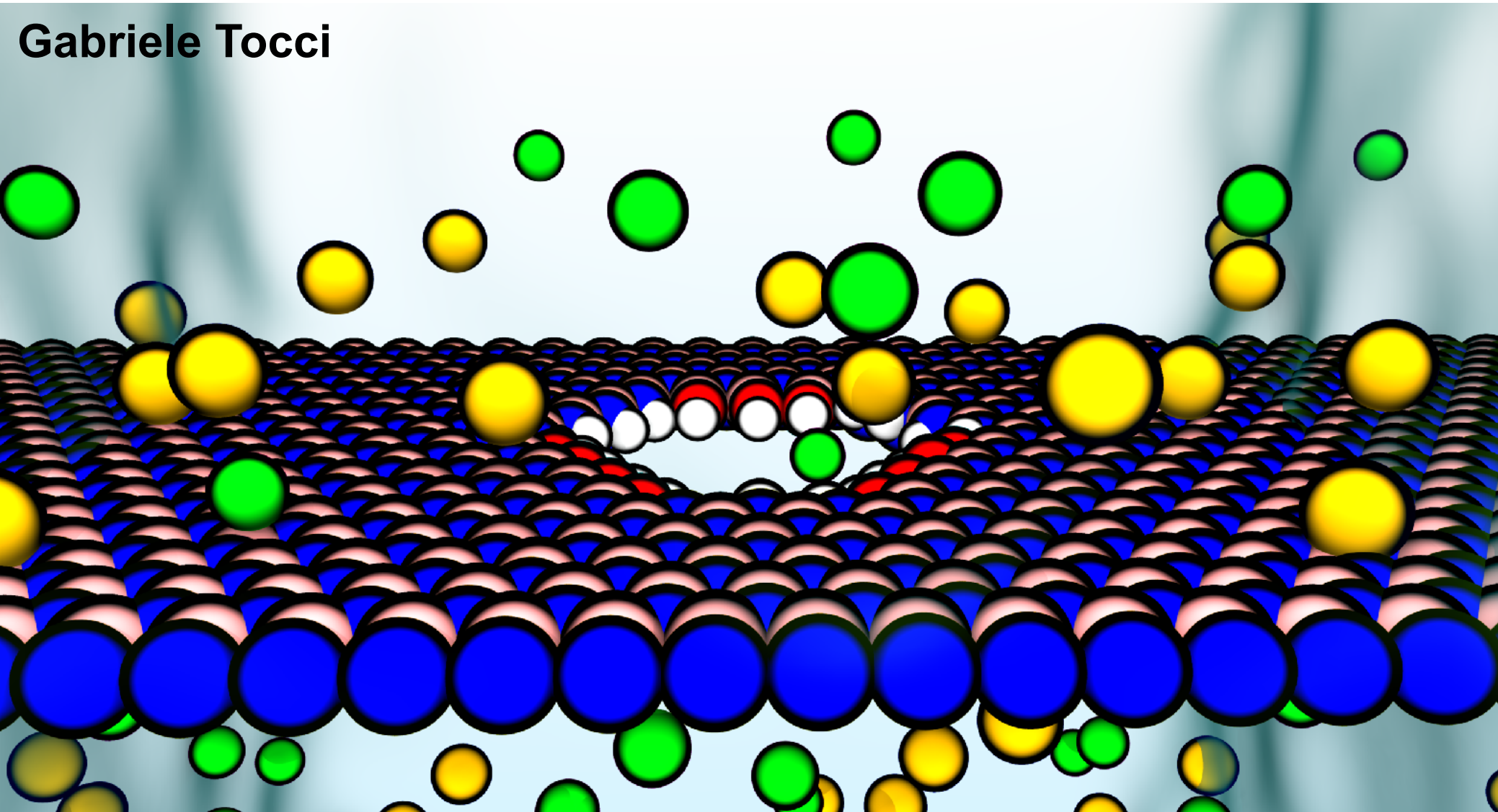
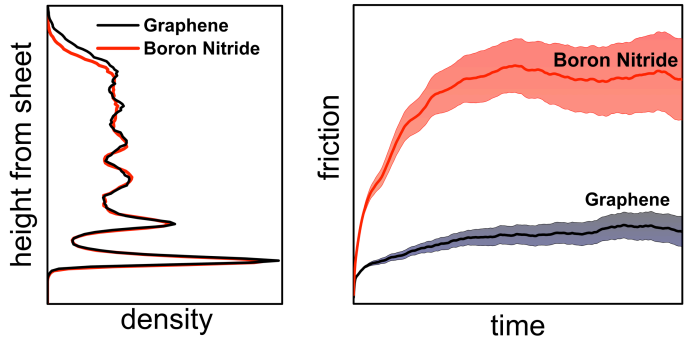
Ab Initio Molecular Dynamics and Machine Learning Interatomic Potentials at Solid Liquid Interfaces

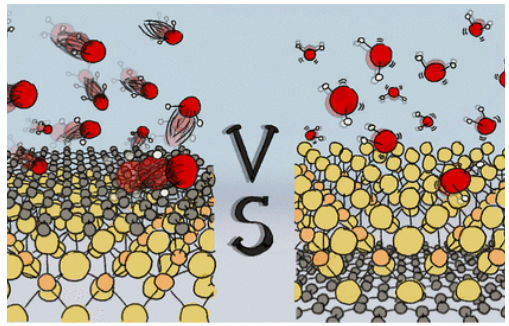
Gabriele Tocci



From *Ab Initio* Molecular Dynamics to Machine Learned Interatomic Potentials

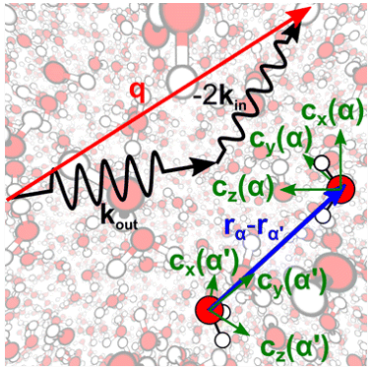


First *Ab Initio* Simulations of Nanofluidic Slip



Led Research in QM and MLIP Simulations for Aqueous Interfaces + NequIP and Allegro Inference Implemented in CP2K

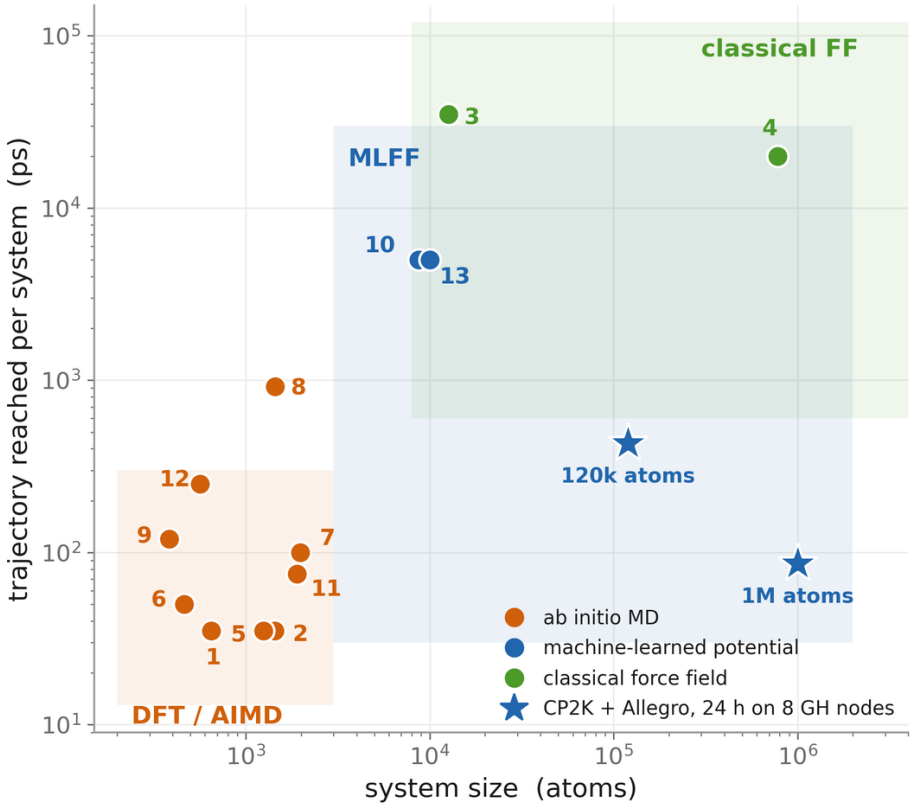
PhD, UCL → Postdoc, EPFL → Senior Scientist, UZH → Senior ML Engineer, Microsoft

