

Machine learning the Hohenberg-Kohn map for molecular excited states

Received: 9 February 2022

Accepted: 25 October 2022

Published online: 17 November 2022

Check for updates

Yuanming Bai ^{1,2,3}, Leslie Vogt-Maranto ³, Mark E. Tuckerman ^{2,3,4,5} & William J. Glover ^{1,2,3} ✉

The Hohenberg-Kohn theorem of density-functional theory establishes the existence of a bijection between the ground-state electron density and the external potential of a many-body system. This guarantees a one-to-one map from the electron density to all observables of interest including electronic excited-state energies. Time-Dependent Density-Functional Theory (TDDFT) provides one framework to resolve this map; however, the approximations inherent in practical TDDFT calculations, together with their computational expense, motivate finding a cheaper, more direct map for electronic excitations. Here, we show that determining density and energy functionals via machine learning allows the equations of TDDFT to be bypassed. The framework we introduce is used to perform the first excited-state molecular dynamics simulations with a machine-learned functional on malonaldehyde and correctly capture the kinetics of its excited-state intramolecular proton transfer, allowing insight into how mechanical constraints can be used to control the proton transfer reaction in this molecule. This development opens the door to using machine-learned functionals for highly efficient excited-state dynamics simulations.

Electronic excitations underlie numerous biological and physical processes of interest, including photosynthesis^{1,2}, DNA damage^{3,4}, photodynamic therapy⁵, photopharmacology^{6,7}, and solar energy conversion^{8,9}. Excited-state dynamics simulations are a powerful tool to uncover the mechanism of photoresponse in these systems, particularly when combined with *ab initio* electronic-structure calculations, since the important molecular motions do not need to be known *a priori* but are revealed by the dynamical trajectories^{10,11}. Given its reasonable accuracy within current approximations, linear-response time-dependent density-functional theory (LR-TDDFT)^{12,13} has become a work-horse method in the field; however, its computational expense (formally scaling with system size as N^3) and the need to solve the electronic-structure problem at every simulation timestep limits the system size and timescale amenable to study. Thus, a cheaper approach to electronic excitations and

excited-state dynamics is highly desirable but poses a significant challenge.

One might question whether the expense of LR-TDDFT is needed in the first place. For the ground state (GS), the first Hohenberg-Kohn (HK1) Theorem¹⁴, which provides the foundation of DFT, establishes the existence of a bijection from the ground-state electron density to the external potential of a many-body system. As a consequence, there exists a formal map from the ground-state electron density not only to the ground-state energy but also to every property of the system. The map from density to ground-state energy is encoded by an unknown universal density functional; however, practical approaches make use of the Kohn-Sham (KS) theory¹⁵, which expresses the functional via a fictitious non-interacting system that shares the density of the true system, allowing an exact treatment of kinetic energy and Coulomb contributions to the functional and approximations to the remaining

¹NYU Shanghai, 1555 Century Avenue, 200122 Shanghai, China. ²NYU-ECNU Center for Computational Chemistry at NYU Shanghai, 3663 Zhongshan Road North, 200062 Shanghai, China. ³Department of Chemistry, New York University, New York, NY 10003, USA. ⁴Simons Center for Computational Physical Chemistry at New York University, New York, NY 10003, USA. ⁵Courant Institute of Mathematical Science, New York University, New York, NY 10012, USA.

✉ e-mail: william.glover@nyu.edu

small exchange and correlation terms. While KS theory provides a practical scheme for solving simultaneously for the ground-state density and energy of a system, functionals that map from the ground-state density to excited-state energies are currently unknown, although, their formal existence is established by the HK1 theorem. As an alternative, LR-TDDFT provides an indirect means to the excited-state energies from knowledge of the ground-state density and its response to an external driving field^{12,13}. Given the expense of LR-TDDFT, a direct map from density to excited-state energies would be preferable.

Theoretical support for a direct route to excited-state energies comes from a generalization of the HK1 theorem to an excited-state density-energy bijection. While a universally general HK1 theorem for excited states with arbitrary external potentials has been disproved¹⁶, a special excited-state HK1 theorem for Coulombic external potentials (i.e., molecular systems) has been argued to exist¹⁷ and has motivated recent orbital-optimized DFT approaches to electronic excited states^{18,19}.

Alongside theoretical developments in excited-state DFT, there has been significant recent progress in using data-driven approaches to machine-learn ground-state density functionals^{20–28}. Related to the current work, there has also been progress in developing machine-learning models to directly predict electronic excited-state energies, gradients, and non-adiabatic couplings given a particular molecular geometry, and the SchNarc approach by Westermayr, Marquetand et al. has been demonstrated to predict non-adiabatic dynamics in agreement with ab initio simulations using a reasonable amount of training data^{29–35}. In the current work, we explore the proposition that a multistate density functional can be machine learned via an excited-state HK map. This new development establishes a framework that is potentially more powerful than directly learning excited-state energies, since densities and density functionals can yield any desired property.

In order to construct an ML excited-state density functional, we work in the framework introduced in ref. 22, which uses a data-driven

ML model with a physically motivated representation of the molecule. The fundamental object in this approach is a potential-to-density map, $n[\nu](\mathbf{r})$, which is machine-learned by computing the external potential and corresponding density at a desired level of theory for a range of nuclear geometries of the system and then inputting these functions as training data. For ground-state densities, this map is known as the machine-learned Hohenberg–Kohn (ML-HK) map; once learned, it can be used to compute any related quantities, including ground-state energies, and other observables at the same level of theory^{22,23}. It is worth noting that the level of theory is not restricted to DFT but can also be performed, for example, using coupled-cluster theory²³. In this work, we generalize the potential-to-density map to excited states in a multistate Hohenberg–Kohn framework (ML-MSHK), allowing us to learn excited-state densities and energies simultaneously with comparable accuracy to the ground state.

In order to test our ML-MSHK model, we consider the excited-state proton transfer (ESPT) reaction in malonaldehyde (MA), a small organic molecule exhibiting a non-trivial internal reaction that is sensitive to electronic excitation. ESPT is at the heart of photoacidity and is a key step in the photocycle of many photoactive proteins, such as green fluorescent protein (GFP)^{36,37}. Intramolecular proton transfer, as in MA, serves as a useful model to study ESPT, since the reaction rate is not limited by the diffusion of the proton donor and acceptor together, and can therefore be probed by ultrafast spectroscopy^{38,39}.

