The Journal of Chemical Physics

Machine learning for vibrational spectroscopy via divide-and-conquer semiclassical initial value representation molecular dynamics with application to *N*-methylacetamide

Cite as: J. Chem. Phys. **153**, 204104 (2020); https://doi.org/10.1063/5.0031892
Submitted: 05 October 2020 . Accepted: 05 November 2020 . Published Online: 23 November 2020

Michele Gandolfi, Alessandro Rognoni, Chiara Aieta, Riccardo Conte, and Michele Ceotto

AIP Publishing

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Machine learning for vibrational spectroscopy via divide-and-conquer semiclassical initial value representation molecular dynamics with application to *N*-methylacetamide

Cite as: J. Chem. Phys. 153, 204104 (2020); doi: 10.1063/5.0031892
Submitted: 5 October 2020 • Accepted: 5 November 2020 •
Published Online: 23 November 2020

Michele Gandolfi, Alessandro Rognoni, Chiara Aieta, Riccardo Conte, and Michele Ceotto[a)]

AFFILIATIONS

Dipartimento di Chimica, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy, https://sites.unimi.it/ceotto/

Note: This paper is part of the JCP Special Topic on Quantum Dynamics with *Ab Initio* Potentials.
[a)]**Author to whom correspondence should be addressed:** michele.ceotto@unimi.it

ABSTRACT

A machine learning algorithm for partitioning the nuclear vibrational space into subspaces is introduced. The subdivision criterion is based on Liouville's theorem, i.e., the best preservation of the unitary of the reduced dimensionality Jacobian determinant within each subspace along a probe full-dimensional classical trajectory. The algorithm is based on the idea of evolutionary selection, and it is implemented through a probability graph representation of the vibrational space partitioning. We interface this customized version of genetic algorithms with our divide-and-conquer semiclassical initial value representation method for the calculation of molecular power spectra. First, we benchmark the algorithm by calculating the vibrational power spectra of two model systems, for which the exact subspace division is known. Then, we apply it to the calculation of the power spectrum of methane. Exact calculations and full-dimensional semiclassical spectra of this small molecule are available and provide an additional test of the accuracy of the new approach. Finally, the algorithm is applied to the divide-and-conquer semiclassical calculation of the power spectrum of 12-atom *trans*-*N*-methylacetamide.

I. INTRODUCTION

When full-dimensional calculations are not computationally feasible, one needs to introduce some sort of approximation. This happens quite often in quantum molecular dynamics because the full-dimensional potential $V(\mathbf{q})$ is generally non-separable. However, if the system potential were made of the sum of a finite number of lower-dimensional terms of the type $V_a(q_1, q_2) + V_b(q_3) + V_c(q_4, q_5, q_6) + \cdots$, then each collection of variables within the several terms would compose a sensible subspace suitable for an independent calculation. Unfortunately, the nuclear potential can not be written as a sum of lower-dimensional terms, however one may wonder which is the decomposition of the full-dimensional vibrational space into subspaces such that the couplings between different subspace modes are minimized. The main goal of this paper is to provide a possible answer to this issue by means of a clustering optimization algorithm.

The algorithm we propose has its roots in the sound ground of Genetic Algorithms (GAs), but it is different for technical definitions and implementation, and we will refer to it with the more general label of evolutionary algorithm. The theoretical background and implementation of GAs could be traced back to the works of Holland, Goldberg, and Henry.[1–3] Since then, they have been successfully applied to various problems in analytical chemistry and chemometrics, such as the analysis of nuclear magnetic resonance (NMR) pulse patterns from complex molecules,[4] the optimal choice of wavelengths to determine the concentration of a sample, and the conformational analysis.[5] GAs have been proven to be the first choice in many cases of feature selection in regression and classification problems in general[6,7] and in quantitative structure activity relationships in particular.[7,8] Evolutionary algorithms have also been used in materials science to study the dynamical properties of molecules on surfaces during molecular dynamics simulations.[9]

We test the accuracy of the new optimization algorithm when applied to the calculation of vibrational power spectra using semiclassical molecular dynamics. Specifically, we employ the divide-and-conquer approach to the time-averaged semiclassical initial value representation (DC-SCIVR) method for nuclear power spectrum calculations.[10] In the DC-SCIVR strategy, the vibrational space is divided into subspaces to overcome the issue of the full-dimensional calculation. So far, to the best of our knowledge, three methods for the partition of the full-dimensional vibrational problem into lower-dimensional subspaces have been proposed. These are the Hessian method,[11] the Wehrle–Šulc–Vaníček method,[12] and the Jacobi method.[11] According to the latter, the residual coupling between subspaces is estimated by recurring to Liouville's theorem. In few words, the best clustering of vibrational modes is the one where each subspace preserves as much as possible its phase-space volume during a classical, full-dimensional trajectory. This approach has the advantage of being based on dynamics, i.e., it depends on both the nuclear kinetic and potential contributions. Originally, we applied this method by searching over all possible subspace combinations. As a matter of fact, the problem scales approximately as a binomial factor times the number of subspaces n_s, i.e., $\sim n_s \times F!/D!(F-D)!$, where F is the number of degrees of freedom and D is the size of a subspace. Furthermore, it is not possible to know in advance how many subspaces one should choose and the size of each one.

In this work, we introduce an evolutionary algorithm and another simplified and less demanding approach. Both are able to automatize these choices and highly reduce the computational cost of the optimization. This paper is organized as follows: In Sec. II, we recap the DC-SCIVR method and present the optimization algorithms. Section III presents our results for spectroscopic calculations, and Sec. IV concludes this paper and offers some perspectives.

II. METHODS

A. Semiclassical spectra

The semiclassical power spectrum $I(E)$ of a system described by the Hamiltonian $\hat{H}$ is equal to the Fourier-transformed wavepacket survival amplitude (in atomic units),[13]

$$I(E) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{iEt}\langle\chi|\chi(t)\rangle dt, \tag{1}$$

where $|\chi(t)\rangle = e^{-i\hat{H}t}|\chi\rangle$ is the quantum time evolution of the arbitrary reference state $|\chi\rangle$. By writing the reference state as a linear combination of the Hamiltonian eigenstates $|\psi_j\rangle$, i.e., $|\chi\rangle = \sum_j c_j|\psi_j\rangle$, it can be shown that

$$I(E) = \sum_j |c_j|^2\delta(E-E_j). \tag{2}$$

Hence, the power spectrum is equal to a sum of delta functions centered at the vibrational frequencies E_j. A convenient way to calculate the formula in Eq. (1) is given by the time-averaging semiclassical initial value representation (TA-SCIVR) method.[14–23] This is obtained by applying a time-averaging filter to the Heller–Herman–Kluk–Kay (HHKK) propagator,[24–36] which defines the approximate quantum time evolution. The final TA-SCIVR expression of Eq. (1) for a system characterized by F degrees of freedom reads

$$I(E) = \left(\frac{1}{2\pi}\right)^F \iint d\mathbf{p}_0 d\mathbf{q}_0 \frac{1}{2\pi T} \times \left|\int_0^T dt\, \langle\chi|\mathbf{p}_t\mathbf{q}_t\rangle e^{i\left(S_t(\mathbf{p}_0,\mathbf{q}_0)+\phi_t(\mathbf{p}_0,\mathbf{q}_0)+Et\right)}\right|^2, \tag{3}$$

where T is the total simulation time, $S_t(\mathbf{p}_0, \mathbf{q}_0)$ is the instantaneous action of the classically evolved trajectory $(\mathbf{p}_t, \mathbf{q}_t)$, and the phase-space integration is performed on the initial trajectory momenta $\mathbf{p}_0$ and positions $\mathbf{q}_0$. In the previous equation, $|\mathbf{p}_t\mathbf{q}_t\rangle$ are coherent states with the following form in position representation:[37]

$$\langle\mathbf{x}|\mathbf{p}_t\mathbf{q}_t\rangle = \left(\frac{\det(\boldsymbol{\gamma})}{\pi^F}\right)^{\frac{1}{4}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{q}_t)^{\mathrm{T}}\boldsymbol{\gamma}(\mathbf{x}-\mathbf{q}_t) + i\mathbf{p}_t^{\mathrm{T}}(\mathbf{x}-\mathbf{q}_t)\right], \tag{4}$$

where $\boldsymbol{\gamma}$ is an $F \times F$ diagonal matrix whose elements are chosen to be numerically equal to the harmonic frequencies of the system. In Eq. (3), $\phi_t(\mathbf{p}_0, \mathbf{q}_0)$ is the phase of the HHKK prefactor,[38]

$$\phi_t(\mathbf{p}_0,\mathbf{q}_0) = phase\left[\sqrt{\frac{1}{2^F}|\mathbf{M}_{\mathbf{qq}} + \boldsymbol{\gamma}^{-1}\mathbf{M}_{\mathbf{pp}}\boldsymbol{\gamma} - i\mathbf{M}_{\mathbf{qp}}\boldsymbol{\gamma} + i\boldsymbol{\gamma}^{-1}\mathbf{M}_{\mathbf{pq}}|}\right], \tag{5}$$

where $\mathbf{M}_{\mathbf{ij}}$ with $\mathbf{i}, \mathbf{j} = \mathbf{p}, \mathbf{q}$ are the elements of the Jacobian (monodromy) matrix,

$$\mathbf{J} = \begin{pmatrix}\mathbf{M}_{\mathbf{pp}} & \mathbf{M}_{\mathbf{pq}}\\ \mathbf{M}_{\mathbf{qp}} & \mathbf{M}_{\mathbf{qq}}\end{pmatrix} = \begin{pmatrix}\frac{\partial\mathbf{p}_t}{\partial\mathbf{p}_0} & \frac{\partial\mathbf{p}_t}{\partial\mathbf{q}_0}\\ \frac{\partial\mathbf{q}_t}{\partial\mathbf{p}_0} & \frac{\partial\mathbf{q}_t}{\partial\mathbf{q}_0}\end{pmatrix}. \tag{6}$$

The determinant $|\det(\mathbf{J})|$ is always equal to 1 along the trajectory, in accordance with Liouville's theorem. The major problem associated with the TA-SCIVR spectral density calculation is represented by the computational cost of the Monte Carlo phase-space integration in Eq. (3).[39] To overcome this problem, the multiple coherent state SCIVR (MC SCIVR) has been introduced. This method relies on the idea that the most important contribution to the spectrum comes from the trajectories whose energies are as close as possible to the quantum mechanical eigenvalues.[40] Thus, in MC SCIVR, the phase-space integral of Eq. (3) is formally replaced by a sum over the most relevant trajectories, i.e., those corresponding to the spectral signals of interest. The MC-SCIVR initial conditions for the jth degree of freedom are[41,42]

$$\begin{cases} q_0^{(j)} = q_{eq}^{(j)}, \\ p_0^{(j)} = \sqrt{(2n_j+1)\omega_j}, \end{cases} \tag{7}$$

with $q_{eq}^{(j)}$ being the equilibrium position of the jth mode and ω_j being its harmonic frequency. Our reference states are combinations of coherent states of the type

$$|\chi\rangle = \prod_{j=1}^{F}\left(|p_0^{(j)}, q_0^{(j)}\rangle + \epsilon_j|-p_0^{(j)}, q_0^{(j)}\rangle\right), \tag{8}$$

where $\epsilon_j = \pm 1$ according to which one wants to enhance the spectroscopic signal. For example, a collection of +1 values allows one to enhance the zero-point energy (ZPE) signal (together with the even transitions), while a selected $\epsilon_j = -1$ and the remaining $\epsilon_{i\neq j} = +1$ enhance the odd transition of the jth mode.[43] The MC SCIVR has been successfully applied to the study of several systems,[44–55] including the different conformers of the glycine amino acid.[44] To improve with respect to the harmonic initial conditions of Eq. (7), a preliminary adiabatic switching warm up can be implemented with the result that frequency estimates are generally more accurate and complications due to deterministic chaos are largely avoided.[56]

