

# Accelerating Equilibration in First-Principles Molecular Dynamics with Orbital-Free Density Functional Theory

Lenz Fiedler,<sup>1,2,\*</sup> Zhandos A. Moldabekov,<sup>1,2,†</sup> Xuecheng Shao,<sup>3,‡</sup> Kaili Jiang,<sup>3,§</sup> Tobias Dornheim,<sup>1,2,¶</sup> Michele Pavanello,<sup>3,4,\*\*</sup> and Attila Cangi<sup>1,2,||</sup>

<sup>1</sup>*Center for Advanced Systems Understanding (CASUS), D-02826 Görlitz, Germany*

<sup>2</sup>*Helmholtz-Zentrum Dresden-Rossendorf, D-01328 Dresden, Germany*

<sup>3</sup>*Department of Chemistry, Rutgers University, 73 Warren St., Newark, NJ 07102, USA*

<sup>4</sup>*Department of Physics, Rutgers University, 101 Warren St., Newark, NJ 07102, USA*

(Dated: September 5, 2022)

We introduce a practical hybrid approach that combines orbital-free density functional theory (DFT) with Kohn-Sham DFT for speeding up first-principles molecular dynamics simulations. Equilibrated ionic configurations are generated using orbital-free DFT for subsequent Kohn-Sham DFT molecular dynamics. This leads to a massive reduction of the simulation time without any sacrifice in accuracy. We assess this finding across systems of different sizes and temperature, up to the warm dense matter regime. To that end, we use the cosine distance between the time series of radial distribution functions representing the ionic configurations. Likewise, we show that the equilibrated ionic configurations from this hybrid approach significantly enhance the accuracy of machine-learning models that replace Kohn-Sham DFT. Our hybrid scheme enables systematic first-principles simulations of warm dense matter that are otherwise hampered by the large numbers of atoms and the prevalent high temperatures. Moreover, our finding provides an additional motivation for developing kinetic and noninteracting free energy functionals for orbital-free DFT.

## I. INTRODUCTION

After decades of advances and methodological improvements, density functional theory (DFT) [1] has become the computational tool of choice for solving the majority of computational materials science problems from first principles. It delivers impactful predictions and insights that propel scientific progress due to its balance of computational cost and accuracy [2]. Yet, new scientific problems challenge even the most efficient DFT implementations. Over the last decades, understanding materials under extreme conditions, i.e., at high temperatures and large pressures, has become an emerging field of research. Most notably, studying *warm dense matter* (WDM) is currently in the spotlight [3–5]. In WDM both the Wigner-Seitz radius  $r_s$  and the reduced temperature  $\theta = \tau/\tau_F$  are close to unity, where  $\tau_F$  denotes the Fermi temperature. This poses a challenge, because Coulomb correlations, fermionic exchange, and thermal excitations are all equally relevant. While modeling WDM using classical approaches is thus inaccurate, the use of quantum mechanical methods, such as DFT, is computationally expensive.

Besides theoretical hurdles in applying Kohn-Sham DFT (KS-DFT) [6] to matter under extreme conditions, one is quickly confronted with seemingly insurmountable scaling limitations. KS-DFT scales unfavorably with growing temperatures  $\tau$ , as illustrated in Fig. 1. For small temperatures, we observe a linear growth of computation time with temperature, most likely due to efficient algorithms used in the employed DFT code. As we move to larger temperatures, the expected cubic scaling of DFT with growing temperature can be observed, which makes simulations in the WDM regime, i.e., at  $\theta \gtrsim 1$

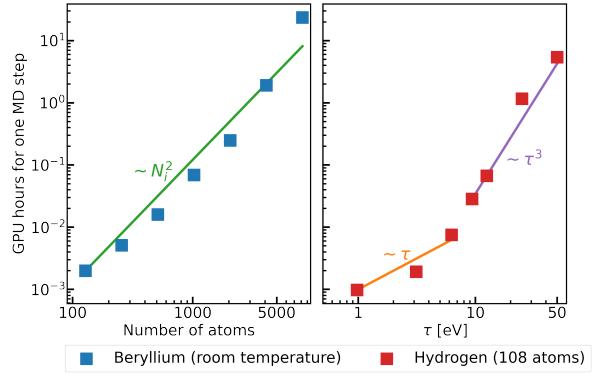


Figure 1. The scaling behavior of KS-DFT as a function of system size for Beryllium at room temperature (left) and as a function of temperature for Hydrogen (right). The computational parameters of the KS-DFT-MD calculations are provided in Sec. II F. Please note that the apparent decrease in performance when going from 4096 to 8192 atoms is caused by the necessity to move from one compute node to two for simulations of this size. Thus, the simulation of 8192 Beryllium atoms is the only one affected by an additional communication penalty.

often impractical. Approaches to circumvent this shortcoming and achieve DFT calculations that scale linearly with temperature are an area of active research [7–9].

Likewise modeling WDM requires extended length scales, i.e., large simulation cells to alleviate potential finite-size problems. While there exist approaches that permit DFT calculations to scale linearly with system size [10–14], usually based on density matrix approaches [15], standard DFT codes are capable of retaining only a quadratic scaling behavior with respect to

number of particles (see Fig. 1). This quadratic scaling becomes intractable quickly, when pushing calculations towards thousands of atoms.

Overcoming these limitations has been subject of active research. One option is orbital-free DFT (OF-DFT) [16], which scales more favorably than KS-DFT both with temperature and system size [17], but suffers from reduced accuracy due to approximating the kinetic energy functional. The quest for increasingly accurate kinetic energy functionals is ongoing [18–21]. Due to these developments, structural properties of certain elements of interest for WDM can now be computed rather accurately with orbital-free DFT molecular dynamics (OF-DFT-MD) both at ambient [17] and extreme conditions [22].

Another emerging research area is the use of machine learning (ML) models [23] in electronic structure theory. They are capable of massively reducing the computational cost for computing properties at KS-DFT accuracy, but can often suffer from generalization errors. In order to achieve high fidelity, these models need to be initialized with ionic configurations close to those trained on.

When standard KS-DFT molecular dynamics (KS-DFT-MD) simulations are carried out at finite temperatures, a significant amount of compute time goes into the equilibration of the ionic configuration, creating an unnecessary computational overhead. We show that this overhead can be avoided by optimizing the initial ionic configuration. This can be achieved in practice in terms of a hybrid approach that combines OF-DFT-MD with KS-DFT-MD. By using OF-DFT-MD to carefully initialize KS-DFT-MD trajectories, we significantly reduce the computational overhead, both for large simulation cells and high temperatures.

The paper is organized as follows. In Section II, we provide the theoretical background and describe our computational and analysis methods. In Section III, we present the central results of our hybrid approach. These include initializing MD trajectories for ML methods, achieving rapid equilibration in MD simulations with large atom counts, and enabling efficient KS-DFT-MD simulations at high temperatures. We have chosen three different systems to represent these fields of application. For the application to ML methods we show results for Aluminum at room temperature (drawing on findings presented in Ref. [24]), for the simulation of extended length scales we chose Beryllium at room temperature, while for the treatment of large temperatures we consider Hydrogen, which is a highly relevant material in the WDM regime. In Section IV, we summarize our findings and conclude with an outlook on future research.

## II. METHODS

### A. Density Functional Theory

In this work we deal with two flavors of DFT, namely KS-DFT and OF-DFT. Both are computational methods for treating a system of  $N_i$  ions at collective positions  $\mathbf{R}$  and  $N_e$  electrons at collective positions  $\mathbf{r}$ , governed by a many-body Schrödinger equation

$$\hat{H}(\mathbf{r}; \mathbf{R})\Psi(\mathbf{r}; \mathbf{R}) = E\Psi(\mathbf{r}; \mathbf{R}) \quad (1)$$

with the Hamiltonian

$$\begin{aligned} \hat{H}(\mathbf{r}; \mathbf{R}) = & \hat{T}_e(\mathbf{r}) + \hat{T}_i(\mathbf{R}) \\ & + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{ei}(\mathbf{r}; \mathbf{R}) + \hat{V}_{ii}(\mathbf{R}) \end{aligned} \quad (2)$$

and the electron-ion wave function  $\Psi(\mathbf{r}; \mathbf{R})$ . In Eq. (2),  $\hat{T}_e$  and  $\hat{T}_i$  denote the kinetic energy of electrons and ions, respectively, while  $\hat{V}_{ee}(\mathbf{r})$ ,  $\hat{V}_{ei}(\mathbf{r}; \mathbf{R})$  and  $\hat{V}_{ii}(\mathbf{R})$  denote the electron-electron, electron-ion and ion-ion interaction, respectively. The nature of these interactions leads to Eq. (1) being computationally intractable. In both KS-DFT and OF-DFT, two important concepts are employed to make the coupled electron-ion problem manageable. Firstly, the Born-Oppenheimer approximation [25] is employed, which separates Eq. (1) into an ionic and an electronic problem. This is feasible since the ions are much heavier than the electrons, resulting in a much larger time scale for ionic motion compared to electronic. The ions are therefore treated as classical point particles, see Sec. II B. This leads to a Born-Oppenheimer Hamiltonian for the electronic problem,

$$\begin{aligned} \hat{H}^{BO}(\mathbf{r}; \mathbf{R}) = & \hat{T}^e(\mathbf{r}) + \hat{V}^{ee}(\mathbf{r}) \\ & + \hat{V}^{ei}(\mathbf{r}; \mathbf{R}) + E^{ii}(\mathbf{R}) \end{aligned} \quad (3)$$

that now depends only parametrically on  $\mathbf{R}$ . Secondly, both flavors of DFT rely on the Hohenberg-Kohn theorems [1] which provide a one-to-one correspondence between the ground-state electronic density  $n_0(\mathbf{r})$  and the external potential generated by the ions  $V_{ei}(\mathbf{r})$ . Therefore, all properties of the system defined by the Born-Oppenheimer Hamiltonian can be determined by knowledge of the ground state density.

The total energy functional in both OF-DFT and KS-DFT is expressed as

$$E_{\text{tot}}[n] = T_s[n] + U[n] + E_{\text{xc}}[n] + V_{ei}[n] + E_{ii}, \quad (4)$$

where  $T_s$  denotes the single-particle kinetic energy,  $U$  the Hartree energy, i.e., the electrostatic interaction of the density with itself,  $E_{\text{xc}}$  the exchange-correlation energy that captures energetic contributions of the electron-electron interaction not included in  $U$ ,  $V_{ei}[n]$  the electron-ion interaction energy, and  $E_{ii}$  the ion-ion interaction energy, which amounts to a constant shift in energy.

In practice,  $E_{\text{xc}}[n]$  has to be approximated, and the accuracy of DFT calculations depends primarily on the

choice of approximation. A plethora of suitable functionals exists. They either depend solely on the electronic density such as the local density approximation (LDA) [6, 26]), incorporate further quantities such as the density gradient in the generalized gradient approximation (GGA) functionals [27–29], or also depend on the kinetic energy density in meta-GGAs such as the SCAN functional [30].

The challenge in DFT is to identify  $n_0$ . This is achieved by minimizing the total energy functional with respect to variations in the density  $n$ . However, a central problem is the lack of an exact expression for  $T_s[n]$  as an explicit density functional, although such an expression does exist in terms of a potential functional for the density [31]. The manner in which  $T_s[n]$  is expressed, is the key difference between OF-DFT and KS-DFT. In KS-DFT,  $T_s[n]$  is not approximated. Instead the minimization of the total energy functional is performed with respect to single-particle orbitals  $\phi_j(\mathbf{r})$  yielding the KS equations [6]

$$\left[ -\frac{1}{2} \nabla^2 + v_s(\mathbf{r}) \right] \phi_j(\mathbf{r}) = \epsilon_j \phi_j(\mathbf{r}), \quad (5)$$

where  $\epsilon_j$  denote the single-particle eigenvalues and  $v_s$  the Kohn-Sham potential. This auxiliary system is restricted to reproduce the density of the interacting system via

$$n(\mathbf{r}) = \sum_{j=0}^{N_e} f_j |\phi_j(\mathbf{r})|^2, \quad (6)$$

with the occupation numbers  $f_j$ . By determining the KS potential  $v_s$  self-consistently, it is ensured that Eq. (6) recovers the density of the interacting system. The KS kinetic energy is determined exactly in terms of the KS orbitals  $\phi_j(\mathbf{r})$  as

$$T_s = \sum_{j=0}^{N_e} \int d\mathbf{r} \phi_j^*(\mathbf{r}) \left( -\frac{1}{2} \nabla^2 \right) \phi_j(\mathbf{r}). \quad (7)$$

OF-DFT [16, 32, 33] follows a different route by approximating the kinetic energy functional *directly* as a functional of the density. Similar to  $E_{\text{XC}}[n]$ , various approximations do exist and are being developed. Commonly used kinetic energy functionals are built using a combination of the Thomas-Fermi functional  $T^{\text{TF}}$  [34, 35], the von Weizsäcker functional  $T^{\text{vW}}$  [36], and a non-local term  $T^{\text{NL}}$ , i.e.,

$$T_s[n] \approx T_{\text{TF}}[n] + T_{\text{vW}}[n] + T_{\text{NL}}[n], \quad (8)$$

with

$$T_{\text{TF}}[n] = \frac{3}{10} (3\pi^2)^{2/3} \int d\mathbf{r} n^{5/3}(\mathbf{r}), \quad (9)$$

$$T_{\text{vW}}[n] = \frac{1}{2} \int d\mathbf{r} \nabla n^{1/2}(\mathbf{r}) \cdot \nabla n^{1/2}(\mathbf{r}). \quad (10)$$

A number of approximations exists for  $T^{\text{NL}}$ , e.g., the Wang-Teter [37], Mi-Genova-Pavanello [38], and Wang-Govind-Carter [39] functionals.

