Conditioning schemes for accurately estimating thermodynamic and transport properties of metallic alloys: application to thermo-elasticity and elasto-diffusion

Manuel Athènes
Anuo Zhong, Clovis Lapointe, Alexandra M. Goryaeva, Mihai-Cosmin Marinica,
Thomas Jourdan, Gilles Adjanor (EDF), Jérôme Creuze (ICMMO) & Vasily V. Bulatov (LLNL)

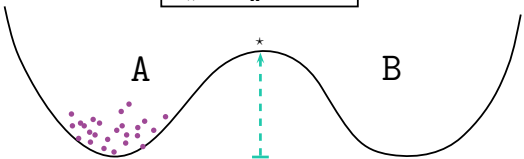
Section de Recherches de Métallurgie Physique
Université Paris Saclay
CEA, France

IPAM Workshop : Increasing the Length, Time, and Accuracy of Materials Modeling Using Exascale Computing
28 March 2023, UCLA,

Metastability, rare events and kinetic trapping

- Metastability

$$\mathcal{F}_\star - \mathcal{F}_\mathrm{A} >> kT$$

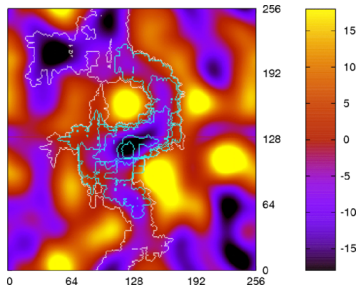


- Free energy difference $\mathcal{F}_\mathrm{B} - \mathcal{F}_\mathrm{A}$ controls the conditions of phase equilibria

- Reaction rates and Transition State Theory

$$k_{\mathrm{A}\to\mathrm{B}} = \nu_0 \exp\left[-\frac{\mathcal{F}_\star - \mathcal{F}_\mathrm{A}}{kT}\right]$$

- Kinetic trapping

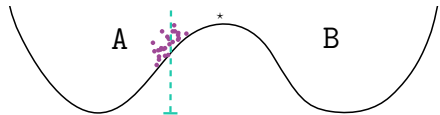


Umbrella sampling and stratification

Extended potential energy

$$\mathcal{U}(\lambda, r) = \beta_{\mathrm{ref}} V(r) + \tfrac{1}{2}\kappa\|\xi(r) - \lambda\|^2 + \epsilon$$

$$\textstyle\sum_{\lambda\in\Lambda} \exp\left[-\frac{1}{2}\kappa\|\xi(r) - \lambda\|^2 - \epsilon\right] = 1$$



- Marginal probability : $\mathrm{p}^{\Lambda}(\lambda) = \exp[-\mathcal{A}(\lambda)]$
- Conditional probability of r given λ : $\pi(r|\lambda) = C_{\Lambda} \exp[\mathcal{A}(\lambda) - \mathcal{U}(\lambda, r)]$
- Law of total probabilities w.r.t $\lambda \in \Lambda$:

$$\mathrm{p}^{\mathcal{R}}(r) = \sum_{\zeta\in\Lambda} \pi(r|\lambda)\mathrm{p}^{\Lambda}(\lambda) = \sum_{\lambda\in\Lambda} \pi(r|\lambda) \exp[-\mathcal{A}(\lambda)]$$

- Marginal probability of r coincides with reference : $\mathrm{p}^{\mathcal{R}}(r) \propto \exp[-\beta_{\mathrm{ref}} V(r)]$
- Law of total expectation (LTE) in joint space $(\lambda, r) \in \Lambda \times \mathcal{R}$ w.r.t $\lambda \in \Lambda$

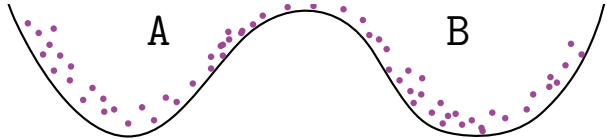
$$\Longrightarrow \qquad \mathbb{E}[\mathbf{1}_{\xi^\star}(\xi_r)] = \textstyle\sum_{\lambda\in\Lambda} \mathbb{E}[\mathbf{1}_{\xi^\star}(\xi_r)|\lambda] \exp[-\mathcal{A}(\lambda)]$$

Expanded ensemble simulations and Bayes formula

Joint probability

$$\mathrm{p}_a(\lambda, r) = \exp\left[a(\lambda) - \mathcal{U}(\lambda, r)\right]$$

$$\mathbb{E}\left[\mathcal{O}\,|\lambda\right] = \frac{\mathbb{E}_a\left[\mathbf{1}_\lambda\, \mathcal{O}(\lambda, \cdot)\right]}{\mathbb{E}_a\left[\mathbf{1}_\lambda\right]}$$



- Conditioning of the total expectations on r

$$\mathbb{E}\left[\mathcal{O}\,|\lambda\right] = \frac{\mathbb{E}_a\left[\mathbb{E}_a\left[\mathbf{1}_\lambda\, \mathcal{O}(\lambda, r)|r\right]\right]}{\mathbb{E}_a\left[\mathbb{E}_a\left[\mathbf{1}_\lambda|r\right]\right]} = \frac{\mathbb{E}_a\left[\pi_a(\lambda|\cdot)\, \mathcal{O}(\lambda, \cdot)\right]}{\mathbb{E}_a\left[\pi_a(\lambda|\cdot\right]}$$

$\Leftrightarrow$ Insertion of conditional probability of λ given r :

$$\pi_a(\lambda|r) = \mathbb{E}_a\left[\mathbf{1}_\lambda|r\right] = \frac{\exp\left[a(\lambda) - \mathcal{U}(\lambda, r)\right]}{\sum_{\zeta\in\Lambda} \exp\left[a(\zeta) - \mathcal{U}(\zeta, r)\right]}$$

- Conditioned estimator : $\Upsilon^M_\Pi(\mathcal{O}\,|\lambda) = \frac{\frac{1}{M}\sum_{m=1}^M \pi_a(\lambda|r_m)\, \mathcal{O}(\lambda, r_m)}{\frac{1}{M}\sum_{m=1}^M \pi_a(\lambda|r_m)}$

Estimation of the expected value of the external parameter λ without and with conditioning on the internal coordinate r

Standard estimator : $\overline{\lambda}^M = \frac{1}{M}\sum_{m=1}^{M} \lambda_m$

$$\mathbb{V}\left[\overline{\lambda}^M\right] = \frac{1}{M}\,\mathbb{V}\left[\lambda\right]$$

Conditioned estimator : $\overline{\lambda}_{\Pi}^M = \frac{1}{M}\sum_{m=1}^{M} \mathbb{E}\left[\lambda|r_m\right]$

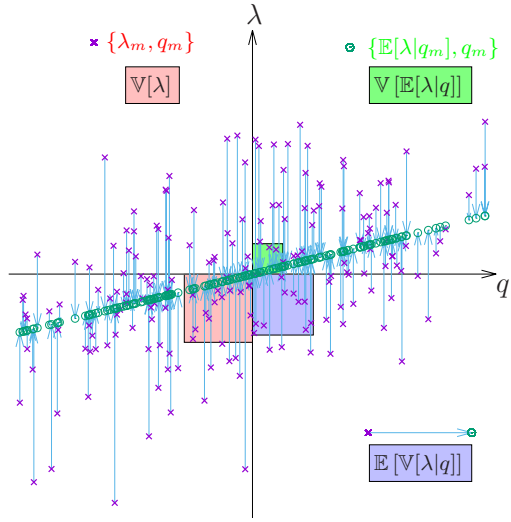
$$\mathbb{V}\left[\overline{\lambda}_{\Pi}^M\right] = \frac{1}{M}\,\mathbb{V}\left[\mathbb{E}\left[\lambda|r\right]\right]$$

Law of total expectation

$$\mathbb{E}\left[\lambda\right] = \mathbb{E}\left[\mathbb{E}\left[\lambda|r\right]\right]$$

Law of total variance

$$\mathbb{V}\left[\lambda\right] = \mathbb{V}\left[\mathbb{E}\left[\lambda|r\right]\right] + \mathbb{E}\left[\mathbb{V}\left[\lambda|r\right]\right]$$

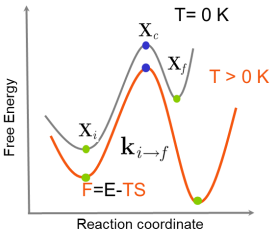


Bayesian Adaptive Biasing Force (BABF) Method

Adaptive Biasing Force (ABF): anharmonic free energy calculation

General potential energy: $U(\zeta, \mathbf{r}) = \zeta U(\mathbf{r}) + (1-\zeta)U_{ref}(\mathbf{r})$

Biasing potential: $U_{A_\star}(\zeta, \mathbf{r}) = U(\zeta, \mathbf{r}) - A_\star(\zeta)$



ζ: coupling parameter

$\mathbf{r}$: configuration

$U(\mathbf{r})$: potential energy of target system

$U_{ref}(\mathbf{r})$: potential energy of reference system

$A_\star(\zeta)$: bias, discretized form: $A_n(\zeta)$

Proven convergence:

$$\lim_{n\to+\infty} \Delta A_n = \Delta A = A(1) - A(0)$$

Bayesian scheme[1] Bayes relation: $p_{A_\star}(\mathbf{r}|\zeta) = \frac{p_{A_\star}(\zeta|\mathbf{r})P_{A_\star}(\mathbf{r})}{P_{A_\star}(\zeta)}$

Mean force: $A'(\zeta) = \int_{T^{3N_a}} \partial_\zeta U(\zeta,\mathbf{r}) p_{A_\star}(\mathbf{r}|\zeta) d\mathbf{r} = \frac{\int_{T^{3N_a}} \partial_\zeta U(\zeta,\mathbf{r}) p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}{\int_{T^{3N_a}} p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}$

Ref: [1] L. Cao, M. Athènes *et al.*, J. Chem. Phys. 140 (2014).