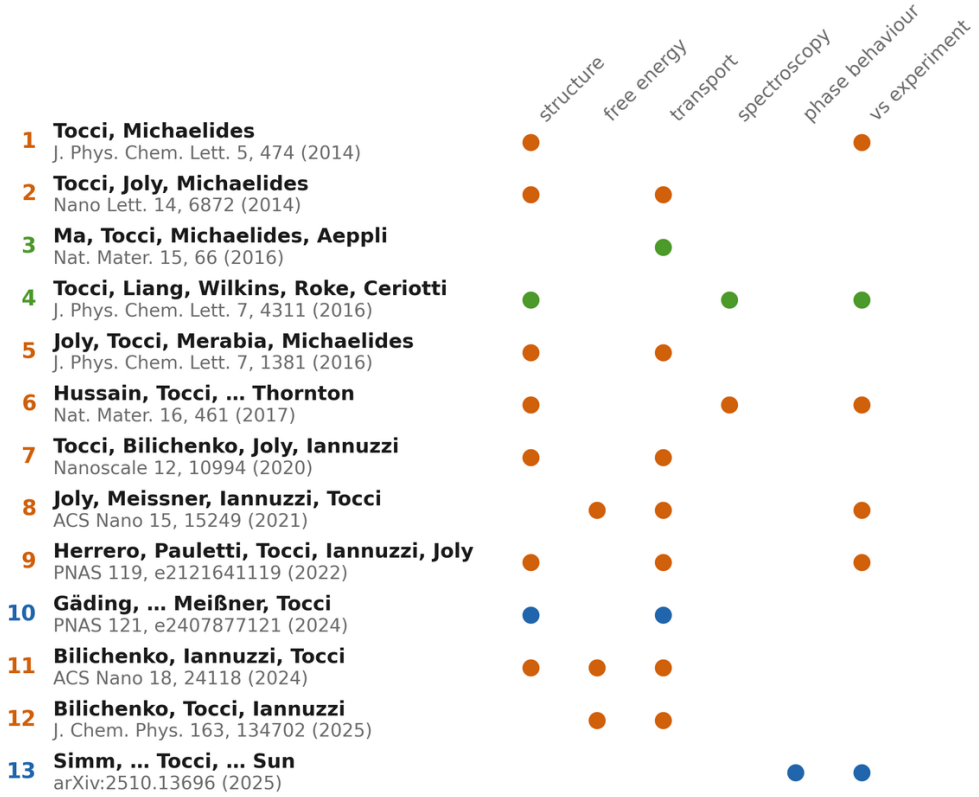


Machine Learning for Molecular Tensors — Equivariance, by Hand

MLIP / QM workflows and platform development on Azure HPC

Accuracy, System Size and Sampling





Interfacial structure, free energies, transport coefficients and spectra all need quantum accuracy. None of them converge until the system is large enough and the trajectory long enough.

Why *Ab Initio* Molecular Dynamics

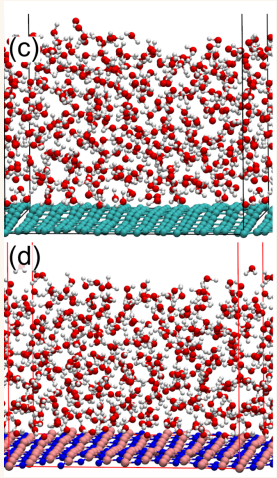
At finite temperature every observable is an average along the trajectory.

$$\langle A \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T A\big(\mathbf{x}(t)\big)\,\mathrm{d}t \;=\; \frac{1}{Z} \int A(\mathbf{x})\, e^{-U(\mathbf{x})/k_\mathrm{B}T}\,\mathrm{d}\mathbf{x}$$

$$\langle A(0)A(t)\rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T A(t')\,A(t'+t)\,\mathrm{d}t' \qquad \text{transport coefficient} \;\propto \int_0^\infty \langle A(0)A(t)\rangle\,\mathrm{d}t$$

Dynamical properties follow from time correlation functions, so the sampling has to be long as well as accurate. Comparing two materials means sampling both well enough that the difference is physical.

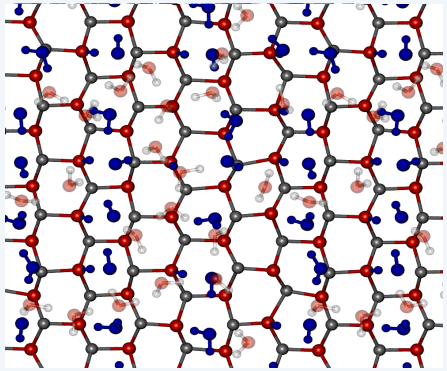
Different interfaces compared on equal footing



The Electronic Structure Dictates Properties related to Wetting and Nanofluidics.

Water on graphene and on hBN. Nano Lett. 14, 6872 (2014).

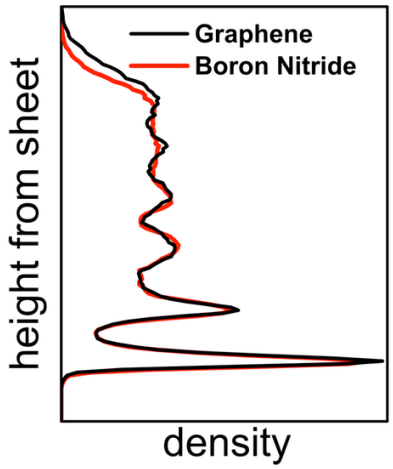
Sampling bond breaking accurately

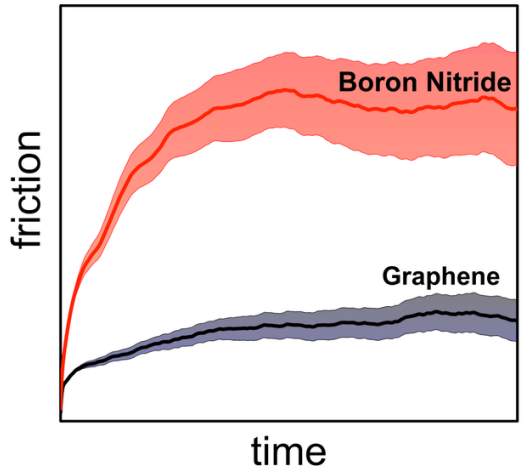


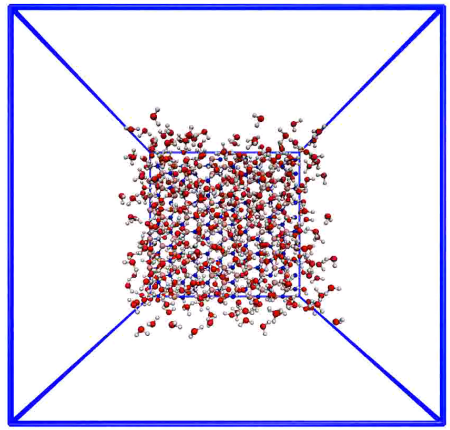
Proton transfer, dissociation and hydroxyl overlayers. Classical force fields cannot describe them.

