboratoire de Spectroscopie Atomique Moléculaire et Applications, Faculté des Sciences de Tunis - Tunisie
[2] Faculté des Sciences de Bizerte, Université de Carthage - Tunisie
[3] Groupe de Spectrométrie Moléculaire et Atmosphérique, Université de Reims Champagne-Ardenne - France
[4] Dassault-Systèmes – Royaume Uni
[5] Centre Régional Informatique et d'Applications Numériques de Normandie (CRIANN) - France

Plants in a situation of stress caused by polluted environments emit chemical compounds known as terpenes, which are then decomposed by molecules present in smog such as ozone. Our aim is to identify the decomposition products as these may serve as biomarkers for oxidative stress.

The reaction mechanism of ozonolysis is highly complex (1, 2), and the resulting coupled differential equations are difficult to solve by conventional methods. We therefore used the free master equation solver MESMER developed by Robertson and colleagues (3) which treats chemical reactions as collisional energy transfers. It is capable of solving the energy grained master equation for a unimolecular system composed of an arbitrary number of wells, transition states, sinks, and reactants. **However it initially required a very long CPU time**, up to several days, for our reaction system, and sometimes it even fails because of the stiffness of the differential equations. Furthermore, the code was not meant to be applied to such a complex system and is not adapted to parallel execution, except for the embarrassingly parallel mode if solutions at several temperatures are sought.

HPC analysis enabling first performance breakthrough

Starting point: some runs did not converge with the maximal batch partition duration (300 h) of CRIANN's HPC resources



Runtime profile of initial code on x86 Genoa (HPE CrayPat analysis tool)

88% of CPU time spent in matrix diagonalization

1. Critical innermost loop not vectorizable (obstructive function call), but made parallel by OpenMP threads

```
NEC C/C++ Compiler (4.0.0) for Vector Engine

LINE          DIAGNOSTIC MESSAGE

 417: vec( 103): Unvectorized loop.
 417: vec( 110): Vectorization obstructive function reference.: renorm

#pragma omp parallel for private(k,km1,f) shared(n,z,s,c,i)

417- | | |+--->       for (k = 1; k <= n; ++k) {
418- | | | |             km1 = k - 1;
419- | | | |             f = z[km1][i];
420- | | | |             z[km1][i] = s*z[km1][i - 1] + c*f;
421- | | | |             z[km1][i - 1] = c*z[km1][i - 1] - s*f;
422- | | |+---         }
```

The NEC SX-Aurora vectorial architecture (Boreale machine of MesoNET at CRIANN) was exploited for compiler vectorization diagnosis. The scalar nature of the code profile made us target a general-purpose processor architecture (x86), AMD Genoa from the Austral cluster at CRIANN.

2. Given the diagonalization algorithm has a CPU-bound profile, gfortran optimization options (-O3) have been tested and numerically validated: 22% gain, and the Genoa processor turbo boost has been activated at runtime: 40% gain (Slurm directive interpreted by its prologue on Austral)

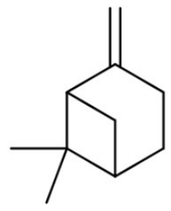
Performance tests (elapsed time) with modified first part of the scheme (dimension of matrix 1504): Option -O2: 2576 s, option -O3 without and with turbo boost: 1552 s and 1209 s, option -O3 with 8 OpenMP threads: 611 s