STEP 1

$$A'_n(\zeta) = \frac{\sum_{s=1}^{n-1} \partial_\zeta U(\zeta, \mathbf{r}_s) p_{A_s}(\zeta|\mathbf{r}_s)}{\sum_{s=1}^{n-1} p_{A_s}(\zeta|\mathbf{r}_s)}$$

STEP 2

$$A_n(\zeta) = \int_0^\zeta A'_n(\tilde{\zeta}) d\tilde{\zeta} + A_n(0)$$

STEP 3

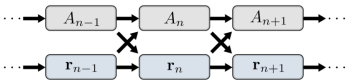
$$p_{A_n}(\zeta|\mathbf{r}_n) = \frac{\exp\left[-\beta U_{A_n}(\zeta, \mathbf{r}_n)\right]}{\int_0^1 \exp\left[-\beta U_{A_n}(\tilde{\zeta}, \mathbf{r}_n)\right] d\tilde{\zeta}}$$

STEP 4

$$\mathbf{F}_{A_n}(\mathbf{r}_n) = -\int_0^1 \nabla_\mathbf{r} U(\zeta, \mathbf{r}_n) p_{A_n}(\zeta|\mathbf{r}_n) d\zeta$$

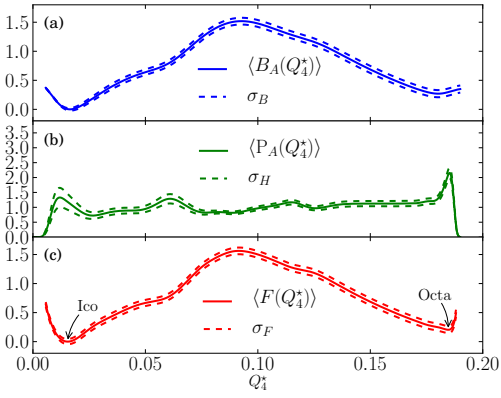
STEP 5 Langevin dynamics

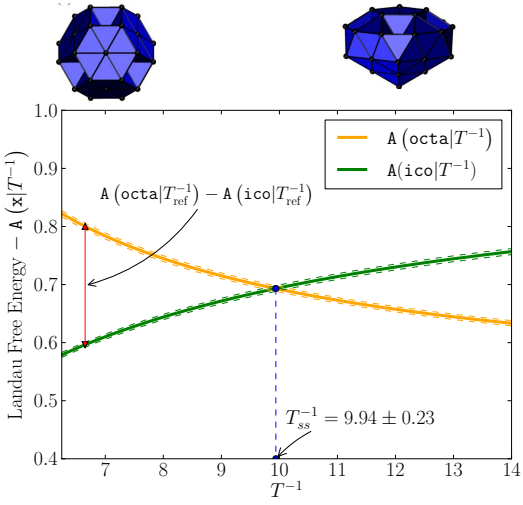
$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{P}\mathbf{F}_{A_n}(\mathbf{r}_n)\delta t + \sqrt{2\beta^{-1}\delta t} B_n$$



Structural Transition in LJ_{38}

- Adaptive biasing force (ABF) simulations in expanded ensemble : $a'(\lambda) \to \mathcal{A}'(\lambda) \implies a(\lambda) \to \mathcal{A}(\lambda)$

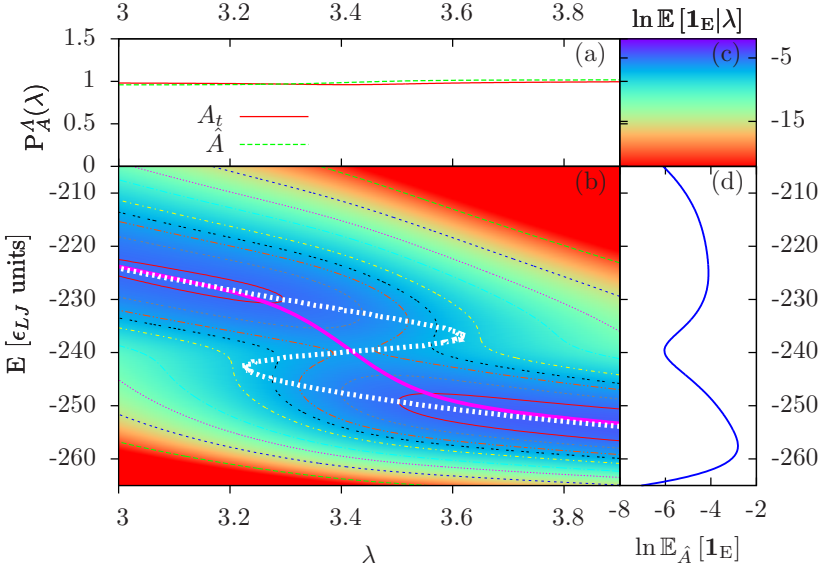




- $A(\lambda) = -\ln \mathbb{E}\left[\mathbf{1}_\lambda\right]$
- $F(Q_4^\star) = -\ln \mathbb{E}\left[\mathbf{1}_{Q_4^\star}\right]$
- $B(Q_4^\star) = \ln \sum_{\lambda \in A} \exp\left[A(\lambda) - |\lambda - Q_4^\star|^2\right]$

MA, P. Terrier, JCP 146 194101 (2015)

Melting temperature of LJ_{55} cluster



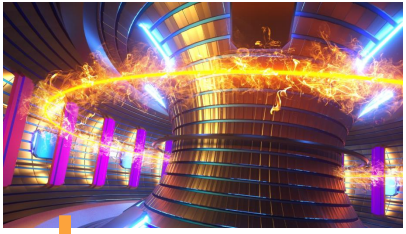
Prediction of High-temperature Elasticity of Tungsten Using Machine Learning and Data-driven Approach

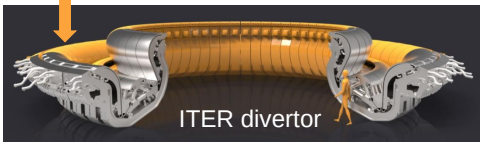
- **Elastic constants are second derivatives of the free energy**
  - Implementation of the Bayesian Adaptive Biasing Force method
  - Development of machine learning interatomic potentials for Tungsten

MiLaDy: Machine Learning Dynamics
https://ai-atoms.github.io/milady/
ASL license, open source software

- **Acknowledges: EUROfusion, GENCI - (CINES/CCRT) computer centre**

Scientific Problems





Nuclear fusion: energy for future generations

Challenge : plasma-facing material of divertor

Very high heat flux: up to 20 MW/m^2

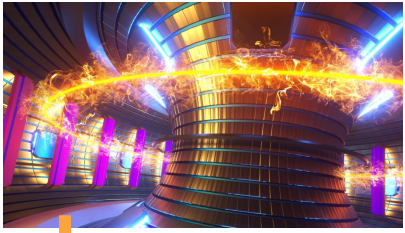
Heat flux
10×

Tungsten: highest melting point of all the metals (3695 K)

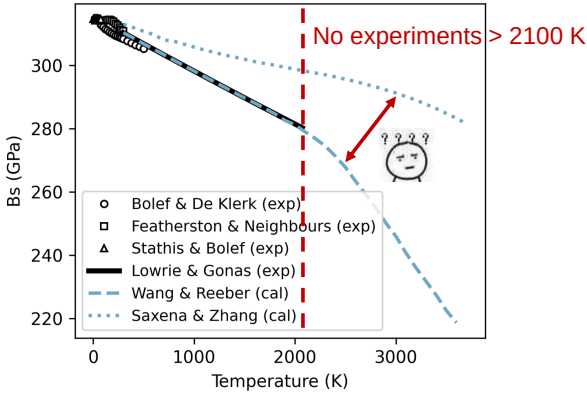
https://www.iter.org/mach/Divertor

Scientific Problems

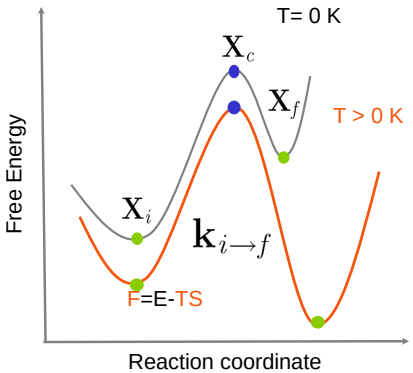
Tungsten: highest melting point of all the metals (3695 K)