The difference in how the kinetic energy is treated results in contrasting differences for KS-DFT and OF-DFT in terms of accuracy and computational cost. KS-DFT scales formally as  $\sim N_e^3$ , although in practice efficient codes scale with  $\sim N_e^2$ , due to need to solve the KS equations self-consistently. On the other hand, OF-DFT scales linearly with  $N_e$ . Generally, KS-DFT outperforms OF-DFT in terms of accuracy.

Finally, often temperature has to be taken into account in DFT calculations. Finite-temperature DFT for  $\tau > 0$  K [22, 40–42] follows the equations outlined above. One significant change is that the role of the total energy in Eq. (4) is replaced by the total free energy

$$A_{\text{tot}}[n] = T_s[n] - k_B \tau S_s[n] + U[n] + E_{\text{xc}}^\tau[n] + V_{ei}[n] + E_{ii}, \quad (11)$$

which includes the single-particle entropy  $S_s[n]$ . In finite-temperature DFT calculations, one seeks to find  $n$  such that  $A_{\text{total}}$  is minimal. OF-DFT and KS-DFT vary in their evaluation of  $S_s$ . In OF-DFT, an explicit density functional is employed to evaluate this term. It is constructed following the same principles as for  $T_s[n]$  [43, 44], while in KS-DFT this term is evaluated exactly as

$$S_s = - \sum_{j=0}^{N_e} \left[ f^\tau(\epsilon_j^\tau) \ln(f^\tau(\epsilon_j^\tau)) + (1 - f^\tau(\epsilon_j^\tau)) \ln(1 - f^\tau(\epsilon_j^\tau)) \right] \quad (12)$$

with the now temperature-dependent occupation number  $f^\tau(\epsilon)$  and the KS energy eigenvalues  $\epsilon_j^\tau$ . Also the XC energy becomes explicitly temperature dependent. That dependence is, however, often neglected.

The computational cost of KS-DFT increases significantly with increasing temperature, because a larger number of KS orbitals needs to be included in the evaluation of the density, which is done via Eq. (6), with the occupation numbers given by the Fermi-Dirac distribution  $f^\tau(\epsilon_j^\tau)$ . Temperature enters this evaluation both through the temperature dependence of  $f^\tau$  as well the Kohn-Sham eigenvalues  $\epsilon_j^\tau$ . Contrarily, in OF-DFT the temperature has virtually no effect on the computational cost.

## B. Molecular Dynamics

In most MD simulations, the ions are considered as classical point particles. In first-principles MD simulations, the electrons are treated on the quantum level using DFT. ML trained interatomic potentials [45–47] are also gaining traction. They encode the potential energy surface obtained from DFT in an efficient manner and, thus, enable accurate MD simulations at DFT accuracy at large scales. In each time-step  $t$ , each ion  $\alpha = 1 \dots N_i$  with mass  $m_\alpha$  is time evolved on the potential energy

surface  $A_{\text{tot}}$  in terms of the Newtonian equations of motion

$$m_\alpha \frac{d^2 \mathbf{R}_\alpha}{dt^2} = -\frac{\partial A_{\text{tot}}}{\partial \mathbf{R}_\alpha}. \quad (13)$$

The right-hand side of Eq. (13) represents the atomic forces that are determined from either OF-DFT or KS-DFT. After the ionic positions have been updated, a subsequent DFT calculation determines new atomic forces, and uses these to update the ionic positions; such a loop is continued until the desired number of time steps has been performed.

In practice, such MD simulations are realized via thermostats, which ensure sampling according to a proper thermodynamic ensemble, by conserving certain physical quantities. Within this work, we perform DFT-MD simulations in the canonical ( $NVT$ ) ensemble, i.e., the number of particles, volume of the simulation cell, and temperature are fixed. In practice, this is realized via the Nosé-Hoover algorithm [48, 49]. Here, one utilizes a heat bath coupled to the system to ensure constant particle number, volume, and temperature. The coupling is realized in terms of additional terms that are added to the (classical) Hamiltonian of the system, i.e.,

$$H_{\text{NH}} = H_0(\underline{\mathbf{R}}, \underline{\mathbf{p}}/s) + N_{\text{df}} k_{\text{B}} \tau \ln s + \frac{\underline{\mathbf{p}}_s^2}{2Q}, \quad (14)$$

where  $H_0$  is the Hamiltonian of the classical system of ions with collective momenta  $\underline{\mathbf{p}}$ ,  $N_{\text{df}}$  the number of degrees of freedom coupled to the thermostat, and  $s$  an artificial variable that represents the heat bath with momentum  $\underline{\mathbf{p}}_s$ . Introducing  $s$  in Eq. (14) implies a coordinate transformation, i.e., a scaling of the momenta in the system. This framework gives rise to updated equations of motions [50, 51] for the  $\alpha$ -th ion

$$\dot{\underline{\mathbf{R}}} = \frac{\underline{\mathbf{p}}_\alpha}{m_\alpha s^2}, \quad (15)$$

$$\dot{\underline{\mathbf{p}}}_\alpha = -\frac{\partial A_{\text{tot}}}{\partial \underline{\mathbf{R}}_\alpha}, \quad (16)$$

$$\dot{s} = \frac{\underline{\mathbf{p}}_s}{Q}, \quad (17)$$

$$\dot{\underline{\mathbf{p}}}_s = \sum_{\alpha=1}^{N_i} \left( \frac{\underline{\mathbf{p}}_\alpha^2}{m_\alpha s^3} \right) - \frac{N_{\text{df}} k_{\text{B}} \tau}{s}, \quad (18)$$

where the generalized coordinates  $\mathbf{q}$  are set to  $\underline{\mathbf{R}}$  for the sake of simplicity. This system of equations shows that the strength of the coupling of the heat bath to the system is governed by  $Q$ , which represents an imaginary mass.  $Q$  is the principal tuning parameter of the thermostat and has to be carefully chosen to ensure proper sampling. The larger  $Q$ , the less strongly coupled heat bath and simulated system become, and the DFT-MD simulation eventually generates a micro-canonical ensemble. Conversely, too small values of  $Q$  introduce unphysical, periodical temperature oscillations [52].

### C. Equilibration analysis

We now explain how OF-DFT can be used to reduce the computational cost of KS-DFT-MD simulations. At the center of attention is the process of *equilibration*, i.e., the period of a KS-DFT-MD trajectory during which a system is brought to the desired temperature. As we aim to specifically shorten this period of the simulation, we need a systematic approach to determine its length. Equilibration is often done heuristically, i.e., if a trajectory has been propagated for a sufficient time, or if configurations appear reasonably equilibrated, one can start sampling thermodynamic observables. Here, we need to make sure that we systematically capture the first time-step that can be considered equilibrated. We do so in a transferable fashion, as to make assertions on the usefulness of OF-DFT-MD trajectories for KS-DFT-MD simulations. We thus introduce an analysis based on the ionic radial distribution function (RDF), which is defined as the average number of ions  $\Xi(r)$  contained in a hollow shell of radius  $r+dr$  and volume  $V_{\text{shell}}$ , for a cell with  $N_i$  ions, normalized by the ionic density  $\rho = N_i/V_{\text{cell}}$  [53–55], i.e.,

$$g(r) = \frac{\Xi(r)}{\rho N_i V_{\text{shell}}}, \quad (19)$$

which can be shown to be equivalent to

$$g(r) = \frac{\Xi(r)}{4\rho\pi dr \left( r^2 + \frac{dr^2}{12} \right)}. \quad (20)$$

Generally, the RDF is useful to distinguish between different phases or structures. Here, we follow the assumption that if two structures have reasonably close RDFs, they belong to similar phase of a material, i.e., if one of them can be considered equilibrated, so can the other. Our analysis, thus, centers on calculating and comparing RDFs across the trajectory. This poses the problem of choosing an appropriate reference point. One possible approach is averaging the RDF over the end of the trajectory. However, this would necessitate the comparison of such an averaged RDF with noisier RDFs at the beginning of the trajectory, requiring additional averaging.

We therefore follow a slightly different approach and assign the very last ionic configuration of the trajectory as the reference point. Thereafter, the cosine distance between the RDFs of all ionic configurations  $g_a = g(r)[\underline{\mathbf{R}}_a]$  (where  $a$  denotes the number of the time-step) within this trajectory and the reference RDF  $g_{\text{ref}} = g(r)[\underline{\mathbf{R}}_{\text{ref}}]$  is calculated in terms of

$$\tilde{d}_C^a(g_{\text{ref}}, g_a) = 1 - \frac{g_{\text{ref}} \cdot g_a}{\|g_{\text{ref}}\| \|g_a\|}. \quad (21)$$

Here,  $g_{\text{ref}} \cdot g_a$  is the dot product of two RDFs expressed as vectors with dimensionality  $r_{\text{max}}$ , where  $r$  is discretized as  $r = 0, dr, 2dr, \dots, r_{\text{max}}$ . Cosine distances (or conversely, cosine similarities) of data points are a standard

analysis technique in data science [56]. The cosine distance is a measure of how different two vectors are. It becomes 0 for identical vectors, and 1 for entirely dissimilar vectors.

The resulting signal  $d_C^a$  is noisy, as raw RDFs are compared to one another. We thus smoothen  $d_C^a$  using a running average of width  $\sigma$ . Each data point is reassigned as an arithmetic average of the  $\sigma - 1$  preceding data points and itself, i.e.,

$$d_C^a = \frac{\sum_{i=0}^{\sigma-1} \tilde{d}_C^{a-i}}{\sigma} \quad (22)$$

for each time-step  $a$  of the trajectory. This necessitates pruning the trajectory. The first  $\sigma$  and last  $\sigma$  time steps have to be included in the average, but cannot be included in the analysis itself, since they cannot be properly averaged. As such, larger  $\sigma$  lead to slightly smaller trajectories being analyzed.

This smooth signal now allows for analyzing the equilibration behavior using these distances, with  $d_C^a$  decreasing as more and more time steps are performed. Naturally, this analysis is build on the assumption that  $\mathbf{R}_{\text{ref}}$  can be considered equilibrated, but this can easily be confirmed by inspecting  $g_{\text{ref}}$ . Furthermore, clear equilibration patterns become apparent when visualizing  $d_C^a$ , which helps in identifying unequilibrated trajectories.

The next step in determining the first equilibrated configuration is to calculate a threshold for  $d_C^a$ . To do so, we assume a certain portion toward the end of the (pruned) trajectory to be equilibrated, as a fraction of the entire trajectory. Averaging  $d_C^a$  across this portion  $p_{\text{equi}} = 0...1$  of the distance metric, with usually  $p_{\text{equi}}$  around 0.1-0.2, yields an equilibration threshold  $d_T$ . Once the distance metric falls below this threshold after  $N_T$  time steps, we assume the trajectory to be equilibrated henceforth.

Calculating  $d_T$  from an averaged distance metric implies that a trajectory may be equilibrated slightly earlier than the algorithm will detect, since  $d_C^a$  regularly exceeds  $d_T$  when fluctuating around the equilibrated average. Alternatively, one may set  $d_T$  according to the upper limit of such fluctuations (i.e., via the standard deviation), but this is problematic, as not fully equilibrated systems could be interpreted as such. We thus employ the average across the equilibrated configurations rather than the upper limit, to ensure that no unequilibrated system is misidentified. The first equilibrated configuration determined can therefore be interpreted as the point where a trajectory is equilibrated beyond doubt. The algorithm discussed here is further visualized in Fig. 2.

