Article

Molecular Dynamics of Artificially Pair-Decoupled Systems: An Accurate Tool for Investigating the Importance of Intramolecular Couplings

Michele Gandolfi and Michele Ceotto*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 6093−6108

Read Online

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

ABSTRACT: We propose a numerical technique to accurately simulate the vibrations of organic molecules in the gas phase, when pairs of atoms (or, in general, groups of degrees of freedom) are artificially decoupled, so that their motion is instantaneously decorrelated. The numerical technique we have developed is a symplectic integration algorithm that never requires computation of the force but requires estimates of the Hessian matrix. The theory we present to support our technique postulates a pair-decoupling Hamiltonian function, which parametrically depends on a decoupling coefficient $\alpha \in [0, 1]$. The closer α is to 0, the more decoupled the selected atoms. We test the correctness of our numerical method on small molecular systems, and we apply it to study the vibrational spectroscopic features of salicylic acid at the Density Functional Theory *ab initio* level on a fitted potential. Our pair-decoupled simulations of salicylic acid show that decoupling hydrogen-bonded atoms do not significantly influence the frequencies of stretching modes, but enhance enormously the out-of-plane wagging and twisting motions of the hydroxyl and carboxyl groups to the point that the carboxyl and hydroxyl groups may overcome high potential energy barriers and change the salicylic acid conformation after a short simulation time. In addition, we found that the acidity of salicylic acid is more influenced by the dynamical couplings of the proton of the carboxylic group with the carbon ring than with the hydroxyl group.

$$\tilde{F}_i = -\int \sum_j \alpha \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right) \frac{p_j}{m_j} \, dt \quad \alpha \in [0, 1]$$
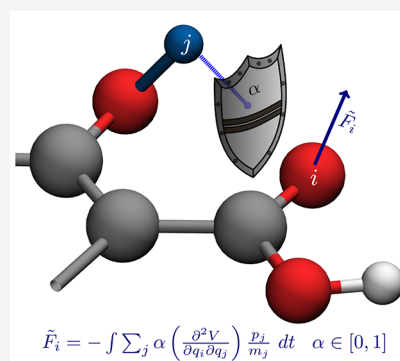
INTRODUCTION

The coupling among atoms in molecular systems is a key concept in chemistry and materials science, especially in organic chemistry, where it allows for an intuitive and qualitative description of the rich reactivity of organic compounds. Direct evidence of the vibrational couplings can be observed in the features of the vibrational (IR and Raman) spectrum, and it can be measured[1−3] in the off-diagonal features of 2D vibrational spectra. In theoretical chemistry, 2D vibrational spectra can be calculated either by using a model coupling Hamiltonian,[1,2] semiclassical approaches,[3] or other trajectory based methods.[4] This field of research is important even outside the realm of spectroscopy, because rationalization of the vibrational spectra allows the prediction, for instance, of selectivity[5,6] and reaction yields.[7]

The correlated/coupled motion of the nuclei in a molecular system is an ultrafast phenomenon, and, as such, it must be studied using either experimental ultrafast techniques, which include pulse probe methods, or computer simulation methods aimed at the interpretation and simulation of two-dimensional spectra.[2,8−10] In theoretical chemistry, the identification of the uncoupled degrees of freedom is useful for computational methodologies that calculate the vibrational spectrum in reduced dimensionality, such as, for instance, semiclassical approaches,[11−22] QM/MM calculations,[23] tensor-trains and sum of products of basis functions methods[24−26] and also the

Multi-Configuration Time-Dependent Hartree method (MCTDH)[27−29] and methods based on MCTDH-like ansatz.[30] Applications of all the aforementioned methods imply either that part of a system is partially independent of another or that the two parts have an artificial interaction. Either way, there is no rigorous method to establish whether the approximation implied is appropriate or not.

A field of research that developed accurate techniques for investigating the couplings is the study of Intramolecular Vibrational energy Redistribution. (IVR)[31−33] An accurate (yet expensive) procedure used to investigate normal mode IVR is the instantaneous normal-mode analysis,[32,34,35] which consists in a rediagonalization of the nonequilibrium Hessian matrix, whose eigenvectors can then be reinterpreted in terms of the traditional normal modes. The main disadvantage of such a technique, apart from the frequent need for the second derivative matrix of the potential, is that it can investigate the couplings only in normal mode coordinates.
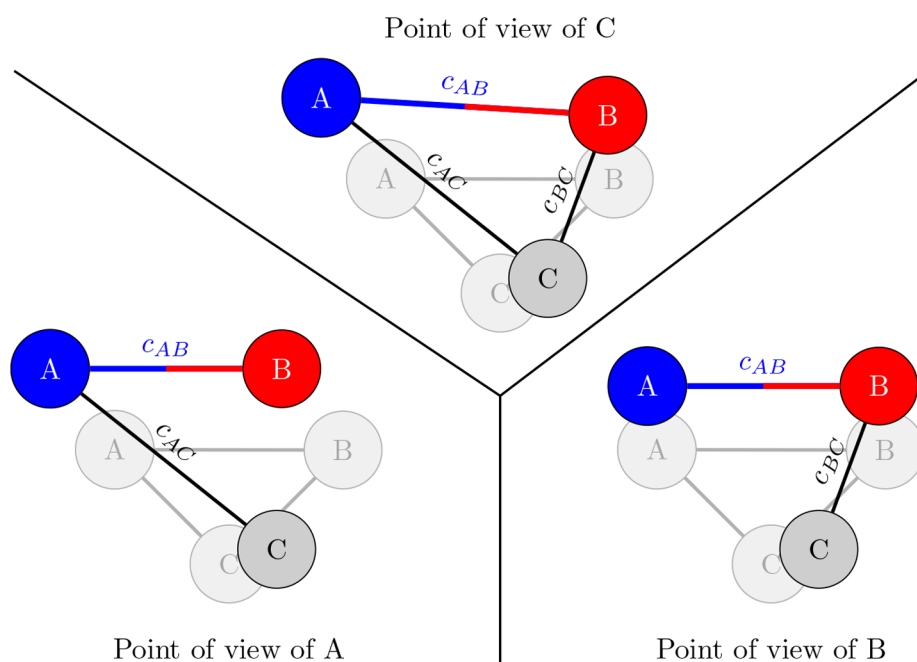
Figure 1. Cartoon of the pair-decoupling idea, representing the points of view of the three atoms compared to their initial geometry (the shaded molecules in the background). The top panel represents a snapshot of the simulation, which corresponds to the (objective) point of view of atom C. The left and right panels represent the points of view of the decoupled ($\alpha = 0$) atoms A and B, respectively, in which either atom B *perceives* atom A as if it never displaced (from its point of view) or vice versa. From the points of view of A and B, connector c_{AB} is set at the initial geometry value.

The aim of this study is to investigate the effects of the couplings from a *dynamical* perspective, accounting for the real time vibrations of the molecules and employing the intuitive Cartesian coordinate system. To that end we provide an entirely new approach to the study of couplings: We rely on a practical description of the coupling between pairs of degrees of freedom that, in its simplicity, allows us to define numerical experiments of artificially *de*coupled atoms in molecules. Specifically, we introduce a method that allows a real time, full dimensional, and numerically accurate simulation of an artificially decoupled system.

We define the atom–atom coupling as the phenomenon in which the force perceived by atom A depends on the position of atom B. In such cases, we would say that the motion of A and B is correlated or that A and B are coupled. First of all, we imagine a molecule represented by a collection of atoms, in which *each* pair of atoms are linked by the end points of a connector. If we put the molecule in its geometrical equilibrium and then abruptly displace a single atom, we would have a force acting on every atom that tries to move the whole system toward the nearest potential minimum. The force acts by either elongating or compressing (compared to their state at equilibrium) each connector and trying to adjust it in response to the deformation. This means that atoms A and B, connected by the connector c_{AB}, would feel either an attractive force that pulls them together or a repulsive one that pushes them apart. However, if we cheat the physics and artificially set this force to its value before the deformation, A and B would perceive each other just as if the system were still in equilibrium, and so the connector c_{AB} would not respond to the deformation, while all the other atoms and connectors would perceive the force and respond accordingly. For instance, another atom C would perceive the deformation, and c_{AB} might either elongate or compress (because of the compression and elongation of c_{AC} and/or c_{BC}). As a result of this

artificial intervention, the motion of decoupled atoms (A and B) in response to the deformation is directly uncorrelated (although it could be indirectly correlated via a third atom). A cartoon of the pair-decoupling idea is shown in Figure 1.

In practice, if f_{AB} is the rate of change of the force perceived by the connector c_{AB} in normal conditions, we could artificially scale it as $\dot{f}_{AB}\cdot\alpha$, where α is a real number between 0 and 1. Of course, the closer α is to 1, the more coupled A and B are and the more realistic the simulation is. On the contrary, the closer α is to 0, the more artificial it is. The practice of artificially modifying the potential is commonly used in accelerated molecular dynamics methods,[36−39] to explore the configuration space faster.

We propose to perform the artificial decoupling in a molecular dynamics simulation, in which the atoms are moved according to their initial velocities and in which we use α to shield the atoms from seeing each other's displacements. To reach this goal, we developed a very simple numerical technique, called the Symplectic Explicit with Force (SEF) integration algorithm that allows an accurate time evolution of pair-decoupled systems. Moreover, we show that the SEF integration of the equations of motion preserves the symplectic symmetry and preserves in a significant amount also time-reversibility. We also show how the time-reversibility and energy conservation properties are exact for harmonic potentials and still accurate when the potential is anharmonic. The reader should notice that the potential between the atoms A and B is not modified when they are in their equilibrium position: the attractive/repulsive contribution of the force on the pair is artificially modified only when they are displaced during time evolution. In fact, when $\alpha = 0$, the pair-decoupling implies that pairs of atoms may perceive each other as if they remain in their initial position. This detail also implies that while the equilibrium properties of the system may be unaltered by the decoupling (as it can be the case of normal mode coordinate decoupling), the dynamical properties

may change very significantly, as we shall see in the section dedicated to the salicylic acid decouplings.

Hereafter, the Theory and Methods section formalizes the pair-decoupling idea and presents the molecular dynamics modified integrator for artificially decoupled pairs of atoms. The Results section presents the benchmarking of the method and its application to salicylic acid. The Discussion and Conclusions section concludes the paper.

■ THEORY AND METHODS

Decoupling Hamiltonian. As anticipated above, we describe molecular systems as composed of some arbitrarily pair-decoupled atoms. We call $H = K + V$ the Hamiltonian of the fully coupled (normal) system and $\tilde{H} = K + \tilde{V}$ the Hamiltonian of the corresponding pair-decoupled system. For the rest of the paper, we will assume that V (and $\tilde{V}$) contains the electron–electron, electron–nucleus, and nucleus–nucleus Coulomb interactions, as well as the electron exchange potential and the electronic kinetic energy. Thus, V (and $\tilde{V}$) is a function of the nuclei positions in the Born–Oppenheimer approximation, and K is the corresponding nuclear kinetic energy. We make the further approximation that the nuclei behave as classical particles so that we can express the classical pair-decoupled Hamiltonian as $\tilde{H} = \sum_\alpha (\tilde{p}_\alpha^2/2m_\alpha) + \tilde{V}(\tilde{q})$, where we assign the tilde ($\sim$) symbol to the canonical coordinates $\tilde{q}$ and $\tilde{p}$ to specify that they are phase-space coordinates of the pair-decoupled Hamiltonian. While we do not have an explicit expression for $\tilde{V}$ in terms of V, we express the relationship in terms of the potential derivatives because we employ these for integrating the equation of motion. Specifically, we assume that the main coupling between pairs is given by the second order derivative terms with respect to coordinates q_i and q_j (i.e., the Hessian matrix elements h_{ij}) and that it can be artificially scaled:

$$\frac{\partial^2 \tilde{V}(\tilde{q})}{\partial \tilde{q}_i \partial \tilde{q}_j} = \alpha \frac{\partial^2 V(\tilde{q})}{\partial \tilde{q}_i \partial \tilde{q}_j} = \alpha h_{ij}(\tilde{q}) \tag{1}$$

In eq 1, we have assumed that the decorrelation is for pairs of degrees of freedom, and α is the amount of decoupling, ranging from 1 (no decoupling) to 0 (fully decoupled). However, one can decouple multiple degrees of freedom at the same time. For instance, the pair-decoupled Hessian matrix for a three degrees of freedom system where two of them are fully coupled and the third one is partially decoupled from them is

$$\tilde{h}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3) = \begin{pmatrix} h_{11} & h_{12} & \alpha h_{13} \\ h_{12} & h_{22} & \alpha h_{23} \\ \alpha h_{13} & \alpha h_{23} & h_{33} \end{pmatrix} \tag{2}$$

Notice that when $\alpha = 0$, eq 2 corresponds to the Hessian matrix of two independent systems, one of which is two-dimensional and the second is monodimensional. When the time-evolution algorithm described in the next sections is applied to such a system, the evolution of the system is artificially separable, and the potential is up to the second order of the type $V(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3) = V_{1,2}(\tilde{q}_1, \tilde{q}_2) + V_3(\tilde{q}_3)$. Notice that in case of a truly separable potential, we could write $V_3(\tilde{q}_3) - V_3(\tilde{q}_3^{eq}) = V(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3) - V(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3^{eq})$ and $V_{1,2}(\tilde{q}_1, \tilde{q}_2) - V_{1,2}(\tilde{q}_1^{eq}, \tilde{q}_2^{eq}) = V(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3) - V(\tilde{q}_1^{eq}, \tilde{q}_2^{eq}, \tilde{q}_3)$. These last expressions correspond to the "projected potentials" used to compute the vibrational spectroscopic features of molecules as large as G-quadruplex in solution[40] with the Divide-and-Conquer SemiClassical Initial

Value Representation (DC-SCIVR) method.[14,15,20] While the DC-SCIVR method simulates the dynamics of a system under a full dimensional potential and then approximates the classical action with a potential projected into subspaces,[14] the algorithm we are presenting here evolves the dynamics entirely under the subspace-projected potential (or partially projected, when $\alpha \neq 0$). Furthermore, the decoupling could be applied to all the degrees of freedom pertaining to two atoms, that is, for instance, to the Cartesian product $(x_1, y_1, z_1) \times (x_2, y_2, z_2)$, to decouple atoms 1 and 2. In this paper, we focus on this Cartesian atom-decoupling scheme, as it appeals to chemical intuition, and we believe it would result as the most interesting for the chemical community. Nonetheless the decoupling idea could be applied to any coordinate system.