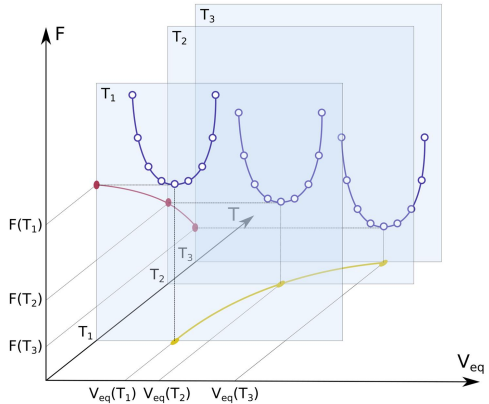
Lack of reliable high-temperature elastic properties for BCC W



Effects of Temperature on Free Energy Profile



Finite-Temperature Properties



Thermodynamic Properties of BCC Tungsten by EAM potentials

EAM potentials: numerically fast but inaccurate

BCC W system of 128 atoms

B^T: isothermal bulk modulus

$C_{11}{}^T$, C^T, $C_{44}{}^T$: isothermal elastic constants

5 most widely used EAM potentials:

WDD: P. M. Derlet *et al.*, Phys. Rev. B 76, 054107 (2007).

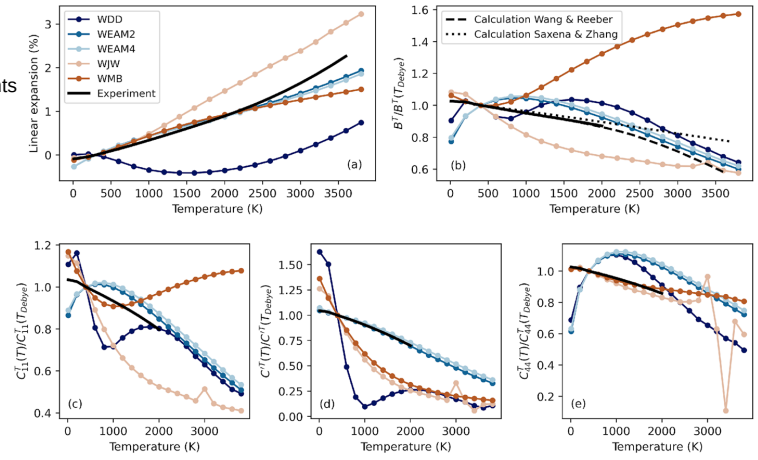
WEAM2 & WEAM4: M.-C. Marinica *et al.*, J. Phys.: Cond. Matter. 25, 395502 (2013).

WJW: N. Juslin *et al.*, J. Nucl. Mater. 432, 61 (2013).

WMB: D. R. Mason *et al.*, J. Phys.: Condens. Matter. 29, 505501 (2017).

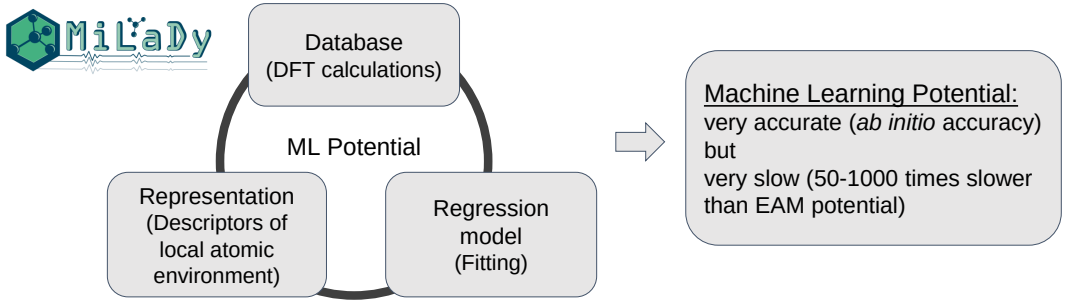
ML potentials

Problem: too slow!
(50-1000 times slower)



Solution: Machine Learning Interatomic Potential?

Construction of Machine Learning (ML) Interatomic Potential



What we want

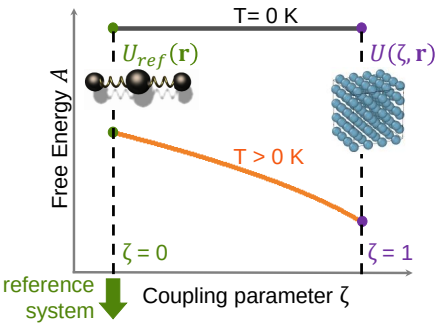
1. Faster simulation method for using computational expensive ML potentials

2. Better ML potential for high-temperature properties

Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Adaptive Biasing Force (ABF): anharmonic free energy calculation

$U(\zeta, \mathbf{r}) = (1-\zeta)U_{ref}(\mathbf{r}) + \zeta U(\mathbf{r})$



Einstein approximation (EA)
Harmonic approximation (HA)

Biasing potential: $U_{A_\star}(\zeta, \mathbf{r}) = U(\zeta, \mathbf{r}) - A_\star(\zeta)$

Estimaton of free energy A with $\lim_{n\to+\infty} \Delta A_n = \Delta A = A(1) - A(0)$

Mean force: $A'(\zeta) = \int_{T^{3N_a}} \partial_\zeta U(\zeta, \mathbf{r}) p_{A_\star}(\mathbf{r}|\zeta) d\mathbf{r}$ — Integration → A

$$= \frac{\int_{T^{3N_a}} \partial_\zeta U(\zeta, \mathbf{r}) p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}{\int_{T^{3N_a}} p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}$$

Bayes relation:

$$p_{A_\star}(\mathbf{r}|\zeta) = \frac{p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r})}{P_{A_\star}(\zeta)}$$

Sampling $P_{A_\star}(\mathbf{r})$

Statistical variance ⬇

L. Cao, M. Athènes et al., J. Chem. Phys. 140 (2014).

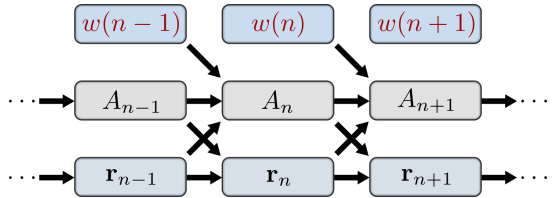
Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Bayesian Adaptive Biasing Force (BABF): anharmonic free energy calculation

Optimization of mean force computation

Introducing a **weighting function w(n)** in mean force calculation at each step n

Mean force $A'_n(\zeta) = \frac{\sum_{s=1}^{n-1} \partial_\zeta U(\zeta,\mathbf{r}_s) p_{A_s}(\zeta|\mathbf{r}_s) w(s)}{\sum_{s=1}^{n-1} p_{A_s}(\zeta|\mathbf{r}_s) w(s)}$



Optimization of reference system

Removal of numerical instabilities from the HA reference system via an SVD filter. ➡ More stable

$\mathbf{D} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$ dynamical matrix $\xrightarrow[\mathbf{U}_r = \left(U_1 \dots U_{3N_a-3}\right) \in \mathbb{R}^{3N_a \times 3N_a-3}]{\text{remove 3 translation eigenvalues}}$ $\mathbf{D}^{\mathrm{SVD}} = \mathbf{U}_r\mathbf{U}_r^\top\mathbf{D}\mathbf{U}_r\mathbf{U}_r^\top$

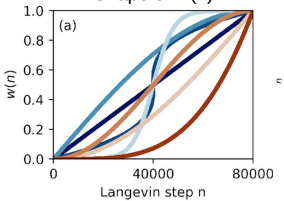
Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Optimization of weighting function w(n)

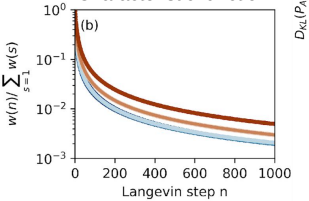
Mean force $A'_n(\zeta) = \frac{\sum_{s=1}^{n-1} \partial_\zeta U(\zeta, \mathbf{r}_s) p_{A_s}(\zeta|\mathbf{r}_s) w(s)}{\sum_{s=1}^{n-1} p_{A_s}(\zeta|\mathbf{r}_s) w(s)}$

Criterion for choosing w(n):
$w(n)/\sum_{s=1}^{n} w(s)$ ↑, convergence speed ↑!

