# Polarons and charge-transfer excitations from grand-canonical neural networks

**Libor Vojáček**[1,2], Cristiano Malica[3], Nicola Marzari[1-4]
[1]Paul Scherrer Institute, Villigen-PSI, Switzerland
[2]National Centre for Computational Design and Discovery of Novel Materials (MARVEL), Villigen-PSI, Switzerland
[3]University of Bremen, Bremen, Germany
[4]École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

libor.vojacek@psi.ch

Magnetite $Fe_3O_4$ (Fig. 1) is the oldest known magnetic material, yet the physics of its Verwey transition is still not fully solved due to its coupled structural, charge, and orbital interactions. The difficulty hinges on the multivalent nature of iron (Fe2+/Fe3+), which is a central motif in redox chemistry and functional oxides.

Machine learning interatomic potentials (MLIP) trained on DFT+$U$+$V$ data have recently proven to efficiently tackle the oxidation state of multivalent metallic ions with near ab initio accuracy [1], adding to their intrinsic efficiency enabling the simulations of large and/or complex systems.

This is especially relevant for magnetite where the low-temperature monoclinic phase requires a large (224 atom) supercell with a combinatorial number of 64 choose 32 ~ $10^{18}$ possible charge orderings, making ab initio calculations expensive.

In this work, we train an oxidation-aware MLIP for magnetite and demonstrate its capabilities side-by-side with a conventional (unaware) fine-tuned foundational model [2]. While the oxidation-aware potential provides the oxidation state directly as a control variable, the conventional potential can capture it intrinsically through the distinct volume of the surrounding oxygen octahedra.

First of all, fine-tuning of a foundational model (`NequIP-OAM-L-0.1`) turns out to be crucial in the case of magnetite to redeem its deficiency and obtain the correct 1:1 ratio of Fe2+ and Fe3+ octahedral sites (Fig. 2).

Even with no explicit oxidation label, the Fe2+ / Fe3+ order can then be reliably derived from the distinct volumes of the Fe octahedral oxygen cages. A machine learning molecular dynamics (ML-MD) using such model naturally exhibits Fe2+ → Fe3+ charge transfer processes (Fig. 3). Moreover, the charge transfer rate across different temperatures corresponds to Arrhenius law with an activation energy $E_a \approx 60$ meV, in the expected experimental range.

In a second step, we train an oxidation-aware MLIP, where the Fe2+ / Fe3+ ox. states are controlled explicitly. This enables us to, among others, quantitatively describe electron transfer reactions (the diabatic free-energy surfaces) from molecular dynamics [3].

The oxidation-aware MLIP also allows for Monte Carlo simulations of the charge ordering in the low-temperature monoclinic cell (Fig. 4), which are able to correctly [4] find a configuration which is several meV/atom below the Verwey charge ordering.

Choosing the paradigmatic $Fe_3O_4$ as an example, this work further highlights the capability of oxidation-aware MLIPs as a practical route to predictive atomistic modeling of redox states. Such capability is broadly relevant for accelerating the design and screening of materials where oxidation chemistry is decisive, including batteries, electrocatalysts, and solid oxide fuel cells [5].

## References

[1] C. Malica and N. Marzari, Npj Comput. Mater. **11** (2025) 212
[2] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, Nat Commun **13** (2022) 2453
[3] P.H.-L. Sit, M. Cococcioni, and N. Marzari, Phys. Rev. Lett. **97** (2006) 028303
[4] N. Fominykh and V. Stegailov, Phys. Rev. B **111** (2025) 115130
[5] L. Vojáček, C. Malica and N. Marzari, under preparation

## Figures

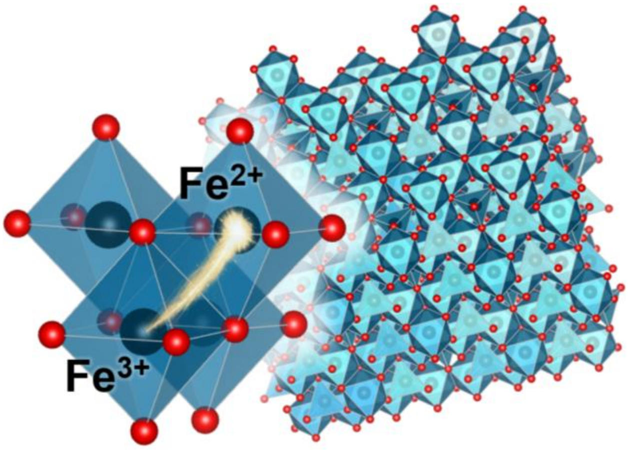

**Figure 1.** Magnetite $Fe_3O_4$ bulk structure and an artist's view of an electron transfer between the Fe octahedral sites.

