Article

Reducing the Cost of Neural Network Potential Generation for Reactive Molecular Systems

Krystof Brezina, Hubert Beck, and Ondrej Marsalek*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Although machine learning potentials have recently had a substantial impact on molecular simulations, the construction of a robust training set can still become a limiting factor, especially due to the requirement of a reference ab initio simulation that covers all the relevant geometries of the system. Recognizing that this can be prohibitive for certain systems, we develop the method of transition tube sampling that mitigates the computational cost of training set and model generation. In this approach, we generate classical or quantum thermal geometries around a transition path describing a conformational change or a chemical reaction using only a sparse set of local normal mode expansions along this path and select from these geometries by an active learning protocol. This yields a training set with geometries that characterize the whole transition without the need for a costly reference trajectory. The performance of the method is evaluated on different molecular systems with the complexity of the potential energy landscape increasing from a single minimum to a double proton-transfer reaction with high barriers. Our results show that the method leads to training sets that give rise to models applicable in classical and path integral simulations alike that are on par with those based directly on ab initio calculations while providing the computational speedup we have come to expect from machine learning potentials.

$\delta_1 = d_{N_1'H} - d_{N_1H}$
$\delta_2 = d_{N_2'H} - d_{N_2H}$
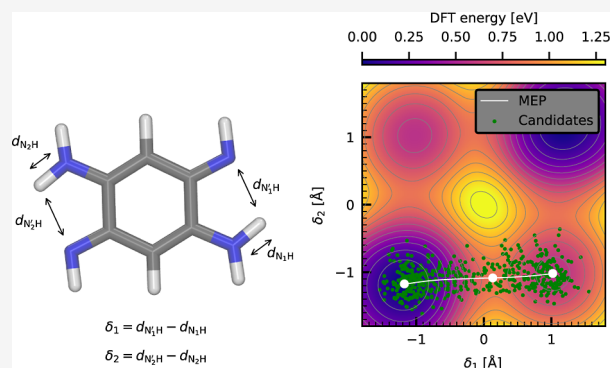
1. INTRODUCTION

Owing to the detailed atomistic insight into the structure and dynamics of molecular systems and materials, the relevance of computer simulations of molecular dynamics (MD) in current research is undeniable. MD simulations represent a valuable analytic and predictive tool in multiple fields of both basic and applied research including physical chemistry, materials science, or drug design.[1−5] They also provide a way to explain and corroborate experimental data that might be difficult to interpret otherwise. For many systems of interest, MD simulations can be routinely performed under the Born–Oppenheimer approximation in the electronic ground state, with the nuclei being treated either classically or quantum-mechanically within the imaginary time path integral formalism. This makes the choice of the potential energy surface (PES) a key decision that determines the accuracy of the resulting simulation. Out of the available options, ab initio molecular dynamics[6] (AIMD) is a state-of-the-art methodology that relies on a full, on-the-fly quantum electronic structure calculation[7,8] at every step of the simulation to evaluate the potential energy and forces. This is most commonly performed at the level of density functional theory[8−10] (DFT), which provides correlated electronic energies at a computational cost accessible in practice, but for smaller systems, the use of correlated wave function methods is feasible as well.[11−13] In any case, the computational

cost of AIMD simulations is typically large—especially so for advanced hybrid DFT functionals in the condensed phase—and can easily become prohibitive in the light of the ever-growing demand for larger time and length scales of the relevant simulations.

This issue can be mitigated by the recent development of the so-called machine learning potentials (MLPs).[14,15] These use various machine learning approaches[16−20] to faithfully approximate the desired ab initio PES by training on a reference data set consisting of a relatively modest number of ab initio geometries and their corresponding energies and, optionally, forces. As such, they indeed combine the best of the two worlds: they are able to maintain the accuracy of the parent ab initio method, but they also circumvent the need for explicit electronic structure calculations at each step of the MD simulation. Thus, they evaluate the potential energy and forces at a significantly reduced computational cost.[21] One particular flavor of MLPs of major practical importance is represented by
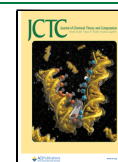
neural network potentials (NNPs), which rely on artificial neural networks combined with a set of appropriate atomic descriptors to accurately represent the molecular geometry-to-energy relationship, including all its symmetries.[16,22] NNPs have repeatedly proved their worth in modeling a plethora of various molecular systems ranging from liquids and solutions to interfaces and solids.[20,23−26] Our recent study,[26,27] building on the findings of previous studies focusing on NNPs,[28−30] shows that rather than using a single NNP to represent the PES, it is advantageous to build a model as a committee[31,32] of NNPs (C-NNP) that comprises a small number of NNPs, each trained individually to a subset of the main training set. The advantage is twofold: first, the energy prediction obtained as the committee average is known to be a better approximation of the ab initio energy than the estimates of the individual members.[29,33] Second, the committee disagreement,[34] represented by the standard deviation of the individual member estimates of energies or forces, serves as a valuable indicator of prediction reliability and can be used to monitor and optionally ensure the stability of the simulation.[27] Crucially, this disagreement can be used as the key ingredient of the active learning process called query by committee[35] (QbC) that systematically builds the training set in a data-driven way.[27,34,36]

An accurate and stable NNP can only be obtained on a foundation of robust, high-quality training data. This is typically based on a reference AIMD trajectory, from which geometries are selected for the training set, together with the corresponding energies and forces. However, the trajectory is highly correlated in time and thus most of the expensive AIMD data do not contribute useful information for the training of the model. This selection has been approached in different ways from random sampling and manual selection to more data-driven procedures,[25,27,30,37−43] with QbC being a particularly efficient method. QbC considers a set of candidate structures, in this case, the whole AIMD trajectory, and iteratively builds up the training set. It starts by training a C-NNP on a very small set of initial configurations and using its disagreement to screen the candidate configurations for those with the most uncertain prediction. A small number of these configurations are then added to the training set, a new C-NNP is trained, and the process is iteratively repeated until some convergence criteria are met. In comparison to random selection, this approach is known to generate more compact training sets that give rise to robust models of similar accuracy.[30,39] Even though the initial AIMD trajectory is typically the most expensive part of the procedure, numerous successful MLPs have been generated on top of reasonably short AIMD simulations.[26]

However, for many purposes, this process involving AIMD is still too expensive to be practical. For instance, the requirements on a high-level electronic structure method can raise the computational demands above a reasonable threshold. One might also be interested in a system that features rare events, such as chemical or conformational changes, which will happen quickly and occur infrequently or not at all in a direct AIMD simulation. In turn, these crucial configurations are under-represented in the set of candidates and enhanced sampling simulations would be required in order to construct a robust training set, which typically raises the computational cost further by one or more orders of magnitude.

In case such a situation occurs, one needs to adhere to an approximate method of candidate generation that relieves some of the computational expenses while maintaining the quality of the resulting candidate set. For simple systems with a single potential energy minimum, the solution is fairly straightforward. In this case, one can benefit from a random sampling of displacements in the directions of a fixed set of normal modes to obtain a set of distorted configurations. This approach, sometimes called normal mode sampling (NMS) in the literature,[18] avoids the cost of a full AIMD simulation by replacing it with a more manageable combination of a Hessian matrix evaluation and a number of single-point electronic structure calculations for the generated geometries. The sampling of the known normal mode distribution itself yields uncorrelated samples by definition and requires no ab initio calculations; therefore, its cost is negligible. Various versions of this approach were successfully used to generate structures for the training of MLPs. Using a scaled uniform random sampling of the normal modes, the method was used to obtain auxiliary structures used in model validation[44] and with approximate thermal distortions in NNP training set generation around configurational minima[18] as well as to construct an NNP model for a gas-phase ammonia molecule.[42] Clearly, the utility of NMS is limited when the harmonic approximation becomes insufficient. This can be the case if individual modes are strongly anharmonic or coupled, or if the system features conformational changes or reactions, where multiple local minima come into play. The need for reactive training data sets was recognized in a recent work introducing the Transition-1x data set,[45] which includes training points along a converged minimum energy path (MEP) obtained through a nudged elastic band calculation[46] and its surrounding arising from prior unconverged iterations of the optimization.

In this work, we propose transition tube sampling (TTS), a robust and general approach to the generation of training sets and models that are able to accurately describe processes that feature transitions over potential energy barriers, which includes both conformational flexibility and chemical reactivity. We achieve this by generating thermally distorted candidate geometries along a reaction pathway with the help of multiple normal mode expansions and screening these candidates using QbC. The role of the minimum geometry in NMS is taken by the MEP that describes the course of the reaction through configuration space. Local harmonic expansions are performed in a small number of relevant configurations along the MEP and physically relevant candidate configurations are generated with uniform distribution along the MEP and with classical or quantum thermal weights in all perpendicular directions based on one of the sets of normal modes. An arbitrary number of these candidate configurations can be generated at a negligible computational cost and submitted to the QbC process, which selects the most important ones to have ab initio calculations performed and to be included in the training set. This results in compact and robust training sets and models that maintain consistent accuracy along the reaction path, making them suitable for MD simulations of the reactive process, including enhanced sampling simulations, while no AIMD trajectories are required as part of this process. We test this method on three different molecules in the gas phase to illustrate its capabilities.

The rest of the paper is organized as follows. In Section 2, we begin by formalizing thermal NMS, which samples the exact classical or quantum canonical distribution under the harmonic approximation. With the obtained framework, we then proceed to introduce the MEP into the picture and

describe the technical details of TTS. In Section 3, we describe how we used TTS to create C-NNP models, the simulations performed with these models, and other related computational details. In Section 4, we apply this approach to three different gas-phase systems of increasing complexity represented by the molecules of benzene, malonaldehyde, and 2,5-diaminobenzo-quinone-1,4-diimine (DABQDI) and discuss its successes and possible pitfalls. Section 5 concludes the paper and offers outlooks concerning the generalization and the limitations of the method beyond the gas phase.

2. THEORY

In this section, we discuss the theoretical basis of the TTS method. In this approach, we rely on the harmonic approximation and the vibrational normal mode formalism to obtain ab initio training data for the construction of C-NNPs for reactive systems without the need to run expensive sampling simulations, such as full AIMD. First, we present the simple key idea behind NMS which relies on the harmonic approximation to describe the underlying PES and thus is expected to work well for systems that are close to harmonic around a single given minimum geometry at the temperature of interest. Clearly, this does not yield a flexible and general method, since the harmonic approximation is readily challenged by many realistic systems, notably those that exhibit more pronounced configurational flexibility or chemical reactivity. Therefore, we propose a more general approach to sampling candidate geometries based on NMS which is applicable even to systems described by multiple minima separated by barriers. This is achieved using the harmonic expansion locally along an MEP in a way that eventually generates a balanced training set.