However, for very high-dimensional systems, the overlap between the initial and the evolved wavepackets of Eq. (1) becomes smaller and smaller as the dimensionality increases, given that the reference states are the direct product of monodimensional coherent states. To overcome this limitation, the Divide-and-Conquer (DC) SCIVR method has been recently introduced.[10,11] The DC basic idea is to project the system onto lower-dimensional subspaces. Within these subspaces, it is possible to calculate reduced dimensionality spectra. Then, the full-dimensional spectrum can be obtained by convolving the subdimensional ones. The DC-SCIVR working equation is

$$\tilde{I}(E) = \left(\frac{1}{2\pi}\right)^D \iint d\tilde{\mathbf{p}}_0 d\tilde{\mathbf{q}}_0 \, \frac{1}{2\pi T} \times \left| \int_0^T dt \, \langle \tilde{\chi} | \tilde{\mathbf{p}}_t \tilde{\mathbf{q}}_t \rangle e^{i\left(\tilde{S}_t(\tilde{\mathbf{p}}_0,\tilde{\mathbf{q}}_0)+\tilde{\phi}_t(\tilde{\mathbf{p}}_0,\tilde{\mathbf{q}}_0)+Et\right)} \right|^2, \tag{9}$$

where the quantities projected onto a D-dimensional subspace have been indicated with the tilde symbol. The DC-SCIVR approach has been successfully applied to complex and fluxional systems, such as small water clusters[57] and the protonated water dimer.[58] It is also possible to implement the multiple coherent states' idea into the DC-SCIVR method by replacing the double integral of Eq. (9) with a sum running on the most relevant trajectories. This method, named MC-DC SCIVR, can deal with very high-dimensional systems. Notable applications of MC-DC SCIVR include dipeptide derivatives,[59] nucleobases[60] and nucleosides,[61] and molecules adsorbed on titania surfaces.[62] The most relevant issue to deal with for a successful DC-SCIVR calculation is the choice of the optimal subspace decomposition. In fact, all the quantities appearing in Eq. (9) can be exactly projected onto subspaces, except for the action, because the potential is not, in general, separable and its coupling terms significantly change the action. To project the action, we adopted the following equation:

$$\tilde{S}_t(\tilde{\mathbf{p}}_0, \tilde{\mathbf{q}}_0) = \int_0^t dt' \left[\frac{1}{2} \tilde{\mathbf{p}}_{t'}^T \tilde{\mathbf{p}}_{t'} - \left(V(\tilde{\mathbf{q}}_{t'}, \mathbf{q}_{t'}^{(F-D)}) - V(\tilde{\mathbf{q}}_{eq}, \mathbf{q}_{t'}^{(F-D)}) \right) \right], \tag{10}$$

which is exact for separable potentials.[10] The projected pre-exponential factor is obtained by substituting the elements of the $2M \times 2M$ sub-block Jacobian matrix of the type of Eq. (6) into the prefactor of Eq. (5).

Among the possible criteria that one can adopt to partition the vibrational space into subspaces, the one that makes each $2D \times 2D$ sub-block Jacobian matrix determinant closer to 1 is the less severe approximation for the reduced dimensionality spectra. We have called this procedure the "Jacobi method."[11] By recalling Liouville's theorem, this vibrational space subdivision is the one that minimizes the energy exchange between subspaces because an ideal partition where each sub-block determinant is equal to unity preserves the energy within each subspace. More specifically, we compute the Jacobian matrix [Eq. (6)], which, in turn, is computed during the dynamics according to the prescription given by Brewer *et al.*[63]

The main goal of this work is to find an efficient method for the subdivision of the Jacobian matrix of Eq. (6) into a number n_s of subdimensional Jacobian matrices, each containing a cluster of normal modes, such that each subspace evolution $\left(\tilde{\mathbf{p}}_t, \tilde{\mathbf{q}}_t\right)$ is the closest possible to satisfy Liouville's theorem. In other words, an optimal choice of the normal mode clustering would lead to $\prod_i^{n_s} \det \tilde{\mathbf{J}}_i$ as close as possible to 1, where $\tilde{\mathbf{J}}_i$ is the Jacobian matrix of subspace i and it is extracted from the full-dimensional Jacobian by simply selecting all the entries involving the modes that belong to subspace i. In our procedure, we compute the Jacobian matrix at every time step along a test trajectory, which starts from the equilibrium position of the atom coordinates and with an initial kinetic energy equal to the harmonic zero-point energy. Originally,[11] we presented a hierarchical search of the most frequently selected subspaces along the test trajectory in a two-step procedure. First, for each possible dimensionality $1 \ldots D < F$, all the $\binom{F}{D}$ possible subspaces were considered and the most frequently chosen along the dynamics were saved. Then, the absolute value of the deviation from one of each subspace Jacobian was computed at each time step, and the optimal subspace for each dimensionality was declared to be the one with the smallest average deviation. The subspace associated with the overall smallest deviation was selected, and the whole procedure reiterated on the remaining degrees of freedom until all modes had been included into a subspace. This approach has still a non-convenient computational scaling cost, which is proportional to $\sum_{D=1}^{F-1} D\binom{F}{D}$, since all subspace combinations need to be tested. Furthermore, it is hierarchical and thus prone to find local optima for the global subdivision. In this study, we develop and test an algorithm that is able to find the global optimal subspaces with lower computational efforts, given a constrained maximum subspace dimensionality $\tilde{D}$. This maximum dimensionality constraint can be freely chosen at the beginning of our proposed procedure. It is useful because, as anticipated, it comes from the necessity to perform semiclassical calculations below a certain dimensionality to get sensible results.

B. Probability graph-evolutionary algorithms (PG-EA)

Here, we introduce a combined Probability Graph and Evolutionary Algorithm (PG-EA) approach to find the best vibrational space subdivision according to Liouville's criterion explained above. Evolutionary algorithms emulate the natural selection of an initial population, where the "fittest" individual is the most likely to survive and its genes to be inherited by the next generation. In GAs' jargon, the population is composed of chromosomes, which are collections of fitness parameters, each one called a gene. There is no obvious or required way to represent the genes, and there are many valid choices. At each epoch, all chromosomes are evaluated and sorted according to their fitness score. First, a fraction of the best individ-

uals gives birth to a set of newborn chromosomes by mixing and mutating genes during the crossover and mutation processes. Then, the new chromosomes take the place of those individuals that are least fit to survive so that the next epoch would be enriched by the more fitted chromosomes.

In our case, each chromosome represents a possible clustering of vibrational degrees of freedom into subspaces to compose the full vibrational space. The collection of all the chromosomes provides many possible subdivisions of the full-dimensional vibrational space. Each chromosome is evaluated by an appropriate score function that rates the individual's fitness. The fitness function is evaluated after the time evolution of the Jacobian matrix along a test trajectory, which, in our case, is the trajectory that evolves from the equilibrium geometry with the energy of the vibrational ground state. Given the consideration at the end of Sec. II A, we propose the following fitness function for a possible collection C of normal modes subdivisions:

$$f(C) = \frac{1}{N}\sum_{\text{steps}}^{N}\sum_{s\in C}\left|1 - \left|\det(\tilde{\mathbf{J}}_s)\right|\right|, \tag{11}$$

where the external sum is over the N time steps. We prefer Eq. (11) with respect to a possible fitness function, such as $|1 - \prod_i^{n_s} \det \tilde{\mathbf{J}}_i|$ as employed in Ref. 11, because in the latter case, there could be a compensation of error that the internal sum in Eq. (11) avoids. More specifically, the optimal criterion in Eq. (11) is satisfied when all the subspaces have the determinant of the Jacobian closest to +1 in the modulus, at every trajectory step. Furthermore, this function rewards preferentially a chromosome made of few large subspaces over one made of several small subspaces because any new term in the internal sum over C is addictive and positive. We prefer to have large subspaces to account for as many normal mode couplings as possible.

Once an initial guess of possible chromosomes is given, we need a probability distribution function to generate the new chromosomes, i.e., the new mode subdivision into subspaces. We propose our own customized evolutionary algorithm inspired to GAs for updating the probability distribution $\mathbf{\Phi}(\tau)$ from which newborn chromosomes are sampled at a certain epoch τ.

First, we represent a chromosome as the adjacency matrix of an unweighted cluster graph, which is a graph where each connected component is a clique (i.e., a fully connected subgraph), as reported in Fig. 1. The vertices are the normal modes that are connected only if they fall into the same subspace. This representation minimizes the redundancy of information since cliques are invariant to vertex permutation and a cluster graph is invariant to clique permutations.

All the information about the subspaces is codified in the adjacency matrix $\mathbf{C}$. It is defined by $C_{ij} = 1$ only if modes i and j are in the same subspace; otherwise, $C_{ij} = 0$. $C_{ii} = 1$ only if mode i is in a one-dimensional subspace. In the end, we have increased the problem variables from the F-dimensional redundant representation [such as in the linear representation (1, 2, 3, 4, 5)(6, 7, 8)(9) = (2, 4, 1, 5, 3)(8, 7, 6)(9) = ...] to the $F(F + 1)/2$-dimensional non-redundant one. The F-dimensional representation is redundant because any mode permutation within a given subspace leaves the subspace unaffected. Our adjacency matrix representation is invariant to row and column permutation within the subspace block. Hence, every subspace configuration has a unique adjacency matrix representation. Furthermore, our adjacency matrix is symmetric and thus completely defined by its F(F + 1)/2 lower (upper) triangular elements.

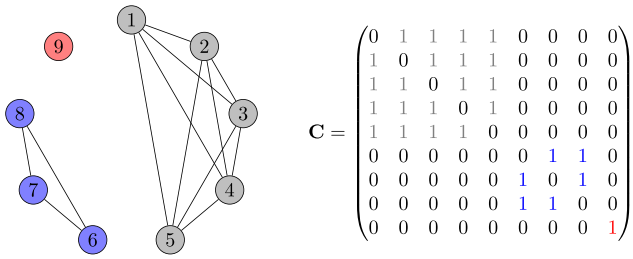

FIG. 1. Chromosome expressed in a cluster graph representation. Here, nine normal modes, corresponding to the graph vertices, are grouped into three subspaces (gray, blue, and red), containing 5, 3, and 1 modes, respectively. Each subspace is a fully connected graph (a clique), and there is no interaction between any two subspaces. On the right, the corresponding adjacency matrix is colored accordingly.

