The Journal of Chemical Physics

Vibrational investigation of nucleobases by means of divide and conquer semiclassical dynamics

Cite as: J. Chem. Phys. **150**, 224107 (2019); https://doi.org/10.1063/1.5100503
Submitted: 17 April 2019 . Accepted: 22 May 2019 . Published Online: 12 June 2019

Fabio Gabas, Giovanni Di Liberto, and Michele Ceotto

AIP Publishing

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

Vibrational investigation of nucleobases by means of divide and conquer semiclassical dynamics

Cite as: J. Chem. Phys. 150, 224107 (2019); doi: 10.1063/1.5100503
Submitted: 17 April 2019 • Accepted: 22 May 2019 •
Published Online: 12 June 2019

Fabio Gabas, Giovanni Di Liberto, and Michele Ceotto[a)]

AFFILIATIONS

Dipartimento di Chimica, Università degli Studi di Milano, via C. Golgi 19, 20133 Milano, Italy

[a)]michele.ceotto@unimi.it

ABSTRACT

In this work, we report a computational study of the vibrational features of four different nucleobases employing the divide-and-conquer semiclassical initial value representation molecular dynamics method. Calculations are performed on uracil, cytosine, thymine, and adenine. Results show that the overall accuracy with respect to experiments is within 20 wavenumbers, regardless of the dimensionality of the nucleobase. Vibrational estimates are accurate even in the complex case of cytosine, where two relevant conformers are taken into account. These results are promising in the perspective of future studies on more complex systems, such as nucleotides or nucleobase pairs.

I. INTRODUCTION

The nucleobases adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U) represent the inner part of nucleotides that are the building blocks of the nucleic acids, DNA and RNA. More specifically, a nucleotide chain constitutes a single strand that interacts with another strand through the bases. These strands compose the secondary and tertiary structures of the nucleic acids as, for example, the famous double helix discovered by Watson and Crick.[1] The interaction between nucleobases follows a specific pairing: A and G, classified as purines, interact, respectively, with T (for DNA) or U (for RNA) and C, called pyrimidines. The nucleobases' structure is particularly relevant since it is directly linked to their functionality.[2] When some of them exist in tautomeric forms, only one conformation is predominant in nature. Also, different tautomers can bring to mispairings between pyrimidines and purines, leading to phenomena known as mutagenesis.[1–6] For these reasons, an accurate investigation of the structure and properties of all the nucleobase conformations is particularly important. Clearly, detailed studies in the condensed phase are the final goal to understand these properties.[7–13] However, a preliminary investigation in vacuum is a mandatory step in order to have a spectroscopic clear picture and to be able to discriminate between more complex condensed phase interactions. Specifically, in the gas phase,[14] it is possible to study the intrinsic properties of such molecules without any intermolecular interaction. Besides, the deep knowledge of biomolecules and their ionic behavior in vacuum is important also for a better understanding and designing of various spectroscopy techniques, such as mass and vibrational spectroscopy. Thus, gas-phase vibrational spectroscopy can certainly help in the characterization and fundamental understanding of nucleobases.

Over the past decades, a lot of experimental work has been done in this direction, starting from the measurement of accurate spectra for adenine, thymine, and uracil,[15–22] and going to the more complex spectra of cytosine and guanine, which show more than a single relevant isomer, due to the facile tautomerism.[23–29] These experimental spectra have been often combined with theoretical calculations for a better interpretation. However, given the molecular size of these systems, in most cases, the theoretical prediction was provided by harmonic frequencies calculated at the equilibrium geometry. Even if this approach is computationally cheap and relatively immediate to apply, it neglects all possible resonances and anharmonicities of the potential. Recently, very detailed vibrational studies of uracil have been presented using the canonical van Vleck second-order vibrational perturbation theory (CVPT2), the fully automated generalized second-order vibrational perturbation (GVPT2) approach, and the Hierarchical Intertwined Reduced-Rank Block Power Method (HIRRBPM), obtaining very accurate results compared to the experiment.[30–32] Also, Perturbation Theory (PT2) has been successful for

adenine[33] and for the oxo isomer of citosine.[34] However, to the best of our knowledge, in the case of thymine, a complete anharmonic computational spectrum is still missing.

In this work, we present a vibrational spectroscopic study of four nucleobases together with their eventual principal isomers, by means of the semiclassical initial value representation (SCIVR) molecular dynamics.[35–49] SCIVR has been proved in recent years to be very powerful for spectroscopic calculations.[50–64] It does not suffer from Zero Point Energy Leakage (ZPEL),[65] and it can be employed for vibrational eigenfunction calculations.[66–68] In particular, the divide-and-conquer semiclassical initial value representation (DC SCIVR) method for vibrational spectroscopy has been employed reliably in several applications, allowing to obtain semiclassical vibrational spectra of variously sized molecules without any relevant loss of accuracy, and with an average deviation of 20 wavenumbers from either exact or experimental results.[69,70] Specifically, DC SCIVR has been applied to study challenging high-dimensional and complex systems such as fullerene, supramolecular glycine-based molecules, and water clusters.[69,71,72]

In the present work, we calculate power spectra by performing *ab initio* molecular dynamics simulations, which has been widely and successfully employed previously for full-dimensional on-the-fly semiclassical applications.[53–55,72,73] More specifically, we want to provide not only an extensive vibrational study of nucleobases but also to open the route to *ab initio* DC SCIVR calculations on DNA-related molecules. For these goals, our quantum mechanical vibrational estimates are compared with both experiments and other theoretical calculations. These calculations will check and prove DC SCIVR method feasibility and reliability and provide the confidence for future calculations of increasing dimensionality up to pairs of bases and higher dimensional sequences.[74–77] Here, we investigate four of the five nucleobases, i.e., uracil, cytosine, thymine, and adenine. Since gas-phase spectra of the remaining nucleobase, guanine, are given by a combination of its four most stable conformers and their signals are all within few wavenumbers, guanine is not a good benchmark for testing the accuracy of our method and it has not been investigated.[25,27,78]

The paper is organized as follows. Section II recalls the DC SCIVR method for vibrational spectra calculations, together with the computational setup. In Sec. III, we present our results about the nucleobases and we compare them to different experiments and other theoretical calculations. Finally, Sec. IV provides some conclusions and future perspectives.

II. DIVIDE AND CONQUER SEMICLASSICAL DYNAMICS

Vibrational power spectra are obtained via Fourier transform of the semiclassical approximation to the survival amplitude $\langle\chi|\chi_t\rangle$,

$$I(E) = \frac{1}{2\pi\hbar}\int dt e^{\frac{i}{\hbar}Et}\langle\chi|e^{-\frac{i}{\hbar}\hat{H}t}|\chi\rangle, \tag{1}$$

where $|\chi\rangle$ is a given reference state, $\hat{H}$ is the Hamiltonian operator, and $\exp\left[-i\hat{H}t/\hbar\right]$ is the quantum mechanical time-evolution operator. We calculate this propagator using the semiclassical approximation.

Semiclassical theory takes advantage of the Feynman path integral representation of the quantum propagator,[79] in which a quantum mechanical amplitude, describing the probability for a particle of mass m to move from a certain initial state $\mathbf{q}(0)$ to a final one $\mathbf{q}(t)$ at time t, is calculated as

$$\langle\mathbf{q}(t)|e^{-\frac{i}{\hbar}\hat{H}t}|\mathbf{q}(0)\rangle = \sqrt{\left(\frac{m}{2\pi i\hbar t}\right)^F}\int \wp[\mathbf{q}(t)]e^{\frac{i}{\hbar}S_t(\mathbf{q}(t),\mathbf{q}(0))}, \tag{2}$$

where the integral of the differential $\wp[\mathbf{q}(t)]$ is given by the summation over all possible paths from $\mathbf{q}(0)$ to $\mathbf{q}(t)$. $S_t(\mathbf{q}(t),\mathbf{q}(0))$ is the action of each path, and F is the number of degrees of freedom. When the stationary phase approximation for the path integration is enforced, only classical paths contribute. This approximation leads to the well known van Vleck propagator,[80] where the integral is replaced by a sum over all classical paths connecting the starting position $\mathbf{q}(0)$ to the ending one $\mathbf{q}(t)$ in an amount of time t,

$$\langle\mathbf{q}(t)|e^{-\frac{i}{\hbar}\hat{H}t}|\mathbf{q}(0)\rangle \sim \sqrt{\left(\frac{1}{2\pi i\hbar}\right)^F}\sum_{Cl\,paths} e^{\frac{i}{\hbar}S_t(\mathbf{q}(t),\mathbf{q}(0))} \times\left|\frac{\partial\mathbf{p}(0)}{\partial\mathbf{q}(t)}\right|^{\frac{1}{2}} e^{-i\pi\nu/2}. \tag{3}$$