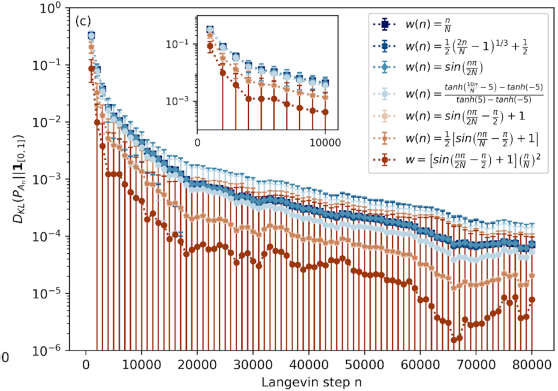
Shape of w(n)



Characteristic function



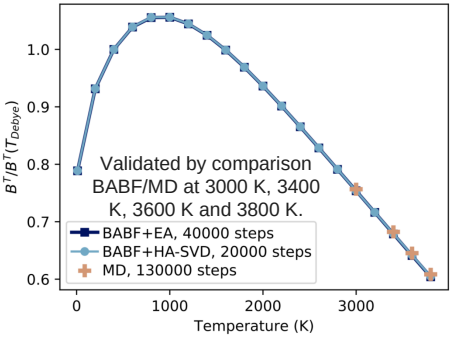
Kullback-Leibler divergence
(smaller ⇔ better sampling)



Accelerated Bayesian Adaptive Biasing Force (BABF) Method

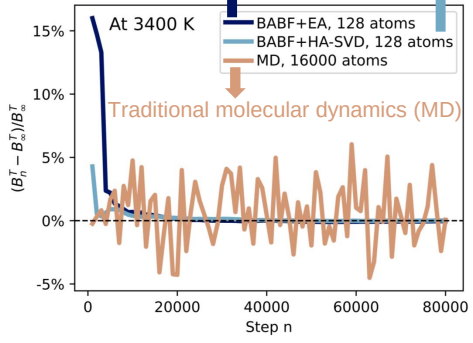
Validation: Comparison of BABF method and stress fluctuation method in molecular dynamics (MD) on bulk modulus calculation

Simulation system: bcc W with WEAM4[1] potential.



EA: Einstein reference

HA-SVD: SVD-filtered harmonic reference (the present method)



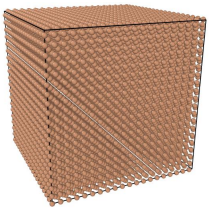
[1] M.-C. Marinica *et al.*, J. Phys.: Cond. Matter. 25, 395502 (2013).

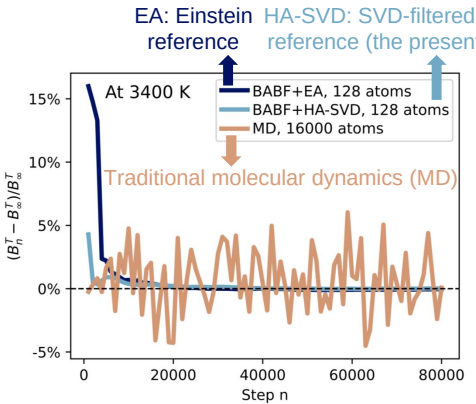
Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Validation: Comparison of BABF method and stress fluctuation method in molecular dynamics (MD) on bulk modulus calculation

The present method:
128 atoms
20000 steps

Traditional MD:
16000 atoms
130000 steps

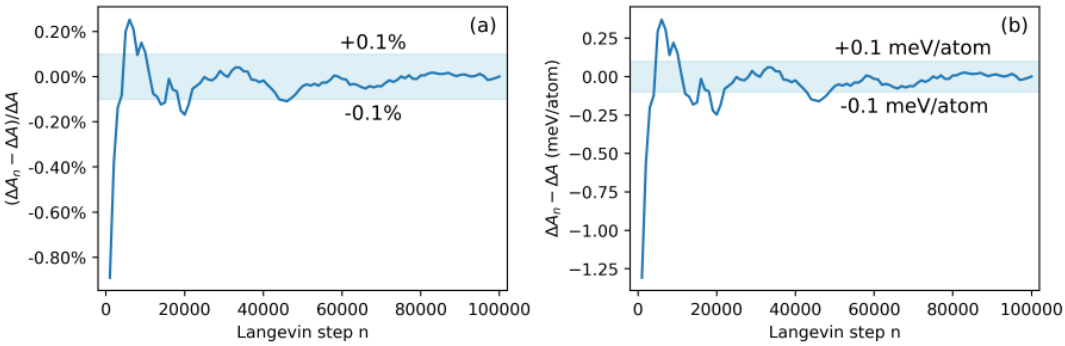




A **hundredfold** increase in computational efficiency compared with MD + Free energy available

Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Runtime errors of free energy difference Δ*A* for BABF method

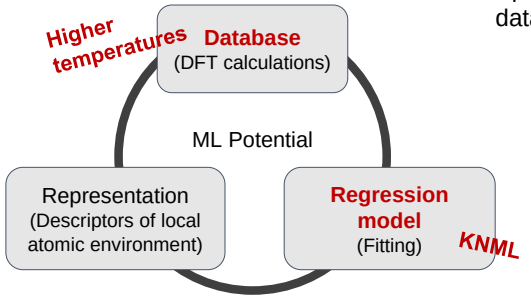


To achieve the accuracy of **0.1 meV/atom**:

- Standard thermodynamic integration: 10^6 steps in Mg system of 490 atoms at 989 K.
- BABF method: 5×10^4 steps in W system of 128 atoms at 3400 K.

Construction of Kernel Noise Machine Learning Potential

Construction of Machine Learning (ML) Interatomic Potential

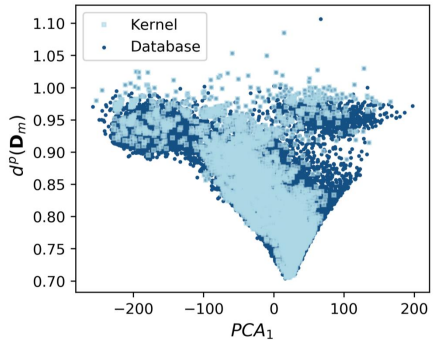


MiLaDy

- Fourth order polynomial kernel
- 6 times faster than GAP[1]

Kernel Noise Machine Learning Potential (KNML)

Sparse points selection: select the most representative data points to build kernel

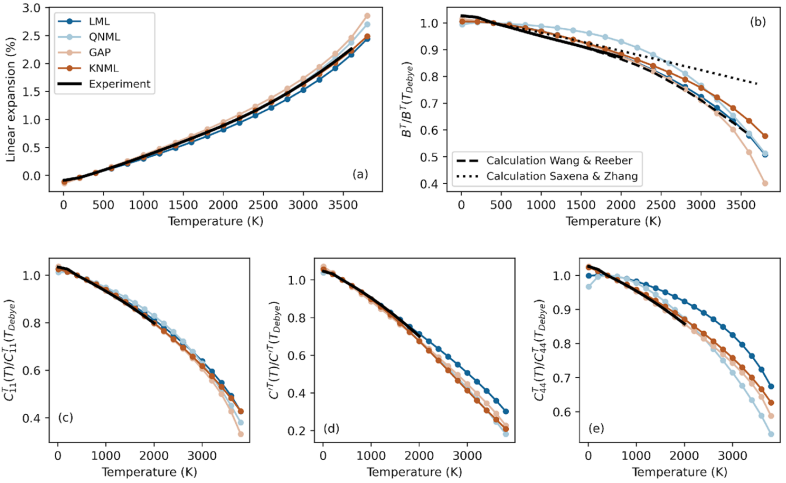


Mahalanobis distance $d(\mathbf{D}_m)$: a statistical distance

[1] W. J. Szlachta *et al.*, Phys. Rev. B 90, 104108 (2014).

Application to Thermodynamic Properties of BCC Tungsten

Performance of machine learning (ML) potentials in a BCC W system of 128 atoms



ML potentials: very slow but very accurate (*ab intio* accuracy)

Existing ML potentials:
LML & QNML: A. M. Goryaeva *et al.*, Phys. Rev. Mater. 5, 103803 (2021).
GAP: W. J. Szlachta *et al.*, Phys. Rev. B 90, 104108 (2014).