In practice, we propose a time-propagation rule that allows us to enforce the pair decoupling idea by making explicit the contributions to the force given by the Hessian matrix. In other words, we obtain the molecular dynamics force by numerically integrating the Hessian matrix over time as $\tilde{F}(t) = \int_0^t \tilde{h}(t') \dot{\tilde{q}}(t') \, dt'$. However, this procedure is tricky because the variables $\tilde{q}(t)$, $\tilde{p}(t)$, and $\tilde{F}(t)$ must be evaluated at the same time for the propagation to take place:

$$q(\tau) = \tilde{q}(0) + \int_0^\tau \tilde{p}(t) \, dt$$

$$\tilde{p}(\tau) = \tilde{p}(0) + \int_0^\tau \tilde{F}(t) \, dt$$

$$\tilde{F}(\tau) = \tilde{F}(0) - \int_0^\tau \tilde{h}(t) \cdot \tilde{p}(t)/m \, dt \tag{3}$$

where $\tilde{h}(t) := \alpha \partial^2 V(\tilde{q}(t))/\partial \tilde{q}_i \partial \tilde{q}_j$ is the pair-decoupled Hessian matrix at time t. The procedure we have developed employs a standard symplectic integration for the coordinates $\tilde{q}$ and $\tilde{p}$, while the force is updated from a time-integration of the Hessian matrix embedded with the symplectic map, consistent with the canonical variables. Specifically, for an integration that is accurate to order n, the practical update is a cycle over the integer k, up until $k = n = 2$ or 4 of the following four simple steps:

$$\tilde{p}_k = \tilde{p}_{k-1} + b_k \tau \tilde{F}_{k-1}$$

$$\tilde{q}_k = \tilde{q}_{k-1} + a_k \tau \frac{\tilde{p}_k}{m}$$

$$q_{aux} = \tilde{q}_k + \sum_{j=1}^k (a_j - b_j) \frac{\tilde{p}_k}{m} \tau$$

$$\tilde{F}_k = \tilde{F}_{k-1} - c_k \tau \left(\tilde{h}(q_{aux}) \cdot \frac{\tilde{p}_k}{m} \right) \tag{4}$$

where m is the mass and $\tilde{p}_k = \tilde{p}(t = \sum_{j=1}^k b_j \tau)$ and $\tilde{q}_k = \tilde{q}(t = \sum_{j=1}^k a_j \tau)$ are the momentum and position (vector) variables at step k of the symplectic map starting from the initial conditions $\tilde{p}_0$, $\tilde{q}_0$, and $\tilde{F}_0 = \partial V(\tilde{q}_0)/\partial \tilde{q}$. The numerical coefficients a_k, b_k, and c_k are universal real numbers that depend only on the order of approximation. Elegant derivations of the a_k and b_k coefficients for high order integrators can be found in the literature, as many authors have worked in the field of symplectic integration.[41−43] Thus, in the next sections, we give only a brief overview for the derivation of the a_k and b_k coefficients, which are solutions of the system of equations given in the Appendix

(we leave a more in depth explanation in the Supporting Information material). In the next sections, we discuss in some more detail how the c_k coefficients can be easily obtained for a second order integrator and how the fourth and higher order integrators can be obtained by composition of second order ones. For a second order integration, we found the unique solution $b_1 = 0$, $b_2 = 1$; $a_1 = c_1 = 1/2$, $a_2 = c_2 = 1/2$, which corresponds to the symplectic leapfrog algorithm with $c_k = a_k$. We then propose a fourth order version of the pair-decoupled algorithm as a symmetric product of three leapfrog algorithms, with coefficients $a_1 = a_4 = (2^{1/3} + 2^{-1/3} + 2)/6$; $a_2 = a_3 = -(2^{1/3} + 2^{-1/3} - 1)/6$; $b_1 = 0$; $b_2 = b_4 = (2^{4/3} + 2^{2/3} + 4)/6$; $b_3 = -(2^{7/3} + 2^{5/3} + 2)/6$, $c_k = a_k$. This choice is the most accurate, according to our numerical tests. However, several other choices arise when the higher order algorithms are not derived as symmetric products of lower order algorithms, as discussed in the next sections and more in detail in the Supporting Information of this paper.

Overview of Hamiltonian Systems Integration with Symplectic Maps. To derive the a_k, b_k, and c_k coefficients, we begin with the formal solution of the equations of motion

$$z(\tau) = e^{-\tau \hat{H}} z(0) \tag{5}$$

where $\hat{H}$ is an operator that transforms the state function $z(t)$ into its time derivative. In quantum mechanics, $z(t)$ is a complex-valued wave function and $\hat{H} = i\hat{H}$, where $\hat{H}$ is the quantum Hamiltonian operator in atomic units. In classical mechanics, $z(t)$ is a real-valued vector of the canonical coordinates and $\hat{H} = \{H, \cdot\}$, where H is the classical Hamiltonian function and $\{H, \cdot\}$ is called the Liouville (or Lie) operator, which works as a Poisson bracket of $z(t)$:

$$\{H, \cdot\} z(t) = \{H, z\}(t)$$
$$= \frac{\partial H}{\partial q} \frac{\partial z}{\partial p}(t) - \frac{\partial H}{\partial p} \frac{\partial z}{\partial q}(t) \tag{6}$$

We restrict our study to the dynamics of molecules, for which the nuclear part of the Hamiltonian can be written as a sum of a kinetic energy and a potential energy; that is, $\hat{H}(q, p) = \hat{K}(p) + \hat{V}(q)$. Hence, the time evolution operator becomes $e^{-\tau(\hat{K}+\hat{V})} z(t)$. Since $\hat{K}$ and $\hat{V}$ do not commute, the exact time evolution operator is not just the product of its kinetic and potential components. A way to link $e^{-\tau(\hat{K}+\hat{V})}$ with products of kinetic and potential evolution operators[42] is via the Baker-Campbell-Housdorff-Dynkin formula (BCHD), which is an infinite series of nested commutators and it can not be implemented directly. Nonetheless, the BCHD formula has been employed to demonstrate the time-invariance properties of even-order symplectic integrators.[41] In case the Hamiltonian is not separable in its kinetic and potential contributions, it would still be possible to use the standard symplectic integration methods, by evolving copies of the system onto an extended phase space, as described by Tao.[44]

Two very common approximations for $e^{-\tau(\hat{K}+\hat{V})}$ are the first order map

$$\mathcal{M}_1(\tau) = e^{-\tau \hat{K}} e^{-\tau \hat{V}} + O(\tau^2)$$

and the second order map

$$\mathcal{M}_2(\tau) = e^{-\tau/2 \hat{K}} e^{-\tau \hat{V}} e^{-\tau/2 \hat{K}} + O(\tau^3)$$

The latter approximation is known by various names, depending on the context (Strang splitting[45] by mathematicians, Trotter-Suzuki splitting[46] in the quantum mechanics community, Symplectic Leapfrog or explicit Verlet in the classical mechanics community).[47] We will refer to the second order map as "SE2" (Symplectic Explicit of 2nd order).

A more general approximation of the time evolution operator is given by

$$e^{-\tau(\hat{K}+\hat{V})} z(0) \approx [e^{-\tau a_1 \hat{K}} e^{-\tau b_1 \hat{V}} e^{-\tau a_2 \hat{K}} e^{-\tau b_2 \hat{V}} \cdots] \tag{7}$$

with the constraints $\sum_k a_k = \sum_k b_k = 1$. Many approximate solutions in the a_k and b_k variables have been found in the '80s and '90s (assuming the classical expressions for $\hat{K}$ and $\hat{V}$).[41,48−54] A major advancement was done by Creutz,[55] Yoshida,[41] and Suzuki,[46] who independently derived (among other things) a general formula for a class of arbitrary even order integrators using a "symmetric product"[41,44,46,50,55] of lower order symplectic maps

$$\mathcal{M}_{n+2}(\tau) = \mathcal{M}_n(\gamma_n \tau) \cdot \mathcal{M}_n((1 - 2\gamma_n)\tau) \cdot \mathcal{M}_n(\gamma_n \tau)$$

with $\gamma_n = \frac{1}{2 - 2^{1/(n+1)}}$. In particular, the symmetric product or SE2 algorithms, gives the well-known fourth order symplectic map with the coefficients reported by Forest and Ruth.[50] Also notice that the symmetric product formula holds true even if we change the form of the $\hat{K}$ and $\hat{V}$ operators, meaning that we can apply it to our customized pair-decoupling Hamiltonian.

We also consider the coefficients presented in ref 56, which provide a very accurate fourth order integrator that, however, cannot be obtained as a symmetric product of second order integrators and it is not time-reversible by construction. In the Supporting Information of this paper, we report a general derivation of symplectic maps up to fourth order, from which we can derive both fourth order integrators along with their pair-decoupled versions. Our approach to derive the a_k, b_k, c_k coefficients consists of the direct application of the operators in eqs 7 or 4 to $z(0)$ for a given truncation of the map (for a given integer n), followed by a comparison of the resulting $z(\tau)$ with the time Taylor series of $z(t)$ centered in $z(0)$ at the same orders of τ.

We point out, for sake of completeness, that there are also other ways of writing a symplectic map, for instance, explicitly including third order terms of the BCHD formula, such as $e^{-\tau d_k [\hat{V}, [\hat{K}, \hat{V}]]}$, into the map in eq 7, where the commutator, assuming classical mechanics, evaluates to

$$[\hat{V}, [\hat{K}, \hat{V}]](q(t), p(t')) = -2 \frac{\partial^2 V(q(t))}{\partial q^2} \cdot \frac{\partial V(q(t))}{\partial q} \cdot \frac{\partial p(t')}{\partial p}$$

This approach was presented by Suzuki, in an attempt to derive an integrator with only positive a_k and b_k coefficients and avoid the unboundedness of the propagators with otherwise positive exponents (which make no sense when applied to diffusion algorithms).[57] The actual integrator with positive-only coefficients was derived and implemented by Chin,[54] and is proven to be extremely accurate, but also more expensive, because it requires evaluation of the Hessian matrix. While this method might be appropriate to integrate a pair-decoupling integrator, we disregarded it because of the difficulties of having the

Table 1. Summary of the a_k, b_k, and c_k Coefficients for Various Versions of the SEF Algorithm