The structural simplicity of MA makes it an especially appealing model of ESPT. In particular, one conformation of malonaldehyde is a ring structure with an internal hydrogen bond that allows for a proton-transfer reaction between the two oxygen atoms (see molecular graphic in Fig. 1a). Excitation of this molecule from the S_0 ground state to the S_2 singlet excited state ($\pi\pi^*$) leads to a substantial reduction of the proton-transfer barrier from several kcal/mol in the ground state to an essentially barrierless reaction in the excited state. However, the presence of a lower-lying singlet $n\pi^*$ state in MA significantly complicates its ESPT reaction, and ultrafast non-adiabatic transitions from S_2 to S_1 are believed to compete with the ESPT reaction^{40–42}. The barrier to

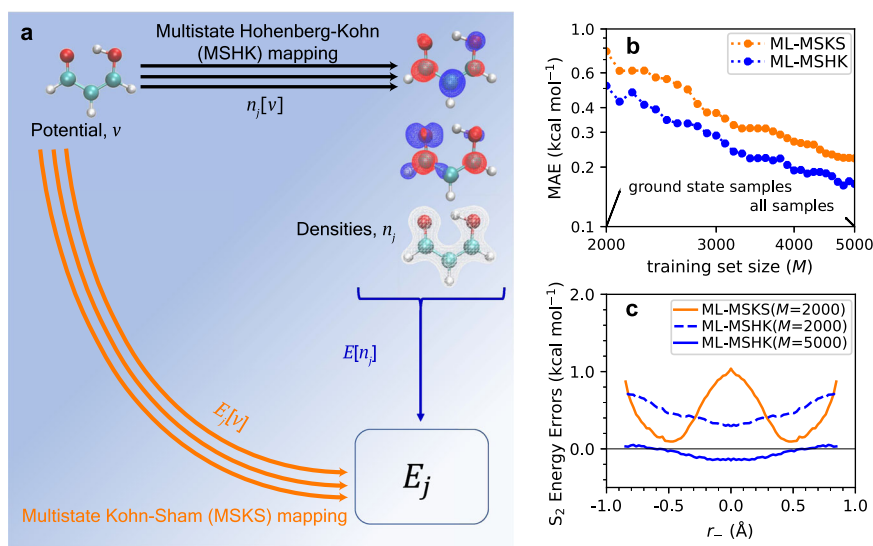


Fig. 1 | Overview and performance of the multistate machine-learning (ML) approach. **a** DFT maps used in this paper. The orange arrows represent a direct ML-MSKS map from the external potential, ν , to total electronic energies, E_j , for each state j . The black arrows represent ML-MSHK maps between ν and electron densities of each state, n_j . The blue arrow represents a single multistate energy functional that maps a density to its energy. The structure of the malonaldehyde (MA) molecule studied here is shown in the top left of this panel. **b** Learning curves (on a logarithmic scale), for MA's S_2 energy predictions from the ML-MSHK model (blue curve) using the lowest three densities (n_0, n_1, n_2) and ML-MSKS model (orange

curve). Training sets are formed starting from 2000 geometries sampled from S_0 dynamics and sequentially adding geometries extracted from ab initio molecular dynamics (AIMD) trajectories on the S_2 state in the manner described in Supplementary Note 1. **c** Errors of energy prediction against a TD-PBE0 benchmark along a minimum energy pathway for proton transfer of MA in the S_2 state predicted with ML-MSKS and ML-MSHK maps using training sets of different sizes. For each fixed value of proton-transfer coordinate, r_- (see text for definition), geometries were optimized on S_2 at the TD-PBE0/aug-cc-pvdz level, subject to a constraint of planarity.

proton transfer on S_1 is even higher than the ground-state^{40,43}, and furthermore, a three-state intersection is predicted to be energetically accessible⁴¹, meaning the electronic excitation can be efficiently quenched to the ground state. As a result, the ESPT reaction in MA is largely hindered by competing processes and has yet to be directly observed experimentally⁴².

One approach to suppressing the competing processes in MA and, thereby, controlling the ESPT reaction is to raise the $n\pi^*$ state energy via functional group modifications. For example, the $n\pi^*$ and $\pi\pi^*$ energetic ordering is reversed in methyl salicylate (MS)⁴⁴, which shares the chelate ring of MA yet exhibits efficient ESPT, as the non-adiabatic transitions to the $n\pi^*$ state are removed^{38,45,46}. In this work, using excited-state dynamics with our ML-MSHK method, we seek to explore whether mechanical restrictions applied to MA can instead be used to promote the ESPT reaction. Recognizing that the S_2/S_1 transition in MA is brought about by torsional motions⁴¹, we propose that a restraint of planarity on the molecule will allow the ESPT reaction to proceed unhindered. This restraint can be viewed as biomimetic of a similar mechanism in GFP whose protein environment enforces planarity of the 4-(p-hydroxybenzylidene)imidazolidin-5-one chromophore, preventing nonadiabatic transitions and enhancing its fluorescence quantum yield⁴⁷.

Results

Theory

There is a growing recognition of the importance of using physics-based principles in choosing descriptors and models for machine-learning molecular properties^{48,49}. The blessing, and perhaps curse, of ML models is their great flexibility to predict almost any quantity given sufficient data. This can lead to a loss of physical insight, and concerns of overfitting must always be addressed. By using physical principles, fundamental constraints are automatically introduced into the model that mitigate these problems. In the context of learning electronic structure, a very natural physical framework relies on the connections between the electron density, $n(\mathbf{r})$, external potential, $v(\mathbf{r})$, and energy, E , of a many-body quantum system. These connections are formalized by the HK1 theorem, which proves the existence of a bijection from the ground-state electron density to the external potential of a many-body system: $n_0(\mathbf{r}) \leftrightarrow v(\mathbf{r})$ ¹⁴. The electron density, which gives the probability for finding an electron at a certain location in space, is a particularly convenient quantity to work with since, unlike the many-body wavefunction, it is a three-dimensional scalar field regardless of the number of electrons and is, therefore, straightforward to represent numerically. For molecular systems, the external potential is simply the Coulombic scalar potential arising from the nuclear charges: $v(\mathbf{r}) = -\sum_a^N Z_a/|\mathbf{r} - \mathbf{r}_a|$. Since the external potential uniquely defines the molecular electronic Hamiltonian (for a given spin state and number of electrons), the electronic energy must also be uniquely determined from knowledge of the external potential or, per HK1, the density¹⁴. These maps from external potential to density and energy are realized as functionals, e.g., $n[v](\mathbf{r})$, which inspired the framework for a previous approach that machine-learned density functionals for ground-state energy predictions (ML-HK)²². Part of the success of the ML-HK approach can be attributed to the uniqueness of the external potential and density as molecular descriptors, a property not shared by some other choices⁵⁰.

We now demonstrate how these ideas can be extended to electronic excited states. As mentioned in the introduction, an excited-state extension of HK1 has been argued to exist for Coulombic external potentials¹⁷, meaning that the map from external potential to each state's density is encoded as a functional, $n_j[v](\mathbf{r})$, where $j = 0, 1, 2, \dots$ labels each excited state. Furthermore, a property of the (unknown) exact universal functional is that each extremal density of the energy functional corresponds to an electronic eigenstate with the functional returning the exact (excited-state) eigenvalue, $E_j = E[n_j]$ ⁵¹.

This motivates a machine-learning approach that leverages an excited-state Hohenberg–Kohn mapping.

Learning excited states with ML-MSHK

The first step of our ML model is to learn multiple electronic state densities. We start by expanding the densities in an orthonormal basis set, $\phi_i(\mathbf{r})$, as $n_j^{\text{ML}}[v] = \sum_{i=1}^L u_j^{(i)}[v] \phi_i(\mathbf{r})$, and we learn the set of basis coefficients $\{u_j^{(i)}\}$ by training on a set of input potentials corresponding to different geometries of a system. We allow for the learning of ground- and excited-state densities with a unique map to each state j . Various ML models, including artificial neural networks and kernel methods, have been used to learn total energies from electronic-structure calculations^{22,23,30,52–59}. However, learning excited-state HK maps is unique to our approach, and we follow the ground-state ML-HK functional's successful use of the kernel ridge regression (KRR) method⁶⁰. In principle, the functionals could instead be learned with neural networks, but a comparison of different methods is beyond the scope of this manuscript.

