Ab Initio Canonical Sampling based on Variational Inference

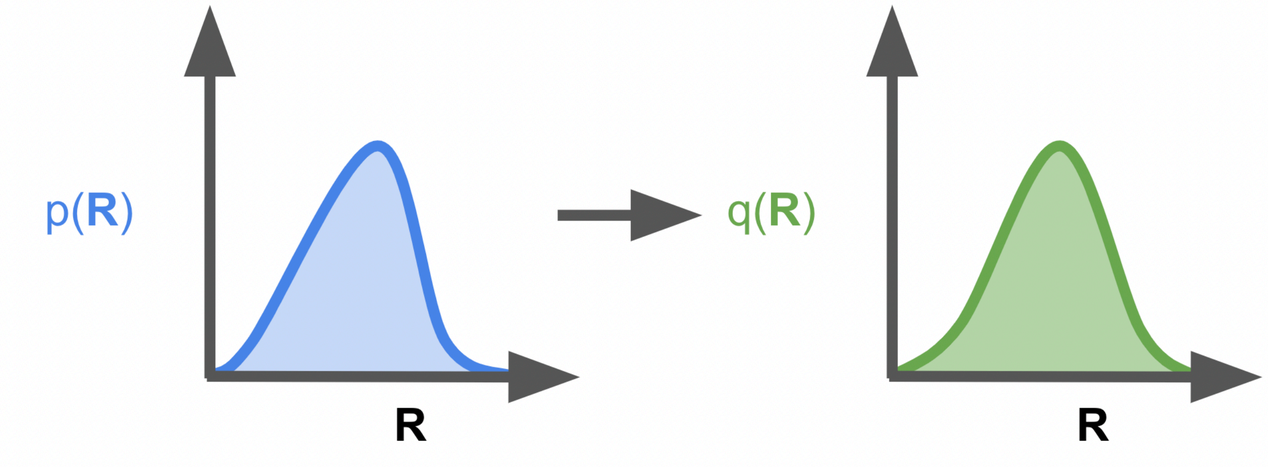
Aloïs Castellano
Nanomat - Université de Liège

François Bottin
CEA DAM/DIF, Bruyères-le-Châtel

Johann Bouchet
CEA DES/IRESNE/DEC, Cadarache

Antoine Levitt
Inria, MATHERIALS, Ecole des Ponts ParisTech

Gabriel Stoltz
Inria, CERMICS, Ecole des Ponts ParisTech



Introduction

Ab Initio Molecular Dynamics (AIMD)

To compute finite temperature properties -> often we need to compute averages of thermalized configurations

Workhorse is *Ab Initio* Molecular Dynamics : iterative algorithm to integrate Newton's equation of motion (often modified with a thermostat)

Huge drawback : very high computational cost -> needs thousands of configurations to be computed with DFT