The index υ is called Morse, or Maslov index, and it accounts for the sign change of the determinant $\left|\frac{\partial\mathbf{p}(0)}{\partial\mathbf{q}(t)}\right|$. This version of the propagator is quite cumbersome and not practical since it requires dealing with a root search problem with fixed boundary conditions. This issue has been overcome by the initial value representation trick proposed by Miller,[37] where the integrand of Eq. (1) is written as

$$\langle\chi|e^{-\frac{i}{\hbar}\widehat{H}t}|\chi\rangle \sim \sqrt{\left(\frac{1}{2\pi i\hbar}\right)^F}\iint d\mathbf{q}(0)d\mathbf{p}(0)\langle\chi|\mathbf{q}(t)\rangle\langle\mathbf{q}(0)|\chi\rangle \times\left|\frac{\partial\mathbf{q}(t)}{\partial\mathbf{p}(0)}\right|^{\frac{1}{2}} e^{\frac{i}{\hbar}S_t(\mathbf{q}(0),\mathbf{p}(0))}e^{-i\pi\nu/2}. \tag{4}$$

In this version of the propagator, the survival amplitude can be obtained by sampling trajectories with different initial conditions $(\mathbf{p}(0),\mathbf{q}(0))$ and it is amenable to the Monte Carlo integration. Later, Heller proposed a very flexible and numerically stable representation of the semiclassical propagator based on coherent states[81,82] of the type

$$\langle\mathbf{q}|\mathbf{p}(t),\mathbf{q}(t)\rangle = \left(\frac{\det(\mathbf{\Gamma})}{\pi^F}\right)^{\frac{1}{4}} e^{-\frac{1}{2}(\mathbf{q}-\mathbf{q}(t))^T\mathbf{\Gamma}(\mathbf{q}-\mathbf{q}(t))+\frac{i}{\hbar}\mathbf{p}^T(t)(\mathbf{q}-\mathbf{q}(t))}, \tag{5}$$

where in our simulations, the $\boldsymbol{\Gamma}$ matrix is chosen to be constant in time and diagonal, with elements equal to the normal mode vibrational frequencies. The wavepacket in Eq. (5) is centered at the classical position $\mathbf{q}(t)$ and momentum $\mathbf{p}(t)$, and it follows the classical trajectory. The expression of the quantum propagator in this representation is the well known Heller-Herman-Kluk-Kay (HHKK) propagator[83–85]

$$e^{-\frac{i}{\hbar}\hat{H}t} = \left(\frac{1}{2\pi\hbar}\right)^F \iint d\mathbf{p}(0)d\mathbf{q}(0)C_t(\mathbf{p}(0),\mathbf{q}(0))e^{\frac{i}{\hbar}S_t(\mathbf{p}(0),\mathbf{q}(0))} \times |\mathbf{p}(t),\mathbf{q}(t)\rangle\langle\mathbf{p}(0),\mathbf{q}(0)|, \quad (6)$$

where $C_t(\mathbf{q}(0),\mathbf{p}(0))$ is the pre-exponential factor and it is equal to

$$C_t(\mathbf{q}(0),\mathbf{p}(0)) = \sqrt{\det\left[\frac{1}{2}\left(\mathbf{M}_{qq} + \boldsymbol{\Gamma}^{-1}\mathbf{M}_{pp}\boldsymbol{\Gamma} + \frac{i}{\hbar}\boldsymbol{\Gamma}^{-1}\mathbf{M}_{pq} + \frac{\hbar}{i}\mathbf{M}_{qp}\boldsymbol{\Gamma}\right)\right]}, \quad (7)$$

and where $\mathbf{M}_{ij}$ are the monodromy (or stability) matrix elements defined as $\mathbf{M}_{ij} = \frac{\partial \mathbf{i}_t}{\partial \mathbf{j}_0}$, where $\mathbf{i}_t$ is calculated at time t and $\mathbf{j}_0$ is calculated at time zero.[86,87] The Monte Carlo integration in Eq. (6) usually needs a high number of trajectories for convergence. For systems having more than a few degrees of freedom, some filtering techniques have been proposed to speed up the convergence.[88–91] A well established procedure is the time-averaging (TA) filtering, in which the semiclassical integrand can be worked out to be positively definite[92] by taking advantage of the so-called separable approximation of the pre-exponential factor, where only the phase is taken into account, i.e., $C_t(\mathbf{q}(0),\mathbf{p}(0)) \sim e^{\frac{i}{\hbar}\varphi_t}$ and $\varphi_t = \text{phase}[C_t(\mathbf{q}(0),\mathbf{p}(0))]$. Using the propagator in Eq. (6), the time-averaging filter and the separable approximation, one obtains the following formula for the spectral density:[92]

$$I(E) = \left(\frac{1}{2\pi\hbar}\right)^F \iint d\mathbf{p}(0)d\mathbf{q}(0)\frac{1}{2\pi\hbar T} \times \left|\int_0^T e^{\frac{i}{\hbar}[S_t(\mathbf{p}(0),\mathbf{q}(0))+Et+\varphi_t]}\langle\chi|\mathbf{p}(t)\mathbf{q}(t)\rangle dt\right|^2. \quad (8)$$

Equation (8) has been employed to perform vibrational estimates of small-sized molecules requiring roughly a 1000 of classical trajectories per degree of freedom to converge.[93–96] Unfortunately when prefitted Potential Energy Surfaces (PESs) are not available, the number of required trajectories is still too computational demanding for on-the-fly or direct dynamics approaches. In this event, only few classical trajectories can be afforded. For these reasons, in recent years, starting from Eq. (8), in our group, the Multiple Coherent State approach (MC SCIVR) has been developed,[53–55,97] in which by properly choosing the initial conditions of the classical trajectories and the reference state it is possible to regain spectra with few or even one classical trajectory and retaining the typical semiclassical accuracy of roughly 20 cm^{-1}. In the single trajectory implementation, the initial conditions are chosen as $(\mathbf{p}(0),\mathbf{q}(0)) = \left(\mathbf{p}_{eq},\mathbf{q}_{eq}\right)$, where $\mathbf{p}_{eq}$ and $\mathbf{q}_{eq}$ stand for the coordinates of the reference coherent state $|\chi\rangle$. The phase space integral of Eq. (8) reduces to a single trajectory formulation for each spectroscopic peak of the type

$$I(E) = \left(\frac{1}{2\pi\hbar}\right)^F \frac{1}{2\pi\hbar T}\left|\int_0^T e^{\frac{i}{\hbar}\left[S_t\left(\mathbf{p}_{eq},\mathbf{q}_{eq}\right)+Et+\varphi_t\left(\mathbf{p}_{eq},\mathbf{q}_{eq}\right)\right]} \times \langle\chi|\mathbf{p}(t),\mathbf{q}(t)\rangle dt\right|^2, \quad (9)$$

where $\mathbf{q}_{eq}$ is usually chosen to be equal to the equilibrium configuration and $\mathbf{p}_{eq}$ is set in order to provide a total initial kinetic energy equal to the harmonic zero point energy (ZPE), which is distributed as $p^i_{eq} = \sqrt{\omega_i}$ among mass-scaled vibrational normal modes, where ω_i is the harmonic frequency of the ith mode. $\mathbf{p}(t)$ and $\mathbf{q}(t)$ are the respective time-evolved quantities. Equation (9) is our working equation for MC SCIVR calculation. Moreover, in the MC SCIVR approach, the reference state can be tailored to decompose the spectrum mode by mode. For an ith mode, we have

$$|\chi\rangle_i = \prod_{j=1}^F\left[|p^i_{j,eq}(0),q^i_{j,eq}(0)\rangle + \varepsilon^i| - p^i_{j,eq}(0),q^i_{j,eq}(0)\rangle\right], \quad (10)$$

where the index j runs over the number of vibrational degrees of freedom F.[54,97] If the ε vector is set equal to one, then the zero point energy peak is obtained (together with other even peaks), while by setting $\varepsilon^i = -1$ only for a certain ith mode, then the ith fundamental excitation will be enhanced, together with the other odd overtones of ith mode.[97] This tool is particularly helpful in presence of crowded spectra, where peaks are very close in energy, as it commonly happens by increasing the dimensionality of the molecule under exam. This strategy successfully allowed to recover accurate spectra of molecules as complex as glycine.[73]