With a set of learned densities, a second KRR model is used to learn ground- or excited-state energies from a set of input training densities. We consider two types of energy functionals, depending on the nature of the input used. The first relies on the proven existence of a single functional that maps ground or excited-state densities to their respective energy eigenvalues⁵¹. We thus learn a single map from density to energy and use multiple states in the training of this map:

$$E^{\text{ML-MSHK}}[n_j] = \sum_{i=1}^M \sum_k \alpha_{i,k} K(\mathbf{u}_{i,j}[v], \mathbf{u}_{i,k}[v_i]), \quad (1)$$

where $K(\mathbf{u}_{i,j}[v], \mathbf{u}_{i,k}[v_i])$ is the kernel, k runs over the states used in the training of the model, and $\{\alpha_{i,k}\}$ are the coefficients learned in the energy functional model. Since energy predictions from this model rely on training with multiple electron densities for a given molecular geometry, we call the combination of the two types of maps the multistate Hohenberg–Kohn approach (ML-MSHK). Learning Eq. (1) comes at the cost of retaining multiple densities for each training example, increasing storage needs; however, the advantage is the resulting energy functional is not state specific. We note that while this work focuses on the planar S_2 state, we also train on densities from the S_1 excited state to demonstrate that including additional densities does not degrade the performance of the ML-MSHK model.

The second energy functional we consider is of the external potential itself, without using the density as an intermediate descriptor. Following the naming of a similar approach for ground-state energies²², we call this the multistate Kohn–Sham map (ML-MSKS):

$$E_j^{\text{ML-MSKS}}[v] = \sum_{i=1}^M \gamma_{i,j} K[v_i, v], \quad (2)$$

where $\{\gamma_{i,j}\}$ are the coefficients learned in the ML-MSKS energy functional model. Such functionals must also exist since the external potential uniquely defines the molecular electronic Hamiltonian and, therefore, also its eigenstates¹⁴. Eq. (2) has the advantage that it uses only external potentials rather than densities, reducing computational storage requirements compared to Eq. (1). However, a disadvantage is that the energy maps in this model are state specific, which we expect will introduce errors near electronic crossings³⁰.

In all cases, models are trained against ab initio densities and energies following the procedure described in section “Methods”. A schematic of the different maps used in this study is shown in Fig. 1a.

Excited-state energy predictions

In order to test the performance of the excited-state machine-learned density functionals, we start by considering learning curves for

training S_2 excited-state energies of MA, shown in Fig. 1b, in which the lowest three densities (n_0, n_1, n_2) were used for the ML-MSHK model. The smallest training set considered contains 2000 molecular conformations of MA generated from a ground-state AIMD trajectory, propagated as described in section “Training and test set generation”. A training set similar to this was previously found to yield a ML-HK potential that accurately described the proton-transfer reaction on the ground state²². From Fig. 1b, however, we see somewhat large out-of-sample mean absolute errors (MAEs) of 0.5 and 0.8 kcal/mol for the ML-MSHK and ML-MSKS models, respectively, with this training set. Furthermore, for the ML-MSKS model, the error is seen to vary strongly with geometry, as revealed by the orange curve in Fig. 1c, which shows energy prediction errors along a minimal-energy pathway for the ESPT reaction as a function of the proton-transfer coordinate, $r_- = r_{\text{HO1}} - r_{\text{HO2}}$, where $r_{\text{HO}i}$ is the distance of the proton from oxygen atom i . Using the same 2000 training geometries, the ML-MSHK error (dashed blue curve) is more uniform than the ML-MSKS predictions, demonstrating the advantage of the former.

In order to use the ML-MSHK model for excited-state dynamics, we must further reduce its prediction errors; however, we found essentially no improvement upon adding more samples from the ground-state trajectory. As we will see below, the remaining source of error in models trained only with ground-state samples is that the S_2 -initiated ESPT involves nuclear responses also in modes orthogonal to the direct proton transfer mode, and these are not adequately sampled by the ground-state trajectory. Thus, we extended the training set by including geometries extracted from 30 S_2 excited-state AIMD trajectories, in a manner described in Supplementary Note 1, yielding the learning curves shown in Fig. 1b.

After including geometries extracted from excited-state trajectories in the training (ground and excited-state samples), both ML models perform significantly better than the ML model constructed using only ground-state samples in the training. Interestingly, the ML-MSKS map (Fig. 1b, orange curve) always performs worse than the MSHK map (blue curve). Similar behavior was noted in the ground-state ML-HK study of ref. 22. The rest of this work pertains to the ML-HK models since learning the densities as an intermediate step yields the models that outperform the direct potential-to-energy approach, and the inclusion of density learning has other benefits such as the capability of density-based delta-learning to add corrections from high-level wavefunction theory²³. Encouragingly, the out-of-sample error for ML-MSHK converges below 0.2 kcal/mol once the sample size reaches $M=5000$. This error is comparable to that found in the previous ground-state study of MA²², suggesting that ML-MSHK will be suitable for molecular dynamics simulations (confirmed below). Having found convergence at $M=5000$, our final training set generation follows the protocol described in section “Training and test set generation”. For the adiabatic excited-state dynamics considered in this work, it is possible to train state-specific ML-HK models to return the S_2 energies using only ground-state or S_2 densities as input (see Supplementary Fig. 1). While we do not include any non-adiabatic transitions in this study, we note that the state-specific models are worse than ML-MSHK in the vicinity of electronic crossings (see Supplementary Note 5). Given the excellent performance of the ML-MSHK model and its ability to capture multiple electronic states simultaneously, we focus exclusively on this model for the remainder of this work.

Excited-state density predictions

Equally important to energy predictions is the question of how well the model reproduces electron densities. This is explored in Fig. 2, which shows electron densities for the C_s -symmetric ground-state optimized geometry of MA (similar results for the proton transfer transition state are shown in Supplementary Fig. 4). In order to highlight the changes in bonding structure, excited-state densities are displayed as density differences with respect to the ground state. For all states considered,

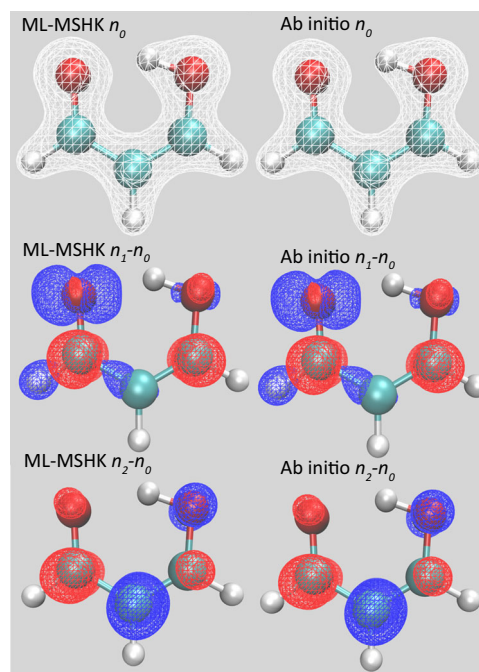


Fig. 2 | Ground and excited-state electron densities for the C_s ground-state minimum energy structure of MA. 1st row: S_0 ground-state densities. 2nd row: density differences between S_1 and S_0 . 3rd row: density differences between S_2 and S_0 . An isosurface of $0.1 e/\text{Bohr}^3$ was used for plotting densities and density differences. Left column: ML predictions, right column: ab initio TD-PBE0 predictions. Each density is represented by an isosurface plot. For density differences, red means an accumulation of electronic charge in the excited state and blue means a depletion.

there is no discernible difference between the ML-MSHK predicted densities (left) and ab initio TD-PBE0 densities (right), even though the ground-state optimized structure was not included in the training set. Indeed, the out-of-sample integrated MAE in the total electron densities is very small ($0.012 e$).