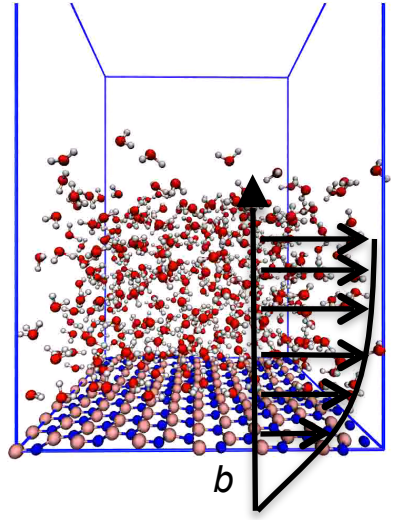
Proton hopping at the water and ZnO(10-10) interface. J. Phys. Chem. Lett. 5, 474 (2014).

Water friction on graphene and hexagonal boron nitride









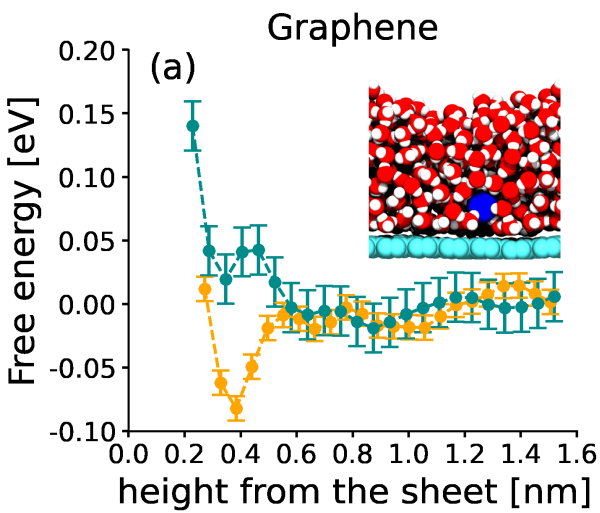
$$\lambda = \frac{1}{2Sk_{\mathrm{B}}T}\int \langle \mathbf{F}(0)\cdot\mathbf{F}(t')\rangle \,\mathrm{d}t \qquad b = \eta/\lambda$$

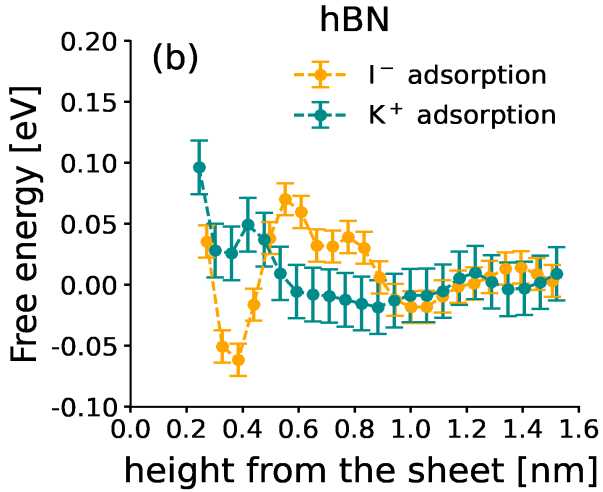
- The interfacial water structure is nearly the same on the two sheets. The friction coefficient is about three times larger on hexagonal boron nitride.
- The difference comes from the corrugation of the water energy landscape, which follows from the electronic structure of the sheet.
- Extended to MoS_2: dynamical density correlations of the liquid give rise to a further slowdown in the flow of water

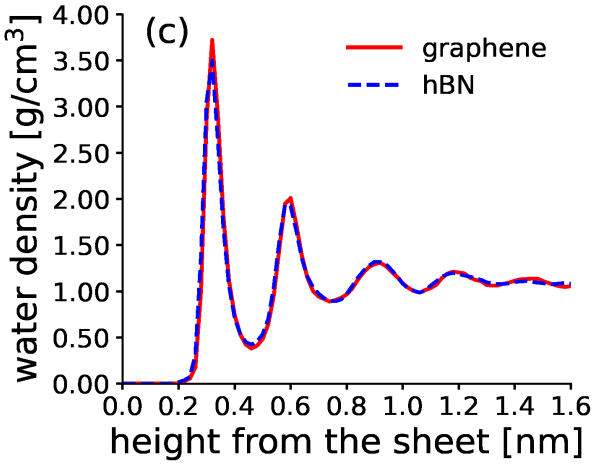
G. Tocci, L. Joly, A. Michaelides, Nano Lett. 14, 6872 (2014). G. Tocci, M. Bilichenko, L. Joly, M. Iannuzzi, Nanoscale 12, 10994 (2020).

Multi-Scale Modelling from *Ab Initio* Umbrella Sampling to Fluid Dynamics

1. Free energy of ion adsorption from ab initio umbrella sampling
2. Electrostatic potential in the double layer from Poisson Boltzmann theory
3. Transport coefficients from the Stokes equation with a slip boundary condition



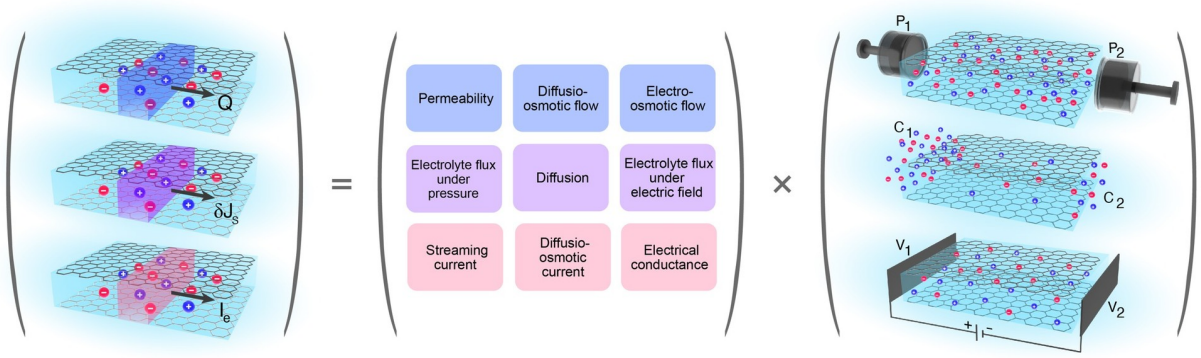


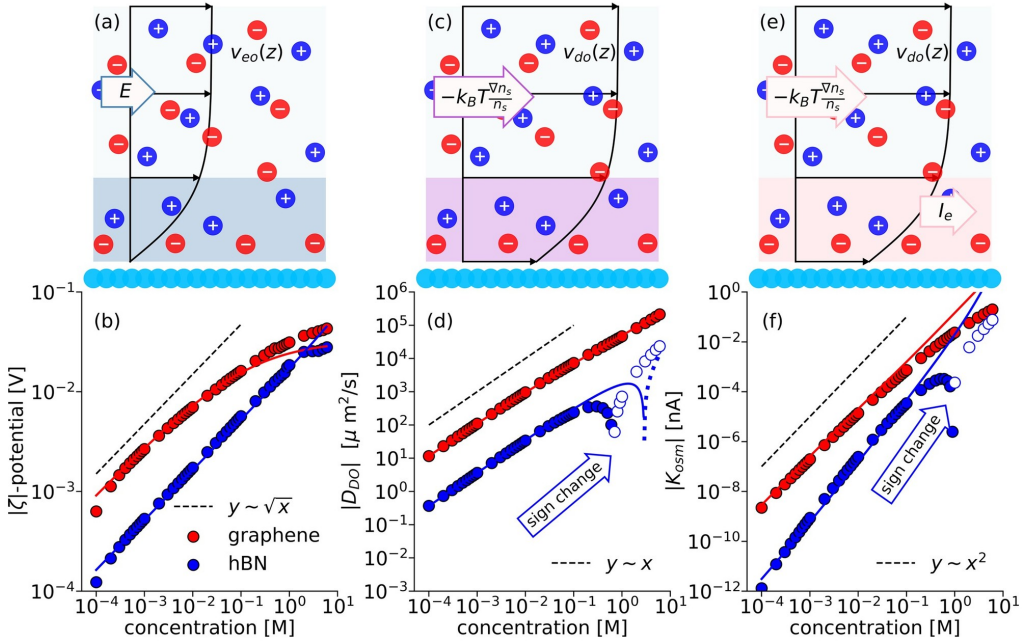


K^+ adsorbs in the same way on graphene and on hBN. I^- does not. The water density profile is almost identical on the two sheets, so the difference is ion specific and not structural.