2.1. NMS for Thermal Sampling around Minimum Geometries. To open the discussion of the theory behind TTS, we first turn our attention to the simple case represented by a PES with a single minimum geometry $\mathbf{R}_0$ on which the nuclear motion is described by classical mechanics. Assuming a reasonable extent of validity of the harmonic approximation to capture the thermally accessible potential energy landscape, the classical thermal probability density ρ_c at temperature T is approximated by

$$\rho_c(\Omega_1, ..., \Omega_{N_{int}}) \propto \prod_{i=1}^{N_{int}} \exp\left(-\frac{1}{2}\beta\omega_i^2\Omega_i^2\right) \tag{1}$$

In this expression, ω_i and Ω_i denote, respectively, the natural frequency and the normal coordinate corresponding to the i-th normalized vibrational normal mode vector $\mathbf{\Omega}_i$, and β is the inverse temperature equal to $1/k_B T$ (with k_B representing the Boltzmann constant). N_{int} is the total number of internal degrees of freedom of the species, typically $3N - 6$ for N atoms. Hence, in the harmonic approximation, the thermal density is described as a multivariate, yet uncoupled normal distribution where each i-th orthogonal degree of freedom has the standard deviation of $\sigma_i = 1/\sqrt{\beta\omega_i^2}$.

As such, it is straightforward to generate completely uncorrelated thermal geometries $\mathbf{R}$ by distorting the minimum geometry $\mathbf{R}_0$ independently in the direction of each of the normal mode vectors. The appropriate magnitude of the distortions is given by a randomly generated value of the corresponding normal coordinate Ω_i from the distribution in eq 1. The instrumental prescription for this procedure is the inverse coordinate transformation from normal modes back to Cartesian coordinates

$$\mathbf{R} = \mathbf{R}_0 + \mathbb{M}^{-1/2}\sum_{i=1}^{N_{int}} \Omega_i\mathbf{\Omega}_i \tag{2}$$

where $\mathbb{M}$ represents the diagonal mass matrix. Thus, by drawing samples of normal coordinates and transforming them, we obtain correctly distributed thermal samples in Cartesian coordinates.

We can now perform thermal NMS by sampling from this auxiliary harmonic ensemble as a source of candidate geometries to be potentially included in the training set of an MLP. The auxiliary ensemble is thus never used directly and no expectation values are calculated over it. It only needs to provide good coverage of the thermally accessible region of the PES, which will be the case as long as the harmonic approximation is reasonably accurate for the system of interest. Specifically, we construct a training set in our active learning procedure by generating a large number of these NMS candidate geometries and screening them in a QbC process using a C-NNP model. In each QbC iteration, electronic structure calculations are performed only for a small number of selected structures to obtain their potential energies and possibly forces, which then comprise the final training set once the process converges. The computational cost is thus determined primarily by the geometry optimization procedure, the Hessian calculation, the C-NNP prediction required for screening, and the electronic structure calculations for the selected geometries. The cost of the sample generation is negligible. This approach is substantially less computationally demanding when compared to the more conventional approach of sampling the candidate geometries for QbC from an AIMD trajectory, which requires a large number of electronic structure calculations for very similar geometries that do not contribute diversity to the training set. In contrast to that, NMS generates fully decorrelated geometries by construction, and electronic structure calculations are only needed for the relatively small number of the most important geometries selected by the subsequent QbC process.

So far, we have focused on the situation where NMS is used to sample a classical distribution on the studied PES. However, since the harmonic approximation describes a molecule as a set of independent one-dimensional harmonic oscillators, we can readily generalize the above classical case to a quantum one as the analytic solution of the quantum harmonic oscillator is known. Specifically, it is straightforward to show (see Section S1 of the Supporting Information) that the canonical thermal density of a quantum harmonic oscillator at a given temperature is Gaussian just as its classical counterpart, but broader. This broadening is encoded in the quantum effective inverse temperature[47]

$$\beta^*(\beta, \omega) = \frac{2}{\hbar\omega}\tanh\left(\frac{\beta\hbar\omega}{2}\right) \tag{3}$$

at which a classical harmonic oscillator would have the same thermal width as a quantum harmonic oscillator at a reference inverse temperature β. Since β^* is by definition a frequency-dependent quantity, one cannot describe the whole molecule by a single quantum effective temperature but instead has to assign one to each individual mode. In turn, the quantum thermal density is given by
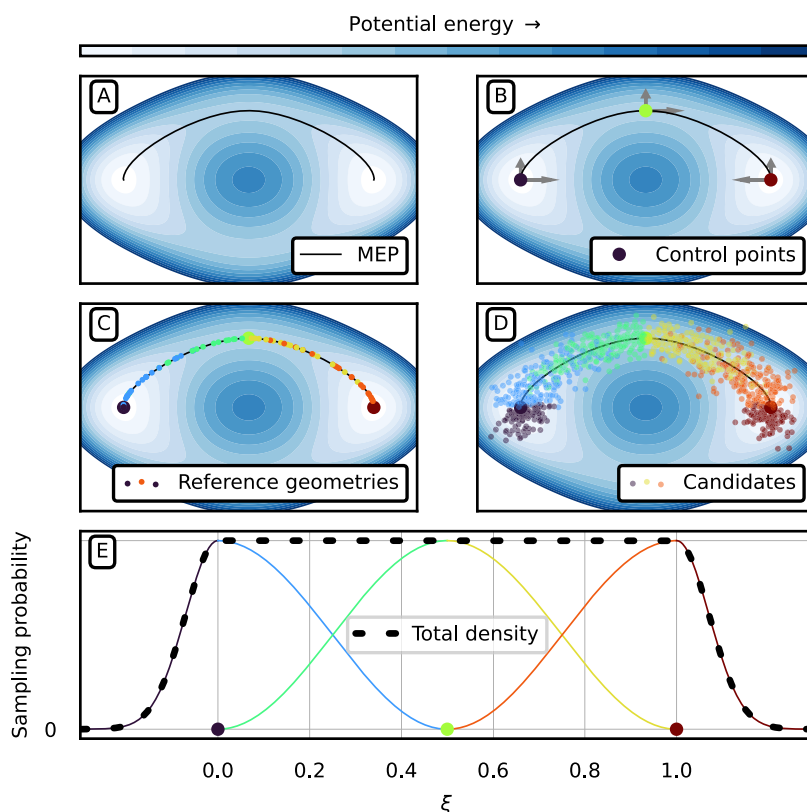
Figure 1. Schematic depiction of the TTS approach proposed for reactive systems. Panel A: an illustrative MEP winding through a model two-dimensional configuration space given by $V(x,y) = (1/6)\{4(1 - x^2 - y^2)^2 + 2(x^2 - 2)^2 + [(x + y)^2 - 1]2 + [(x - y)^2 - 1]^2 - 2\}$. Panel B: control points are selected and their local normal modes (gray arrows) are calculated. Here, the control points are taken as the two end-point minima on the MEP (purple and brown dots) and the transition state (yellow−green). Panel C: a much denser set of uniformly distributed reference geometries is generated along the MEP. Panel D: each of the reference geometries is distorted using the local modes of their assigned control point (as detailed in panel E) to become a candidate geometry. Panel E: a detailed view showing how each reference geometry is assigned to a set of local modes. A set of reference geometries on the MEP is assigned to each control point following eq 5. At the MEP edges, standard Gaussian thermal NMS is performed outside of the reaction coordinate (decaying purple and brown tails of the total density).

$$\rho_q(\Omega_1, ..., \Omega_{N_{int}}) \propto \prod_{i=1}^{N_{int}} \exp\left[-\frac{1}{2}\beta^*(\beta, \omega_i)\omega_i^2\Omega_i^2\right] \quad (4)$$

This simple modification allows one to generate an auxiliary quantum ensemble at practically the same cost as the classical one that would otherwise need to be approached from a significantly more demanding perspective, perhaps based on sampling techniques using the imaginary time path integral formalism.

2.2. Transition Tube Sampling. Up to this point, we have relied on the ability of the harmonic expansion around a single minimum to approximate the real PES so that the generated samples cover sufficiently all the relevant regions for the purpose of generating an MLP. Arguably, this is a reasonable requirement for most stable molecules with a single minimum geometry where the onset of the anharmonic region connected to the dissociation of the molecule is not thermally accessible. On the other hand, it is a stringent requirement for molecular systems which display conformational changes or chemical reactivity and, therefore, are represented by multiple PES minima connected by MEPs: features not captured by a single harmonic expansion. However, in such cases, it is desirable for the resulting model to be able to describe the potential energy landscape not only around local potential energy minima but also in the transition regions. This is vital in the case of low-$k_B T$ barriers, where spontaneous transitions occur during direct

MD. Nonetheless, it cannot be omitted even in the case of high-$k_B T$ barriers where an enhanced sampling simulation would be required to cross the barrier. Even if the transition does not actually occur, the presence of the transition state may introduce substantial anharmonicity within the original PES basin that an eventual MLP should learn. However, in the case of barrier transitions it is not desirable to attempt to build the C-NNP model starting from a candidate set representing the true thermal ensemble, even if we could obtain it, since this would lead to a possibly detrimental under-representation of the high-energy configurations close to the transition state in the resulting candidate set and, in turn, to poor performance of the resulting model in the transition regions.

Therefore, we propose the TTS method: a generalization of NMS for systems with transitions that employ local normal modes along an MEP to sample uniformly along the path and with proper thermal weights in all perpendicular directions. This leads to an auxiliary harmonic ensemble that differs significantly from the true thermal one but enables the construction of MLPs with uniform accuracy along the whole transition. The TTS method naturally reduces to thermal NMS as described above for single-minimum systems in the zero MEP length limit. The process, illustrated in Figure 1, starts by finding the MEP $\mathbf{R}(\xi)$ on the given PES (panel A). Here, ξ is a dimensionless reaction coordinate along the MEP curve through configuration space normalized to the interval from

0 to 1. In the following, we shall assume that the MEP is available as a continuous, differentiable function of the parameter ξ. In practice, this can be achieved by spline fitting of the discretized representation of the MEP originating from, for instance, a nudged elastic band calculation.[46] Note that by definition, the MEP is a minimum of the PES in all directions perpendicular to it. Once the relevant MEP is known, we continue by selecting a sparse set of control points $\mathbf{R}_c$, $c = 1, ..., N_p$ along the MEP at positions ξ_c for which the Hessian matrices are calculated and diagonalized to give the set of local normal mode vectors $\Omega_{c,i}$ and their corresponding frequencies. For instance, this can be the two end-point minima and the transition state between them (Figure 1, panel B), although there is no constraint on how densely one might select the control points along the MEP other than the limiting computational expense of the Hessian matrix calculation. The selection of the control points is performed by hand by the user, ensuring a homogeneous coverage of the MEP. Formally, the expansion of the PES along the MEP becomes exact under the harmonic approximation in the limit of a large number of control points N_p. Since we want to achieve uniform sampling along the MEP, we now proceed to the generation of reference geometries on the MEP that do have this property. Specifically, to each control point $\mathbf{R}_c$, we first assign a probability distribution $p_c(\xi)$ defined on the interval $[\xi_{c-1}, \xi_{c+1}]$ (Figure 1, panel E, interval between purple and brown control points) as

$$
p_c(\xi) = \begin{cases} \sin^2\left[\dfrac{\pi}{2|\xi_c - \xi_{c-1}|}(\xi - \xi_{c-1})\right] & \xi_{c-1} \leq \xi \leq \xi_c \\[2mm] \cos^2\left[\dfrac{\pi}{2|\xi_c - \xi_{c+1}|}(\xi - \xi_c)\right] & \xi_c \leq \xi \leq \xi_{c+1} \\[2mm] 0 & \text{otherwise} \end{cases}
\tag{5}
$$