Knowledge of the density differences also provides insight into the nature of the electronic excitations. From Fig. 2, S_1 clearly corresponds to an $n\pi^*$ transition with a depletion of density from the proton-accepting oxygen's p_x orbital and an accumulation on the backbone π^* orbital. S_2 , on the other hand, is a $\pi\pi^*$ transition, as identified by a depletion of density on the central conjugated carbon's p_z orbital.

Excited-state MD using ML potentials

Having demonstrated that our ML-MSHK model provides a quantitative prediction of excited-state electron densities and total energies within our test set of geometries, we next consider its ability to generate a functional for use in excited-state MD. We thus initiated non-equilibrium MD trajectories of MA on its S_2 state following vertical excitation from the ground state. 1000 ML-MSHK and 50 AIMD trajectories were run and compared. It should be noted that the computational effort associated with propagating dynamics on the ML-MSHK surface is negligible compared to AIMD dynamics. Thus, for the cost of 50 excited-state AIMD trajectories used to generate the training and test sets, we gain the ability to perform excited-state dynamics with 1000 ML-MSHK trajectories using the resulting machine-learned functional to converge the excited-state non-equilibrium averaged properties.

We consider three metrics to assess the quality of the machine-learned functional. The first concerns how well ML-MSHK is able to interpolate the energy between geometries explicitly included in the training. This is accomplished by evaluating ML-MSHK energies along

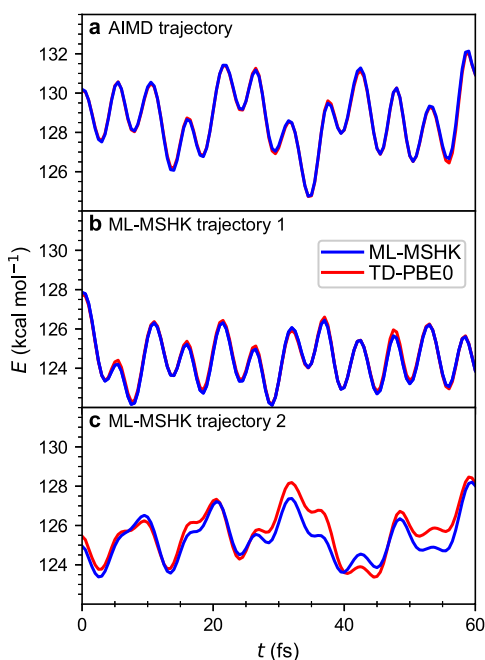


Fig. 3 | Predictions of electronic energies along trajectories. ML-MSHK predictions (blue curves) and TD-PBE0 reference values (red curves) are shown along three different representative trajectories taken from: **a** an ab initio molecular dynamics (AIMD) trajectory and **b** an ML-generated excited-state trajectory, both initiated from geometries in the training set, and **c** an ML-generated excited-state trajectory initiated from a geometry out of the training set.

an AIMD excited-state trajectory from which some geometries were included in the training. The result is shown in Fig. 3a, where we see almost perfect agreement between ML-MSHK and the AIMD TD-PBE0 energies. Note that only 32 geometries out of 481 from this trajectory were used in the training set of the ML-MSHK model, highlighting the fidelity to which ML-MSHK is able to interpolate the functional between training points. Second, we assess how well ML-MSHK propagates the dynamics starting from a geometry included in the training set, shown in panel b, where again we see almost perfect agreement with AIMD energies evaluated on the ML trajectory snapshots. Finally, we check how well the machine-learned functional does on arbitrary geometries by evaluating ab initio TD-PBE0 energies along an ML-MSHK trajectory where no geometry was included in the training set, shown in panel c. Here, slightly larger deviations between ML-MSHK and TD-PBE0 energies are seen; however, the difference is within the expected MAE bounds from the learning curve and does not grow with time, showing that ML-MSHK faithfully reproduces the excited-state energy functional, at least in the configurational space sampled during non-equilibrium dynamics. The error could be reduced by further increasing the training set; however, we found this unnecessary since this small error does not influence the dynamical predictions as we show below.

ESPT in planar MA

Having seen that ML-MSHK provides an accurate description of the S_2 excited-state functional of MA, we next consider the predictions of excited-state dynamics of the ESPT reaction using this functional. To that end, we compute the following non-equilibrium response function:

$$S(t) = \overline{(r_-(0)r_-(t))/|r_-(0)r_-(t)|}, \quad (3)$$

where $t=0$ corresponds to the instance of vertical excitation from an equilibrium configuration on the S_0 state to the S_2 state, and the

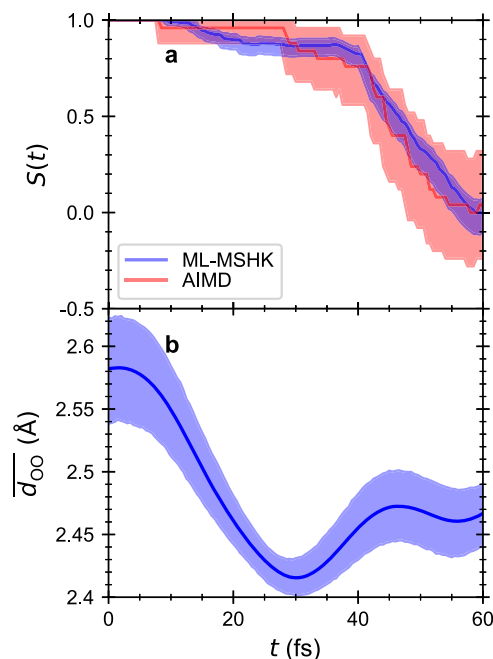


Fig. 4 | The non-equilibrium response of the excited-state proton transfer reaction and its relationship to the time varying oxygen–oxygen distance of MA. **a** Non-equilibrium response function of the proton's location (reactant or product basin according to Eq. (3)), $S(t)$, that reflects the progress of the excited-state proton transfer reaction. The solid curves are the expectation values of $S(t)$ and the shaded area shows the 95% confidence intervals estimated from bootstrap analysis using 50 trajectories for ab initio molecular dynamics (AIMD) (red) and 1000 trajectories (blue) for ML-MSHK. The converged $S(t)$ of the ML-MSHK model shows good agreement with the result from AIMD simulations. **b** The average oxygen–oxygen distance is shown as a function of time following photoexcitation.

overbar indicates a non-equilibrium average over initial conditions. $S(t)$ thus measures the memory of which reaction basin the proton is in (reactant or product). A value of $S=1$ means the proton is in the reactant basin, a value of $S=-1$ means the proton is in the product basin, and a value of $S=0$ means memory has been lost. Since MA has a symmetric proton donor and acceptor, $S=0$ indicates the reaction is complete.