L. Joly, R. H. Meissner, M. Iannuzzi, G. Tocci, ACS Nano 15, 15249 (2021). 1400 atoms, 23 umbrella windows of 40 ps for each ion and each sheet.

Ion Adsorption and Osmotic Transport

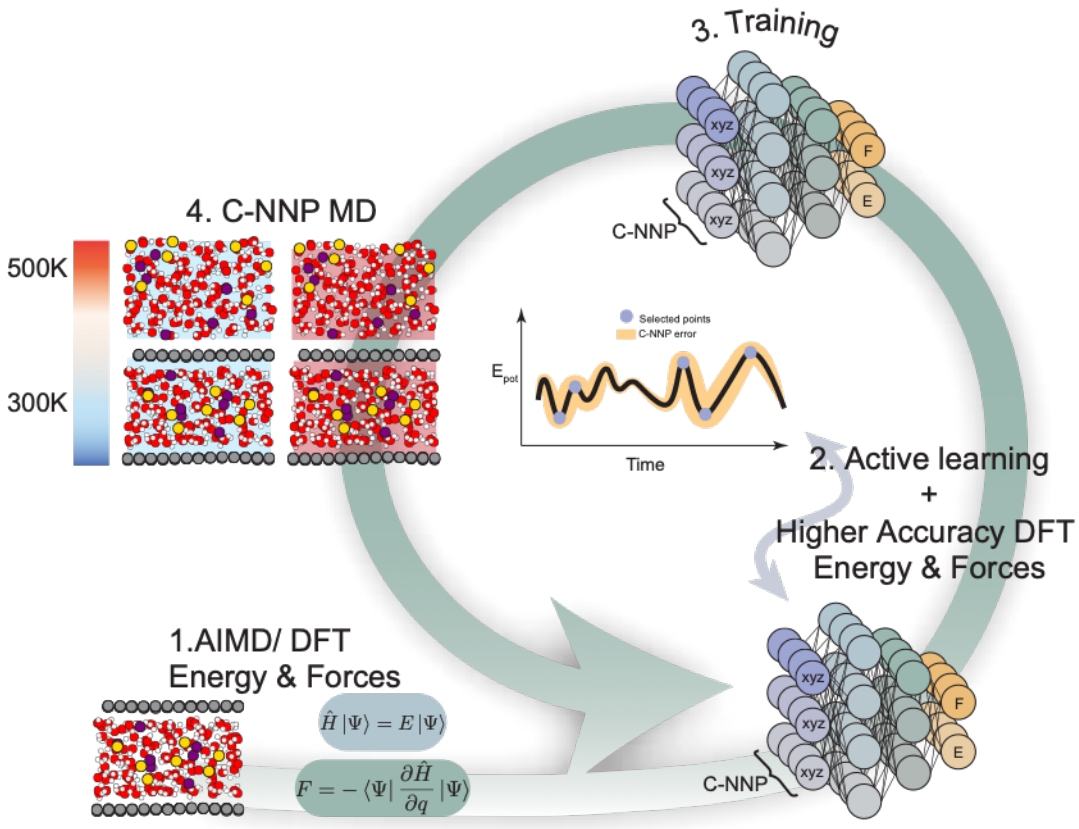


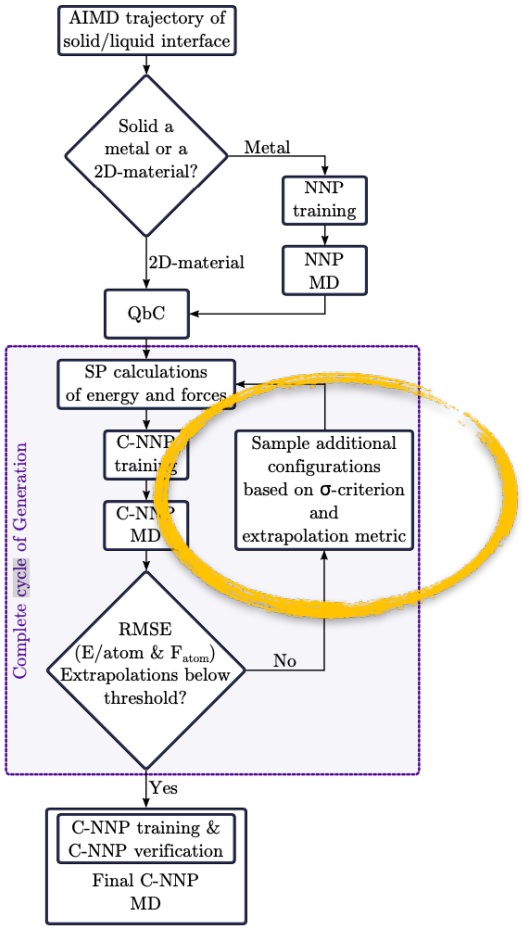


K^+ adsorbs in the same way on graphene and on hBN. I^- does not.

Each osmotic transport coefficient then follows a different scaling law with concentration.

Active Learning for Machine Learning Force Fields





A committee of neural network potentials estimates its own uncertainty along a trajectory. Ab initio calculations performed only where the committee disagrees, or where a symmetry function leaves the range seen in training.

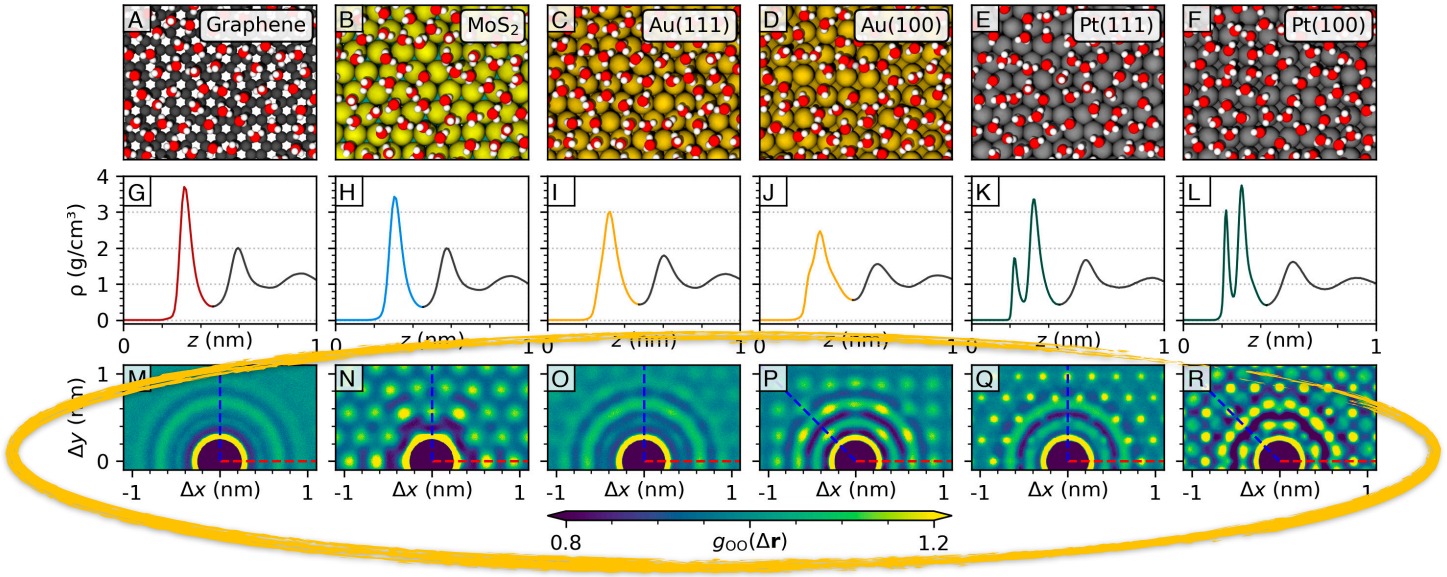
C. Schran, K. Brezina, O. Marsalek, *J. Chem. Phys.* 153(10), (2020)
M. Eckhoff, J. Behler, *npj Comput Mater*, 170 (2021)