Please note that in its current form this technique is limited to analyzing fully equilibration patterns for trajectories which are fully equilibrated. Dynamical application of this method, i.e., determining whether a trajectory is equilibrated during run time, should be possible in principle, but require further development, exceeding the scope of this work.

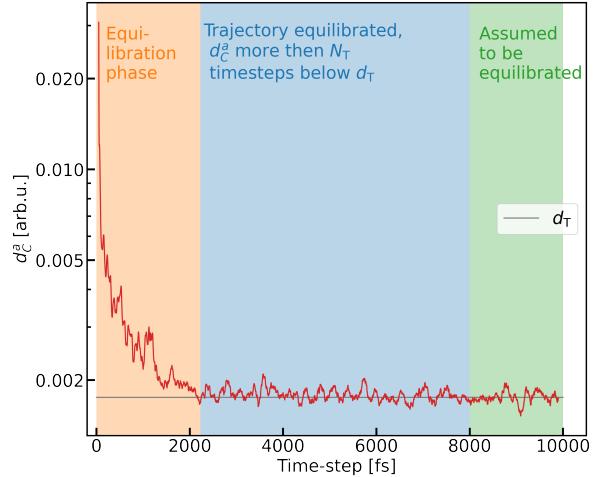


Figure 2. Convergence analysis employ throughout our work, shown for a trajectory of 512 Beryllium atoms at room temperature.

#### D. DFT Surrogate models

ML and data-driven methodologies are becoming increasingly important for DFT applications [23]. Such approaches encompass interatomic potentials, which replace DFT as a means to evaluate the potential energy surface  $A_{\text{tot}}$  and forces  $-\partial A_{\text{tot}}/\partial \mathbf{R}_\alpha$  in MD simulations, as well as property mappings, which learn specific physical and chemical properties from large DFT databases. We have recently [24] introduced an ML based workflow to reproduce DFT calculations in terms of both energy and electronic structure predictions, the Materials Learning Algorithms package (MALA) [57]. MALA predicts the electronic structure of an ionic configuration via a local mapping that predicts the local density of states (LDOS)  $d$  based on descriptors encoding local ionic configuration around points in real space. As such it is in principle capable of predicting the electronic structure of extended systems [58]. This prediction is realized by feed-forward neural networks (NNs).

Our current research is focused on applying such surrogate models to dynamical simulations or use them to perform thermodynamic sampling via Monte Carlo methods. However, it is well known that NNs perform poorly in extrapolation [59], and should mostly be used for interpolation tasks. Therefore, optimal results require initialization of such calculations with configurations for which the model predictions will constitute an interpolation. This can be challenging when predicting the electronic structure of systems of extended size, for which no KS-DFT-MD data can be acquired, yet model predictions *may* be interpolatory, given that physically sound ionic configurations are used.

### E. Hybrid MD workflow

As the central methodological development, we present a hybrid OF-DFT-MD and KS-DFT-MD workflow for dynamical materials modelling. Usually, one has to rely on ideal crystal structures to initialize KS-DFT-MD trajectories or DFT surrogate model sampling routines. As we show in Sec. III, OF-DFT-MD trajectories, which can be calculated at almost negligible computational cost and with favorable scaling behavior, serve as an alternative that improves performance in either case. Especially at larger temperatures, the actual ionic configurations and electronic densities differ greatly from ideal crystal structures, as shown in Fig. 3.

The task of OF-DFT-MD in our hybrid workflow is thus not to *replace* KS-DFT-MD fully, as KS-DFT-MD trajectories are always built on top of the OF-DFT-MD generated structures. Rather, by providing reliable approximations of the actual ionic configurations, OF-DFT-MD (1) reduces the equilibration period for KS-DFT-MD trajectories and (2) improves surrogate model prediction accuracy. There is still some “equilibration” required to transition from slightly differing equilibrated configurations, but the associated number of time-steps and computational cost is vastly reduced. In practice, we simply choose the last configuration of the OF-DFT-MD trajectory as the initial configuration for KS-DFT-MD or DFT surrogate model evaluation.

This workflow is further illustrated in Fig. 4. There, Fig. 4a details the equilibration process in both KS-DFT-MD and OF-DFT-MD starting from the ideal crystal structure; both methods equilibrate to slightly different averages. However, the difference between the equilibrated ionic configurations is drastically smaller than the difference between equilibrated configuration and ideal crystal structure in either case, leading to the aforementioned speed up. In Fig. 4b, which visualizes the entirety of the hybrid workflow, this short “equilibration” phase in going from OF-DFT-MD to KS-DFT-MD is shown in a red background color. In our calculations, we have kept the number of OF-DFT-MD time-steps purposefully high to ensure that no unequilibrated structure enters the analysis. As a result, the OF-DFT-MD propagation of the equilibrated configurations shown in green is longer than necessary. In practice, it can be shortened significantly, because only the initial equilibration of the system (shown in blue) is relevant for the hybrid workflow.

### F. Computational details

All KS-DFT-MD simulations were performed using the highly efficient GPU implementation of VASP [62–66] at the  $\Gamma$ -point. In all cases, a plane-wave basis set with a PAW pseudopotential [67, 68] and the PBE exchange-correlation functional [27–29] were employed. As cutoff energies for the plane wave basis set 440 eV, 248 eV, and 350 eV were used for Aluminum, Beryllium,

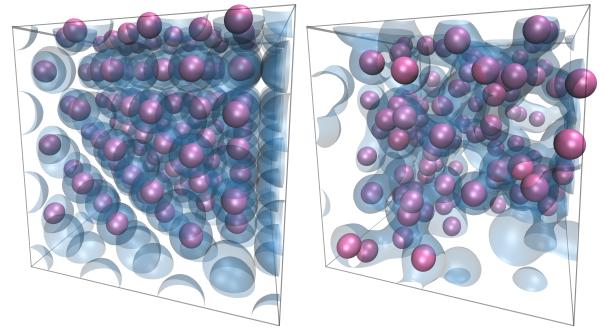
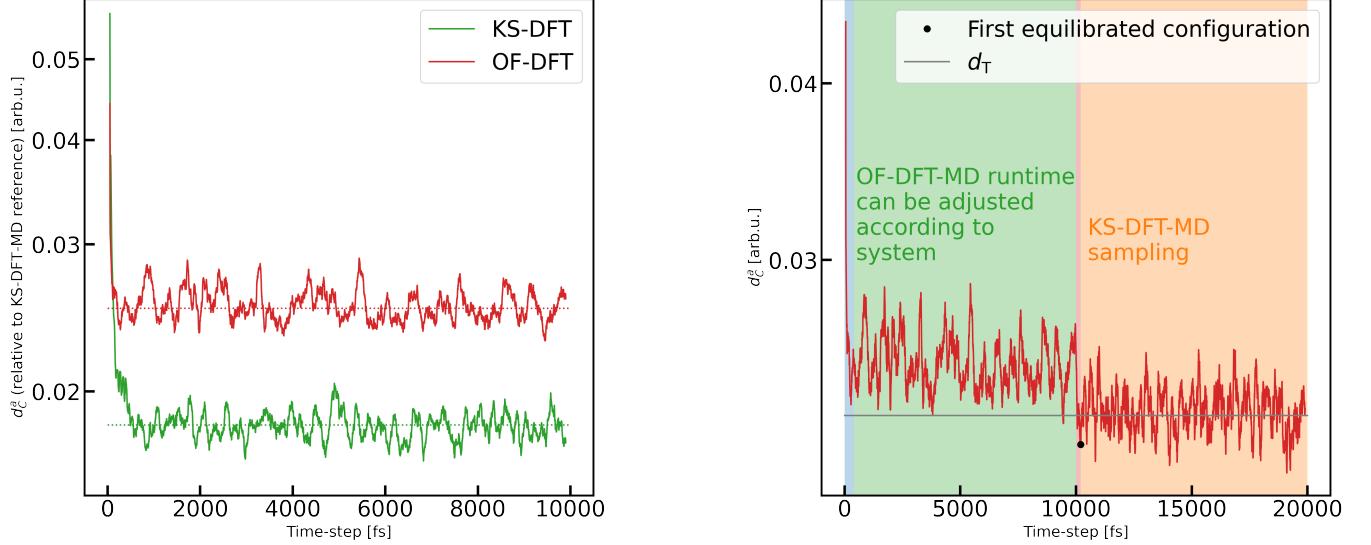


Figure 3. Ideal crystal structure (left) and equilibrated ionic configuration (right) of Hydrogen at  $\tau = 6.26$  eV, including an isosurface of the electronic density, created using VMD [60, 61].

lum, and Hydrogen, respectively. For the single-point KS-DFT calculations in Sec. III A, we employ Quantum ESPRESSO [69–71] with parameters consistent as described in Ref. [24]. Likewise, the same ML model is used. All Beryllium and Aluminium calculations have been carried out at room temperature and ambient mass density, 1.896 g/cc and 2.699 g/cc, respectively, while the Hydrogen simulations have been performed at  $\theta = [0.007, 0.25, 0.5, 0.75]$ , which equals  $\tau = [0.977 \text{ eV}, 3.132 \text{ eV}, 6.264 \text{ eV}, 9.390 \text{ eV}]$ , and a mass density corresponding to  $r_s = 2$ . In the initialization of these calculations for Aluminum, Beryllium and Hydrogen, an fcc, hcp and fcc structure has been used, respectively. The number of orbitals in KS-DFT simulations are adjusted with the change in temperature to ensure the smallest occupation number is not higher than  $10^{-6}$ .

If not otherwise noted, KS-DFT-MD simulations have been performed for 10,000 time-steps with either a time-step of 1 fs (Aluminum and Beryllium) or 0.01 fs (Hydrogen). For Beryllium, different values of  $Q$  have been tested to confirm our findings across different thermostat settings, while for Hydrogen,  $Q$  has been kept fixed at 0.5. For Aluminum,  $Q = 0.001$  has been used. Beryllium trajectories have been analyzed with the aforementioned method and an assumed equilibrated portion of the trajectory  $p_{\text{equi}} = 0.2$ , a running average width  $\sigma = 100$ , and required number of time-steps below the distance threshold for equilibration  $N_T = 50$ . As the higher temperature in the Hydrogen simulations lead to noisier signals,  $N_T$  and  $\sigma$  were both increased to 200. We provide full equilibration graphs in App. B and App. C to further verify the choice of these parameters.

For Aluminum and Beryllium, the Wang-Teter kinetic energy functional [72] with PBE as exchange-correlation functional was used. Optimized effective potentials [73] were used as local pseudopotentials. Given the emphasis on higher temperatures, for Hydrogen we used the finite temperature TF free energy functional with the von Weizsäcker gradient correction in combination with LDA XC functional. In order to perform finite-temperature OF-DFT-MD simulations, we implemented the finite-



(a) Distance metrics for KS-DFT-MD and OF-DFT-MD, both starting from an ideal crystal structure. In either case, distance metrics are given relative to the equilibrated KS-DFT-MD configuration.

(b) Full visualization of the hybrid workflow approach. Blue: Equilibration from ideal crystal structure using OF-DFT-MD; Green: OF-DFT-MD propagation of equilibrated configuration; Red: Short “equilibration” phase in going from OF-DFT-MD to KS-DFT-MD. Orange: KS-DFT-MD sampling of configurations or thermodynamic observables.

Figure 4. Sketch of hybrid workflow based on the data for 128 Beryllium atoms. The data depicted here is the same as analyzed in Sec. III B.

temperature TF functional in the OF-DFT code DFTpy, which was previously introduced for the ground state calculations [17]. The version of DFTpy with the newly added option for the finite-temperature TF free energy functional is now available with the most recent open-source version of DFTpy. The underlying theoretical details of our implementation are outlined in App. A. In all OF-DFT-MD runs, the MD time step was 2 fs and all other simulation parameters were kept consistent with those used for the KS-DFT-MD calculations.

All DFT-MD data, example input scripts, and processing scripts can be obtained from Ref. [74]. All ML experiments and equilibration analysis have been carried out with the MALA code version 1.1.0 [57].

### III. RESULTS

The results are divided into three parts, reflecting three areas of materials modeling for which our workflow is highly useful: (1) the application of ML methods trained on DFT data, (2) simulations with large numbers of atoms with KS-DFT-MD, and (3) simulations of systems at high temperatures using KS-DFT-MD. The utility of the latter two applications is twofold, as improved KS-DFT-MD performance does not only alleviate the cost of direct sampling of thermodynamic observables, but further helps with data generation for future ML applications.