Once this is done for all N_p control points, the identity

$$
p(\xi) = \sum_{c=1}^{N_p} p_c(\xi) = 1
\tag{6}
$$

holds over the whole length of the MEP (Figure 1, panel E over the range of ξ). Note that the choice of squares of harmonic functions is only one out of many possibilities, and any other pair of complementary functions that sum up to unity would work in this case. Next, we generate an arbitrary number of reference geometries $\mathbf{R}_0(\xi)$ at a chosen linear density by drawing random values of ξ from the above distributions and passing them to the continuous prescription of the MEP, all while keeping track of the parent c-th control point (Figure 1, panel C). Analogously to the distortion of the minimum geometry in the single-minimum case through eq 2, we distort each of these reference geometries using the normal modes and frequencies of its parent control point using

$$
\mathbf{R} = \mathbf{R}_0(\xi) + \mathbb{M}^{-1/2} \sum_i{}' \Omega_{c,i}[1 - \mathbb{P}(\xi)]\Omega_{c,i}
\tag{7}
$$

where the normal coordinate values $\Omega_{c,i}$ are sampled thermally according to eq 1 or 4 (Figure 1, panel D); the prime indicates that modes with imaginary frequencies are omitted from the sum. The matrix $\mathbb{P}(\xi)$ is the projector on the tangent direction at the point ξ which can be constructed analytically from

$\mathrm{d}\mathbf{R}(\xi)/\mathrm{d}\xi$. This is used to obtain distortions strictly perpendicular to the MEP and thus to correct for the approximate validity of the normal mode expansion calculated at ξ_c for all the displaced geometries. However, the use of the decaying probability distributions (eq 5) favors the use of the local modes close to their origin. Through this procedure, one obtains a set of candidate geometries distributed inside a tube around the MEP the width of which is given thermally. At this point, this tube still has open ends cut sharply by the planes defined by normal vectors equal to the MEP tangent vector at the end points of the MEP. Since these points are (usually) also well-defined minima on the PES, the presence of these sharp edges is easily sanitized by appending the usual thermal NMS samples at these minima, although only adding the configurations away from the MEP (Figure 1, decaying purple and brown lines). In other words, just one-half of the multivariate Gaussian is appended to the tube that does not overlap with it. In our TTS implementation, we ensure that the uniform density of the sampling along the MEP and the one at the peak of the half-Gaussian are seamlessly matched (as described in Section S1 of the Supporting Information).

Using the described sampling approach leads to an auxiliary ensemble of candidates that does not correspond to the true thermal ensemble, but contains a balanced selection of geometries distributed uniformly along the MEP with classical or quantum thermal displacements around it. Just like with plain NMS, we submit these samples as candidates to the QbC procedure, where in each iteration a large number of them is screened and a small number of those is selected to be included in the training set. Ab initio calculations are only required for these selected geometries. This enables the building of diverse training sets in which all representative structures that might be encountered in a future simulation are contained so that the resulting model is, in fact, able to accurately describe the PES along the whole MEP, even in regions that have negligible thermal populations. Similar to NMS, the computational cost of TTS is determined primarily by the MEP optimization procedure, the Hessian calculation, the C-NNP prediction required for screening, and the electronic structure calculations for the selected geometries. In general, this can be expected to be substantially less computationally demanding than executing direct, or even enhanced sampling, classical, or path integral AIMD simulations and sampling from their trajectories.

3. COMPUTATIONAL DETAILS

3.1. Ab Initio Electronic Structure. Two different levels of electronic structure theory were used in the simulations presented in this work. In both cases, we used the implementation provided by the CP2K software package[48] with its Quickstep DFT module.[49,50] We described the electronic structure of the benzene molecule in the gas phase at the self-consistent charge density-functional tight binding[51] (SCC-DFTB) level with third-order diagonal corrections. The system was enclosed in a 10 Å wide cubic box with open boundary conditions. For malonaldehyde and DABQDI systems, we used the revPBE0-D3 hybrid density functional[52−55] combined with the TVZ2P Gaussian basis set[49,56,57] to represent the molecular orbitals and a plane wave basis with a 600 Ry cutoff to represent the density. The core electrons of the heavy atoms were represented using Goedecker−Tetter−Hutter pseudopotentials.[58] In addition, we used the auxiliary density matrix method[59] with the cpFIT3

fitting basis set for the DABQDI molecule. Both systems using hybrid DFT were centered in a 15 Å wide cubic box with open boundary conditions and the wavelet Poisson equation solver.

3.2. C-NNP Model Generation. Throughout all of our investigations, we used committees consisting of eight different Behler−Parrinello NNPs,[16] where for each of them, a different initialization of weights and a different 90% subset of the full training data set was used to ensure a diverse committee. The models consisted of two hidden layers of 20 nodes each and were trained using the multistream[60] adaptive extended Kalman[61,62] filter with 32 streams. The input features were a standard set of atom-centered symmetry functions.[26] The training of the individual models was done using the n2p2 package[60] and the selection of training structures by QbC was done with a development version of our AML package[63] following the procedures outlined in refs 26 and 27. For benzene, 20 structures were randomly sampled initially and in each of the 40 QbC iterations, 10 new structures with the highest committee force disagreement were added to the data set, for a total of 420 training geometries. The final NNPs were trained for 2000 epochs. For the first generation of malonaldehyde and DABQDI models, 20 initial structures were sampled randomly and then 15 structures were selected and added to the data set in each of the 40 QbC iterations, for a total of 620 training geometries. The selection process is based on the standard deviation of the force prediction of the individual committee members and in each iteration, a predetermined number of structures with the highest disagreement of the candidate structures is selected. In our case, adding 15 structures at each iteration forms a compromise between choosing only the optimal structure and limiting the number of QbC iterations (which each require training a new committee) necessary to get to a sufficient training set size. For malonaldehyde, where additional generations of models were required (as detailed in Section 4), the original training set was supplemented by additional structures QbC-sampled from an MD trajectory which was produced using the previous C-NNP model. Here, 15 structures were added in each iteration until the force committee disagreements for the selected structures and the remaining candidate structures were similar. Like in our previous work, we chose this stopping criterion over a predetermined force disagreement threshold because it works well on its own, whereas an additional calibration against the force error would have been necessary to determine a suitable threshold. All reference calculations were done using CP2K and the electronic structure settings described above.

3.3. Geometry Optimization and Vibrational Analysis. The optimization of the minimum reference geometries for benzene and DABQDI was executed natively in the CP2K software. It was performed using the BFGS optimizer[64] combined with threshold criteria of 0.07 eV Å^{-1} for the maximum change in force components, 0.009 Å for the change in atomic positions, and 0.13 eV for the change in total energy. For the malonaldehyde molecule, we employed the Atomic Simulation Environment (ASE)[65] together with CP2K and performed the optimization using the FIRE optimizer[66] while specifying only a force criterion of 0.01 eV Å^{-1}. Additional constrained optimizations in the case of DABQDI needed for the relaxed PES scan were performed using the constraint functionality provided by ASE together with its FIRE optimizer. The Hessian matrix evaluation on the optimized structures was performed in each case using CP2K and a Cartesian atomic displacement of 0.0005 Å.

3.4. Nudged Elastic Band Calculations. The relevant MEPs needed for the TTS procedure were obtained through the climbing-image[67] nudged elastic band[46] (CI-NEB) optimization procedure as implemented in CP2K. The initial band geometries in this work consisted of 15 replicas of the molecule in question including the two fixed, preoptimized endpoints, and were obtained through linear interpolation. The spring constant of the harmonic links between the neighboring replicas was kept constant at the value of 4.86 eV Å^{-2}. We used a force convergence criterion of 0.007 eV Å^{-1} and the minimization of the band energy was performed using a DIIS optimizer.

3.5. MD Simulations. All MD simulations involving both ab initio as well as C-NNP potentials[33] were run using the CP2K package. The simulations with the classical representation of the nuclei were propagated at a temperature of 300 K using a time step of 0.5 fs to numerically integrate the Langevin equation with the friction coefficient γ of 0.02 fs^{-1} to achieve canonical sampling. The path integral simulations that include nuclear quantum effects were performed using imaginary-time ring polymers consisting of 64 replicas using the RPMD propagator. The canonical distribution at 300 K was sampled using the local path integral Langevin equation thermostat[68] (PILE-L) with the time constant for the centroid motion of 200 fs while the integration time step was kept at 0.25 fs.

3.6. Umbrella Sampling. The initial conditions for each umbrella sampling window were extracted from a steered MD trajectory, which was performed in the CP2K v2022.1 software package combined with the PLUMED plugin.[69−71] In this case, the value of the proton-sharing coordinate δ_1 (as detailed in Section 4) was biased from −1.2 to 1.2 Å during a 10 ps long simulation using a moving harmonic restraint with the force constant κ of 500.0 kJ mol^{-1} Å^{-2}. The simulation was performed classically with an integration time step of 0.5 fs in the canonical ensemble at 300 K using a local CSVR thermostat[72] with a time constant of 50 fs.

30 equidistant umbrella sampling windows separated by 0.08 Å were set up from the above steered MD simulation. Individually in each window, the value of δ_1 was biased by a static harmonic restraint of 500.0 kJ mol^{-1} Å^{-2} and simulated for 50 ps using the same setup as for the steered MD simulation above. The overlap of the corresponding histograms of δ_1 values observed in each simulation window is shown in Section S2 of the Supporting Information. The value of δ_2 was kept unbiased in each simulation window but was monitored for use in the following analysis. The biased configurations were reweighed to the unbiased ensemble using a Python implementation of the multistate Bennet acceptance ratio[73,74] (MBAR) procedure to obtain both a one-dimensional free energy profile for the proton-transfer along δ_1 as well as a two-dimensional free energy surface showing the dependence on both proton-sharing coordinates. This was done by determining the thermal weight associated with each configuration in the biased simulations and using these to obtain the probability distribution in the δ_1, δ_2 subspace, and from that the corresponding free energy surface.

4. RESULTS AND DISCUSSION

To showcase the performance of the TTS procedure in the creation of models for realistic potentials, we select three different gas-phase molecules with an increasing complexity of their PES. We begin with benzene, which represents a single-minimum system with a close-to-harmonic potential at room

temperature and thus allows us to illustrate the simple thermal NMS procedure. This is followed by a study of the enol form of 1,3-propanedial (malonaldehyde), which exhibits reactivity by sharing the acidic proton between the two oxygen atoms spontaneously at ambient conditions.[75] Finally, we focus on a more involved proton-sharing system represented by 2,5-diaminobenzoquinone-1,4-diimine (DABQDI), which has two proton-sharing sites.[76] Spontaneous proton transfer is hindered by a barrier thermally insurmountable at room temperature, and an enhanced sampling simulation is necessary to determine the free energy profile.