Unfortunately, MC SCIVR runs out of steam when the system dimensionality overcomes 25–30 degrees of freedom because of the so-called curse of dimensionality problem. Recently, we have addressed this issue by proposing the DC SCIVR method,[69,70] where full dimensional classical simulations are projected onto subdimensional spaces for semiclassical subdimensional spectroscopic calculations. The same working formula of Eq. (8) is employed but formulated in terms of subspace coordinates. More specifically, the spectral density for a M-dimensional subspace is

$$\widetilde{I}(E) = \left(\frac{1}{2\pi\hbar}\right)^M \iint d\tilde{\mathbf{p}}(0)d\tilde{\mathbf{q}}(0)\frac{1}{2\pi\hbar T}\left|\int_0^T e^{\frac{i}{\hbar}\left[\tilde{S}_t(\tilde{\mathbf{p}}(0),\tilde{\mathbf{q}}(0))+Et+\tilde{\phi}_t\right]} \times \langle\tilde{\boldsymbol{\chi}}|\tilde{\mathbf{p}}(t),\tilde{\mathbf{q}}(t)\rangle dt\right|^2, \quad (11)$$

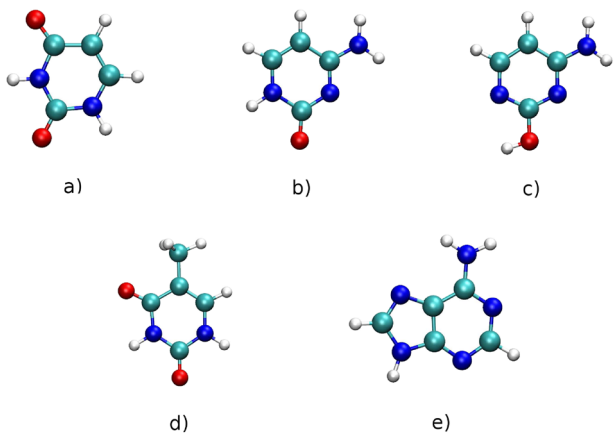

FIG. 1. Molecular structures of the simulated nucleobases. (a) uracil, (b) and (c), respectively, oxocytosine and hydroxycytosine conformers, (d) thymine, and (e) adenine. Red: O, gray: H, blue: N, and dark green: C.

where ~ denotes the projected quantities. The full-dimensional spectrum is recovered as a combination of reduced dimensionality ones. $\langle \tilde{\mathbf{x}}|\tilde{\mathbf{p}}\tilde{\mathbf{q}}\rangle = \prod_{i=1}^{M}\langle x_i|p_i q_i\rangle$ are the projected coherent states written as direct product of monodimensional ones and involving only the degrees of freedom in the subspace. $\tilde{S}_t$ is the projected action, and $\tilde{\phi}_t$ is the phase of the projected pre-exponential factor. This latter term is obtained from the reduced dimensionality monodromy matrix elements upon a singular value decomposition of the full-dimensional ones.[98,99] The projected action functional is written as

$$\tilde{S}_t(\tilde{\mathbf{p}}(0),\tilde{\mathbf{q}}(0)) = \int_0^t \left[\frac{1}{2}m\dot{\tilde{\mathbf{q}}}^2(t') - V_S\big(\tilde{\boldsymbol{q}}(t')\big)\right]dt', \qquad (12)$$

where the kinetic part is trivially projected since it is naturally separable. For the potential part, we assume that the projected potential depends on the degrees of freedom contained in the subspace and the remaining ones are downgraded to parameters. We include a time-dependent external scalar field $\lambda(t)$ to ensure that the equation of the projected potential is exact in the limit of a separable system and it accounts for the contribution arising from the degrees of freedom not belonging to the subspace,

$$V_S(\tilde{\mathbf{q}}(t)) = V\big(\tilde{\mathbf{q}}(t);\mathbf{q}^{eq}_{N_{vib}-M}\big) + \lambda(t), \qquad (13)$$

$$\lambda(t) = V\big(\tilde{\mathbf{q}}(t);\mathbf{q}_{N_{vib}-M}(t)\big) - \big[V\big(\tilde{\mathbf{q}}(t);\mathbf{q}^{eq}_{N_{vib}-M}\big) + V\big(\mathbf{q}^{eq}_{M};\mathbf{q}_{N_{vib}-M}(t)\big)\big]. \qquad (14)$$

Clearly, the definition of the subspaces is a critical issue within this approach since it is responsible for the accuracy of the action and the monodromy matrix calculation. As for Eq. (9), Eq. (11)

TABLE I. Vibrational fundamental excitations of uracil. Values are reported in cm^{-1}. The first column reports the label of the excitation, while the second one reports the experimental values from Refs. 18 and 103. Third and fourth columns contain our harmonic estimates and the hybrid harmonic/anharmonic CCSD(T)/B3LYP values, respectively. GVPT2 and CVPT2 estimates of Refs. 30 and 31 are reported under the VPT2 columns. The last two columns report the energy levels calculated with MC SCIVR and DC SCIVR. "MAE" stands for Mean Absolute Error.

		Harmonic	Harmonic/anharmonic	VPT2		Semiclassical	
Label	Exp	B3LYP/ aug-cc-pvdz	CCSD(T)/ B3LYP[30]	GVPT2	CVPT2	MC SCIVR	DC SCIVR
1		168	140	132	139.7	156	160
2	185	181	159	155	157.9	164	170
3	391	395	387	386	385.6	380	370
4	411	405	388	384	384.4	396	390
5	516	522	517	510	510.8	524	510
6	537	543	541	530	531.1	520	535
7	562	560	545	549	535.3	552	550
8	551	582	559	555	549.4	544	560
9	662	698	670	654	651.4	676	680
10	718	729	728	711	715.8	708	710
11	757	764	765	746	756.1	680	750
12	759	771	773	751	751.8	748	750
13	804	821	814	793	803.2	788	820
14	987	963	973	954	955.9	940	940
15	958	966	968	940	947.5	952	950
16	980	988	995	978	979.9	968	970
17	1075	1083	1084	1061	1060.8	1052	1060
18	1185	1194	1205	1176	1179.9	1104	1200
19	1217	1222	1248	1221	1214.4	1208	1200
20	1359	1379	1394	1384	1382.0	1336	1340
21	1389	1402	1414	1355	1351.0	1340	1340
22	1400	1412	1427	1388	1394.1	1376	1370
23	1472	1497	1505	1466	1462.7	1464	1460
24	1643	1673	1678	1643	1642.8	1644	1630
25	1706	1757	1762	1733	1729.5	1720	1720
26	1764	1788	1790	1761	1761.2	1764	1760
27		3210	3218	3072	3081.0	2980	3070
28		3250	3253	3117	3133.6	3144	3160
29	3435	3585	3602	3436	3435.2	3480	3510
30	3485	3631	3653	3485	3486.3	3536	3550
MAE		26	30	10	9	22	20

can be reduced to a single trajectory formulation and this formulation will be employed in our calculations. Next, one desires to group in the same subspace the most interacting degrees of freedom. We described several strategies in Ref. 70. Here, we employ the one based on the Hessian matrix since the off-diagonal terms indicate the level of coupling between degrees of freedom. The procedure to separate the full-dimensional space starts by defining a time-averaged Hessian matrix **H** along a test trajectory, where $\overline{H}_{ij} = \frac{1}{N_{steps}} \sum_{i=1}^{N_{steps}} H_{ij}$.[69,70] We fix a threshold parameter that is comparable with the normal mode mass-scaled off-diagonal terms of $\overline{\boldsymbol{H}}$. If $\overline{H}_{ij} \geq \varepsilon$, then the level of coupling is considered significant and the degrees of freedom are enrolled in the same subspace. Instead, if $\overline{H}_{ij} \leq \varepsilon$, i and j modes are on different subspaces, unless there exists a third mode k such that $\overline{H}_{ki} \geq \varepsilon$ and/or $\overline{H}_{kj} \geq \varepsilon$. The best possible scenario is when there are few subspaces and they are big enough to retain most of the couplings. However, in practice, the threshold choice is often a trade-off between accuracy and feasibility of the semiclassical calculation.

