

MACE: a Machine-learning Approach to Chemistry Emulation

Silke Maes^{1¶}, Frederik De Ceuster^{1,5}, Marie Van de Sande^{2,3}, and Leen Decin^{1,4}

¹ Institute of Astronomy, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium ² Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands ³ School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, United Kingdom ⁴ School of Chemistry, University of Leeds, Leeds LS2 9JT, United Kingdom ⁵ Leuven Gravity Institute, KU Leuven, Celestijnenlaan 200D, Leuven, Belgium ¶ Corresponding author

DOI: [10.21105/joss.07148](https://doi.org/10.21105/joss.07148)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: [Josh Borrow](#) ↗ 

Reviewers:

- [@richings](#)
- [@trappitsch](#)

Submitted: 10 June 2024

Published: 04 April 2025

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Astrochemistry is the study of chemical species in astrophysical objects. By combining observations and theoretical predictions, the physical conditions of the object can be analysed as well as its chemical composition and yields, since chemistry is closely coupled to the dynamics of the environment. Most often, such astrophysical environments are known to be complex and out of thermodynamic equilibrium. For that reason, the chemical evolution is usually modelled using a chemical kinetics approach, in which a set of non-linear coupled ordinary differential equations (ODEs) is solved for a given network of chemical species and reactions. For large chemical networks, however, this method is computationally slow. Therefore, we developed MACE, a *Machine-learning Approach to Chemistry Emulation*, inspired by similar work in the literature, e.g., Holdship et al. (2021); Grassi et al. (2022); Sulzer & Buck (2023). Their code is publicly and can be found in their respective papers. However, not all of their code can be easily adapted or used as such.

MACE is a PyTorch module that offers a trainable surrogate model that is able to emulate chemical kinetics in only 5% of the computation time of its classical analogue, in the current implementation. More speed up is expected when using GPUs instead of only CPUs. MACE provides a machine-learning architecture, consisting of an autoencoder and a trainable ODE, implemented in PyTorch and [torchode](#) (Lienen & Günnemann, 2022). Mathematically, MACE is given by

$$\hat{n}(t) = \mathcal{D}\left(g(\mathcal{E}(\mathbf{n}, \mathbf{p}), t)\right), \quad (1)$$

where $\hat{n}(t)$ represents the predicted chemical abundances at time t . The symbols \mathbf{n} and \mathbf{p} represent the chemical abundances and physical parameters, respectively. The autoencoder consists of an encoder \mathcal{E} , able to map \mathbf{n} and \mathbf{p} to a latent space representation \mathbf{z} , and decoder \mathcal{D} , working the other way around. G is a function that describes the evolution of the latent variables such that $\mathbf{z}(\Delta t) = G(\mathbf{z}, \Delta t) = \int_0^{\Delta t} g(\mathbf{z}) dt$, where g represents the trainable ODE. A schematic representation of the architecture and flow of the emulator is given in Fig. [Figure 1](#) for a time step Δt . More details can also be found in Maes et al. (2024). MACE is developed in the context of circumstellar envelopes (CSEs) of evolved stars, but is flexible to be retrained on data of other astrophysical environments.

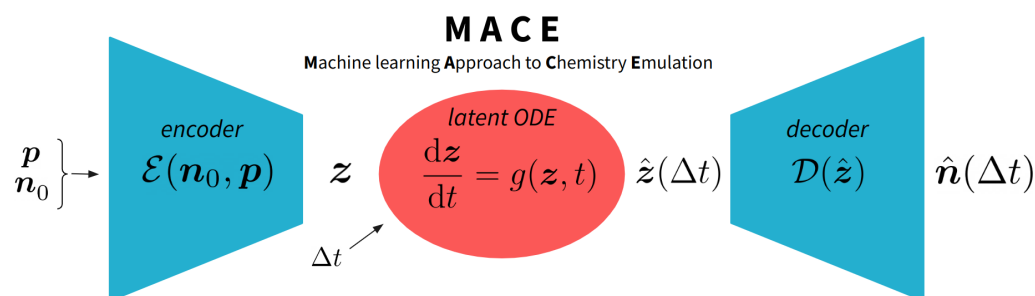


Figure 1: Schematic representation of the architecture of MACE. The autoencoder is shown in blue, the time evolution happens in the red part of the emulator. (Adapted from [Maes et al., 2024](#).)

