Classical effective potentials – what to do and what to avoid

Peter Brommer

Warwick Centre for Predictive Modelling
University of Warwick

Molecular and Materials Modelling Summer School
28 September 2016

THE UNIVERSITY OF
WARWICK

WCPM

Motivation

Quantum theory of electrons and nuclei

QM (QED) is a theory with enormous predictive power:

- Energy levels of hydrogen atom to a few ppm.

Solution to all our modelling needs?

Motivation

Quantum theory of electrons and nuclei

QM (QED) is a theory with enormous predictive power:

- Energy levels of hydrogen atom to a few ppm.

Solution to all our modelling needs? . . . No.

Motivation

Quantum theory of electrons and nuclei

QM (QED) is a theory with enormous predictive power:

- Energy levels of hydrogen atom to a few ppm.

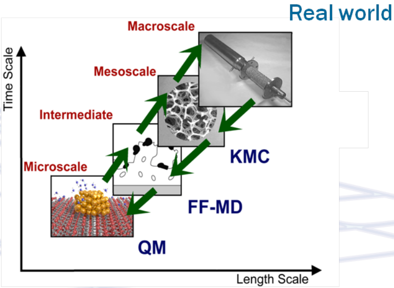
Solution to all our modelling needs? . . . No.

“Sequential” Multiscale Modelling

Use proper simulation tool for each scale.

- Parameterise from small to large.
- No direct coupling between models.

Propagate uncertainty from one scale to the next.



essenceofescience.se/nobel-2013/

Outline

1. Standard Sequential QM/MM Multiscale Modelling
   - Density Functional Theory
   - Molecular Dynamics
   - Effective Potentials
   - Force Matching

2. Uncertainty
   - Sources
   - Testing potentials
   - Quantification
   - Example

(Plane Wave) Density Functional Theory

The briefest DFT recap

Schrödinger Equation:

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle$$

$$\hat{H} = T_N + T_e + V_{NN}(\boldsymbol{R}) + V_{ee}(\boldsymbol{r}) + V_{Ne}(\boldsymbol{r}, \boldsymbol{R})$$

Born-Oppenheimer approximation:

$$[T_e + V_{ee}(\boldsymbol{r}) + V_{Ne}(\boldsymbol{r}, \boldsymbol{R})]\,\Psi(\boldsymbol{x}, \boldsymbol{R}) = \epsilon_n(\boldsymbol{R})\Psi(\boldsymbol{x}, \boldsymbol{R})$$

$$\frac{\partial \boldsymbol{P}_I}{\partial t} = -\nabla_I(\epsilon_0(\boldsymbol{R}) + V_{NN}(\boldsymbol{R})).$$

What if we don't care about what the electrons do, only about their effects on the nuclei?

Molecular Dynamics

What is Molecular Dynamics?

Equations of motion of a system of interacting particles are integrated numerically.

- Direct simulation of the basic laws of physics:
  Newton’s (or Hamilton’s) equations.

Molecular Dynamics

What is Molecular Dynamics?

Equations of motion of a system of interacting particles are integrated numerically.

- Direct simulation of the basic laws of physics:
  Newton's (or Hamilton's) equations.

Needed

- Initial condition: structure model
- Equation of motion: model of the interactions

Molecular Dynamics

What is Molecular Dynamics?

Equations of motion of a system of interacting particles are integrated numerically.

- Direct simulation of the basic laws of physics: Newton’s (or Hamilton’s) equations.

Needed

- Initial condition: structure model
- Equation of motion: model of the interactions

Big systems or long simulation times are feasible only with classical effective potentials.

Effective Potentials

Eliminate electronic degrees of freedom

Re-write equation of motion:

$$\frac{\partial \boldsymbol{P}_I}{\partial t} = -\nabla_I(\epsilon_0(\boldsymbol{R}) + V_{NN}(\boldsymbol{R})) = \nabla_I(\tilde{V}_{\mathrm{eff}}(\boldsymbol{R}))$$

Usually, V_{eff} much simpler. Expand in manybody contributions:

$$V_{\mathrm{eff}}(\boldsymbol{R}) = \sum_i \phi_1(\boldsymbol{r}_i) + \tfrac{1}{2}\sum_{\substack{i,j\\ i\neq j}} \phi_2(\boldsymbol{r}_i, \boldsymbol{r}_j) + \tfrac{1}{6}\sum_{\substack{i,j,k\\ i\neq j\neq k\neq i}} \phi_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \ldots$$

