An information-theoretic approach to uncertainty quantification in atomistic modelling of crystalline materials

Maciej Buze

23 June 2022

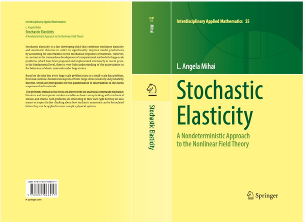
Joint work with Angela Mihai (Cardiff), Thomas Woolley (Cardiff), Geneviève Dusson (CNRS & Université Bourgogne Franche-Comté)

CECAM Workshop: Error control in first-principles modelling
Lausanne, 20-24 June 2022

Motivation

- Given an interatomic potential, is there a systematic way to form a meaningful prior for the parameters based on physical constraints?

- Plenty of recent literature on continuum stochastic elasticity, including information-theoretic framework for deriving least-biased prior distributions for model parameters*.



- Given an atomistic model, one can link it to a continuum one via the Cauchy–Born rule†.

* J. Guilleminot and C. Soize. On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties. Journal of Elasticity, 111(2):109–130, 2013.

† J.L. Ericksen. On the Cauchy–Born Rule. Mathematics & Mechanics of Solids, 13 (3–4): 199–220, 2008.

Setup

- The position of M atoms $\boldsymbol{R} = \{\boldsymbol{r}_j\}_{j=1}^{M}$.
- The total energy is expressed as a sum of body terms

$$E(\boldsymbol{R}) = \sum_{k=1}^{N} \frac{1}{k!} \sum_{i_1 \neq \ldots \neq i_k} E_k(\boldsymbol{r}_{i_1}, \ldots, \boldsymbol{r}_{i_k}),$$

  truncated at some $N \in \mathbb{N}$.
- The 1-body term, E_1, typically related to external forces, is disregarded.
- Two toy model cases:
  1. $N = 2$, E_2 is the Lennard-Jones potential;
  2. $N = 3$ and E_2, E_3 are basic prototype ML potentials following the aPIP framework*.

*C. van der Oord, G. Dusson, G. Csanyi, and C. Ortner. Regularised atomic body-ordered permutation-invariant polynomials for the construction of interatomic potentials. Mach. Learn.: Sci. Technol., 1, 2020.

Setup

- Given $\boldsymbol{R} = \{\boldsymbol{r}_j\}_{j=1}^{M}$ and using $\boldsymbol{r_{ji}} = \boldsymbol{r_j} - \boldsymbol{r_i}$, useful to rewrite as

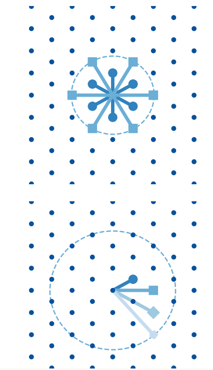
$$
\begin{aligned}
E(\boldsymbol{R}) &= \sum_{k=1}^{N} \frac{1}{k!} \sum_{i_1 \neq \ldots \neq i_k} E_k(\boldsymbol{r}_{i_1}, \ldots, \boldsymbol{r}_{i_k}) \\
&= \sum_{j=1}^{M} V(\boldsymbol{R_j}), \quad \boldsymbol{R_j} = \{\boldsymbol{r_{ji}}\}_{i \in N(j)},
\end{aligned}
$$

where $V : \mathbb{R}^{d \times (M-1)} \to \mathbb{R}$ is the site potential and there is a invariances-induced correspodence, e.g.

$$
\begin{aligned}
E_2(\boldsymbol{r}_1, \boldsymbol{r}_2) &\equiv V_2(|\boldsymbol{r}_{12}|), \\
E_3(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) &\equiv V_3(|\boldsymbol{r}_{12}|, |\boldsymbol{r}_{13}|, |\boldsymbol{r}_{23}|).
\end{aligned}
$$

Deformations and displacements

- We decompose $\boldsymbol{R} = \boldsymbol{R}^0 + \boldsymbol{R}^1$, where $\boldsymbol{R}^0$ is a fixed reference configuration, a Bravais lattice $\mathrm{A}\mathbb{Z}^d$, and $\boldsymbol{R}^1$ a displacement from it.