Statement of need

Astrophysical objects or environments entail 3D dynamical processes, as well as chemical and radiation processes. Thus, in order to properly model such an environment, these three constituents need to be included in a single simulation. However, modelling each constituents separately already poses a computational challenge for a modest parameter space, let alone running a simulation with the three constituents integrated, and on an extended parameter space. Specifically on the chemistry side, it is generally established that, from a mathematical point of view, chemical kinetics is a stiff problem. When working with an extended chemical network ($\gtrsim 100$ species, connected by $\gtrsim 1000$ reactions), the system becomes computationally very slow to solve. Therefore, it is essential to speed up the solver in an alternative way, also because the solver is applied multiple times when calculating the chemical evolution of a certain environment. A fast and accurate chemistry solver would make it computationally feasible to run it alongside a 3D hydrodynamical and/or radiation model.

Example case

In the context of circumstellar envelopes, we aim to build a comprehensive model, consisting of 3D hydrodynamics, chemistry, and radiation. This is needed since recently, complex structure (such as spiral, arcs, and bipolarity) are observed in CSEs around evolved, low- and intermediate mass stars, currently attributed to the presence of (sub-)stellar companions ([Decin et al., 2020](#); [Gottlieb et al., 2022](#); [Nordhaus & Blackman, 2006](#)). In order to study these stars and their CSEs, taking into account companions, accurate 3-dimensional models are needed. Currently, only hydrodynamical modelling is done for these systems in 3D (e.g., [Maes et al., 2021](#); [Malfait et al., 2021](#)), though the aim is to include chemistry as well as radiation ([Maes et al., 2022](#)). Starting with integrating chemistry and 3D hydrodynamics, this coupling will only be computationally feasible when a faster alternative to chemical kinetics exists.

The development of MACE is a first step towards this coupling. Using the MACE architecture, we are able to accurately reproduce the 1D abundance profiles of 468 chemical species. In order to verify this, we defined the following error metric:

$$\text{error} = \frac{\log_{10} n - \log_{10} \hat{n}}{\log_{10} n}, \quad (2)$$

which is executed element-wise and subsequently summed over the different chemical species to come to a global error on a MACE model. More details on this error metric and the accuracy of the MACE models can be found in [Maes et al. \(2024\)](#). For MACE models with different sets of hyperparameters, a global error between 10 and 18 was found, averaged over 3000

test datasets, where a lower number indicates a better model. [Figure 2](#) shows the abundance profiles of seven chemical species, where the full curves indicate the MACE test of model *int4* from Maes et al. (2024), and the dashed curves give the result for the classical model. On average, the MACE routine provides a speed-up of a factor 26, not taking into account the potential extra speed-up factor due to efficient vectorisation when coupling MACE with an SPH hydrodynamical model, or making it compatible with GPU calculations.

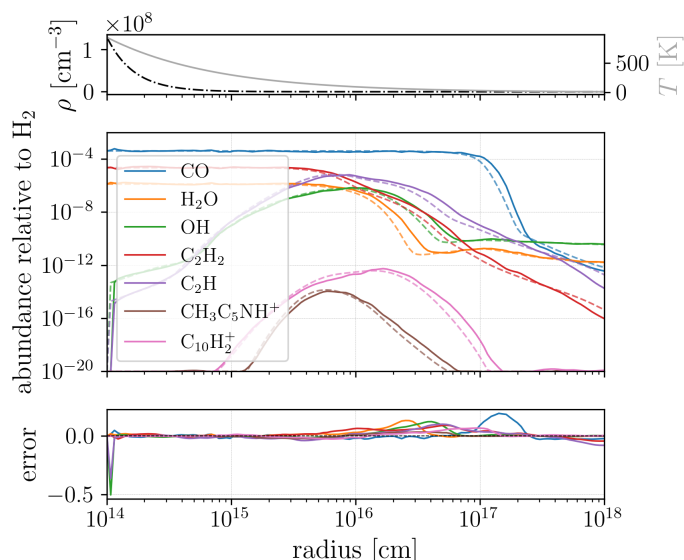


Figure 2: Chemical abundance profiles from a test of trained MACE model model *int4* (full curves), compared to the classical model (dashed curves). The error on the MACE model is calculated according to [Equation 2](#) per chemical species. More details in Maes et al. (2024).