SEF version	coefficients	k = 1	k = 2	k = 3	k = 4
SEF2[a]	a_k	1/2	1/2	0	0
	b_k	0	1	0	0
	c_k	1/2	1/2	0	0
SEF4[b]	a_k	$(2^{1/3}+2^{-1/3}+2)/6$	$-(2^{1/3}+2^{-1/3}-1)/6$	$-(2^{1/3}+2^{-1/3}-1)/6$	$(2^{1/3}+2^{-1/3}+2)/6$
	b_k	0	$(2^{4/3}+2^{2/3}+4)/6$	$-(2^{7/3}+2^{5/3}+2)/6$	$(2^{4/3}+2^{2/3}+4)/6$
	c_k	$(2^{1/3}+2^{-1/3}+2)/6$	$-(2^{1/3}+2^{-1/3}-1)/6$	$-(2^{1/3}+2^{-1/3}-1)/6$	$(2^{1/3}+2^{-1/3}+2)/6$
SEF4-I[c]	a_k	$\sqrt{3}/6+1/2$	$-\sqrt{3}/3$	$\sqrt{3}/3$	$-\sqrt{3}/6+1/2$
	b_k	0	$-\sqrt{3}/6+1/4$	1/2	$\sqrt{3}/6+1/4$
	c_k	5/26	$8\sqrt{3}/39+4/13$	$-8\sqrt{3}/39+4/13$	5/26
SEF4-II[c]	a_k	$\sqrt{3}/6+1/2$	$-\sqrt{3}/3$	$\sqrt{3}/3$	$-\sqrt{3}/6+1/2$
	b_k	0	$-\sqrt{3}/6+1/4$	1/2	$\sqrt{3}/6+1/4$
	c_k	$5\sqrt{3}/48$	$\sqrt{3}/8+1/2$	$-\sqrt{3}/8+1/2$	$-5\sqrt{3}/48$
SEF4-III[c]	a_k	$\sqrt{3}/6+1/2$	$-\sqrt{3}/3$	$\sqrt{3}/3$	$-\sqrt{3}/6+1/2$
	b_k	0	$-\sqrt{3}/6+1/4$	1/2	$\sqrt{3}/6+1/4$
	c_k	1/4	$\sqrt{3}/6+1/4$	$-\sqrt{3}/6+1/4$	1/4
SEF4-IV[c]	a_k	$\sqrt{3}/6+1/2$	$-\sqrt{3}/3$	$\sqrt{3}/3$	$-\sqrt{3}/6+1/2$
	b_k	0	$-\sqrt{3}/6+1/4$	1/2	$\sqrt{3}/6+1/4$
	c_k	$\sqrt{3}/6$	1/2	1/2	$-\sqrt{3}/6$

[a]Using Symplectic Leapfrog a_k and b_k coefficients. [b]Using Forest and Ruth a_k and b_k coefficients.[50] [c]Using Brewer, Hulme, and Manolopoulos a_k and b_k coefficients.[56]

derivatives of the potential evaluated at the same time t, as should be clear from the next section of this paper.

The developments in this work can also be applied in different contexts, although they might require different interpretations. We collect here some of the recent developments in the use of geometric and symplectic integrators in the context of simulating quantum mechanical systems using classical trajectory methods, that could make use of the pair-decoupling integrators straightforwardly.[58−64]

Integration of the Pair-Decoupled System. We construct our algorithm to be of the type of a nth order symplectic map

$$\mathcal{M}_n = e^{-\tau b_n\{\tilde{V},\cdot\}}e^{-\tau a_n\{K,\cdot\}}\cdots e^{-\tau b_1\{\tilde{V},\cdot\}}e^{-\tau a_1\{K,\cdot\}} \quad (8)$$

which consists of a time evolution of a free system, followed by a time evolution of the pair-decoupled system with zero velocity, followed by evolution of the free system and so on. Notice that, in comparison with the standard symplectic map in eqs 7, we modify only the form of the potential energy operator $\hat{\mathcal{V}}$, without changing the structure of the map, which remain symplectic, independently of how we modify $\hat{\mathcal{V}}$. In fact, as long as eq 8 can be written as a single product of time evolution operators, we are sure that symplectic structure is preserved, contrary, for instance, to standard Runge-Kutta-Nystrom algorithms, which cannot be written as a single product, as explained by Chin.[43] We use the definition given in eq 1 to integrate the Hessian and get the locally harmonic approximated expression for the (pair-decoupled) force

$$\tilde{F}(\tau) = \tilde{F}(0) - \sum_{k=1}^{n} \tilde{h}(\tilde{q}(b_k\tau))\cdot\dot{\tilde{q}}(b_k\tau)c_k\tau \quad (9)$$

The positions and momenta are those resulting from the application of the operators in the symplectic map of eq 8. Notice that in eq 9, one needs $\tilde{q}$ and $\dot{\tilde{q}}$ to be evaluated at the same time, $t = b_k\tau$. However, after the application of the two rightmost operators of eq 8, one obtains the position and the velocity at

different time values, i.e., $\tilde{q}(a_1\tau)$ and $\tilde{p}/m = \dot{\tilde{q}}(b_1\tau)$, since, in general, $a_1 \neq b_1$. Hence, we introduce an auxiliary position variable

$$q_{1,aux} := \tilde{q}(a_1\tau) - (a_1 - b_1)\tau\dot{\tilde{q}}(b_1\tau) \approx q(b_1\tau)$$

which is the position estimate at the same instant of time of the conjugated momentum variable. In general,

$$q_{k,aux} := \tilde{q}(a_k\tau) - \sum_{j=1}^{k} (a_j - b_j)\tau\dot{\tilde{q}}(b_k) \approx q(b_k\tau)$$

In this way, the integration of the force in eq 9 is consistent with the rest of the algorithm. Finally, given eq 9, we can compare the explicit form of the evolution operator for the pair-decoupled potential with the evolution operator of the original potential. To do that, we first evolve $z(0)$ until we get $(p(t_p), q(t_q)) = e^{-\tau a_k\{K,\cdot\}}\prod_i^{k-\tau b_i\{V,\cdot\}}e^{-\tau a_i\{K,\cdot\}})z(0)$, where we call $t_p = \sum_i^k a_i\tau$ and $t_q = \sum_i^{k-1} b_i\tau$ the time arguments of the momentum and position variables. Then, the additional application of the evolution operator $e^{-\tau b_k\{V,\cdot\}}$ gives

normal: $\quad e^{-\tau b_k\{V,\cdot\}}z(t_p, t_q) = [1 + F(q(b_k\tau))]\dfrac{\partial z}{\partial p}(t_p, t_q)$

decoupled:

$$e^{-\tau b_k\{\tilde{V},\cdot\}}\tilde{z}(t_p, t_q) = \left[1 + \left(\tilde{F}(0) - \sum_{i=1}^{k}\tilde{h}(q_{i,aux})\cdot\dot{\tilde{q}}(a_i\tau)c_i\tau\right)\right]\dfrac{\partial\tilde{z}}{\partial\tilde{p}}(t_p, t_q) \quad (10)$$

We call $\tilde{F}_k := \left(\tilde{F}(0) - \sum_{i=1}^{k}\tilde{h}(q_{i,aux})\cdot\dot{\tilde{q}}(a_i\tau)c_i\tau\right)$, $\tilde{p}_k := \tilde{p}_{k-1} + b_k\tau\tilde{F}_{k-1}$, and $\tilde{q}_k := \tilde{q}_k + a_k\tau\tilde{p}_k$ the variables we need to store to implement the algorithm. We determine the coefficients c_k by considering that in the case of no decoupling ($\alpha = 1$) the pair decoupled algorithm must provide a good estimate of the original force, i.e., $\tilde{F}_k \approx F_k$, and of the pair-decoupled position and momentum. Hence, a route to find the c_k coefficients is to assume that a_k and b_k coefficients are equal to those derived for the fully coupled system and then to choose the c_k coefficients so

that the errors on $\tilde{q}_n$ and $\tilde{p}_n$ (and $\tilde{F}_n$) are of a given order of τ when $\alpha = 1$. Notice that, in the particular case of a quadratic potential, $\tilde{h}$ in eq 10 is a constant (and $\tilde{F}(0) = 0$, assuming that at $t = 0$ the system is in equilibrium). Thus, for a quadratic potential, we have

$$F(q(b_k\tau)) \approx -\tilde{h} \cdot \sum_{i=1}^{k} \dot{\tilde{q}}(a_i\tau) c_i \tau$$

This is a good estimate of the force when the sum $\sum_{i=1}^{k} \dot{\tilde{q}}(a_i\tau) c_i \tau$ is a good estimate of the position, that is, when $c_i = a_i$. In fact, the appealing choice $c_k = a_k$ is appropriate also for non-quadratic potentials when using a second order integrator (that is, either $a_1 = a_2 = 1/2$, $b_1 = 0$, $b_2 = 1$, or $b_1 = b_2 = 1/2$, $a_1 = 0$, $a_2 = 1$). This is not surprising, because the expanded expression for $e^{-\tau/2\hat{K}} e^{-\tau\hat{V}} e^{-\tau/2\hat{K}} z(t)$ contains the derivatives of $V(q)$ and $K(p)$ up to the second order, meaning that it is, in fact, a propagator of the system under a local quadratic approximation of the exact functions $K(\tilde{p})$, which is actually quadratic, and $\tilde{V}(\tilde{q})$, which is not. We call this choice the "SEF2" method (Symplectic Explicit with Force integration of the second order). The full set of coefficients is reported in the corresponding lines of Table 1.

The obvious way to derive an integrator of order four and higher is by composing lower order integrators, as mentioned above. The fourth order SEF4 integrator can be obtained as a symmetric product of the second order SEF2, with a_k and b_k coefficients equal to those derived by Forest and Ruth (also, independently derived by Campostrini and Rossi, and Candy and Rozmus),[49,53] and with $c_k = a_k$. All of the coefficients are reported in the SEF4 row of Table 1. Integrators of order 6 and higher can be easily obtained in the same way, resulting in the coefficients reported by Yoshida,[41] with $c_k = a_k$. However, the method of composing lower order integrators does not generate all the solutions to the fourth order symplectic map in eq 7. All such solutions can be found by solving the system in eq 19 of the Appendix, where, however, time reversible symmetry is not enforced. In particular, the coefficients reported in Appendix of ref are solutions to eq 16, but do not enforce time reversibility, despite providing an integrator that is more accurate than Forest and Ruth's (in terms of energy conservation). While it is possible to build a pair-decoupled integrator using the a_k and b_k coefficients by Brewer et al., it is not possible to reach fourth order accuracy, and there are multiple possible choices of the c_k coefficients that we discuss in the Supporting Information. All such coefficients are reported in the rows SEF4-I to SEF4-IV of Table 1, as possible variants of the SEF4 integrator.

In the Numerical tests section we show numerically that our modification of the symplectic algorithm, which accounts for the pair-decoupling concept, preserves the properties of symplectic integration. Considering the Jacobian matrix

$$J_{ij}(t, t') = \frac{\partial(\tilde{p}_i(t), \tilde{q}_j(t))}{\partial(\tilde{p}_i(t'), \tilde{q}_j(t'))} \tag{11}$$

and the canonical symplectic matrix

$$\mathcal{J} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{12}$$

we measure how much the relation $J^T(t, t') \cdot \mathcal{J} \cdot J(t, t') = \mathcal{J}$ holds true for the special case $t' = 0$. In the case of a quadratic

potential, the time evolution is exactly time reversible. Instead, in the case of a generic potential, the time-reversibility property is only approximate.

Computational Details. In order to test our algorithm of eq 4, we employ accurate quartic force fields[65,66] for the simulations of water and formaldehyde molecules. Instead, for salicylic acid calculations, we use the fitted potential energy DFT surface[67,68] provided alongside the sGDML software.[67−73] The sGDML PES is given already trained[67,68] on 1000 training points, and it showed, with the inclusion of the Tkatchenko-Scheffler correction to account for the van der Waals interactions,[74] a mean absolute error (MAE) which is less than 0.12 kcal/mol with respect to the pVDZ/DFT-PBE values.

To test the accuracy of the integration technique, we run 4 types of tests with $\alpha = 1$. First we check the energy conservation along the simulation. Then, we checked the symplectic property of the Jacobian matrix. Then, we check the time-reversibility of the integrator, and eventually we compute also the classical power spectrum. The check of energy conservation may appear redundant because a symplectic integration, by definition, implies that all the constants of motion are preserved. However, our approximation of the force implies that the system evolved under an approximation of the potential. As a matter of fact, even if $\alpha = 1$, if the force estimate in eq 9 is not accurate, the energy (of the fully coupled system) might not be accurately conserved, while the integration remains symplectic on the approximated potential.

To prove that the Jacobian matrix $J(t, t')$ is symplectic we use the relation $J^T(t, t') \cdot \mathcal{J} \cdot J(t, t') = \mathcal{J}$ for the special case $t' = 0$, as anticipated above. The Jacobian with $t' = 0$ is called the monodromy matrix $M(t)$, which can be computed numerically with the extended version of the algorithm described in the Supporting Information. Hence, we asses the stringent condition[12]

$$\Upsilon(t) = \sqrt{\sum_i \sum_j |(M(t)^T \cdot \mathcal{J} \cdot M(t))_{ij} - \mathcal{J}_{ij}|^2} \approx 0 \tag{13}$$

Although eq 13 proves the symplectic property of the Jacobian matrix only for the special case $t' = 0$, this is the most stringent test from the numerical point of view.