#### A. ML-DFT trajectory initialization

As mentioned above, one potential pitfall of transferable ML surrogate models for DFT calculations is their application to ionic configurations that are too dissimilar from observed training data for the model to perform well. Quantifying uncertainties in model predictions is naturally a crucial tool for detecting such behavior, but it does not solve the problem of providing initial configurations on which the model can be expected to perform reasonably well. One may incorporate ideal crystal structure data as well as data from equilibration phases into the training set to alleviate this problem – but ultimately the interest lies with equilibrated configurations, and as such, one would unnecessarily complicate model training. Any such model would have to perform well on ideal crystals, slowly equilibrating systems as well as the actually equilibrated system. While such models are by no means unachievable (in fact, interatomic potentials often follow this approach), the amount of additional training data can become challenging if the actual electronic structure has to be learned rather than the potential energy surface. We thus propose to circumvent the need to do so by employing ionic configurations from OF-DFT-MD as initial configurations for extended DFT surrogate model driven simulations. Given that OF-DFT-MD captures the KS-DFT geometries on which the ML model has been trained, one can rely on ML predictions to equilibrate the system fully. We have investigated this behavior for Alu-

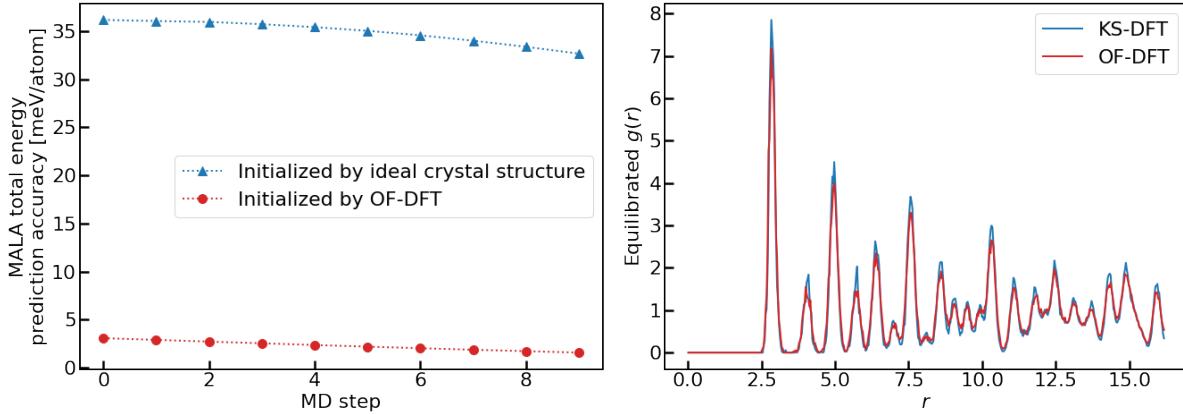


Figure 5. Left: Accuracy of ML surrogate models predictions on KS-DFT-MD trajectories that are initialized with the ideal crystal geometry and OF-DFT-MD. Right: Comparison of the RDF of equilibrated ionic configurations for 256 Aluminum atoms obtained from OF-DFT-MD and from KS-DFT-MD starting from the ideal crystal structure (RDF calculated for one configuration per method).

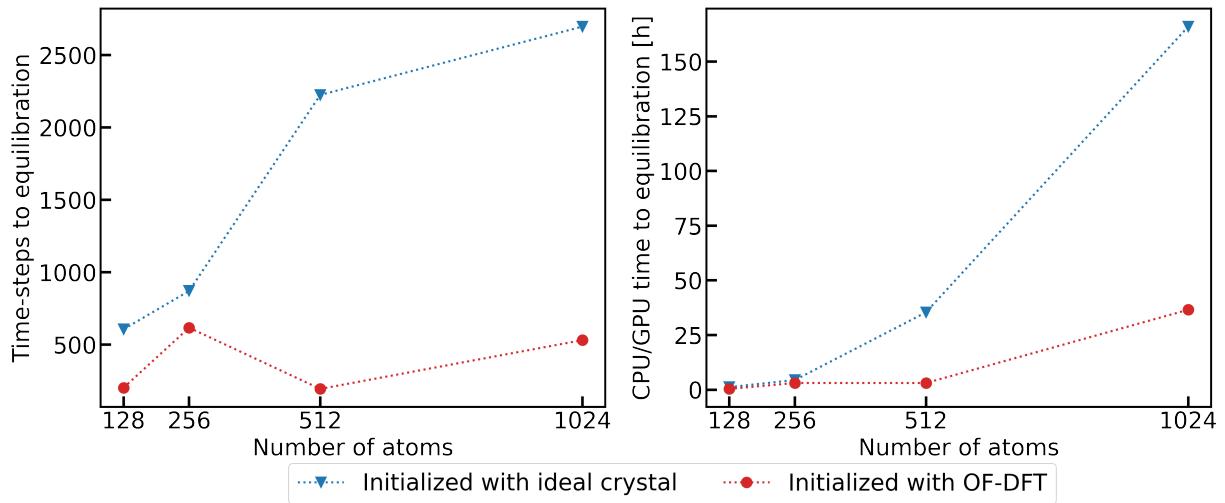


Figure 6. Equilibration of Beryllium simulation cells of increasing size according to our distance metric analysis.

minimum at room temperature. The results are shown in Fig. 5, where we have used our MALA framework introduced above as a representative ML surrogate model. Two different KS-DFT-MD trajectories, each for 10 time steps, are investigated. One is initialized by OF-DFT-MD and the other by the ideal crystal structure. For each configuration in these trajectories, a full KS-DFT calculation is performed alongside a MALA inference (using the 298 K model from Ref. [24]) to determine the total free energies. Additional KS-DFT calculations are necessary, since the model was trained on data generated with Quantum ESPRESSO, while we perform the MD simulation with VASP to be consistent with the other results throughout this investigation.

As shown in Fig. 5, within the first ten time steps, the configurations (and therefore prediction accuracies) per trajectory change only slightly. This means that for the

ideal crystal, we have configurations consistent with 0 K, while for the OF-DFT-MD trajectory, we have configurations consistent with 298 K, with the model having only been trained on 298 K data. It can be seen that this temperature difference amounts to an error of around 35 meV/atom in the total free energy for the ideal crystal trajectory, as the configurations at 0 K represent an out-of-distribution sampling for the ML model, while the OF-DFT-MD initialized trajectory leads to in-distribution sampling. It is well known that NN based approaches excel at the latter task, while performing poorly at the former. In either case, the error reduces as the KS-DFT-MD trajectory equilibrates. When OF-DFT-MD is used to initialize the ionic configuration, the error is below 5 meV/atom from the beginning, which is consistent with the errors reported in Ref. [24] for a full KS-DFT-MD trajectory.

Naturally, the merit of this workflow mainly depends on the ability of OF-DFT-MD to converge to useful and comparable geometries. One way to visually confirm this is in terms of the RDF. As shown in Fig. 5, both KS-DFT-MD and OF-DFT-MD generally agree quite well. While the amplitude of certain peaks is under- or overestimated by the latter, the overall positions match, which is the basis for the performance observed in Fig. 5. As shown in the following sections, this behavior can be supported for increasingly larger systems and temperatures, two important dimensions in which one seeks to build transferable ML models.

Overall, using OF-DFT-MD allows us to create models based solely on equilibrated data and apply them on extended scales using approximately correct initial configurations, thus saving model training time while maintaining high prediction accuracy.

### B. Large length scales

Numerical calculations with KS-DFT-MD are prone to finite-size errors [75], if small simulation cells are used. Due to the aforementioned scaling behavior of DFT, increasingly larger simulations are computationally challenging. This is especially problematic given the fact that for large simulation cells a larger number of time steps is usually required to properly equilibrate the system, yet each time-step is more costly by itself. Decreasing the time to equilibration would therefore drastically improve the performance of such simulations, and can be achieved with OF-DFT-MD. To test our workflow in this setting, we equilibrate Beryllium simulation cells of increasing size at room temperature and ambient mass density starting from the ideal crystal structure and OF-DFT-MD trajectories. For each trajectory we determine the time required to equilibration using the algorithm outlined in Sec. II C. The results of this analysis are given in Fig. 6. Generally, OF-DFT-MD yields an almost perfectly equilibrated geometry. Equilibration is detected after a number of time-steps almost of the same magnitude as  $N_T$ , i.e., the minimum resolution our analysis can provide. Our analysis thus unveils no significant overall equilibration to be necessary for the OF-DFT-MD initialized trajectories, as is exemplified by the full equilibration curves provided in App. B. Furthermore, while a clear increase in initialization time can be seen for trajectories initialized with ideal crystal structure, no such trend is observed for those initialized with OF-DFT-MD. As DFT calculations become increasingly costly with system size, this grows to large amounts of computational time being saved. In the case of 1024 Beryllium atoms, we are able to save more than 100 GPU hours, which correspond to approximately 20% of the entire simulation time. The increase in time-steps required for equilibration is linear with the number of particles in the case of ideal crystal initialization.

Equilibration depends on the thermostat used during

the DFT-MD run. As we use the Nosé-Hoover thermostat, the equilibration time is chiefly influenced by the choice of the Nosé mass  $Q$ . We have verified that our findings are not impacted by the choice of this technical parameter, see App. D

These speed ups are offset by the time needed to perform OF-DFT-MD simulations. These simulation however accrue to only a small computational overhead compared to KS-DFT-MD simulations. For example, a single time-step costs  $t \simeq 3.8$  s ( $t \simeq 13$  s) for 128 (1024) Beryllium atoms using a single CPU, or about 11 (36) hours using one CPU without parallelization for the entire trajectory of 10,000 time steps. While these times may seem excessive by itself, they are easily explained by two important considerations. Firstly, as can be seen above for the case of KS-DFT-MD, 10,000 times-steps are not required for equilibrated configurations under these conditions, we have chosen such a number to ensure that our investigations are not distorted by unequilibrated trajectories. Furthermore, we have obtained all OF-DFT-MD results using a single CPU; such timings do not compare to GPU hours reported for KS-DFT-MD trajectories, even when reporting core-hours rather than wall-time hours. We thus omit OF-DFT-MD timings from Fig. 6 and 15.

### C. High temperatures

As discussed above, KS-DFT is known to scale unfavorably with temperature. To investigate our hybrid workflow for increasingly larger temperatures, we thus choose a computationally more tractable system than Beryllium, namely 108 Hydrogen atoms under a range of temperatures up to  $\theta = 0.75$ , i.e.,  $\tau = 9.39$  eV. We then undertake a similar investigation as in the preceding section, by performing KS-DFT-MD simulations at increasing temperatures, initialized either by the ideal crystal structure or OF-DFT-MD ionic configurations, and analyzing the equilibration patterns. The results are illustrated in Fig. 7.

The temperature dependence of the increase in time steps required to equilibrate a system is not as straightforward as the dependence on the number of particles reported in Fig. 6. Larger temperature fluctuations lead to noisier trajectories, making equilibration, and determination of achieving it, more difficult. The necessity to use smaller simulation cells at the large temperatures investigated further leads to additional noise. This noise is witnessed in the equilibration curves App. C. Fig. 7 shows a slight temperature-dependence of the number of time-steps required to equilibrate a trajectory from an ideal crystal structure. For the OF-DFT-MD initialized trajectory, such a dependence cannot be clearly deduced. Rather, we observe that as temperature increases, the number of time-steps saved by the utilizing an OF-DFT-MD initialized trajectory appears to be roughly constant (with the exception of the result for  $\tau = 6.264$  eV, which

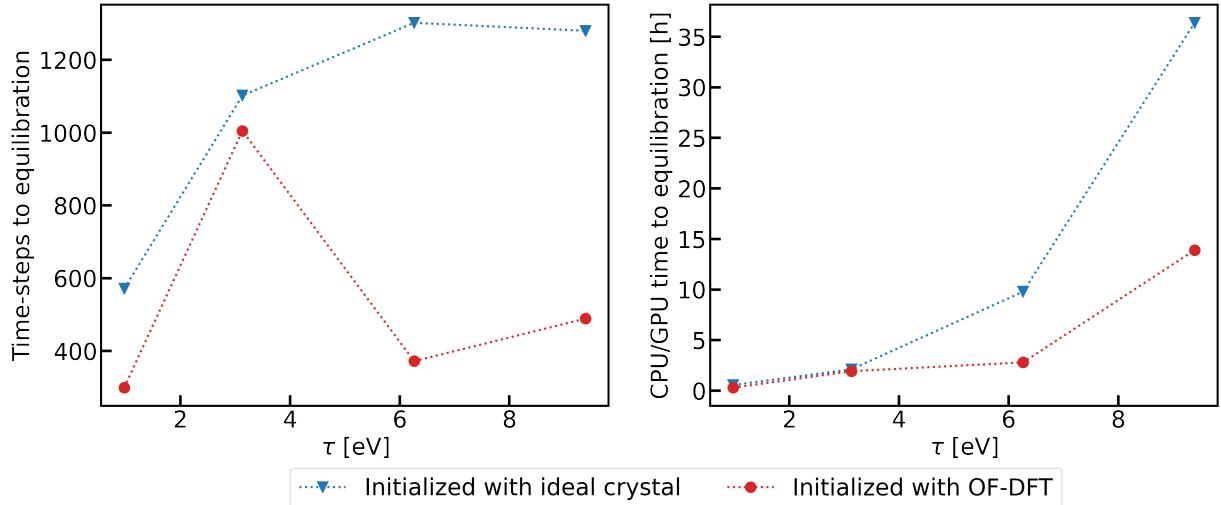


Figure 7. Equilibration of Hydrogen simulations for increasing temperature according to our distance metric analysis.

seems to be an outlier). As the computational cost per individual time step increases with temperature, OF-DFT-MD initialization leads to sizable savings in computation time. As the influence of different thermostat parameters has amply been investigated in App. D, we restrict ourselves to only one thermostat parameter here.