4.1. Benzene. To lead off the discussion of the ability of TTS to seed a training set for the creation of C-NNP models for realistic systems in the gas phase, we focus on the benzene molecule. It represents an ideal example to illustrate the basic idea of thermal NMS using a single normal mode expansion at an optimal geometry since it features a single configurational minimum and the surrounding PES exhibits almost no anharmonic effects.

To prepare the ground for comparison with the relevant C-NNP data, we initially performed one 250 ps AIMD simulation of gas-phase benzene at 300 K at the DFTB level using a classical representation of the atomic nuclei as well as a 100 ps PIMD simulation using 64 replicas to approximate the imaginary time path. Two C-NNP models were then based on candidate sets obtained from a thermal NMS of gas-phase benzene using a Hessian matrix calculated at the same DFTB level of theory as the (PI)-AIMD simulations at 300 K for the classical model and with the appropriate effective temperatures at 300 K for the quantum one. The resulting models were evaluated on test sets consisting of 1000 structures sampled from the two AIMD trajectories. Both models performed very well with an energy root-mean-square error (RMSE) of 1.66 and 5.90 meV for the model constructed for the use without and with path integral structures, respectively. The RMSE for a single force component was 14.9 and 30.4 meV A^{-1}. Subsequently, the models were used to obtain new 500 ps long MD and 100 ps PIMD simulations at 300 K.

The comparison of the C-NNP models to the corresponding (PI)-AIMD trajectories in terms of molecular geometry properties is summarized in Figure 2. In general, we can see the expected broadening of probability distributions due to nuclear quantum effects when we compare the left and right columns of Figure 2. In both the classical and quantum case, we observe a perfect match between the ab initio (green shading) and C-NNP distributions (blue dashed lines) in C-C bond lengths (panels A and B), C-H bond lengths (panels D and E), C-C-C angles (panels E and F), and C-C-C-C dihedrals (panels G and H). The two types of covalent bonds have expected distributions; the mean of the C-C-C angle is located at 120° which shows the average hexagonal arrangement of the aromatic ring subject to planarity, which is, in turn, demonstrated by the (signed) C-C-C-C dihedral angle peaking at 0° as expected. This level of agreement suggests that the final models used for production MD represent excellent approximations of the original DFTB PES. The negligible deviations between the C-NNP and the (PI)-AIMD results are quantified by the differences shown in the small sub-panels in Figure 2 in blue. Additionally, we show the distributions of the NMS structures (orange dotted lines) alongside the anharmonic distributions. These exhibit significant overlap with both the (PI)-AIMD and C-NNP data. This suggests that the harmonic approximation to the original ensemble is
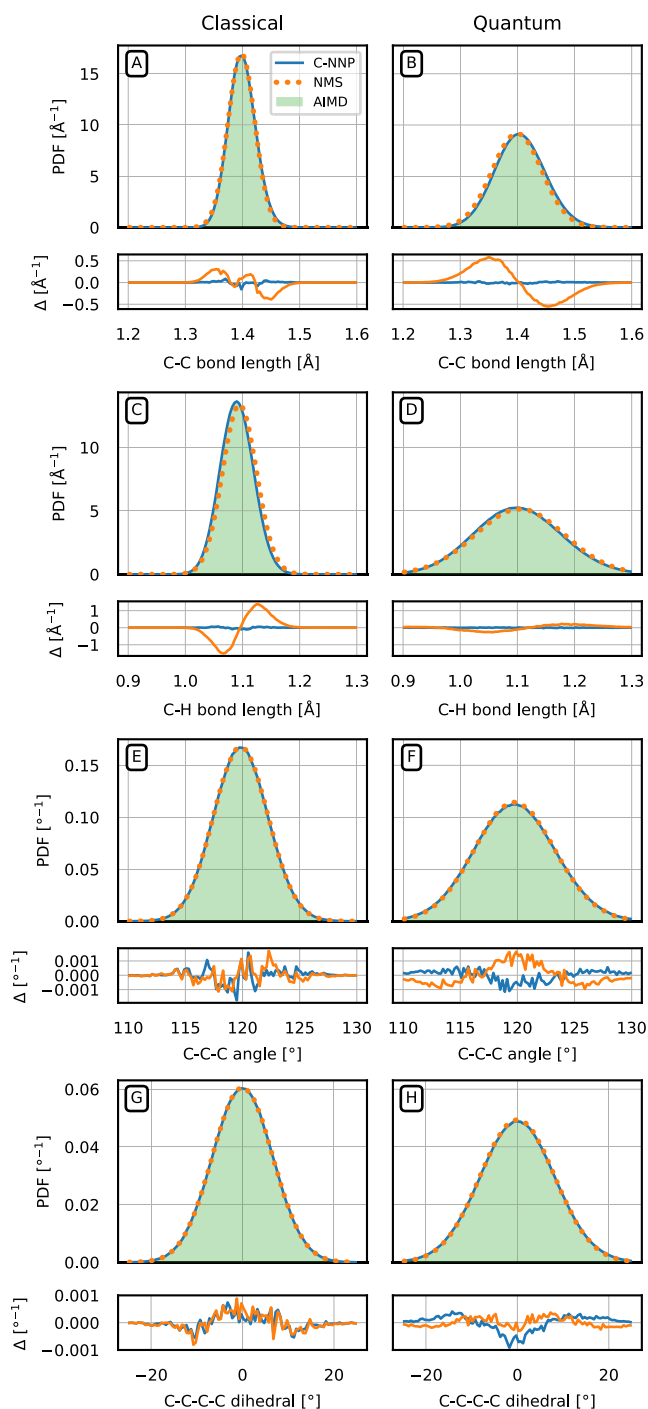

Figure 2. Thermal geometry properties of benzene in the gas phase at 300 K from classical MD (left column) and path integral MD (right column) compared between simulations using the reference DFTB potential, the harmonic TTS ensemble, and simulations using a C-NNP model building on the thermal NMS geometries. Panels A and B show the distribution of C-C bond lengths, panels C and D the distribution of C-H bond lengths, panels E and F the distribution of C-C-C angles, and, finally, panels G and H the distribution of the C-C-C-C dihedral angles. The smaller panels below each labeled panel show the deviations of the NMS and C-NNP data from the DFTB reference, using the same color coding as in panel A.

relatively good and confirms the assumed high degree of harmonicity of the 300 K gas-phase benzene PES, even in the quantum case. However, note that the match of the NMS data
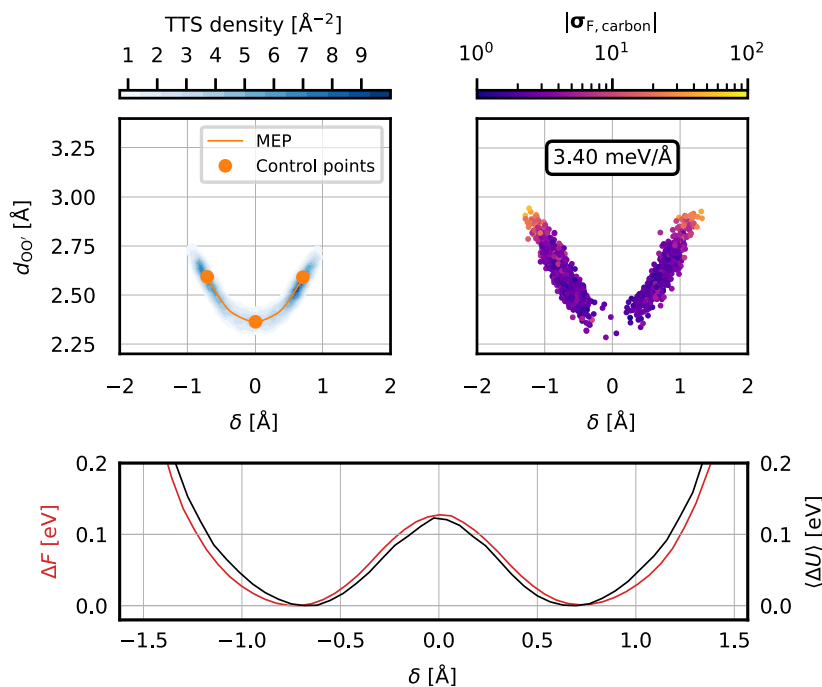
Figure 3. TTS sampling of the malonaldehyde proton transfer and the MD simulation with the resulting C-NNP model. The top left panel shows the relevant MEP (orange) with the three selected control points in the two configurational minima and the transition state highlighted and the distribution of the TTS geometries (blue). The top right panel shows a scatter plot of a subset of geometries (selected with a stride of 37.5 fs) obtained during a 250 ps MD simulation using the C-NNP model built on top of the TTS ensemble. Each point is colored by the norm of the force committee disagreement on the carbon atoms and the mean of the quantity is shown in the box. Note the high force disagreement in the high δ tails of the distribution. The bottom panel shows the Boltzmann-inverted free energy profile (red) and the corresponding binned average potential energy of the system (black, aligned to zero) along the proton transfer reaction as observed in the MD simulation. The error of the free energy, obtained by block averaging, is ≤ 2 meV, which roughly corresponds to the thickness of the red line.
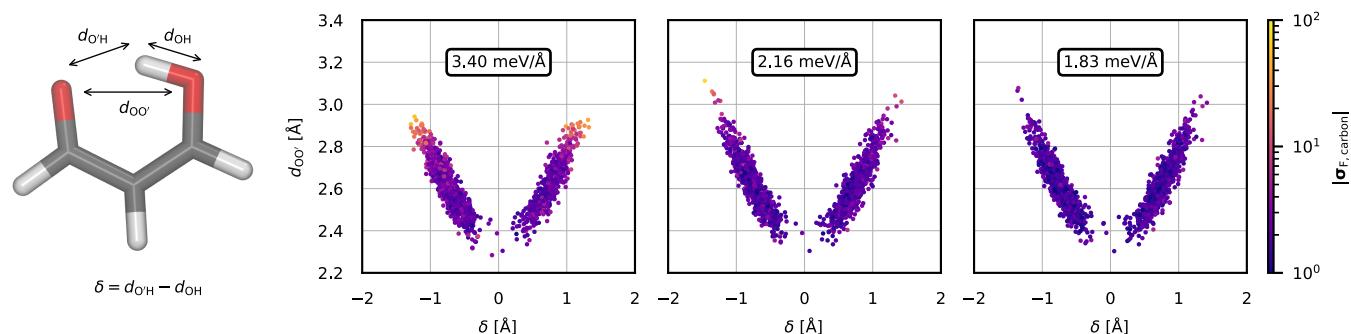

Figure 4. Evolution of the force disagreement of the carbon atoms through multiple instances of QbC. The left panel shows a subset of configurations originating from an MD simulation using a model trained on data selected directly from the TTS candidate set (identical data as in the top right panel of Figure 3 are shown). The force disagreement (depicted using the color scale) in the vicinity of the two configurational minima and along the proton-sharing reaction is adequately low; however, for structures with a high absolute δ, it is more than 1 order of magnitude higher. The central panel shows configurations and disagreements obtained from an MD simulation using a model trained on the initial training data augmented by QbC-selected high-disagreement configurations from the data in the left panel. In turn, the right panel shows data obtained by improving the model using the new data sampled in the simulation shown in the central panel. Most notably, structures at the tails of the populated configuration space are substantially improved. The mean force disagreement over all configurations in each data set is shown in the framed box in each panel. The δ and $d_{OO'}$ coordinates are illustrated in the snapshot to the left of the panels.

with the (PI)-AIMD data is not nearly as perfect as that of the C-NNP data and certain deviations are, in fact, present. As discussed in Section 2, these are to be expected since the NMS ensemble is only auxiliary and its goal is to provide sufficient coverage of the accessible PES which ultimately leads to an accurate C-NNP model. The differences of the NMS ensemble from the ab initio reference are again quantified as differences in Figure 2. Using thermal NMS, we were thus able to construct a C-NNP that accurately describes the original PES

of benzene based on a single Hessian evaluation and 420 single-point electronic structure calculations.