- The deformation and the displacement:

$$y^u : \mathrm{A}\mathbb{Z}^d \to \mathbb{R}^d, \quad y^u(\underbrace{m}_{\equiv j}) = \underbrace{m}_{\equiv \boldsymbol{r}_j^0} + \underbrace{u(m)}_{\equiv \boldsymbol{r}_j^1}.$$

- Finite differences and discrete gradients:

$$D_\rho y^u(m) := \underbrace{y^u(m+\rho) - y^u(m)}_{\equiv \boldsymbol{r}_{ji}}, \quad Dy^u(m) := \{D_\rho y^u(m)\}_{\underbrace{\rho \in \mathcal{R}}_{\equiv i \in N(j)}}.$$

- The energy:

$$\mathcal{E}(u) := \sum_m V(Dy^u(m)) \quad \equiv \quad E(\boldsymbol{R}).$$

Lattice constant and the Cauchy–Born rule

- Formal Taylor expansion around $u = 0$ $(\equiv \boldsymbol{R}_0)$:

$$\mathcal{E}(u) = \mathcal{E}(0) + \langle \delta\mathcal{E}(0), u\rangle + \langle \delta^2\mathcal{E}(0)u, u\rangle + \mathrm{h.o.t},$$

$$\begin{aligned}
\langle \delta\mathcal{E}(0), u\rangle &= \sum_m \nabla V(\{\rho\}) : Du(m) = \sum_m \sum_{i,\sigma} \partial_{i\sigma} V(\{\rho\}) D_\sigma u_i(m), \\
\langle \delta^2\mathcal{E}(0)u, u\rangle &= \sum_m \nabla^2 V(\{\rho\}) Du(m) : Du(m) \\
&= \sum_m \sum_{i,\sigma,j,\lambda} \partial^2_{i\sigma j\lambda} V(\{\rho\}) D_\sigma u_i(m) D_\lambda u_j(m).
\end{aligned}$$

- Typical assumption: perfect lattice is a stable equilibrium, that is,

$$\delta\mathcal{E}(0) = 0, \quad \langle \delta^2\mathcal{E}(0)u, u\rangle \geq C\|Du\|^2_{\ell^2}.$$

Lattice constant and the Cauchy–Born rule

- Restrict to uniform displacements $u(m) = Fm$. Then $D_\sigma u(m) = \nabla u(m)\sigma = F\sigma$ and

$$\langle \delta\mathcal{E}(0), u\rangle = \sum_m \sum_{i\alpha} L_{i\alpha} \partial_\alpha u_i,$$

$$\langle \delta^2\mathcal{E}(0)u, u\rangle = \sum_m \sum_{i\alpha j\beta} S_{i\alpha}^{j\beta} \partial_\alpha u_i \partial_\beta u_j.$$

- Perfect lattice an equlibrium:

$$\langle \delta\mathcal{E}(0), u\rangle = 0 \implies L_{i\alpha} = 0 \quad \text{(determines lattice constant } \ell\text{)}.$$

- Cauchy–Born rule and elasticity tensor:

$$W(F) := \frac{1}{\det(\ell \mathrm{A})} V(\{F\sigma\}),$$

$$\mathbb{C}_{i\alpha}^{j\beta} := \partial_{F_{i\alpha}F_{j\beta}} W(\mathrm{Id}) = \frac{1}{\det(\ell \mathrm{A})} S_{i\alpha}^{j\beta}.$$

Continuum linearised elasticity

- The strain tensor, $\varepsilon(x) = \left[\nabla u(x) + \nabla u(x)^\top\right]/2$, and the constitutive stress-strain relation is

$$\sigma(x) = \mathbb{C} : \varepsilon(x).$$

- Equilibria: $\sigma_{ij,j} = 0$ for $i = 1, 2, 3$.
- Up to to 21 independent entries in $\mathbb{C}$ with a symmetric matrix representation $[\mathbb{C}] \in \mathbb{R}^{6\times 6}$ admitting a decomposition*

$$[\mathbb{C}] = \sum_{i=1}^{n} c_i M_i.$$

- The parameters are $\boldsymbol{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$.

* J. Guilleminot and C. Soize. On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties. Journal of Elasticity, 111(2):109–130, 2013.