Even simpler: Central (homogeneous and isotropic) pair potentials:

$$V(\boldsymbol{R}) = \tfrac{1}{2}\sum_{\substack{i,j\\ i\neq j}} \phi_{ij}(r_{ij}), \qquad \text{with } r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$$

Interactions

Choice of interaction model depends on material to simulate

(different ways to truncate manybody expansion, different calculation prescriptions)

- Central pair potentials, EAM potentials for metals.
- Angular dependent potential (ADP), MEAM.
- Covalent potentials (Tersoff, Stillinger-Weber, . . .) for ceramics.
- Coulomb potential (Ewald method, Wolf summation).
- Dipolar interaction for oxides.
- Bond-order potentials, charge-optimising force fields.
- Temperature-dependent potentials (laser ablation).
- Coarse-grained potentials.
- Simulation of organic molecules: Force fields for polymer chains, water, amino acids,. . .

Example: EAM potentials

EAM potentials

(Semi-)empirical potentials:

$$E = \sum_{ij} \Phi_{ij}(r_{ij}) + \sum_{i} F_i(\rho_i), \quad \text{where} \quad \rho_i = \sum_{j} \Psi_j(r_{ij})$$

Bond strength depends on environment –
better suited to describe vacancies and other defects.

Example: EAM potentials

EAM potentials

(Semi-)empirical potentials:

$$E = \sum_{ij} \Phi_{ij}(r_{ij}) + \sum_i F_i(\rho_i), \quad \text{where} \quad \rho_i = \sum_j \Psi_j(r_{ij})$$

Bond strength depends on environment –
better suited to describe vacancies and other defects.

Analytic vs. tabulated potentials

- Φ, F, ρ are the potential functions.
- Can be represented by analytic functions, e.g. $F = c\sqrt{\rho}$.
- Alternative: Tabulated at sampling points.

Potential Development

How to obtain effective potentials?

Potential serves to determine energies and forces
→ determines the physics of the system!

Potential Development

How to obtain effective potentials?

Potential serves to determine energies and forces
→ determines the physics of the system!

- Depending on the system (metal, oxide, etc.), a suitable potential type must be chosen.
- Within such a potential family, the potential parameters determine the physical properties a particular material.
- The parameters are chosen such that the desired material properties are correctly reproduced.

Potential Development

How to obtain effective potentials?

Potential serves to determine energies and forces
→ determines the physics of the system!

- Depending on the system (metal, oxide, etc.), a suitable potential type must be chosen.
- Within such a potential family, the potential parameters determine the physical properties a particular material.
- The parameters are chosen such that the desired material properties are correctly reproduced.
- The material properties to be reproduced are often computed ab-initio, instead of measured experimentally.

Reference quantities

Experimental Quantities of Interest

e.g. phase diagrams, elastic properties, diffusivities. . .

- often expensive to calculate ⇒ “simple” systems only.

Reference quantities

Experimental Quantities of Interest

e.g. phase diagrams, elastic properties, diffusivities. . .

- often expensive to calculate ⇒ “simple” systems only.

Literature and first-principles quantities

e.g. bond lengths, heats of formation, bond angles.

- Sampling configurational space?

Example: ReaxFF.

Reference quantities

Experimental Quantities of Interest

e.g. phase diagrams, elastic properties, diffusivities. . .

- often expensive to calculate ⇒ “simple” systems only.

Literature and first-principles quantities

e.g. bond lengths, heats of formation, bond angles.

- Sampling configurational space?

Example: ReaxFF.

First-principles forces, stresses, energies

- easily calculated (1 MD step).
- sample configurational space.

Force Matching!

Ercolessi & Adams, Europhys. Lett. **26**, 583 (1994)

Force Matching with *potfit*

Open source force matching code ***potfit***

- Flexible and modular.
- Supports pair, (M)EAM, ADP potentials (metals).
- Oxide potentials.
- Electron-temperature dependent potentials (laser ablation).
- Interfaces to DFT and MD codes.

Widely used code

- 40 downloads/month,
- 60 citations with potentials,
- from more than ten distinct groups around the globe.