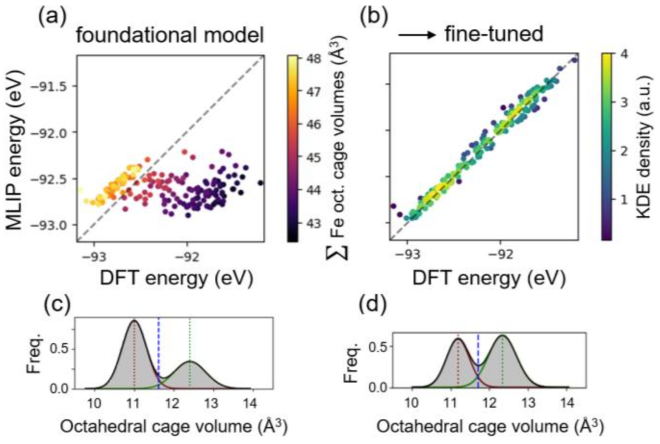

**Figure 2.** (a) A foundational model incorrectly predicts magnetite configurations with unproportionally large number of small (Fe3+) octahedral cages as low-energy structures. (b) Fine-tuning on DFT+*U*+*V* data redeems a correct 1:1 ratio of Fe2+ / Fe3+ octahedral sites. This can be inferred from the octahedral oxygen cage volumes (c,d).

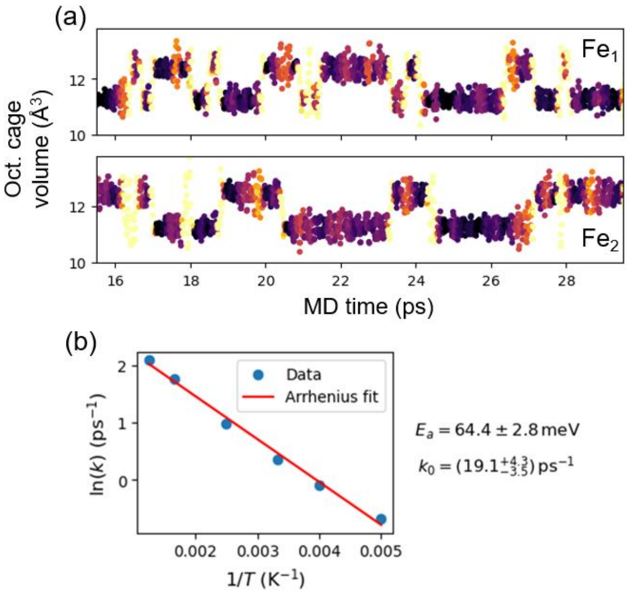

**Figure 3.** (a) Small polaron hopping during an ML-MD ($T$=300 K) using the conventional fine-tuned MLIP. (b) Their activation energy corresponds to Arrhenius law with an activation energy of ≈ 60 meV, in the expected experimental range.

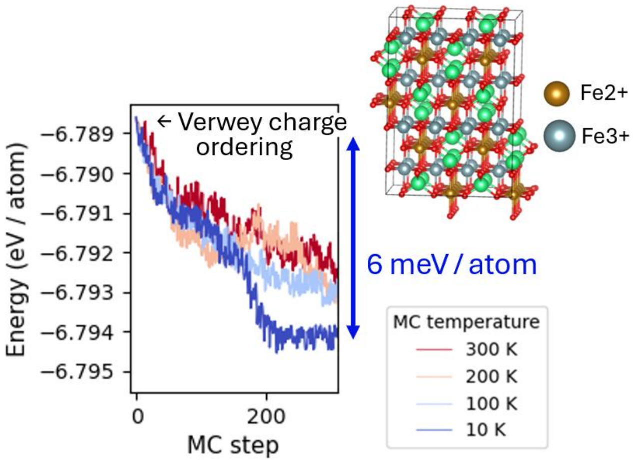

**Figure 4.** Monte Carlo simulations of the monoclinic magnetite cell using the oxidation-aware MLIP reveal a charge ordering several meV/atom below the Verwey charge order.