To measure the degree of time reversibility of the integrator, we run a simulation until time T, with a 10 au time step. After that, we invert the sign of the momentum variable, and we continue the propagation for another time lapse equal to T backward, until a total simulation time of $2T$ is reached. Finally, we measure the quantity

$$\tau(t) = \frac{1}{F} \sum_{j=1}^{F} |x_j(2T - t) - x_j(t)| \approx 0 \tag{14}$$

where $F = 3N_{at}$ (N_{at} is the number of atoms) and x_j is the jth element of the F-dimensional Cartesian geometry vector. In all our tests, $T = 6000$ au.

Finally, we apply our integration technique for the calculation of the vibrational spectra, ranging from small molecules up to the salicylic acid molecule in the gas phase. We use a numerically convenient formula[75] to evaluate the power spectrum of the jth mass-scaled normal mode,

$$I_j(\omega_j) = \frac{1}{T} \left| \int_0^T p_j(t) e^{i\omega_j t} \, dt \right|^2 \tag{15}$$

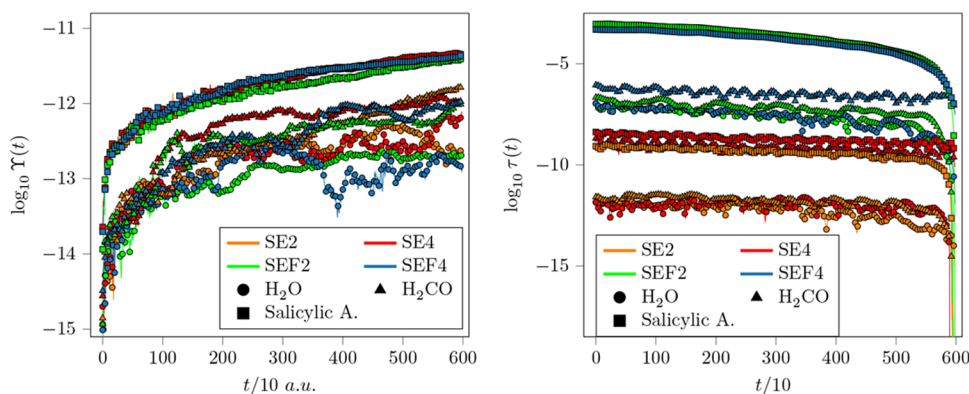**Figure 2.** $\Upsilon(t)$ and $\tau(t)$ for a 3000 time-step simulation of 10 au each. SE2 (orange), SE4 (red), SEF2 (green), and SEF4 (blue) integration methods for H_2O (circles), H_2CO (triangles), and salicylic acid (squares). SEF2 and SEF4 are tested without decoupling.

This formula provides a resolved power spectrum with short ($\approx$ 0.6 ps) simulations. In fact, eq 15 is the classical analogue of the time-averaging method employed in semiclassical spectroscopy.[76,77] Since eq 15 computes a power spectrum from the velocity correlation function, all vibrational frequencies are reproduced and they can be compared with either infrared or Raman experimental frequencies. Instead, the intensities $I(\omega_j)$ are not comparable with IR or Raman experiments because they depend only on the number of times the vibrational mode with frequency ω_j has occurred during the simulation time, which depends ultimately on the trajectory initial conditions. On the other hand, the experimental infrared and Raman intensities depend, for example, on the transition dipole moments and on the polarizabilities. To represent the power spectrum intensity of a multidimensional system we simply compute the sum of the power spectra of eq 15, that is

$$I(\omega) = \sum_j I_j(\omega_j) \tag{16}$$

The definite integral of eq 16 over a frequency domain can be interpreted as the average kinetic energy of the modes of vibration within that frequency domain.[33] Thus, when the intensity $I(\omega)$ of the pair-decoupled simulation is different from the nondecoupled one, there must be a shift in the vibrational frequency or an intramolecular vibrational energy redistribution caused by the decoupling. The two effects may occur at the same time.

Although all the integrators described in this paper allow the system to evolve in any full-dimensional coordinate system, we always employ mass-scaled normal modes, which have the advantage of discarding translational and rotational motion. To decouple the Cartesian degrees of freedom, we just rotate the normal mode Hessian matrix to Cartesian coordinates, apply the decoupling, and rotate the decoupled matrix back to normal modes. We use this procedure, instead of evolving in Cartesian coordinates, because the SEF algorithm cannot accurately describe free translations and rotations or other types of motion that have a very flat (in general very anharmonic) potential landscape.

All of the simulations described in this paper are full dimensional and start from the equilibrium geometry of the fully coupled system. The normal mode coordinates are constructed using the fully coupled Hessian matrix. Also, in case the SEF algorithm is used, the initial force is assumed to be 0, just as if the initial geometry were an energy minimum for the

pair-decoupled system, as well. The initial momentum in normal mode coordinates is set equal to the square root of the corresponding harmonic frequency so that the initial kinetic energy is equal to the harmonic zero point energy. In the simulations with $\alpha \neq 1$ we follow the same recipe, but, when the effect of the decoupling is weak, we run the simulations for longer time (5000 time steps), and discard the initial 2000 steps (which is about 0.5 ps), to allow the decoupled fragments to actually decorrelate. In some cases, such as when decoupling all the functional groups of salicylic acid, the decoupling effect is very strong, and the decorrelation effects can be seen already from the very beginning of the simulation. In such cases, we run the simulation for only 3000 steps and discard none. Anyway, if the pairs of atoms are naturally independent, the normal spectrum and the pair-decoupled spectrum would be exactly the same.

■ RESULTS

Numerical Tests. We start by testing the accuracy of our algorithms. In Figure 2 we show how much the SEF2 and SEF4 integrators with $\alpha = 1$ preserve the symplectic symmetry of the monodromy matrix $M(t)$ and the time reversibility property. These are compared with the well established Symplectic Leapfrog (SE2) and fourth order SE4 method, that is the Symplectic Explicit integration method with the coefficients of Forest and Ruth.[50] Independently of the integrator, the larger is the system, the quicker $\Upsilon(t)$ and $\tau(t)$ deteriorate, and this is mainly due to the fact that more operations are carried out in a finite precision arithmetic. However, when switching from the SE to the SEF algorithms, no significant further errors in $\Upsilon(t)$ are introduced, while the time reversibility accuracy is decreased by orders of magnitude. This is expected for two main reasons. First, the calculation of the force in the SEF algorithms is performed by time integration and it requires four sums of matrix multiplications. The second, and the most important one, is that the calculation of the force is based on a local harmonic approximation of the potential landscape. Our approximate evolution of the force within the local harmonic approximation is not a time-reversible process, except for quadratic potentials. These limitations are clearly amplified with the dimensionality.

The symplectic properties of the SEF integration are preserved also in the case of the decoupled pairs of degrees of freedom, i.e., $\alpha = 0$. We show in the Supporting Information (Figure S1) that $\Upsilon(t)$ and $\tau(t)$ have the same shape even for the pair-decoupled system.
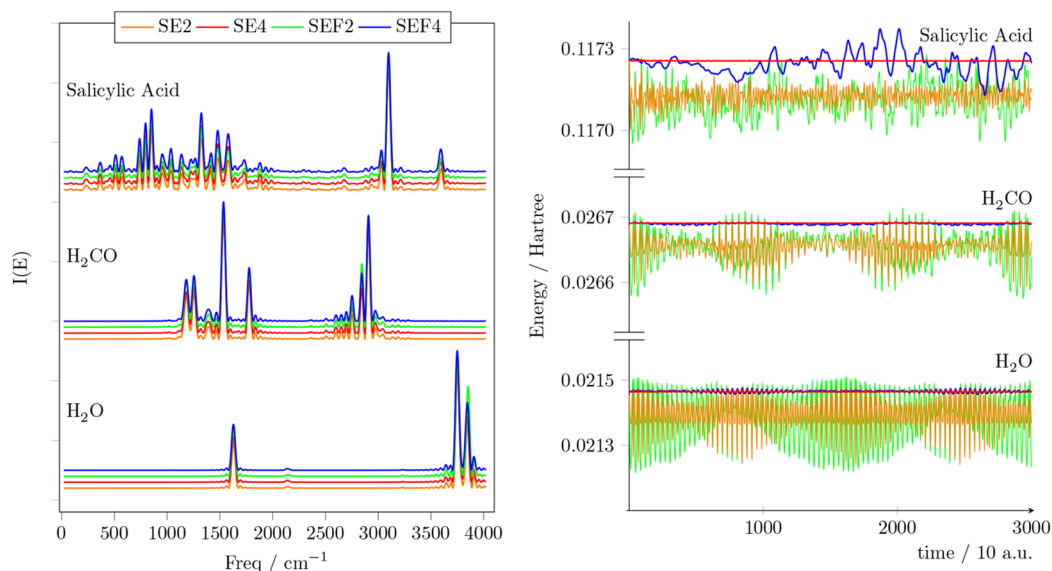
Figure 3. Power spectra and energy conservation. The left panel shows the power spectra of H_2O, H_2CO, and salicylic acid molecules computed with SE2 (orange), SE4 (red), SEF2 (green), and SEF4 (blue) integrators (always with $\alpha = 1$). The right panel shows the energy profiles of the corresponding simulations.
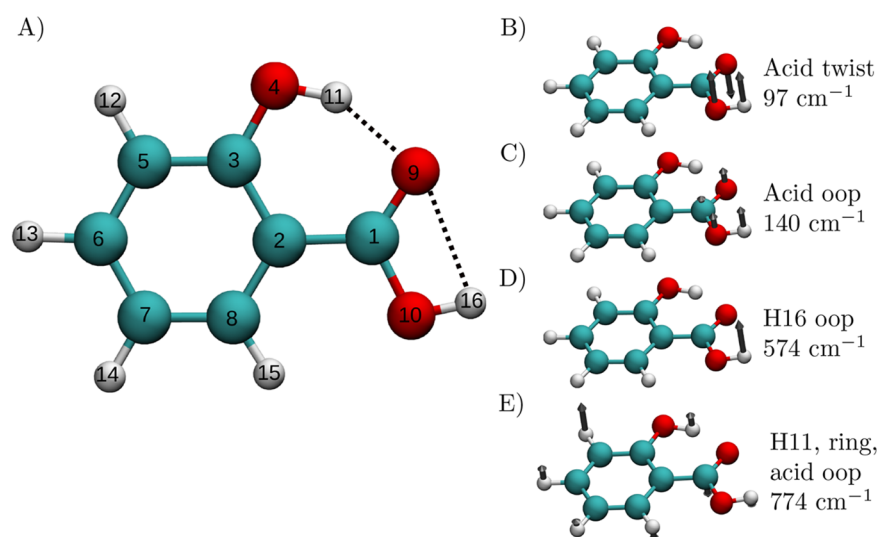

Figure 4. (A) Picture of salicylic acid with labeled atoms and (B, C, D, E) some relevant types of motion[83] involving the acid group, H11, and H16 with their relative harmonic frequencies.

In Figure 3 we show that, for all systems, the spectroscopic features are perfectly captured by the entire integration method. SEF4, SE2, and SE4 provide spectra that are almost quantitatively equivalent when applied to all systems. The total energy of the H_2O and H_2CO systems is well conserved by the SEF algorithms, with SEF2 having an oscillation that is about ∼10% larger than SE2, and SEF4 having an oscillation that is ∼20% to ∼30% larger than SE4. When the system includes floppy modes, however, such as the salicylic acid, these modes induce a slow oscillatory pattern in the energy profile that is not well captured by the SEF algorithm. In fact, SEF can not predict very accurately the strongly anharmonic contributions to the force. Nevertheless, this is not an issue, because the SEF energy does not display a systematic drift, but only a slow oscillatory pattern that follows the oscillation of the low energy modes.

Decoupling the Salicylic Acid Fragments. Since the pair decoupling is, by definition, an artificial procedure, we

rationalize the following results in a *reductio ad absurdum* style, where first we enforce that some fragments of the molecule are independent and simulate the corresponding system, and then we see how much the vibrational features are affected by the decoupling. In this way, we can observe that decoupling some fragments of the salicylic acid does not lead to significant conformational changes within a short simulation time, while decoupling other fragments quickly leads to unrealistic phenomena. However, given a long enough simulation time, decoupling any pairs of molecular fragments will eventually lead to unphysical behaviors.