Figure 4a shows the ESPT reaction's non-equilibrium response function, $S(t)$, computed for 50 AIMD trajectories (red) and 1,000 ML-MSHK trajectories (blue). Shaded areas represent the 95% confidence intervals computed with the bootstrap method⁶¹, as implemented in SciPy v. 1.7.1⁶². The results show that the AIMD and ML-MSHK agree within statistical certainty, highlighting the predictive power of ML-MSHK. Furthermore, an interesting feature of $S(t)$ becomes clear in the ML-MSHK results: apart from a ~10% fraction of trajectories that undergo almost immediate proton transfer, there is a waiting period of ~40 fs before the remainder of ESPT reaction proceeds, which is then rapidly completed by 60 fs. This behavior is hinted at in the AIMD results but becomes much more apparent in the ML-MSHK predictions due to averaging over many more trajectories.

The observed 40-fs waiting period arises from important responses of the heavy atoms following excitation to S_2 . This can be seen for example in the average oxygen–oxygen distance, $\overline{d_{OO}}$, plotted in Fig. 4b, which exhibits a decrease from ~2.57 Å to ~2.42 Å in the first 30 fs following excitation. This timescale is comparable to the waiting period of the ESPT reaction, suggesting that the oxygen–oxygen motion gates the reaction, which does not proceed rapidly until a distance of $d_{OO} \leq 2.45$ Å is reached.

The response of d_{OO} following excitation to S_2 can be understood from changes in the electron density shown in Fig. 2. The reduction in π -bonding character lessens the rigidity of the conjugated system,

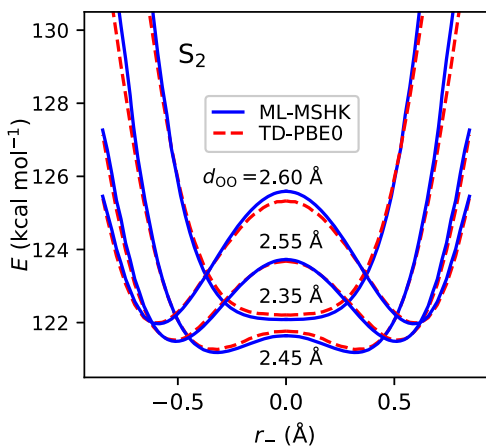


Fig. 5 | S_2 potential energy surfaces of MA along the H-transfer coordinate, r_- , for a series of fixed oxygen–oxygen distances, d_{oo} . All coordinates except r_- and d_{oo} are optimized at the TD-PBE0/aug-cc-pvdz level for S_2 , subject to a constraint of planarity, following the procedure described in Supplementary Note 3. We see excellent agreement between ab initio results (dashed red curves) and ML-MSHK predictions (solid blue curves).

allowing the chelating ring to contract and the oxygen atoms to become closer. Bringing together the proton donor and acceptor oxygen atoms then reduces the barrier to ESPT, thus explaining the gating mechanism. This is explored explicitly in Fig. 5, which plots the potential energy profile of the minimum energy pathway of ESPT on S_2 for a series of fixed d_{oo} distances. Encouragingly, quantitative agreement between ML-MSHK (blue curves) and TD-PBE0 (dashed red curves) is seen, despite none of the minimum energy pathway geometries being included explicitly in the training set. For the initial average d_{oo} distance (-2.57 Å), a proton-transfer barrier of >2 kcal/mol is seen, explaining why only a small fraction of excited trajectories undergo ESPT in the first 20 fs. The S_2 potential is downhill for motions that reduce d_{oo} to a value of 2.45 Å (following the minimum of the potential energy profile as a function of d_{oo}), which also brings about a reduction of the barrier to proton transfer, having a value of 0.3 kcal/mol at $d_{oo} = 2.45$ Å. The barrier disappears completely for $d_{oo} = 2.35$ Å; however, this oxygen–oxygen distance is disfavored due to an overall increase in the potential, explaining why the initial reduction in d_{oo} is seen to reverse at $t = 30$ fs in Fig. 4b.

Discussion

In this paper, we developed a multistate Hohenberg–Kohn machine-learning framework for the accurate prediction of excited-state densities and energies. We demonstrated that excited-state energies, which are expressible as a single functional of multiple state densities, can be learned as accurately as the ground-state energy, even for the moderately sized molecule malonaldehyde, which exhibits a non-trivial excited-state proton transfer reaction. The resulting machine-learned functional is faithful to the underlying excited-state AIMD trajectories on which the model was based, allowing accurate non-equilibrium excited-state molecular dynamics trajectories to be performed with a 10-fold computational saving compared to AIMD. Furthermore, our study yielded new insight into the excited-state proton transfer reaction in malonaldehyde. This was aided by the low-cost of ML-MSHK energy predictions that allowed us to converge the non-equilibrium dynamics with 1,000 excited-state trajectories, revealing the 40-fs waiting period following photoexcitation, after which ESPT promptly completed.

The observed gating mechanism of proton transfer in planar MA by heavy atom motion has been seen previously in other intramolecular ESPT reactions which are not precluded by ultrafast non-adiabatic transitions. In particular, the lack of isotope effect for ESPT

in methyl salicylate was explained by a delocalization of the reaction coordinate to modes other than the donor OH stretch³⁸. Later studies in related ESPT molecules found the reaction path was dominated by anharmonic low-frequency backbone modes that modulated the donor-acceptor distance^{63–65}. Thus, by enforcing planarity in MA, we have suppressed nonadiabatic transitions that compete with ESPT and find the mechanism of proton transfer is indeed multidimensional in nature and follows the consensus picture that has emerged for intramolecular chelate ring structures. That we correctly captured the multidimensional nature of ESPT from a limited training of excited-state trajectories is a testament to the power of our ML-MSHK method.

This work can be extended in a number of directions. As an example, we removed the restraint of planarity on malonaldehyde (see Supplementary Note 5). Models using additional non-planar training geometries show that the multistate density functional outperforms the state-specific functionals near S_2/S_1 state crossings, which is vital for running future non-adiabatic MD trajectories. In addition, the LR-TDDFT training data could be replaced with high-level wavefunction-based electronic structure for input to train the ML-MSHK density and energy maps. A similar approach was recently shown to be successful for the ground-state ML-HK method to learn coupled-cluster energies²³. Based on the promising results presented in this work, we are optimistic that ML-MSHK will find practical use in excited-state non-adiabatic simulations with the costly electronic-structure steps done only once to build the training set.