Gäding, J., Della Balda, V., Lan, J., Konrad, J., Iannuzzi, M., Meißner, R.H,. Tocci, G., PNAS 121(38), e2407877121 (2024)

PNAS RESEARCH ARTICLE | CHEMISTRY OPEN ACCESS

The role of the water contact layer on hydration and transport at solid/liquid interfaces

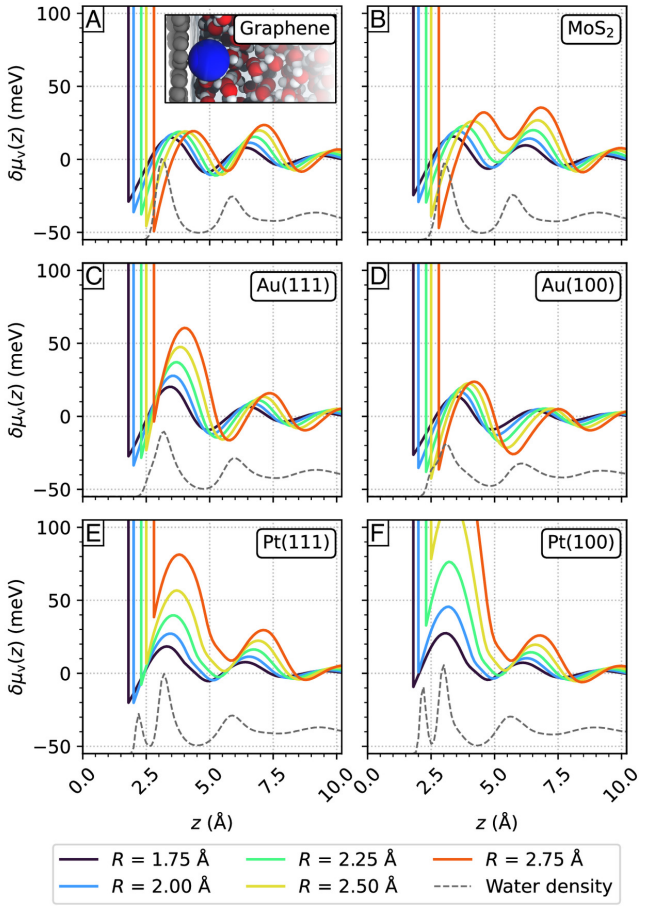
J. Gäding [a,b], V. Della Balda [c], J. Lan [d,e], J. Konrad [a], M. Iannuzzi [c,1], R. H. Meißner [a,b,1], and G. Tocci [c]

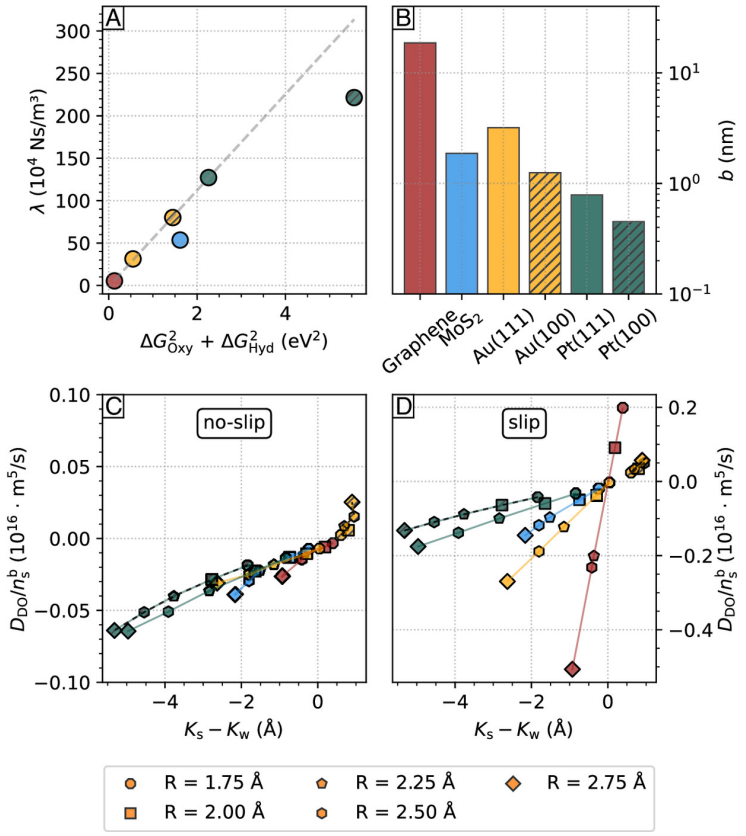


$$g_{OO}(\Delta\mathbf{r}) = \frac{1}{N\rho}\left\langle \sum_{i\neq j} \delta(\Delta\mathbf{r} - \Delta\mathbf{r}_{ij}(t)) \right\rangle$$

The water contact layer form a rich variety of structures on different surfaces, with a strong impact on transport and hydration, as shown by ab initio molecular dynamics accelerated with machine learning force fields obtained from an active-learning workflow

Hydrophobic Solutes and Interfacial Transport





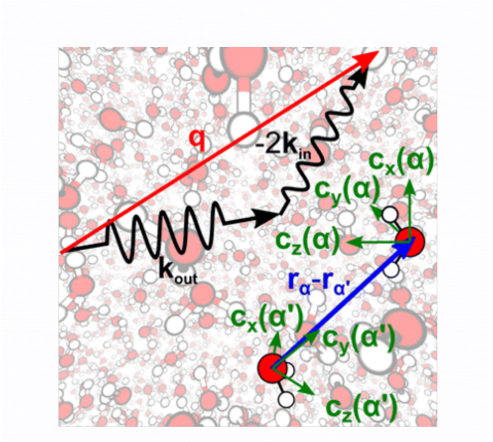
- Insertion free energy of a hydrophobic solute against height above the surface, for six surfaces and solute radii from 1.75 to 2.75 Å.
- Friction follows the hydration free energy of the surface.
- The diffusio-osmotic coefficient changes sign once slip is included.
- Transport of a hydrophobe is set by the contact layer, not by the bulk liquid.