Continuum stochastic linearised elasticity

- Invoke the maximum entropy principle (MaxEnt) on a minimal set of constraints*:
  1. The mean value of the tensor is known.
  2. The elasticity tensor $\mathbb{C}$ and its inverse both have finite second-order moments (physical consistency).
- Write them as expectations

$$\mathbb{E}\{\boldsymbol{f}(\boldsymbol{c})\} = \boldsymbol{h},$$

  where $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^q$ and $\boldsymbol{h} \in \mathbb{R}^q$.
- Then the least informative prior on $\boldsymbol{c}$ has probability density

$$\rho(\mathfrak{c}) := 1_S(\mathfrak{c}) \exp\{-\langle \boldsymbol{\lambda}, \boldsymbol{f}(\mathfrak{c})\rangle_{\mathbb{R}^q}\},$$

  where $S \subset \mathbb{R}^n$ represents all admissable choices for $\boldsymbol{c} \in \mathbb{R}^n$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)$ are Lagrange multipliers.
- Key point: study the separability of $\phi(\boldsymbol{c}) = \log(\det[\mathbb{C}])$.

*J. Guilleminot and C. Soize. On the statistical dependence for the components of random elasticity tensors exhibiting material symmetry properties. Journal of Elasticity, 111(2):109–130, 2013.

Back to the atomistic model

- Recall that

$$\mathbb{C}_{i\alpha}^{j\beta} = \partial_{F_{i\alpha}F_{j\beta}} W(\mathrm{Id}) = \frac{1}{\det(\ell \mathrm{A})} S_{i\alpha}^{j\beta},$$
$$S_{i\alpha}^{j\beta} = \sum_{\sigma,\lambda} \partial^2_{i\sigma j\lambda} V(\{\rho\}) \sigma_\alpha \lambda_\beta.$$

.

- Invoke the maximum entropy principle (MaxEnt) on a minimal set of constraints:
  1. The mean value of the interatomic potential parameters are known.
  2. The Cauchy–Born elasticity tensor $\mathbb{C}$ and its inverse both have finite second-order moments (physical consistency).
- Key point: study the separability of $\log(\det[\mathbb{C}])$, which is now a function of the parameters of the potential.

Lennard–Jones potential example

- The site potential is

$$V(Dy^u(m)) = \sum_\rho V_2(|D_\rho y^u(m)|),$$

$$V_2(r) = 4a_1 \left[\left(\frac{1}{a_2 r}\right)^{12} - \left(\frac{1}{a_2 r}\right)^{6} \right].$$

- Hence two parameters $\boldsymbol{a} = (a_1, a_2)$.
- For pair potentials

$$L_{i\alpha} = \sum_\rho \frac{V_2'(|\rho|)}{|\rho|}\rho_i\rho_\alpha, \text{ hence here } L_{i\alpha} = 0 \implies \ell = \left(\frac{B}{A}\right)^{1/6} {a_2}^{-1},$$

$$\mathbb{C}_{i\alpha}^{j\beta} = \frac{1}{\det(\ell \mathrm{A})} \sum_\rho \left[\left(\frac{V_2''(|\rho|)}{|\rho|^2} - \frac{V_2'(|\rho|)}{|\rho|^3} \right) \rho_i\rho_j + \delta_{ij} \frac{V_2'(|\rho|)}{|\rho|} \right] \rho_\alpha\rho_\beta.$$

Lennard–Jones potential example

Proposition[*]

In an in-plane model on a triangular lattice the MaxEnt probability density function of the random variable $\boldsymbol{a} = (a_1, a_2)$ *is given by*

$$\begin{aligned} \rho_{\boldsymbol{a}}(\boldsymbol{a}) &= \rho_{a_1}(a_1) \times \rho_{a_2}(a_2), \\ \rho_{a_1}(a_1) &= \mathbb{1}_{\mathbb{R}_+}(a_1) k_1 a_1^{-\tau} \exp\{-\lambda_1 a_1\}, \\ \rho_{a_2}(a_2) &= \mathbb{1}_{\mathbb{R}_+}(a_2) k_2 a_2^{-2\tau} \exp\{-\lambda_2 a_2\}, \end{aligned}$$

with k_1 *and* k_2 *positive normalization constants, and* λ_1 *and* λ_2 *Lagrange multipliers. The parameter* τ *controls the level of statistical fluctuations and is required to satisfy* $\tau \in (-\infty, 1/2)$*.*

In other words: a_1, a_2 statistically independent and Gamma distributed.