Second, we customize the crossover and mutation operators to mix the chromosomes with simple arithmetic rules and store the genetic information in a matrix of weights $\mathbf{\Phi}(\tau)$, which represents the probability distribution of the mixed chromosomes at the evolution epoch τ. We define the crossover $\mathcal{X}$ of a couple of chromosomes C and C' as a weighted average of their adjacency matrices,

$$\mathcal{X} = \frac{1}{w_C + w_{C'}}\left(w_C\mathbf{C} + w_{C'}\mathbf{C}'\right), \tag{12}$$

where $\mathbf{C}$ and $\mathbf{C}'$ are adjacency matrices and the weights w_C and $w_{C'}$ depend on the chromosome fitnesses. The pure mutation $\mathcal{M}$ of a chromosome C with probability μ is

$$\mathcal{M} = \frac{1}{1 + F\mu}(\mathbf{1}\mu + \mathbf{C}), \tag{13}$$

where $\mathbf{1}$ is the square matrix of ones and F is the number of genes. In our approach, each vibrational normal mode corresponds to a gene, as illustrated in the previous example of Fig. 1. Both crossover and mutation have the basic property of scattering the gene probability. More specifically, the crossover distributes the probability among the genes expressed in C and C', depending on their fitness, and the mutation distributes it among every possible outcome, independently of the fitness and the expression. The combination of the crossover and mutation processes is obtained by the subsequent application of the two operations: the mutation can be equivalently applied to the single chromosomes prior to undergoing crossover or to the result of the crossover process. In addition, considering that we use m chromosomes for the optimization, only an elite fraction $\eta < 1$ of these are the fittest chromosomes that undergo the crossover and mutation processes. Eventually, the resulting probability distribution at epoch τ, $\mathbf{\Phi}(\tau)$, is

$$\mathbf{\Phi}(\tau) = \frac{1}{1 + F\mu}\left[\mathbf{1}\mu + \frac{1}{\sum_i^{m\eta} w_i}\sum_{i=1}^{m\eta} w_i\mathbf{C}_i\right], \tag{14}$$

where $\mathbf{C}_i$ is the ith chromosome adjacency matrix and the new probability distribution is essentially generated by suitably mixing the

adjacency matrices of the previous generation. This is the machine learning part, where the algorithm, epoch by epoch, learns the optimal probability distribution from an evolving population of chromosomes. For this work, we use the simple weighting scheme $w_i = (m\eta - i)/m\eta$ with the resulting normalization constant $\sum_i^{m\eta} w_i = (m\eta - 1)/2$, and the elite fraction is $\eta = 0.4$, while the mutation probability μ and the number of chromosomes will be specified below case by case. $\mathbf{\Phi}(\tau)$ is updated at every epoch and contains the average genetic material of the previous generation of chromosomes according to Eq. (14).

To sample new chromosomes from the probability distribution, $\mathbf{\Phi}(\boldsymbol{\tau})$ must be normalized. This means that $\mathbf{\Phi}(\boldsymbol{t})$ of Eq. (14) has to be symmetric and doubly stochastic, i.e., with rows and columns summing up to 1, so that we can consider $\mathbf{\Phi}(\boldsymbol{\tau})$ a weighted undirected graph to sample from. To enforce the doubly stochastic property and, at the same time, retain the symmetry, we rely on Sinkhorn's theorem,[64] which ensures that there exist two diagonal matrices $\mathbf{R}$ and $\mathbf{S}$ such that $\mathbf{R\Phi}(\tau)\mathbf{S}$ is doubly stochastic. R_{ii} and S_{jj} are found by repeatedly and alternatively normalizing the rows and the columns of $\mathbf{\Phi}(\tau)$, according to the updates

$$\begin{aligned} R_{ii} &= \frac{1}{\sum_j \Phi_{ij}(\tau) S_{jj}} \quad \forall\, i, \\ S_{jj} &= \frac{1}{\sum_i \Phi_{ij}(\tau) R_{ii}} \quad \forall\, j, \end{aligned} \tag{15}$$

with S_{jj} initialized to 1 for all j. $\mathbf{R}$ and $\mathbf{S}$ will converge, up to a small threshold ε, after an unspecified number of iterations.[64] In all the applications described below, we use $\varepsilon = 10^{-8}$ on each element of $\mathbf{R}$ and $\mathbf{S}$, which is always satisfied in less than 100 iterations.

Third, we need to elaborate a procedure for obtaining the newborn chromosomes from the symmetric and doubly stochastic probability matrix $\mathbf{\Phi}(\tau)$. To sample representative cluster graphs from $\mathbf{\Phi}(\tau)$, we propose a sampling procedure to generate a population, which reflects the original distribution:

1. generate the random numbers r_i, $i = 1, 2, \ldots, F$ and sample independently the chances of each mode to be in a subspace alone. If $r_i < \Phi_{ii}(\tau)$, then the normal mode i is in a subspace alone;
2. iterate on the leftover modes in a random order: if the kth mode is already joined with another mode, then continue with the next one; otherwise, sample the edge between modes k and $j \neq k$ with a random number and join them with a probability given by the matrix element $\Phi_{jk}(\tau)$. If the kth mode cannot be joined to any j (for instance, because each of them is in a one-dimensional subspace), it stays in a subspace by itself;
3. identify the connected components of the sampled graph and complete them, obtaining the cluster graph for a newborn chromosome.

Note that before step 3, the procedure samples tree graphs, which means that there are no redundant sampling steps. Each chromosome sampled with this procedure may be weakly biased anyway, and the random shuffle of the mode order in step 2 is required to make the sampled population representative and the overall sampling unbiased. To achieve step 3, we look for a basis of the Laplacian matrix Kernel, with the Laplacian matrix defined as $L_{ij} = -C_{ij} + \delta_{ij} \sum_j C_{ij}$. Since $\sum_j L_{ij} = 0$ by definition, the vector of ones always belongs to $Ker(\mathbf{L})$, i.e., to the collection of vectors $\mathbf{x}$, such that $\mathbf{Lx} = \mathbf{0}$. Furthermore, if the graph is disconnected, $\mathbf{L}$ can be rearranged to be the block diagonal by swapping row and column indices, with each block being the Laplacian of the corresponding connected subgraph. Hence, each basis vector of $Ker(\mathbf{L})$ is 1 on the entries of the connected vertices and 0 elsewhere. The sum of all basis vectors is the vector of ones. To practically find a basis for $Ker(\mathbf{L})$, we solve the linear equation $\mathbf{Lx} = \mathbf{0}$ by applying Gaussian elimination[65] to the augmented matrix $\mathbf{L}|\mathbf{I}$, with output $\mathbf{L_{rref}}|\mathbf{B}$, where $\mathbf{L_{rref}}$ is the reduced row echelon form of $\mathbf{L}$. The rows $\mathbf{x}$ in $\mathbf{B}$ corresponding to the row indices where $\mathbf{L_{rref}} = \mathbf{0}$ do solve the linear equation and hence form a basis for the kernel.

To measure the likelihood of a subspace s of size D sampled from $\mathbf{\Phi}(\tau)$, we sum the edge products of all the possible trees that span the clique (subspace), as

$$p(s,\tau) = \frac{(D-1)^{D-1}}{D^{D-2}} \sum_{\mathbf{T} \in span(\tilde{\mathbf{\Phi}}_s(\tau))} \prod_{e=1}^{D} T_e, \tag{16}$$

where T_e is the edge of the tree graph $\mathbf{T}$, which spans the subspace probability distribution $\tilde{\mathbf{\Phi}}_s(\tau)$ (that is, the probability distribution considering only the modes in s). The first factor is a normalization constant so that $p(s, \tau)$ does not depend on the subspace size, and it is maximized to 1 when $\tilde{\mathbf{\Phi}}_s(\tau)$ is uniform. D^{D-2} is the number of spanning trees for a clique according to Cayley's formula.[66] $p(s, \tau)$ measures the degree of convergence toward the chosen subspace s such that, as $p(s, \tau)$ approaches unity, the population becomes more and more uniform and eventually the algorithm stops learning. The brute force application of Eq. (16) is out of reach for large subspaces ($D > \approx 10$); therefore, we use it only to check the algorithm progression toward an optimal solution of the small systems described below. Furthermore, GAs, in general, and PG-EA, in particular, do not require a full convergence of the population for the solution to be satisfactory. On the contrary, if the solution is unknown or hard to find, a homogeneous population is undesirable as it kills diversity and damps the optimization.

In Fig. (2), we report a four normal mode example to show how the PG-EA algorithm works in practice. First, as we do in all our simulations, the initial probability distribution is set as $\Phi_{ij}(0) = 1/F \;\forall\, i, j$. Then, we generate the initial chromosome population according to the sampling procedure described at points 1–3 above. Each chromosome graph is reported together with the corresponding adjacency matrix in panel 1 in the left side of Fig. (2). The chromosomes C_j are ordered by the fitness score, which is calculated using the function $f(C)$ [pay attention that according to our definition of the fitness function, Eq. (11), the preferred chromosomes are those with a lower fitness score]. The chromosome fraction $m\eta$ undergoes crossings and mutations. We reject (i.e., apply an infinite penalty) to any subspace that has a dimension larger than the largest subspace value $\bar{D}$ that one fixes *a priori*. Since assigning the fitness score is the most expensive step, it is advisable to build a score database, i.e., a list of already known chromosomes with their fitness value, so that whenever a known chromosome is encountered, its score does not have to be recomputed. After undertaking crossovers and mutations according to Eq. (14), a new (non-normalized) probability distribution is generated (panel 2 on the right side of Fig. 2). After transforming the new probability distribution into a symmetric and

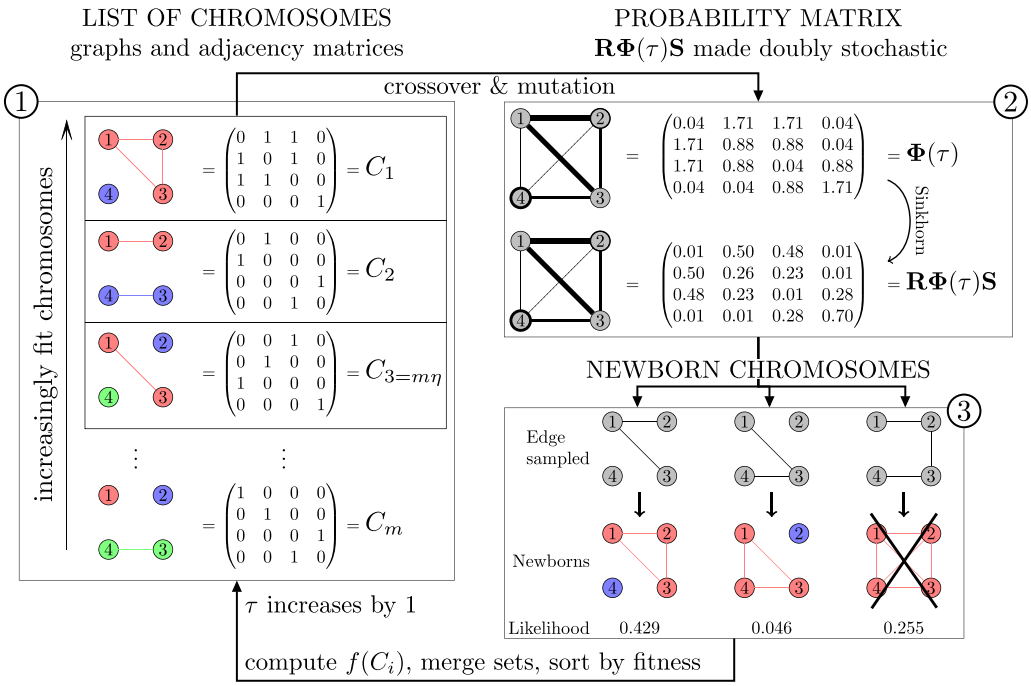

FIG. 2. A numerical example of the PG-EA for a four dimensional system. The procedure is broken down into three steps. The inner rectangle in panel 1 includes the $m\eta$ elite chromosomes, which will be employed for mutation and crossover in panel 2.

doubly stochastic distribution matrix using Sinkhorn's algorithm, we generate the newborn chromosomes according to the sampling procedure of points 1–3 described above. A likelihood coefficient is calculated according to Eq. (16). In the lower right part of Fig. 2, the four-dimensional solution is rejected because its dimensionality is greater than the largest subspace value $\bar{D}$, which, in this numerical example, has been fixed to be $\bar{D} = 3$. Finally, a fitness coefficient is attributed to each chromosome, and we are back to panel 1 for the next iteration. At each epoch, the fittest ith chromosome, i.e., the one with the lowest $f(C_i)$ value, provides the graph with the so far optimal normal mode arrangement into subspaces.

