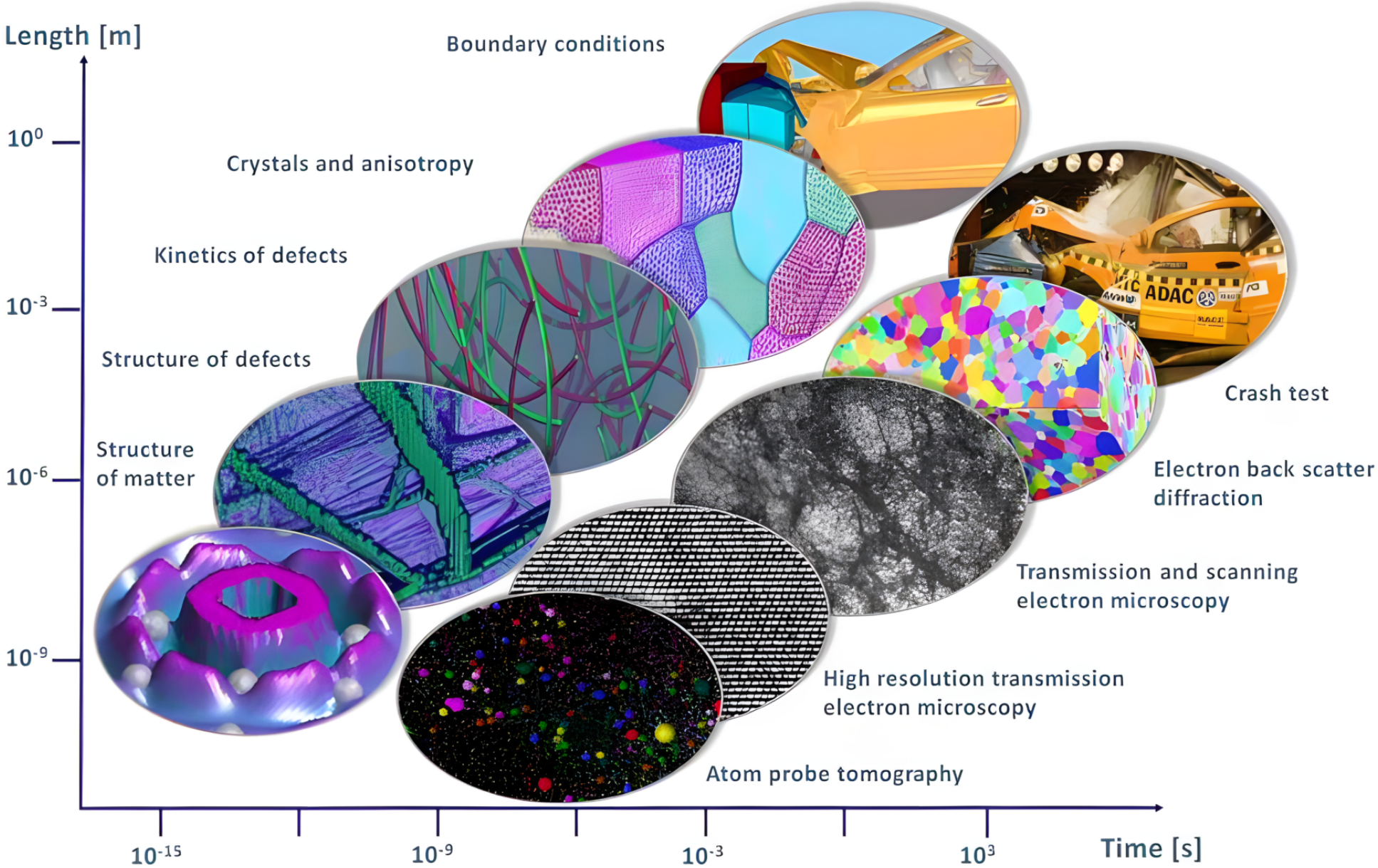
INTRODUCTION TO ATOMIC CLUSTER EXPANSION