4.2. Malonaldehyde. The enol form of malonaldehyde is a simple organic molecule that has been used in MD simulations[75] to illustrate a simple intramolecular proton-transfer reaction

$$\text{(8)}$$
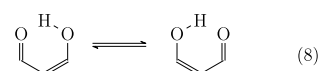
in which the proton is moved from the enol oxygen to the aldehyde oxygen with a simultaneous electronic rearrangement which altogether causes the reactant and the product to become symmetrically mirrored, chemically identical structures (eq 8). As such, malonaldehyde is a convenient molecule to demonstrate the ability of TTS to describe a simple reaction.

With the aim to describe the proton-sharing process accurately, we decided to model the original ab initio PES at the hybrid DFT level using the revPBE0-D3 functional. Thanks to the ability to use quantum normal coordinate distributions in the TTS method, we produced both classical and quantum models and classical and PIMD trajectories for malonaldehyde to test and showcase this functionality. However, we focus mostly on the classical case in the main text and discuss the complementary quantum results, for which qualitatively similar conclusions arise, in the Supporting Information, Section S2. The starting point of the TTS procedure is the proton-sharing MEP, which was discretized into 15 replicas and optimized using the CI-NEB procedure and the revPBE0-D3 density functional. For illustration purposes, we decided to use the full MEP with both the reactant and product explicitly represented: this is strictly speaking not necessary since the symmetry of the reaction allows the use of only one nontrivial half of the MEP for TTS. Out of the optimized full-length MEP, three control points were selected in the two minima (reactant and product) and in the transition state. For the visualization of the multidimensional configuration data, we choose the reduction into a 2D space of two geometric parameters: the proton-sharing coordinate $\delta(\mathbf{R}) = |\mathbf{R}_O - \mathbf{R}_H| - |\mathbf{R}_{O'} - \mathbf{R}_H|$ and the oxygen−oxygen $d_{OO'}(\mathbf{R}) = |\mathbf{R}_O - \mathbf{R}_{O'}|$, where O and O′ denote the two oxygen atoms that share the proton H. The obtained optimized MEP and the selected control points in this representation are shown in the top left panel of Figure 3 in orange. The chosen parameters, illustrated in the snapshot on the left of Figure 4, are not relevant for the execution of TTS itself, which takes place in the full dimension, but allow to conveniently show the results of the sampling in a reduced-dimensionality parameter space that is physically meaningful and suitable for the characterization of a proton transfer process. The TTS classical candidate structures were then generated using the procedure outlined in Section 2 at the temperature of 300 K, linear sampling density along the MEP of 1×10^5 Å^{-1}, and matched-density sampling at the minima. The distribution of the obtained configurations is shown in the top left panel of Figure 3 as blue contours. The same distribution colored by the assignment of each candidate to the individual control points (corresponding to the situation shown in panel D of Figure 1) is shown in Section S2 of the Supporting Information. Note that this particular presentation of the data does not do justice to the uniformity of the sample distribution along the MEP as the regions around the minima seem to be more populated than the transition state. This is an effect of the deformation of the configuration space by the projection on the selected subspace; the samples are distributed uniformly in the full dimension. After passing the resulting set of candidates through the QbC selection and training a C-NNP model on the obtained training set, the model was used to run a direct 250 ps MD simulation of gas-phase malonaldehyde at 300 K. A subset of the obtained configurations is shown in the form of a scatter plot in the top right panel of Figure 3 colored by the decadic logarithm of the norm of the force committee disagreement on carbon atoms in

the usual δ and $d_{OO'}$ representation. Additionally, a 1D free-energy profile obtained by a Boltzmann inversion of the probability density of configurations along the δ-axis is shown in the bottom panel of Figure 3; the height of the barrier is approximately 120 meV which corresponds to roughly 5 $k_B T$ at 300 K. This accounts for the expected low, but existing population surrounding the transition state at $\delta = 0$ Å. We estimated the error of the free energy by the block-averaging method followed by extrapolation to infinite block size. This gives errors lower than 2 meV over the studied range of δ, which corresponds to the fact that the transition is sampled often during the simulation and the fact that the raw profile is already symmetric. Alongside the free energy profile, we show the corresponding average potential energy as a function of δ.

An important observation can be made from the presented data. By comparison of the two distributions in the top panels of Figure 3, it is clear that the TTS distribution populates a smaller volume of the configuration space than the data obtained from the MD simulation. In this particular case, it means that TTS does not directly provide good enough coverage and the resulting model is undertrained in the missing, yet thermally accessible regions. Specifically, the C-NNP model performs poorly in the large $d_{OO'}$ tails of the shown distribution, as quantified by the larger force disagreement values. On the contrary, in regions around the proton-sharing MEP, the coverage is good and the force disagreement remains small, despite the tiny thermal population. Regardless of the elevated model uncertainty for some configurations, these MD simulations remain stable. The observed increased disagreement can be interpreted in the following way: going in the opposite direction from the minima as the proton-sharing MEP, the true anharmonic PES has a potential wall that grows slower than the harmonic wall captured by the TTS distribution and, therefore, the thermal coverage of the TTS configurations cannot reach far enough. In principle, this behavior is caused by either the strongly anharmonic character of the chemical bonds leading to bond dissociation or, more likely in this case, the presence of another reactive process leading to a new transition state. In the following two paragraphs, we present two possible solutions to the issue. The first one relies on an active-learning-based iterative improvement of the model which has the advantage of requiring no knowledge of the origin of the anharmonicity but is tedious to perform since several intermediate simulations and model generations need to be created. Meanwhile, the other solution relies on the chemical intuition of the user to identify the reactive nature of the issue with the aim to extend the initial TTS, which leads to a fully capable C-NNP model straight away.

The QbC process can be used to fill in an already existing training data set that has gaps, perhaps due to an incomplete TTS candidate set in the QbC selection for the initial model. We illustrate this process in Figure 4, where the left panel shows the same data as the top right panel of Figure 3 as a starting point. Regions of configuration space not covered well in the training set of this generation 1 model can be easily identified by the high committee disagreement, as can be seen in the tails of the distribution. Hence, one can start a new QbC using the existing training data set augmented by structures from an MD simulation performed with the initial C-NNP. Depending on the size of the gaps in the initial training data set, only a few iterations of QbC are typically necessary. However, adding these structures to the training data set can
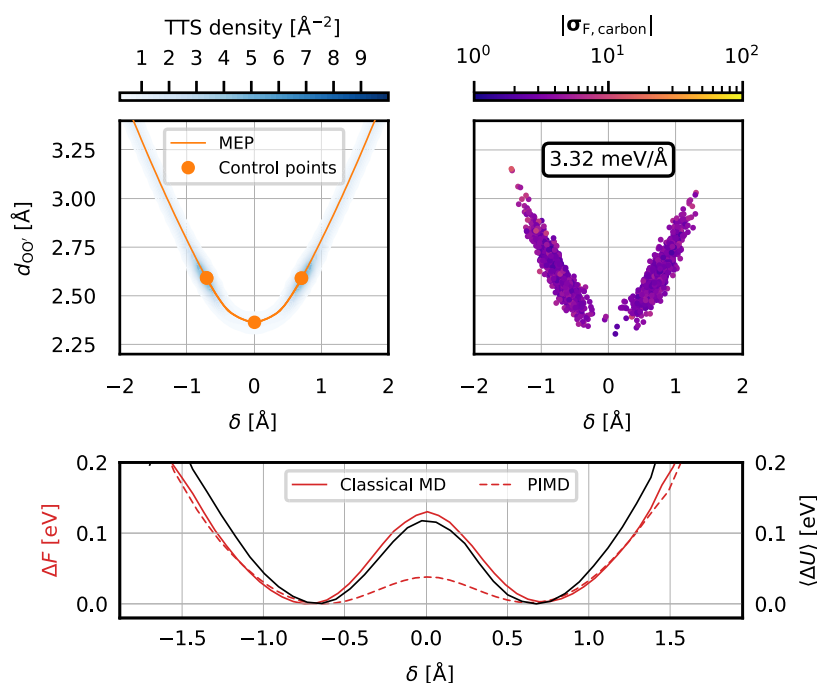
Figure 5. TTS sampling of the extended malonaldehyde MEP containing the proton transfer reaction as well as the single C–C bond torsion and the MD simulation with the resulting C-NNP model. Identical quantities as in Figure 3 are shown with the free energy being plotted here for both classical (solid line) and path-integral (dashed line) data. The orange curve shown in the top left panel is a union of the MEPs corresponding to the proton transfer and the C–C bond torsion; the projection of the latter into the δ, $d_{OO'}$ subspace is not, strictly speaking, a physically meaningful concept, but clearly visualizes the fact that MEP is the continuation in the desired direction. As such, the C-NNP model trained on the combined TTS structures has no further deficiencies as shown by the overall uniform force disagreement in the top right panel.

lead to substantial changes in the previously inaccurate regions of the PES, resulting in an MD simulation that again reaches new regions of the configuration space where the shape of the PES is yet unknown to the model and the committee disagreement is high. This can be seen in the generation 2 model shown in the middle panel of Figure 4. Therefore, multiple repetitions of the MD–QbC cycle might be necessary until a highly accurate model that exhibits low and uniform disagreements over the sampled data is reached, as is the case for the generation 3 model in the right panel of Figure 4. As such, the approach could become practically cumbersome when the regions of high disagreement coincide with regions of high free energy and long MD simulations are needed to uncover these structures, but nonetheless represents a general solution to the anharmonicity problem. Overall, repeating the cycle of sampling MD configurations with a given generation of a C-NNP model followed by training a new generation on a training set enhanced by high-disagreement QbC-selected structures from the previous MD simulations leads to a force disagreement that is lower in the problematic PES regions and, therefore, more uniform overall. In addition, we observe a decrease in the mean of the force disagreement of the sampled configurations due to the fact that the size of the training set increases in each generation. Specifically, 620 structures were used for the original model in the left panel of Figure 4, 800 structures for the model in the middle panel, and 950 structures for the last model in the right panel. This approach can be beneficial in systems where it is difficult to identify the origin of the anharmonicity of the original PES but is rather demanding from the point of view of both computational requirements and user involvement due to the need for the semi-supervised iterative procedure.