Methods

Machine-learning model

In order to predict the excited-state energies of a molecule using electron densities via the ML-MSHK map introduced in section “Learning excited states with ML-MSHK”, we start by using a basis expansion of the densities:

$$n_j^{\text{ML}}[v](\mathbf{r}) = \sum_{l=1}^L u_j^{(l)}[v]\phi_l(\mathbf{r}), \quad (4)$$

where j indexes an electronic state ($j = 0$ is the ground state, $j = 1$ is the first excited state, etc). Following the ground-state ML-HK method²², we choose a Fourier basis. In Eq. (4), l indexes a basis function, of which there are L . In this work 50 functions are used in each dimension ($125,000 = 50 \times 50 \times 50$ in total). The coefficients of the basis set expansion are learned using KRR. In particular, $u_j^{(l)}$ are represented as a kernel expansion of the form

$$u_j^{(l)}[v] = \sum_{i=1}^M \beta_{ij}^{(l)} \kappa[v_i, v]. \quad (5)$$

Here, $\beta_{ij}^{(l)}$ parameterizes the kernel model, and the kernel functional κ has a Gaussian form

$$\kappa[v_i, v] = \exp\left[-\frac{1}{\sigma^2} \int d\mathbf{r} (v(\mathbf{r}) - v_i(\mathbf{r}))^2\right], \quad (6)$$

where σ is a kernel width hyperparameter. The external potentials, $v_i(\mathbf{r})$, $i = 1, \dots, M$, are unique to each of the M training geometries, and are paired with M training densities, $n_i(\mathbf{r})$, for each state j . Like the density, we use a basis representation of the external potential; however, care must be taken to avoid the Coulomb singularities in $v(\mathbf{r})$. We therefore use a Gaussian representation of the potential:

$$v(\mathbf{r}) = \sum_a^N Z_a e^{-|\mathbf{R}_a - \mathbf{r}|^2 / 2\sigma_{\text{pot}}^2}, \quad (7)$$

where \mathbf{R}_a is the position of the a th nucleus, Z_a is its corresponding nuclear charge, and σ_{pot} is a width parameter^{22,66}. The potential from

Eq. (7) is then formed on a 3D grid surrounding the molecule and stored as a vector, \mathbf{v}_i , for each sample i to be used in Eq. (6). $60 \times 50 \times 30$ grid points were used with a spacing of 0.2 Å, commensurate with the shape of the molecule. We used previously optimized values of $\sigma_{\text{pot}} = 0.2$ Å and a grid spacing for MA from ref. 22.

As a final step, we train a multistate total energy functional:

$$E^{\text{ML-MSHK}}[n_j] = \sum_{i=1}^M \sum_k \alpha_{i,k} \kappa[\mathbf{u}_{i,k}, \mathbf{u}_j], \quad (8)$$

where k sums over the states of the densities used in the training and κ is another Gaussian kernel.

Reference ab initio electronic structure

Reference calculations on MA's excited states were performed at the LR-TDDFT level using the PBE0^{67,68} approximate exchange and correlation functional. This level of theory, which we abbreviate as TD-PBE0, was chosen as a compromise between accuracy and efficiency in this first demonstration of our ML-MSHK method. In particular, PBE0 provides qualitatively correct ground-state thermochemical properties of MA, yielding proton transfer barrier height of 2.0 kcal/mol compared to the predicted barrier from high-level theory of 4.1 kcal/mol^{69,70}. In addition, TD-PBE0 provides reasonable spectroscopic quantities for MA, with a S_0 - S_2 vertical excitation of 5.3 eV compared to the experimental value of 4.7 eV⁷¹. Electronic structure calculations were performed in CPMD v. 4.3.0⁷² using a plane-wave basis with a kinetic energy cutoff of 90 Rydberg. Core electrons were replaced with Troullier-Martins norm-conserving pseudopotentials⁷³.

AIMD excited-state molecular dynamics

Ab initio excited-state Born-Oppenheimer MD simulations of a gas-phase malonaldehyde molecule were performed in CPMD using the same TD-PBE0 level of theory discussed above. 50 independent non-equilibrium trajectories were initiated on the S_2 state following vertical excitations spaced every 100 fs from an AIMD ground-state trajectory sampled at 300 K taken from ref. 22. Excited-state dynamics were propagated in the microcanonical ensemble with a 0.25 fs timestep for 120 fs, which takes 5027 min for one simulation with all cores (24) of dual Intel Xeon E5 2650 v4 (2.2 GHz) CPUs. Planarity was maintained with the following added restraining potential: $V^{\text{rest.}} = \sum_a^N \frac{1}{2} k_a (z_a)^2$, where k_a is the strength of each restraint and is 9 kcal/(mol bohr²) for hydrogen and 40 kcal/(mol bohr²) for heavy atoms. z_a is the z-coordinate of each atom. The planarity restraints were chosen to be stiffer for the heavy atoms, since backbone torsions have been identified as bringing about the S_2/S_1 electronic crossing in MA⁴¹.

Training and test set generation

It has been shown that exploiting molecular point group symmetries helps to increase the effective dataset size significantly without performing additional quantum chemical calculations²³. We thus generated training sets for malonaldehyde that included reflection about the mirror symmetry plane perpendicular to the molecular axis, σ_{ν} , of an idealized planar C_{2v} structure, since this symmetry operation ensures the equivalence of the two oxygen atoms as proton donors/acceptors. To achieve a symmetrization of the training sets while avoiding unnecessary duplication of samples, we combined symmetrization with a clustering of molecular geometries. For example, the training set for MD runs was formed by starting with 2000 structures taken from a previous ground-state AIMD trajectory²² and 14,400 structures taken every timestep from 30 independent ab initio excited-state trajectories described in section "AIMD excited-state molecular dynamics". Next, all geometries were aligned to a reference geometry representing the planar proton-transfer transition state with C_{2v} symmetry (see Supplementary Note 4). Then, following alignment,

any molecular structures with a negative proton-transfer coordinate, r_- , were reflected in the σ_{ν} plane, i.e., to ensure $r_- > 0$. The training set was then clustered with K -Means⁷⁴ to make a set of 2500 samples, using a metric related to the L^2 deviation of the external potential,

$$\sqrt{\int d\mathbf{r} (v_i(\mathbf{r}) - v_j(\mathbf{r}))^2},$$

where v_i represents the potential from Eq. (7). Finally, the training set was doubled in size by applying the reflection operator in the σ_{ν} plane to yield the production training set of 5000 geometries. The test set contains 240 aligned snapshots extracted every 10 fs from 20 independent ab initio excited-state trajectories described in section "AIMD excited-state molecular dynamics".

The electronic-structure calculations for excited-state energies and densities take 40 min with 8 cores of an Intel Xeon E5 2650 v4 (2.2 GHz) CPU for one geometry and the training takes 10 min using the same number of cores with hyperparameters given.

ML excited-state molecular dynamics

To generate initial conditions to perform excited-state dynamics on the ML S_2 potential, a ground-state AIMD trajectory was run in CPMD using the PBE exchange and correlation functional⁶⁷ with the same kinetic energy cutoff and pseudopotentials discussed above. This trajectory was run in the canonical ensemble at 300 K using massive Nosé-Hoover chain thermostats with a timestep of 0.5 fs⁷⁵. 1000 independent non-equilibrium trajectories were initiated on the S_2 state following vertical excitations spaced every 100 fs from the ground-state trajectory. Excited-state dynamics were propagated in the microcanonical ensemble with a 0.25 fs timestep using the atomistic simulation environment v.3.19.0⁷⁶. Atomic forces were evaluated numerically using central differences with a step length of $dx = 0.001$ Å. Dynamics were propagated for 60 fs, which was sufficient to observe the ESPT reaction. The propagation of one 60-fs dynamics simulation takes 20 min with 2 cores of an Intel Xeon E5 2650 v4 (2.2 GHz) CPU.

Data availability

The molecular coordinates, electronic densities and energies for ground and excited states used in this study are available in Zenodo [<https://doi.org/10.5281/zenodo.7064211>]. The source data of the figures are provided in the Source Data file. Source data are provided with this paper.

Code availability

Codes used in this work are available at <https://doi.org/10.5281/zenodo.7064211>.