Variational inference

Sampling using a surrogate distribution

True distribution

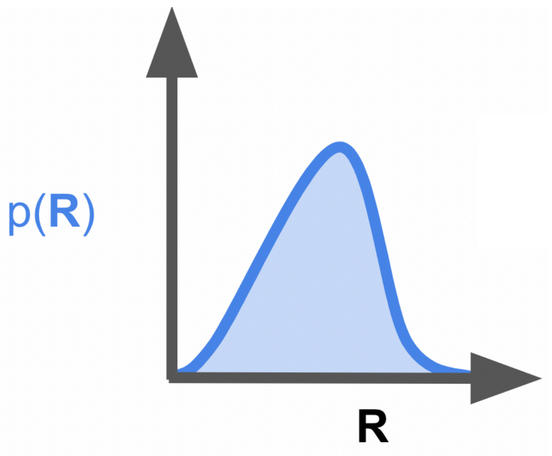
$p(\mathbf{R})$

Surrogate distribution

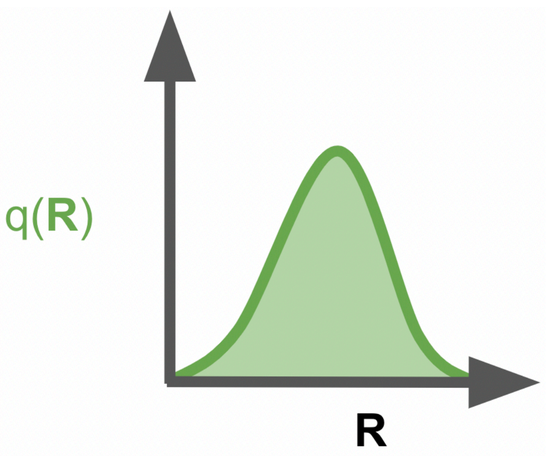
$q_\gamma(\mathbf{R})$

Kullback-Leibler divergence

$$D_{\mathrm{KL}}(p\|q_\gamma) = \int d\mathbf{R} \ln\left(\frac{q_\gamma(\mathbf{R})}{p(\mathbf{R})}\right) q_\gamma(\mathbf{R}) \geq 0$$



The lower $D_{\mathrm{KL}}(p\|q_\gamma)$, the closer $p(\mathbf{R})$ and $q_\gamma(\mathbf{R})$ are



We want to find the $q_\gamma(\mathbf{R})$ that minimize $D_{\mathrm{KL}}(p\|q_\gamma)$

Variational inference

Sampling using a surrogate distribution

True distribution

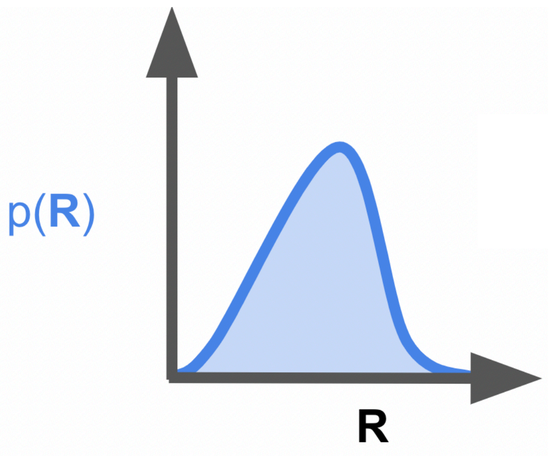
$$p(\mathbf{R}) = \frac{1}{Z} e^{-\beta V(\mathbf{R})}$$

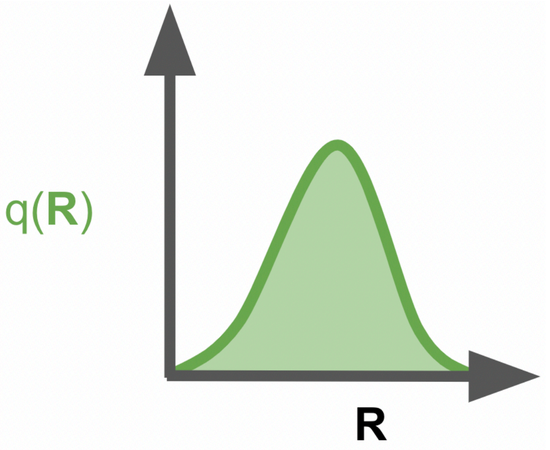
Surrogate distribution

$$q_\gamma(\mathbf{R}) = \frac{1}{\widetilde{Z}} e^{-\beta \widetilde{V}(\mathbf{R})}$$

Kullback-Leibler divergence

$$D_{\mathrm{KL}}(p \| q_\gamma) = \int d\mathbf{R} \ln\left(\frac{q_\gamma(\mathbf{R})}{p(\mathbf{R})}\right) q_\gamma(\mathbf{R}) \geq 0$$





The lower $D_{\mathrm{KL}}(p \| q_\gamma)$, the closer $p(\mathbf{R})$ and $q_\gamma(\mathbf{R})$ are

We want to find the $q_\gamma(\mathbf{R})$ that minimize $D_{\mathrm{KL}}(p \| q_\gamma)$

Variational inference

Sampling using a surrogate distribution

True distribution

$$p(\mathbf{R}) = \frac{1}{Z} e^{-\beta V(\mathbf{R})}$$

Surrogate distribution

$$q_\gamma(\mathbf{R}) = \frac{1}{\widetilde{Z}} e^{-\beta \widetilde{V}(\mathbf{R})}$$

Kullback-Leibler divergence

$$D_{\mathrm{KL}}(p \| q_\gamma) = \int d\mathbf{R} \ln\left(\frac{q_\gamma(\mathbf{R})}{p(\mathbf{R})}\right) q_\gamma(\mathbf{R}) \geq 0$$