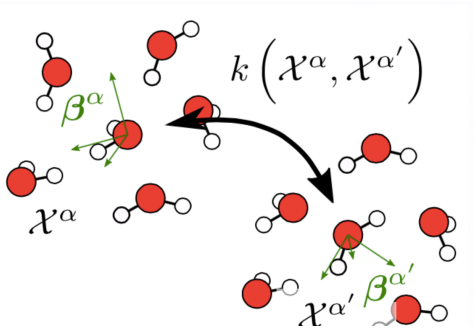
$$\delta\mu(z) = -\,k_{\mathrm{B}}T\,\ln\frac{P(z, N_R = 0)}{P(z^b, N_R = 0)}$$

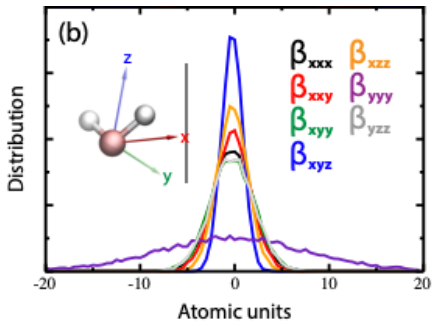
$$D_{DO} = \frac{k_{\mathrm{B}}T \times n_{\mathrm{s}}^{b}}{\eta}\left[K_{\mathrm{s}}L_{\mathrm{s}} - K_{\mathrm{w}}L_{\mathrm{w}} + b(K_{\mathrm{s}} - K_{\mathrm{w}})\right]$$

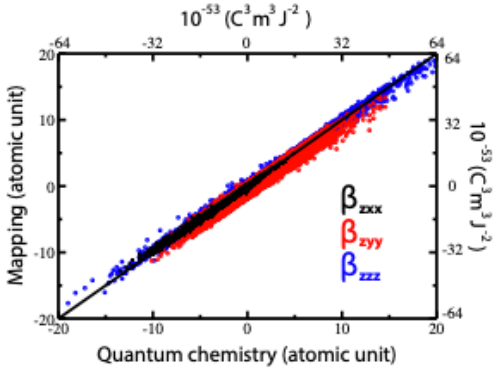
J. Gäding et al., PNAS 121, e2407877121 (2024). M. Bilichenko, M. Iannuzzi, G. Tocci, ACS Nano 18, 24118 (2024).

Tensor Properties from Kernel Ridge Regression









- Second harmonic scattering depends on the hyperpolarisability of every molecule and on the orientational correlations between molecules.
- β is learned from a descriptor carrying the local structure and the electric field.
- Training and prediction done by rotating the molecule and local environment to a common molecular reference frame to preserve the underlying symmetry

$$\beta_{abc}(\mathbf{x}) = \bar{\beta}_{abc} + \sum_{i}^{N_t} c_{i\,abc}\, K(\mathbf{x}, \mathbf{x}_i), \qquad K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-(\mathbf{x} - \mathbf{x}_i)^2/\sigma^2\right)$$

$$\mathbf{x} = \{\mathbf{E}_i,\ \rho_i^{\mathrm{O}},\ \rho_i^{\mathrm{H}}\}, \qquad \mathbf{E}_i = -\sum_j \nabla_i [q_j\, \mathrm{erf}(r_{ij}/\sigma_E)/r_{ij}]\,, \qquad \rho_i^X = \sum_{j\in X} e^{-r_{ij}^2/2\sigma_\rho^2}$$

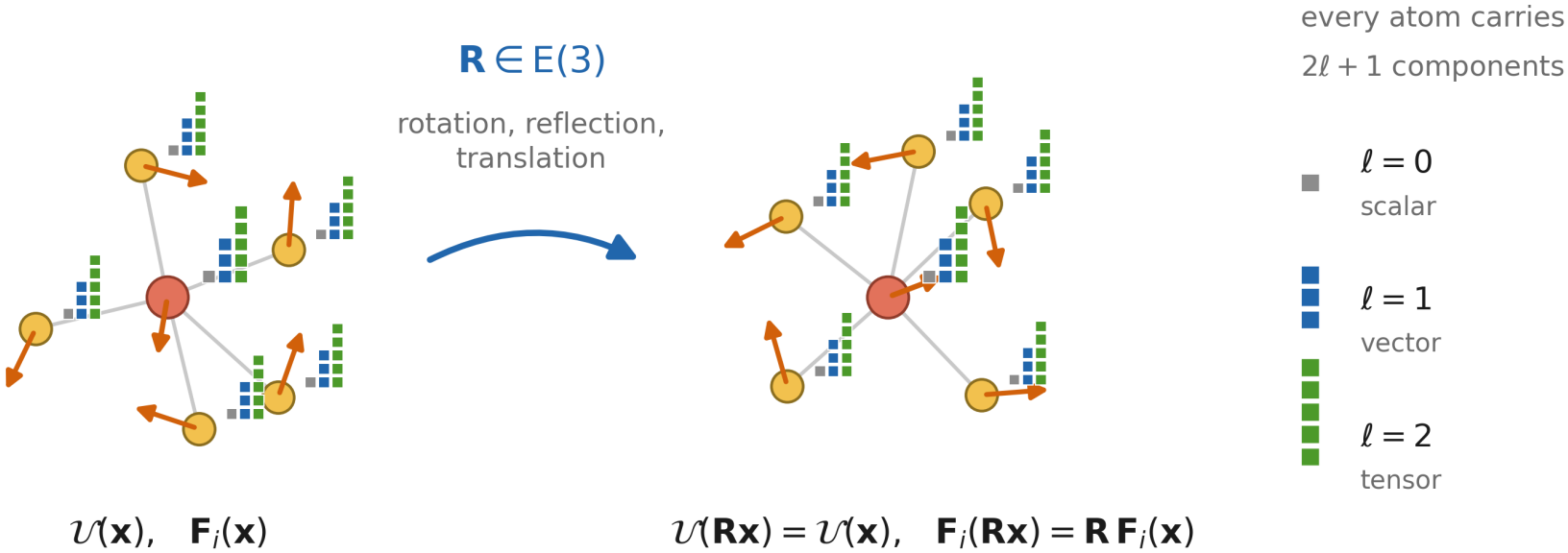
Rotations handled by predicting the components in a molecular frame and rotating back.

$$\beta'_{abc} = R_{aa'} R_{bb'} R_{cc'}\, \beta_{a'b'c'}$$

Symmetry adapted kernels, and later equivariant neural networks, hand-chosen frame with one built into the model.

G. Tocci, C. Liang, D. Wilkins, S. Roke, M. Ceriotti, J. Phys. Chem. Lett. 7, 4311 (2016). C. Liang, G. Tocci, D. Wilkins, A. Grisafi, S. Roke, M. Ceriotti, Phys. Rev. B 96, 041407(R) (2017). λ-SOAP: A. Grisafi, D. Wilkins, G. Csányi, M. Ceriotti, Phys. Rev. Lett. 120, 036002 (2018).

From Molecular Frames to Equivariant Features



$\mathcal{U}(\mathbf{x}), \quad \mathbf{F}_i(\mathbf{x})$

$\mathcal{U}(\mathbf{Rx}) = \mathcal{U}(\mathbf{x}), \quad \mathbf{F}_i(\mathbf{Rx}) = \mathbf{R}\,\mathbf{F}_i(\mathbf{x})$

Invariance

$$f\left(D_X(g)\,\mathbf{x}\right) = f(\mathbf{x})$$

The total energy is a scalar, $\ell = 0$. Here D_Y is the identity, so the energy is unchanged by every element of E(3).

Equivariance

$$f\left(D_X(g)\,\mathbf{x}\right) = D_Y(g)\,f(\mathbf{x})$$

The forces are vectors, $\ell = 1$. Here D_Y is the rotation R, so the forces turn with the system.

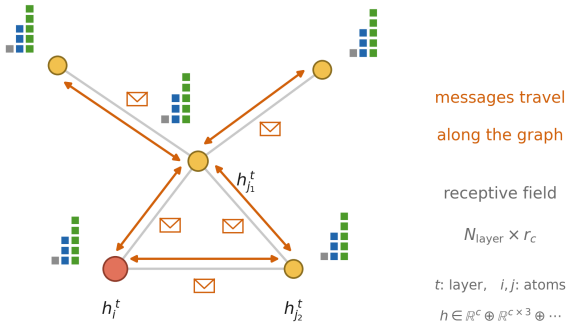
In Equivariant MLIPs symmetry is built in, not learned. Orders of magnitude fewer training configurations needed.