C. Two-mode interaction method

As an alternative, we propose an approximate and computationally cheaper method to deal with large molecules when the computational cost of PG-GA is prohibitive or in instances in which one can reasonably assume that for each normal mode, the coupling is mainly due to the interaction with just a second mode.

In this alternative approach, we first compute a two-mode coupling network, where each vertex i, j is a normal mode and each edge is weighted by the two-mode Jacobian determinant $G_{ij} = \det(\tilde{\mathbf{J}}_{ij})$. $\tilde{\mathbf{J}}_{ij}$ is a 4×4 matrix containing all the partial derivatives between the phase-space momenta and positions of modes i and j with respect to the initial conditions. For each full-dimensional Jacobian matrix $\mathbf{J}$ along the trajectory, we evaluate the determinant of every two-mode combination, G_{ij}. Then, we compute the distance matrix $\mathbf{E} = |\mathbf{G} - \mathbf{1}|$, which measures how large is the error done by assuming that the relevant interaction is only between the couple (i, j) of modes, while other interactions are disregarded. Specifically, when $E_{ij} = 0$, then modes i and j are fully correlated and uncoupled to any other mode. $\mathbf{E}$ is computed at every time step of the test trajectory, and all $\mathbf{E}$ matrices are averaged into a single distance matrix representative of the whole trajectory.

Then, we employ an agglomerative hierarchical clustering technique called Weighted Pair Group Method with Arithmetic Mean (WPGMA)[67] to cluster the normal modes using the information encoded in $\mathbf{E}$. The algorithm produces a dendrogram where each branching is an optimal subspace. Among the several hierarchical clustering techniques available in the literature, we choose WPGMA because it provides results that are the closest to the exact ones for the model systems considered below. WPGMA clustering is iterative and hierarchical. To start, each mode is in a subspace by itself. Then, at every iteration, the two "closest" subspaces are merged into one, and the dendrogram profile shows a link. The distance between the newly formed subspace $(j \cup k)$ and a given subspace i is calculated as the arithmetic mean of the distances from the newly merged subspaces j and k,

$$E'_{i(j\cup k)} = \frac{E_{ij} + E_{ik}}{2}. \tag{17}$$

The procedure goes on until a maximum distance criterion has been met, i.e., until all modes fall into one large subspace.

The whole process is represented by using a dendrogram, where each node corresponds to a subspace and each edge represents the link between two subspaces. At the root of the dendrogram, there is the full-dimensional system, which contains all modes. The leaves are the subspaces containing one mode only. The distance E_{ij} of

every update is a measure of how close the linked subspaces are. Finally, this process generates a number of arrangements of normal modes at different levels of the tree, and for every such arrangement, we measure the fitness score, i.e., the full-dimensional Jacobian factorization error with Eq. (11), along with E_{ij}.

III. RESULTS

This section presents our results, and it can be divided into three parts. In Subsection III A, we show how PG-EA and the two-mode interaction method are effective when applied to model systems such as coupled Morse oscillators with non-trivial coupling topologies but with obvious mode separations. In Subsection III B, we show that we can improve spectral accuracy with respect to previous calculations where the hierarchical subspace optimization originally proposed was adopted.[11] Finally, in Subsection III C, we show that PG-EA allows us to apply the DC-SCIVR method and select the subspaces with the Jacobi criterion even for the simulation of mid-large molecules such as the 12-atom *trans*-*N*-Methylacetamide (NMA). Remarkably, it would not have been possible to accomplish this task with a brute force combinatorial approach.

A. Model systems

To preliminarily test our algorithms, we consider the arrangements of F coupled Morse oscillators **A** and **B** in Fig. 3. Each oscillator experiences the following Morse-type potential:

$$V^{morse} = \sum_{i=1}^{F} D_e\left(1 - e^{-\omega_i (2D_e)^{-1/2}(q_i - q_{eq,i})}\right)^2 + \sum_{i=1}^{F-1}\sum_{j>i}^{F} \lambda_{ij}(q_i - q_{eq,i})(q_j - q_{eq,j}), \quad (18)$$

where the dissociation energy $D_e = 38\,293$ cm^{-1} and the equilibrium position $q_{eq} = 1.4$ a.u. are valid for all F degrees of freedom. According to the coupling graphs and matrices $\mathbf{\Lambda}$ schematically represented in Fig. 3, the oscillators might be uncoupled (no edge), weakly coupled ($\lambda = 10^{-7}$ a.u., dashed edge), or strongly coupled ($\lambda = 10^{-5}$ a.u., solid edge). We devise two topologies (**A** and **B**) to provide non-trivial examples. **A** has an oscillator frequency of $\omega = 3000$ cm^{-1} for oscillators 1 through 6 and $\omega = 4000$ cm^{-1} for oscillators 7 and 8; **B** has all oscillators with the same frequency of $\omega = 3000$ cm^{-1}. In both cases, the correct separation into subspaces is unique.

Both PG-EA and the two-mode interaction method separate system **A** correctly. In PG-EA, we use $m = 50$ chromosomes, a mutation probability of $\mu = 0.001$, a crossover fraction of $\eta = 0.4$, and 70 epochs, with the constraint that the maximum dimension is $\bar{D} = 6$. The likelihood of the optimal subspaces calculated using Eq. (16) is plotted against the epochs in panel (a) at the top left of Fig. 4, showing that the population quickly converges to the unique global optimum represented by the continuous lines. The two-mode interaction method provides three choices for the subspaces, the best of which is the global optimum (1, 2, 3, 4, 5, 6)(7, 8), reported in the first branching of the dendrogram in panel (b) at the top right of Fig. 4. This global optimum has a Jacobian factorization error of about $3.37 \cdot 10^{-6}$. The example of topology **B** is much more challenging as it is an 18-dimensional system divided into four loosely connected regions, one of which containing a single mode. As expected, it turns out that the best subspace division has oscillator 10 (black subspace in Fig. 3) joined together with five oscillators (7–12, green subspace) since there is one term less in the fitness function summation with respect to the case in which oscillator 10 is left isolated. In panel (c), at the bottom left of Fig. 4, PG-EA provides the optimal desired solution using $m = 300$ chromosomes, $\mu = 0.1$, $\eta = 0.4$, and 1000 epochs, with the constraint that the largest subspace dimension is $\bar{D} = 10$. In panel (d), at the bottom right of Fig. 4, the two-mode interaction method also provides the optimal solution, as shown in the upper part of the dendrogram with the smallest Jacobian factorization error.

Now, one may wonder if these subdivisions are, indeed, the most suitable ones for DC-SCIVR spectroscopic calculations. In Fig. 5, we show that DC SCIVR can account properly for most of the spectral features of these systems if the subspaces are chosen according to the algorithms described above. For example, when choosing the subspace separation (1, 3, 5, 8)(2, 4, 6, 7), which is the least fit for case **A,** i.e., it has the largest fitness score in the case of requiring two subspaces only, the corresponding spectra are quite noisy,

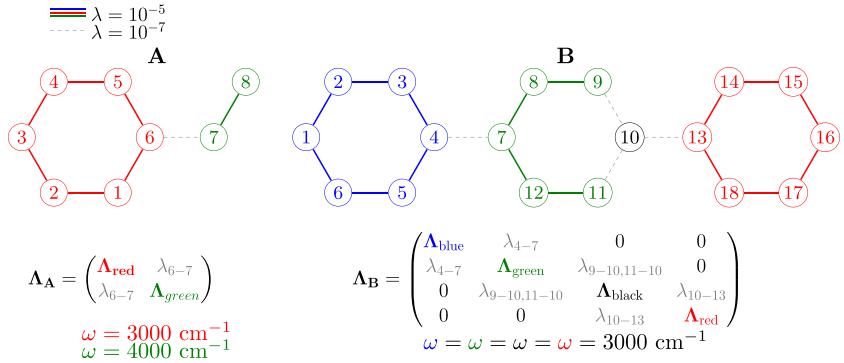

FIG. 3. Toy model systems with different coupling topologies. The circled numbers represent Morse oscillators, and the edges represent their couplings. When two oscillators are not connected by an edge, the relative coupling matrix element λ_{ij} is zero. The oscillator frequencies ω, reported below each coupling matrix, are 3000 cm^{-1} except for modes 7 and 8 of topology **A**, for which $\omega = 4000$ cm^{-1}. Note that for topology **A** reported on the left, oscillators 1 and 5 are equivalent and they produce the same signal, and 2 and 4. Similar symmetry considerations can be applied to topology **B**.

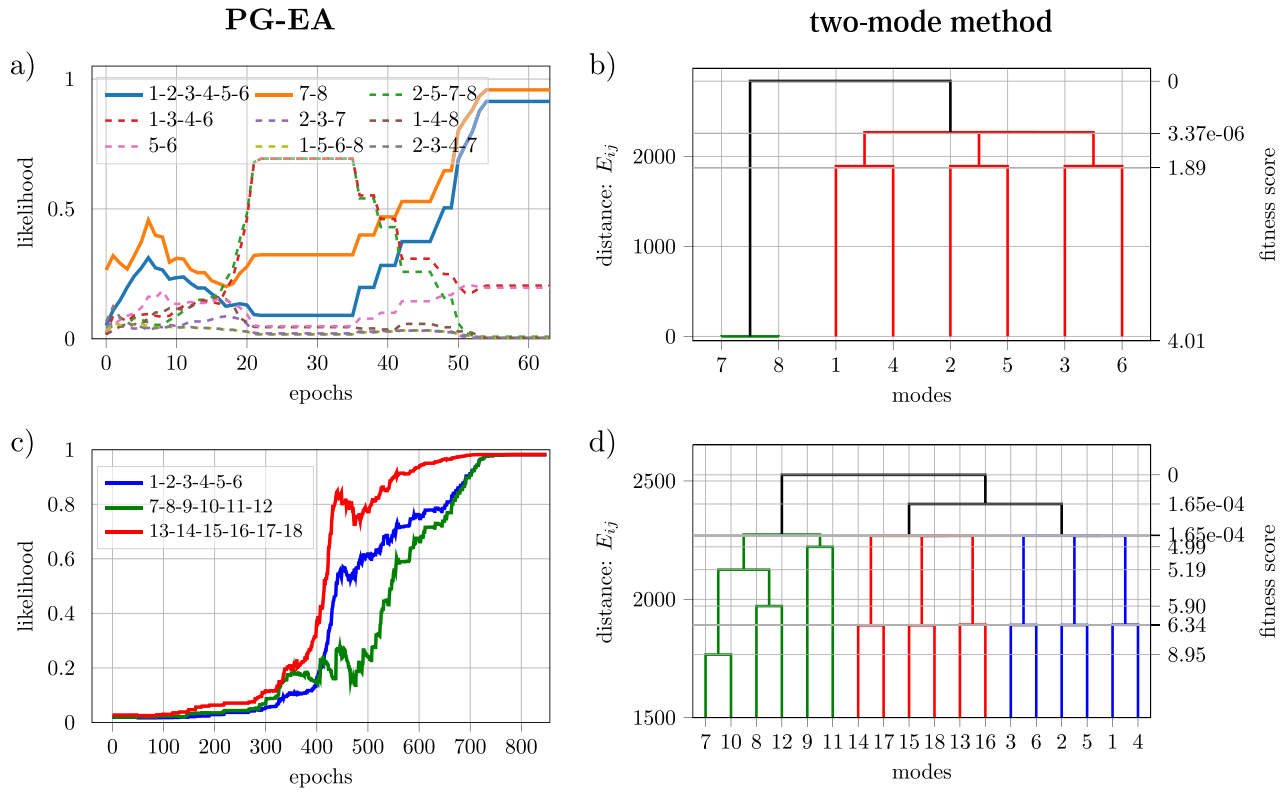

FIG. 4. Vibrational modes' subdivision optimizations of the model systems in Fig. 3 using PG-EA [left panels (a) and (c)] and the two-mode interaction method [right panels (b) and (d)]. In the top panels [(a) and (b)], the subspace optimizations of topology **A** are effectively achieved using both methods. The dendrogram (b) is colored to highlight the least error branching. In the bottom panels [(c) and (d)], the subspace optimization is effectively reached by both methods for topology **B**.

as shown in Fig. 5. Furthermore, phantom signals are observed, for example, at 2686 cm^{-1}. Conversely, the spectra of the subspaces suggested by both our algorithms, which are reported with green and red lines, are without noise to the naked eye. However, we note that the signal originated from a combination band of modes from different subspaces at 7006 cm^{-1} is too weak at this scale to be observed. This is not a drawback of the algorithms proposed in this work but is a known feature of the DC-SCIVR method in predicting mixed overtones originated from modes belonging to different subspaces.

