CECAM - EPFL - 20-24 June 2022

Interatomic Potentials from First Principles

Christoph Ortner UBC



Quantum Chemistry





Acknowledgements: L Zhang (UBC), M Bachmayr (Mainz), H Chen (BNU), G Dusson (CNRS), J Kermode, Jack Thomas, Reinhard Maurer, (Warwick); G Csanyi, C van der Oord, D Covacs (Cambridge), R Drautz + group (Bochum)

DFT





QM



Interatomic potentials



Coarse grained molecular

ML for coarse-graining materials and molecules, POV, review, outlook









Molecular Mechanics Simulation

 $\mathbf{R} = {\{\mathbf{r}_i\}}_{i=1}^N \dots$ particle positions $E(\mathbf{R})$... potential energy

Which level of theory?







Interatomic potentials



 \Rightarrow

Coarse grained molecular



Finite elements

Interatomic Potential?

 $\sum_{i} F\left(\sum_{j} \rho(r_{ij})\right)$





Dimension Reduction Coarse-Graining



Quantum Chemistry

Electronic Structure $H(\mathbf{R}, \{\psi_k\})\psi_k = \epsilon_k \psi_k$ OFFLINE Suppose we only want mechanics ... $E^{qm}(\mathbf{R}) = \mathscr{F}(\mathbf{R}, \{\epsilon_k\}, \{\psi_k\})$

Today's talk: how (multi-scale) analysis techniques can help understand or even support ML-backed coarse-graining? Towards an end-to-end theory ...





Interatomic Potentials



ML: $\varepsilon(\mathbf{R}_i) :=$ universal approximator



until ca 1995: ad hoc modelling, some formal asymptotics

NOT SYSTEMATIC!

2005 onwards: Machine-Learning revisit interatomic potentials from this new perspective!



Transferrability via Spatial Decomposition

We want to perform large-scale materials simulations that cannot be done with DFT; O(10k) to O(1M) atoms

Parameterisation of Site Energy:

 $E(\mathbf{R}) = \sum \varepsilon(\mathbf{R}_i) = \sum \varepsilon(\mathbf{c}; \mathbf{R}_i)$

can train on small configurations!!!







Rigorous Approximation Result

$H(\mathbf{R}, \{\psi_k\})\psi_k = \epsilon_k \psi_k$ $E^{\text{qm}}(\mathbf{R}) = \mathscr{F}(\mathbf{R}, \{\epsilon_k\}, \{\psi_k\})$

Our Claim:

$$E^{\text{QM}}(\boldsymbol{R}) = \sum_{i} \varepsilon^{\text{QM}}(\boldsymbol{R}_{i})$$
$$\left|\frac{\partial \varepsilon^{\text{QM}}(\{\boldsymbol{r}_{ij}\}_{j})}{\partial \boldsymbol{r}_{im}}\right| \leq C \exp(-cr_{im})$$

NB: $C = C(\beta, gap)$ as well,

Theorem [Chen, Thomas, CO, 2019] [Thomas, 2020] For a wide range of tight-binging models (implicitly assuming charge screening) (I) c ~ gap + β^{-1} for $\beta \in (1, \infty]$ (2) Point spectrum (defect) does not affect the result. (3) For nonlinear models, provided dielectric response stable





Curse of Dimensionality

 $E^{\text{QM}}(\boldsymbol{R}) = \sum_{i} \varepsilon^{\text{QM}}(\boldsymbol{R}_{i})$

Corollary: We can approximate

In practice: $\#R_i$ contains ca 30 to 100 atoms high-dimensional approximation



$$\frac{\partial \varepsilon^{\text{QM}}(\{\boldsymbol{r}_{ij}\}_{j})}{\partial \boldsymbol{r}_{im}} \leq C \exp(-cr_{im})$$

$$\varepsilon^{\text{QM}}\left(\left\{\boldsymbol{r}_{ij}\right\}_{j}\right) = \varepsilon\left(\left\{\boldsymbol{r}_{ij}\right\}_{r_{ij} < r_{\text{cut}}}\right) + O\left(e^{-cr_{\text{cut}}}\right)$$

Body-order Expansion (HDMR, ANOVA, ...)