Brommer, Gähler, Model. Simul. Mater. Sci. Eng. **15**, 295 (2007).
Brommer *et al.*, Model. Simul. Mater. Sci. Eng. **23**, 074002 (2015)
`http://potfit.net/`

Potential Generation

1. Select potential model, starting potential.
2. Select reference structures (100–200 atoms, MD simulation at various temperatures, strained structures).
3. Calculate forces, stresses, energies with ab-initio code.
4. Optimize starting potential with *potfit*.
5. Generate reference structures with new potential.
   ⇒ more realistic configurations.
6. Test potential.

Potential Generation

1. Select potential model, starting potential.
2. Select reference structures (100–200 atoms, MD simulation at various temperatures, strained structures).
3. Calculate forces, stresses, energies with ab-initio code.
4. Optimize starting potential with *potfit*.
5. Generate reference structures with new potential.
   $\Rightarrow$ more realistic configurations.
6. Test potential.

If results are not satisfying

- use more/different reference configurations,
- replace insufficient potential model.

and iterate procedure.

Outline

1. Standard Sequential QM/MM Multiscale Modelling
   - Density Functional Theory
   - Molecular Dynamics
   - Effective Potentials
   - Force Matching
2. Uncertainty
   - Sources
   - Testing potentials
   - Quantification
   - Example

Uncertainties

Sources of uncertainty for force-matched potentials

Generic errors:

- “Imported” uncertainty: cannot beat DFT.
- Algorithmic uncertainty: global optimum?

Force Matching specific (structural & parameter) uncertainties:

- Bad reference data selection (parameter uncertainty).
- Wrong functional form (model bias).
- Overfitting (parameter uncertainty).
- Wrong potential model (model bias).

Uncertainties

Sources of uncertainty for force-matched potentials

Generic errors:

- “Impor
- Algorit

Force Matc

- Bad re
- Wrong
- Overfit
- Wrong

Target function $Z(N)$, N parameters in model

Uncertainties

Sources of uncertainty for force-matched potentials

Generic errors:

- “Imported” uncertainty: cannot beat DFT.
- Algorithmic uncertainty: global optimum?

Force Matching specific (structural & parameter) uncertainties:

- Bad reference data selection (parameter uncertainty).
- Wrong functional form (model bias).
- Overfitting (parameter uncertainty).
- Wrong potential model (model bias).

Uncertainties

Sources of uncertainty for force-matched potentials

Generic errors:

- “Imported” uncertainty: cannot beat DFT.
- Algorithmic uncertainty: global optimum?

Force Matching specific (structural & parameter) uncertainties:

- Bad reference data selection (parameter uncertainty).
- Wrong functional form (model bias).
- Overfitting (parameter uncertainty).
- Wrong potential model (model bias).

Properties of force-matched potentials:

- (Generally) good representability.
- Limited transferability.

Caveat emptor! Don’t just trust any potential.

Interpolated potentials

Special case: tabulated or interpolated potential

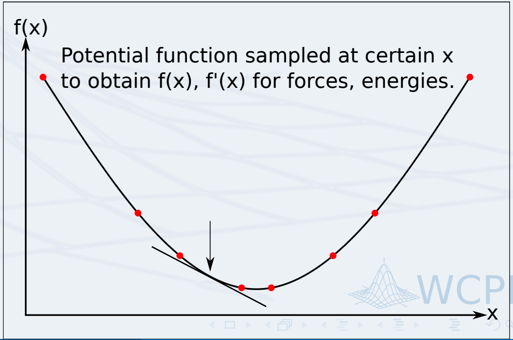
Interpolated potentials can have many parameters (>100).

- No bias from particular functional form.
- Parameters have no meaning.

No formal uncertainty propagation in literature.

Confidence of sampling point values

Forces: evaluate potential functions and gradients.



Potential function sampled at certain x to obtain f(x), f'(x) for forces, energies.

Interpolated potentials

Special case: tabulated or interpolated potential

Interpolated potentials can have many parameters (>100).

- No bias from particular functional form.
- Parameters have no meaning.

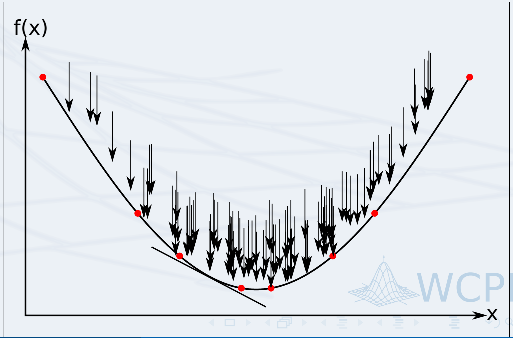
No formal uncertainty propagation in literature.