However, in the case of malonaldehyde, the general active learning iterative procedure to improve the model might be excessive. The possible reasons for the softer-than-harmonic wall due to conformational flexibility are few and the particular direction against which the first generation of the model is pushing can be easily identified with the s-cis and s-trans torsional isomerism

$$O \overset{H}{\sim} O \rightleftharpoons \overset{H \sim O}{\underset{O}{}} \qquad (9)$$
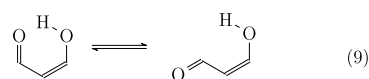
which is mediated by rotation around the C–C single bond in the propane backbone (eq 9). The optimized MEP corresponding to this torsion displays a perfect continuation in the correct direction when projected into the δ and $d_{OO'}$ subspace, as shown in the top left panel of Figure 5 in orange. Although this pair of descriptors is not appropriate for the whole torsion MEP, which entails a more complicated motion, it is accurate enough at small deviations from the equilibrium geometry. To include structures along this MEP into the initial (first generation) proton-sharing TTS, two control points were chosen in the minimum (shared by the two MEPs) and in the new transition state (not shown in Figure 5, as it is around $d_{OO'} = 3.8$ Å). We do not need to use the s-trans minimum at all, as we are not interested in including the transition itself, only the shape of the PES on the side of the global minimum. A new TTS was performed between these control points with the same parameters as the initial one and the resulting distribution of the combined sets of configurations is shown in blue in the top left panel of Figure 5. Running a 250 ps long MD simulation with a new C-NNP model trained on the QbC-selected training set from this combined candidate set leads to the distribution shown in the top right panel in Figure 5.
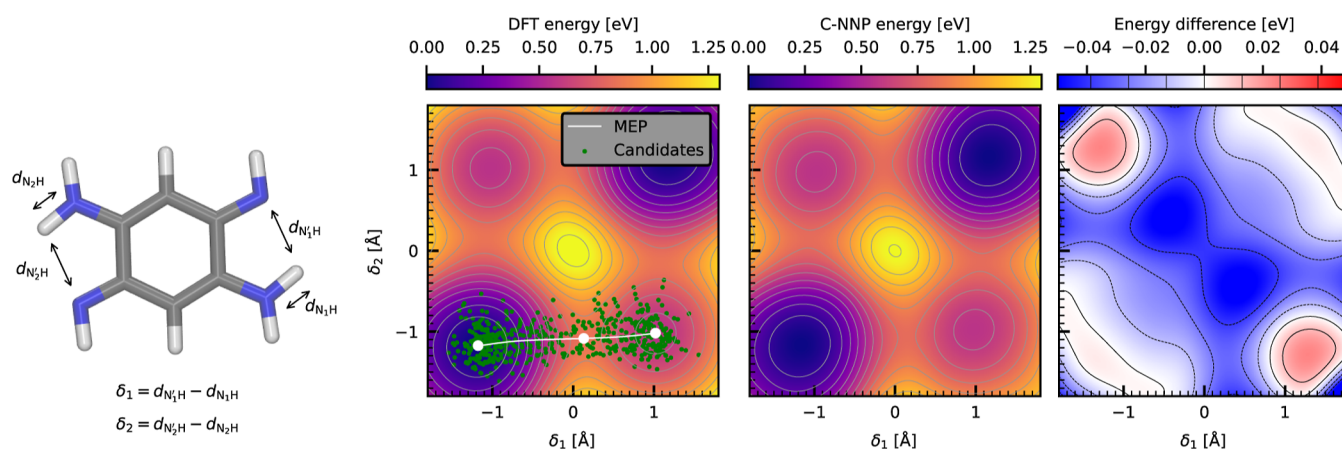
Figure 6. Comparison between the reference ab initio revPBE0-D3 and the C-NNP proton-sharing PESs of the DABQDI molecule. The left panel shows the projection of the reference DFT PES into the δ_1, δ_2 subspace using a color scale as well as individual isoenergetic contours. Furthermore, the minimal nontrivial MEP which describes a single proton transfer is depicted in white with the selected control points highlighted. A sparse subset of the TTS geometries sampled around the MEP at 300 K is shown in green. The middle panel shows the corresponding PES projection calculated with the resulting C-NNP. Finally, the right panel shows the difference between the two PESs aligned to the global minima. The black contours range from $-1.5\ k_BT$ to $1.5\ k_BT$ and are spaced by $0.5\ k_BT$ at 300 K. The two δ coordinates are illustrated in the snapshot of the DABQDI molecule to the left of the panels.

Clearly, the distribution reaches all the expected thermally accessible regions, the force disagreement is evened out across the configurations, and the high-disagreement tails are no longer present. The mean value of the disagreement is comparable to that of the first-generation model, as these two models are based on training sets of the same size. The bottom panel of Figure 5 shows the potential energy and the free energy profile along δ, where the latter is shown for both classical and path-integral data. In all three curves, one can observe a softening of the barrier in the high $|\delta|$ regions in comparison to the data shown in Figure 3 resulting from the present C-NNP being aware of the anharmonic nature of the PES in these regions. The transition-state free energy of 43 meV for the quantum simulation is ~4 times lower than in the classical case, which is a manifestation of proton tunneling through the barrier. This is qualitatively consistent with existing literature but quantitatively different from the results reported for the BLYP functional, where the classical free energy barrier is somewhat higher and the quantum effect is substantially smaller, decreasing the barrier by a factor of only ~2.[75]

For further insights, a test set independent of the training set data was created by generating 500 structures using TTS and sampling 500 structures from the classical MD trajectory shown in Figure 5, and evaluating their energies and forces with the revPBE0-D3 reference method. The generation 3 C-NNP of the iterative approach as well as the extended TTS C-NNP performs well with an energy RMSE of 1.80 and 3.44 meV and a force component RMSE of 18.3 and 24.4 meV A^{-1}, respectively. The slightly elevated RMSE of the extended TTS approach is due to the broader coverage of the training set. It includes a range of geometries along the C−C single bond torsion, even in regions that are not populated during MD, leading to a less dense coverage of the rest of the configuration space. The validation errors of the intermediate models of the iterative approach and the distribution of errors within configuration space are discussed in more detail in the Supporting Information, Section S2.

Both of the approaches above thus yield highly accurate models for the description of the proton-sharing reaction in malonaldehyde for classical and quantum nuclei. The advantage of one over the other therefore depends mostly on the specific situation in which the need for any of them should arise: if the reaction coordinate of the complementary transition can be identified, then the latter approach using the extended TTS reaches the desired result with higher efficiency. Note that this approach can also be used when multiple different transitions are to be included in a single model. Finally, it is worth noting that an MLP trained on structures selected only from the quantum formulation of TTS and PIMD trajectories performs well for classical MD simulations of malonaldehyde, too, as detailed in Section S2 of the Supporting Information.

4.3. DABQDI. The most complex reactive system used to demonstrate the performance of the TTS method is represented by the DABQDI molecule. This nitrogenated benzoquinone derivative can exchange two protons between the neighboring amine and imine groups

$$\text{(10)}$$
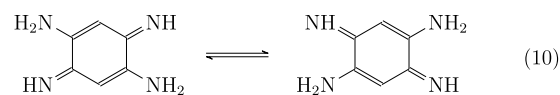
again accompanied by an electronic rearrangement that maintains the π-electron conjugation throughout the process (eq 10). However, this time, the proton-sharing does not take place at ambient conditions, which suggests high barriers to the process.

The corresponding PES reduced to the relevant δ_1 and δ_2 subspace (illustrated in the snapshot in Figure 6), where each proton-sharing coordinate describes a single proton-sharing site, was obtained at the revPBE0-D3 level of electronic structure theory through a relaxed scan of the molecular potential energies while applying appropriate constraints and is shown in the left panel of Figure 6. The shown data was aligned so that the global minimum of the PES corresponds to the zero-energy level. The typical structure of the PES featuring four distinct configurational minima and four transition states corresponds to a sequential double proton transfer at the level of an MEP. Here, one proton is first fully exchanged to reach an intermediate state located at a higher
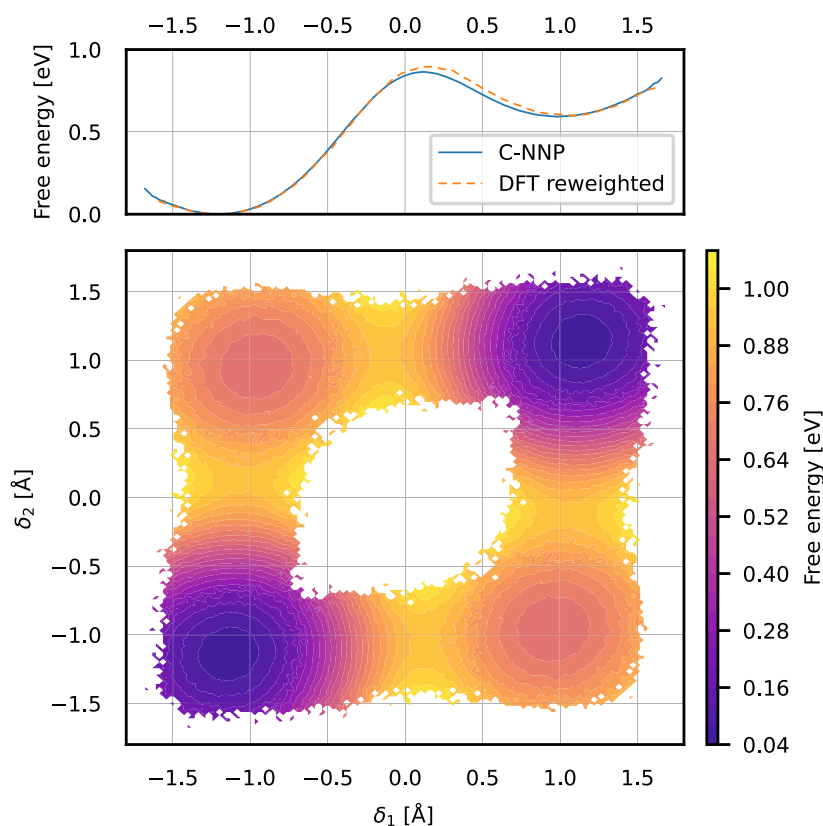
Figure 7. Umbrella sampling simulation of the single proton transfer in the DABQDI molecule along the δ_1 collective variable using the C-NNP potential. The top panel shows the obtained free energy profile in blue. For validation purposes, we also show the DFT free energy profile (orange, dashed) obtained by reweighting the C-NNP configurations as described in the text. Note that no umbrella sampling using the DFT potential was performed to obtain the DFT free energy profile. The bottom panel shows the full 2D free energy surface obtained by weighting the distribution in the two proton-sharing coordinates by the thermal Boltzmann factors extracted from the biased simulation and symmetrizing the resulting histogram.