In the case of Hydrogen, one OF-DFT-MD step for 108 particles requires about  $t \simeq 2.5$  s regardless of the temperature value. This results in approximately 7 hours on a single CPU for 10,000 time-steps. Apart from this, the same considerations as discussed in Sec. III B hold true, i.e., less time steps might be necessary in actuality.

Since OF-DFT-MD calculation times do not increase with temperature, while KS-DFT scales so unfavorably due to increase in the number of recurred bands (orbitals), OF-DFT-MD initializations can be crucial when attempting to investigate larger temperature ranges, i.e., in the WDM regime. We see already from the results in Fig. 7, that the saved computational time increases massively for moving to  $\tau = 9.39$  eV, for which the main cause is the drastically increased cost per time-step as temperature increases.

#### IV. CONCLUSION AND OUTLOOK

We have shown that initializing ionic configurations based on OF-DFT-MD greatly reduce the computational cost for equilibrating KS-DFT-MD trajectories. Given that OF-DFT is in principle able to capture the physics of the systems at hand, KS-DFT-MD thereafter requires only very little computational overhead to reach equilibration. We have verified our results on different systems and have shown that our findings hold true as both the system size and temperature increase. Our hybrid workflow is thus widely applicable to simulations of matter under extreme conditions, especially given the favorable

scaling properties of OF-DFT-MD with temperature and system size.

OF-DFT-MD is an especially crucial addition to the toolkit of first-principles simulations when it comes to ML-DFT applications. Not only does it improve performance of data acquisition, as outlined in Sec. III B and III C, but furthermore it directly helps with ensuring that models are used in an interpolative manner, by providing initial configurations on which surrogate model inference yields to accurate predictions.

Yet, there is some work to be done. OF-DFT-MD workflows have to be integrated into larger frameworks, such as MALA. It is further well known that OF-DFT is not universal in the types of systems that can be treated. While metals can generally be treated to high accuracy [17], other systems evade accurate treatment. Theoretical development and the advent of ML based kinetic energy functionals may alleviate these problems. Our presented hybrid approach, where accuracy of the OF-DFT part is essential for the acceleration of the simulations (but not for the accuracy of the final KS-DFT results), represents another motivation for further developing kinetic energy and entropy functionals for OF-DFT applicable to various materials.

#### ACKNOWLEDGMENTS

This work was partially supported by the Center for Advanced Systems Understanding (CASUS) which is financed by Germany's Federal Ministry of Education and Research (BMBF) and by the Saxon state government out of the State budget approved by the Saxon State Parliament.

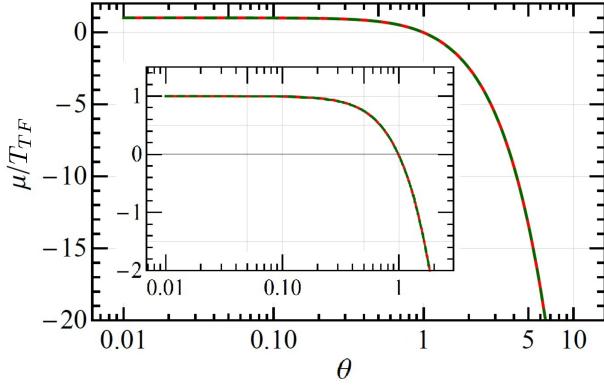


Figure 8. Chemical potential inversion according to the numerical solution of the inverse problem in Eq. (A5) (red) and obtained from the approximation in Eq. (A6) (green).

### Appendix A: Finite-Temperature Thomas-Fermi Free Energy

The TF approximation is particularly useful at high temperature. We therefore provide the theoretical basis of the finite-temperature TF free energy functional, which we have implemented in the OF-DFT code DFTpy [17].

In general, the noninteracting free energy and the free energy density are defined as

$$F_s[n] \equiv T_s[n] - k_B\tau S_s[n] = \int d\mathbf{r} f_s[n; \tau](\mathbf{r}). \quad (\text{A1})$$

$$\mu[n(\vec{r}); \tau] = \begin{cases} \alpha n(\vec{r})^{2/3} \left( 1 + \sum_i^4 a_i \times \left( \frac{\tau}{\alpha n(\vec{r})^{2/3}} \right)^i \right) & \text{if } \theta = \frac{\tau}{\alpha n(\vec{r})^{2/3}} < 1.36 \\ \tau \ln \left( C \left( \frac{\tau}{\alpha n(\vec{r})^{2/3}} \right)^{-3/2} \right) + \tau \ln \left( 1 + C \left( 2 \frac{\tau}{\alpha n(\vec{r})^{2/3}} \right)^{-3/2} \right) & \text{if } \theta = \frac{\tau}{\alpha n(\vec{r})^{2/3}} \geq 1.36 \end{cases} \quad (\text{A6})$$

where  $\alpha = \frac{(3\pi^2)^{2/3}}{2}$ ,  $a_1 = 0.016$ ,  $a_2 = -0.957$ ,  $a_3 = -0.293$ ,  $a_4 = 0.209$ , and  $C = \frac{2}{3\Gamma(3/2)} = 0.752252778063675$ .

Note that the dimensionless potential  $v_s/(\alpha n(\vec{r})^{2/3}) = \mu/(\alpha n(\vec{r})^{2/3})$  depends only on the parameter  $\theta[n(\vec{r}); \tau] = \frac{\tau}{\alpha n(\vec{r})^{2/3}}$ , the local degeneracy parameter. This dependence is shown in Fig. (8), where we compare the exact solution of the inverse problem Eq.(A5) for  $\mu$  with the approximation in Eq. (A6). This comparison demonstrates the high accuracy of the approximation in Eq. (A6).

We also observe in Fig. 8, that in the limit of low temperatures  $\theta = \frac{\tau}{\alpha n(\vec{r})^{2/3}} \ll 1$ , Eq. (A6) follows the TF model

$$v_s = \mu[n(\vec{r})] = T_{TF}[n] = \alpha n(\vec{r})^{2/3}. \quad (\text{A7})$$

At high temperatures,  $\theta > 1$ , the TF potential yields

In TF theory, we have  $f_s \equiv f_{TF}$ , where  $f_{TF}$  is the Thomas-Fermi free energy density [76, 77] defined as

$$f_{TF}([n], \theta) = \frac{\sqrt{2}m^{3/2}}{\hbar^3 \pi^2 \beta^{5/2}} \left( \eta I_{1/2}(\eta) - \frac{2}{3} I_{3/2}(\eta) \right), \quad (\text{A2})$$

where  $I_\nu$  is the Fermi integral of order  $\nu$ , and  $\eta = \mu\beta$  is a constant following from the normalization  $N = \int n(\vec{r})d\vec{r}$  and  $\beta = 1/(k_B\tau)$ . The Fermi-Dirac integral of order  $\nu$  is defined as

$$I_\nu = \int_0^\infty dx \frac{x^\nu}{(1 + \exp(x - \eta))}. \quad (\text{A3})$$

The functional derivative of the free energy with respect to the electron density yields [78]

$$v_s[n(\vec{r}); \tau] = \frac{\delta F_{TF}}{\delta n} = \mu[n(\vec{r}); \tau]. \quad (\text{A4})$$

In order to evaluate the effective potential in Eq. (A4), we need to find the chemical potential of the free, non-interacting electron gas at finite temperature which is formally given by the inverse of [79]

$$\frac{2}{3} (\theta[n(\vec{r})])^{-3/2} = \int_0^\infty dx \frac{\sqrt{x}}{(1 + \exp(x - \eta))}. \quad (\text{A5})$$

From Eq. (A5), we see that  $\eta[n] = \beta\mu[n]$  is defined by the degeneracy parameter given by  $\theta[n] = k_B\tau/T_{TF}[n]$ . To compute  $\mu[n(\vec{r}); \tau]$ , one can use the following expression (in the Hartree atomic units) as the solution of the inverse problem Eq. (A5) [80]:

negative values. This can lead to numerical instabilities in the OF-DFT scheme, where the total free energy is minimized. This problem is solved by introducing a gradient correction which is strictly positive and blocks the emergence of large density gradients.

For the simulation of Hydrogen at high temperatures, we used the von Weizsäcker functional defined in Eq. (10) as a gradient correction. One can use different versions of a functional representing first order gradient correction to the local density approximation (e.g., see Refs [77, 78, 81–83]). However, at the considered high temperatures, both the ion-electron and ion-ion coupling are weak and the ion-ion interaction is not so sensitive to the corrections beyond the LDA [84–87]. In this case, the role of a gradient correction with OF-DFT merely amounts as a numerical trick for stabilizing the numerics.

In the simulations involving non-zero temperatures the

von Weizsäcker and nonlocal terms are borrowed from the corresponding zero temperature expressions.

## Appendix B: Equilibration curves for Beryllium

In the following, we provide the full equilibration curves for the Beryllium based experiments presented above, to confirm the correct application of the algorithm outlined in Sec. II C and Fig. 2. The given figures follow Fig. 2 in their general outline, i.e. the smoothed cosine distance between configurations and reference configurations is shown, along with equilibration thresholds and first equilibrated configuration. The equilibration behavior for Beryllium is shown in Fig. 9 for DFT-MD simulations starting from the ideal crystal structure and in Fig. 10 for those started from OF-DFT-MD configurations. In Fig. 9 the gradual equilibration process is neatly pronounced, it can be seen how the system is slowly thermalized, and how thermalization requires more and more time as one moves to larger temperatures. Please note that raw numeric values of  $d_C^a$  should not be compared between individual systems, as they are relative quantities. Absolute values of these metrics differ across system sizes since the small deviations in the RDFs become increasingly small relative to the overall RDF. This however does not negatively affect performance, as only relative values are compared. Equilibration trends are notably less pronounced in Fig. 9 as the initial configurations are already very close to the equilibrated ones. A slight equilibration can be seen for 1024 atoms, where the first few hundred time-steps are mostly above the equilibration threshold. Generally however, the OF-DFT-MD trajectories are very close to equilibrium to begin with, which is reflected in the similar (and rather small) number of time-steps required for equilibration throughout. The distance metrics essentially fluctuate around the average almost from the beginning of the simulation. Fig. 11 corroborates that both methods of initialization lead to equilibrated trajectories that fluctuate around the same average.

## Appendix C: Equilibration curves for Hydrogen

Similar to App. B, here we provide the full equilibration curves for the Hydrogen based experiments presented in Sec. III C. The equilibration behavior for Hydrogen is shown in Fig. 12 for DFT-MD simulations starting from the ideal crystal structure and in Fig. 13 for those starting from OF-DFT-MD simulations. The overall equilibration behavior is similar to those observed in the preceding section. In contrast to the results for Beryllium, it is noticeable that the trajectories appear to be noisier, even with additional smoothing, obfuscating the equilibration process to a degree, especially visible in Fig. 13. While this in principle *can* lead to a slight overestimation of the overall equilibration phase, it does

not infringe on our analysis, as the same parameters are employed to analyze both the ideal crystal and OF-DFT-MD initialized trajectories. Furthermore we minimize such effects by increasing  $N_T$  and  $\sigma$  for larger  $\tau$ . Generally, the Hydrogen trajectories equilibrate faster, which is mostly due to the smaller number of atoms. Fig. 14 shows that either method of initialization converges to the same fluctuating average.