Confidence of sampling point values

Forces: evaluate potential functions and gradients.

- Both in training and use.

Training set and application: sample similarly.



Testing potentials

. . . or why is my simulation doing freaky things

Fe monovacancy migration

Iron: Well studied system (FP6: Perfect project)

Family of EAM potentials for iron (Ackland, Mendelev, Marinica)

- Optimised to DFT energies of defects (available from 2000 onwards).

Surprise:

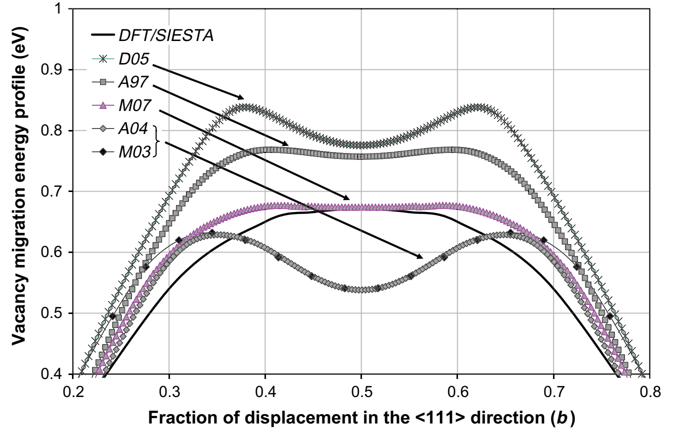
Fe monovacancy migration

Iron: Well studied system (FP6: Perfect project)

Family of EAM potentials for iron (Ackland, Mendelev, Marinica)

- Optimised to DFT energies of defects (available from 2000 onwards).

Surprise: Vacancy migration barrier.



Testing potentials

. . . or why is my simulation doing freaky things

You want to do a MD/MC/. . . simulation with a potential, but you don't yet have a potential. So you

- look on Google Scholar
- ask your supervisor/colleague
- run *potfit*

But does the potential actually do what you want it to do?

Testing potentials

. . . or why is my simulation doing freaky things

You want to do a MD/MC/. . . simulation with a potential, but you don't yet have a potential. So you

- look on Google Scholar
- ask your supervisor/colleague
- run *potfit*

But does the potential actually do what you want it to do?

What could go wrong?

Testing potentials

…or why is my simulation doing freaky things

You want to do a MD/MC/…simulation with a potential, but you don't yet have a potential. So you

- look on Google Scholar
- ask your supervisor/colleague
- run *potfit*

But does the potential actually do what you want it to do?

What could go wrong?

Surprisingly much (something frustratingly trivial):

- Potential not available for your MD/MC/…tool (⇒ conversion errors).
- Differences in implementation (cutoff handling, interpolation).
- Typos in publications (!!!).
- Not optimised for your situation

Two-fold problem

Implementation problems

Potential not working as designed.

- “Unit” testing? Not always trivial.
- Ask authors for potential?
- Potential databases?[a]

---

[a] e.g. `www.ctcms.nist.gov/potentials`

Potential problems

Potential not designed for use case.

- Validation of QoI.
- What are good QoI for my application?

Potential solution: OpenKIM Framework

API that provides

- Models (i.e. potentials) including implementation.
- Tests (i.e. QoI calculated from models) including reference data.
- Workflow to cross-reference (new models/tests trigger new test results).
- Interface to MD and *potfit*.

Standardised access to potentials with guaranteed results.

LAMMPS and potentials

LAMMPS comes free with a selection of potentials

There is a whole directory full of potentials.

Ag_u3 AlCu AlFe_mm AlO AlSiMgCuFe Al_jnp Al_mm Al_zhou Au_u3 BNC C CCu_v2 CH CdTe CdTeSe CdTeZnSeHgS0 CdZnTe_v1 CdZnTe_v2 CoAl Cu CuH CuNi CuTa CuTa_eam CuZr_mm Cu_mishin1 Cu_smf7 Cu_u3 Cu_u6 Cu_zhou FeCr FeP_mm Fe_mm GaAs GaN GaN_sw GaN_tersoff InP MOH Mg_mm Mo5 Ni NiAlH_jea Ni_smf7 Ni_u3 Pd_u3 Pt_u3 Si SiC SiCGe SiC_1989 SiC_1990 SiC_1994 SiC_Erhart-Albe SiO Si_1 Si_2 Ta06A Ta06A_pot Ta4 Ta6 Ti V6 VFe_mm W_zhou Zr_mm

LAMMPS and potentials

LAMMPS comes free with a selection of potentials

There is a whole directory full of potentials.