KNML (Kernel Noise Machine Learning Potential): constructed with Milady package.
(https://ai-atoms.github.io/milady/)

⬇

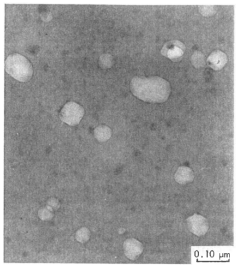
Polynomial model for B^T, $C_{11}{}^T$, C'^T and $C_{44}{}^T$ of W derived from KNML

A. Zhong *et al.*, Phys. Rev. Mater. 7.2 (2023): 023802.

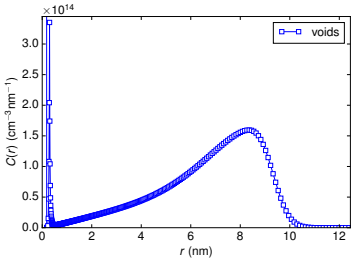
Importance of transport coefficients in materials modeling : example of cavity formation in irradiated aluminium

✓ Vacancy agglomeration is observed experimentally under ion irradiation

(a) [Farrell et al., Scripta Metall., 1977]



(b) Simulated cavity distribution

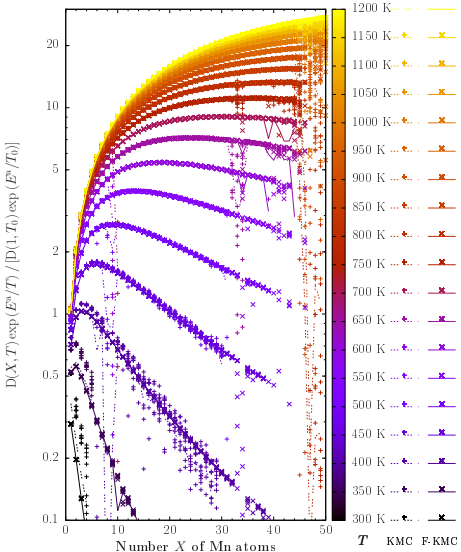


✓ involved chemical reactions : $V_n + V_1 \underset{}{\overset{F_n(D_n+D_1)}{\rightleftharpoons}} V_{n+1}$ and $V_n \overset{K_n}{\rightleftharpoons} V_{n-1} + V_1$

$$\frac{d[V_n]}{dt} = F_{n-1}(D_{n-1} + D_1)[V_{n-1}][V_1] - F_n(D_{n-1} + D_1)[V_n][V_1] + K_{n+1}[V_{n+1}] - K_n[V_n]$$

✓ how to efficiently compute rate constants K_n, absorption efficiencies F_n and diffusion coefficients D_n ?

Diffusion of manganese in α iron via a vacancy mechanism



Conventional versus advanced kinetic Monte Carlo simulations

MA, S. Kaur, G. Adjanor, T. Vanacker, T. Jourdan, Phys. Rev. Mat. 3, 103802 (2019)

Computations of linear mass transport coefficients using kinetic Monte Carlo methods

In KMC simulations, diffusion coefficients are estimated from the mean (expectation) of the atomic square displacements $(\mathbf{r}_{0\to\ell}) \otimes (\mathbf{r}_{0\to\ell})$ at equilibrium, for large numbers of KMC steps ℓ.

$$\mathbb{E}\left[(\mathbf{r}_{0\to\ell}) \otimes (\mathbf{r}_{0\to\ell})\right] = \mathbb{V}\left[\mathbf{r}_{0\to\ell}\right]$$

Mean displacements $\mathbb{E}\left[\mathbf{r}_{0\to\ell}\right]$ being zero at equilibrium, the MSD's are variances

The diffusion matrix is half the asymptotic variance w.r.t to the elapsed physical time

$$\mathbb{D}_\infty\left[\mathbf{r}_{0\to 1}\right] = \lim_{\ell\to\infty} \mathbb{D}_\ell\left[\mathbf{r}_{0\to 1}\right] \text{ with } \mathbb{D}_\ell\left[\mathbf{r}_{0\to 1}\right] = \frac{\mathbb{V}\left[\mathbf{r}_{0\to\ell}\right]}{2\mathbb{E}\left[t_{0\to\ell}\right]}$$

If $\bar{\tau}$ is the mean waiting time for a KMC jump, the diffusion matrix $\mathbb{D}_\infty\left[\mathbf{r}_{0\to 1}\right]$ is estimated from a set of K trajectories

$$\mathbf{D}_\ell = \frac{1}{\ell K}\frac{1}{2\bar{\tau}}\sum_{h=1}^{K}(\mathbf{r}_{h,0\to\ell})^{2\otimes}$$

Displacement conditioning and correlation splitting

- The variance satisfying $\mathbb{V}\left[\mathbf{r}_{0\to 1}\right] = 2\bar{\tau}\mathbb{D}_1\left[\mathbf{r}_{0\to 1}\right]$, the law of total variance for the uncorrelated contribution to diffusion is

$$\mathbb{D}_1\left[\mathbf{r}_{0\to 1}\right] = \mathbb{E}\left[\mathbb{D}_1\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right] + \mathbb{D}_1\left[\mathbb{E}\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right]$$

- Consecutive atomic correlations define the intra-correlated contribution to diffusion :

$$\mathbb{D}_2\left[\mathbf{r}_{0\to 1}\right] = \mathbb{E}\left[\mathbb{D}_1\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right]$$

- The sum of the remaining correlations define the extra-correlated contribution to diffusion. Remarkably, it has a very simple form $-\mathbb{D}_\infty\left[\mathbb{E}\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right]$.

- This yields a law of total diffusion, i.e. a relation for the asymtotic variance of reversible stochastic processes

$$\mathbb{D}_\infty\left[\mathbf{r}_{0\to 1}\right] = \mathbb{E}\left[\mathbb{D}_1\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right] - \mathbb{D}_\infty\left[\mathbb{E}\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right]$$

Green-Kubo formula and Poisson equation

Any linear transport coefficient can be expressed as a Green-Kubo formula. For mass transport :

$$\mathbb{D}_\infty \left[\mathbf{r}_{0\to 1}\right] = \mathbb{D}_1 \left[\mathbf{r}_{0\to 1}\right] + \frac{1}{\bar{\tau}} \mathbb{E}\left[\mathbf{r}_{0\to 1} \otimes \boldsymbol{\epsilon}(\chi_1)\right]$$

The relaxation vector $\boldsymbol{\epsilon}(\chi_1) = \mathbb{E}\left[\mathbf{r}_{1\to\infty} | \chi_1\right]$ satisfies a Poisson equation :

$$\boldsymbol{\epsilon}(\chi_1) = \mathbb{E}\left[\boldsymbol{\epsilon}(\chi_2) | \chi_1)\right] + \mathbf{e}(\chi_1)$$

where the mean displacement vector is defined by

$$\mathbf{e}(\chi_1) = \mathbb{E}\left[\mathbf{r}_{1\to 2} | \chi_1\right] = -\mathbb{E}\left[\mathbf{r}_{0\to 1} | \chi_1\right].$$

Estimating the mass transport coefficients from the law of total diffusion

Law of total diffusion :

$$\mathbb{D}_\infty\left[\mathbf{r}_{0\to 1}\right] = \mathbb{E}\left[\mathbb{D}_1\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right] - \mathbb{D}_\infty\left[\mathbb{E}\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right]$$

where formally :

Intra-correlated part is estimated from the mean of the sampled conditional variances :

$$\mathbb{E}\left[\mathbb{D}_1\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right] \simeq \frac{1}{2\tau\ell}\sum_{k=1}^{K}\left(\sum_{l=1}^{\ell-1}\mathbf{V}(\chi_l)\right)$$

Extra-correlated part is estimated via its conditioned expression and with $\ell \to \infty$

$$\mathbb{D}_{\ell-1}\left[\mathbb{E}\left[\mathbf{r}_{0\to 1}|\chi_0\right]\right] \simeq \frac{1}{2\tau(\ell-1)}\sum_{k=1}^{K}\left(\sum_{l=2}^{\ell}\mathbf{e}(\chi_l)\right)^{2\otimes}$$

MA, G. Adjanor, J. Creuze, Phys. Rev. Mat. 6, 013805, (2022)

Kinetic Monte Carlo algorithm

- The evolution of the system is governed by its transition rate matrix $\mathbf{K}$ and a master equation that is continuous in time and discrete in space :

$$\mathbf{P}(t, t+\tau) = \exp(\tau \mathbf{K}) \quad \forall \beta \; K_{\beta\beta} = -\sum_{\gamma \neq \beta} K_{\beta\gamma} \qquad \Longrightarrow \; \forall \beta, t, \tau \; \sum_{\gamma} P_{\beta\gamma} = 1$$