## Appendix D: Influence of Nosé mass

As the performance of DFT-MD simulations that employ the Nosé-Hoover thermostat depend on  $Q$  we have to verify the applicability of the results presented in Fig. 6 across the range of applicable  $Q$ . More precisely, for Beryllium, we determine the outer boundaries of values for  $Q$  for which the KS-DFT-MD trajectories do not diverge (i.e., do not find equilibrium in the NVT ensemble or produce temperature oscillations) as  $Q = [0.1, 0.005]$ . Comparing to the boundaries of this interval,  $Q = 0.01$ , which has been used for the simulations shown in Fig. 6, gives reasonably good performance for the equilibration. This comparison is shown in Fig. 15, and illustrates that our choice of  $Q = 0.01$  reflects standard production quality. It further shows that OF-DFT-MD reduces computational time needed to equilibrate a system across different values of  $Q$ . This trend can be expected to hold true for other thermostats as well.

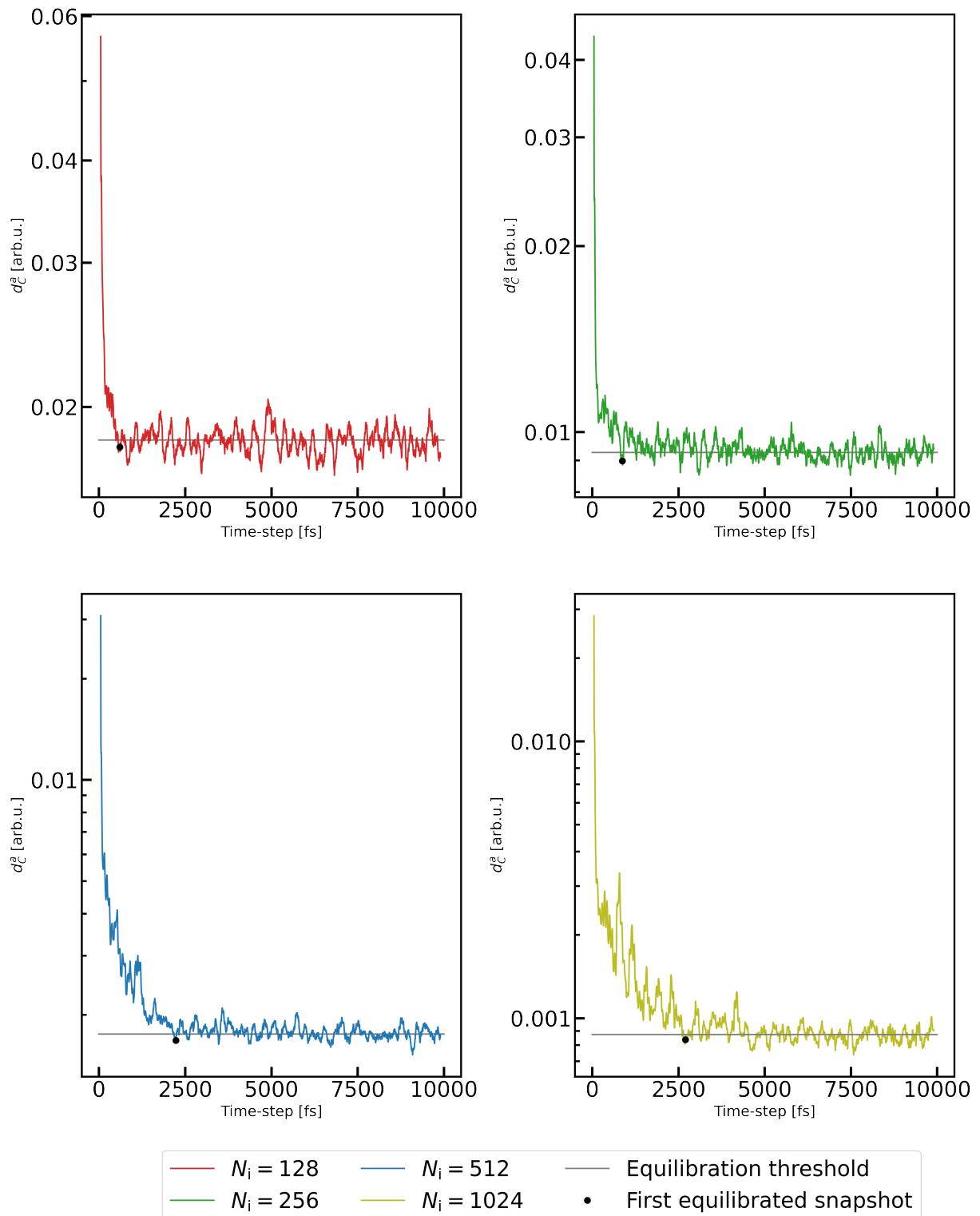


Figure 9. Equilibration curves for Beryllium, starting from the ideal crystal structure.

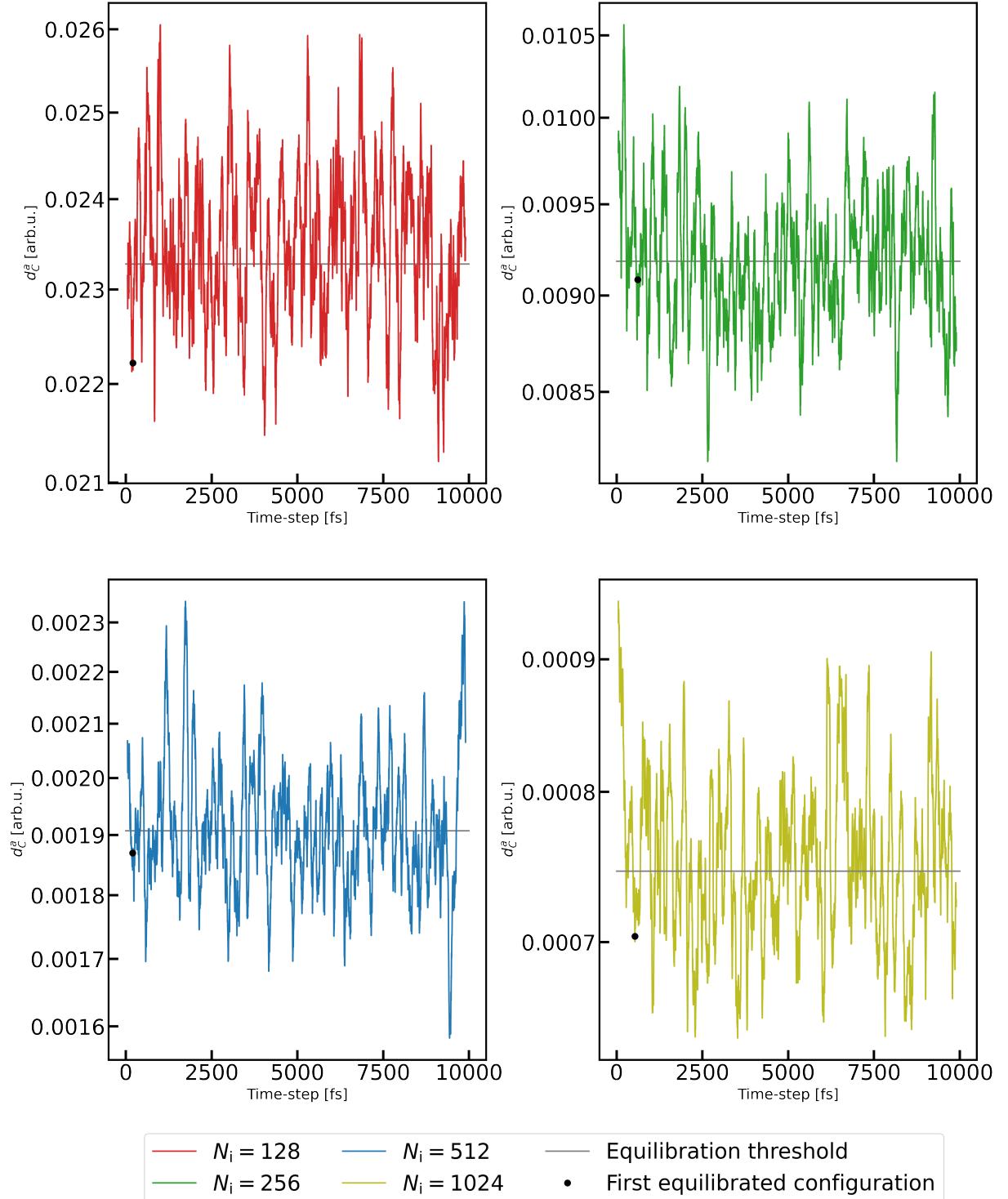


Figure 10. Equilibration curves for Beryllium, starting from the OF-DFT-MD structure.

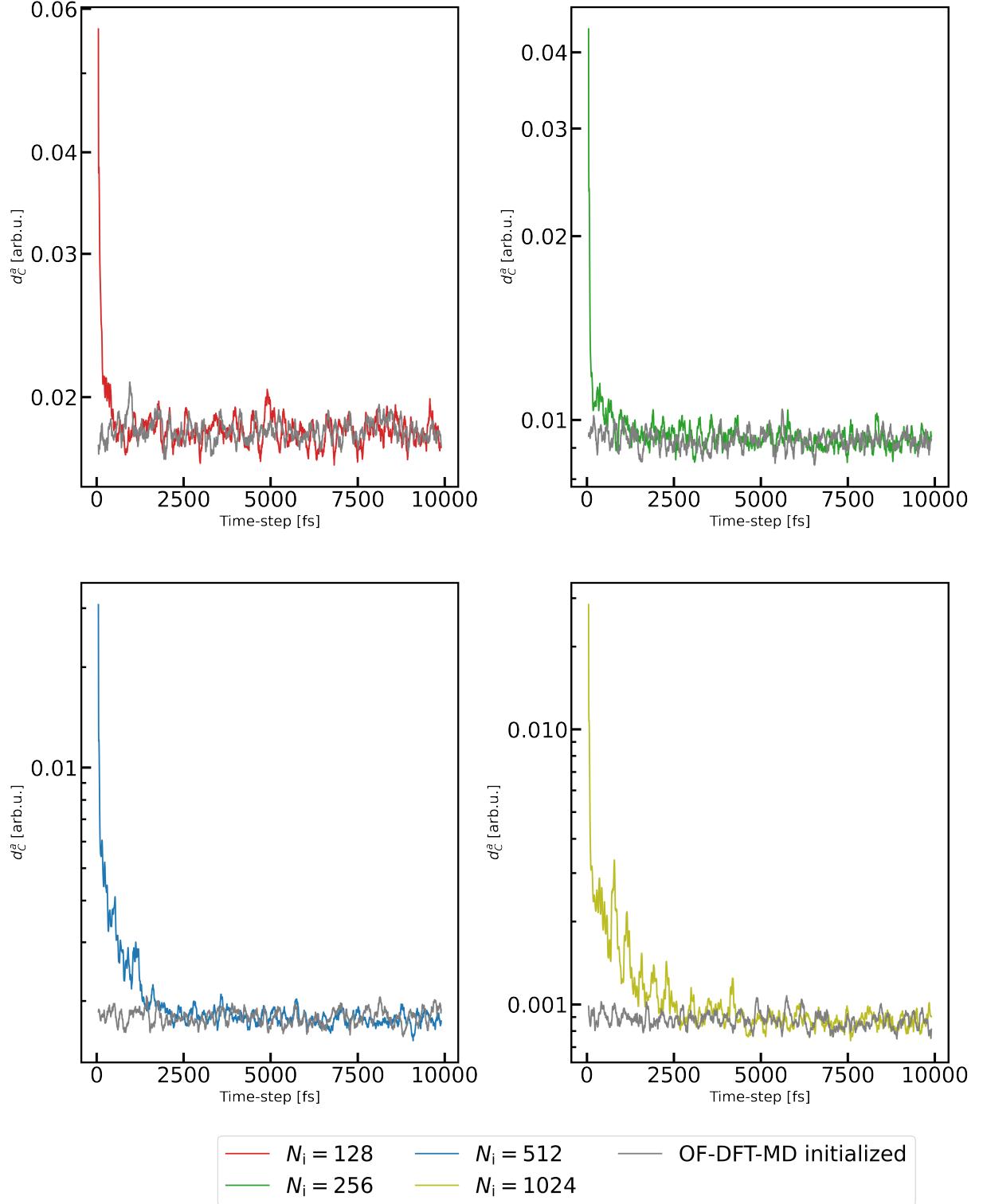


Figure 11. Equilibration curves for Beryllium, with the OF-DFT-MD initialized trajectories superimposed over those starting from ideal crystal structure. Please note that to be consistent, both distance metrics were calculated with the reference configuration of the ideal crystal structure trajectory, and thus, the OF-DFT-MD initialized metrics differ slightly from those shown in Fig. 10.