Overall performance gain (keeping initial physical model): > 4 factor

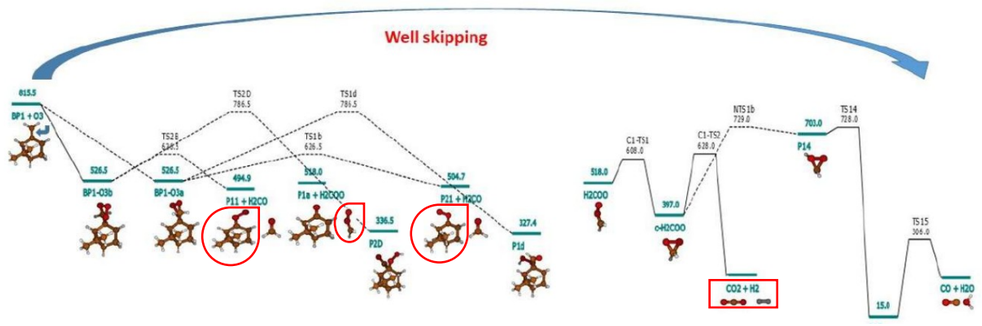
Physical model improvement

Given the breakthrough achieved in high-performance computing with the initial physical model, the motivation to further reduce CPU time was reinforced.

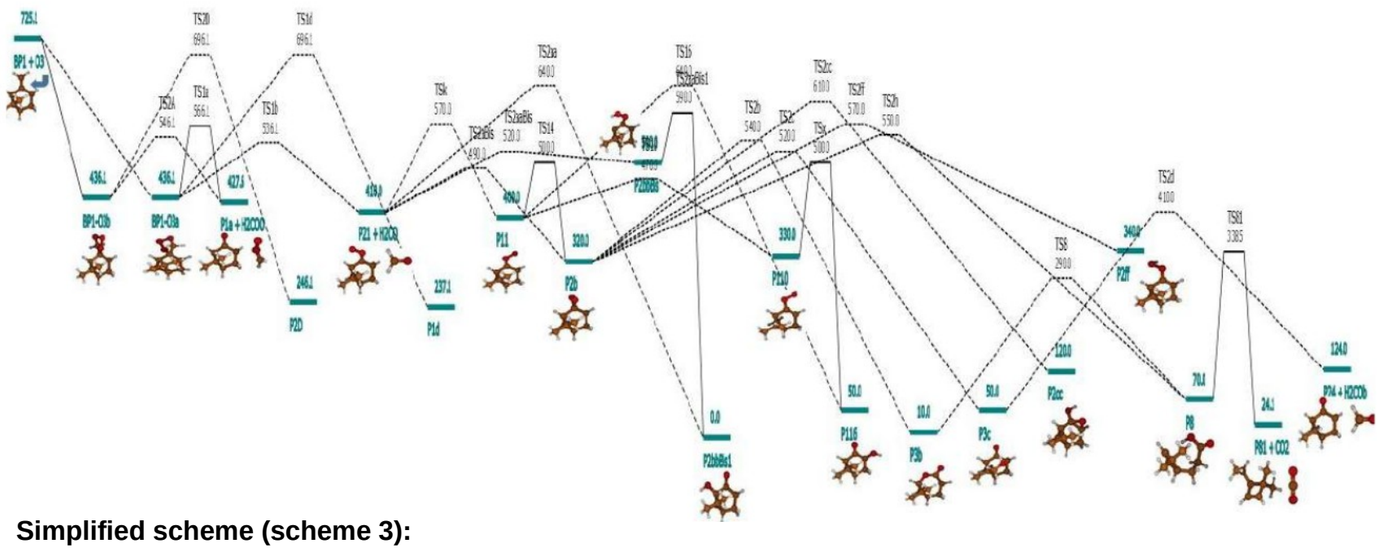
As an example we consider the ozonolysis of β-pinene, $C_{10}H_{16}$ Ozone breaks up the C=C double bond, leading to highly reactive so-called Criegee intermediates (P11, P21, H_2COO). The full scheme and its energy profile is,