References

- Meech, S. R., Hoff, A. J. & Wiersma, D. A. Role of charge-transfer states in bacterial photosynthesis. *Proc. Natl. Acad. Sci. USA* **83**, 9464 (1986).
- Wahadoszamen, M., Margalit, I., Ara, A. M., van Grondelle, R. & Noy, D. The role of charge-transfer states in energy transfer and dissipation within natural and artificial bacteriochlorophyll proteins. *Nat. Commun.* **5**, 5287 (2014).
- Simons, J. How do low-energy (0.1–2 eV) electrons cause DNA-strand breaks? *Acc. Chem. Res.* **39**, 772 (2006).
- Alizadeh, E., Sanz, A. G., García, G. & Sanche, L. Radiation damage to DNA: The indirect effect of low-energy electrons. *J. Phys. Chem. Lett.* **4**, 820 (2013).
- Benov, L. Photodynamic therapy: current status and future directions. *Med. Princ. Pract.* **24**, 14 (2015).
- Lerch, M. M., Hansen, M. J., van Dam, G. M., Szymanski, W. & Feringa, B. L. Emerging targets in photopharmacology. *Angew. Chem. Intl. Ed.* **55**, 10978 (2016).
- Hüll, K., Morstein, J. & Trauner, D. In vivo photopharmacology. *Chem. Rev.* **118**, 10710 (2018).

8. Clarke, T. M. & Durrant, J. R. Charge photogeneration in organic solar cells. *Chem. Rev.* **110**, 6736 (2010).
9. Zhao, Y. & Liang, W. Charge transfer in organic molecules for solar cells: Theoretical perspective. *Chem. Soc. Rev.* **41**, 1075 (2012).
10. Virshup, A. M., Chen, J. & Martínez, T. J. Nonlinear dimensionality reduction for nonadiabatic dynamics: the influence of conical intersection topography on population transfer rates. *J. Chem. Phys.* **137**, 22A519 (2012).
11. Pieri, E. et al. The non-adiabatic nanoreactor: towards the automated discovery of photochemistry. *Chem. Sci.* **12**, 7294 (2021).
12. Gross, E. & Kohn, W. in *Density Functional Theory of Many-Fermion Systems, Vol. 21 Advances in Quantum Chemistry* (ed. Löwdin, P.-O.) 255–291 (Academic Press, 1990).
13. Casida, M. E. in *Recent Advances in Density Functional Methods* (ed. Chong, D. P.) 155–192 (World Scientific Publishing, 1995).
14. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
15. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
16. Gaudoin, R. & Burke, K. Lack of Hohenberg-Kohn theorem for excited states. *Phys. Rev. Lett.* **93**, 173001 (2004).
17. Ayers, P. W., Levy, M. & Nagy, A. Time-independent density-functional theory for excited states of Coulomb systems. *Phys. Rev. A* **85**, 042518 (2012).
18. Hait, D. & Head-Gordon, M. Excited state orbital optimization via minimizing the square of the gradient: general approach and application to singly and doubly excited states via density functional theory. *J. Chem. Theory Comput.* **16**, 1699 (2020).
19. Hait, D. & Head-Gordon, M. Orbital optimized density functional theory for electronic excited states. *J. Phys. Chem. Lett.* **12**, 4517 (2021).
20. Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R. & Burke, K. Finding density functionals with machine learning. *Phys. Rev. Lett.* **108**, 253002 (2012).
21. Mardirossian, N. & Head-Gordon, M. ω B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **16**, 9904 (2014).
22. Brockherde, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **8**, 872 (2017).
23. Bogojeski, M., Vogt-Maranto, L., Tuckerman, M. E., Müller, K.-R. & Burke, K. Quantum chemical accuracy from density functional approximations via machine learning. *Nat. Commun.* **11**, 5223 (2020).
24. Dick, S. & Fernandez-Serra, M. Machine learning accurate exchange and correlation functionals of the electronic density. *Nat. Commun.* **11**, 3509 (2020).
25. Nagai, R., Akashi, R. & Sugino, O. Completing density functional theory by machine learning hidden messages from molecules. *npj Comput. Mater.* **6**, 43 (2020).
26. Cuevas-Zuñiría, B. & Pacios, L. F. Analytical model of electron density and its machine learning inference. *J. Chem. Inf. Model.* **60**, 3831 (2020).
27. Gedeon, J. et al. Machine learning the derivative discontinuity of density-functional theory. *Mach. Learn. Sci. Technol.* **3**, 015011 (2021).
28. Cuevas-Zuñiría, B. & Pacios, L. F. Machine learning of analytical electron density in large molecules through message-passing. *J. Chem. Inf. Model.* **61**, 2658 (2021).
29. Dral, P. O., Barbatti, M. & Thiel, W. Nonadiabatic excited-state dynamics with machine learning. *J. Phys. Chem. Lett.* **9**, 5660 (2018).
30. Westermayr, J. & Marquetand, P. Machine learning for electronically excited states of molecules. *Chem. Rev.* **121**, 9873 (2020).
31. Hu, D., Xie, Y., Li, X., Li, L. & Lan, Z. Inclusion of machine learning kernel ridge regression potential energy surfaces in on-the-fly nonadiabatic molecular dynamics simulation. *J. Phys. Chem. Lett.* **9**, 2725 (2018).
32. Westermayr, J. et al. Machine learning enables long time scale molecular photodynamics simulations. *Chem. Sci.* **10**, 8100 (2019).
33. Westermayr, J., Gastegger, M. & Marquetand, P. Combining SchNet and SHARC: The SchNarc machine learning approach for excited-state dynamics. *J. Phys. Chem. Lett.* **11**, 3828 (2020).
34. Li, J., Stein, R., Adrion, D. M. & Lopez, S. A. Machine-learning photodynamics simulations uncover the role of substituent effects on the photochemical formation of cubanes. *J. Am. Chem. Soc.* **143**, 20166 (2021).
35. Li, J. et al. Automatic discovery of photoisomerization mechanisms with nanosecond machine learning photodynamics simulations. *Chem. Sci.* **12**, 5302 (2021).
36. Chattoraj, M., King, B. A., Bublitz, G. U. & Boxer, S. G. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proc. Natl. Acad. Sci. USA* **93**, 8362 (1996).
37. Tsien, R. Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509 (1998).
38. Herek, J. L., Pedersen, S., Bañares, L. & Zewail, A. H. Femtosecond real-time probing of reactions. IX. Hydrogen-atom transfer. *J. Chem. Phys.* **97**, 9046 (1992).
39. Formosinho, S. J. & Arnaut, L. G. Excited-state proton transfer reactions II. Intramolecular reactions. *J. Photochem. Photobiol. A* **75**, 21 (1993).
40. Sobolewski, A. L. & Domcke, W. Photophysics of malonaldehyde: An ab initio study. *J. Phys. Chem. A* **103**, 4494 (1999).
41. Coe, J. D. & Martínez, T. J. Ab initio molecular dynamics of excited-state intramolecular proton transfer around a three-state conical intersection in malonaldehyde. *J. Phys. Chem. A* **110**, 618 (2006).
42. List, N. H., Dempwolff, A. L., Dreuw, A., Norman, P. & Martínez, T. J. Probing competing relaxation pathways in malonaldehyde with transient X-ray absorption spectroscopy. *Chem. Sci.* **11**, 4180 (2020).
43. do Monte, S. A., Dallos, M., Müller, T. & Lischka, H. MR-CISD and MR-AQCC calculation of excited states of malonaldehyde: geometry optimizations using analytical energy gradient methods and a systematic investigation of reference configuration sets. *Collect. Czech. Chem. Commun.* **68**, 447 (2003).
44. Aquino, A. J. A., Lischka, H. & Hättig, C. Excited-state intramolecular proton transfer: a survey of TDDFT and RI-CC2 excited-state potential energy surfaces. *J. Phys. Chem. A* **109**, 3201 (2005).
45. Coe, J. D., Levine, B. G. & Martínez, T. J. Ab initio molecular dynamics of excited-state intramolecular proton transfer using multireference perturbation theory. *J. Phys. Chem. A* **111**, 11302 (2007).
46. Coe, J. D. & Martínez, T. J. Ab initio multiple spawning dynamics of excited state intramolecular proton transfer: the role of spectroscopically dark states. *Mol. Phys.* **106**, 537 (2008).
47. Park, J. W. & Rhee, Y. M. Electric field keeps chromophore planar and produces high yield fluorescence in green fluorescent protein. *J. Am. Chem. Soc.* **138**, 13619 (2016).
48. Bereau, T., DiStasio, R. A., Tkatchenko, A. & von Lilienfeld, O. A. Non-covalent interactions across organic and biological subsets of chemical space: physics-based potentials parametrized from machine learning. *J. Chem. Phys.* **148**, 241706 (2018).
49. Zubatiuk, T. & Isayev, O. Development of multimodal machine learning potentials: toward a physics-aware artificial intelligence. *Acc. Chem. Res.* **54**, 1575 (2021).