Acknowledgements

S.M. and L.D. acknowledge support from the Research Foundation Flanders (FWO) grant G099720N. F.D.C. is a Postdoctoral Research Fellow of the Research Foundation - Flanders (FWO), grant number 1253223N, and was previously supported for this research by a Postdoctoral Mandate (PDM) from KU Leuven, grant number PDMT2/21/066. M.V.d.S. acknowledges support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 882991 and the Oort Fellowship at Leiden Observatory. L.D. also acknowledges support from KU Leuven C1 MAESTRO grant C16/17/007, KU Leuven C1 BRAVE grant C16/23/009, and KU Leuven Methusalem grant METH24/012.

References

- Decin, L., Montargès, M., Richards, A. M. S., Gottlieb, C. A., Homan, W., McDonald, I., El Mellah, I., Danilovich, T., Wallström, S. H. J., Zijlstra, A., Baudry, A., Bolte, J., Cannon, E., De Beck, E., De Ceuster, F., de Koter, A., De Ridder, J., Etoka, S., Gobrecht, D., ... Yates, J. (2020). (Sub)stellar companions shape the winds of evolved stars. *Science*, 369(6510), 1497–1500. <https://doi.org/10.1126/science.abb1229>
- Gottlieb, C. A., Decin, L., Richards, A. M. S., De Ceuster, F., Homan, W., Wallström, S. H. J., Danilovich, T., Millar, T. J., Montargès, M., Wong, K. T., McDonald, I., Baudry, A., Bolte, J., Cannon, E., De Beck, E., de Koter, A., El Mellah, I., Etoka, S., Gobrecht, D., ...

- Zijlstra, A. (2022). ATOMIUM: ALMA tracing the origins of molecules in dust forming oxygen rich M-type stars. Motivation, sample, calibration, and initial results. *Astronomy & Astrophysics*, 660, A94. <https://doi.org/10.1051/0004-6361/202140431>
- Grassi, T., Nauman, F., Ramsey, J. P., Bovino, S., Picogna, G., & Ercolano, B. (2022). Reducing the complexity of chemical networks via interpretable autoencoders. *Astronomy & Astrophysics*, 668, A139. <https://doi.org/10.1051/0004-6361/202039956>
- Holdship, J., Viti, S., Haworth, T. J., & Ilee, J. D. (2021). Chemulator: Fast, accurate thermochemistry for dynamical models through emulation. *Astronomy & Astrophysics*, 653, A76. <https://doi.org/10.1051/0004-6361/202140357>
- Lienen, M., & Günnemann, S. (2022). Torchode: A parallel ODE solver for PyTorch. *The Symbiosis of Deep Learning and Differential Equations II, NeurIPS*. <https://openreview.net/forum?id=uiKVKTiUYB0>
- Maes, S., De Ceuster, F., Van de Sande, M., & Decin, L. (2024). MACE: A Machine-learning Approach to Chemistry Emulation. *The Astrophysical Journal*, 969(2), 79. <https://doi.org/10.3847/1538-4357/ad47a1>
- Maes, S., Homan, W., Malfait, J., Siess, L., Bolte, J., De Ceuster, F., & Decin, L. (2021). SPH modelling of companion-perturbed AGB outflows including a new morphology classification scheme. *Astronomy & Astrophysics*, 653, A25. <https://doi.org/10.1051/0004-6361/202140823>
- Maes, S., Siess, L., Homan, W., Malfait, J., De Ceuster, F., Ceulemans, T., Donné, D., Esseldeurs, M., & Decin, L. (2022). Route towards complete 3D hydro-chemical simulations of companion-perturbed AGB outflows. In L. Decin, A. Zijlstra, & C. Gielen (Eds.), *The origin of outflows in evolved stars* (Vol. 366, pp. 227–233). <https://doi.org/10.1017/S1743921322000217>
- Malfait, J., Homan, W., Maes, S., Bolte, J., Siess, L., De Ceuster, F., & Decin, L. (2021). SPH modelling of wind-companion interactions in eccentric AGB binary systems. *Astronomy & Astrophysics*, 652, A51. <https://doi.org/10.1051/0004-6361/202141161>
- Nordhaus, J., & Blackman, E. G. (2006). Low-mass binary-induced outflows from asymptotic giant branch stars. *Monthly Notices of the Royal Astronomical Society*, 370(4), 2004–2012. <https://doi.org/10.1111/j.1365-2966.2006.10625.x>
- Sulzer, I., & Buck, T. (2023). Speeding up astrochemical reaction networks with autoencoders and neural ODEs. *arXiv e-Prints*, arXiv:2312.06015. <https://doi.org/10.48550/arXiv.2312.06015>