S. Batzner et al., Nat. Commun. 13, 2453 (2022).

NequIP, Allegro and Vivace

NequIP

message passing



$$E = \sum_i E_i$$

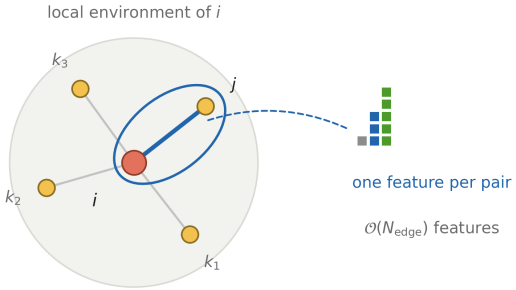
$$h_i^L = \sum_{k \in \mathcal{N}(i)} h_k^{L-1} \otimes w_{ik}^L \vec{Y}_{ik}$$

The neighbour feature sits inside the sum.

Message passing implies that scaling is cubic in N_{atoms} so the model does not parallelise well.

Allegro

strictly local, features on pairs



$$E = \sum_i \sum_{j \in \mathcal{N}(i)} E_{ij}$$

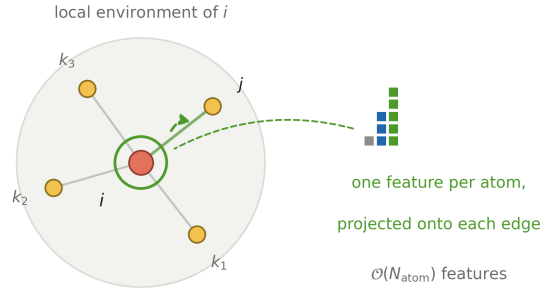
$$h_{ij}^L = h_{ij}^{L-1} \otimes \sum_{k \in \mathcal{N}(i)} w_{ik}^L \vec{Y}_{ik}$$

The pair feature comes outside the sum.

Strictly local. Each rank evaluates its own edges.

Vivace

strictly local, features on atoms



$$E = \sum_i \sum_{j \in \mathcal{N}(i)} E_{ij}$$

$$\tilde{\xi}_i^L = \tilde{\xi}_i^{L-1} \otimes \sum_{k \in \mathcal{N}(i)} w_{ik}^L \vec{Y}_{ik}$$

$$e_{ij} = \text{Linear}\left(\tilde{\xi}_i^L \cdot \vec{Y}_{ij}\right) \text{MLP}(\rho_{ij})$$

The same as Allegro, applied to the atom.

Strictly local. Equivariant operations up to 3.8 Å, invariant ones up to 6.5 Å.

S. Batzner et al., Nat. Commun. 13, 2453 (2022). A. Musaelian et al., Nat. Commun. 14, 579 (2023). G. N. C. Simm et al., arXiv:2510.13696 (2025).

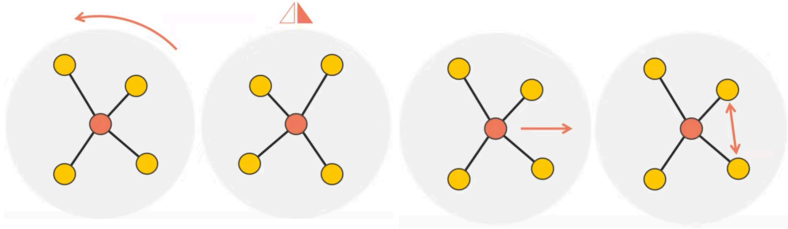
Nequip and Allegro Inference in CP2K

Nequip and Allegro: Equivariant GNNs

Atoms represented as nodes embedded in their local environment connected to neighbours (edges)

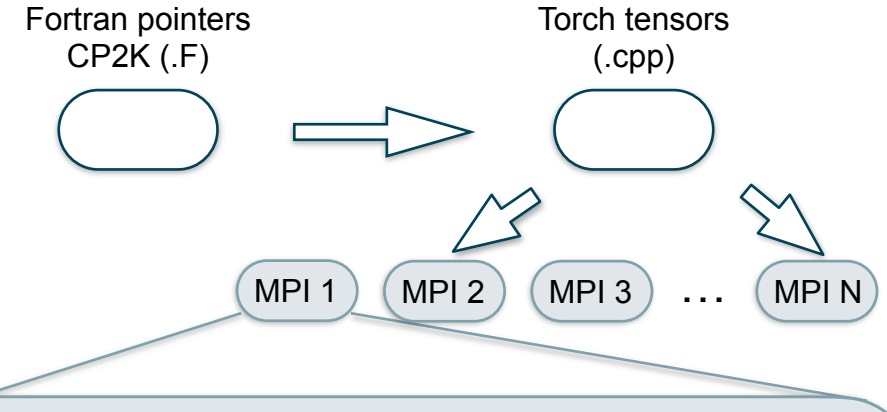
The total energy: $\mathcal{U} = \sum_i \sum_{j \in \mathcal{N}(i)} U_{ij}$

Forces: $\mathbf{F}_i = -\nabla_i \mathcal{U}$



API Abstraction in CP2K:

1. Mapped Fortran pointers to torch tensors using a custom dictionary built into the C-API, that handles the CP2K - LibTorch interface
2. Atoms (nodes) and neighbours (edges) divided into separate MPI ranks, with the model copied over ranks
3. Fwd pass calculates atomic energies and forces on separate ranks (on GPUs)
4. Energies and Forces obtained from sum/reduce on each rank, aggregating local contributions



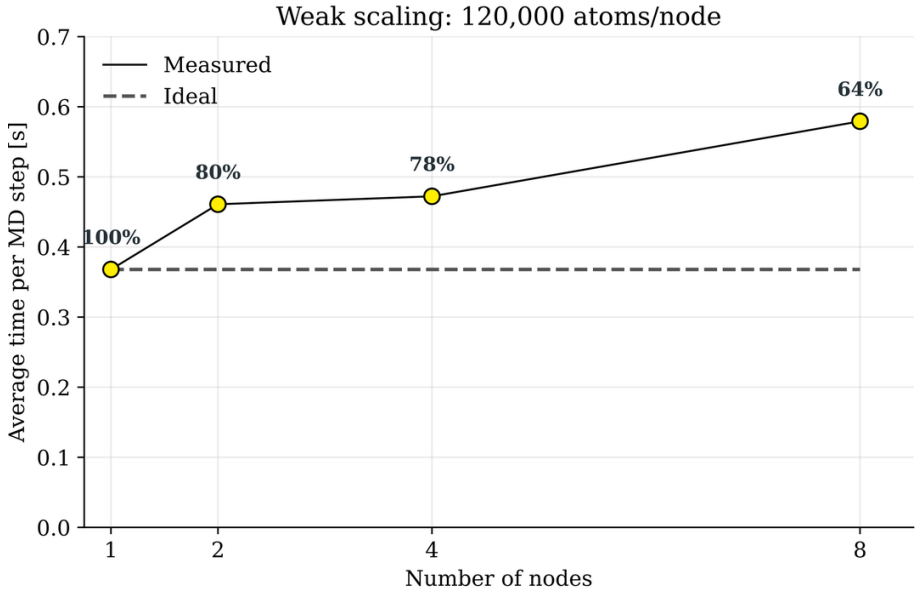
API Call on a single rank for a subset of atoms