potential energy and only then the other proton follows. The height of the potential energy barrier for this sequential process of roughly 0.8 eV indicates that its thermal rate should be negligible. The alternative concerted proton transfer path that is seen in other species including carboxylic acid dimers[77] is classically disallowed in this case by a tall (>1.2 eV) potential barrier in the middle of the presented PES which represents a second-order saddle point and, as such, no MEP can go through it. The symmetry of the DABQDI PES allows us to explicitly address only a single proton transfer: unlike in the previous example, we exploit this feature here for the C-NNP model generation. The relevant nontrivial MEP connecting the two chemically distinct minima was obtained using the CI-NEB optimization at the revPBE0-D3 level of theory and is shown in white in the left panel of Figure 6. From there, the straightforward TTS candidate generation was performed using the three usual control points in the two minima and in the transition state at 300 K with the linear sampling density of 1×10^3 Å^{-1}. A subset of the candidate geometries is shown in the left panel of Figure 6 as green points. The obtained C-NNP model was used to recreate the 2D proton transfer PES which is shown in the middle panel of Figure 6 with the energies aligned in the same way as in the DFT case. Qualitatively, the C-NNP model captures all the features of the original ab initio PES including the position of the minima, the transition state, and the barriers, as well as the potential energy values. Note that the good agreement in the representation of the central barrier in spite of the lack of corresponding

geometries in the TTS candidates is due to the successful extrapolation by the model. The quantitative difference between the ab initio and C-NNP potential energy landscape is captured in the right panel of Figure 6 which shows the difference between them relative to the global minima, with contours in multiples of 0.5 $k_B T$ for reference. It is important to view this deviation in the context of the height of the barrier, which is 0.71 eV (27.3 $k_B T$, deviation of ~−1.25 $k_B T$), and the relative energy of the two minima, which is 0.45 eV (17.3 $k_B T$, deviation of ~0.5 $k_B T$). A test set calculated over the TTS auxiliary ensemble quantifies the error in the thermal vicinity of the reaction coordinate to 0.88 meV/atom for energies and 47.6 meV/A for forces. These errors are comparable to test set errors of accurate nonreactive models for water using the same NNP architecture.[27,60] This difference could be decreased further, if desired, by optimizing hyperparameters of the model, by completely changing the architecture of the MLP, or by increasing the size of the training set beyond the current intentionally rather small set of 620 structures.

Since DABQDI features barriers that are not practically accessible by direct MD, it serves as a useful example to illustrate the power of the TTS-based C-NNP model to perform an enhanced sampling calculation to correctly estimate the free energy profile of the double proton transfer at 300 K. This was obtained using an umbrella sampling simulation in the coordinate δ_1 with the C-NNP PES (as described in Section 3) followed by a multistate Bennet acceptance ratio (MBAR) reweighing of the biased config-

urations. The obtained 1D free energy profile in δ_1 is shown in blue in the top panel of Figure 7. The transition state is located at roughly 0.8 eV above the global minimum. Comparing this with the value of the corresponding potential energy suggests that the entropic contribution in the gas-phase system is small and that the population at the barrier is clearly negligible at 300 K. The maximum error of the free energy profile is ±2 meV, which was estimated by the error analysis infrastructure of the PyMBAR implementation of MBAR[73] (shown in Figure S6 of the Supporting Information). To validate the obtained free energy profile, we perform a reweighing of a subset of the obtained configurations in each umbrella window to the original DFT ensemble. This is achieved by additionally multiplying each MBAR-obtained unbiased weight by the corresponding factor $e^{-\beta\Delta E}$, where the energy difference ΔE is the difference between the C-NNP and DFT potential energy for each configuration. For this purpose, we used a total of 3000 configurations obtained by selecting 100 geometries evenly spaced in time from each of the 30 umbrella sampling windows. The resulting profile, which is a good approximation to the full-DFT free energy profile, is displayed as the orange dashed line in Figure 7 and shows very good correspondence with the profile obtained using the C-NNP model alone. This procedure thus at the same time validates the C-NNP model and provides DFT data for a fraction of the cost of the hypothetical purely ab initio enhanced sampling simulation. Monitoring the values of the collective variable δ_2 along the umbrella sampling simulation and using the thermal weights obtained from the MBAR treatment of the biased simulations also allows for recovering the 2D free energy surface in δ_1 and δ_2. Its symmetrized version is shown in the bottom panel of Figure 7.

4.4. Computational Efficiency.
The computational speedup due to using TTS rather than the traditional sampling based on AIMD simulations is not straightforward to measure as it depends on the particular system in question. At least, one can give a rough estimate of the order of magnitude of the difference between the number of single-point evaluations required by both approaches. For TTS, its dominant computational requirements arise due to the parts of the algorithm that require ab initio calculations: the MEP optimization, the calculation of the Hessians at the chosen control points, and the evaluation of training energies and forces during QbC. Therefore, the total number of single-point energy and force evaluations is

$$N_{SP} = N_{SP}^{MEP} + N_{SP}^{Hessian} + N_{SP}^{QbC} \qquad (11)$$

where

$$N_{SP}^{Hessian} = 6NN_p \qquad (12)$$

when analytic forces are available. For the MEP optimization, a reasonable expectation is an NEB with low tens of replicas that converges to the desired path on the order $\sim 10^2$ steps, which means the total number of single-point evaluations N_{SP}^{MEP} on the order of 10^3. For the Hessians, we typically require three control points (maybe five in more challenging cases not encountered in this work), which yields $N_{SP}^{Hessian}$ on the order of 10^2. Similarly, for the QbC we find that a converged training set size is as small as several hundred thermal geometries. In total, the number of single-point evaluations required for TTS is thus on the order of 10^3. For simple NMS, where the MEP is replaced by optimization of a single structure to potential

energy minimum and where only one Hessian is required, we expect the number of single points to be an order of magnitude less than TTS. The number of single-point evaluations for the AIMD alternative is again not clear, as it depends on the system-dependent correlation time scale. For a small gas-phase molecule with a single minimum such as benzene, a reasonable (although still optimistic) estimate is that a 20 ps long classical direct AIMD trajectory with a 0.5 fs integration time step, which allows us to extract 400 geometries with a 50 fs stride, should be sufficient to provide decorrelated data for the training of a C-NNP. This already requires $\sim 10^4$ single-point evaluations. However, more complex systems typically require longer simulations to provide sufficient sampling. Crucially, reactivity increases the computational cost of a hypothetical reference AIMD simulation dramatically. Already in the reactive case of malonaldehyde, adequately sampling the barrier regions (Figure 5) required $\sim 10^6$ MD steps of direct sampling, which would have to be performed at the ab initio level without TTS. Systems with higher barriers that will not be crossed by direct AIMD on reasonable time scales would require the use of enhanced sampling techniques, and the construction of a model compatible with quantum nuclei would require path integral simulations. These would, again, raise the cost of the simulations by at least an order of magnitude each—think, for instance, low tens of umbrella sampling windows to cover a reaction coordinate and tens to hundreds of path integral replicas to converge quantum properties. All in all, we expect TTS to be always computationally superior to running AIMD simulations. Especially when enhanced sampling and path integral simulations are needed for an appropriate description of the studied system, we expect the difference between the methods to be three to 4 orders of magnitude.

Another facet of the considerations of computational efficiency is the simulation performance of the resulting C-NNP in comparison to the original ab initio electronic structure method itself, which we illustrate on the DABQDI enhanced sampling simulation, where the acceleration of the C-NNP umbrella simulation in comparison to the naive execution with the original DFT method is substantial. To illustrate the computational savings, we can compare the times required for one MD step with the implementations in CP2K used in this work. With the hybrid functional, one MD step takes 272 s on a single core or 17 s on 32 cores (a full node) of our EPYC-based cluster. With the C-NNP, one step takes 0.006 s and does not scale meaningfully to more cores due to the small system size. This yields a speedup of $\sim 45{,}000\times$ on identical resources or $\sim 2800\times$ with more resources given to the DFT calculation. Obviously, the specific numbers will depend on the details of the electronic structure setup and the MLP architecture used, as well as the specific implementations and hardware used, but this behavior of our particular setup should provide a general idea.

5. CONCLUSIONS

In this work, we have introduced the TTS method to sample thermal geometries around MEPs that describe barrier-crossing transitions in molecular systems. The goal of the method is to provide a physically meaningful set of candidate structures for the creation of MLPs without the need to run computationally demanding ab initio simulations. In our case specifically, we submit these geometries to QbC and construct a C-NNP model, but the same candidates could be used for

other types of models as well. The execution of the TTS protocol as a whole entails a relatively modest computational cost with respect to the original ab initio method that is given by the MEP optimization, several Hessian evaluations, and a small number of single-point ab initio calculations for the QbC-selected geometries. In terms of application to realistic systems, the TTS method yields highly accurate C-NNP models in all studied cases. This was achieved either by using the generated candidate set directly or by letting the resulting C-NNP model undergo additional active learning generations to compensate for a pronounced anharmonic effect as seen in the case of malonaldehyde. As such, the performance of TTS in the presented test systems demonstrates its robustness and efficiency and suggests applicability in most gas-phase systems, including highly anharmonic cases.

A noteworthy feature of the TTS method is its ability to provide thermal geometries sampled from the quantum thermal distribution at essentially the computational cost of the classical case. As such, models that are appropriate for use in path integral simulations are made readily available without the need to run expensive PI-AIMD simulations at all. However, it is important to recognize that although the present formulation of TTS can address quantum behavior, it has limitations in this regard that derive from the fundamentally classical nature of the MEP. Nuclear quantum effects, in particular quantum tunneling through the potential barrier, can cause the configuration-space probability density of the system to deviate from the transition tube around the MEP in a way that renders the coverage by TTS samples insufficient.

To account for this, the above formulation of TTS can be straightforwardly generalized from sampling around classical MEPs to ring-polymer instantons,[78] which represent the paths of optimal tunneling. While this modification requires essentially no adaptation of the TTS theory and implementation itself, one can anticipate an elevated computational cost due to the required instanton optimization at the explicit ab initio level. The approach will find applications beyond the gas phase, in systems where vibrational normal modes are a meaningful concept, such as in the study of materials, molecular crystals, or in surface science for the description of growth and molecular adsorption. Disordered condensed phase, including liquids, represents a more challenging case in which TTS alone is not applicable for efficient thermal sampling of geometries. However, the auxiliary use of the TTS protocol in obtaining more diverse thermal structures of liquids, for instance with the help of local normal modes, should be explored. Our research anticipates the need to address some of these condensed-phase systems in the near future and we expect TTS to be a valuable tool in the creation of accurate, yet computationally accessible potentials that will enable the accurate description of these more complex systems at unprecedented sizes and simulation time scales.

ASSOCIATED CONTENT

ⓈI Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00391.

> Additional theory details, additional results, and C-NNP validation (PDF)

AUTHOR INFORMATION

Corresponding Author

Ondrej Marsalek − *Charles University, Faculty of Mathematics and Physics, 121 16 Prague 2, Czech Republic;* ⓞ orcid.org/0000-0002-8624-8837; Email: ondrej.marsalek@mff.cuni.cz

Authors

Krystof Brezina − *Charles University, Faculty of Mathematics and Physics, 121 16 Prague 2, Czech Republic;* ⓞ orcid.org/0000-0003-0285-1282

Hubert Beck − *Charles University, Faculty of Mathematics and Physics, 121 16 Prague 2, Czech Republic;* ⓞ orcid.org/0000-0002-7191-2192

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.3c00391

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Tian, P. Molecular dynamics simulations of nanoparticles. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2008**, *104*, 142−164.