Maciej Buze

21 May 2025

Lancaster AI Reading Group

LENGTH- AND TIME-SCALES IN MULTISCALE MODELLING

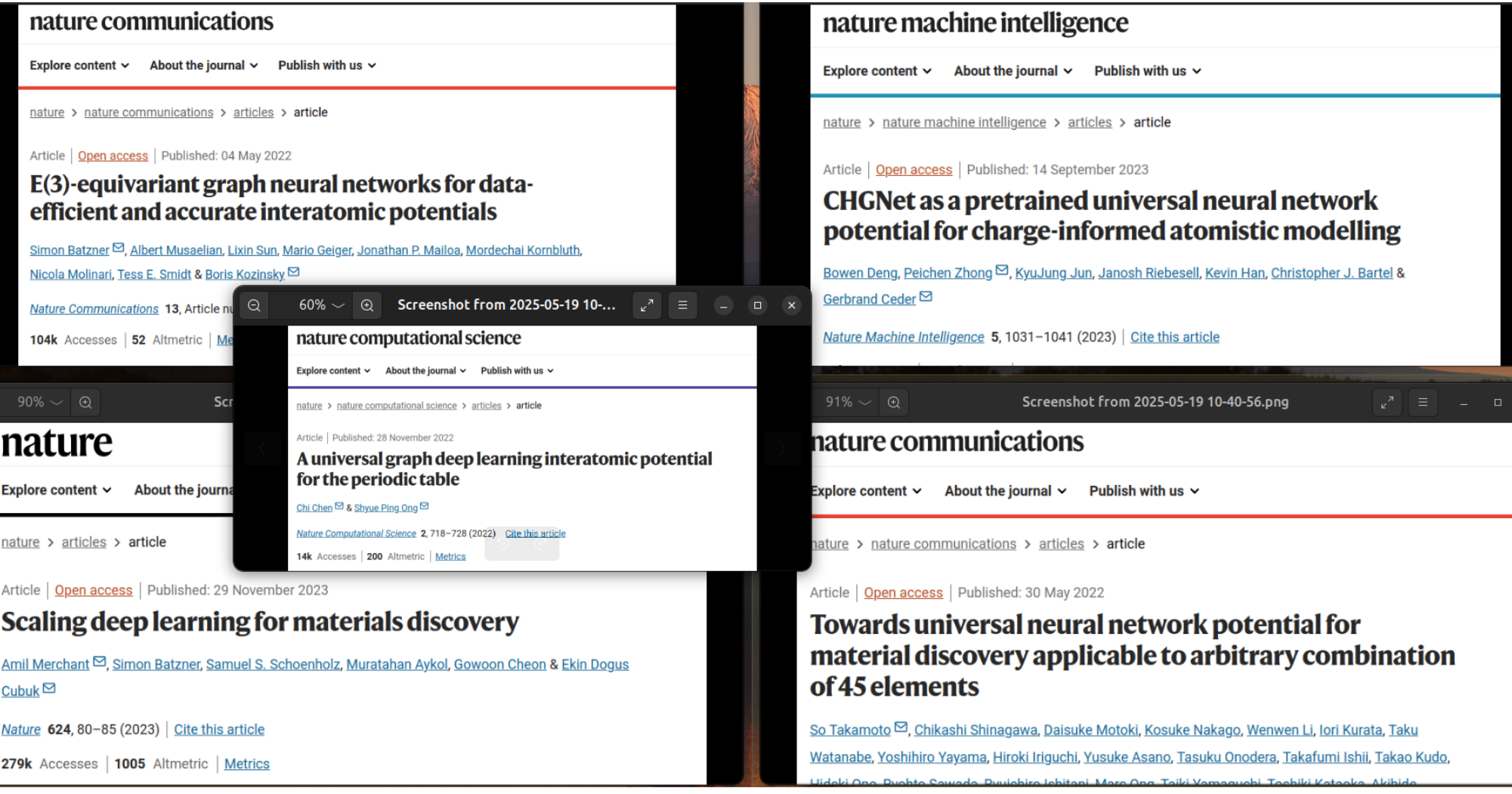


F. Roters, P. Eisenlohr, L. Hantcherli, D.D. Tjahjanto, T.R. Bieler, D. Raabe: Acta Materialia 58 (2010), adapted from dierk-raabe.com

Motivation: atomistic scale modelling with interatomic potentials

(Live demonstration in Ovito)