Ag_u3 AlCu AlFe_mm AlO AlSiMgCuFe Al_jnp Al_mm Al_zhou Au_u3 BNC C CCu_v2 CH CdTe CdTeSe CdTeZnSeHgS0 CdZnTe_v1 CdZnTe_v2 CoAl Cu CuH CuNi CuTa CuTa_eam CuZr_mm Cu_mishin1 Cu_smf7 Cu_u3 Cu_u6 Cu_zhou FeCr FeP_mm Fe_mm GaAs GaN GaN_sw GaN_tersoff InP MOH Mg_mm Mo5 Ni NiAlH_jea Ni_smf7 Ni_u3 Pd_u3 Pt_u3 Si SiC SiCGe SiC_1989 SiC_1990 SiC_1994 SiC_Erhart-Albe SiO Si_1 Si_2 Ta06A Ta06A_pot Ta4 Ta6 Ti V6 VFe_mm W_zhou Zr_mm

There is also a README (which is actually a disclaimer)

- Executive summary: Use at your own risk.

Q: How many simulations were made using those potentials just because they are there?

Formal UQ: Potential derivation as an inverse problem

Forward problem: data (e.g. forces) δ from parameters u

$$\delta = G(u)$$

- u: Parameter vector
- δ: Result vector

Well-posed problem.

Formal UQ: Potential derivation as an inverse problem

Forward problem: data (e.g. forces) δ from parameters u

$$\delta = G(u)$$

- u: Parameter vector
- δ: Result vector

Well-posed problem.

Inverse problem: parameters u from noisy data δ

$$\delta = G(u) + \eta$$

Ill-posed. Optimisation problem:

$$\min_u \frac{1}{2}||\delta - G(u)||^2$$

Potential derivation as a Bayesian inverse problem

Interpret u, η, δ as RV/random fields

Attack problem using “standard” techniques:

- MCMC methods
- MAP estimators
- Sparse deterministic approximations

⇒ Account naturally for noisy reference data.

Computationally involved (to say the least).

Potential derivation as a Bayesian inverse problem

Interpret u, η, δ as RV/random fields

Attack problem using “standard” techniques:

- MCMC methods
- MAP estimators
- Sparse deterministic approximations

$\Rightarrow$ Account naturally for noisy reference data.

Computationally involved (to say the least).

In the following:

- How has this been used so far (selection)?

The problem is two-fold:

1. Quantifying uncertainty in potential (inverse problem).
2. Propagate uncertainty to MD simulations (forward problem).

“Sloppy model” MD

Replace best-fit effective potential by ensemble

Conditional probability of parameter set u:

$$P(u|\delta, G) \propto \exp\left[-\frac{C(u)}{T}\right],$$

- cost function $C(u) = \frac{1}{2}||\delta - G(u)||^2$
- formal temperature T, with $T_0 = 2C_0/N_p$, $C_0 = \min_u C(u)$

For any observable O, calculate mean $\langle O\rangle|_{T,D,M}$, variance $\sigma_O^2|_{T,D,M}$.

Frederiksen *et al.*, PRL **93** 165501 (2004)

Sloppy model pros and cons

Advantages

- Integrates seamlessly into force matching/MD software stack:
  - Generate ensemble of potentials.
  - Run *N* independent MD simulations.
- Accounts for model & parameter errors.
- Computational cost: factor *N*.

Disadvantages

- No error propagation from DFT (DFT=truth).
- Sampling of parameter space can be tricky (anisotropic).

PhD project Sarah Wishart.

Conclusions

Force Matching

Extending atomistic simulations to new materials:

- Preserve DFT precision to larger systems, longer times.
- Foundation for other atomistic and meso-scale problems.

Essential part of multi-scale modelling stack.

Conclusions

Force Matching

Extending atomistic simulations to new materials:

- Preserve DFT precision to larger systems, longer times.
- Foundation for other atomistic and meso-scale problems.

Essential part of multi-scale modelling stack.

Effective potentials

(Semi-)empirical potentials have their limitations.

- Know their limits (don't guess them).
- Predictive simulations: Quantify Uncertainty,

Materials properties require materials potentials.