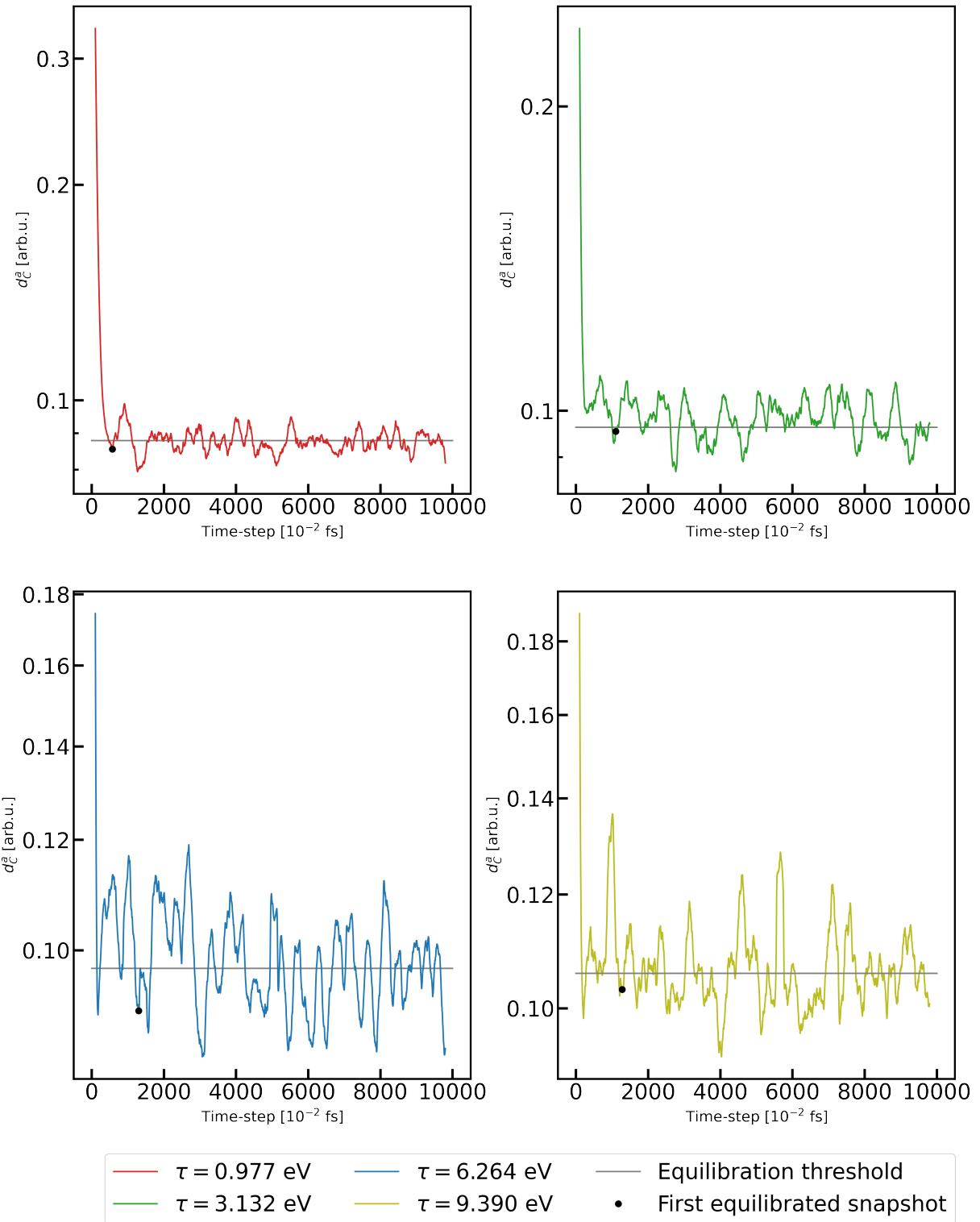


Figure 12. Equilibration curves for Hydrogen, starting from the ideal crystal structure.

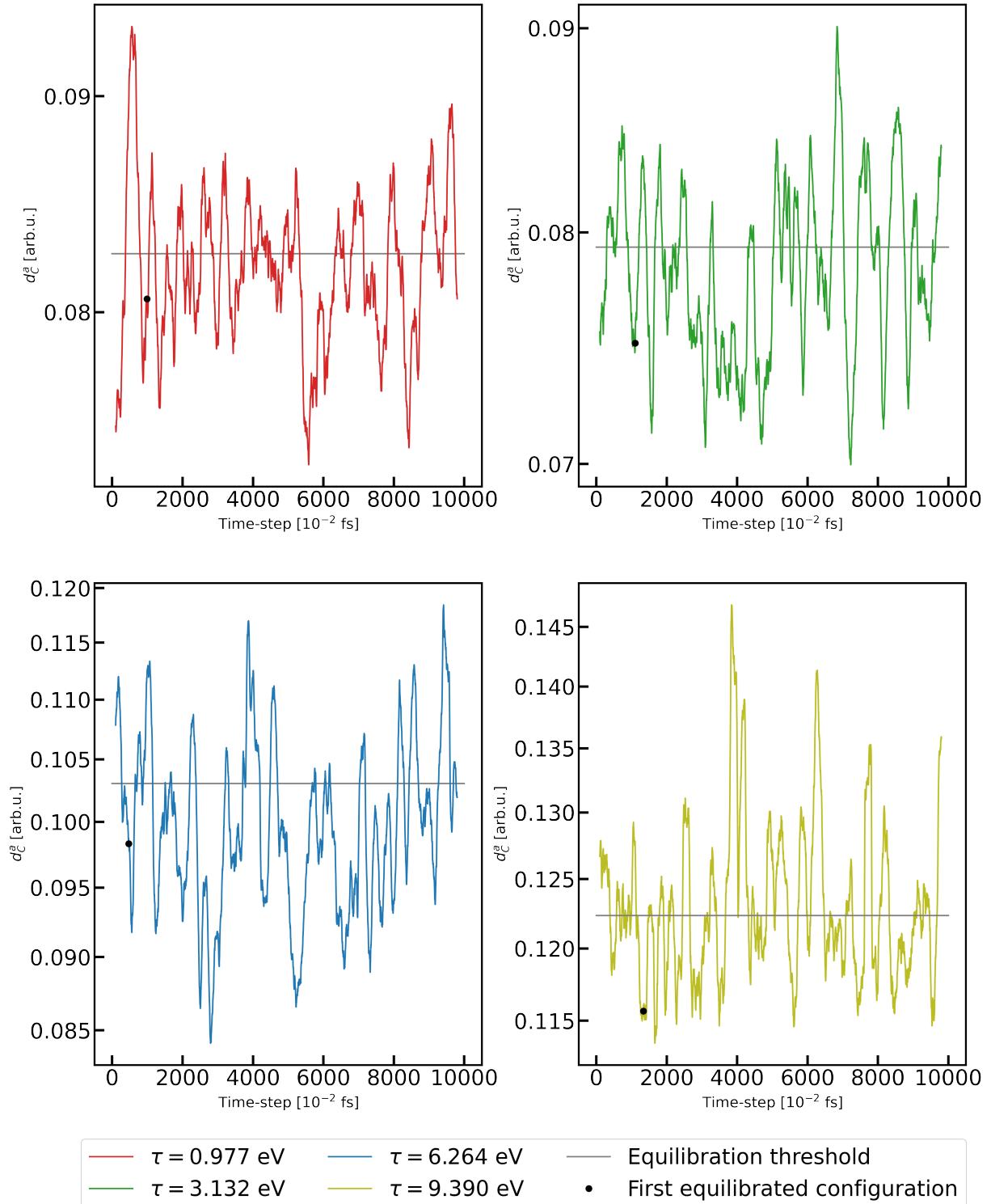


Figure 13. Equilibration curves for Hydrogen, starting from the OF-DFT-MD structure.

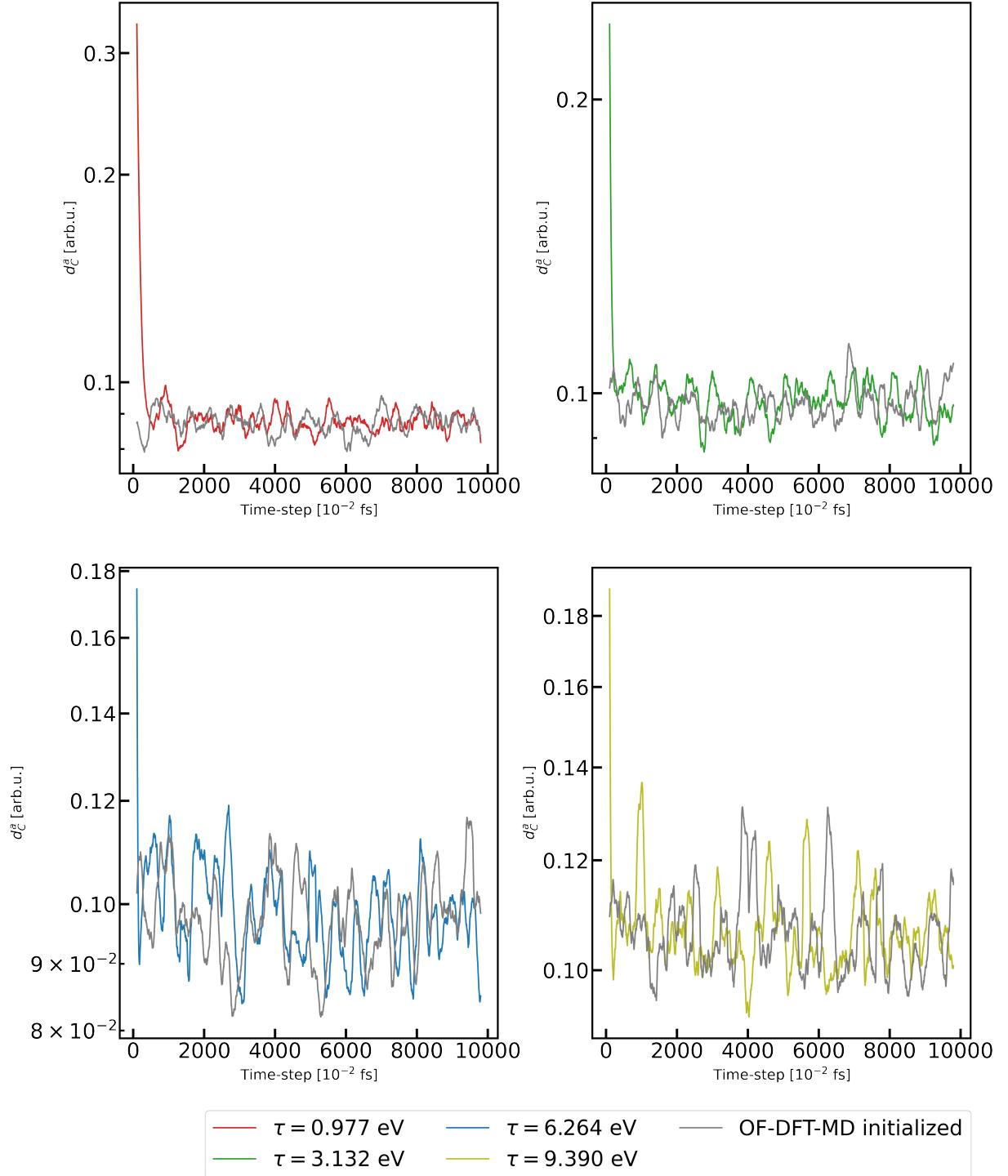


Figure 14. Equilibration curves for Hydrogen, with the OF-DFT-MD initialized trajectories superimposed over those starting from ideal crystal structure. Please note that to be consistent, both distance metrics were calculated with the reference configuration of the ideal crystal structure trajectory, and thus, the OF-DFT-MD initialized metrics differ slightly from those shown in Fig. 13.