Molecular dynamics simulations in this work are performed using the NWChem package[100] at a density functional theory (DFT) B3LYP/aug-cc-pvdz level of theory,[101] which is a typical setup for semiclassical *ab initio* calculations and represents a good trade-off between accuracy and computational overhead. For each conformer, we evolve a single trajectory for a total of 25 000 a.u., starting from $(\mathbf{p}_{eq}, \mathbf{q}_{eq})$, where $\mathbf{q}_{eq}$ is the equilibrium configuration and $\mathbf{p}_{eq}$ is the momenta vector, setting each momentum to have a total initial energy equal to the harmonic ZPE of the molecule. With the exception of the uracil molecule, we calculate the Hessian every two steps and approximate it otherwise.[86,87] This is a typical setup that is often enough to recover MC SCIVR and DC SCIVR vibrational spectra with an average accuracy within 20–25 wavenumbers.[53,73,96,102]

III. RESULTS AND DISCUSSION

We start our investigation with the lowest dimension nucleobase, i.e., uracil. For this system, we perform both MC SCIVR and DC SCIVR calculations to prove once more the reliability of the DC SCIVR method. Results are compared with experiments and high level VPT2 calculations. Then, we move to cytosine, for which there are two conformers which are spectroscopic relevant, and to thymine and adenine vibrational spectra and compare them to the available experimental results. With the exception of uracil, full-dimensional MC SCIVR calculations cannot be afforded for the other nucleobases. In these cases, only DC SCIVR simulations will be performed. The molecular structures of these molecules are reported in Fig. 1.

A. Uracil

Uracil is made of one pyrimidine ring resulting into 12 atoms. The molecular structure of the global minimum is reported in panel (a) of Fig. 1. For this molecular system, the energy difference between the global minimum (oxo form) and its tautomer (hydroxy form) is around 45 kJ/mol.[21] For this reason, we perform our semiclassical study only on the structure reported in panel (a) of Fig. 1 since it is expected to be by far the most representative one. In the past years, this molecule was the subject of several studies, both with experimental and theoretical approaches. Specifically, we will compare our semiclassical vibrational estimates with experimental values[18,103,104] and GVPT2 and CVPT2 calculated energy levels.[30,31] As a first evaluation of the level of theory, we report in Table I the computed harmonic frequencies of the molecule at B3LYP/aug-cc-pvdz level of theory together with those of Ref. 30, calculated using a hybrid CCSD(T)/B3LYP force field, where harmonic CCSD(T) estimates have been corrected using GVPT2 estimates at the level

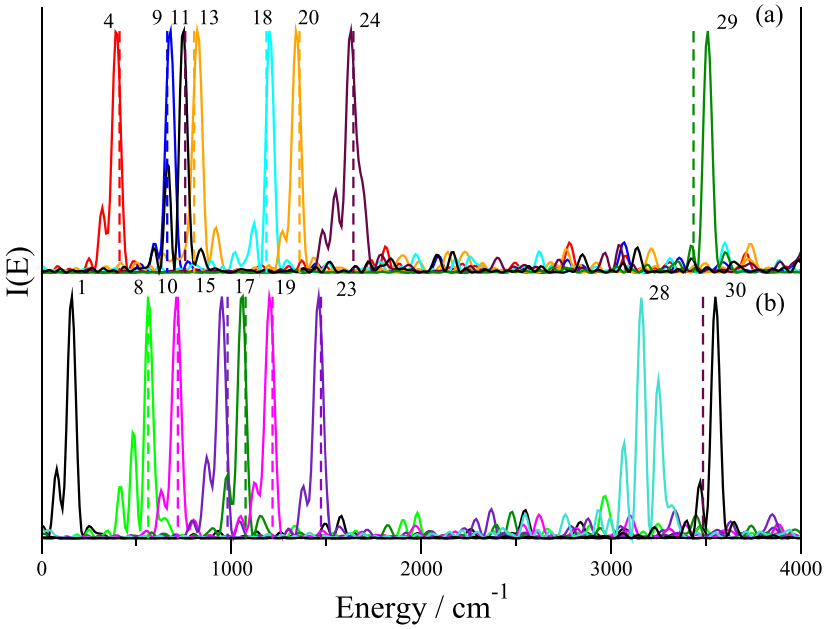

FIG. 2. Panel (a) and (b): DC SCIVR fundamental energy levels respect to the ZPE of the 17-dimensional subspace. Each fundamental excitation was obtained by tailoring the reference state $|\chi\rangle$ according to Eq. (10). The labels of the energy levels are given according to the nomenclature reported in Table I. Vertical lines are centered at experimental values.

of B3LYP theory for anharmonicities. We can observe a good agreement between these two column values, suggesting that our computational setup is a good harmonic estimate for a possible accuracy and feasibility of the semiclassical simulations. When moving to the semiclassical dynamics results, each uracil fundamental frequency is evaluated by tailoring the reference state of the semiclassical integrand according to the MC SCIVR approach of Eq. (10). Given the above mentioned dimensional limits, uracil represents also a good benchmark for comparison between MC SCIVR and DC SCIVR. The full-dimensional space is partitioned employing the Hessian matrix method.[69] A threshold value equal to $\varepsilon = 4 \times 10^{-7}$ leads to a 17-dimensional subspace, and all other subspaces are monodimensional. In Table I, the computed MC SCIVR and DC SCIVR energy levels are reported. In the fifth and sixth columns of the same table, the GVPT2 and CVPT2 estimates are slightly more accurate than the semiclassical ones probably because of the higher level of *ab initio* theory employed.

TABLE II. Vibrational fundamental excitations of the oxocytosine isomer. Values are reported in cm^{-1}. The first column reports the label of the excitation, and the second and the third ones report the experimental values taken from Table III of Ref. 23 measured in Ne and Ar matrices, respectively. Fourth column is for VPT2 energy levels.[34] Fifth column is for Harmonic results using a B3LYP/aug-cc-pvdz level of theory, and the last column reports the energy levels calculated with DC SCIVR at the same level of theory. The MAEs of DC SCIVR with respect to both experiments is reported into the last row, where the value in parentheses is for the comparison with the experimental values in the Ar matrix.

Label	Exp/Ne	Exp/Ar	VPT2	Harmonic	DC SCIVR
8				546	520
9	571	575		578	560
10	614	614		643	630
11	717	716		726	710
12	749	747		768	750
13	767	747		771	760
14	784	818		786	780
15				921	890
16				964	940
17				984	980
18		1088		1082	1060
19	1103	1090	1106	1120	1110
20	1198	1196	1193	1208	1190
21	1237	1244	1227	1258	1220
22	1340	1337	1335	1351	1320
23	1423	1423	1408	1441	1420
24	1475	1475	1472	1497	1480
25	1540	1539	1521	1566	1550
26	1602	1598	1588	1629	1590
27	1659	1656	1647	1685	1660
28	1730	1733	1736	1755	1750
29			3037	3202	3140
30			3092	3227	3120
31	3457	3441	3460	3596	3510
32	3474	3472	3477	3611	3510
33	3575	3564	3610	3741	3580
MAE			10(13)	38(41)	13(18)

MC SCIVR and DC SCIVR values are very similar and close to the experimental values as well. The MAE obtained by comparison with the experimental results is, respectively, 22 and 20 wavenumbers, which is pretty much the accuracy of other semiclassical simulations. Actually, also the MAE for the harmonic approximation at the same level of theory is quite similar. We think this good accuracy is accidental because higher level *ab initio* harmonic estimates (see Table I of Ref. 30) provide frequencies which are systematically higher than B3LYP ones. Going back to the semiclassical results, we think that the DC SCIVR MAE is slightly smaller than the more accurate MC SCIVR one either because of a compensation of errors or because MC SCIVR has been pushed too close to the dimensional limit of the method. These results confirm that the main advantage of the DC SCIVR approach is given by its portability and applicability to higher-dimensional molecules by retaining a comparable accuracy to that one of the MC SCIVR.

Interestingly, MC SCIVR is able to recover some experimental aspects which are very hard to be reproduced, such as the exchange in energy levels between modes 7 and 8. Namely, by employing normal mode or GVPT2 normal mode analysis (see, respectively, columns three and four in Table I) or VPT2 approaches (columns five and six of the same table), mode 7 results lower in energy than mode 8, while according to the experimental assignment mode,

TABLE III. The same as Table II but for the hydroxycytosine isomer.