FOUNDATIONAL MODELS



MACE

Physics > Chemical Physics

[Submitted on 29 Dec 2023 (v1), last revised 1 Mar 2024 (this version, v2)]

A foundation model for atomistic materials chemistry

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Fabio Falcioni, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, Gábor Csányi

Machine-learned force fields have transformed the atomistic modelling of materials by enabling simulations of ab initio quality on unprecedented time and length scales. However, they are currently limited by: (i) the significant computational and human effort that must go into development and validation of potentials for each particular system of interest; and (ii) a general lack of transferability from one chemical system to the next. Here, using the state-of-the-art MACE architecture we introduce a single general-purpose ML model, trained on a public database of 150k inorganic crystals, that is capable of running stable molecular dynamics on molecules and materials. We demonstrate the power of the MACE-MP-0 model - and its qualitative and at times quantitative accuracy - on a diverse set problems in the physical sciences, including the properties of solids, liquids, gases, chemical reactions, interfaces and even the dynamics of a small protein. The model can be applied out of the box and as a starting or "foundation model" for any atomistic system of interest and is thus a step towards democratising the revolution of ML force fields by lowering the barriers to entry.

Comments: 119 pages, 63 figures, 37MB PDF
Subjects: **Chemical Physics (physics.chem-ph)**; Materials Science (cond-mat.mtrl-sci)
Cite as: arXiv:2401.00096 **[physics.chem-ph]**
(or arXiv:2401.00096v2 **[physics.chem-ph]** for this version)
https://doi.org/10.48550/arXiv.2401.00096

SETUP

- Configuration space Ω for particles $\boldsymbol{x}_j \in \Omega$. Examples:
  - Atoms with positions and chemical species:
    $\boldsymbol{x}_j = (\boldsymbol{r}_j, Z_j) \in \Omega = \mathbb{R}^3 \times \mathbb{Z}$
  - Electrons with positions and spins:
    $\boldsymbol{x}_j = (\boldsymbol{r}_j, \sigma_j) \in \Omega = \mathbb{R}^3 \times \{+1, -1\}$
  - Agents with beliefs/descriptors:
    $\boldsymbol{x}_j = (\boldsymbol{r}_j, \boldsymbol{b}_j) \in \Omega = \mathbb{R}^3 \times B$ (?)
- Generic configuration $\boldsymbol{X} = \{\boldsymbol{x}_j\}_{j=1}^M \subset \mathbb{R}^3 \times \mathbb{R}^d$ (today $d = 0$).
- We want to derive an energy $\boldsymbol{X} \mapsto E(\boldsymbol{X})$ that will match "the truth" very well [more on that later]
- Often natural to expand E as a sum of site energies

$$E(\boldsymbol{X}) = \sum_{i=1}^{M} E_i(\boldsymbol{X}_i), \quad \boldsymbol{X}_i = \{\boldsymbol{x}_{ij}\}_{j=1}^M, \quad \boldsymbol{x}_{ij} := \boldsymbol{x}_i - \boldsymbol{x}_j$$

  (translation invariance)
- Let G be a Lie group acting on Ω. Function E_i is invariant under G:

$$E_i(gX_i) = E_i(X_i), \qquad \forall g \in G$$

MANY-BODY (CLUSTER) EXPANSION

$E(X) = \sum_{i=1}^{M} E_i(X_i)$

- Naive / historical expansion:

$$E_i(X_i) = V_0(\boldsymbol{x}_i) + \sum_{j_1=1}^{M} V_1(\boldsymbol{x}_{ij_1}) + \sum_{j_2=1}^{M} \sum_{j_1=1}^{j_2-1} V_2(\boldsymbol{x}_{ij_1}, \boldsymbol{x}_{ij_2})$$
$$+ \ \ldots \ + \sum_{j_1<\cdots<j_N}^{M} V_N(\boldsymbol{x}_{ij_1}, \ldots, \boldsymbol{x}_{ij_N})$$

- ACE expansion:

$$E_i(X_i) = U_0(\boldsymbol{x}_i) + \sum_{j_1=1}^{M} U_1(\boldsymbol{x}_{ij_1}) + \sum_{j_2=1}^{M} \sum_{j_1=1}^{M} U_2(\boldsymbol{x}_{ij_1}, \boldsymbol{x}_{ij_2})$$
$$+ \ \ldots \ + \sum_{j_1,\ldots,j_N}^{M} U_N(\boldsymbol{x}_{ij_1}, \ldots, \boldsymbol{x}_{ij_N})$$

(spurious self-interactions added, worse scaling, so why bother?)