$$\varepsilon^{(N)}(\mathbf{R}_{i}) := \varepsilon^{(0)} + \sum_{j_{1}} \varepsilon^{(1)}(\mathbf{r}_{ij}) + \sum_{j_{1} < j_{2}} \varepsilon^{(2)} + \sum_{j_{1} < j_{2} < j_{3}} \varepsilon^{(3)}(\mathbf{r}_{ij_{1}}, \mathbf{r}_{ij_{2}}, \mathbf{r}_{ij_{3}}) + \sum_{j_{1} < j_{2} < j_{3}} \varepsilon^{(3)}(\mathbf{r}_{ij_{1}}, \mathbf{r}_{ij_{2}}, \mathbf{r}_{ij_{3}})$$

- might converge very slowly ... this is based on the vacuum cluster expansion

e.g. Stillinger-Weber, MEAM, ... always truncated and N = 2i.e. body-order = 3

 $(1)(r_{ij_1}r_{ij_2},\cos\theta_{ij_1j_2})$

+···+ $\sum \varepsilon^{(N)}(\mathbf{r}_{ij_1}, \dots, \mathbf{r}_{ij_N})$ body-order = N + 1 $j_1 < \cdots < j_N$

- For MLIPs: truncating at $N \ll \#\mathbf{R}_i$ reduces dimensionality of the parameterisation - Much easier to reason about approximation of $\varepsilon^{(N)}$, since defined on \mathbb{R}^{dN} .

- conflicting folklore ... general belief this expansion need not always converge and often









Rapid Converging Many-Body Expansion

$$|\varepsilon^{\mathrm{QM}}(\mathbf{R}_i) - \varepsilon^{(N)}(\mathbf{R}_i)| \leq Ce^{-cN}$$

Theorem [Chen, Thomas, CO, 2021] For linear tight-binding models $\mathbf{c} \sim \mathbf{gap} + 1/\beta$ for all configurations R_i within one crystalline phase of the material

NB: $C = C(\beta, gap)$ as well, more complex.



Classical potential theory (e.g., Ransford, Sat generalisation to non-interval sets (closely related to BOP theory)

Modelling: Atomic Body-Order Expansion

$$E(\mathbf{R}) = \sum_{i} \varepsilon(\mathbf{R}_{i})$$

$$\varepsilon^{(N)}(\mathbf{R}_{i}) := \varepsilon^{(0)} + \sum_{j_{1}} \varepsilon^{(1)}(\mathbf{r}_{ij}) + \sum_{j_{1} < j_{2}} \varepsilon^{(2)}(\mathbf{r}_{ij_{1}}\mathbf{r}_{ij_{2}}, \cos \theta_{ij_{1}j_{2}}) \qquad \begin{array}{l} \textbf{3-bod}\\ \text{e.g. Stil} \end{array}$$

$$+ \sum_{j_{1} < j_{2} < j_{3}} \varepsilon^{(3)}(\mathbf{r}_{ij_{1}}, \mathbf{r}_{ij_{2}}, \mathbf{r}_{ij_{3}}) + \dots + \sum_{j_{1} < \dots < j_{N}} \varepsilon^{(N)}(\mathbf{r}_{ij_{1}}, \dots, \mathbf{r}_{ij_{N}})$$

$$\operatorname{cost} \sim (\#\mathbf{R}_{i})^{3} \qquad \operatorname{cost} \sim (\#\mathbf{R}_{i})^{\nu}/\nu!$$

- the exponential scaling severely limits the utility of the body-order expansion! traditionally truncated at 3-body, with highly simplified 4-body for molecules

3-body : cost ~ $(\#R_i)^2$ llinger-Weber, MEAM, ...