$\Longrightarrow$ the evolution operator is a stochastic matrix

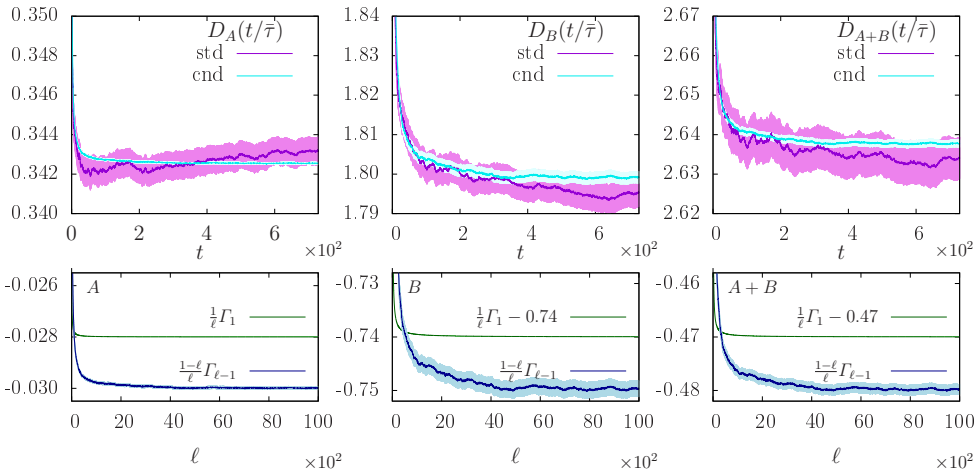
- transition rates are given by TST : $K_{\beta\gamma} = \nu \exp\left[-\frac{E_{\beta\gamma}^{\mathrm{saddle}} - E_{\beta}}{kT}\right]$

- Kinetic Monte Carlo (kMC) algorithm :

1. moves are drawn using transition matrix $\tilde{\mathbf{P}} = \mathbf{I} + \mathbf{K}\tau$ with $\tau = -\min_{\beta} K_{\beta\beta}^{-1}$
2. The elapsed time is drawn from the exponential distribution $\Delta t \sim \exp(-\Delta t/\tau)$

Atomic diffusivities measured by KMC simulations

Random solid solution with jump frequencies : $\nu_A = 1$ and $\nu_B = 5$
Composition $C_B = 0.59$ (at.)

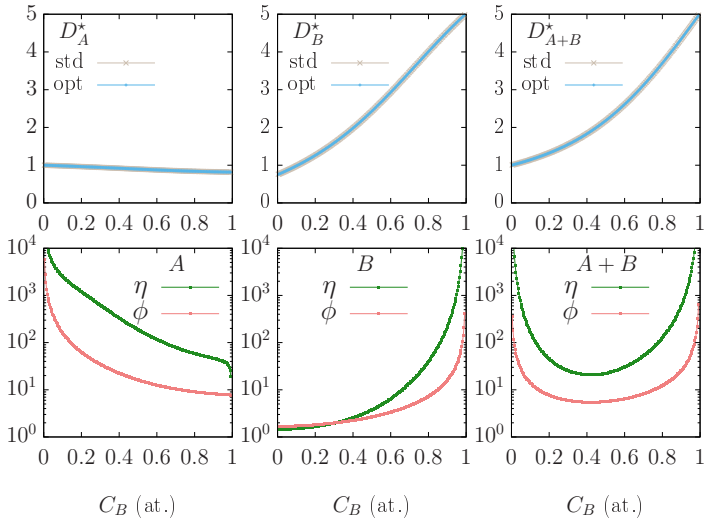


Atomic diffusivities measured in kMC : $\nu_A = 1$, $\nu_B = 5$

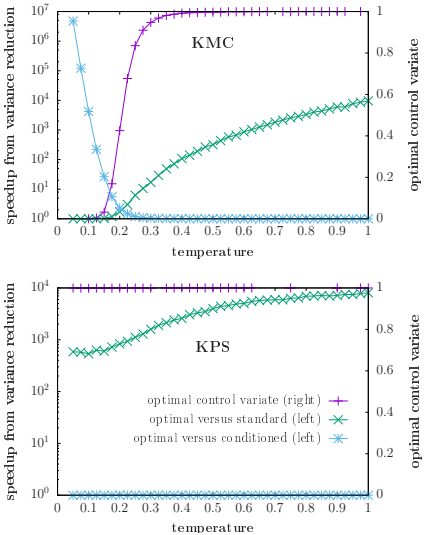
Diffusion coefficients

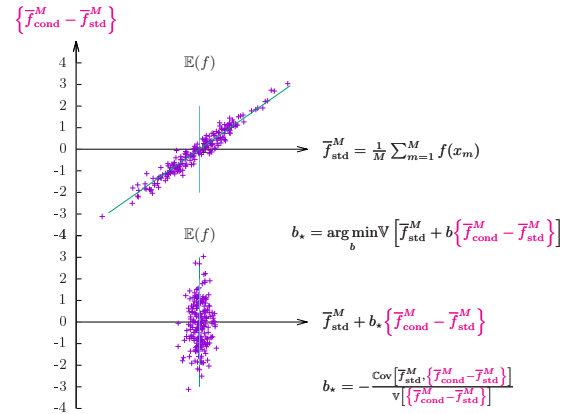
η : speed-up factor

ϕ : intra-to-extra correlated ratio



Optimal combination of standard and conditioned estimators through a control variate



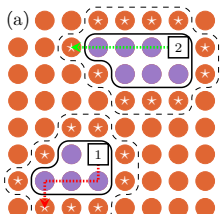


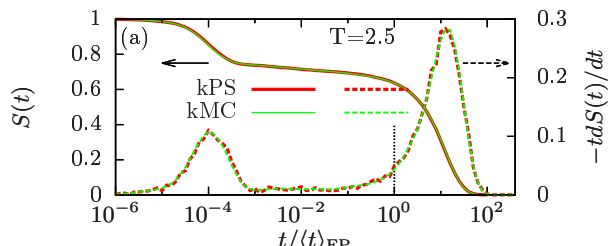
Kinetic trapping of simulated KMC trajectories in small basins of sizes $N < 10^4$

- $\mathcal{T}_\beta$ Mean first passage time from state β satisfies a discrete Poisson equation :

$$\mathcal{T}_\beta = \mathbb{E}\left[\mathbb{E}\left[\tau + \mathcal{T}_\gamma | \gamma\right]\right] = \tau + \sum_{\gamma \in \mathrm{trap}} \left(I_{\beta\gamma} + \tau K_{\beta\gamma}\right) \mathcal{T}_\gamma$$

- linear system : $\mathbf{Ax} = \mathbf{b}$ with $A_{\beta\gamma} = -K_{\beta\gamma}$, $x_\beta = \mathcal{T}_\beta$ and $b_\gamma = 1$
- first passage time are drawn through randomization based on $\mathbf{LUx} = \mathbf{b}$

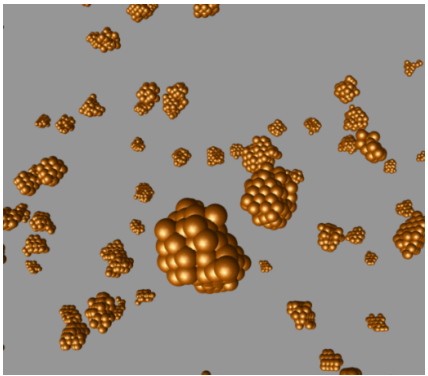


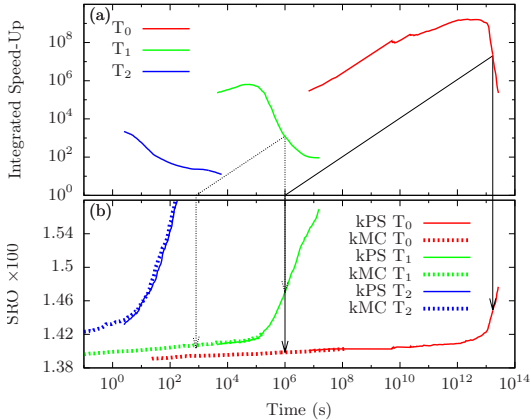


- In kinetic path sampling simulations, exit probabilities are formally the marginal probabilities to reach the corresponding absorbing states, MA, V. Bulatov, PRL 113, 230601 (2014)

Copper precipitation in α-iron

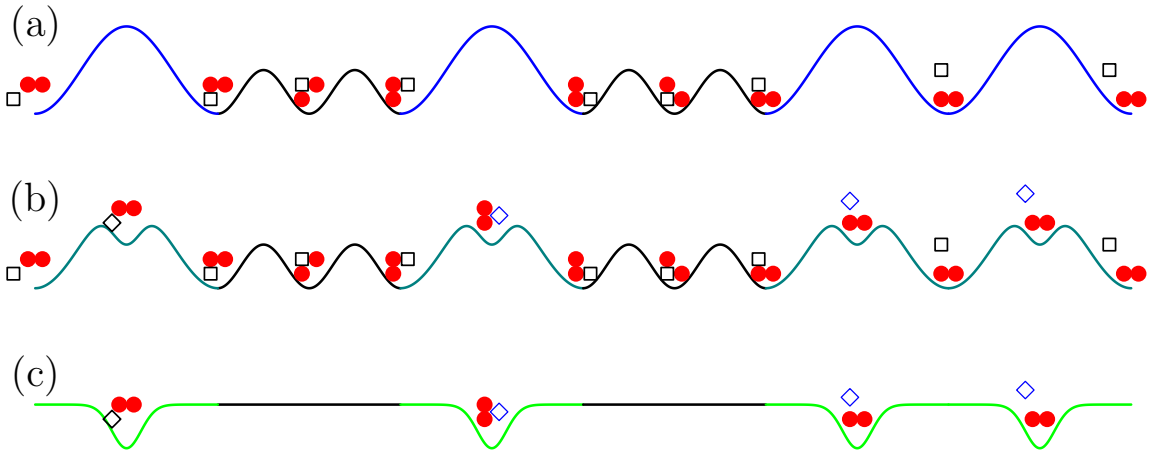
- Parametrization based on electronic structure calculations *F. Soisson and C. Fu, Phys. Rev. B, 76, 214102 (2007).* System with a single vacancy and 1.34 at% Cu
- KPS simulations of Cu precipitation in α-iron, *MA and Bulatov, PRL, 113, 230601, (2014)*