The lower $D_{\mathrm{KL}}(p \| q_\gamma)$, the closer $p(\mathbf{R})$ and $q_\gamma(\mathbf{R})$ are

We want to find the $q_\gamma(\mathbf{R})$ that minimize $D_{\mathrm{KL}}(p \| q_\gamma)$

Gibbs-Bogoliubov inequality

$$\widetilde{\mathscr{F}} = \mathscr{F}_\gamma + \langle V(\mathbf{R}) - \widetilde{V}(\mathbf{R}) \rangle \geq \mathscr{F}$$

We want to find the $\widetilde{V}(\mathbf{R})$ that minimize $\widetilde{\mathscr{F}}$

Variational inference

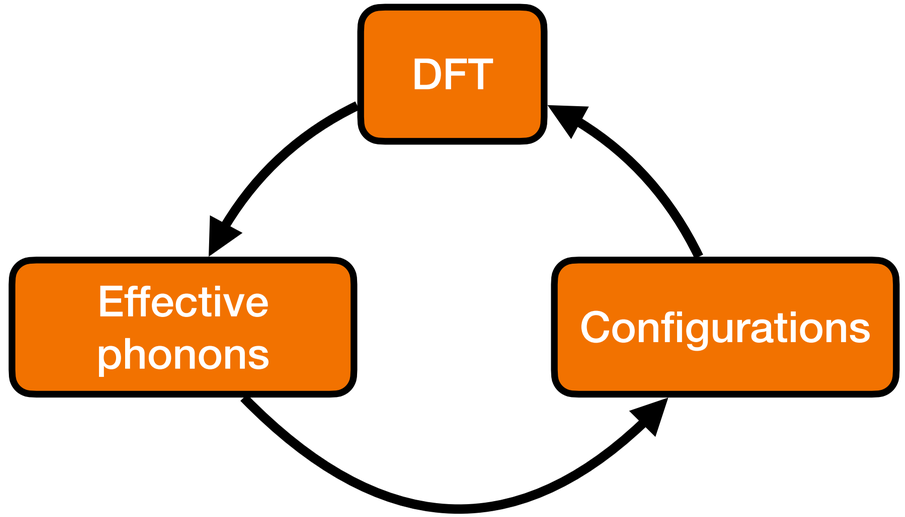
Similar approach in condensed matter physics

Self-consistent harmonic approximation

$$\widetilde{V}(\mathbf{R}) = \frac{1}{2}\mathbf{u}\mathbf{\Phi}\mathbf{u} \qquad q_\gamma(\mathbf{R}) = e^{-\beta\frac{1}{2}\mathbf{u}\mathbf{\Phi}\mathbf{u}}$$

$$\underbrace{\widetilde{\mathscr{F}} = \mathscr{F}_\gamma + \langle V(\mathbf{R}) - \widetilde{V}(\mathbf{R})\rangle}$$

To be minimized self-consistently



SSCHA : L. Monacelli *et al* 2021 J. Phys.: Condens. Matter 33 363001 (2021)

sTDEP : N. Shulumba *et al* Phys. Rev. B 95, 014302 (2017)

QSCAILD : A. van Roekeghem *et al* Comput. Phys. Commun. 263 107945 (2021)

Variational inference

Similar approach in condensed matter physics

Self-consistent harmonic approximation

$$\widetilde{V}(\mathbf{R}) = \frac{1}{2}\mathbf{u}\mathbf{\Phi}\mathbf{u} \qquad q_\gamma(\mathbf{R}) = e^{-\beta\frac{1}{2}\mathbf{u}\mathbf{\Phi}\mathbf{u}}$$

$$\underbrace{\widetilde{\mathscr{F}} = \mathscr{F}_\gamma + \langle V(\mathbf{R}) - \widetilde{V}(\mathbf{R})\rangle}$$

To be minimized self-consistently

Shape of the distribution is limited to gaussians

Machine-Learning Assisted Canonical Sampling

Using Machine-Learning potential for variational inference

Linear MLIP

$$\widetilde{V}(\mathbf{R}) = \sum_i \sum_k \gamma_k \widetilde{D}_{k,i}(\mathbf{R}) \qquad q_\gamma(\mathbf{R}) = e^{-\beta \widetilde{V}(\mathbf{R})}$$

$$\widetilde{\mathscr{F}} = \mathscr{F}_\gamma + \langle V(\mathbf{R}) - \widetilde{V}(\mathbf{R}) \rangle$$

To be minimized

Or equivalently, minimize

$D_{\mathrm{KL}}(p \| q_\gamma)$

Machine-Learning Assisted Canonical Sampling

How to minimize the effective free energy ?