Previous infrared spectroscopic studies of salicylic acid (SA) focused primarily on the intramolecular H-bond between hydrogen 11 and oxygen 9[78−82] and hydrogen 11 and oxygen 10[78] (see the atom numbering in Figure 4) in the ground and first excited electronic states. These studies are mainly about the proton transfer process and deactivation of the excited electronic
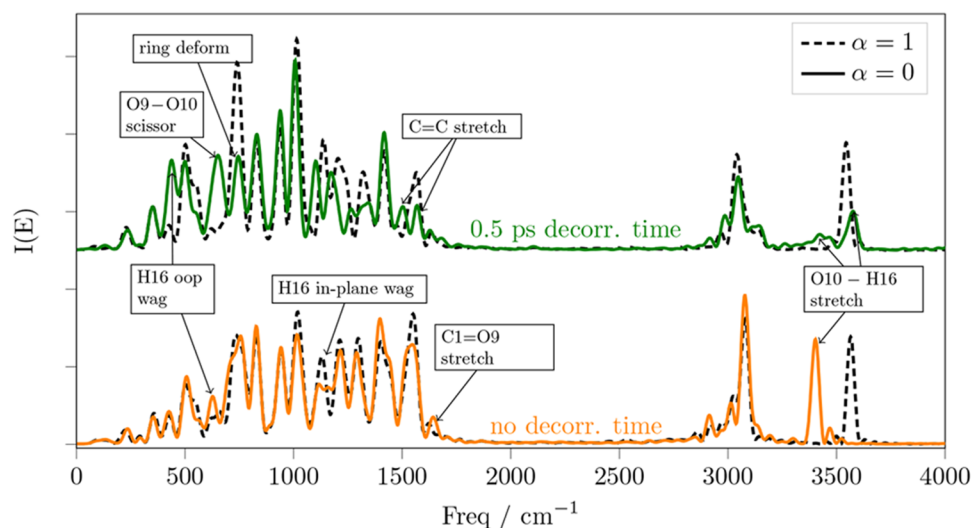
Figure 5. Power spectra of the salicylic acid (black dashed lines) and of the decoupled ($\alpha = 0$) C1—O10 and O10—H16 stretching modes, without waiting for any decorrelation time (solid orange line) and after 0.5 ps of decorrelation time (solid green line). The $\alpha = 1$ spectra are taken over the same time intervals of the corresponding $\alpha = 0$ spectra, accounting for the decorrelation time.

state via a radiationless mechanism, which is possible only if O9 and H11 are close enough, as shown in panel A of Figure 4, which is at the equilibrium geometry of the ground state. Below we show that such a configurational arrangement is stable over time only if the motion of O9 and H11, as well as of O9 and H16, is correlated. Furthermore, we investigate how much the carboxyl O − H stretching motion changes after artificially decoupling the different functional groups of the molecule.

Modes of Vibration of the Salicylic Acid. The salicylic acid molecule in its minimum energy geometry and within the harmonic approximation has 42 normal modes of vibration. Only some of them show a significant displacement of H11 or H16. However, beyond this approximation, the vibration of the hydroxyl and carboxylic acid fragments implies the significant displacement of H11 and H16 from their equilibrium condition. More specifically, approximating the O4−H11 and O10−H16 stretching modes with the harmonic approximation implies neglecting the coupling between these stretching modes and the twist, wag, and other complex motions that involve the whole OH and acid fragments. Four low frequency normal modes which are crucial for the salicylic acid vibrational motion are reported in panels B, C, and D of Figure 4. These modes are the acid group twist and the acid group out-of-plane (oop) modes, which involve, respectively, a twist and oop wagging of the carboxyl group with respect to the ring and the H16 oop motion, which is an out of plane wag of the H16 hydrogen. In addition, panel E of Figure 4 represents an out-of-plane mode that is delocalized over the three functional groups, involving H11, the ring, and the carboxylic acid group. We find that these four types of motion are those that are most significantly influenced by the pair decoupling of the hydrogen-bonded fragments and of the functional groups of SA. There are two main reasons for this. One is that they involve flexible regions of the molecule that easily couple with many other types of motion, and the other is that they break the directionality of the intramolecular hydrogen bonds.

An estimate of the vibrational frequencies of the pair decoupled system cannot be straightforwardly made within the harmonic approximation. In fact, while scaling the off-diagonal entries of the Hessian does not change the trace, which

is conserved in the diagonalization, it might change the magnitude of the eigenvalues. These changes imply that the pair-decoupled Hessian does not correspond to a stationary point configuration anymore. Consider, for instance, the case when the carboxyl group is decoupled from the hydroxyl groups; that is, all of the atoms in the carboxyl group are fully decoupled from the O and H atoms in the hydroxyl group. The normal-mode analysis of such a system at the original equilibrium geometry shows that the now unhindered acid twist mode has a frequency of about 3650 cm^{-1}. This is clearly not realistic. As mentioned above, the reason a normal-mode analysis can not be employed for an artificially decoupling analysis is that the equilibrium geometry of the system, at which the Hessian matrix is computed, is not a stationary point for the pair-decoupled system. Instead, computing the vibrational spectrum from the velocity correlation function does not suffer from this problem, and it can account for both anharmonicities, nonequilibrium, and dynamical couplings, which are lacking in the harmonic approximation.

Decoupling the O10—H16 and C1═O9 Stretching Modes. In this section, we apply the pair-decoupling idea in normal mode coordinates. In particular, we decorrelate the O−H stretching mode from the C═O stretching mode of the SA. Both the O−H and C═O stretchings are localized, meaning that we can interpret the O−H and C═O as two oscillators, where the O and H, and C and O, atoms are each connected by a spring. Even though the two oscillators are defined as independent when the molecule is at equilibrium, outside of equilibrium, the two oscillators are coupled, and each one depends on the other one's displacement. Furthermore, both oscillators also couple with all of the other oscillators that compose the SA vibrations. All the observations that we can make about the spectra in Figure 5 originate from the in-plane oscillations localized on the carboxylic acid group.

In Figure 5 we see that if we do not wait for any decorrelation time, the anharmonic vibrational frequency of the O−H stretch is red-shifted by nearly 160 cm^{-1}. This effect is evidently localized on the O−H, because the rest of the spectrum is only slightly changed by the decoupling. However, after 0.5 ps of decorrelation time, the decoupling also affects the O−C−O
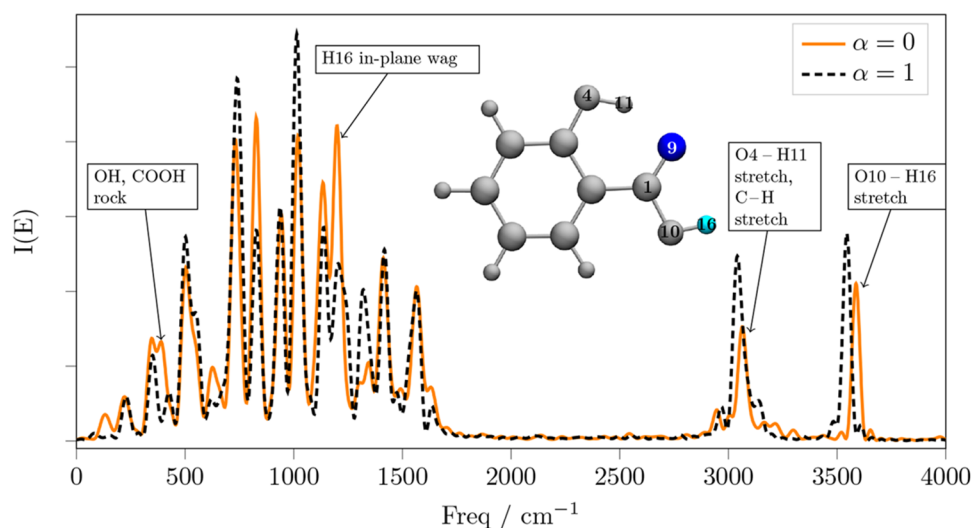
Figure 6. Power spectra of the salicylic acid (dashed black line) and of the O9−H16 decoupled ($\alpha = 0$) salicylic acid (solid orange line). Both spectra are recorded after discarding the first 0.5 ps of simulation.

scissoring mode, as well as modes that involve deformations and C=C stretchings. Moreover, after 0.5 ps of decorrelation time, the O−H stretching frequency becomes again similar to that of the nondecoupled system. Let us first focus on the bottom part of Figure 5. O9 and H16 are connected by a nondirectional hydrogen bond, that is weakened when the C=O and O−H bonds are stretched, because the two stretchings move the two atoms further apart. As the two oscillators stretch, they do not retain their reciprocal phase, because of the difference in mass between O and H. However, when we apply the decoupling, atom H16, would keep feeling the effect of a nonstretched C1−O9 oscillator, while oscillating back and forth. Thus we can see that the O−H stretching is hampered by the stable O9−H16 hydrogen bond. And we can quantify the importance of this effect by measuring the red-shift, which amounts to about 160 cm^{-1}.

Let us now focus on the top part of Figure 5. It shows that after 0.5 ps of decorrelation time, also the normal modes localized on the ring are affected by the decoupling. The involvement of the other modes implies a structural deformation of the entire molecule, compared to the nondecoupled dynamics. This brings the carboxylic acid O−H stretching frequency to about 3570 cm^{-1}, which is slightly blue-shifted from the nondecoupled spectrum (dashed line). This effect cannot be explained with simple arguments, because it evidently involves the whole molecule. As a matter of fact, the only portions of the spectrum that appear almost unaffected by the decoupling are C−H and hydroxyl O−H stretching modes, as well as the 700 and 1050 cm^{-1} region, which involves some ring deformation and breathing modes.

Decoupling the Carboxyl O9−H16 Hydrogen Bond. The first physical insight provided by the artificially decoupled O9−H16 hydrogen bond is that the O9 and H16 atoms do not oscillate synchronously anymore. This asynchronous motion induces an angular momentum that enhances the acid twist mode to the point that, after less than 300 fs, the carboxyl group attempts a 180° rotation around the C1−C2 axis. This 180 deg rotation has a potential barrier (computed as energy of the transition state minus energy of the minimum) in the original fully coupled system of 6366 cm^{-1} at pVDZ/DFT-PBE level of theory, and it should be an extremely rare event for the fully

coupled system, considering that the acid twist motion is initialized with less than 100 cm^{-1} of kinetic energy. To avoid this artificial twist, which is not in a fitted region of the given sGDML potential energy surface, we run simulations where the acid twist normal mode is kept at equilibrium and the O9−H16 hydrogen bond is still decoupled. The power spectrum of this simulation after a 0.5 ps decorrelation time is shown in Figure 6. In this case, we can observe a rather weak decorrelation effect in terms of the enhanced H16 rock and in-plane wag. The effects of such enhanced motion are given by the more intense bands in the 400 to 500 cm^{-1} and 1100 to 1300 cm^{-1} regions, as well as by the blueshift of the O10−H16 stretching mode, indicating a slightly weaker bond between O10 and H16.

From a fixed nuclei picture of the SA in the minimum energy configuration, one assumes that the hydrogen bond between O9 and H16 ensures that the carboxyl group remains confined in a plane and that H16 is oriented toward O9. In a dynamical picture instead, H16 oscillates out of the O9−C1−O10 plane and, given enough energy, it might overcome the potential barrier and get oriented toward C8 in a 180° rotation around the C1−O10 axis. In the original fully coupled system, this barrier height is about 4355 cm^{-1} at the pVDZ/DFT-PBE level of theory, and thus, the 180° rotation around the C1−O10 axis is a rare event, considering that the H16 oop wagging is initialized with 574 cm^{-1} of kinetic energy. The lack of synchronization in the out-of-plane motion of O9 and H16 redistributes some of the stretching vibrational energy to the out-of-plane modes to the point that the rotations become allowed. In fact, if one keeps the carboxyl twist mode at equilibrium, the H16 oop wag begins to oscillate significantly after about 1.2 ps of simulation, and we observe an attempt of 180° rotation of O10−H16 around the C1−O10 axis.