 $+O(e^{-cN})$

 $COST \sim (\#\mathbf{K}_i)^2 / \mathcal{U}!$

+ additional cost of approx. in $\mathbb{R}^{3\nu}$

11



Atomic Cluster Expansion (1) Drautz, 2019; Seko et al 2019; Dusson et al, 2019 ...

$$\varepsilon^{(N)}(\mathbf{R}_{i}) = \varepsilon^{(0)} + \sum_{j_{1}} \varepsilon^{(1)}(r_{ij}) + \sum_{j_{1} < j_{2}} \varepsilon^{(2)}(r_{ij_{1}})$$
$$= U^{(0)} + \sum_{j_{1}} U^{(1)}(r_{ij}) + \sum_{j_{1}, j_{2}} U^{(2)}(\mathbf{r}_{j_{1}}),$$
(permutation invariance)

 $\bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet$

"canonical formulation"

(an aside) $r_{ij_2}, \cos \theta_{ij_1j_2}) + \dots + \sum_{j_1 < \dots < j_N} \varepsilon^{(N)}(r_{j_1}, \dots, r_{j_N})$ $(r_{j_2}) + \cdots + \sum_{j_1,\dots,j_N} U^{(N)}(r_{j_1},\dots,r_{j_N})$



"self-interaction formulation" can exploit the • • • • • • • tensor product structure!



Atomic Cluster Expansion (2)

 $U^{(N)}(\mathbf{r}_{j_1},...,\mathbf{r}_{j_N})$ j_1,\ldots,j_ν $\approx \sum c_{k_1 \cdots k_N} \prod \phi_{k_\alpha}(\mathbf{r}_{j_\alpha})$ $j_1, ..., j_N k_1, ..., k_N$ $\alpha = 1$ $= \sum c_{k_1 \cdots k_N} \sum \prod \phi_{k_\alpha}(\mathbf{r}_{j_\alpha})$ $k_1, ..., k_N$ $j_1, ..., j_N \alpha = 1$ $= \sum c_{k_1 \cdots k_N} \prod \sum \phi_{k_\alpha}(\mathbf{r}_j)$ k_1,\ldots,k_N $\alpha=1$ j



Expand $U^{(N)}(r)$ in a tensor product basis

$$\mathbf{cost} \sim \#(k_{\alpha})^{N} + \#k \times \#\mathbf{R}_{i}$$

Exploit tensor product structure to exchange \sum and \Box

$$cost \sim \#(k_{\alpha}) \times N + \#k \times \#R_i$$



Atomic Cluster Expansion (3)

[1] Atomic basis / density projection:

$$A_{v} := \sum_{j} \phi_{v}(\mathbf{r}_{ij})$$

[2] Symmetric basis / N-Correlations:

$$A_{v_1,\ldots,v_N} := \prod_t A_{v_t}$$

[3] ACE Basis / Symmetry-adapted N-Correlations:

$$X_{v_1,\ldots,v_N} := \int_{O(3)} A_{v_1,\ldots,v_N} \circ Q \ dQ$$

Equivariance: if we require $\phi \circ Q =$ then simply modify st

Bartok, Csanyi, Kondor, 2010 Ceriotti et al, ..., 2020, 2021 Drautz, 2019 Dusson, ..., CO, 2019

 \mathbb{R}^3 basis functions $\phi_{v}(\mathbf{r}) := \phi_{nlm}(\mathbf{r}) = p_{n}(r)Y_{l}^{m}(\hat{\mathbf{r}})$

(to explicitly resolve the O(3) integral)

Thm: basis of permuation-invariant multi-set functions

$$\Rightarrow \quad X_i = \mathscr{C} \cdot A(R_i)$$

where \mathscr{C} is sparse (Clebsch-Gordan coefficients) classical representation theory of O(3)

$$\begin{array}{l} Q[\varphi] \\ ep_{14}[3]: \end{array} \quad X_{v} := \int_{O(3)} Q[e_{\alpha}] A_{v} \circ Q \ dQ \end{array}$$





Approximation with Invariant Polynomials (earlier results by Shapeev, 2016)

Approx.Thy.: Bachmayr, Chen, Dusson, Thomas, CO [2021] Under the assumptions of our previous results ... $\|\varepsilon^{\mathrm{qm}} - \varepsilon^{\mathrm{ace}}\|_{\infty} \lesssim e^{-a_1 r_{\mathrm{cut}}} + e^{-a_2 N} + e^{-a_3 D}$ how to balance errors, ... cost estimates, $\|\varepsilon^{qm} - \varepsilon^{ace}\|_{\infty} \lesssim \exp\left(-\gamma \frac{[\log PARAMS]^{4/3}}{\log \log PARAMS}\right)$