Label	Exp/Ne	Exp/Ar	Harmonic	DC SCIVR
8	553	557	560	540
9	525	520	563	540
10	600	601	602	570
11	711	710	721	730
12	784	751	790	790
13	796	781	794	770
14	809	807	818	830
15	948	955	986	970
16	948	955	993	970
17	989	980	997	980
18	1085	1083	1092	1090
19	1113	1110	1120	1100
20	1198	1196	1241	1220
21	1258	1257	1298	1270
22	1338	1333	1342	1330
23	1382	1379	1399	1380
24	1441	1439	1459	1440
25	1495	1496	1517	1490
26	1576	1575	1606	1570
27	1592	1589	1630	1610
28	1625	1622	1656	1640
29			3165	3080
30			3207	3090
31	3461	3446	3596	3480
32	3575	3564	3736	3600
33	3618	3592	3770	3670
MAE			41(36)	16(18)

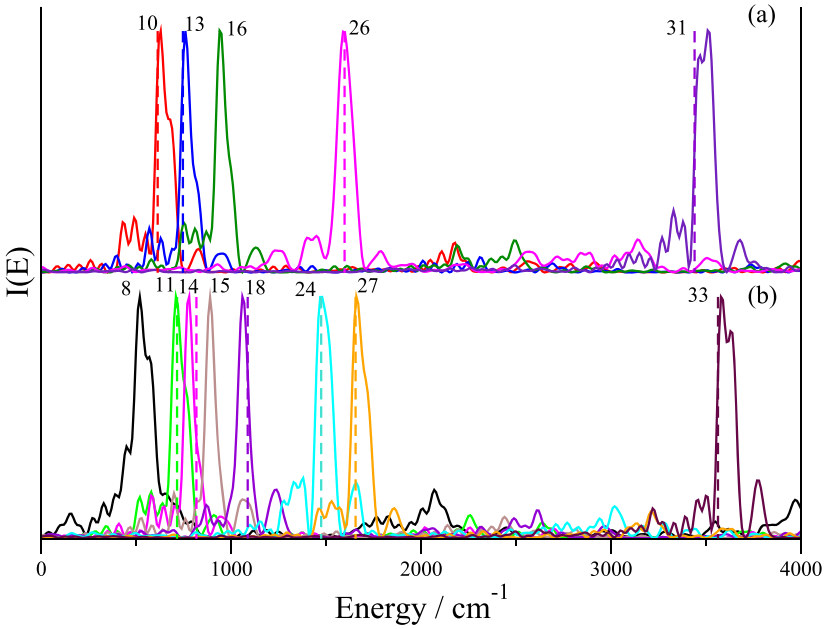

FIG. 3. Panel (a) and (b): The same as in Fig. 2 but for the 20-dimensional subspace of the oxocytosine isomer. Labels of the energy levels follow the nomenclature reported in Table II.

7 is slightly higher in energy than mode 8. MC SCIVR estimates agree with the experimental picture by locating mode 7 at 552 cm^{-1} and mode 8 at 544 cm^{-1}. We believe this peak inversion can be reproduced only by including the relevant anharmonic effects experienced by semiclassical trajectories far from the equilibrium configuration.

Figure 2 reports the spectrum of the 17-dimensional uracil subspace. Each fundamental excitation was obtained by tailoring the reference state $|\chi\rangle$ of the semiclassical integrand according to Eq. (10). Vertical dashed lines in the figures are centered at the available experimental levels. The peaks are labeled according to the first column of Table I, and the values are reported in the same table.

In summary, our results for uracil molecule show that all relevant anharmonic effects can be reproduced and that the full dimensional MC SCIVR calculations lead to accurate results. DC SCIVR

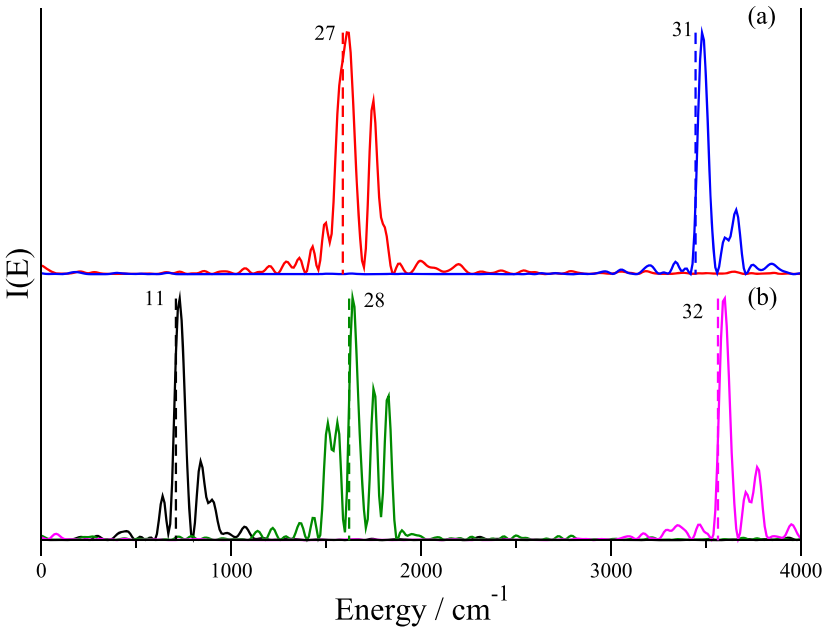

FIG. 4. Panel (a) and (b): The same as Fig. 3 but for the 11-dimensional hydroxycytosine isomer subspaces. Missing peaks in the figure are those with initial momenta set to 0. Labels of the energy levels follow the nomenclature reported in Table III.

results are comparable and enough accurate to pursue other nucleobase investigation, where full-dimensional MC SCIVR calculations cannot be longer afforded.

B. Cytosine

The next nucleobase we treat is cytosine, which is constituted by a pyrimidine ring. It differs from uracil by the presence of the amino group in place of an oxygen atom. This molecule is made by 13 atoms, resulting into 33 vibrational degrees of freedom. In addition to the increased dimension, the vibrational spectroscopic investigation gets complicated by the tautomerism between the oxocytosine and the hydroxycytosine. Both forms are spectroscopic relevant, given the very small minimum energy difference of few kilojoule/mole.[23] This difference is 2.1 kJ/mol at our level of DFT theory[21] and in favor of the oxo form. Thus, we perform molecular dynamics simulations of both tautomers by running two classical trajectories, one starting from the oxocytosine isomer and the other from the hydroxycytosine one [respectively, panels (b) and (c) of Fig. 1]. We observe that there is no isomeric change along the entire simulation time of 25 000 a.u. For this nucleobase, we observe the presence of a strong coupling between hindered rotations and other vibrational modes. For this reason, the initial kinetic energy of the first seven lowest frequency vibrational modes is set to zero since rotational contributions would jeopardize the numerical convergence of the spectra. This strategy is similar to what has been done in previous semiclassical calculations,[70,71] and it does not represent a bias since the initial kinetic energy of the trajectory is reduced by only 5% below the harmonic ZPE value, which is obviously in excess with respect to the actual ZPE. The columns of Tables II and III, labeled "Harmonic," report the harmonic frequencies of both isomers. We observe from the experimental values that most of the fundamentals of the two isomers are very similar in energy and this can be seen already at the harmonic level. A common strategy to discriminate between the two isomers is to look at the region around 3400–3700 wavenumbers. In that region are present for both isomers the symmetric and asymmetric NH_2 stretches, around 3500–3600 cm^{-1}. Additionally the oxo form shows a peak around 3450 cm^{-1} corresponding to the NH stretching, which is missing in hydroxycitosine spectrum that instead shows a signal above 3600 cm^{-1}, due to the OH stretching. This trend is well reproduced by the DC SCIVR results that can effectively discriminate between the two tautomers. Tables II and III report the DC SCIVR computed energy levels and show the comparison with available experimental data and other theoretical results. The full dimensional vibrational space is divided into subspaces using a threshold value of $\varepsilon = 7 \times 10^{-7}$ for Hessian elements. This choice generates for the oxocytosine one 20 dimensional subspace, leaving all the others monodimensional. In the case of hydroxycytosine, the threshold is fixed at $\varepsilon = 2 \times 10^{-6}$ and the full-dimensional space is fragmented into one 11-dimensional, one 7-dimensional, and all remaining monodimensional. We observe that for both isomers, the agreement with the experiment is very strict, giving a MAE around 18 cm^{-1} for both systems, even if there are some frequencies which are occasionally quite off the mark. Harmonic estimates show a double MAE deviation.

Figure 3 shows DC SCIVR spectra for a 20-dimensional subspace of the oxocytosine isomer, while Fig. 4 reports the computed spectra for a 11-dimensional subspaces of the hydroxycytosine isomer. We observe that the number of peaks in Figs. 3 and 4 is less than the subspace-dimension since such subspace contains one low index mode (1–7). These figures show neat spectroscopic signals with several overtones reproduced, especially Fig. 4.

So far we have looked at the lowest dimensional couple of nucleobases made by a pyrimidine ring. Our results agree with

TABLE IV. Table for the vibrational frequencies of thymine. First column reports the label of the excitation. Columns two and three are for gas phase spectra of the isolated[15] and Ar-tagged thymine.[18,19] Column four reports the harmonic estimate at the level of B3LYP/aug-cc-pvdz, and the last column reports our computed DC SCIVR energy levels at the same level of theory. MAE is reported on the last row in comparison with both series of experiments. Values are reported in cm^{-1}.