To sum up, the simulations of the O9−H16 decoupled SA provide two main physical insights. First of all, the artificial decoupling allows one to appreciate the importance of the synchronous oop vibration without which the carboxyl group would rotate, leading to a less stable minimum configuration. Second, the asynchronous motion of O9 and H16 leads to a fast vibrational energy redistribution in favor of the out of plane modes. As a secondary result, we see that the pair decoupling
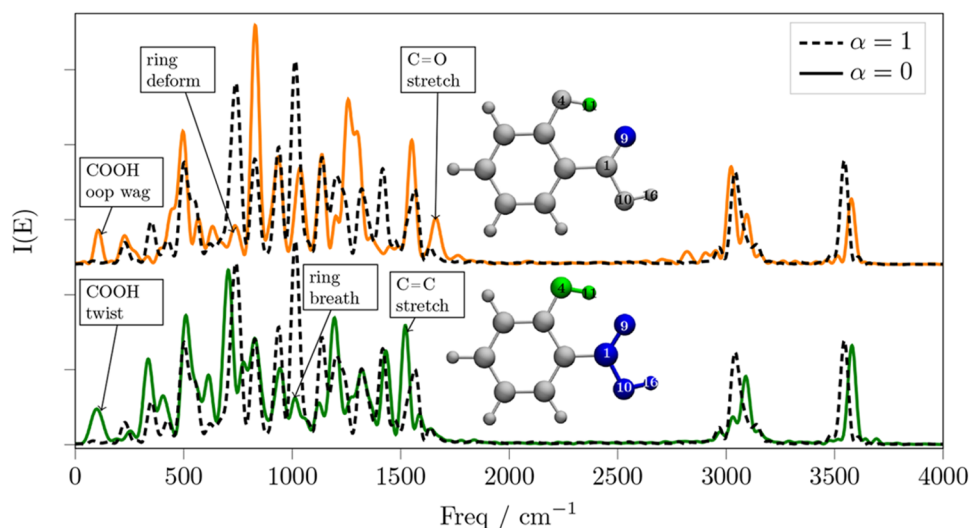
Figure 7. Power spectrum of the O9−H11 (solid orange line) and carboxyl−hydroxyl (solid green line) decoupled SA. The black dashed line is the power spectrum of the fully coupled simulation for comparison. The O9−H11 simulation (orange line) was obtained without evolving the acid twist mode, and both spectra are recorded after a 0.5 ps decorrelation time.
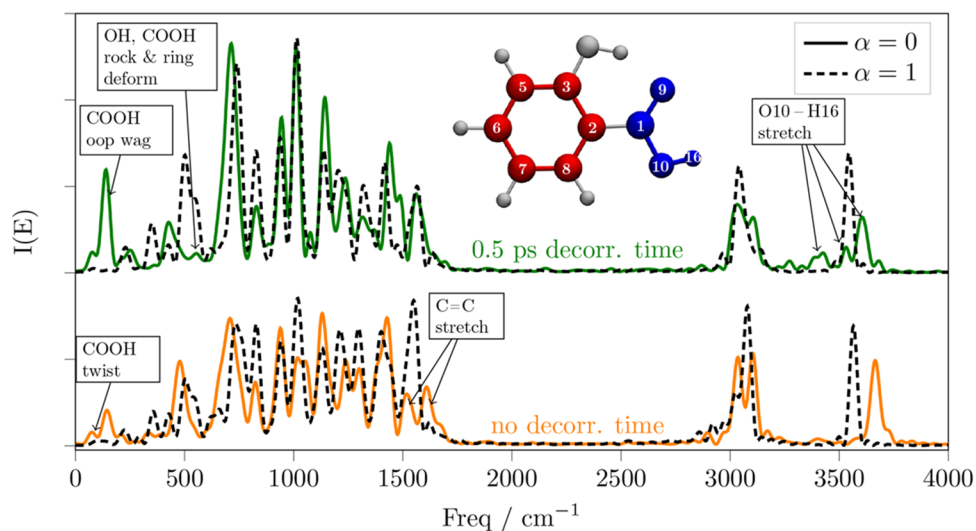

Figure 8. Spectrum of the SA with the carbon ring decoupled from the carboxyl group with $\alpha = 0$, without waiting for any decorrelation time (orange line) and after 0.5 ps of decorrelation time (green line). The dashed black lines are the spectra of the fully coupled system after discarding the corresponding amount of simulation time to match the decorrelation time.

allows one to quickly explore otherwise almost forbidden configurational regions of the potential surface.

Decoupling the Hydroxyl and Carboxylic Acid Groups. Here we show with our simulations why the O9−H11 hydrogen bond is fundamental for the planar shape of SA. As a consequence of the O9−H11 decoupling, the carboxyl group quickly initiates a large amplitude twist around the C1−C2 axis. This effect is similar to that one we have described in the previous section for the O9−H11 decoupling, where the decoupling artificially augments the kinetic energy of the acid twist motion represented in panel B of Figure 4. However, in this case, the role of the H-bond is very different. When O9 and H11 are decoupled, we argue that, even if the out-of-plane motions are not synchronized anymore, the oop wag of H11 remains coupled to O10 (and H16) and this coupling stimulates the oop motion of O10 (given the planarity of the carboxyl group). Eventually, the artificially enhanced oop wagging of H11 induces

an attempted 180° rotation of the carboxyl group around the C1−C2 axis. We deduce that there must be a strong synchronized interaction of each atom composing the *whole* hydroxyl group with each atom composing the *whole* carboxyl group. This interpretation of the importance of the H11 interaction with each singular atom composing the carboxyl group is validated by the fact that the acid twist motion is not enhanced when the whole carboxyl group is decoupled from the entire hydroxyl group. In Figure 7 we show both the spectrum when the O9−H11 interaction is decoupled and the acid twist mode is kept at equilibrium, and also the spectrum when it is the carboxyl−hydroxyl entire groups to be decoupled. Both spectra are recorded after 0.5 ps of decorrelation time, and both spectra show that the decoupling effect is quite significant in terms of vibrational energy redistribution. In fact, the ring breathing and ring deformation modes donate vibrational energy to the out-of-plane and C═O stretching modes at about 100 to 400, and 1600
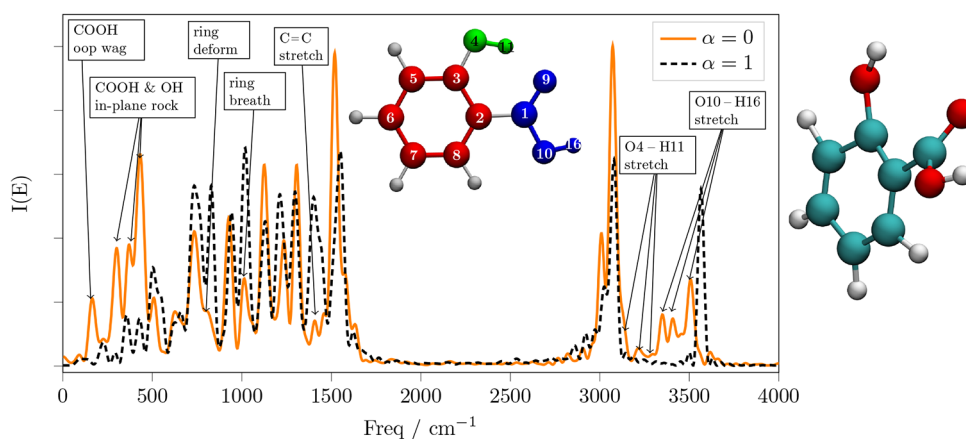
Figure 9. Spectrum of the SA with the carbon ring, carboxyl, and hydroxyl fragments all decoupled from each other with $\alpha = 0$ (orange line). The dashed black line is the spectrum of the fully coupled system for comparison. The inset on the right shows a representative frame of the simulation, as a result of the asynchronous motion of the three decoupled fragments.

cm^{-1}, respectively. More specifically, both the O9−H11 (orange line) and the carboxyl−hydroxyl decoupled spectra (green line) show that the decouplings induce significant vibrational energy redistributions in both the low frequency and fingerprint regions of the spectra, especially from the ring breathing and ring deformation modes, while the ring C−H and O−H stretching signals retain their kinetic energy on average. In both the O9−H11 and carboxyl−hydroxyl decouplings, the O−H stretching signals are mildly blue-shifted, indicating slightly weaker hydrogen bonds. In conclusion, these results clearly show that the intuitive picture of the independent functional groups in ortho position on the aromatic ring is partial to describing the appropriate vibrational dynamics of the SA and that the single atom−atom instantaneous couplings are essential for an accurate description of the interactions between the two functional groups. Our results also show that, surprisingly, the hydroxyl−carboxyl decoupled system provides a more realistic simulation of the O9−H11 decoupled one, because it does not induce the acid twist rotation.

Decoupling the Carboxyl and Ring Fragments. We find that the strongest decorrelation effects that impact the carboxyl O−H stretching mode occur when we decouple the motion of the ring from that of the carboxylic acid group. Specifically, the stretching O−H mode of the carboxylic acid is blueshifted by 103 cm^{-1} (from 3573 cm^{-1} of the fully coupled system to 3675 cm^{-1} of the $\alpha = 0$ decoupled system), as shown by the solid orange line in Figure 8. This effect is accompanied by an increased amplitude H16 oop wagging motion, which shortly after about 0.7 ps induces a 180° rotation of the O−H around the C1−O10 axis. If we keep the H16 oop wag at its equilibrium geometry to avoid this artificial rotation and record the power spectrum after 0.5 ps of decorrelation time, the O−H stretching of the acid is still blueshifted, although only by about 50 cm^{-1}, as shown by the solid green line of Figure 8. We conclude that the ring and carboxyl group decoupling have a strong effect on the carboxyl O−H motion. In fact, the C−H and hydroxyl O−H stretching signals around 3100 cm^{-1} are split, although not one of those hydrogens is decoupled.

The ring-carboxyl decoupling is a good example of how the two fragments can be interpreted as independently vibrating fragments just after the decoupling effect is turned on but not after 0.5 ps of decorrelation time. As a matter of fact, after the decorrelation time, the carboxyl O−H stretching signal is

smeared over nearly 500 cm^{-1}, mainly because of the enhanced oop wagging motion of the carboxyl group at 150 cm^{-1}. Nonetheless, most of the signals that involve ring stretching and other motions delocalized over the ring are still well recognizable in the fingerprint region of the spectrum, even after the decorrelation time. This observation shows that the effects of substituents on the vibrational features of the ring are mostly static, i.e., decorrelating the ring and carboxyl group vibrations does not induce very significant changes in the ring vibration frequencies. In fact, the fingerprint region of the spectra, between 700 and 1500 cm^{-1}, displays only some mild vibrational energy redistribution, in favors of the COOH twist and oop wag.

Decoupling the Entire SA into a Ring Part and Its Substituents. We conclude the results section by describing the scenario in which the SA molecule is decomposed into a ring and its substituents.

Figure 9 shows in orange the power spectrum of the SA where its functional groups are artificially decoupled. These groups are the aromatic carbon ring, the hydroxyl group, and the carboxyl group. After only 180 fs the decoupled system attempts a rotation of the O4−H11 group around the C3−O4 axis. Therefore, in this case, we decided to keep at the equilibrium position the mode which involves simultaneously the H11, ring, and acid oop displacements, which is indicated in panel E of Figure 4. Then, we record the spectrum without waiting for any decorrelation time. The strongest decorrelation effects observed in this dynamics consists of the enhanced acid twist, acid oop wag, and O−H oop wags, represented in panels B, C, and D of Figure 4. All of these effects quickly bring the molecule into very energetic regions of the PES, such as shown in the inset of Figure 9. From Figure 9 we can also see the same smearing of the carboxyl O−H stretching signal observed in the green spectrum of Figure 8, which occurs without any decorrelation time. This smearing effect is mainly due to the fact that the acid twists to a staggered position, and simultaneously, it also bends toward the ring (see the inset of Figure 9). In such a distorted configuration, the carboxyl O−H stretching motion strongly depends on the orientation of the O−H, as well as on the amount of twisting. In addition, the hydroxyl group is anomalously stretched apart from the ring and the C1−O10−H16 angle is highly increased, and similarly the signals of the hydroxyl O−H stretchings in the 3100 to 3300 cm^{-1} frequency interval are smeared. On the contrary the C−H stretching signals at about 3000 and 3100

cm^{-1} are only slightly shifted but not smeared. The redshift of the carboxyl stretching can also be interpreted from the distorted geometry of the SA. When the carboxyl group is twisted, its electron withdrawing effect on the ring is weakened and it becomes a weaker acid.

Analogously to what we observed in the ring–carboxyl decoupling of Figure 8, the fingerprint region of the spectrum remains reasonably similar (in terms of frequencies) to that of the fully coupled system. However, some of the ring modes, in particular the C=C stretchings, ring deformation, and breathing modes, donate vibrational energy to the hydroxyl and carboxyl groups in the 100 to 500 cm^{-1} region of the spectrum.

■ DISCUSSION AND CONCLUSIONS

In this paper, we introduce a pair-decoupling idea that offers a novel perspective for the study of relationships among groups of atoms or, more generally, of degrees of freedom in a molecule. The pair-decoupling idea is based on a simple, yet always applicable, mathematical definition: the Hessian matrix of a pair decoupled system is equal to the Hessian matrix of a normal system where some of the off-diagonal elements are weighted with an arbitrary coefficient α. The SEF algorithm that we introduce enforces the pair-decoupling idea for the molecular dynamics simulations of small and medium sized organic molecules. The simulations faithfully preserve the properties of symplectic symmetry of classical dynamics, in particular, the phase space conservation, in agreement with Liouville's theorem. The SEF method is effectively a symplectic integration technique of a system under a locally harmonic, "pair-decoupled" potential. The main disadvantage of the method is the requirement of 2 or 4 Hessian matrix calculations per time-step. However, this limitation could be alleviated by suitable numerical techniques.[84,85] This unavoidable feature limits the employment to middle-sized molecules and imposes the use of a computationally affordable potential for the electronic structure.