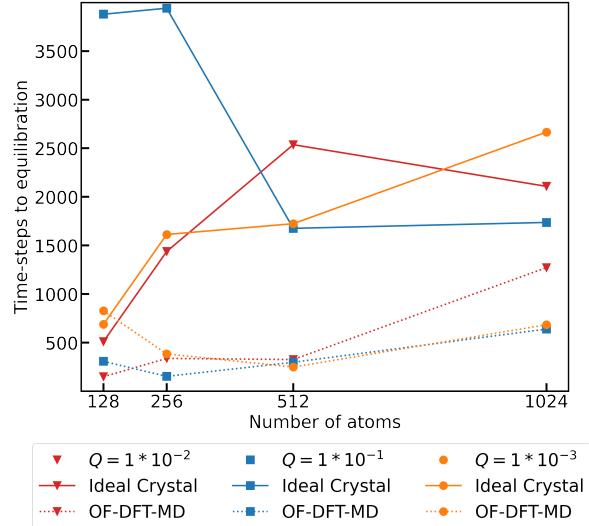


Figure 15. Comparison of MD runs with different  $Q$ . In contrast to Fig 6, all MD trajectories were run for 5000 time-steps of 1 fs for computational feasibility. Please note that the less systematic behavior of the  $Q = 0.1$  and  $Q = 0.005$  trajectories can be explained by the trajectories not being fully converged after 5000 time-steps, and that the results for  $Q = 0.01$  vary slightly compared to Fig. 6, since in Fig. 6 the full 10000 time-steps were analyzed. Here only 5000 time-steps are analyzed, in order to be consistent with the other trajectories.

- 
- \* l.fiedler@hzdr.de  
 † z.moldabekov@hzdr.de  
 ‡ xuecheng.shao@rutgers.edu  
 § kaili.jiang@rutgers.edu  
 ¶ t.dornheim@hzdr.de  
 \*\* m.pavanello@rutgers.edu  
 †† a.cangi@hzdr.de
- [1] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).  
 [2] P. Verma and D. G. Truhlar, *Trends Chem.* **2**, 302 (2020).  
 [3] M. Bonitz, T. Dornheim, Z. A. Moldabekov, S. Zhang, P. Hamann, H. Kähler, *et al.*, *Phys. Plasmas* **27**, 042710 (2020).  
 [4] V. E. Fortov, *Phys.-Usp.* **52**, 615 (2009).  
 [5] P. K. Shukla and B. Eliasson, *Rev. Mod. Phys.* **83**, 885 (2011).  
 [6] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).  
 [7] S. Zhang, H. Wang, W. Kang, P. Zhang, and X. T. He, *Phys. Plasmas* **23**, 042707 (2016).  
 [8] P. Suryanarayana, P. P. Pratapa, A. Sharma, and J. E. Pask, *Comput. Phys. Commun.* **224**, 288 (2018).  
 [9] Y. Cytter, E. Rabani, D. Neuhauser, and R. Baer, *Phys. Rev. B* **97**, 115207 (2018).  
 [10] A. Nakata, J. S. Baker, S. Y. Mujahed, J. T. L. Poult, S. Arapan, J. Lin, *et al.*, *J. Chem. Phys.* **152**, 164112 (2020).  
 [11] E. Hernández and M. J. Gillan, *Phys. Rev. B* **51**, 10157 (1995).  
 [12] J. VandeVondele, U. Borštník, and J. Hutter, *J. Chem. Theory Comput.* **8**, 3565 (2012).  
 [13] S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker, *et al.*, *Phys. Chem. Chem. Phys.* **17**, 31360 (2015).  
 [14] F. Mauri and G. Galli, *Phys. Rev. B* **50**, 4316 (1994).  
 [15] R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960).  
 [16] V. L. Lignères and E. A. Carter, in *Handbook of Materials Modeling: Methods* (Springer Netherlands, Dordrecht, 2005) pp. 137–148.  
 [17] X. Shao, K. Jiang, W. Mi, A. Genova, and M. Pavanello, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **11**, e1482 (2021).  
 [18] J. P. Perdew and L. A. Constantin, *Phys. Rev. B* **75**, 155109 (2007).  
 [19] S. Laricchia, L. A. Constantin, E. Fabiano, and F. Della Sala, *J. Chem. Theory Comput.* **10**, 164 (2014).  
 [20] W. C. Witt and E. A. Carter, *Phys. Rev. B* **100**, 125106 (2019).  
 [21] W. C. Witt and E. A. Carter, *Phys. Rev. B* **100**, 125107 (2019).  
 [22] V. V. Karasiev, D. Chakraborty, O. A. Shukruto, and S. B. Trickey, *Phys. Rev. B* **88**, 161108 (2013).  
 [23] L. Fiedler, K. Shah, M. Bussmann, and A. Cangi, *Phys. Rev. Materials* **6**, 040301 (2022).  
 [24] J. A. Ellis, L. Fiedler, G. A. Popoola, N. A. Modine, J. A. Stephens, A. P. Thompson, *et al.*, *Phys. Rev. B* **104**, 035120 (2021).  
 [25] M. Born and R. Oppenheimer, *Ann. Phys.* **389**, 457 (1927).  
 [26] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980).  
 [27] J. P. Perdew and W. Yue, *Phys. Rev. B* **33**, 8800 (1986).  
 [28] J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**, 13244 (1992).  
 [29] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).  
 [30] J. Sun, A. Ruzsinszky, and J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).  
 [31] A. Cangi, D. Lee, P. Elliott, K. Burke, and E. K. U. Gross, *Phys. Rev. Lett.* **106**, 236404 (2011).  
 [32] T. A. Wesolowski and Y. A. Wang, *Recent Progress in Orbital-free Density Functional Theory*, Recent Advances in Computational Chemistry, Vol. 6 (WORLD SCIENTIFIC, Singapore, 2012).  
 [33] V. V. Karasiev, D. Chakraborty, and S. B. Trickey, in *Many-Electron Approaches in Physics, Chemistry and Mathematics: A Multidisciplinary View* (Springer International Publishing, Cham, 2014) pp. 113–134.  
 [34] L. H. Thomas, *Math. Proc. Camb. Philos. Soc.* **23**, 542–548 (1927).  
 [35] E. Fermi, *Z. Physik* **48**, 73 (1928).  
 [36] C. v. Weizsäcker, *Z. Physik* **96**, 431 (1935).  
 [37] L.-W. Wang and M. P. Teter, *Phys. Rev. B* **45**, 13196 (1992).  
 [38] W. Mi, A. Genova, and M. Pavanello, *Chem. Phys.* **148**, 184107 (2018).  
 [39] Y. A. Wang, N. Govind, and E. A. Carter, *Phys. Rev. B* **58**, 13465 (1998).  
 [40] N. D. Mermin, *Phys. Rev.* **137**, A1441 (1965).  
 [41] S. Pittalis, C. R. Proetto, A. Floris, A. Sanna, C. Bersier, K. Burke, *et al.*, *Phys. Rev. Lett.* **107**, 163001 (2011).  
 [42] A. Pribram-Jones, S. Pittalis, E. K. U. Gross, and K. Burke, in *Lecture Notes in Computational Science and Engineering*, Vol. 96 (Springer International Publishing, Cham, 2014) pp. 25–60.  
 [43] V. V. Karasiev, T. Sjostrom, and S. Trickey, *Phys. Rev. B* **86**, 115101 (2012).  
 [44] J. Bartel, M. Brack, and M. Durand, *Nucl. Phys. A* **445**, 263 (1985).  
 [45] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, *J. Comput. Phys.* **285**, 316 (2015).  
 [46] K. T. Schütt, H. E. Saucedo, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, *Chem. Phys.* **148**, 241722 (2018).  
 [47] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Phys. Rev. Lett.* **104**, 136403 (2010).  
 [48] W. G. Hoover, *Phys. Rev. A* **31**, 1695 (1985).  
 [49] S. Nosé, *Chem. Phys.* **81**, 511 (1984).  
 [50] D. J. Evans and B. L. Holian, *Chem. Phys.* **83**, 4069 (1985).  
 [51] N. H. Siboni, *Molecular dynamics studies of thermodynamical consistency and non-locality of effective temperature*, Ph.D. thesis, Aachen, Techn. Hochsch., 2014 (2016).  
 [52] A. Guy and A. Bowling, *J. Comput. Nonlinear Dyn.* **12** (2017).  
 [53] B. J. Alder, S. P. Frankel, and V. A. Lewinson, *Chem. Phys.* **23**, 417 (1955).  
 [54] K. Zahn, G. Maret, C. Ruf, and H.-H. von Grünberg, *Phys. Rev. Lett.* **91**, 115502 (2003).  
 [55] M. Gottschalk, *AIP Adv.* **11**, 045026 (2021).  
 [56] A. Singhal *et al.*, *IEEE Data Eng. Bull.* **24**, 35 (2001).  
 [57] A. Cangi, J. A. Ellis, L. Fiedler, D. Kotik, N. A. Modine, *et al.*, “MALA,” (2021).

- [58] L. Fiedler, S. Schmerler, N. Modine, G. Popoola, J. Vogel, *et al.*, "In preparation,".
- [59] P. Haley and D. Soloway, in *IJCNN International Joint Conference on Neural Networks*, Vol. 4 (1992) pp. 25–30.
- [60] W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graph.* **14**, 33 (1996).
- [61] J. Stone, *An Efficient Library for Parallel Ray Tracing and Animation*, Master's thesis, Computer Science Department, University of Missouri-Rolla (1998).
- [62] G. Kresse and J. Hafner, *Phys. Rev. B* **47**, 558 (1993).
- [63] G. Kresse and J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- [64] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- [65] M. Hacene, A. Anciaux-Sedrakian, X. Rozanska, D. Klahr, T. Guignon, and P. Fleurat-Lessard, *J. Comput. Chem.* **33**, 2581 (2012).
- [66] M. Hutchinson and M. Widom, *Comput. Phys. Commun.* **183**, 1422 (2012).
- [67] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
- [68] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
- [69] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, *et al.*, *J. Chem. Phys.* **152**, 154105 (2020).
- [70] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, *et al.*, *J. Phys.: Condens. Matter* **21**, 395502 (2009).
- [71] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, *et al.*, *J. Phys. Condens. Matter* **29**, 465901 (2017).
- [72] L.-W. Wang and M. P. Teter, *Phys. Rev. B* **45**, 13196 (1992).
- [73] W. Mi, S. Zhang, Y. Wang, Y. Ma, and M. Miao, *J. Chem. Phys.* **144**, 134108 (2016).
- [74] L. Fiedler, Z. Moldabekov, X. Shao, K. Jiang, T. Dornheim, *et al.*, "Data and scripts for "accelerating equilibration in first-principles molecular dynamics with orbital-free density functional theory"," (2022).
- [75] P. R. C. Kent, R. Q. Hood, A. J. Williamson, R. J. Needs, W. M. C. Foulkes, *et al.*, *Phys. Rev. B* **59**, 1917 (1999).
- [76] L. G. Stanton and M. S. Murillo, *Phys. Rev. E* **91**, 033104 (2015).
- [77] F. Perrot, *Phys. Rev. A* **20**, 586 (1979).
- [78] Z. A. Moldabekov, M. Bonitz, and T. S. Ramazanov, *Phys. Plasmas* **25**, 031903 (2018).
- [79] N. R. Arista and W. Brandt, *Phys. Rev. A* **29**, 1471 (1984).
- [80] A. G. Sotnikov, K. V. Sereda, and Y. V. Slyusarenko, *Low Temp. Phys.* **43**, 144 (2017).
- [81] W. Jones and W. H. Young, *J. Phys. C: Solid State Phys.* **4**, 1322 (1971).
- [82] D. A. Kirzhnits, Y. E. Lozovik, and G. V. Shpatakovskaya, *Sov. Phys. Usp.* **18**, 649 (1975).
- [83] Z. Moldabekov, T. Schoof, P. Ludwig, M. Bonitz, and T. Ramazanov, *Phys. Plasmas* **22**, 102104 (2015).
- [84] Z. A. Moldabekov, S. Groth, T. Dornheim, H. Kählert, M. Bonitz, and T. S. Ramazanov, *Phys. Rev. E* **98**, 023207 (2018).
- [85] Z. A. Moldabekov, H. Kählert, T. Dornheim, S. Groth, M. Bonitz, and T. S. Ramazanov, *Phys. Rev. E* **99**, 053203 (2019).
- [86] Z. A. Moldabekov, T. Dornheim, and M. Bonitz, *Contrib. to Plasma Phys.* **62**, e202000176 (2022).
- [87] Z. Moldabekov, S. Groth, T. Dornheim, M. Bonitz, and T. Ramazanov, *Contrib. to Plasma Phys.* **57**, 532 (2017).