Mode	Gas	Gas/Ar	Harmonic	DC SCIVR
1			130	110
2			162	150
3			168	150
4		280	281	290
5			302	310
6		391	394	380
7		394	407	390
8	462	455	461	460
9		540	546	530
10	541	545	583	570
11			603	600
12	689	662	700	670
13		727	735	710
14	755	754	757	750
15	767	764	779	770
16	804	800	802	780
17	885	890	911	900
18	963	959	962	950
19		1005	1018	1000
20	1031	1046	1060	1040
21	1078	1087	1153	1120
22	1178	1184	1196	1170
23		1221	1232	1190
24		1357	1367	1380
25		1395	1397	1360
26	1393	1389	1408	1390
27	1409	1405	1419	1380
28		1437	1448	1430
29		1451	1473	1460
30	1463	1472	1497	1460
31	1668	1684	1700	1690
32	1725	1712	1741	1750
33	1772	1768	1784	1760
34		2855	3038	2955
35	2941	2940	3097	2970
36	2984	2971	3121	2990
37	3076	2992	3202	3080
38	3437	3434	3588	3470
39	3484	3485	3632	3530
MAE			48(38)	17(21)

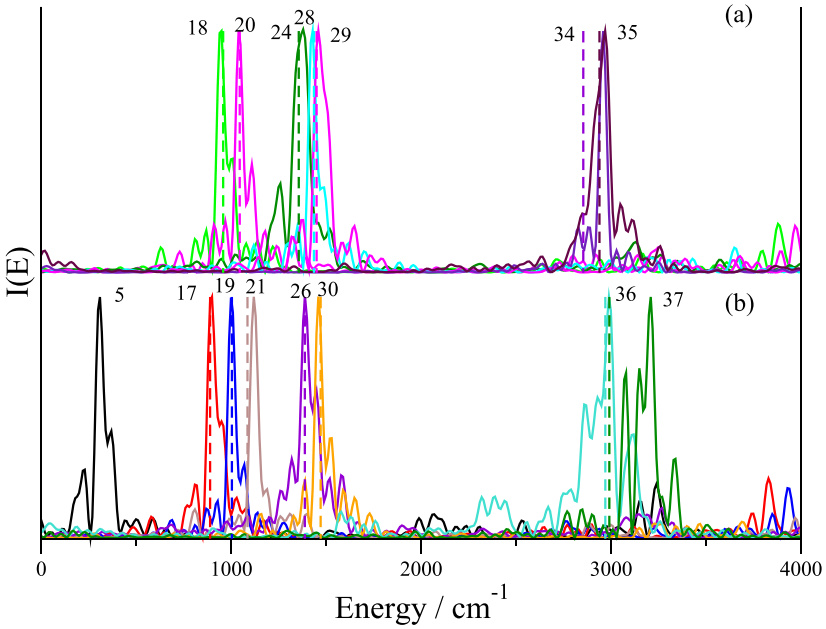

FIG. 5. Panel (a) and (b): The same as in Fig. 2 but for the 18-dimension subspace of thymine. Labels of the energy levels follow the nomenclature reported in Table IV.

experiments with an average deviation of 15–20 wavenumbers, in line with previous semiclassical simulations, and are comparable with other theoretical methods. In Sec. III C, we look at thymine, the last pyrimidine-based nucleobase.

C. Thymine

Thymine is the highest dimension pyrimidine nucleobase. The structure is similar to uracil, where one hydrogen atom is substituted by a methyl group. Because of their resemblance, thymine and uracil show some similarities in biological processes and both link to adenine. The molecule is made of 15 atoms leading to 39 vibrational degrees of freedom. We calculate the vibrational spectra using classical trajectories starting from the global minimum structure, as we have done for uracil in Sec. III B, since the hydroxy tautomer minimum is about 44 kJ/mol less stable,[20] a value similar to the one between uracil tautomers (45 kJ/mol).[21] This is expected given their structural similarity. The geometry of the molecule at

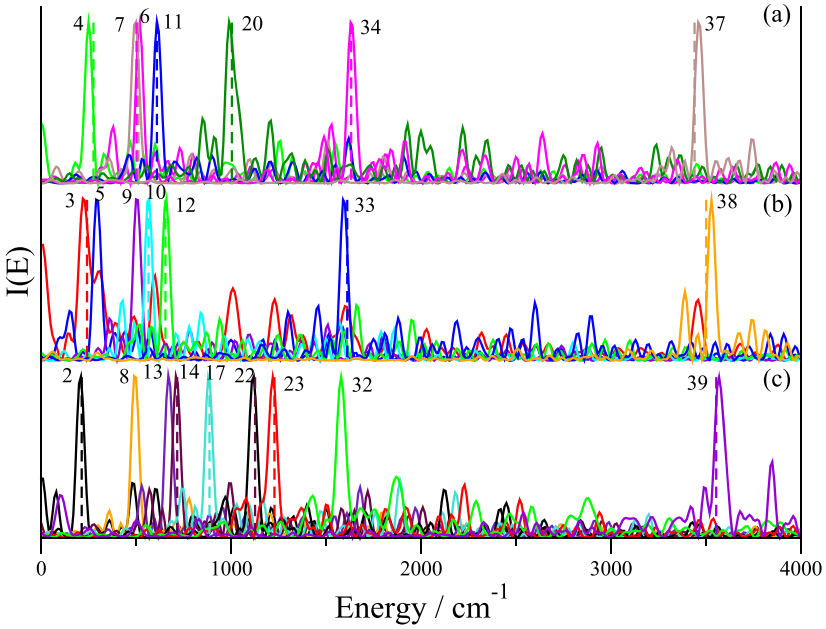

FIG. 6. Panel (a)–(c): The same as in Fig. 2 but for the 23-dimensional subspace of adenine. Frequency labels follow the nomenclature reported in Table IV.

equilibrium configuration is reported in Panel (d) of Fig. 1. The Hessian partitioning method leads to a 18-dimensional subspace, a seven-dimensional one and all the others are mono-dimensional, when employing a threshold parameter $\varepsilon = 8 \times 10^{-7}$. DC SCIVR frequencies are calculated for each subspace.

Table IV reports the vibrational frequencies at different levels of calculation and compares them with the experimental ones. In the table, DC SCIVR values are compared with the experimental energies of gas phase isolated and Ar-tagged thymine. We also report modes 1–3 even if no experimental data are available at the best of our knowledge. From the table, we observe that the MAE of DC SCIVR estimates is equal to 17 and 21 wavenumbers in comparison with gas phase and Ar-tagged levels, respectively, which is less than half the harmonic estimates. These values are in line once again with the typical semiclassical accuracy and are comparable with what was found from previous pyrimidine bases.

Figure 5 shows the computed DC SCIVR 18 dimension spectra, and one can note how the thymine fingerprint region around 1000 wavenumbers is very crowded. However, by using Eq. (10), we are able to selectively report each peak, as shown in Fig. 5. In this way, it is possible to resolve and attribute peaks which would be otherwise hardly distinguishable, since they are within few wavenumbers in energy, as in the case of modes 18–20 and 28–30.

D. Adenine

Now, we move our attention to the remaining nucleobase, adenine. Adenine is a nucleobase made of a purine ring, composed of a pyrimidine condensed with an imidazole. The molecular structure at the equilibrium configuration is reported in panel (e) of Fig. 1, showing 15 atoms and resulting into 39 vibrational degrees of freedom, the same as thymine. In this study, we focus our attention only on the global minimum structure since previous works suggested that the other tautomers of the molecule are located around 30 kJ/mol above the global minimum, an amount of energy difference that is close to that of thymine and uracil, rather than cytosine.

The 39 vibrational degree of freedom space is separated into one 23-dimensional, two bidimensional, and all other are mono-dimensional subspaces, by employing a threshold parameter for the Hessian method equal to $\varepsilon = 9 \times 10^{-7}$. Figure 6 reports the spectrum of the highest dimensional subspace for this molecule.

Once again, our reference state choice according to Eq. (10) helps us to resolve different peaks which are very close in energy, like modes 37–39 or 32–34. Figure 6 presents several overtone peaks, in comparison with previous ones. We believe this is due to the presence of the NH_2 hindered rotation, which gets easily vibrational excited during the dynamics. Nevertheless, the full width at half maximum (FWHM), which denotes the peak definition, is quite small. For a more detailed comparison with experiments, Table V reports the computed energy levels and shows the comparison with available theoretical and experimental results, measured in the isolated gas phase and Ar matrix. The average deviation of the DC SCIVR results from both experiments is quite small, only 16 and 14 wavenumbers, almost three times more accurate than harmonic estimates. This accuracy is comparable with previous system investigations and with PT2 estimates. Overall, the DC SCIVR spectrum of adenine is further proof of the invariance of accuracy of our approach with the increase in nucleobase dimensionality. We believe that the investigation of this purine nucleobase strengthens previous considerations about pyrimidine-based nucleobases, in terms of spectroscopic quality and accuracy of the energy levels, when compared with the experiments.