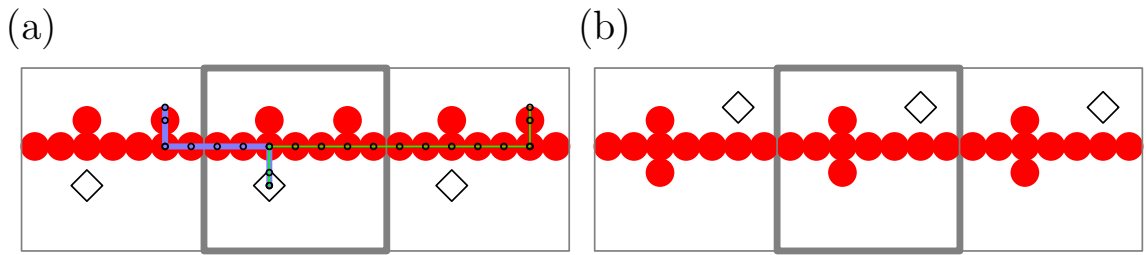


- vacancies tend to get trapped in copper precipitates
- acceleration is more important at low temperatures where trapping is more severe

Trapping ($\nu_B >> \nu_A$), kinetic path sampling and reversibility



Kinetic path sampling with traps percolating through supercell



- KPS handles periodic boundary conditions
- First moment of total displacement is a relaxation vector solution of a PE
- Second moment is solution of a Poisson equations involving first moment
- Moments are computed iteratively by applying Green function. They are derivatives of a moment generating function

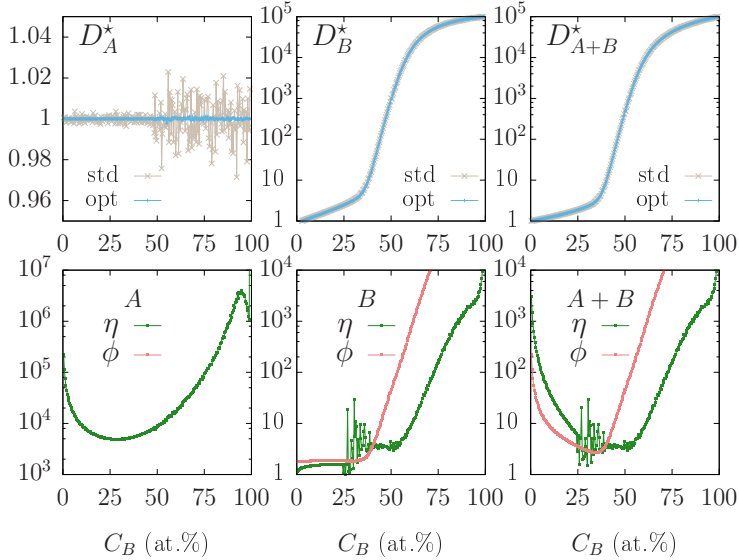
Swinburne and Perez, Nat. Comput. Materials (2020)

Atomic diffusivities measured in kPS with $\nu_1 = 1$ and $\nu_B = 10^5$

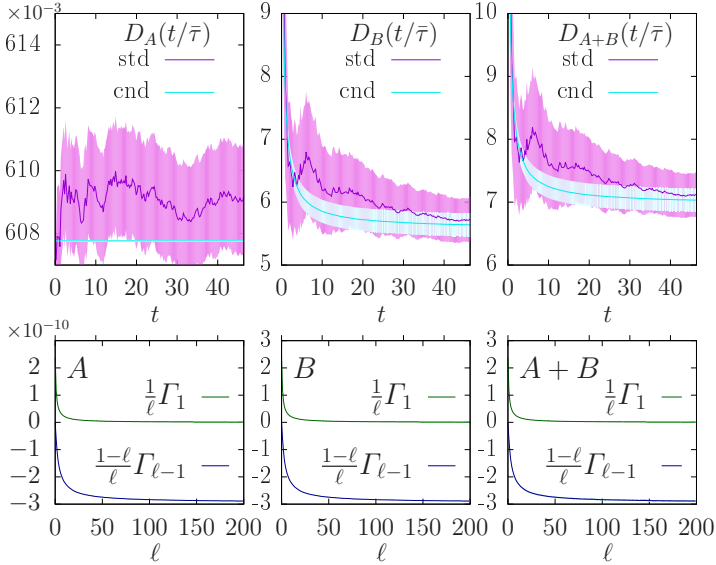
Diffusion coefficients

η : speed-up factor

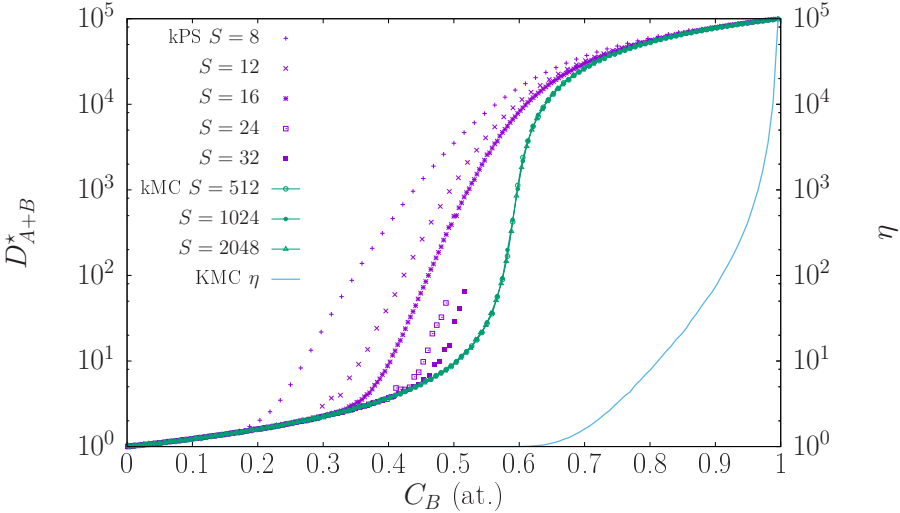
ϕ : intra-to-extra correlated ratio



Atomic diffusivities measured in kPS with $\nu_1 = 1$ and $\nu_B = 10^5$

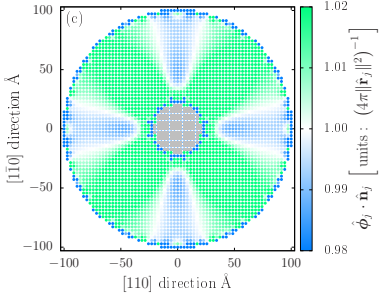


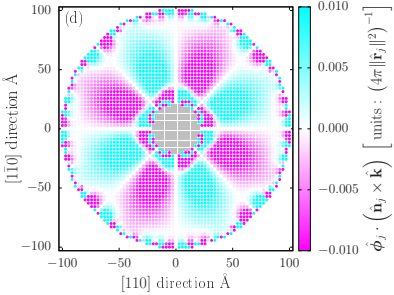
Finite size analysis

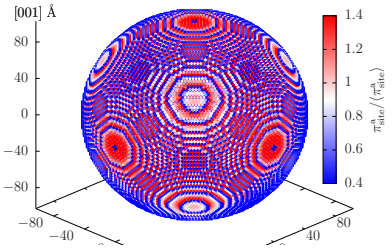


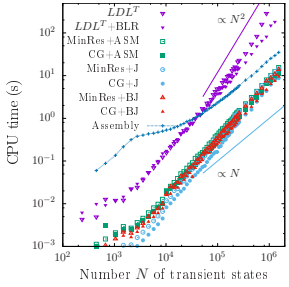
- Beyond percolation extra-correlated part vanishes. Kinetic cluster expansion can be implemented, which consists in retaining the intra-correlated part.