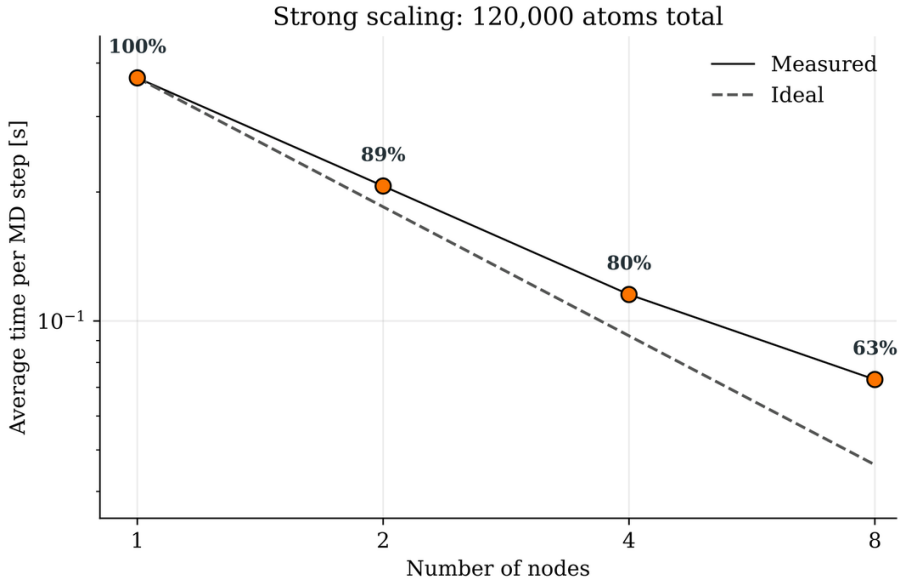
```
CALL torch_dict_create(inputs)
CALL torch_dict_insert(inputs, "pos", pos_t)
CALL torch_dict_insert(inputs, "edge_index", idx_t)
CALL torch_dict_insert(inputs, "edge_cell_shift", shift_t)
CALL torch_dict_insert(inputs, "cell", cell_t)
CALL torch_dict_insert(inputs, "atom_types", types_t)

CALL torch_dict_create(outputs)
CALL torch_model_eval(allegro_data%model, inputs, outputs)
CALL torch_dict_get(outputs, "atomic_energy", atomic_energy)
CALL torch_dict_get(outputs, "forces", forces)
```

Parallel Scaling of Allegro Inference in CP2K

- Bulk water from 1.2×10^5 to about 10^6 atoms, on 1 to 8 Grace Hopper nodes of CSCS Alps.
- Above 60 per cent efficiency in both weak and strong scaling on 8 nodes.
- About 0.5 s per molecular dynamics step at 10^6 atoms, and under 0.1 s per step at 1.2×10^5 atoms.

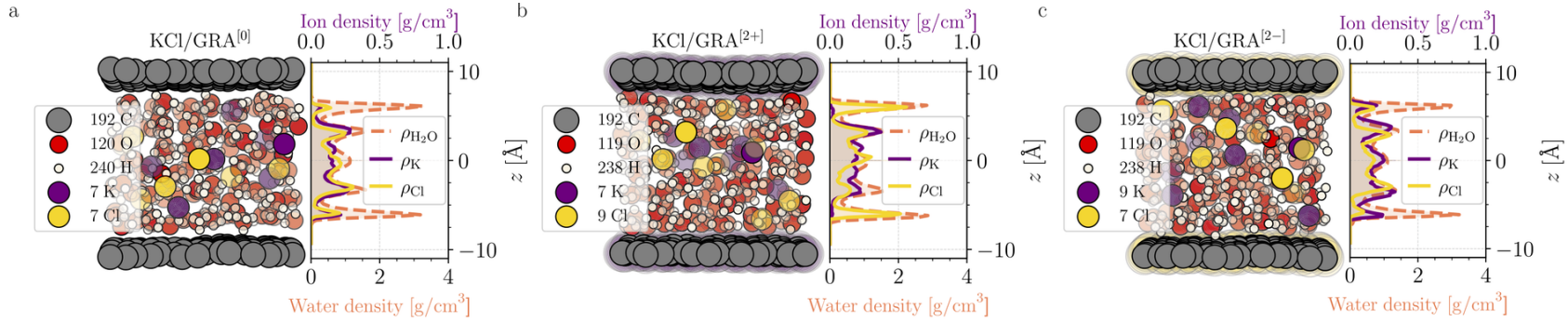




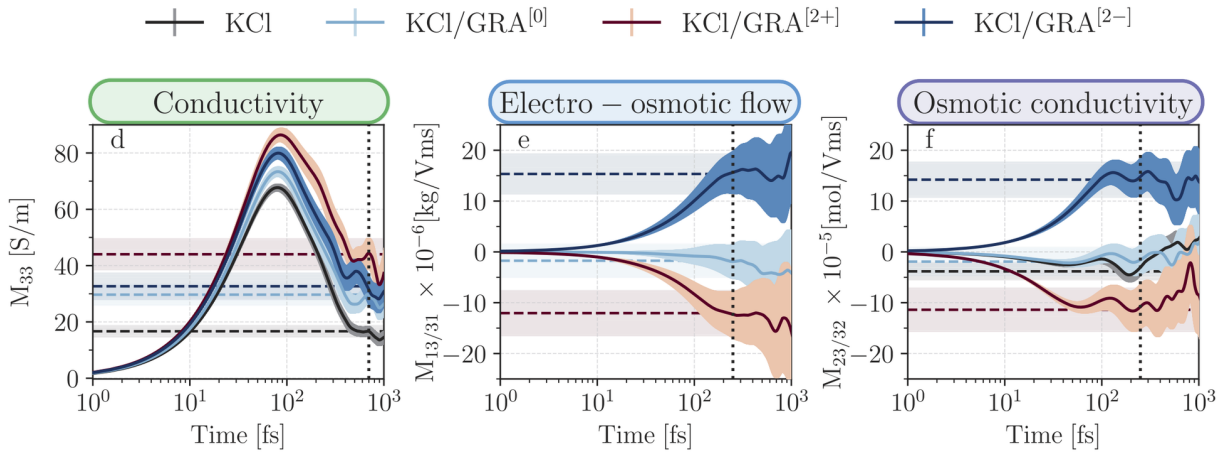
Allegro inference through the CP2K interface. Benchmarks on CSCS Alps, Daint Grace Hopper partition.

Transport Coefficients from an Allegro Potential

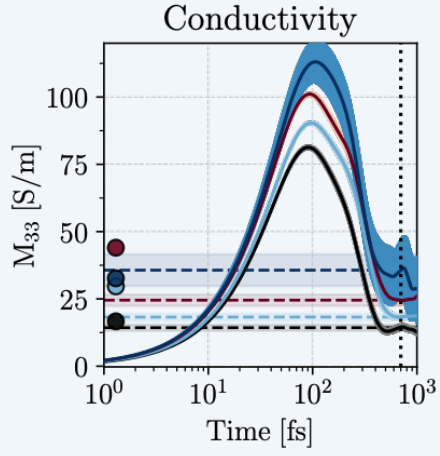
Concentrated KCl, confined between neutral and charged graphene sheets



Ab initio molecular dynamics



Allegro potential, in progress

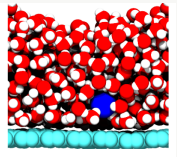


M. Bilichenko, G. Tocci, M. Iannuzzi, J. Chem. Phys. 163, 134702 (2025)

Conclusions and Acknowledgements

Accuracy

Ab initio molecular dynamics combined with hydrodynamics of complex interfaces reveal that subtle electronic structure effects have dramatic consequences on transport

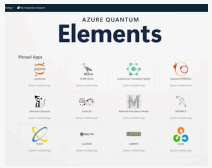


Scale

NequIP and Allegro inference in CP2K. 10^6 atoms on Daint at 0.5 s / step enabling scalable MLIPs in CP2K

End-to-End Workflows

Committee active learning made six surfaces feasible instead of one. Similar types of workflows built on Azure HPC, from dataset generation to deployed and benchmarked models.



Funding Swiss National Science Foundation, Ambizione. University of Zurich.

Computing PRACE and CSCS. Benchmarks on CSCS Daint Alps.