Gibbs-Bogoliubov free energy

$$\widetilde{\mathscr{F}} = \mathscr{F}_{\gamma} + \langle V(\mathbf{R}) - \widetilde{V}(\mathbf{R}) \rangle$$

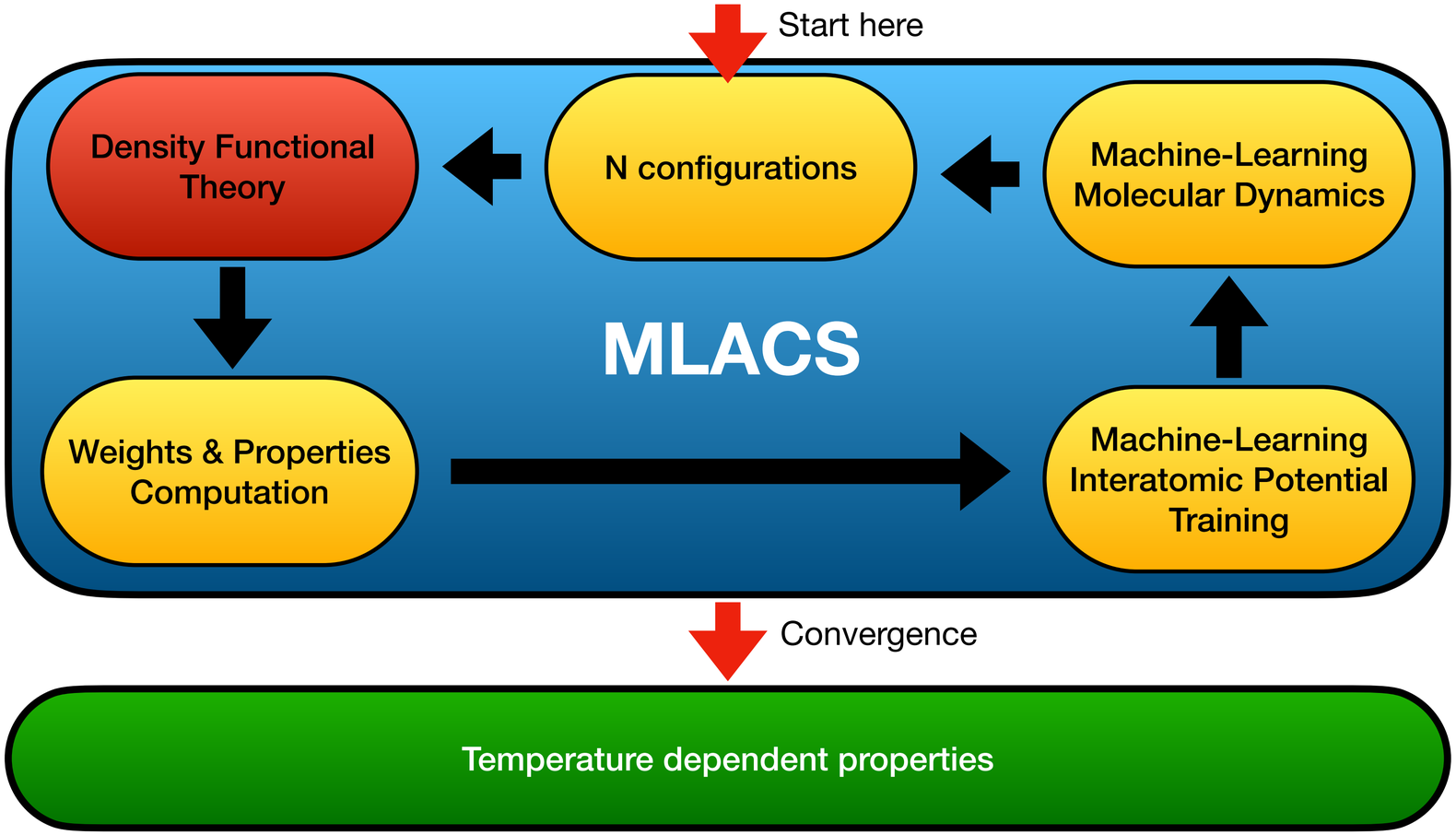
We minimize this quantity with respect to the MLIP coefficients γ