On the right panel of Fig. 5, we show the optimal subspace spectra for system **B**. In this case, the system is too large to have a well-converged semiclassical full-dimensional TA-SCIVR spectrum as shown by the black continuous line spectrum.[43] Instead, it is

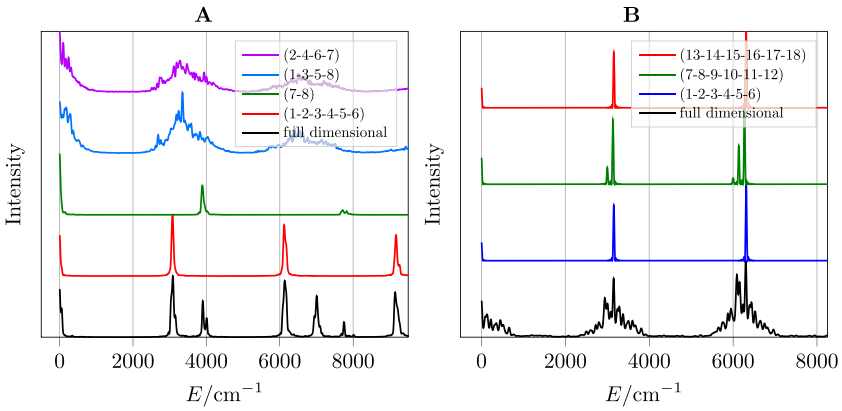

FIG. 5. Spectra of the coupling topologies **A** and **B** of Morse oscillators with ZPE signals shifted to 0 cm^{-1}. Left panel for system **A**: black represents the full-dimensional spectrum; red and green represent the best subspaces (ring and segment, respectively, with reference to Fig. 3), while blue and purple represent the two worst subspaces. Right panel for system **B**: blue, dark green, and red represent the best subspaces, and black represents the full-dimensional calculation.

possible to recover the most significant spectroscopic features of the system with a DC-SCIVR calculation based on the optimal subspaces suggested by the algorithms.

B. The CH_4 molecule

This section further confirms the ability of the proposed algorithms to find optimal subspace separations for DC-SCIVR calculations when applied to real systems. We show that our techniques can reproduce and improve the DC-SCIVR spectra even for small molecules. We consider CH_4 as the case system. The methane vibrational spectrum is a tough challenge for DC SCIVR because the molecule is characterized by highly chaotic dynamics and high symmetry, which is difficult to recover if a proper subspace partition is not implemented. We simulate 180 000 trajectories for 30 000 a.u. long, and each trajectory is rejected during the dynamics $||\det(\mathbf{J}^T\mathbf{J})| - 1| > 10^{-5}$. The initial trajectory conditions are sampled from the Husimi distribution centered in phase space at $(\sqrt{\boldsymbol{\omega}}, \mathbf{q}_{eq})$, while gradients and Hessian matrices are computed by finite differences with infinitesimal displacements equal to 10^{-3} a.u. for all modes.

We use the force field by Lee, Martin, and Taylor,[68] which takes into account the symmetry relations of cubic and quartic force constants.[69,70] The same potential energy surface (PES) and the hierarchical subspace optimization with the Jacobi method were employed in a previous work of the group.[10] For this system, PG-EA successfully converges with the constraint $\tilde{D} \leq 7$, which leads to the optimal couple of subspaces (2, 5)(1, 3, 4, 6, 7, 8, 9), with a fitness score of about 0.71. The three subspaces (1)(2, 3)(4, 5, 6, 7, 8, 9) were selected in a previous work of the group[10] using the Jacobi criterion but looking for optimal subspaces with a brute force hierarchical approach and constraining the largest subspace to be six dimensional. These three subspaces have an associated fitness score of about 0.91. For methane, PG-EA converges using $m = 100$ chromosomes, $\eta = 0.4$, $\mu = 0.01$, and 50 epochs. The likelihood plot is represented in panel (a) in the left part of Fig. 6.

As shown in Fig. 6 [panel (b)], the two-mode approximation leads, in this case, to a poor subspace separation: the best branch in the dendrogram is the colored (1, 3, 4)(2, 5, 6)(7)(8)(9) subspace with a score of 4.84, leading to a bigger error than PG-EA for the Jacobian factorization.

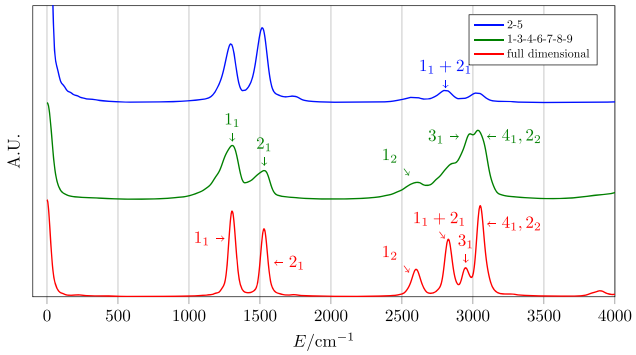

FIG. 7. Spectrum of methane: the full-dimensional spectrum is in red, and the reduced dimensionality spectra chosen according to PG-EA are in green and blue.

Methane, in the absence of a preliminary adiabatic switching sampling,[56] is known to be characterized by highly chaotic dynamics;[68] thus, we employed 180 000 trajectories to provide convergent results, with a rejection rate of about 90%, keeping nearly 2000 trajectories per degree of freedom. Figure 7 reports both the full-dimensional spectrum (red continuous line) and the partial dimensional ones (green and blue), according to the PG-EA vibrational space sub-division found above. All spectroscopic features are properly reproduced. Note that degenerate modes belonging to different subspaces give rise to spectral lines at the same energy (see, for instance, 1_1 and 2_1 signals displayed in both subspace spectra of Fig. 7). In these cases, we consider more accurately the peaks appearing in the largest subspace as more mode interactions are taken into account, even if frequencies of degenerate modes in different subspaces are very similar and cannot be distinguished by the naked eye. In Fig. 7, we employ an incremental notation for the spectral features so that degenerate signals are collected together under the same label. For a deeper insight, we report in Table I the value in wavenumbers of each spectral peak frequency.

In conclusion, PG-EA provides a subdivision of the vibrational space appropriate for DC-SCIVR spectroscopic calculations, and we can move to apply it to larger systems where previous recipes for the vibrational space subdivision are impractical.

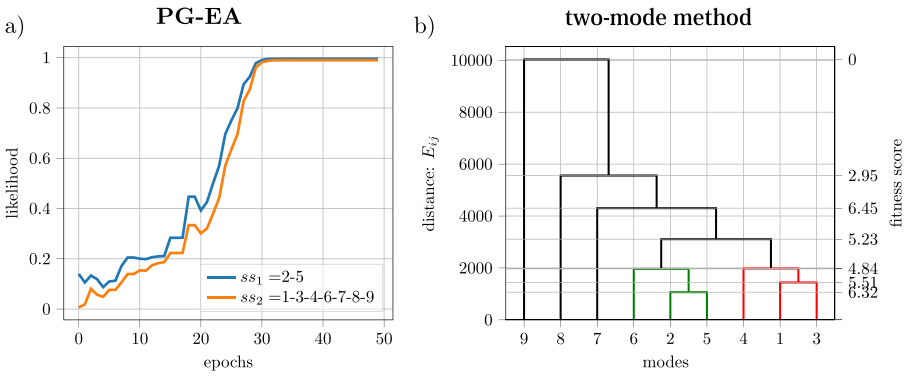

FIG. 6. Likelihood of the optimal subspaces of methane during the PG-EA optimization (a) and dendrogram for separation with the two-mode approximation method (b). Note that the two-mode approximation leads to a very different result.

TABLE I. Quantum frequencies of vibration of the methane molecule in cm^{-1} calculated on the PES by Lee, Martin, and Taylor[68] using the full-dimensional TA SCIVR, DC-SCIVR based on PG-EA subspace partition, and discrete variable representation calculations (Exact). MAE stands for mean absolute error, calculated using exact [MAE (Exact)] or full-dimensional semiclassical values [MAE (TA SCIVR)] as reference. In the fourth column, we report the DC-SCIVR frequencies obtained from the subdivision proposed in Ref. 11 where a different approach for the Jacobi method was employed.

Incremental label	Modes (symmetry)	Exact[71]	TA SCIVR	DC SCIVR (1)(2, 3)(4–9)[11]	DC SCIVR PG-EA (2, 5) (1, 3, 4, 5–9) [sub][a]
1_1	1, 2, 3 (F_2)	1313	1304	1287	1305 (G)
2_1	4, 5 (E)	1535	1529	1534	1530 (G)
1_2	1, 2, 3	2624	2600	2562	2610 (G)
$1_1 + 2_1$	1, 2, 3, 4, 5	2836	2827		2807 (B)
3_1	6 (A_1)	2949	2948	2960	2980 (G)
4_1	7, 8, 9 (F_2)	3053	3051	3044	3036 (G)
2_2	4, 5	3067	3051	3044	3036 (G)
MAE (exact)			9.6	22.0	19.3
MAE (TA SCIVR)				14.3	13.4

[a]Subspace from which the wavenumber is taken: G for the 7D green one and B for 2D blue one with reference to Fig. 7.

C. *Trans*-*N*-methylacetamide

Here, we present the subspace optimization and the associated DC-SCIVR spectroscopic calculations for the 30-mode *trans*-*N*-Methylacetamide (NMA) molecule represented in Fig. 8. NMA has been studied thoroughly both computationally[73–75] and experimentally[73,75–78] as it is one of the simplest examples of a molecule featuring the HNCO peptide bond. We use the full-PES by Qu and Bowman,[72] which has been designed for both *cis*- and *trans*-NMA and accounts for the three-fold symmetry of the methyl rotors. The PES is permutationally invariant and was fitted to thousands of *ab initio* calculated energies and gradients at the B3LYP/cc-pVDZ level of theory.[72] In this case, we can compare our DC-SCIVR spectroscopic results with harmonic frequencies and gas phase IR and Raman experimental values.[76]

The two methyl rotational frequencies, i.e., the two lowest-frequency normal mode values, are not considered to be part of the vibrational space. Thus, the dimensionality of the vibrational space we consider is 28. We use a simulation time step of 5 a.u. The finite difference displacement of the ith normal mode for Hessian and gradients is rescaled by $10^{-3}\sqrt{\max(\boldsymbol{\omega})/\boldsymbol{\omega}_i}$,[62] to account for different PES curvatures along each one of the normal mode coordinates. We use the signal obtained from a single 30 000 a.u. long trajectory with the initial conditions $(\sqrt{\boldsymbol{\omega}}, \mathbf{q}_{eq})$. We do not observe any conformational change from the *trans* to the *cis* potential energy basin during our simulations.