TABLE V. Vibrational frequencies for adenine. First column reports the label of the excitation. Columns two and three are for gas phase spectra of isolated[15] and Ar-tagged adenine.[16] Column four reports the Perturbation Theory (PT2) energy levels.[33] Column five report the harmonic estimate at the level of B3LYP/aug-cc-pvdz, and the last column reports our computed DC SCIVR energy levels at the same level of theory. MAE is reported on the last row, where the value in parentheses refers to the experiment in the Ar matrix. Values are reported in cm^{-1}.

Mode	Gas	Gas/Ar	PT2	Harmonic	DC SCIVR
1	162		139	168	160
2		214	181	210	210
3	244	242	209	267	220
4	270	276	276	279	250
5			299	302	290
6	506	503	491	512	520
7	515	513	516	529	500
8			518	531	490
9			529	542	500
10	563	566	570	583	570
11	600	610	610	617	610
12	650	655	655	668	660
13			677	688	670
14		717	717	723	710
15	801	802	811	815	800
16	847	848	846	858	840
17		887	885	896	880
18	926	927	931	942	930
19	957	958	969	979	960
20		1005	1018	1015	990
21	1065	1061	1054	1076	1050
22	1126	1127	1125	1141	1120
23	1234	1229	1230	1242	1220
24		1246	1243	1263	1250
25	1280	1290	1291	1332	1300
26	1326	1328	1325	1356	1330
27	1346	1345	1338	1365	1330
28		1389	1376	1414	1390
29	1415	1419	1406	1432	1400
30	1468	1474	1466	1498	1485
31		1482	1481	1512	1500
32			1577	1607	1580
33		1612	1591	1635	1590
34	1625	1633	1616	1660	1630
35		3041	3049	3172	3070
36	3061	3057	3102	3245	3160
37	3434	3441	3432	3588	3460
38	3501	3502	3497	3641	3530
39	3552	3555	3539	3727	3570
MAE			10(9)	42(38)	16(14)

IV. SUMMARY AND CONCLUSIONS

In this paper, we have presented a semiclassical investigation of the vibrational features of uracil, cytosine, thymine, and adenine nucleobases. The investigation on uracil has shown that MC SCIVR energy levels are on average around 20 wavenumbers away from experimental levels, a typical value for semiclassical spectroscopic calculations. Then, the DC SCIVR method leads to values very close to the full-dimensional MC SCIVR ones, proving its reliability for the calculation of similar systems. Moving to higher dimensional nucleobases, DC SCIVR energy levels of cytosine retain their typical accuracy with respect to the experimental results. Despite the presence of more than one comparable minimum, the method still reproduces the spectra of different isomers retaining the standard accuracy of semiclassical simulations. Then, we focus on the last pyrimidine molecule, thymine, which is the highest dimensional nucleobase of this type. We have found that DC SCIVR retains its accuracy despite the increased dimensionality. Similar considerations hold for the adenine case, where the MAE is around 15 wavenumbers and comparable with PT2 estimates. Overall, we always find the accuracy of the DC SCIVR method to be comparable to other state of the art theoretical spectroscopy methods.

These outcomes are promising for a future exploitation of the method. Since the accuracy is seemingly insensitive to the increase in the molecule dimensionality, we will exploit DC SCIVR also to study more complex systems, like nucleotides and nucleobase pairs. This will possibly pave the way toward the assessment of important structural features of nucleoacids that lead to the formation of secondary and tertiary structures.

ACKNOWLEDGMENTS

The authors thank Dr. Riccardo Conte for useful discussions. The authors acknowledge support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 647107—SEMICOMPLEX—ERC-2014-CoG). The authors acknowledge also CINECA (Italian Supercomputing Center) for providing high performance computational resources under the IscraB grant (QUASP). All authors thank Università degli Studi di Milano for further computational time at CINECA.

REFERENCES

[1] J. D. Watson and F. H. Crick, Nature **171**, 964 (1953).
[2] V. Gabelica, *Nucleic Acids in the Gas Phase* (Springer, 2016).
[3] R. Rein, M. Shibata, R. Garduno-Juarez, and T. Kieber-Emmons, *Structure and Dynamics: Nucleic Acids and Proteins* (Adenine Press, New York, 1983), pp. 269–288.
[4] V. Danilov and G. Kventsel, *Electronic Representations in the Point Mutation Theory* (Naukova Dumka, Kiev, 1971).
[5] C. Colominas, F. J. Luque, and M. Orozco, J. Am. Chem. Soc. **118**, 6811 (1996).
[6] R. D. Wells, Trends Biochem. Sci. **32**, 271 (2007).
[7] M.-P. Gaigeot and M. Sprik, J. Phys. Chem. B **107**, 10344 (2003).
[8] M.-P. Gaigeot and M. Sprik, J. Phys. Chem. B **108**, 7458 (2004).
[9] P. López-Tarifa, M.-P. Gaigeot, R. Vuilleumier, I. Tavernelli, M. Alcamí, F. Martín, M.-A. Hervé du Penhoat, and M.-F. Politis, Angew. Chem. **125**, 3242 (2013).
[10] T. Gustavsson, Á. Bányász, E. Lazzarotto, D. Markovitsi, G. Scalmani, M. J. Frisch, V. Barone, and R. Improta, J. Am. Chem. Soc. **128**, 607 (2006).
[11] M. Baer, D. Marx, and G. Mathias, Angew. Chem., Int. Ed. **49**, 7346 (2010).
[12] G. Mathias, S. D. Ivanov, A. Witt, M. D. Baer, and D. Marx, J. Chem. Theory Comput. **8**, 224 (2011).
[13] O. Asvany, P. Kumar, B. Redlich, I. Hegemann, S. Schlemmer, and D. Marx, Science **309**, 1219 (2005).
[14] A. M. Rijs and J. Oomens, *Gas-Phase IR Spectroscopy and Structure of Biological Molecules* (Springer, 2015), Vol. 364.
[15] P. Colarusso, K. Zhang, B. Guo, and P. F. Bernath, Chem. Phys. Lett. **269**, 39 (1997).
[16] M. J. Nowak, L. Lapinski, J. S. Kwiatkowski, and J. Leszczyński, J. Phys. Chem. **100**, 3527 (1996).
[17] D. Picconi, F. J. A. Ferrer, R. Improta, A. Lami, and F. Santoro, Faraday Discuss. **163**, 223 (2013).
[18] M. Graindourze, J. Smets, T. Zeegers-Huyskens, and G. Maes, J. Mol. Struct. **222**, 345 (1990).
[19] A. Leś, L. Adamowicz, M. J. Nowak, and L. Lapinski, Spectrochim. Acta, Part A **48**, 1385 (1992).
[20] J. Rejnek, M. Hanus, M. Kabeláč, F. Ryjáček, and P. Hobza, Phys. Chem. Chem. Phys. **7**, 2006 (2005).
[21] S. Tian, C. Zhang, Z. Zhang, X. Chen, and K. Xu, Chem. Phys. **242**, 217 (1999).
[22] M. Y. Choi and R. E. Miller, J. Phys. Chem. A **111**, 2475 (2007).
[23] J. S. Kwiatkowski and J. Leszczyński, J. Phys. Chem. **100**, 941 (1996).
[24] E. Nir, C. Janzen, P. Imhof, K. Kleinermanns, and M. De Vries, J. Chem. Phys. **115**, 4604 (2001).
[25] M. Y. Choi and R. E. Miller, J. Am. Chem. Soc. **128**, 7320 (2006).
[26] M. Nowak, L. Lapinski, and J. Fulara, Spectrochim. Acta, Part A **45**, 229 (1989).
[27] M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel, and M. Elhanine, J. Phys. Chem. A **106**, 5088 (2002).
[28] M. Mons, F. Piuzzi, I. Dimicoli, L. Gorb, and J. Leszczynski, J. Phys. Chem. A **110**, 10921 (2006).
[29] C. M. Marian, J. Phys. Chem. A **111**, 1545 (2007).
[30] C. Puzzarini, M. Biczysko, and V. Barone, J. Chem. Theory Comput. **7**, 3702 (2011).
[31] S. V. Krasnoshchekov, N. Vogt, and N. F. Stepanov, J. Phys. Chem. A **119**, 6723 (2015).
[32] P. S. Thomas, T. Carrington, Jr., J. Agarwal, and H. F. Schaefer III, J. Chem. Phys. **149**, 064108 (2018).
[33] M. Biczysko, P. Panek, and V. Barone, Chem. Phys. Lett. **475**, 105 (2009).
[34] R. Wałęsa, T. Kupka, and M. A. Broda, Struct. Chem. **26**, 1083 (2015).
[35] W. H. Miller and T. F. George, J. Chem. Phys. **56**, 5637 (1972).
[36] W. H. Miller, Proc. Natl. Acad. Sci. U. S. A. **102**, 6660 (2005).
[37] W. H. Miller, J. Chem. Phys. **53**, 1949 (1970).
[38] E. Heller, J. Chem. Phys. **95**, 9431 (1991).
[39] K. G. Kay, Annu. Rev. Phys. Chem. **56**, 255 (2005).
[40] D. J. Tannor, *Introduction to Quantum Mechanics* (University Science Books, 2007).
[41] J. Shao and N. Makri, J. Phys. Chem. A **103**, 7753 (1999).
[42] E. Pollak, "The semiclassical initial value series representation of the quantum propagator," in *Quantum Dynamics of Complex Molecular Systems*, edited by D. A. Micha and I. Burghardt (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007), pp. 259–271.
[43] S. Zhang and E. Pollak, J. Chem. Theory Comput. **1**, 345 (2005).
[44] D. V. Shalashilin and M. S. Child, J. Chem. Phys. **115**, 5367 (2001).
[45] D. V. Shalashilin and M. S. Child, Chem. Phys. **304**, 103 (2004).
[46] G. Tao, Theor. Chem. Acc. **133**, 1448 (2014).
[47] J. Liu, Int. J. Quantum Chem. **115**, 657 (2015).
[48] M. Church, S. V. Antipov, and N. Ananth, J. Chem. Phys. **146**, 234104 (2017).
[49] Z. Ma and D. Coker, J. Chem. Phys. **128**, 244108 (2008).
[50] J. Tatchen and E. Pollak, J. Chem. Phys. **130**, 041103 (2009).
[51] M. Wehrle, M. Šulc, and J. Vaníček, J. Chem. Phys. **140**, 244114 (2014).
[52] M. Wehrle, S. Oberli, and J. Vaníček, J. Phys. Chem. A **119**, 5685 (2015).
[53] R. Conte, A. Aspuru-Guzik, and M. Ceotto, J. Phys. Chem. Lett. **4**, 3407 (2013).