For linear MLIP

Self-consistent ordinary least-squares

Machine-Learning Assisted Canonical Sampling

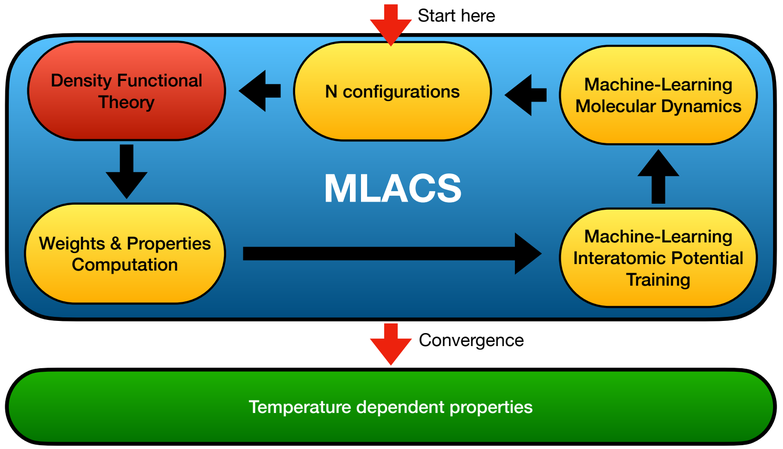
Algorithm



- **Small number of DFT calculations needed**
- **DFT calculations can be done in parallel**

Machine-Learning Assisted Canonical Sampling

- **Results are the configurations**
- **Properties are computed with DFT observables**
- **The MLIP is a tool to create the configurations, but can still be used to obtain properties/ statistics (but with low transferability)**



Results

To check the accuracy of the method we will compute finite temperature phonons

TDEP

$$\mathbf{\Phi}_{ij} = -\langle \mathbf{u}_i \mathbf{u}_j \rangle^{-1} \langle \mathbf{u}_i \mathbf{F}_j \rangle$$

$\Omega_\lambda(T)$

Average on thermalized configurations

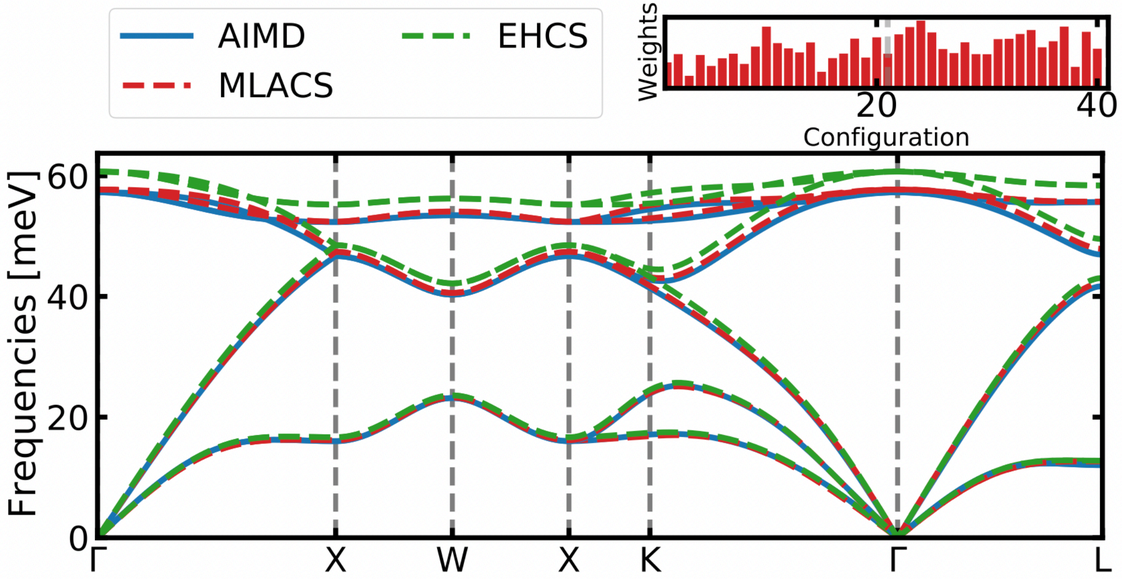
O. Hellman *et al* Phys. Rev. B 84, 180301 (2011)