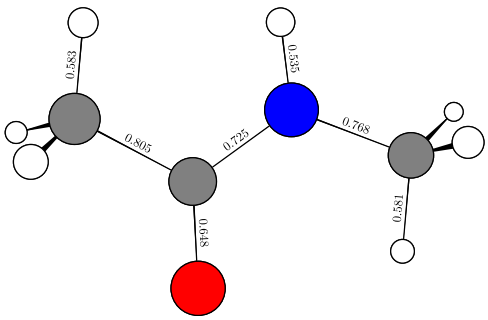

FIG. 8. *Trans*-*N*-methylacetamide equilibrium geometry on the potential energy surface by Qu and Bowman.[72] Representative bond lengths of the equilibrium geometry are shown in Åström. The HNCO peptide bond is of fundamental importance for the dynamics of peptides.

We apply PG-EA, the two-mode interaction method, and the Hessian method[11] to divide the vibrational space into subspaces. The results are quite different. For PG-EA, we run a thorough optimization using 10 000 epochs, $m = 300$ chromosomes, $\mu = 0.01$, and $\eta = 0.4$, with the largest subspace constraint set to $\tilde{D} = 15$, and we obtain the following three subspaces: $\mathbf{A} = (3, 5, 7, 10, 11, 12, 15, 21, 23, 26, 28, 30)$, $\mathbf{B} = (4, 9, 14, 16, 17, 18, 19, 22, 24, 25, 27, 29)$, and $\mathbf{C} = (6, 8, 13, 20)$. The fitness score of the chromosome, i.e., the score of the Jacobian factorization, is 1.96. When applying the basic average Hessian criterion[11] with a coarse-graining parameter equal to $\mathbf{8 \cdot 10^{-6}}$, we obtain the following subspaces: $a = (3, 5, 7, 22)$, $b = (10, 11, 13, 16–18, 20, 24–29)$, $c = (4)$, $d = (6)$, $e = (8)$, $f = (9)$, $g = (12)$, $h = (14)$, $i = (15)$, $j = (19)$, $k = (21)$, $l = (23)$, and $m = (30)$, which can be associated with a fitness score equal to 4.49. Finally, we apply the two-mode dendogram approach and obtain the following subspaces: $\alpha = (3, 4, 6–8, 10–17, 20–24, 29)$, $\beta = (5, 9, 22, 30)$, $\gamma = (18, 19, 26)$, and $\delta = (25, 27, 28)$ with a fitness score of 2.18. The two-mode interaction method produced the dendrogram reported in Fig. 9, which has a slightly worse score than the PG-EA result. However, the presence of an 18-dimensional subspace makes this subdivision not convenient for Monte Carlo phase-space integration convergence. The Hessian method is, instead, clearly penalized by many 1D subspaces found. These considerations suggest that the PG-EA subdivision into three subspaces is, indeed, the best choice. Based on PG-EA, we calculate the spectra reported in Fig. 10, where the different spectral regions

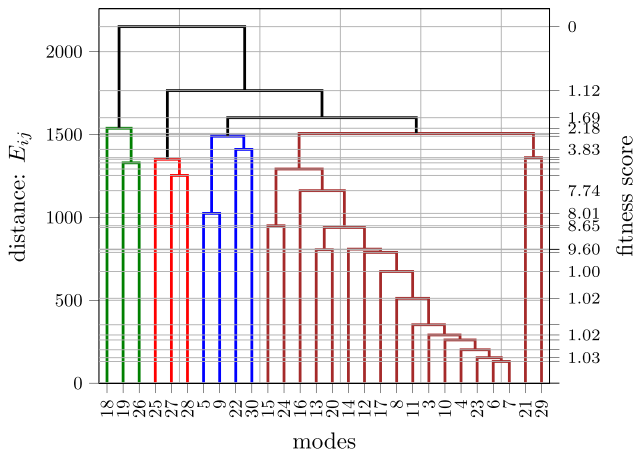

FIG. 9. Two-mode interaction dendrogram of *trans*-NMA. Four subspaces are generated with a Jacobian factorization score equal to 2.18. As an alternative, one might prefer the five-subspace option with an error of 3.83. We do not show the likelihood vs epochs plot since the convergence of the whole population is not achieved nor desired. Furthermore, computing the likelihood in Eq. (16) for a 12 dimensional subspaces is not feasible as it requires the computation of 12^{10} products.

are highlighted by different colors according to the experimentalists' denomination. Overall, the DC-SCIVR spectra are reproducing well all the spectroscopic features of this molecule. Actually, there are more spectral features in the DC-SCIVR simulation than there are in the experiment because DC SCIVR calculates the power spectrum, which is made of all vibrational levels (that we scale with respect to the zero-point energy), even those associated with transitions that are not IR active. An IR spectrum simulation with related intensities would require calculation of the vibrational eigenfunctions. This feature is not implemented yet in our divide-and-conquer approach, but we are planning to do it soon. In addition, according to our simulations and referring to the experimentalists' denomination, mode 23 is the only one responsible for the amide I band, and subspace A contributes the most to amide III and A bands, while amide II is mostly localized in subspace B.

All these subspace choices produce spectra with almost the same MAEs, as reported in Table II. This suggests that the subspace choice is flexible, as well as the choice of the subdivision criterion. However, PG-EA is the method that minimizes the number of subspaces and prevents from having many 1D subspaces, which could result in a very noisy and not resolved spectrum.[62]

A closer inspection of the vibrational frequency values in Table II allows us to better understand the physical meaning of the MAE of the different methods, in particular of the Harmonic vs the DC-SCIVR one. In the case of the harmonic frequencies reported in the second column of Table II, 22 out of 26 vibrational frequencies are higher than the experimental values. Thus, the MAE value of 48 cm^{-1} is because of estimates by excess. In the case of the DC-SCIVR calculations, it is the other way around. For example, 20 PG-EA values are underestimating the experimental frequencies, and the MAE value of 30 cm^{-1} is given by estimates by defect. Thus, the amount of anharmonic contribution introduced by the DC-SCIVR calculations is on average per mode ~78 cm^{-1}. This amount is comparable with the MAE under the column HO/MP2,

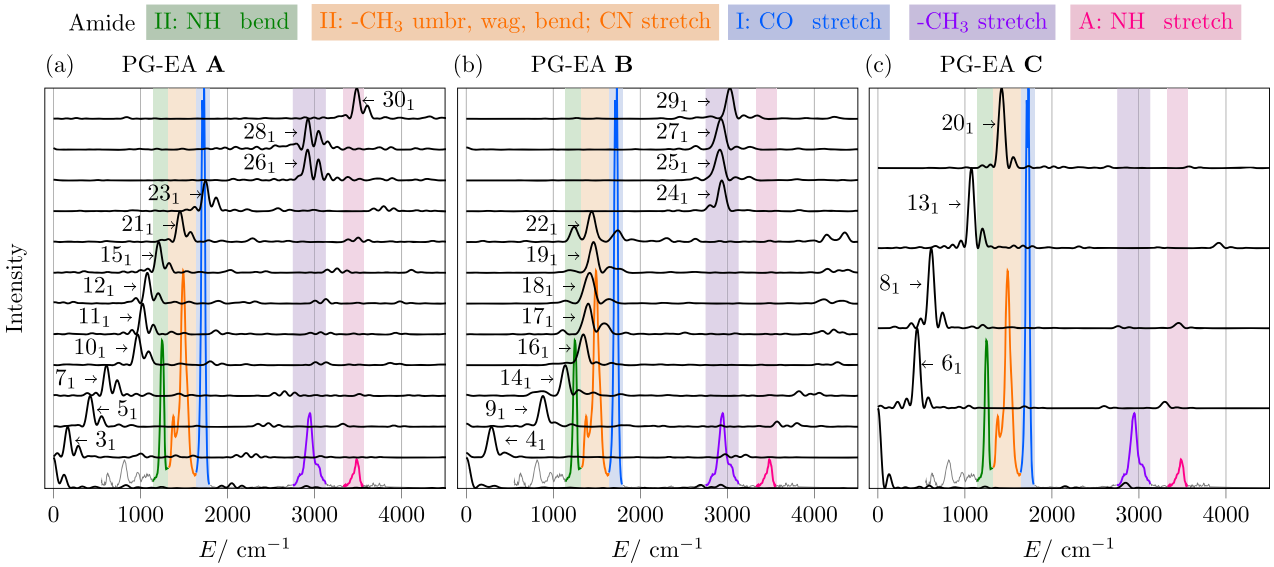

FIG. 10. Panels (a)–(c) are the three partial spectra for the three **A**, **B**, and **C** PG-EA subspaces. The colored spectrum is the experimental IR spectrum.[79] The black continuous spectra are DC-SCIVR spectra calculated using different ϵ_j values [see Eq. (8)] for each normal mode of the subspace. Colored windows are for different spectroscopic regions as denominated by experimentalists. These windows are commonly labeled amide II (green and orange), amide I (blue), and amide A (pink). The purple window corresponds to CH_3 stretching. Amide III, IV, and V bands are located below ~1200 cm^{-1} and are very dependent on the side chains.

TABLE II. Vibrational fundamental frequencies for *trans-N*-Methylacetamide (NMA). The first column denominates the vibrational modes. In the second column, the fundamental frequencies in the harmonic approximation (HO) are reported. The DC-SCIVR fundamental frequency of vibrations is obtained on the basis of subspace partition by means of PG-EA (third column), the Hessian method (Hess, fourth column), and the two-mode interaction method (two-mode, fifth column). The results are sorted by increasing the value of the harmonic frequencies and assigned by comparing the associated vibrational motion to the corresponding experimental description. Superscripts refer to the subspace that each mode belongs to. The sixth column reports the *ab initio* harmonic frequencies at the MP2/aug-cc-pVTZ level of theory (HO/MP2).[80] All data are compared by calculating the Mean Absolute Error (MAE) with respect to the experimental values (last column, exp.) by Ataka, Takeuchi, and Tasumi.[76]