[*] M. B., T.E. Woolley, and L.A.Mihai. A stochastic framework for atomistic fracture. SIAM Journal on Applied Mathematics, 82(2):526–548, 2022.

Translating the framework to MLIPs

- A toy ML potential with aPIP

$$E(\boldsymbol{R}) = \sum_{k=2}^{3} \sum_{\boldsymbol{i}} a_{k\boldsymbol{i}} B_{k\boldsymbol{i}}(\boldsymbol{R}),$$

where the sum is over all admissible tuples $\boldsymbol{i} = (i_j)_{j=1}^{k(k-1)/2}$ given a prescribed polynomial degree and $\{B_{k\boldsymbol{i}}\}$ is a polynomial basis.

- The parameters $\boldsymbol{a} = \{a_{k\boldsymbol{i}}\}$.
- Equivalent formulation:

$$\mathcal{E}(u) = \sum V(Dy^u(m)), \quad \text{where } V = V_2 + V_3,$$

$$V_3(Dy^0(m)) = \sum_{\sigma \neq \rho} \psi(|\rho|, |\sigma|, |\rho - \sigma|),$$

$$\psi(x, y, z) = \hat{\psi}(x + y + z, xy + xz + yz, xyz),$$

$$\hat{\psi}(x, y, z) = a_{31} + a_{32}x + a_{33}x^2 + a_{34}x^3 + a_{35}y + a_{36}xy + a_{37}z.$$

Translating the framework to MLIPs

In [13]:
```
S2 = SubstituteNumericalApprox(digits=2)(S)
S2
```

Out[13]:
```
4.5*(9.1*a24*l + 2400.*a34*l + 690.*a36*l + 53.*a37*l + 2.0*a23 + 180.*a33 + 49.*a35 + 0.14*a22/l^1.0 + 6.8*a32/l^1.
0)*l^2.0 + 4.5*(9.1*a24*l + 2300.*a34*l + 690.*a36*l + 53.*a37*l + 2.0*a23 + 180.*a33 + 49.*a35 + 0.14*a22/l^1.0 +
6.8*a32/l^1.0)*l^2.0 + 2.0*(6.0*a24*l + 1800.*a34*l + 530.*a36*l + 40.*a37*l + 2.0*a23 + 170.*a33 + 42.*a35 + 8.2*a3
2/l^1.0)*l^2.0 + 1.0*(3.8*a24*l + 1700.*a34*l + 540.*a36*l + 56.*a37*l + 2.0*a23 + 230.*a33 + 71.*a35 + 0.75*a22/l^
1.0 + 22.*a32/l^1.0)*l^2.0 + 18.*(140.*a34*l + 44.*a36*l + 5.4*a37*l + 7.0*a33 + 2.5*a35 + 0.072*a32/l^1.0)*l^2.0 +
18.*(140.*a34*l + 44.*a36*l + 5.2*a37*l + 7.0*a33 + 2.5*a35)*l^2.0 + 24.*(120.*a34*l + 39.*a36*l + 4.8*a37*l + 7.5*a
33 + 2.7*a35 + 0.040*a32/l^1.0)*l^2.0 + 4.0*(96.*a34*l + 32.*a36*l + 4.0*a37*l + 8.0*a33 + 3.0*a35)*l^2.0 + 12.*(80.
*a34*l + 25.*a36*l + 2.7*a37*l + 5.8*a33 + 1.9*a35 + 0.16*a32/l^1.0)*l^2.0 + 12.*(80.*a34*l + 25.*a36*l + 2.6*a37*l
+ 5.6*a33 + 1.8*a35)*l^2.0 + 8.0*(63.*a34*l + 20.*a36*l + 2.3*a37*l + 6.2*a33 + 2.1*a35 + 0.14*a32/l^1.0)*l^2.0 + 2.
0*(42.*a34*l + 13.*a36*l + 1.4*a37*l + 5.0*a33 + 1.5*a35 + 0.38*a32/l^1.0)*l^2.0 + 2.0*(40.*a34*l + 13.*a36*l + 1.2*
a37*l + 4.5*a33 + 1.2*a35)*l^2.0 - 12.*(25.*a34*l + 5.7*a36*l - 0.22*a37*l + 3.6*a33 + 0.77*a35 + 0.38*a32/l^1.0)*l^
2.0 + 18.*(23.*a34*l + 9.1*a36*l + 1.3*a37*l + 1.5*a33 + 0.75*a35 - (4.1e-1024610093)*a32/l^1.0)*l^2.0 - 12.*(21.*a3
4*l + 4.5*a36*l - 0.43*a37*l + 1.9*a33 - 0.067*a35)*l^2.0 - 8.0*(20.*a34*l + 6.0*a36*l + 0.50*a37*l + 4.5*a33 + 1.2*
a35 + 0.75*a32/l^1.0)*l^2.0 - 6.0*(16.*a34*l + 4.1*a36*l + (2.2e-16)*a37*l + 4.2*a33 + 1.1*a35 + 0.75*a32/l^1.0)*l^
2.0 - 18.*(16.*a34*l + 4.1*a36*l + (1.1e-16)*a37*l + 4.2*a33 + 1.1*a35 + 0.75*a32/l^1.0)*l^2.0 + 2.0*(5.6*a34*l + 2.
3*a36*l + 0.43*a37*l + 0.50*a33 + 0.25*a35 - (4.1e-1024610093)*a32/l^1.0)*l^2.0
```