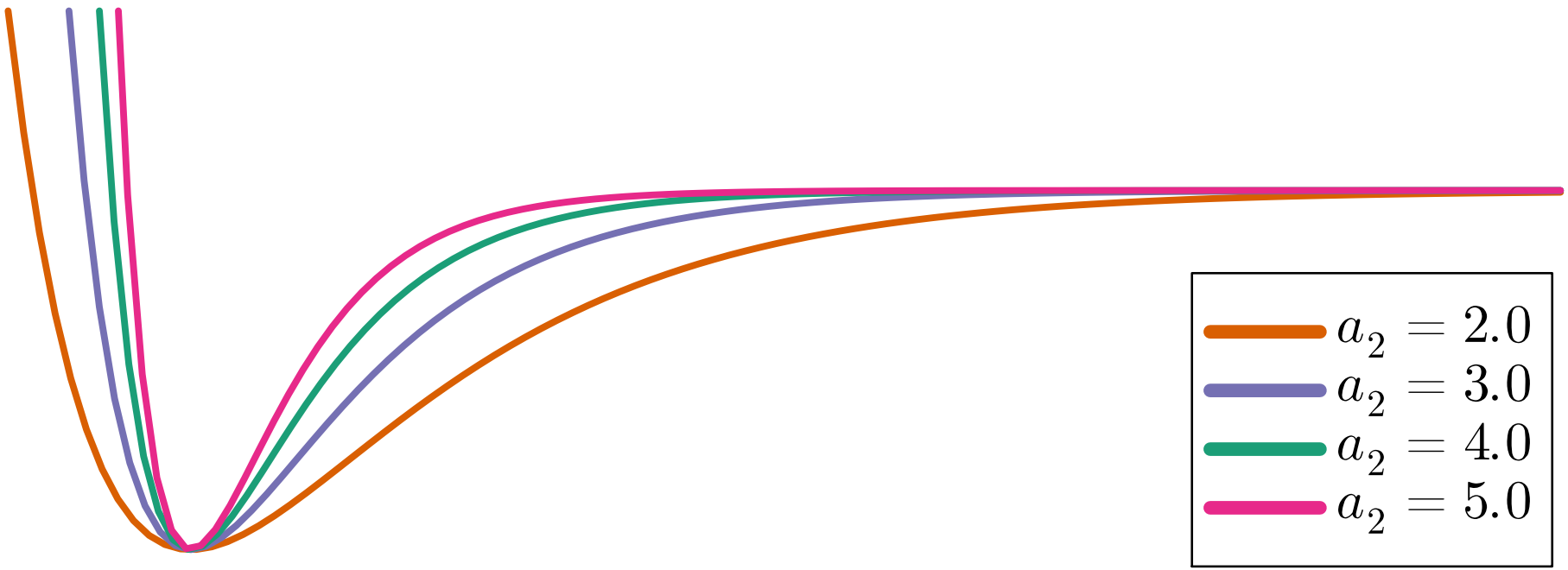
HISTORICAL PERSPECTIVE: EMPIRICAL POTENTIALS

- Pair potential (Morse):

$$E_i(X_i) = \sum_{j=1}^{M} V_1(\boldsymbol{x}_{ij}), \quad V_1(\boldsymbol{x}) = \phi(|\boldsymbol{x}|),$$

$$\phi(r) = a_1 \left(\exp\Big(-2a_2(r - a_3)\Big) - 2\exp\Big(-a_2(r - a_3)\Big) \right),$$

$$\boldsymbol{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$$



HISTORICAL PERSPECTIVE: EMPIRICAL POTENTIALS

- Many body potential (Stillinger-Weber):

$$E_i(X_i) = \sum_{j=1}^{M} V_1(\boldsymbol{x}_{ij}) + \sum_{j=1}^{M}\sum_{k=1}^{j-1} V_2(\boldsymbol{x}_{ij}, \boldsymbol{x}_{ik}),$$