Results quantify how approx. symmetric functions is much easier. Intuitively $(N!)^{-1}$ less information, but technically one needs to work with generalisations of integer partitions [Hardy, Ramanujan, 1918]





Systematic Convergence Test - Si training set



Potential for Si (I)





GAP - PRX 2018 DFT - CASTEP Cvd Oord, 2021

T (K)





Molecules



David Kovacs et al, 2021





Example: MoNbTaW





Cas van der Oord 14:30 to 15:00 - Hyperactive Learning (HAL)

- Phase transitions found using Nested Sampling
- Nested Sampling integrates partition function by exploring entire configurational space
- Requires potentials to be accurate, robust and fast (~ ms/atom per force call)
- ACE potential using HAL database



Physics-Informed Nonlinearities

 $\varepsilon(\boldsymbol{R}_i) := \sum V(r_{ij}) + F\left(\sum_i \rho(r_{ij})\right)$ EAM:

Drautz (2019) : (EAM, FS, BOP inspired)

Anton Bochkarev[®], Yury Lysogorskiy[®], Sarath Menon[®], Minaam Qamar, Matous Mrovec[®], and Ralf Drautz[®] ICAMS, Ruhr-Universität Bochum, Bochum, Germany

 $\varepsilon(\boldsymbol{R}_i) = \varphi_1(\boldsymbol{R}_i) - \sqrt{\varphi_2(\boldsymbol{R}_i)}$

where each φ_i is an ACE (linear parameterisation of an atomic many-body expansion)

overarching idea: decompose a property into local, low-body-order, smooth components

Efficient parametrization of the atomic cluster expansion





Physics-Informed Nonlinearities overarching idea: decompose a property into local, low-body-order, smooth components

Drautz (2019) : (EAM, FS, BOP inspired)

Recall the truncation error of body-order expansion:

 $|\varepsilon^{\mathrm{QM}}(\mathbf{R}_i) - \varepsilon^{(N)}(\mathbf{R}_i)| \leq Ce^{-cN}$

Theorem [Chen, Thomas, CO, 2021] $\varepsilon^{(N)} = \text{linear ACE model, then}$ $c \sim gap + 1/\beta$ for all configurations R_i within one crystalline phase



Performant implementation of the atomic cluster expansion (PACE): Application to copper and silicon



Reproducing a benchmark of

Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. P. Ong, J. Phys. Chem. A 124, 731 (2020).

Yury Lysogorskiy,¹ Cas van der Oord,² Anton Bochkarev,¹ Sarath Menon,¹ Matteo Rinaldi,¹ Thomas Hammerschmidt,¹ Matous Mrovec,¹ Aidan Thompson,³ Gábor Csányi,² Christoph Ortner,⁴ and Ralf Drautz¹



 nonlinear ACE: pacemaker (Drautz group) Iinear:ACEI.jl

Coarse-graining across all the scales

 $\ddot{u} = -\nabla F(u) + \Gamma(u)\Gamma(u)^T \dot{u} + \Gamma(u)\dot{W}$



e.g. tight-binding approximation



DFT



e.g. mean field approximation

Quantum

Monte Carlo



Empirical QM



Interatomic potentials



Quantum Chemistry



Coarse-graining across all the scales: Self-Consistent Hamiltonian

$H = H(\mathbf{R})$

e.g. tight-binding approximation

QM





Monte Carlo





Quantum Chemistry

Interatomic potentials



L. Zhang, B. Onat, G. Dusson, G. Anand, R.J. Maurer, CO and JRK, arXiv:2111.13736

Analysis - Jack Thomas, H Chen, CO, 2021



Finite elements





Coarse grained molecular

Example: Self-Consistent Hamiltonian

Start from KS-DFT with Atomic Orbital Basis

 $H = -\Delta + V_{\text{eff}}$

 $H\psi_a = \epsilon_a \psi_a$

 $\rho_i = \sum_a [\psi_a]_i^2$ Yukawa $[V_{\text{eff}}]_{i} = \mathscr{V}[\rho_{i}] + \text{Coulomb}$