50. von Lilienfeld, O. A., Ramakrishnan, R., Rupp, M. & Knoll, A. Fourier series of atomic radial distribution functions: a molecular fingerprint for machine learning models of quantum chemical properties. *Int. J. Quant. Chem.* **115**, 1084 (2015).
51. Perdew, J. P. & Levy, M. Extrema of the density functional for the energy: Excited states from the ground-state theory. *Phys. Rev. B* **31**, 6264 (1985).
52. Rupp, M., Tkatchenko, A., Müller, K.-R. & von Lilienfeld, O. A. Fast and accurate modeling of molecular atomization energies with machine learning. *Phys. Rev. Lett.* **108**, 058301 (2012).
53. Behler, J. Perspective: machine learning potentials for atomistic simulations. *J. Chem. Phys.* **145**, 170901 (2016).
54. Schütt, K. T., Arbabzadah, F., Chmiela, S., Müller, K. R. & Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nat. Commun.* **8**, 13890 (2017).
55. Chmiela, S. et al. Machine learning of accurate energy-conserving molecular force fields. *Sci. Adv.* **3**, 5 (2017).
56. Chmiela, S., Sauceda, H. E., Müller, K.-R. & Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **9**, 3887 (2018).
57. Faber, F. A., Christensen, A. S., Huang, B. & von Lilienfeld, O. A. Alchemical and structural distribution based representation for universal quantum machine learning. *J. Chem. Phys.* **148**, 241717 (2018).
58. Schütt, K. T., Gastegger, M., Tkatchenko, A., Müller, K.-R. & Maurer, R. J. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.* **10**, 5024 (2019).
59. Christensen, A. S. & von Lilienfeld, O. A. On the role of gradients for machine learning of molecular energies and forces. *Mach. Learn. Sci. Technol.* **1**, 045018 (2020).
60. Muller, K.-R., Mika, S., Ratsch, G., Tsuda, K. & Scholkopf, B. An introduction to kernel-based learning algorithms. *IEEE Trans. Neural. Netw.* **12**, 181 (2001).
61. Efron, B. & Tibshirani, R. The bootstrap method for assessing statistical accuracy. *Behaviormetrika* **12**, 1 (1985).
62. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* **17**, 261 (2020).
63. Chudoba, C., Riedle, E., Pfeiffer, M. & Elsaesser, T. Vibrational coherence in ultrafast excited state proton transfer. *Chem. Phys. Lett.* **263**, 622 (1996).
64. Lochbrunner, S., Szeghalmi, A., Stock, K. & Schmitt, M. Ultrafast proton transfer of 1-hydroxy-2-acetonaphthone: Reaction path from resonance Raman and transient absorption studies. *J. Chem. Phys.* **122**, 244315 (2005).
65. Takeuchi, S. & Tahara, T. Coherent nuclear wavepacket motions in ultrafast excited-state intramolecular proton transfer: Sub-30-fs resolved pump-probe absorption spectroscopy of 10-hydroxybenzo[h]quinoline in solution. *J. Phys. Chem. A* **109**, 10199 (2005).
66. Bartók, A. P., Payne, M. C., Kondor, R. & Csányi, G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.* **104**, 136403 (2010).
67. Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982 (1996).
68. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: the PBE0 model. *J. Chem. Phys.* **110**, 6158 (1999).
69. Barone, V. & Adamo, C. Proton transfer in the ground and lowest excited states of malonaldehyde: A comparative density functional and post-Hartree-Fock study. *J. Chem. Phys.* **105**, 11007 (1996).
70. Sadhukhan, S., Muñoz, D., Adamo, C. & Scuseria, G. E. Predicting proton transfer barriers with density functional methods. *Chem. Phys. Lett.* **306**, 83 (1999).
71. Seliskar, C. J. & Hoffman, R. E. Electronic spectroscopy of malonaldehyde. *Chem. Phys. Lett.* **43**, 481 (1976).
72. Marx, D. et al. CPMD, IBM Corporation 1990–2019 and MPI für Festkörperforschung Stuttgart 1997–2001. <http://www.cpmd.org> (2019).
73. Troullier, N. & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, 1993 (1991).
74. Steinhaus, H. Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci* **1**, 801 (1956).
75. Martyna, G. J., Klein, M. L. & Tuckerman, M. Nosé-Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.* **97**, 2635 (1992).
76. Larsen, A. H. et al. The atomic simulation environment—a python library for working with atoms. *J. Phys. Condens. Matter* **29**, 273002 (2017).

Acknowledgements

W.J.G. acknowledges financial support from the National Natural Science Foundation of China (NSFC, grant no. 22173060), the NSFC Fund for International Excellent Young Scientists (grant no. 22150610466), the Ministry of Science and Technology of the People's Republic of China (MOST) National Foreign Experts Program Fund (grant No. QN2021013001L), and the MOST Foreign Young Talents Program (grant no. WGZX2022006L). M.E.T. acknowledges support from the National Science Foundation (grant no. CHE-1955381). L.V.-M. acknowledges support from the NYU University Research Challenge Fund. This project was supported in whole (or in part) by the NYU Shanghai Boost Fund. Computational resources were supported by the NYU-ECNU Center for Computational Chemistry and a start-up fund from NYU Shanghai. We thank Mihail Bogojeski for sharing an early version of the ML-DFT code.

Author contributions

Y.B. developed the ML python code, performed TDDFT calculations, ab initio MD simulations, and ML experiments. L.V.-M. assisted with developing the ML python code and dataset clustering. L.V.-M., M.E.T., and W.J.G. initiated the work, and M.E.T. and W.J.G. contributed to the theory and experiments. All authors contributed to the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-34436-w>.

Correspondence and requests for materials should be addressed to William J. Glover.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022