Vacancy emission from cavity in Aluminium









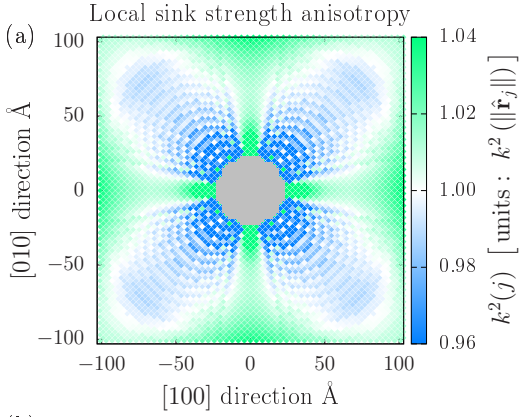
Sink strengths of cavity in aluminum

- Single cavity acting like absorbing sink
- Absorbing Markov chain in periodic cell
- Mean first passage vector $\boldsymbol{\tau}$ is solution of Poisson equation
- Local sink strength

$$k^2(j) = \frac{1}{\tau_j D_V}$$

- Sink strength coefficient

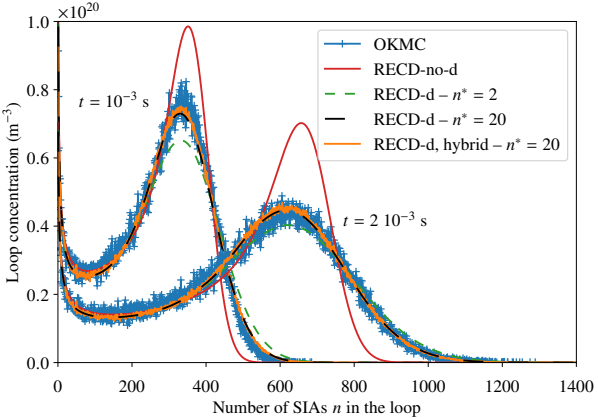
$$k^2 = \frac{1}{\langle \tau_j \rangle D_V}$$



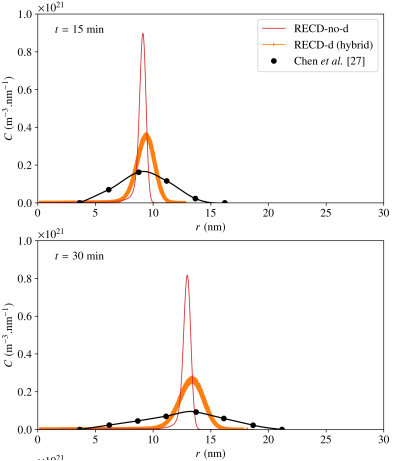
S. Kaur, MA, J. Creuze, J. Comp. Phys. 454, 110987 (2022)

Effect of sink force dispersion on cluster distributions simulated by RECD

KMC and RECD simulations



RECD simulations versus experiments



D. Carpentier, T. Jourdan, MA, Y. Lebouar, J. Nucl. Mat. 533, (2020)

Conclusion

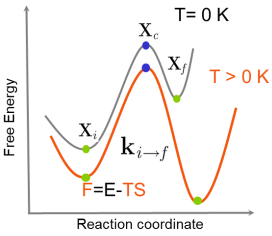
- In free energy computations, conditioning allows to construct efficient estimators within the adaptive biasing force method
  - Application to the calculation of the thermo-elastic properties of Tungstem up to the melting point using data-driven force fields
- In mass transport computations, conditioning leads to a splitting of correlation and to the formulation of a law of total diffusion
  - optimal estimator based on a control variate
  - amenable to conventional and advanced KMC simulations at equilibrium
- Direct computation of **transient** transport properties of defects towards absorbing sinks
  - numerical solution of Poisson equation using sparse solvers
  - characterization of absorbing efficiencies and sink strengths
  - inclusion of the dispersion of sink strengths in rate-equation cluster dynamic simulations yields better agreement with KMC simulations

Appendix: Accelerated Bayesian Adaptive Biasing Force (BABF) Method

Adaptive Biasing Force (ABF): anharmonic free energy calculation

General potential energy: $U(\zeta, \mathbf{r}) = \zeta U(\mathbf{r}) + (1-\zeta)U_{ref}(\mathbf{r})$

Biasing potential: $U_{A_\star}(\zeta, \mathbf{r}) = U(\zeta, \mathbf{r}) - A_\star(\zeta)$



ζ: coupling parameter

$\mathbf{r}$: configuration

$U(\mathbf{r})$: potential energy of target system

$U_{ref}(\mathbf{r})$: potential energy of reference system

$A_\star(\zeta)$: bias, discretized form: $A_n(\zeta)$

Proven convergence:

$$\lim_{n\to+\infty} \Delta A_n = \Delta A = A(1) - A(0)$$

Bayesian scheme[1] Bayes relation: $p_{A_\star}(\mathbf{r}|\zeta) = \frac{p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r})}{P_{A_\star}(\zeta)}$

Mean force: $A'(\zeta) = \int_{T^{3N_a}} \partial_\zeta U(\zeta,\mathbf{r}) p_{A_\star}(\mathbf{r}|\zeta) d\mathbf{r} = \frac{\int_{T^{3N_a}} \partial_\zeta U(\zeta,\mathbf{r}) p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}{\int_{T^{3N_a}} p_{A_\star}(\zeta|\mathbf{r}) P_{A_\star}(\mathbf{r}) d\mathbf{r}}$

Ref: [1] L. Cao, M. Athènes *et al.*, J. Chem. Phys. 140 (2014).

STEP 1

$$A'_n(\zeta) = \frac{\sum_{s=1}^{n-1} \partial_\zeta U(\zeta, \mathbf{r}_s) p_{A_s}(\zeta|\mathbf{r}_s) w(s)}{\sum_{s=1}^{n-1} p_{A_s}(\zeta|\mathbf{r}_s) w(s)}$$

weighting function

STEP 2

$$A_n(\zeta) = \int_0^\zeta A'_n(\tilde{\zeta}) d\tilde{\zeta} + A_n(0)$$

STEP 3

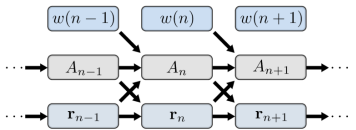
$$p_{A_n}(\zeta|\mathbf{r}_n) = \frac{\exp\left[-\beta U_{A_n}(\zeta, \mathbf{r}_n)\right]}{\int_0^1 \exp\left[-\beta U_{A_n}(\tilde{\zeta}, \mathbf{r}_n)\right] d\tilde{\zeta}}$$

STEP 4

$$\mathbf{F}_{A_n}(\mathbf{r}_n) = -\int_0^1 \nabla_\mathbf{r} U(\zeta, \mathbf{r}_n) p_{A_n}(\zeta|\mathbf{r}_n) d\zeta$$

STEP 5 Langevin dynamics

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{P}\mathbf{F}_{A_n}(\mathbf{r}_n)\delta t + \sqrt{2\beta^{-1}\delta t} B_n$$



Appendix: Application to Thermodynamic Properties of BCC Tungsten

Polynomial model for B^T, C_{11}^T, C'^T and C_{44}^T of W derived from KNML

$$\frac{B^T(T)}{B^T(T_{\text{Debye}})} = -4.434 \times 10^{-12} T^3 - 2.082 \times 10^{-9} T^2 - 4.042 \times 10^{-5} T + 1.013$$

$$\frac{C_{11}^T(T)}{C_{11}^T(T_{\text{Debye}})} = -3.018 \times 10^{-13} T^3 - 2.209 \times 10^{-8} T^2 - 6.875 \times 10^{-5} T + 1.029$$

$$\frac{C'^T(T)}{C'^T(T_{\text{Debye}})} = 5.800 \times 10^{-12} T^3 - 5.166 \times 10^{-8} T^2 - 1.103 \times 10^{-4} T + 1.054$$

$$\frac{C_{44}^T(T)}{C_{44}^T(T_{\text{Debye}})} = -2.592 \times 10^{-12} T^3 - 1.343 \times 10^{-10} T^2 - 6.616 \times 10^{-5} T + 1.026$$

Kernel Noise Machine Learning Potential (KNML)

MiLaDy

Energy of system s containing atom a:

$$E_s = \sum_{a \in s} \epsilon_{s,a}$$

STEP 1. Linear regression of descriptors $\mathbf{D}_{s,a}$

$$\epsilon_{s,a}^{\text{KNML}} = \epsilon_{s,a}^{\text{LML}} + \epsilon_{s,a}^{\text{KML}}$$

STEP 2. Linear regression of $\mathbf{k}(\mathbf{D}_{s,a})$

$$\mathbf{k}(\mathbf{D}_{s,a}) = \begin{pmatrix} k(\mathbf{D}_{s,a}, \mathbf{Z}_1) \\ k(\mathbf{D}_{s,a}, \mathbf{Z}_2) \\ \vdots \\ k(\mathbf{D}_{s,a}, \mathbf{Z}_K) \end{pmatrix} \in \mathbb{R}^{K \times 1}$$

K sparse points

Normalize

Polynomial kernel

$$\tilde{k}(\mathbf{D}_{s,a}, \mathbf{Z}_k) = \left(\sigma^2 + \frac{\mathbf{D}_{s,a} \cdot \mathbf{Z}_k}{2l^2} \right)^p$$