(2) De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **2016**, *59*, 4035−4061.

(3) Hollingsworth, S. A.; Dror, R. O. Molecular Dynamics Simulation for All. *Neuron* **2018**, *99*, 1129−1143.

(4) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31*, 1902765.

(5) Yao, N.; Chen, X.; Fu, Z. H.; Zhang, Q. Applying Classical, Ab Initio, and Machine-Learning Molecular Dynamics Simulations to the Liquid Electrolyte for Rechargeable Batteries. *Chem. Rev.* **2022**, *122*, 10970−11021.

(6) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: New York, 2009.

(7) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, Inc.: Mineola, N.Y., 1996.

(8) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(9) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864−B871.

(10) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133−A1138.

(11) Marsalek, O.; Chen, P. Y.; Dupuis, R.; Benoit, M.; Méheut, M.; Bačić, Z.; Tuckerman, M. E. Efficient calculation of free energy differences associated with isotopic substitution using path-integral molecular dynamics. *J. Chem. Theory Comput.* **2014**, *10*, 1440−1453.

(12) Spura, T.; Elgabarty, H.; Kühne, T. D. On-the-fly" coupled cluster path-integral molecular dynamics: impact of nuclear quantum effects on the protonated water dimer. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14355−14359.

(13) Kapil, V.; VandeVondele, J.; Ceriotti, M. Accurate molecular dynamics and nuclear quantum effects at low cost by multiple steps in

real and imaginary time: Using density functional theory to accelerate wavefunction methods. *J. Chem. Phys.* **2016**, *144*, 054111.

(14) Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.

(15) Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.* **2021**, *20*, 750−761.

(16) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.

(17) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(18) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.* **2017**, *8*, 3192−3203.

(19) Schütt, K. T.; Kindermans, P. J.; Sauceda, H. E.; Chmiela, S.; Tkatchenko, A.; Müller, K. R. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 992−1002.

(20) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 2453.

(21) Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *J. Phys. Chem. Lett.* **2020**, *11*, 2336−2347.

(22) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, *134*, 074106.

(23) Natarajan, S. K.; Behler, J. Neural network molecular dynamics simulations of solid-liquid interfaces: Water at low-index copper surfaces. *Phys. Chem. Chem. Phys.* **2016**, *18*, 28704−28725.

(24) Schütt, K. T.; Sauceda, H. E.; Kindermans, P. J.; Tkatchenko, A.; Müller, K. R. SchNet - A deep learning architecture for molecules and materials. *J. Chem. Phys.* **2018**, *148*, 241722.

(25) Sivaraman, G.; Krishnamoorthy, A. N.; Baur, M.; Holm, C.; Stan, M.; Csányi, G.; Benmore, C.; Vázquez-Mayagoitia, Á. Machine-learned interatomic potentials by active learning: amorphous and liquid hafnium dioxide. *npj Comput. Mater.* **2020**, *6*, 104.

(26) Schran, C.; Thiemann, F. L.; Rowe, P.; Müller, E. A.; Marsalek, O.; Michaelides, A. Machine learning potentials for complex aqueous systems made simple. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2110077118.

(27) Schran, C.; Brezina, K.; Marsalek, O. Committee neural network potentials control generalization errors and enable active learning. *J. Chem. Phys.* **2020**, *153*, 104105.

(28) Artrith, N.; Behler, J. High-dimensional neural network potentials for metal surfaces: A prototype study for copper. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 045439.

(29) Gastegger, M.; Behler, J.; Marquetand, P. Machine learning molecular dynamics for the simulation of infrared spectra. *Chem. Sci.* **2017**, *8*, 6924−6935.

(30) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is more: Sampling chemical space with active learning. *J. Chem. Phys.* **2018**, *148*, 241733.

(31) Hansen, L. K.; Salamon, P. Neural Network Ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 993−1001.

(32) Cooper, A. M.; Hallmen, P. P.; Kästner, J. Potential energy surface interpolation with neural networks for instanton rate calculations. *J. Chem. Phys.* **2018**, *148*, 094106.

(33) Schran, C.; Behler, J.; Marx, D. Automated Fitting of Neural Network Potentials at Coupled Cluster Accuracy: Protonated Water Clusters as Testing Ground. *J. Chem. Theory Comput.* **2020**, *16*, 88−99.

(34) Krogh, A.; Vedelsby, J. Neural Network Ensembles, Cross Validation, and Active Learning. *Adv. Neural Inf. Process. Syst.* **1995**, *7*, 231−238.

(35) Seung, H. S.; Opper, M.; Sompolinsky, H. Query by committee. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory-COLT'92*: New York, New York, USA, 1992; pp 287−294.

(36) Chen, L.; Sukuba, I.; Probst, M.; Kaiser, A. Iterative training set refinement enables reactive molecular dynamics via machine learned forces. *RSC Adv.* **2020**, *10*, 4293−4299.

(37) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How van der Waals interactions determine the unique properties of water. *Proc. Natl. Acad. Sci.* **2016**, *113*, 8368−8373.

(38) Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. Machine learning for the structure-energy-property landscapes of molecular crystals. *Chem. Sci.* **2018**, *9*, 1289−1300.

(39) Vandermause, J.; Torrisi, S. B.; Batzner, S.; Xie, Y.; Sun, L.; Kolpak, A. M.; Kozinsky, B. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Comput. Mater.* **2020**, *6*, 20.

(40) Wang, W.; Yang, T.; Harris, W. H.; Gómez-Bombarelli, R. Active learning and neural network potentials accelerate molecular screening of ether-based solvate ionic liquids. *Chem. Commun.* **2020**, *56*, 8920−8923.

(41) Miksch, A. M.; Morawietz, T.; Kästner, J.; Urban, A.; Artrith, N. Strategies for the construction of machine-learning potentials for accurate and efficient atomic-scale simulations. *Mach. Learn.: Sci. Technol.* **2021**, *2*, 031001.

(42) Schwalbe-Koda, D.; Tan, A. R.; Gómez-Bombarelli, R. Differentiable sampling of molecular geometries with uncertainty-based adversarial attacks. *Nat. Commun.* **2021**, *12*, 5104−5112.

(43) Vandenhaute, S.; Cools-Ceuppens, M.; DeKeyser, S.; Verstraelen, T.; Van Speybroeck, V. Machine learning potentials for metal-organic frameworks using an incremental learning approach. *npj Comput. Mater.* **2023**, *9*, 19.

(44) Rupp, M.; Ramakrishnan, R.; Von Lilienfeld, O. A. Machine Learning for Quantum Mechanical Properties of Atoms in Molecules. *J. Phys. Chem. Lett.* **2015**, *6*, 3309−3313.

(45) Schreiner, M.; Bhowmik, A.; Vegge, T.; Busk, J.; Winther, O. Transition1x - a dataset for building generalizable reactive machine learning potentials. *Sci. Data* **2022**, *9*, 779.

(46) Jónsson, H.; Mills, G.; Jacobsen, K. W. Nudged elastic band method for finding minimum energy paths of transitions. *Classical and Quantum Dynamics in Condensed Phase Simulations*; World Scientific, 1998; pp 385−404.

(47) Ceriotti, M.; Bussi, G.; Parrinello, M. Nuclear Quantum Effects in Solids Using a Colored-Noise Thermostat. *Phys. Rev. Lett.* **2009**, *103*, 030603.

(48) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; Vandevondele, J. Cp2k: Atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 15−25.

(49) Vandevondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.

(50) Kühne, T. D.; Iannuzzi, M.; Del Ben, M.; Rybkin, V. V.; Seewald, P.; Stein, F.; Laino, T.; Khaliullin, R. Z.; Schütt, O.; Schiffmann, F.; et al. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *J. Chem. Phys.* **2020**, *152*, 194103.

(51) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *51*, 12947−12957.

(52) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(53) Zhang, Y.; Yang, W. Comment on "generalized gradient approximation made simple". *Phys. Rev. Lett.* **1998**, *80*, 890.

(54) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.

(55) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics,

and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670−6688.

(56) Guidon, M.; Schiffmann, F.; Hutter, J.; Vandevondele, J. Ab initio molecular dynamics using hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 214104−214115.

(57) Guidon, M.; Hutter, J.; VandeVondele, J. Robust periodic Hartree-Fock exchange for large-scale simulations using Gaussian basis sets. *J. Chem. Theory Comput.* **2009**, *5*, 3010−3021.

(58) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 1703−1710.

(59) Guidon, M.; Hutter, J.; Vandevondele, J. Auxiliary density matrix methods for Hartree-Fock exchange calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2348−2364.

(60) Singraber, A.; Morawietz, T.; Behler, J.; Dellago, C. Parallel Multistream Training of High-Dimensional Neural Network Potentials. *J. Chem. Theory Comput.* **2019**, *15*, 3075−3092.

(61) Shah, S.; Palmieri, F.; Datum, M. Optimal filtering algorithms for fast learning in feedforward neural networks. *Neural Network.* **1992**, *5*, 779−787.

(62) Blank, T. B.; Brown, S. D. Adaptive, global, extended Kalman filters for training feedforward neural networks. *J. Chemom.* **1994**, *8*, 391−407.

(63) AML public GitHub repository. https://github.com/MarsalekGroup/aml (accessed March 20, 2023).

(64) Broyden, C. G. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J. Appl. Math.* **1970**, *6*, 76−90.

(65) Hjorth Larsen, A.; Jørgen Mortensen, J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; et al. The atomic simulation environment - A Python library for working with atoms. *J. Phys. Condens. Matter.* **2017**, *29*, 273002.

(66) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. Structural relaxation made simple. *Phys. Rev. Lett.* **2006**, *97*, 170201.

(67) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901−9904.

(68) Ceriotti, M.; Parrinello, M.; Markland, T. E.; Manolopoulos, D. E. Efficient stochastic thermostatting of path integral molecular dynamics. *J. Chem. Phys.* **2010**, *133*, 124104.

(69) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.

(70) The PLUMED consortium; et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.

(71) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.

(72) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(73) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.

(74) WHAM. https://github.com/apallath/WHAM (accessed March 20, 2023).

(75) Tuckerman, M. E.; Marx, D. Heavy-Atom Skeleton Quantization and Proton Tunneling in "Intermediate-Barrier" Hydrogen Bonds. *Phys. Rev. Lett.* **2001**, *86*, 4946−4949.

(76) Cahlík, A.; Hellerstedt, J.; Mendieta-Moreno, J. I.; Švec, M.; Santhini, V. M.; Pascal, S.; Soler-Polo, D.; Erlingsson, S. I.; Výborný, K.; Mutombo, P.; Marsalek, O.; Siri, O.; Jelínek, P. Significance of Nuclear Quantum Effects in Hydrogen Bonded Molecular Chains. *ACS Nano* **2021**, *15*, 10357−10365.

(77) Smedarchina, Z.; Siebrand, W.; Fernández-Ramos, A.; Meana-Pañeda, R. Mechanisms of double proton transfer. Theory and applications. *Z. Phys. Chem.* **2008**, *222*, 1291−1309.

(78) Kästner, J. Theory and simulation of atom tunneling in chemical reactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 158−168.