For the following examples, we will use the SNAP potential as a MLIP:

SO(4) descriptor with linear functional

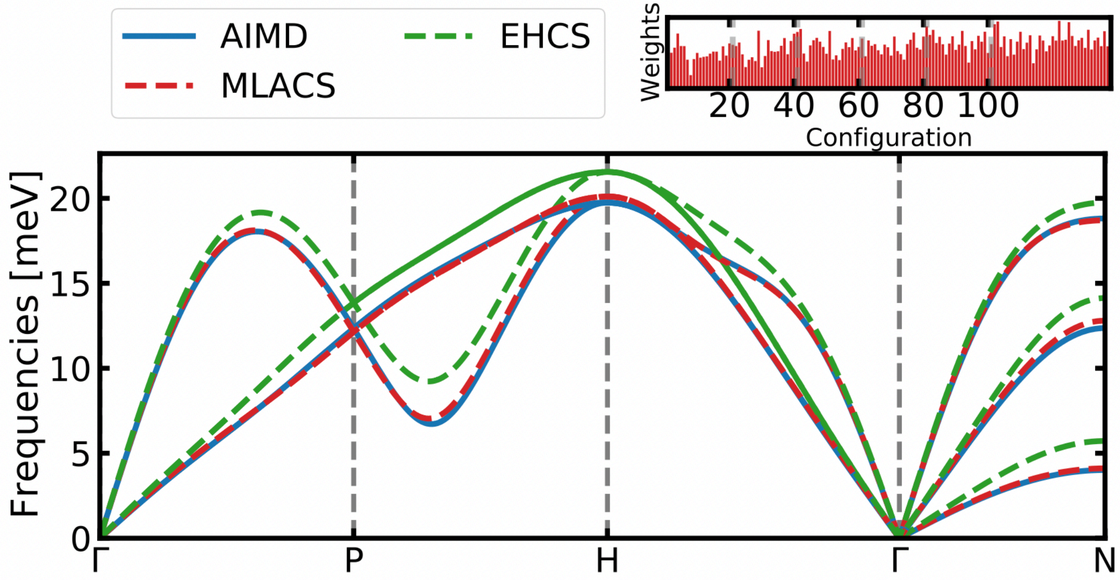
A.P. Thompson *et al* J. Comput. Phys. 285, 316 (2015)

Results - Finite temperature phonons



Silicium 900K

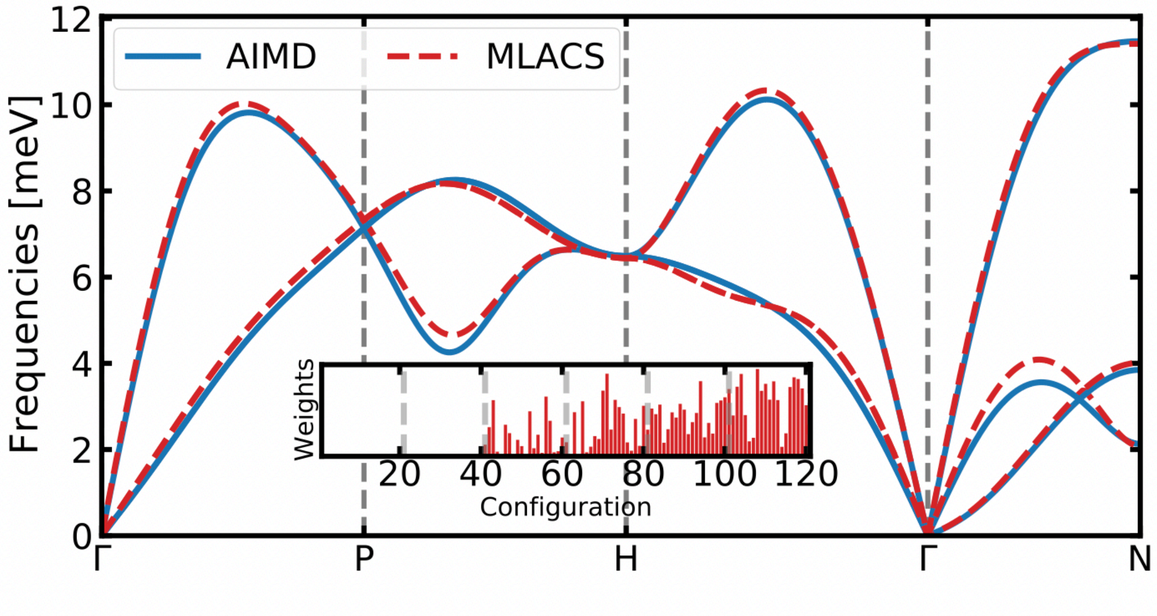
- **AIMD 4000 configurations**
- **MLACS 40 configurations**