$$V_1(\boldsymbol{x}_{ij}) = \psi(|\boldsymbol{x}_{ij}|), \quad \psi(r) = (a_1 r^{-a_2} - a_3 r^{-a_4}) \exp\left(\frac{1}{r - a_5}\right)$$

$$V_2(\boldsymbol{x}_{ij}, \boldsymbol{x}_{ik}) = \eta(|\boldsymbol{x}_{ij}|, |\boldsymbol{x}_{ik}|, \theta_{ijk})$$

$$\eta(r, t, \theta) = a_6 \exp\left(\frac{a_7}{r - a_8} + \frac{a_9}{t - a_{10}}\right) (\cos\theta_{ijk} + a_{11})^2$$

$$\boldsymbol{a} = (a_1, \ldots, a_{11}) \in \mathbb{R}^{11}$$

MANY-BODY (CLUSTER) EXPANSION

$E(X) = \sum_{i=1}^{M} E_i(X_i)$

- Naive expansion:

$$E_i(X_i) = V_0(\boldsymbol{x}_i) + \sum_{j_1=1}^{M} V_1(\boldsymbol{x}_{ij_1}) + \sum_{j_2=1}^{M} \sum_{j_1=1}^{j_2-1} V_2(\boldsymbol{x}_{ij_1}, \boldsymbol{x}_{ij_2})$$
$$+ \ \ldots \ + \sum_{j_1<\cdots<j_N}^{M} V_N(\boldsymbol{x}_{ij_1}, \ldots, \boldsymbol{x}_{ij_N})$$

- ACE expansion:

$$E_i(X_i) = U_0(\boldsymbol{x}_i) + \sum_{j_1=1}^{M} U_1(\boldsymbol{x}_{ij_1}) + \sum_{j_2=1}^{M} \sum_{j_1=1}^{M} U_2(\boldsymbol{x}_{ij_1}, \boldsymbol{x}_{ij_2})$$
$$+ \ \ldots \ + \sum_{j_1,\ldots,j_N}^{M} U_N(\boldsymbol{x}_{ij_1}, \ldots, \boldsymbol{x}_{ij_N})$$

(spurious self-interactions added, worse scaling, so why bother?)

EXPLOITING TENSOR PRODUCT STRUCTURE

- Each $U_N : \mathbb{R}^{3N} \to \mathbb{R}$ is to be represented as a linear combination in a tensor product basis:

$$U_N(\boldsymbol{x}_{ij_1}, \dots, \boldsymbol{x}_{ij_N}) = \sum_{\boldsymbol{k}} a_{\boldsymbol{k}} \prod_{t=1}^{N} \phi_{k_t}(\boldsymbol{x}_{ij_t})$$

  where $\boldsymbol{k} = (\boldsymbol{n}, \boldsymbol{l}, \boldsymbol{m}) \in \mathbb{R}^{3N}$ and

$$\phi_{(n_t,l_t,m_t)}(\boldsymbol{x}_{ij_t}) = \underbrace{P_{n_t}(|\boldsymbol{x}_{ij_t}|)}_{\text{radial part}} \underbrace{Y_{l_t}^{m_t}\left(\frac{\boldsymbol{x}_{ij_t}}{|\boldsymbol{x}_{ij_t}|}\right)}_{\text{angular part}}$$

- Angular part is given by complex spherical harmonics

$$Y_l^m(\hat{x}) = P_l^m(\cos\theta)\exp(im\phi), \quad \hat{x} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta) \in \mathbb{R}^3$$

- $n_t, l_t \in \{0, 1, 2, \dots\}$ and $m_t \in \{-l_t, \dots, 0, \dots, l_t\}$
- Truncation: sum over all $\boldsymbol{k}$ such that $|\boldsymbol{k}|_1 \leq T$ (defines set $\mathcal{A}^N$)

DENSITY TRICK FOR COMPUTATIONAL TRACTABILITY

- Fubini's Theorem aka "density trick":

$$\sum_{j_1,\dots,j_N=1}^{J} \sum_{\boldsymbol{k}\in\mathcal{A}^N} a_{\boldsymbol{k}} \prod_{t=1}^{N} \phi_{k_t}(\boldsymbol{x}_{ij_t}) = \sum_{\boldsymbol{k}\in\mathcal{A}^N} a_{\boldsymbol{k}} \prod_{t=1}^{N} \sum_{j=1}^{J} \phi_{k_t}(\boldsymbol{x}_{ij})$$