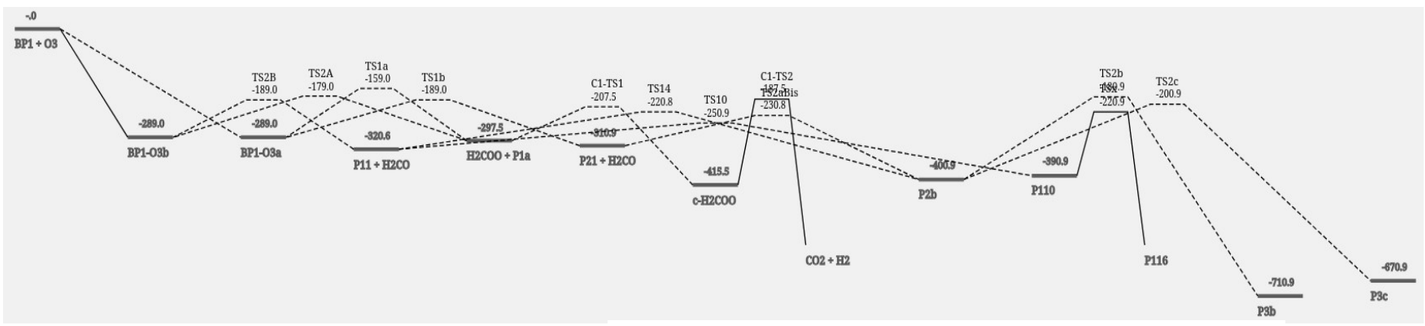
Part1:



Part 2:

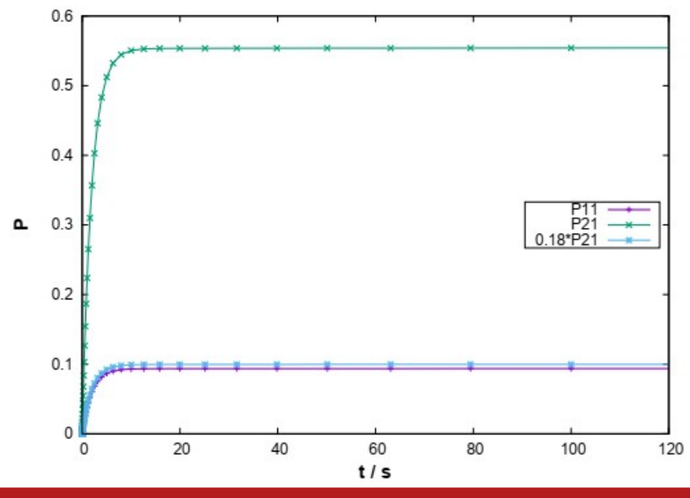


Simplified scheme (scheme 3):



Formation probabilities of Criegee intermediates P11 and P21 as function of time

Excellent agreement with experiment (4): [P11]/[P21] = 0.18 at *T*=298K



Conclusion

- Integration of the complete scheme, parts 1 and 2, is numerically unstable even at high-temperature conditions. Removal of slow reactions leads to scheme 3 which can be treated efficiently with the optimized code
- Efficient MESMER computations for large systems can be performed on super computers

References

(1) N. Derbel, C. Kalalian, A. Alijah, S. H. Robertson, A. Chakir and E. Roth, "Ozonolysis of 2-Methyl-2-pentenal: New Insights from Master Equation Modeling". J. Phys. Chem. A 128, 2534 (2025),
(2) N. Derbel, A. Alijah, S. H. Robertson, T. Lauvaux and L. Joly, "First-Generation Prod- ucts of Trans-2-Hexenal Ozonolyis: A New Look at the Mechanism". J. Phys. Chem. A 129, 3272 (2025)
(3) D. R. Glowacki, C.-H. Liang, C. Morley, M. J. Pilling and S. H. Robertson, "MESMER: An Open-Source Master Equation Solver for Multi-Energy Well Reactions". J. Phys. Chem. A 116, 9545 (2012)
(4) J. Ahrens, P. T. M. Carlson, N. Hertl, M. Olzmann, M. Pfeifle, J. L. Wolf and T. Zeuch, "Infrared detection of Criegee Intermediates …", Ang. Chem. Int. Ed. 55, 715 (2014)

Acknowledgements

Computer time was provided by the Regional HPC center of Normandy (CRIANN).
We are grateful for financial support from PHC Utique, project number 24G1301.
This work was supported by a French government grant managed by the Agence Nationale de la Recherche under the "Investissements d'avenir" program (reference "ANR-21-ESRE-0051").