Same principles apply as for site potentials



\Rightarrow Parameterize the Hamiltonian by an Equivariant ACE Model

ACE Param. of a Self-Consistent Hamiltonian

E.g. KS-DFT (FHlaims) "Learn" the self-consistent Hamiltonian in atomic orbital basis

$$H = -\Delta + V_{\rm eff}$$

 $H\psi_i = \epsilon_i \psi_i$ $V_{\text{eff}} = V_{\text{eff}} [(\psi_i)_i]$

 $\mathscr{H}^{\alpha\beta} \circ Q = D'$



L. Zhang, B. Onat, G. Dusson, G. Anand, R.J. Maurer, C. Ortner ang/JRK, arXiv:2111.13736

$$H_{IJ}^{\alpha\beta} = \mathcal{H}^{\alpha\beta}(\mathbf{R}_{IJ})$$
 where \mathbf{R}_{IJ} = bond environmed

$$\alpha \mathcal{H}^{\alpha\beta}(D^{\beta})* \Rightarrow$$
 use equivariant ACE expansion

same principles apply: locality & many-body expansion \Rightarrow Expand each block as an ACE

$$\mathscr{H}^{\alpha\beta}(\mathbf{R}_{IJ}) = \sum_{v} c_{v} X_{v}^{\alpha\beta}(\mathbf{R}_{IJ})$$

equivariant ACE basis

ment ansion

Prediction Accuracy



L. Zhang, B. Onat, G. Dusson, G. Anand, R.J. Maurer, C. Ortner and JRK, arXiv:2111.13736

Fit AI hamiltonian to bcc and fcc MD data, predict band structure and DOS (fit to FHIaims)





Y Wang, CO, in prep

predict on domain $[0,M]^3$, M large

[1] A general (but useless) result: force error, simulation domain $\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \lesssim \|\nabla \boldsymbol{E}_{M}^{\text{QM}}(\boldsymbol{R}_{M}^{\text{QM}}) - \nabla \boldsymbol{E}_{M}^{\text{ML}}(\boldsymbol{R}_{M}^{\text{QM}})\|$

remember we train on small $[0,L]^3$ domain, $L \ll M$



+ perturbations



Error in Terms of Training Data

- [1] A general but useless result:
- [2] Split Force Error into Training Error + Interaction:

Dislocations

Point Defects

$$\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \lesssim \|\nabla \boldsymbol{E}_{L}^{\text{QN}}\|_{E}$$



force error, simulation domain $\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \lesssim \|\nabla \boldsymbol{E}_{M}^{\text{QM}}(\boldsymbol{R}_{M}^{\text{QM}}) - \nabla \boldsymbol{E}_{M}^{\text{ML}}(\boldsymbol{R}_{M}^{\text{QM}})\|$





Error in Terms of Training Data

- [1] A general but useless result:
- [2] Split Force Error into Training Error + Interaction:

Dislocations

Point Defects

$$\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \lesssim \|\nabla \boldsymbol{E}_{L}^{\text{QN}}\|_{E}$$



force error, simulation domain $\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \leq \|\nabla \boldsymbol{E}_{M}^{\text{QM}}(\boldsymbol{R}_{M}^{\text{QM}}) - \nabla \boldsymbol{E}_{M}^{\text{ML}}(\boldsymbol{R}_{M}^{\text{QM}})\|$





Preliminary Example

Point Defects

 $\|\boldsymbol{R}_{M}^{\text{QM}} - \boldsymbol{R}_{M}^{\text{ML}}\|_{E} \lesssim \|\nabla \boldsymbol{E}_{L}^{\text{QM}}(\boldsymbol{R}_{L}^{\text{QM}}) - \nabla \boldsymbol{E}_{L}^{\text{ML}}(\boldsymbol{R}_{L}^{\text{QM}})\| + \boldsymbol{\mathcal{V}}^{3/2}$

Vacancies & Interstitials in EAM-W



remaining uncontrolled error

force error, training domain

$+L^{-3/2} \|FCM^{QM} - FCM^{ML}\| + L^{-9/2}$ force constants error





Conclusion: Coarse-graining with ML

- Classic analytic modelling
- keep track of approximation steps
- Use ML to fill in the gaps









DFT

Quantum Chemistry

Main GAP: Long-range charge equilibration - but see forthcoming work from Jack Thomas...

\Rightarrow end-to-end error control





Interatomic potentials



Coarse grained molecular

Empirical QM

empirical no longer ???