- meaning that overall

$$\begin{aligned} E_i(X_i) &= \sum_{N=1}^{\mathcal{N}} \sum_{j_1,\dots,j_N=1}^{M} U_N(\boldsymbol{x}_{ij_1},\dots,\boldsymbol{x}_{ij_N}) \\ &= \sum_{N=1}^{\mathcal{N}} \sum_{\boldsymbol{k}\in\mathcal{A}^N} a_{\boldsymbol{k}} \underbrace{\prod_{t=1}^{N} \sum_{j=1}^{J} \phi_{k_t}(\boldsymbol{x}_{ij})}_{=:A_{\boldsymbol{k}}^{(i)}} \\ &= \boldsymbol{a} \cdot A^{(i)} \end{aligned}$$

$\rightarrow$ a linear model (!) with parameters $\boldsymbol{a}$

$(\boldsymbol{a} \in \mathbb{R}^{\alpha},\ \alpha \sim 1000 - 100000)$

LAST STEP: ENFORCEMENT OF SYMMETRIES

- Recall that $A^{(i)}_{\boldsymbol{k}} = \prod_{t=1}^{N} \sum_{j=1}^{J} \phi_{k_t}(\boldsymbol{x}_{ij})$
- Basic idea is to "average over G":

$$B^{(i)}_{\boldsymbol{k}} := \int_G A^{(i)}_{\boldsymbol{k}} \circ gH(dg).$$

- Possible analytically (or with simple numerical scheme) if the basis is compatible with group G:

$$\phi_{k_t}(g(\boldsymbol{x}_{ij_t})) = \sum_{k'} \underbrace{\alpha_{k_t k'_t}(g)}_{\in \mathbb{R}} \phi_{k'_t}(\boldsymbol{x}_{ij_t})$$

  (applying group element to "relative particle" the same as a linear combination of the original basis).
- Bottom line:

$$B^{(i)} = \mathcal{C} A^{(i)}$$

  where $\mathcal{C}$ are (generalised) Clebsch-Gordan coefficients ($\rightarrow$ can be precomputed, typically very sparse)

GROUND TRUTH: A "PRIMER" ON QUANTUM MECHANICS

- In the Born-Oppenheimer approximation and further Hartree-Fock approximation, the ground energy of a system of M atomic nuclei at positions $\boldsymbol{R} = \{R_I\}_{I=1}^{M}$ with N electrons is given by

$$\mathcal{E}^{\mathrm{QM}}(\boldsymbol{R}) \approx \min_{\{\Psi_1,\dots,\Psi_N\}} \left\{ \langle \hat{H}\psi^{\mathrm{HF}}, \psi^{\mathrm{HF}} \rangle \mid \langle \psi_i, \psi_j \rangle = \delta_{ij} \right\}.$$

- Hamiltonian operator $\hat{H}$ is given by

$$\hat{H} = \Big(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{\substack{i,j=1 \\ j\neq i}}^{N} \frac{1}{|r_i - r_j|} - \sum_{i=1}^{N}\sum_{I=1}^{M} \frac{Z_I}{|r_i - R_I|} + \frac{1}{2} \sum_{\substack{I,J=1 \\ J\neq I}}^{M} \frac{Z_I Z_J}{|R_I - R_J|} \Big)$$

- The Euler-Lagrange equation is a nonlinear eigenvalue problem. By-product: Forces $\mathcal{F}^{\mathrm{QM}}(\boldsymbol{R}) = \nabla \mathcal{E}^{\mathrm{QM}}(\boldsymbol{R})$.