In [8]:
```
SubstituteNumericalApprox(digits=2)(L_aut)
```

Out[8]:
```
(56.*a24 + 24000.*a34 + 7500.*a36 + 770.*a37)*l^3.0 + (24.*a23 + 3100.*a33 + 970.*a35)*l^2.0 + (8.2*a22 + 310.*a32)*
l
```

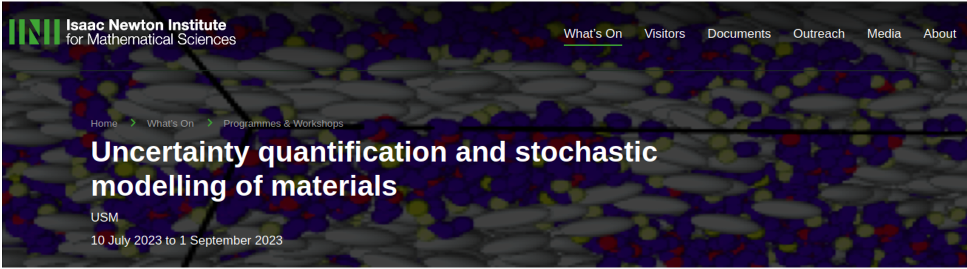
In [9]:
```
soln = solve(L_aut,l)
soln1 = soln[0].full_simplify()
soln2 = SubstituteNumericalApprox(digits=2)(soln1)
soln2
```

Out[9]:
```
l == -(3.2e-11)*(3.1e35*a23 + 3.9e37*a33 + 1.2e37*a35 + sqrt(9.3e70*a23^2.0 - 3.0e71*a22*a24 - 1.1e73*a24*a32 + 2.4e
73*a23*a33 + 1.5e75*a33^2.0 - 2.7e60*(4.6e13*a22 + 1.7e15*a32)*a34 + 5.6e66*(1.4e6*a23 + 1.7e8*a33)*a35 + 1.5e74*a35
^2.0 - 8.7e59*(4.6e13*a22 + 1.7e15*a32)*a36 - 8.8e58*(4.6e13*a22 + 1.7e15*a32)*a37))/(4.5e25*a24 + 1.9e28*a34 + 6.1e
27*a36 + 6.3e26*a37)^1.0
```

Conclusions

- Perhaps a way forward is to build a framework with MaxEnt in mind? Specific formulation of constraints, specific change of variables, specific basis functions, specific forms of nonlinearities to guarantee statistiscal indepedence?
- Perhaps not worth it - need to check how the choice of the prior influences the posterior.

- Maximum Entropy Principle in principle provides a systematic way of recasting physical constraints in a probabilistic language.
- Easier to use if the constraints can be cast in a form of expectation.
- In practice, beyond simplest of cases it can easily get intractably messy.



10 July 2023 to 1 September 2023

https://www.newton.ac.uk/event/usm/