	Frequency $^{\text{subID}}$ (cm^{-1})					
Modes No.	HO	DC SCIVR PG-EA (3 subs)	DC SCIVR hess (12 subs)	DC SCIVR two-mode (4 subs)	HO/MP2[80]	Expt.[76]
3	150	163^{A}	159^{a}	156^{α}	151	
4	290	289^{B}	285^{c}	283^{α}	259	279
5	393	421^{A}	417^{a}	416^{β}	347	429
6	433	451^{C}	451^{d}	448^{α}	423	439
7	621	609^{A}	608^{a}	606^{α}	630	619
8	629	613^{C}	611^{e}	612^{α}	633	658
9	866	881^{B}	875^{f}	871^{β}	883	857
10	995	967^{A}	970^{b}	972^{α}	1003	980
11	1038	1024^{A}	1028^{b}	1025^{α}	1058	1037
12	1112	1078^{A}	1080^{g}	1080^{α}	1119	1089
13	1132	1075^{C}	1069^{b}	1072^{α}	1169	
14	1166	1137^{B}	1138^{h}	1134^{α}	1195	1168
15	1260	1208^{A}	1210^{i}	1205^{α}	1290	1266
16	1391	1345^{B}	1344^{b}	1347^{α}	1402	1370
17	1415	1401^{B}	1388^{b}	1388^{α}	1460	1419
18	1434	1418^{B}	1409^{b}	1415^{γ}	1487	1432
19	1474	1462^{B}	1463^{j}	1461^{γ}	1494	1446
20	1485	1422^{C}	1426^{b}	1430^{α}	1499	1432
21	1491	1453^{A}	1453^{k}	1453^{α}	1529	1472
22	1550	1441^{B}	1444^{a}	1472^{β}	1561	1511
23	1772	1774^{A}	1747^{l}	1745^{α}	1749	1707
24	3019	2936^{B}	2934^{b}	2920^{α}	3088	2915
25	3040	2914^{B}	2923^{b}	2919^{δ}	3091	2958
26	3072	2920^{A}	2903^{b}	2908^{γ}	3165	2916
27	3125	2926^{B}	2933^{b}	2924^{δ}	3188	3008
28	3126	2924^{A}	2912^{b}	2912^{δ}	3188	3008
29	3137	3030^{B}	3044^{b}	3037^{α}	3197	2973
30	3630	3485^{A}	3484^{m}	3485^{β}	3703	3498
MAE exp	47.9	30.0	29.7	27.5	78.5	

where harmonic frequencies are calculated at a higher level of *ab initio* theory than the DFT-B3LYP/cc-pVDZ one. These considerations are suggesting that most probably, for this molecule, the discrepancies with respect to the experimental values are mainly due to the DFT level of *ab initio* theory. Conversely, only at a lower degree, the inaccuracy can be related to the semiclassical approximation or the quality of the potential energy surface fitting, as previously shown on other systems.[81] Unfortunately, a DC-SCIVR simulation at the MP2/aug-cc-pVTZ level of *ab initio* theory is out of reach at time of writing due to its computational burden.

IV. SUMMARY AND CONCLUSIONS

We have presented a machine learning algorithm based on a probability graph representation and an evolutionary algorithm procedure. The algorithm is able to find the best subdivision of the full-dimensional vibrational space into subspaces for model systems in which the best subspace division is known. Our approach is able to preserve Liouville's theorem for each subspace as much as possible and for a given maximum dimensionality of the subspaces. We proved that the clustering provided by PE-GA is, indeed, one of

the possible solutions that minimize the energy exchange between subspaces during the vibrational dynamics and thus the most convenient for DC-SCIVR and spectroscopic calculations in general. As an alternative, we have proposed a two-mode coupling scheme, which is not only less computationally intense but also less accurate. Application to the DC-SCIVR power spectrum calculation of *trans-N*-Methylacetamide is made manageable under Liouville's criterion restrictions only by means of these algorithms. The calculation of the DC-SCIVR power spectrum of *trans*-NMA with the subspace division selected with Liouville's criterion is manageable only employing these two algorithms.

The choice of the PG-EA parameters is arbitrary to some extent. As a matter of fact, the method is bound to look for new solutions at every epoch, and hence, it will get the global optimum, eventually. However, a sensible choice of the number of chromosomes and the mutation probability may significantly enhance the optimization. Assuming that the number of epochs is fixed, increasing the number of (elite) chromosomes means that the population evolves more slowly, enhancing the chances of eventually hitting the global optimum. However, as the evaluation of the fitness function is the most expensive step, a large population requires significantly more computational time. Conversely, using a small population means a fast evolution, and therefore, it is likely to obtain a fast local minimum. We suggest that a sensible choice of the mutation probability is in the interval of [0.001, 0.2]. This parameter makes sure that the algorithm does not get stuck in a local minimum, even if the whole population is homogeneous. It becomes less and less important as the pool of possible solutions and the number of elite chromosomes increase.

There are several quantum methods that can take advantage from partitioning the nuclear vibrational degrees of freedom. Clearly, dividing the vibrational space into putative independent subspaces is an approximation. However, if this subdivision is performed according to Jacobi's criterion, it may turn out not to be a rough approximation, especially for high-dimensional and loosely connected systems. We believe that at the affordable cost of a single adiabatic classical trajectory with Hessian calculation, the PG-EA algorithm can be useful to assist any method that has to deal with increasing computational costs with system dimensionality but is able to perform accurate spectroscopic calculations for each subspace independently. This may be the case, for example, of the local mode variant of multimode[82–84] or other semiclassical wavepacket propagation methods developed by other groups.[85–87]

The work we have presented provides a rigorous rationalization of the simplification of a larger dimensional problem into a set of lower-dimensional ones. The examples illustrated in this paper demonstrate that reliable spectroscopic results are obtained if a rigorous strategy is employed to get to a reasonable subspace partition while a non-educated, unwise choice of subspaces may lead to inaccurate or unreliable results. Our algorithms might serve as a powerful tool for advancing the computational spectroscopy of large molecules.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the European Research Council [Grant Agreement No. (647107)—SEMICOMPLEX—ERC-2014-CoG] under the European Union's Horizon 2020 research and innovation programme and from the Italian Ministry of Education, University, and Research (MIUR) (FARE Programme No. R16KN7XBRB-Project QURE).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] J. H. Holland, "Outline for a logical theory of adaptive systems," J. ACM **9**, 297–314 (1962).
[2] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," Mach. Learn. **3**, 95 (1988).
[3] J. H. Holland *et al.*, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (MIT Press, 1992).
[4] R. Freeman and W. Xili, "Design of magnetic resonance experiments by genetic evolution," J. Magn. Reson. **75**, 184–189 (1987).
[5] D. B. Hibbert, "Genetic algorithms in chemistry," Chemom. Intell. Lab. Syst. **19**, 277–293 (1993).
[6] R. Leardi, "Genetic algorithms in chemometrics and chemistry: A review," J. Chemom. **15**, 559–569 (2001).
[7] A. Niazi and R. Leardi, "Genetic algorithms in chemometrics," J. Chemom. **26**, 345–351 (2012).
[8] A. Beheshti, E. Pourbasheer, M. Nekoei, and S. Vahdani, "QSAR modeling of antimalarial activity of urea derivatives using genetic algorithm–multiple linear regressions," J. Saudi Chem. Soc. **20**, 282–290 (2016).
[9] B. Bhattacharya, G. R. Dinesh Kumar, A. Agarwal, Ş. Erkoç, A. Singh, and N. Chakraborti, "Analyzing Fe–Zn system using molecular dynamics, evolutionary neural nets and multi-objective genetic algorithms," Comput. Mater. Sci. **46**, 821–827 (2009).
[10] M. Ceotto, G. Di Liberto, and R. Conte, "Semiclassical 'divide-and-conquer' method for spectroscopic calculations of high dimensional molecular systems," Phys. Rev. Lett. **119**, 010401 (2017).
[11] G. Di Liberto, R. Conte, and M. Ceotto, "'Divide and conquer' semiclassical molecular dynamics: A practical method for spectroscopic calculations of high dimensional molecular systems," J. Chem. Phys. **148**, 014307 (2018).
[12] M. Wehrle, M. Šulc, and J. Vaníček, "On-the-fly *ab initio* semiclassical dynamics: Identifying degrees of freedom essential for emission spectra of oligothiophenes," J. Chem. Phys. **140**, 244114 (2014).
[13] E. J. Heller, "The semiclassical way to molecular spectroscopy," Acc. Chem. Res. **14**, 368–375 (1981).
[14] A. L. Kaledin and W. H. Miller, "Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels," J. Chem. Phys. **118**, 7174–7182 (2003).
[15] A. L. Kaledin and W. H. Miller, "Time averaging the semiclassical initial value representation for the calculation of vibrational energy levels. II. Application to H_2CO, NH_3, CH_4, CH_2D_2," J. Chem. Phys. **119**, 3078–3084 (2003).
[16] W. H. Miller, "Uniform semiclassical approximations for elastic scattering and eigenvalue problems," J. Chem. Phys. **48**, 464–467 (1968).
[17] W. H. Miller, "Semiclassical nature of atomic and molecular collisions," Acc. Chem. Res. **4**, 161–167 (1971).
[18] W. H. Miller, "Spiers memorial lecture quantum and semiclassical theory of chemical reaction rates," Faraday Discuss. **110**, 1–21 (1998).
[19] W. H. Miller, "The semiclassical initial value representation: A potentially practical way for adding quantum effects to classical molecular dynamics simulations," J. Phys. Chem. A **105**, 2942–2955 (2001).
[20] W. H. Miller, "Quantum dynamics of complex molecular systems," Proc. Natl. Acad. Sci. U. S. A. **102**, 6660–6664 (2005).