The application of our technique to salicylic acid has shown both intuitive behaviors of the pair-decoupled system, such as the rotation of the carbonyl in response to a decoupling of the hydroxyl-carboxyl H-bond, and less intuitive and surprising effects, such as the blueshift of the carboxyl O–H stretching frequency when the acid hydrogen is decoupled from the aromatic ring. We also showed that the synchronous vibrations of the atoms in the carboxyl and hydroxyl fragments are essential for the equilibrium configuration to be stable over time. As a consequence, in the absence of such couplings, the proton transfer photochemistry of the salicylic acid would be impossible. Ultimately, our simulations of the pair-decoupled salicylic acid show that the picture of the molecule as composed of independent vibrating fragments is partial and often unreliable. Since this intuitive picture is at the origin of the functional groups definition, we think that these results show how there are important exceptions to the functional group picture. In fact, an artificial decoupling of apparently unrelated groups of atoms may induce evident changes in the vibrational spectroscopy of the whole molecule. We think that these considerations are applicable to many other chemical systems and that our results open the path to further investigations thanks to the computational tool that we have presented. Furthermore, the pair-decoupled simulation technique can be used to validate applications that assume that a portion of the system is partially uncoupled from another such as in MCTDH and QM/MM calculations. The most practical way to do that in the case of QM/MM, for example, is to simulate the chosen pair-decoupled system at the MM level and verify whether it is an appropriate partition for the QM/MM calculation.

We hope that the pair-decoupling idea can inspire other less computationally expensive methods that can assess the importance of couplings in molecules. Finally, we believe that the SEF algorithm can help increase the sensibility of chemists toward the (unexpected) effects of approximations that involve artificially decoupled systems.

■ APPENDIX

The solutions of the fourth order map $\mathcal{M}_{n=4}$ are obtained from the following system of equations, which we derived, as explained in the Supporting Information, with the help of the SageMath[86] computer algebra system:

$$
\begin{cases}
a_1 + a_2 + a_3 + a_4 = 1 \\
b_1 + b_2 + b_3 + b_4 = 1 \\
(a_1 + a_2 + a_3 + a_4)b_1 + (a_2 + a_3 + a_4)b_2 + (a_3 + a_4)b_3 + a_4 b_4 = \frac{1}{2} \\
a_1 b_2 + (a_1 + a_2)b_3 + (a_1 + a_2 + a_3)b_4 = \frac{1}{2} \\
(a_1 + a_2 + a_3)a_4 b_4 + (a_1 a_2 + a_1 a_3 + a_1 a_4)b_2 \\
\quad + ((a_1 + a_2)a_3 + (a_1 + a_2)a_4)b_3 = \frac{1}{6} \\
a_1 b_1 b_2 + ((a_1 + a_2)b_1 + a_2 b_2)b_3 + ((a_1 + a_2 + a_3)b_1 + (a_2 + a_3)b_2 \\
\quad + a_3 b_3)b_4 = \frac{1}{6} \\
\frac{1}{2}a_1^2 b_2 + \frac{1}{2}(a_1^2 + 2a_1 a_2 + a_2^2)b_3 + \frac{1}{2}(a_1^2 + 2a_1 a_2 + a_2^2 \\
\quad + 2(a_1 + a_2)a_3 + a_3^2)b_4 = \frac{1}{6} \\
(a_1 a_2 + a_1 a_3 + a_1 a_4)b_1 b_2 + (((a_1 + a_2)a_3 + (a_1 + a_2)a_4)b_1 \\
\quad + (a_2 a_3 + a_2 a_4)b_2)b_3 + ((a_1 + a_2 + a_3)a_4 b_1 + (a_2 + a_3)a_4 b_2 \\
\quad + a_3 a_4 b_3)b_4 = \frac{1}{24} \\
\frac{1}{2}(a_1^2 + 2a_1 a_2 + a_2^2 + 2(a_1 + a_2)a_3 + a_3^2)a_4 b_4 \\
\quad + \frac{1}{2}(a_1^2 a_2 + a_1^2 a_3 + a_1^2 a_4)b_2 + \frac{1}{2}((a_1^2 + 2a_1 a_2 + a_2^2)a_3 \\
\quad + (a_1^2 + 2a_1 a_2 + a_2^2)a_4)b_3 = \frac{1}{24} \\
a_1^2 b_1 b_2 + ((a_1^2 + 2a_1 a_2 + a_2^2)b_1 + (a_1 a_2 + a_2^2)b_2)b_3 \\
\quad + ((a_1^2 + 2a_1 a_2 + a_2^2 + 2(a_1 + a_2)a_3 + a_3^2)b_1 \\
\quad + (a_1 a_2 + a_2^2 + (a_1 + 2a_2)a_3 + a_3^2)b_2 \\
\quad + ((a_1 + a_2)a_3 + a_3^2)b_3)b_4 = \frac{1}{8} \\
\frac{1}{6}a_1^3 b_2 + \frac{1}{6}(a_1^3 + 3a_1^2 a_2 + 3a_1 a_2^2 + a_2^3)b_3 + \frac{1}{6}(a_1^3 + 3a_1^2 a_2 + 3a_1 a_2^2 \\
\quad + a_2^3 + 3(a_1 + a_2)a_3^2 + a_3^3 + 3(a_1^2 + 2a_1 a_2 + a_2^2)a_3)b_4 = \frac{1}{24} \\
a_1 a_2 b_2 b_3 + ((a_1 + a_2)a_3 b_3 + (a_1 a_2 + a_1 a_3)b_2)b_4 = \frac{1}{24}
\end{cases}
\tag{17}
$$

It is important to notice that this system is undetermined. We need to choose another constraint to effectively get numerical values for a_k and b_k coefficients. A reasonable choice to saturate the system is to set $b_1 = 0$, which leads to three possible sets of real coefficients: the first set was first published by Campostrini and Rossi,[49] then independently by Forest and Ruth,[50] Yoshida (who attributes the origin to Neri),[41] and Candy and Rozmus.[53]

$$\begin{cases} a_1 = \dfrac{2^{1/3} + 2^{-1/3} + 2}{6} \\[2mm] a_2 = -\dfrac{2^{1/3} + 2^{-1/3} - 1}{6} \\[2mm] a_3 = -\dfrac{2^{1/3} + 2^{-1/3} - 1}{6} \\[2mm] a_4 = \dfrac{2^{1/3} + 2^{-1/3} + 2}{6} \\[2mm] b_1 = 0 \\[2mm] b_2 = \dfrac{2^{4/3} + 2^{2/3} + 4}{6} \\[2mm] b_3 = -\dfrac{2^{7/3} + 2^{5/3} + 2}{6} \\[2mm] b_4 = \dfrac{2^{4/3} + 2^{2/3} + 4}{6} \end{cases} \tag{18}$$

while the second was reported by Brewer et al.,[56]

$$\begin{cases} a_1 = -\sqrt{3}/6 + 1/2 \\[1mm] a_2 = \sqrt{3}/3 \\[1mm] a_3 = -\sqrt{3}/3 \\[1mm] a_4 = \sqrt{3}/6 + 1/2 \\[1mm] b_1 = 0 \\[1mm] b_2 = \sqrt{3}/6 + 1/4 \\[1mm] b_3 = 1/2 \\[1mm] b_4 = -\sqrt{3}/6 + 1/4 \end{cases} \tag{19}$$

■ ASSOCIATED CONTENT

Data Availability Statement

Data and computer codes are available upon reasonable request to the corresponding author.

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00553.

> A more detailed description of the strategy we used to derive the system in eq 16 of the Appendix; a more detailed description of the modified symplectic map, and of the choices of the c_k coefficients for SEF4-I to SEF4-IV variants; equations to implement the monodromy matrix time evolution alongside the SEF algorithm; pictures that show the $\Upsilon(t)$, $\tau(t)$ functions for the decoupled ($\alpha = 0$) water molecule (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Michele Ceotto − *Dipartimento di Chimica, Università degli Studi di Milano, 20133 Milano, Italy;* ⓘ orcid.org/0000-0002-8270-3409; Email: michele.ceotto@unimi.it

Author

Michele Gandolfi − *Dipartimento di Chimica, Università degli Studi di Milano, 20133 Milano, Italy;* ⓘ orcid.org/0000-0001-8319-3773

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.3c00553

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Hamm, P.; Lim, M.; Hochstrasser, R. M. Structure of the amide I band of peptides measured by femtosecond nonlinear-infrared spectroscopy. *J. Phys. Chem. B* **1998**, *102*, 6123−6138.

(2) Ghosh, A.; Ostrander, J. S.; Zanni, M. T. Watching proteins wiggle: Mapping structures with two-dimensional infrared spectroscopy. *Chem. Rev.* **2017**, *117*, 10726−10759.

(3) Ramesh, P.; Loring, R. F. Thermal population fluctuations in two-dimensional infrared spectroscopy captured with semiclassical mechanics. *J. Phys. Chem. B* **2018**, *122*, 3647−3654.

(4) Jansen, T. l. C.; Saito, S.; Jeon, J.; Cho, M. Theory of coherent two-dimensional vibrational spectroscopy. *J. Chem. Phys.* **2019**, *150*, 100901.

(5) Milo, A.; Bess, E. N.; Sigman, M. S. Interrogating selectivity in catalysis using molecular vibrations. *Nature* **2014**, *507*, 210−214.

(6) Bess, E. N.; Guptill, D. M.; Davies, H. M.; Sigman, M. S. Using IR vibrations to quantitatively describe and predict site-selectivity in multivariate Rh-catalyzed C−H functionalization. *Chemical science* **2015**, *6*, 3057−3062.

(7) Yada, A.; Nagata, K.; Ando, Y.; Matsumura, T.; Ichinoseki, S.; Sato, K. Machine learning approach for prediction of reaction yield with simulated catalyst parameters. *Chem. Lett.* **2018**, *47*, 284−287.

(8) Mukamel, S. Multidimensional femtosecond correlation spectroscopies of electronic and vibrational excitations. *Annu. Rev. Phys. Chem.* **2000**, *51*, 691−729.

(9) Sertcan, B.; Mousavi, S. J.; Iannuzzi, M.; Hamm, P. Low-Frequency Anharmonic Couplings in Crystalline Bromoform: Theory. *J. Chem. Phys.* **2023**, 014203.

(10) Buhrke, D.; Michael, N.; Hamm, P. Vibrational couplings between protein and cofactor in bacterial phytochrome Agp1 revealed by 2D-IR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, No. e2206400119.

(11) Bonnet, L.; Rayez, J. Quasiclassical trajectory method for molecular scattering processes: necessity of a weighted binning approach. *Chem. Phys. Lett.* **1997**, *277*, 183−190.

(12) Wehrle, M.; Sulc, M.; Vanicek, J. On-the-fly ab initio semiclassical dynamics: Identifying degrees of freedom essential for emission spectra of oligothiophenes. *J. Chem. Phys.* **2014**, *140*, 244114.

(13) Antipov, S. V.; Ye, Z.; Ananth, N. Dynamically consistent method for mixed quantum-classical simulations: A semiclassical approach. *J. Chem. Phys.* **2015**, *142*, 184102.

(14) Ceotto, M.; Di Liberto, G.; Conte, R. Semiclassical "Divide-and-Conquer" Method for Spectroscopic Calculations of High Dimensional Molecular Systems. *Phys. Rev. Lett.* **2017**, *119*, No. 010401.

(15) Di Liberto, G.; Conte, R.; Ceotto, M. "Divide and conquer" semiclassical molecular dynamics: A practical method for spectroscopic

calculations of high dimensional molecular systems. *J. Chem. Phys.* **2018**, *148*, No. 014307.

(16) Gabas, F.; Di Liberto, G.; Conte, R.; Ceotto, M. Protonated glycine supramolecular systems: the need for quantum dynamics. *Chem. Sci.* **2018**, *9*, 7894−7901.

(17) Church, M. S.; Ananth, N. Semiclassical dynamics in the mixed quantum-classical limit. *J. Chem. Phys.* **2019**, *151*, 134109.

(18) Bertaina, G.; Di Liberto, G.; Ceotto, M. Reduced rovibrational coupling Cartesian dynamics for semiclassical calculations: Application to the spectrum of the Zundel cation. *J. Chem. Phys.* **2019**, *151*, 114307.

(19) Cazzaniga, M.; Micciarelli, M.; Moriggi, F.; Mahmoud, A.; Gabas, F.; Ceotto, M. Anharmonic calculations of vibrational spectra for molecular adsorbates: A divide-and-conquer semiclassical molecular dynamics approach. *J. Chem. Phys.* **2020**, *152*, 104104.

(20) Gandolfi, M.; Rognoni, A.; Aieta, C.; Conte, R.; Ceotto, M. Machine learning for vibrational spectroscopy via divide-and-conquer semiclassical initial value representation molecular dynamics with application to N-methylacetamide. *J. Chem. Phys.* **2020**, *153*, 204104.

(21) Rognoni, A.; Conte, R.; Ceotto, M. How many water molecules are needed to solvate one? *Chemical Science* **2021**, *12*, 2060−2064.

(22) Gabas, F.; Conte, R.; Ceotto, M. Quantum vibrational spectroscopy of explicitly solvated thymidine in semiclassical approximation. *J. Phys. Chem. Lett.* **2022**, *13*, 1350−1355.