[VERY EXPENSIVE COMPUTATION]

- Training data: $\{\boldsymbol{R}^{(i)}, \mathcal{E}^{\mathrm{QM}}(\boldsymbol{R}^{(i)}), \mathcal{F}^{\mathrm{QM}}(\boldsymbol{R}^{(i)}), \dots)\}_i$

TRAINING

- Let $\mathcal{R} = \{\boldsymbol{R}^{(i)}\}_{i=1}^{J}$ be the training set of atomic configurations and $\mathcal{E}_{\mathcal{R}} = \{\mathcal{E}_i^{\mathrm{QM}}\}_{i=1}^{J}$ the corresponding "true" energy and $\mathcal{F}_{\mathcal{R}} = \{\mathcal{F}_i^{\mathrm{QM}}\}_{i=1}^{N}$ the "true" forces.
- Least squares $\min_{\boldsymbol{a}} I(\boldsymbol{a})$ where

$$I(\boldsymbol{a}) = \sum_{i=1}^{J} \left(W_E^2 |E_{\boldsymbol{a}}(\boldsymbol{R}^{(i)}) - \mathcal{E}_i|^2 + W_F^2 |F_{\boldsymbol{a}}(\boldsymbol{R}^{(i)}) - \mathcal{F}_i|^2 \right)$$

- Rewrite as $I(\boldsymbol{a}) = \|\mathrm{B}\boldsymbol{a} - Y\|_{Q^{-1}}^2$, where $\mathrm{B} : \mathbb{R}^{\alpha} \to \mathbb{R}^{(3M+1)J}$ and $\{\mathrm{B}\boldsymbol{a}\}_{i=1}^{M} = \{E_{\boldsymbol{a}}(R^{(i)})\}_{i=1}^{M}$ and the remaining entries are forces and

$$Q^{-1} = \mathrm{diag}(\underbrace{W_E^2, \ldots, W_E^2}_{J-\text{times}}, \underbrace{W_F^2, \ldots, W_F^2}_{3MJ-\text{times}})$$

- Tikhonov regularisation $\min_{\boldsymbol{a}} \tilde{I}(\boldsymbol{a})$ where

$$\tilde{I}(\boldsymbol{a}) = \|\mathrm{B}\boldsymbol{a} - Y\|_{Q^{-1}}^2 + \|\boldsymbol{a} - \overline{\boldsymbol{a}}\|_{P^{-1}}^2.$$

- (connection to Bayesian inference)

SUMMARY

- Atomic Cluster Expansion is an instance of a Geometric Shallow Learning framework on graphs.
- It uses a decades-old idea of a body order expansion and makes it computationally tractable in the age of enormous computational resources.
- It prescribes geometric prior information which allows an explicit enforcement of symmetries and physical constraints and assumptions.
- It works amazingly well for atom-based systems!
- Perhaps there is scope to employ it in agent-based models?

- Robust UQ is surprisingly difficult (my on-going work on this)
  [MARS Summer Internship]

REFERENCES

ACE:

1. Drautz, R., 2019. Atomic cluster expansion for accurate and transferable interatomic potentials. Physical Review B, 99(1), p.014104.
2. Dusson, G., Bachmayr, M., Csányi, G., Drautz, R., Etter, S., van Der Oord, C. and Ortner, C., 2022. Atomic cluster expansion: Completeness, efficiency and stability. Journal of Computational Physics, 454, p.110946.
3. Batatia, I., Benner, P., Chiang, Y., Elena, A.M., Kovács, D.P., Riebesell, J., Advincula, X.R., Asta, M., Avaylon, M., Baldwin, W.J. and Berger, F., 2023. A foundation model for atomistic materials chemistry. arXiv preprint arXiv:2401.00096.
4. Ortner, C., 2023. On the Atomic Cluster Expansion: interatomic potentials and beyond. arXiv preprint arXiv:2308.06462.
5. Batatia, I., Kovacs, D.P., Simm, G., Ortner, C. and Csányi, G., 2022. MACE: Higher order equivariant message passing neural networks for fast and accurate force fields. Advances in neural information processing systems, 35, pp.11423-11436.

Other:

1. Müser, M.H., Sukhomlinov, S.V. and Pastewka, L., 2023. Interatomic potentials: Achievements and challenges. Advances in Physics: X, 8(1), p.2093129.
2. Morse, P.M., 1929. Diatomic molecules according to the wave mechanics. II. Vibrational levels. Physical Review, 34(1), p.57.
3. Stillinger, F.H. and Weber, T.A., 1985. Computer simulation of local order in condensed phases of silicon. Physical Review B, 31(8), p.5262.