[21]X. Sun, H. Wang, and W. H. Miller, "On the semiclassical description of quantum coherence in thermal rate constants," J. Chem. Phys. **109**, 4190–4200 (1998).
[22]M. Thoss, H. Wang, and W. H. Miller, "Generalized forward–backward initial value representation for the calculation of correlation functions in complex systems," J. Chem. Phys. **114**, 9220–9235 (2001).
[23]T. Yamamoto and W. H. Miller, "Semiclassical calculation of thermal rate constants in full Cartesian space: The benchmark reaction D + H_2 → DH + H," J. Chem. Phys. **118**, 2135–2152 (2003).
[24]E. J. Heller, "Frozen Gaussians: A very simple semiclassical approximation," J. Chem. Phys. **75**, 2923–2931 (1981).
[25]M. F. Herman and E. Kluk, "A semiclassical justification for the use of non-spreading wavepackets in dynamics calculations," Chem. Phys. **91**, 27–34 (1984).
[26]W. H. Miller, "An alternate derivation of the Herman–Kluk (coherent state) semiclassical initial value representation of the time evolution operator," Mol. Phys. **100**, 397–400 (2002).
[27]E. Kluk, M. F. Herman, and H. L. Davis, "Comparison of the propagation of semiclassical frozen Gaussian wave functions with quantum propagation for a highly excited anharmonic oscillator," J. Chem. Phys. **84**, 326–334 (1986).
[28]K. G. Kay, "Semiclassical propagation for multidimensional systems by an initial value method," J. Chem. Phys. **101**, 2250–2260 (1994).
[29]K. G. Kay, "Integral expressions for the semiclassical time-dependent propagator," J. Chem. Phys. **100**, 4377–4392 (1994).
[30]K. G. Kay, "Numerical study of semiclassical initial value methods for dynamics," J. Chem. Phys. **100**, 4432–4445 (1994).
[31]K. G. Kay, "The Herman–Kluk approximation: Derivation and semiclassical corrections," Chem. Phys. **322**, 3–12 (2006).
[32]F. Grossmann and A. L. Xavier, "From the coherent state path integral to a semiclassical initial value representation of the quantum mechanical propagator," Phys. Lett. A **243**, 243–248 (1998).
[33]S. V. Antipov, Z. Ye, and N. Ananth, J. Chem. Phys. **142**, 184102 (2015).
[34]M. S. Church, S. V. Antipov, and N. Ananth, "Validating and implementing modified Filinov phase filtration in semiclassical dynamics," J. Chem. Phys. **146**, 234104 (2017).
[35]M. S. Church and N. Ananth, "Semiclassical dynamics in the mixed quantum-classical limit," J. Chem. Phys. **151**, 134109 (2019).
[36]M. Buchholz, E. Fallacara, F. Gottwald, M. Ceotto, F. Grossmann, and S. D. Ivanov, "Herman–Kluk propagator is free from zero-point energy leakage," Chem. Phys. **515**, 231–235 (2018).
[37]E. J. Heller, J. Chem. Phys. **94**, 2723–2729 (1991).
[38]G. Di Liberto and M. Ceotto, "The importance of the pre-exponential factor in semiclassical molecular dynamics," J. Chem. Phys. **145**, 144107 (2016).
[39]X. Ma, G. Di Liberto, R. Conte, W. L. Hase, and M. Ceotto, "A quantum mechanical insight into S_N2 reactions: Semiclassical initial value representation calculations of vibrational features of the $Cl^-\cdots CH_3Cl$ pre-reaction complex with the venus suite of codes," J. Chem. Phys. **149**, 164113 (2018).
[40]N. De Leon and E. J. Heller, "Semiclassical quantization and extraction of eigenfunctions using arbitrary trajectories," J. Chem. Phys. **78**, 4005–4017 (1983).
[41]M. Ceotto, S. Atahan, G. F. Tantardini, and A. Aspuru-Guzik, "Multiple coherent states for first-principles semiclassical initial value representation molecular dynamics," J. Chem. Phys. **130**, 234113 (2009).
[42]M. Ceotto, S. Atahan, S. Shim, G. F. Tantardini, and A. Aspuru-Guzik, "First-principles semiclassical initial value representation molecular dynamics," Phys. Chem. Chem. Phys. **11**, 3861–3867 (2009).
[43]M. Ceotto, G. F. Tantardini, and A. Aspuru-Guzik, "Fighting the curse of dimensionality in first-principles semiclassical calculations: Non-local reference states for large number of dimensions," J. Chem. Phys. **135**, 214108 (2011).
[44]F. Gabas, R. Conte, and M. Ceotto, "On-the-fly *ab initio* semiclassical calculation of Glycine vibrational spectrum," J. Chem. Theory Comput. **13**, 2378 (2017).
[45]M. Ceotto, D. Dell'Angelo, and G. F. Tantardini, "Multiple coherent states semiclassical initial value representation spectra calculations of lateral interactions for CO on Cu(100)," J. Chem. Phys. **133**, 054701 (2010).
[46]R. Conte, A. Aspuru-Guzik, and M. Ceotto, "Reproducing deep tunneling splittings, resonances, and quantum frequencies in vibrational spectra from a handful of direct *ab initio* semiclassical trajectories," J. Phys. Chem. Lett. **4**, 3407–3412 (2013).
[47]M. Micciarelli, F. Gabas, R. Conte, and M. Ceotto, "An effective semiclassical approach to IR spectroscopy," J. Chem. Phys. **150**, 184113 (2019).
[48]M. Micciarelli, R. Conte, J. Suarez, and M. Ceotto, "Anharmonic vibrational eigenfunctions and infrared spectra from semiclassical molecular dynamics," J. Chem. Phys. **149**, 064115 (2018).
[49]D. Tamascelli, F. S. Dambrosio, R. Conte, and M. Ceotto, "Graphics processing units accelerated semiclassical initial value representation molecular dynamics," J. Chem. Phys. **140**, 174109 (2014).
[50]M. Buchholz, F. Grossmann, and M. Ceotto, "Mixed semiclassical initial value representation time-averaging propagator for spectroscopic calculations," J. Chem. Phys. **144**, 094102 (2016).
[51]M. Buchholz, F. Grossmann, and M. Ceotto, "Application of the mixed time-averaging semiclassical initial value representation method to complex molecular spectra," J. Chem. Phys. **147**, 164110 (2017).
[52]M. Buchholz, F. Grossmann, and M. Ceotto, "Simplified approach to the mixed time-averaging semiclassical initial value representation for the calculation of dense vibrational spectra," J. Chem. Phys. **148**, 114107 (2018).
[53]Y. Zhuang, M. R. Siebert, W. L. Hase, K. G. Kay, and M. Ceotto, "Evaluating the accuracy of Hessian approximations for direct dynamics simulations," J. Chem. Theory Comput. **9**, 54–64 (2012).
[54]M. Ceotto, Y. Zhuang, and W. L. Hase, "Accelerated direct semiclassical molecular dynamics using a compact finite difference Hessian scheme," J. Chem. Phys. **138**, 054116 (2013).
[55]C. Aieta, M. Micciarelli, G. Bertaina, and M. Ceotto, "Anharmonic quantum nuclear densities from full dimensional vibrational eigenfunctions with application to protonated glycine," Nat. Commun. **11**, 4348 (2020).
[56]R. Conte, L. Parma, C. Aieta, A. Rognoni, and M. Ceotto, "Improved semiclassical dynamics through adiabatic switching trajectory sampling," J. Chem. Phys. **151**, 214107 (2019).
[57]G. Di Liberto, R. Conte, and M. Ceotto, "'Divide-and-conquer' semiclassical molecular dynamics: An application to water clusters," J. Chem. Phys. **148**, 104302 (2018).
[58]G. Bertaina, G. Di Liberto, and M. Ceotto, "Reduced rovibrational coupling Cartesian dynamics for semiclassical calculations: Application to the spectrum of the Zundel cation," J. Chem. Phys. **151**, 114307 (2019).
[59]F. Gabas, G. Di Liberto, R. Conte, and M. Ceotto, "Protonated glycine supramolecular systems: The need for quantum dynamics," Chem. Sci. **9**, 7894–7901 (2018).
[60]F. Gabas, G. Di Liberto, and M. Ceotto, "Vibrational investigation of nucleobases by means of divide and conquer semiclassical dynamics," J. Chem. Phys. **150**, 224107 (2019).
[61]F. Gabas, R. Conte, and M. Ceotto, "Semiclassical vibrational spectroscopy of biological molecules using force fields," J. Chem. Theory Comput. **16**, 3476–3485 (2020).
[62]M. Cazzaniga, M. Micciarelli, F. Moriggi, A. Mahmoud, F. Gabas, and M. Ceotto, "Anharmonic calculations of vibrational spectra for molecular adsorbates: A divide-and-conquer semiclassical molecular dynamics approach," J. Chem. Phys. **152**, 104104 (2020).
[63]M. L. Brewer, J. S. Hulme, and D. E. Manolopoulos, "Semiclassical dynamics in up to 15 coupled vibrational degrees of freedom," J. Chem. Phys. **106**, 4832–4839 (1997).
[64]R. Sinkhorn and P. Knopp, "Concerning nonnegative matrices and doubly stochastic matrices," Pac. J. Math. **21**, 343–348 (1967).
[65]N. J. Higham, "Gaussian elimination," Wiley Interdiscip. Rev.: Comput. Stat. **3**, 230–238 (2011).
[66]M. Aigner and G. M. Ziegler, "Cayley's formula for the number of trees," in *Proofs from The Book*," (Springer, Berlin, Heidelberg, 1998).
[67]R. R. Sokal and C. D. Michener, "A statistical method for evaluating systematic relationships," Univ. Kans. Sci. Bull. **38**, 1409–1438 (1958).
[68]T. J. Lee, J. M. L. Martin, and P. R. Taylor, "An accurate *ab initio* quartic force field and vibrational frequencies for CH_4 and isotopomers," J. Chem. Phys. **102**, 254–261 (1995).

[69]D. L. Gray and A. G. Robiette, "The anharmonic force field and equilibrium structure of methane," Mol. Phys. **37**, 1901–1920 (1979).
[70]W. T. Raynes, P. Lazzeretti, R. Zanasi, A. J. Sadlej, and P. W. Fowler, "Calculations of the force field of the methane molecule," Mol. Phys. **60**, 509–525 (1987).
[71]S. Carter, H. M. Shnider, and J. M. Bowman, "Variational calculations of rovibrational energies of CH_4 and isotopomers in full dimensionality using an *ab initio* potential," J. Chem. Phys. **110**, 8417–8423 (1999).
[72]C. Qu and J. M. Bowman, "A fragmented, permutationally invariant polynomial approach for potential energy surfaces of large molecules: Application to *N*-methyl acetamide," J. Chem. Phys. **150**, 141101 (2019).
[73]X. G. Chen, R. Schweitzer-Stenner, S. A. Asher, N. G. Mirkin, and S. Krimm, "Vibrational assignments of *trans*-*N*-methylacetamide and some of its deuterated isotopomers from band decomposition of IR, visible, and resonance Raman spectra," J. Phys. Chem. **99**, 3074–3083 (1995).
[74]J. Kubelka and T. A. Keiderling, "*Ab initio* calculation of amide carbonyl stretch vibrational frequencies in solution with modified basis sets. 1. *N*-methyl acetamide," J. Phys. Chem. A **105**, 10922–10928 (2001).
[75]H. Torii, T. Tatsumi, T. Kanazawa, and M. Tasumi, "Effects of intermolecular hydrogen-bonding interactions on the amide I mode of *N*-methylacetamide: Matrix-isolation infrared studies and *ab initio* molecular orbital calculations," J. Phys. Chem. B **102**, 309–314 (1998).
[76]S. Ataka, H. Takeuchi, and M. Tasumi, "Infrared studies of the less stable *cis* form of *N*-methylformmaide and *N*-methylacetamide in low-temperature nitrogen matrices and vibrational analyses of the *trans* and *cis* forms of these molecules," J. Mol. Struct. **113**, 147–160 (1984).
[77]L. C. Mayne and B. Hudson, "Resonance Raman spectroscopy of *N*-methylacetamide: Overtones and combinations of the carbon-nitrogen stretch (amide II′) and effect of solvation on the carbon-oxygen double-bond stretch (amide I) intensity," J. Phys. Chem. **95**, 2962–2967 (1991).
[78]N. E. Triggs and J. J. Valentini, "An investigation of hydrogen bonding in amides using Raman spectroscopy," J. Phys. Chem. **96**, 6922–6931 (1992).
[79]W. E. Wallace, "Infrared spectra," in *NIST Chemistry WebBook*, NIST Standard Reference Database Number 69, edited by P. J. Linstrom and W. G. Mallard (National Institute of Standards and Technology, Gaithersburg, MD, 2020).
[80]A. L. Kaledin and J. M. Bowman, "Full dimensional quantum calculations of vibrational energies of *N*-methyl acetamide," J. Phys. Chem. A **111**, 5593–5598 (2007).
[81]R. Conte, G. Botti, and M. Ceotto, "Sensitivity of semiclassical vibrational spectroscopy to potential energy surface accuracy: A test on formaldehyde," Vib. Spectrosc. **106**, 103015 (2020).
[82]Y. Wang and J. M. Bowman, "Coupled-monomers in molecular assemblies: Theory and application to the water tetramer, pentamer, and ring hexamer," J. Chem. Phys. **136**, 144113 (2012).
[83]H. Liu, Y. Wang, and J. M. Bowman, "Quantum calculations of intramolecular IR spectra of ice models using *ab initio* potential and dipole moment surfaces," J. Phys. Chem. Lett. **3**, 3671–3676 (2012).
[84]Y. Wang and J. M. Bowman, "IR spectra of the water hexamer: Theory, with inclusion of the monomer bend overtone, and experiment are in agreement," J. Phys. Chem. Lett. **4**, 1104–1108 (2013).
[85]T. Begusic, J. Roulet, and J. Vanicek, "On-the-fly *ab initio* semiclassical evaluation of time-resolved electronic spectra," J. Chem. Phys. **149**, 244115 (2018).
[86]A. Patoz, T. Begušić, and J. Vaníček, "On-the-fly *ab initio* semiclassical evaluation of absorption spectra of polyatomic molecules beyond the condon approximation," J. Phys. Chem. Lett. **9**, 2367–2372 (2018).
[87]M. Wehrle, S. Oberli, and J. Vaníček, "On-the-Fly *ab initio* semiclassical dynamics of floppy molecules: Absorption and photoelectron spectra of ammonia," J. Phys. Chem. A **119**, 5685–5690 (2015).