[54]M. Ceotto, S. Atahan, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **130**, 234113 (2009).
[55]M. Ceotto, S. Atahan, S. Shim, G. F. Tantardini, and A. Aspuru-Guzik, Phys. Chem. Chem. Phys. **11**, 3861 (2009).
[56]M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **147**, 164110 (2017).
[57]M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **148**, 114107 (2018).
[58]M. Buchholz, F. Grossmann, and M. Ceotto, J. Chem. Phys. **144**, 094102 (2016).
[59]T. Begušić, A. Patoz, M. Šulc, and J. Vaníček, Chem. Phys. **515**, 152 (2018).
[60]T. Begušić, J. Roulet, and J. Vaníček, J. Chem. Phys. **149**, 244115 (2018).
[61]J. Vaníček, CHIMIA Int. J. Chem. **71**, 283 (2017).
[62]M. Wehrle, M. Šulc, and J. Vaníček, CHIMIA Int. J. Chem. **65**, 334 (2011).
[63]E. Zambrano, M. Šulc, and J. Vaníček, J. Chem. Phys. **139**, 054109 (2013).
[64]M. Šulc and J. Vaníček, Mol. Phys. **110**, 945 (2012).
[65]M. Buchholz, E. Fallacara, F. Gottwald, M. Ceotto, F. Grossmann, and S. D. Ivanov, Chem. Phys. **515**, 231 (2018).
[66]M. Micciarelli, R. Conte, J. Suarez, and M. Ceotto, J. Chem. Phys. **149**, 064115 (2018).
[67]M. Micciarelli, G. Fabio, R. Conte, and M. Ceotto, J. Chem. Phys. **150**, 184113 (2019).
[68]M. Ceotto, S. Valleau, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **134**, 234103 (2011).
[69]M. Ceotto, G. Di Liberto, and R. Conte, Phys. Rev. Lett. **119**, 010401 (2017).
[70]G. Di Liberto, R. Conte, and M. Ceotto, J. Chem. Phys. **148**, 014307 (2018).
[71]G. Di Liberto, R. Conte, and M. Ceotto, J. Chem. Phys. **148**, 104302 (2018).
[72]F. Gabas, G. Di Liberto, R. Conte, and M. Ceotto, Chem. Sci. **9**, 7894 (2018).
[73]F. Gabas, R. Conte, and M. Ceotto, J. Chem. Theory Comput. **13**, 2378 (2017).
[74]C. Plützer, I. Hünig, K. Kleinermanns, E. Nir, and M. S. de Vries, Chem. Phys. Chem. **4**, 838 (2003).
[75]A. Abo-Riziq, L. Grace, E. Nir, M. Kabelac, P. Hobza, and M. S. de Vries, Proc. Natl. Acad. Sci. U. S. A. **102**, 20 (2005).
[76]E. Nir, C. Janzen, P. Imhof, K. Kleinermanns, and M. De Vries, Phys. Chem. Chem. Phys. **4**, 732 (2002).
[77]T. Fornaro, D. Burini, M. Biczysko, and V. Barone, J. Phys. Chem. A **119**, 4224 (2015).
[78]E. Nir, C. Plützer, K. Kleinermanns, and M. De Vries, Eur. Phys. J. D **20**, 317 (2002).
[79]R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, 1965).
[80]J. H. Van Vleck, Proc. Natl. Acad. Sci. U. S. A. **14**, 178 (1928).
[81]E. J. Heller, J. Chem. Phys. **75**, 2923 (1981).
[82]E. J. Heller, Acc. Chem. Res. **14**, 368 (1981).
[83]M. F. Herman and E. Kluk, Chem. Phys. **91**, 27 (1984).
[84]K. G. Kay, Chem. Phys. **322**, 3 (2006).
[85]F. Grossmann and A. L. Xavier, Phys. Lett. A **243**, 243 (1998).
[86]Y. Zhuang, M. R. Siebert, W. L. Hase, K. G. Kay, and M. Ceotto, J. Chem. Theory Comput. **9**, 54 (2012).
[87]M. Ceotto, Y. Zhuang, and W. L. Hase, J. Chem. Phys. **138**, 054116 (2013).
[88]K. G. Kay, J. Chem. Phys. **100**, 4432 (1994).
[89]K. G. Kay, J. Chem. Phys. **100**, 4377 (1994).
[90]K. G. Kay, J. Chem. Phys. **101**, 2250 (1994).
[91]H. Wang, D. E. Manolopoulos, and W. H. Miller, J. Chem. Phys. **115**, 6317 (2001).
[92]A. L. Kaledin and W. H. Miller, J. Chem. Phys. **118**, 7174 (2003).
[93]A. L. Kaledin and W. H. Miller, J. Chem. Phys. **119**, 3078 (2003).
[94]D. Tamascelli, F. S. Dambrosio, R. Conte, and M. Ceotto, J. Chem. Phys. **140**, 174109 (2014).
[95]M. Ceotto, D. Dell'Angelo, and G. F. Tantardini, J. Chem. Phys. **133**, 054701 (2010).
[96]G. Di Liberto and M. Ceotto, J. Chem. Phys. **145**, 144107 (2016).
[97]M. Ceotto, G. F. Tantardini, and A. Aspuru-Guzik, J. Chem. Phys. **135**, 214108 (2011).
[98]K. Hinsen and G. R. Kneller, Mol. Simul. **23**, 275 (2000).
[99]B. B. Harland and P.-N. Roy, J. Chem. Phys. **118**, 4791 (2003).
[100]M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, Comput. Phys. Commun. **181**, 1477 (2010).
[101]A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[102]X. Ma, G. Di Liberto, R. Conte, W. L. Hase, and M. Ceotto, J. Chem. Phys. **149**, 164113 (2018).
[103]M. Szczesniak, M. Nowak, H. Rostkowska, K. Szczepaniak, W. B. Person, and D. Shugar, J. Am. Chem. Soc. **105**, 5969 (1983).
[104]S. Chin, I. Scott, K. Szczepani, and W. B. Person, J. Am. Chem. Soc. **106**, 3415 (1984).