Zirconium 1000K

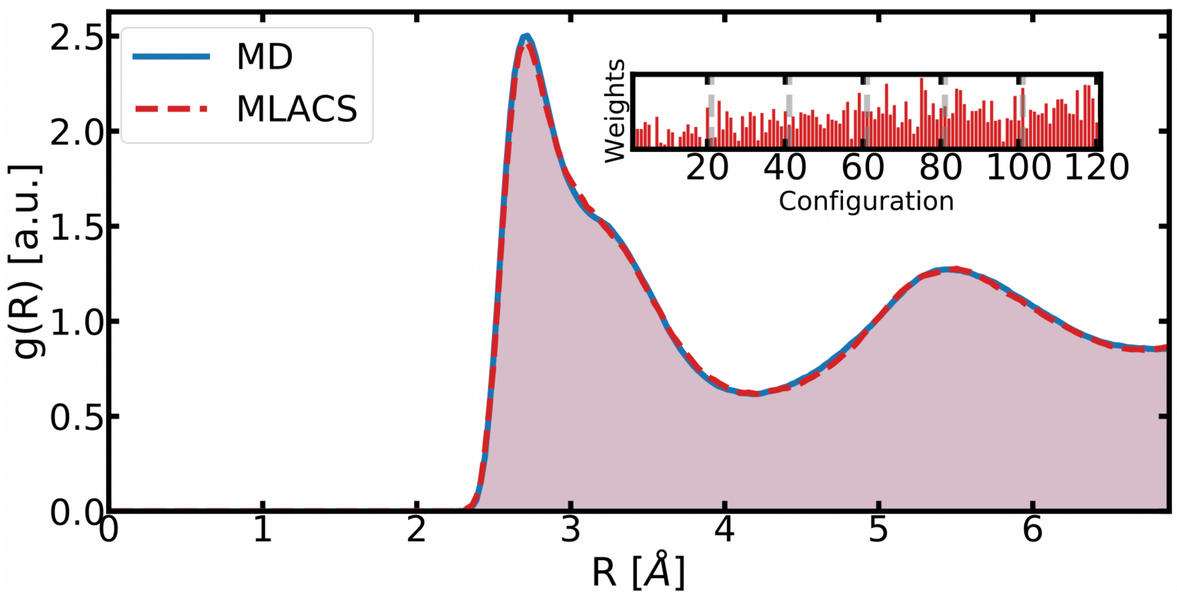
- **AIMD 8000 configurations**
- **MLACS 160 configurations**

Results - Finite temperature phonons



Uranium 1200K

- AIMD 6000 configurations
- MLACS 120 configurations



Uranium (MEAM) 2500K

- AIMD 15000 configurations
- MLACS 120 configurations

Some results - Free energy

Free energy reference system

(Einstein crystal, LJ)

↓ Thermodynamic integration

Free energy MLIP

↓ Thermodynamic perturbation

Free energy of the real potential

	$\widetilde{\mathcal{F}}$ [eV/at]	$\mathcal{F}$ [eV/at]	$\Delta\mathcal{F}$ [meV/at]
Silicon T=1500K	-5.0845	-5.0843	-0.2
$Al_{0.5}Cu_{0.5}$ **T=600K**	-3.7076	-3.7073	-0.3
U liquide T=2500K	-7.4786	-7.4790	0.4

No new DFT computation necessary !

Conclusion

Machine-Learning Assisted Canonical Sampling

- **Variational inference method based on machine-learning potential**
- **DFT accuracy with a fraction of the cost of AIMD**

- **Can be extended to other ensembles (multi-thermal, path-integral,…)**

Thank you for your attention !