(23) Morzan, U. N.; Alonso de Armino, D. J.; Foglia, N. O.; Ramirez, F.; Gonzalez Lebrero, M. C.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in complex environments from QM−MM simulations. *Chem. Rev.* **2018**, *118*, 4071−4113.

(24) Thomas, P. S.; Carrington, T., Jr Using Nested Contractions and a Hierarchical Tensor Format To Compute Vibrational Spectra of Molecules with Seven Atoms. *J. Phys. Chem. A* **2015**, *119*, 13074−13091.

(25) Rakhuba, M.; Oseledets, I. Calculating vibrational spectra of molecules using tensor train decomposition. *J. Chem. Phys.* **2016**, *145*, 124101.

(26) Thomas, P. S.; Carrington, T., Jr; Agarwal, J.; Schaefer, H. F., III Using an iterative eigensolver and intertwined rank reduction to compute vibrational spectra of molecules with more than a dozen atoms: Uracil and naphthalene. *J. Chem. Phys.* **2018**, *149*, No. 064108.

(27) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H.-D. The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. *Physics reports* **2000**, *324*, 1−105.

(28) Mendive-Tapia, D.; Meyer, H.-D.; Vendrell, O. Optimal Mode Combination in the Multiconfiguration Time-Dependent Hartree Method through Multivariate Statistics: Factor Analysis and Hierarchical Clustering. *J. Chem. Theory Comput.* **2023**, *19*, 1144−1156.

(29) Hino, K.; Kurashige, Y. Matrix Product State Formulation of the MCTDH Theory in Local Mode Representations for Anharmonic Potentials. *J. Chem. Theory Comput.* **2022**, *18*, 3347−3356.

(30) Larsson, H. R. Computing vibrational eigenstates with tree tensor network states (TTNS). *J. Chem. Phys.* **2019**, *151*, 204102.

(31) Uzer, T.; Miller, W. Theories of intramolecular vibrational energy transfer. *Physics reports* **1991**, *199*, 73−146.

(32) Buchner, M.; Ladanyi, B. M.; Stratt, R. M. The short-time dynamics of molecular liquids. Instantaneous-normal-mode theory. *J. Chem. Phys.* **1992**, *97*, 8522−8535.

(33) Zhang, R. M.; Xu, X.; Truhlar, D. G. Observing Intramolecular Vibrational Energy Redistribution via the Short-Time Fourier Transform. *J. Phys. Chem. A* **2022**, *126*, 3006−3014.

(34) Bastida, A.; Soler, M. A.; Zúñiga, J.; Requena, A.; Kalstein, A.; Fernández-Alberti, S. Instantaneous normal modes, resonances, and decay channels in the vibrational relaxation of the amide I mode of N-methylacetamide-D in liquid deuterated water. *J. Chem. Phys.* **2010**, *132*, 1.3435212.

(35) Alfonso-Hernandez, L.; Athanasopoulos, S.; Tretiak, S.; Miguel, B.; Bastida, A.; Fernandez-Alberti, S. Vibrational energy redistribution during donor−acceptor electronic energy transfer: criteria to identify

subsets of active normal modes. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18454−18466.

(36) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(37) Voter, A. F. A method for accelerating the molecular dynamics simulation of infrequent events. *J. Chem. Phys.* **1997**, *106*, 4665−4677.

(38) Ceotto, M.; Ayton, G. S.; Voth, G. A. Accelerated superposition state molecular dynamics for condensed phase systems. *J. Chem. Theory Comput.* **2008**, *4*, 560−568.

(39) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 826−843.

(40) Moscato, D.; Gabas, F.; Conte, R.; Ceotto, M. Vibrational spectroscopy simulation of solvation effects on a G-quadruplex. *J. Biomol. Struct. Dyn.* **2023**, 1.

(41) Yoshida, H. Construction of higher order symplectic integrators. *Phys. Lett. A* **1990**, *150*, 262−268.

(42) McLachlan, R. I.; Quispel, G. R. W. Splitting methods. *Acta Numerica* **2002**, *11*, 341−434.

(43) Chin, S. A. Structure of numerical algorithms and advanced mechanics. *American Journal of Physics* **2020**, *88*, 883−894.

(44) Tao, M. Explicit symplectic approximation of nonseparable Hamiltonians: Algorithm and long time performance. *Phys. Rev. E* **2016**, *94*, No. 043303.

(45) Strang, G. On the construction and comparison of difference schemes. *SIAM journal on numerical analysis* **1968**, *5*, 506−517.

(46) Suzuki, M. Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations. *Phys. Lett. A* **1990**, *146*, 319−323.

(47) See the historical development section in the work of McLachlan and Quispel: McLachlan, R I.; Quispel, G. R. W. Splitting methods. *Acta Numerica* **2002**, 341.

(48) Ruth, R. D. A canonical integration technique. *IEEE Trans. Nucl. Sci.* **1983**, *30*, 2669−2671.

(49) Campostrini, M.; Rossi, P. A comparison of numerical algorithms for dynamical fermions. *Nuclear Physics B* **1990**, *329*, 753−764.

(50) Forest, E.; Ruth, R. D. Fourth-order symplectic integration. *Physica D: Nonlinear Phenomena* **1990**, *43*, 105−117.

(51) McLachlan, R. I.; Atela, P. The accuracy of symplectic integrators. *Nonlinearity* **1992**, *5*, 541.

(52) Gray, S. K.; Noid, D. W.; Sumpter, B. G. Symplectic integrators for large scale molecular dynamics simulations: A comparison of several explicit methods. *J. Chem. Phys.* **1994**, *101*, 4062−4072.

(53) Candy, J.; Rozmus, W. A symplectic integration algorithm for separable Hamiltonian functions. *J. Comput. Phys.* **1991**, *92*, 230−256.

(54) Chin, S. A. Symplectic integrators from composite operator factorizations. *Phys. Lett. A* **1997**, *226*, 344−348.

(55) Creutz, M. Lattice fields and strong interactions. Lectures at the NATO International Advanced Course on the Nuclear Equation of State, Pensicola, Spain, May 21-June 3, 1989.

(56) Brewer, M. L.; Hulme, J. S.; Manolopoulos, D. E. Semiclassical dynamics in up to 15 coupled vibrational degrees of freedom. *J. Chem. Phys.* **1997**, *106*, 4832−4839.

(57) Suzuki, M. Hybrid exponential product formulas for unbounded operators with possible applications to Monte Carlo simulations. *Phys. Lett. A* **1995**, *201*, 425−428.

(58) Choi, S.; Vaníček, J. High-order geometric integrators for representation-free Ehrenfest dynamics. *J. Chem. Phys.* **2021**, *155*, 124104.

(59) Church, M. S.; Hele, T. J.; Ezra, G. S.; Ananth, N. Nonadiabatic semiclassical dynamics in the mixed quantum-classical initial value representation. *J. Chem. Phys.* **2018**, *148*, 102326.

(60) Richardson, J. O.; Meyer, P.; Pleinert, M.-O.; Thoss, M. An analysis of nonadiabatic ring-polymer molecular dynamics and its application to vibronic spectra. *Chem. Phys.* **2017**, *482*, 124−134.

(61) Kelly, A.; van Zon, R.; Schofield, J.; Kapral, R. Mapping quantum-classical Liouville equation: Projectors and trajectories. *J. Chem. Phys.* **2012**, *136*, No. 084101.

(62) Choi, S.; Vaníček, J. Efficient geometric integrators for nonadiabatic quantum dynamics. I. The adiabatic representation. *J. Chem. Phys.* **2019**, *150*, 204112.

(63) Roulet, J.; Choi, S.; Vaníček, J. Efficient geometric integrators for nonadiabatic quantum dynamics. II. The diabatic representation. *J. Chem. Phys.* **2019**, *150*, 204113.

(64) Vaníček, J. Family of Gaussian wavepacket dynamics methods from the perspective of a nonlinear Schrodinger equation. *arXiv Preprint* **2023**, No. 2302.10221.

(65) Dressler, S.; Thiel, W. Anharmonic force fields from density functional theory. *Chem. Phys. Lett.* **1997**, *273*, 71−78.

(66) Martin, J.; Lee, T. J.; Taylor, P. R. An accurate ab initio quartic force field for formaldehyde and its isotopomers. *J. Mol. Spectrosc.* **1993**, *160*, 105−116.

(67) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine Learning of Accurate Energy-conserving Molecular Force Fields. *Science Advances* **2017**, *3*, No. e1603015.

(68) Chmiela, S.; Sauceda, H. E.; Müller, K.-R.; Tkatchenko, A. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.* **2018**, *9*, 3887.

(69) Chmiela, S.; Sauceda, H. E.; Poltavsky, I.; Müller, K.-R.; Tkatchenko, A. sGDML: Constructing accurate and data efficient molecular force fields using machine learning. *Comput. Phys. Commun.* **2019**, *240*, 38−45.

(70) Sauceda, H. E.; Chmiela, S.; Poltavsky, I.; Müller, K.-R.; Tkatchenko, A. Molecular force fields with gradient-domain machine learning: Construction and application to dynamics of small molecules with coupled cluster forces. *J. Chem. Phys.* **2019**, *150*, 114102.

(71) Chmiela, S.; Sauceda, H. E.; Tkatchenko, A.; Müller, K.-R. *Machine Learning Meets Quantum Physics*; Springer International Publishing, 2020; pp 129−154.

(72) Sauceda, H. E.; Chmiela, S.; Poltavsky, I.; Müller, K.-R.; Tkatchenko, A. *Machine Learning Meets Quantum Physics*; Springer International Publishing, 2020; pp 277−307.

(73) Sauceda, H. E.; Gastegger, M.; Chmiela, S.; Müller, K.-R.; Tkatchenko, A. Molecular force fields with gradient-domain machine learning (GDML): Comparison and synergies with classical force fields. *J. Chem. Phys.* **2020**, *153*, 124109.

(74) Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Physical review letters* **2009**, *102*, No. 073005.

(75) Rognoni, A.; Conte, R.; Ceotto, M. Caldeira−Leggett model vs ab initio potential: A vibrational spectroscopy test of water solvation. *J. Chem. Phys.* **2021**, *154*, No. 094106.

(76) Kaledin, A. L.; Miller, W. H. Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels. *J. Chem. Phys.* **2003**, *118*, 7174−7182.

(77) Kaledin, A. L.; Miller, W. H. Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels. II. Application to H2CO, NH3, CH4, CH2D2. *J. Chem. Phys.* **2003**, *119*, 3078−3084.

(78) Yahagi, T.; Fujii, A.; Ebata, T.; Mikami, N. Infrared spectroscopy of the OH stretching vibrations of jet-cooled salicylic acid and its dimer in S0 and S1. *J. Phys. Chem. A* **2001**, *105*, 10673−10680.

(79) Sobolewski, A. L.; Domcke, W. Intramolecular hydrogen bonding in the S1 (π π*) excited state of anthranilic acid and salicylic acid: TDDFT calculation of excited-state geometries and infrared spectra. *J. Phys. Chem. A* **2004**, *108*, 10917−10922.

(80) Abd El-Hakam Abou El-Nasr, E.; Fujii, A.; Ebata, T.; Mikami, N. Substitution effects on the excited-state intramolecular proton transfer of salicylic acid: an infrared spectroscopic study on the OH stretching vibrations of jet-cooled 5-methoxysalicylic acid. *Chemical physics letters* **2003**, *376*, 788−793.

(81) Abd El-Hakam Abou El-Nasr, E.; Fujii, A.; Ebata, T.; Mikami, N. Infrared and ultraviolet laser spectroscopy of jet-cooled substituted salicylic acids; substitution effects on the excited state intramolecular proton transfer in salicylic acid. *Mol. Phys.* **2005**, *103*, 1561−1572.

(82) Raeker, T.; Hartke, B. Full-dimensional excited-state intramolecular proton transfer dynamics of salicylic acid. *J. Phys. Chem. A* **2017**, *121*, 5967−5977.

(83) Notice that in in panel B of Figure 4 the acid twist and acid oop normal modes involve most of the atoms in the molecule, not just the indicated ones. As a matter of fact, the normal modes must keep the center of mass fixed and the angular momentum 0. We chose not to put arrows on all the atoms for the sake of clarity, as in another frame of reference we can imagine the normal modes to follow approximately only the indicated motions.

(84) Zhuang, Y.; Siebert, M. R.; Hase, W. L.; Kay, K. G.; Ceotto, M. Evaluating the Accuracy of Hessian Approximations for Direct Dynamics Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 54−64.

(85) Ceotto, M.; Zhuang, Y.; Hase, W. L. Accelerated direct semiclassical molecular dynamics using a compact finite difference Hessian scheme. *J. Chem. Phys.* **2013**, *138*, No. 054116.

(86) The Sage Developers *SageMath, the Sage Mathematics Software System*, ver. 9.5; 2022; https://www.sagemath.org.