On the ground state calculation of a many-body system using a self-consistent basis and quasi-Monte Carlo. An application to water hexamer.

Ionuţ Georgescu,[1, a)] Svetlana Jitomirskaya,[2] and Vladimir A. Mandelshtam[1]
[1)] *Chemistry Department, University of California, Irvine, CA 92697, USA*
[2)] *Department of Mathematics, University of California, Irvine, CA 92697, USA*

Given a quantum many-body system, the Self-Consistent Phonons (SCP) method provides an optimal harmonic approximation by minimizing the free energy. In particular, the SCP estimate for the vibrational ground state (zero temperature) appears to be surprisingly accurate. We explore the possibility of going beyond the SCP approximation by considering the system Hamiltonian evaluated in the harmonic eigenbasis of the SCP Hamiltonian. It appears that the SCP ground state is already uncoupled to all singly- and doubly-excited basis functions. So, in order to improve the SCP result at least triply-excited states must be included, which then reduces the error in the ground state estimate substantially. For a multidimensional system two numerical challenges arise, namely, evaluation of the potential energy matrix elements in the harmonic basis, and handling and diagonalizing the resulting Hamiltonian matrix, whose size grows rapidly with the dimensionality of the system. Using the example of water hexamer we demonstrate that such calculation is feasible, i.e. constructing and diagonalizing the Hamiltonian matrix in a triply-excited SCP basis, without any additional assumptions or approximations. Our results indicate particularly that the ground state energy differences between different isomers (e.g., cage and prism) of water hexamer are already quite accurate within the SCP approximation.

I. INTRODUCTION

The Self Consistent Phonons (SCP) approach[1–3] arose in an effort to include anharmonic effects in the nuclear dynamics of condensed phase systems. Given a quantum many-body system at equilibrium, localized in an energy minimum, SCP maps it variationally to a reference harmonic system by optimizing the free energy. Recently SCP has been implemented successfully for finite systems, for example, for estimating the absorption spectrum of aromatic hydrocarbons[4], or for computing the ground states of very large Lennard-Jones clusters[5,6]. Most recently Brown *et al*[7] have shown how to circumvent the key numerical bottleneck of the SCP method, namely, the numerical evaluation of Gaussian averages of the potential energy surface (PES) and its derivatives without making any assumptions about the PES or approximations, thus expanding the applicability range of SCP to *ab-initio* potentials.

While SCP is successful in providing a substantial improvement to the standard normal mode analysis, i.e., that based on the Taylor expansion of the PES at its minimum, an apparent drawback of the method seems to be the lack of a natural procedure to systematically improve it further, or even provide a reliable assessment of its accuracy, besides the obvious cases where the exact results are available. (Note though ref. 3, where a diagrammatic technique was suggested to include higher order corrections into the SCP estimate of the free energy. Unfortunately, this development had so far no follow-up in the literature.) In this paper we address both issues. We consider a harmonic basis emerging from the eigenfunctions of the reference (i.e., SCP) Hamiltonian, which from now on we will refer to as the **SCP basis**. The vibrational ground state is then estimated by the lowest eigenvalue of the physical Hamiltonian evaluated in this basis. It may occur that there is nothing new in this strategy, which is implemented in, for example, MULTIMODE algorithm of Carter *et al*[8]. However, the devil is in details. Firstly, as will be shown below, the SCP basis, though also harmonic, turns out to be much more efficient than the standard normal mode harmonic basis. Furthermore, as far as we aware, it has always been thought that for a many-dimensional system an accurate evaluation of the potential energy matrix elements in a harmonic (or any other delocalized) basis is a very difficult problem, which for a general PES could hardly be handled by standard means, e.g., by using quadrature; neither did it seem feasible to compute the corresponding integrals by Monte Carlo (MC). Indeed, the size of a quadrature grid usually scales exponentially with the dimensionality of the system, while the $S^{-1/2}$ convergence rate of a standard MC integration is too slow with respect to the size, S, of the random sequence.

Note also that accurate numerical evaluation of the potential energy matrix elements in a harmonic basis is also a key problem in perturbative approaches. One way to avoid high-dimensional integrals is to replace the true PES by special representations such that matrix elements can be evaluated analytically. In the n-mode representation[9], the potential energy is expanded in normal mode coordinates and partitioned into n terms, with the first term being a sum of single-mode potentials, the second term, a sum of two-mode potentials, etc., up to the sum of n-mode potentials. This representation is currently most "famously" implemented in

[a)] Electronic mail: ionutg@gmail.com

MULTIMODE[8,9], a VSCF/VCI code that computes the eigenvalues and eigenfunctions of a rovibrational Hamiltonian. Another widely used representation is *potfit*, an expansion in a product basis of single-particle potentials developed for MCTDH[10–12]. Thus, even though all matrix elements in question can be evaluated analytically, the problem of accurate numerical evaluation of multidimensional integrals is replaced by that of obtaining a relatively compact and accurate parametrization of the PES, and what is worst, this introduces an element of ambiguity, as the accuracy of such an approximation becomes hard to control.

The key breakthrough of ref. 7 was to demonstrate that the multivariate Gaussian integrals involving a general PES can be computed accurately and efficiently, avoiding any additional assumptions and approximations, with the help of a quasi-MC approach, in which the integration error decays almost as S^{-1}, rather than $S^{-1/2}$. Since the potential energy matrix elements in SCP basis have a structure similar to that of the SCP method itself, i.e. Gaussian averages of relatively smooth functions, the said quasi-MC integration can be extended to calculations of the corresponding matrix elements. (Note though, that integrals involving higher-order harmonic excitations are expected to converge slower.)

We will first use an example of a 1D Morse oscillator to confirm that the SCP basis is indeed superior to the standard normal mode basis. Then the method will be applied to a very challenging problem, the ground state calculation of isomers of water hexamer. This system has received considerable attention in the past, both experimentally and theoretically. Since our goal here is primarily to demonstrate the numerical technique, for a relatively comprehensive overview of the literature on water hexamer the interested reader is referred to a recent paper by Babin and Paesani[13], which explores the effects of different water models on the equilibrium properties of this system. Briefly, water hexamer corresponds to the smallest water cluster whose low-energy isomers show three-dimensional structures reminiscent of the hydrogen bond network in bulk water. Two of these isomers, the cage and the prism, are believed to be the main candidates for the low-temperature equilibrium structure. Most empirical and *ab-initio*-based PES's predict the prism to be the global energy minimum. However, the quantum effects, especially for the light hydrogen atoms, may easily change the energy ordering. CCSD(T) calculations with extrapolation to the complete basis set limit and zero-point energies (ZPEs) calculated within the harmonic approximation predict the prism to be the most stable isomer, followed closely by the cage isomer[14]. By contrast MP2 calculations with harmonic ZPE corrections predicted the chair isomer to be the most stable structure, followed by the cage, book and prism isomers[15]. Wang *et al*[16] considered the WHBB PES of Bowman and co-workers[17], in which the 3-body interactions were fit at MP2 level. Based on both Diffusion Monte Carlo (DMC) and Replica-Exchange Path Integral Molecular Dynamics (RE-PIMD) calculations the authors of ref. 16 conclude that the WHBB cage and prism have approximately the same ground state energies. The SCP result of Brown *et al*[7] disagreed with this prediction, finding the prism ground state about $100\,\mathrm{cm}^{-1}$ below the cage. However, the DMC results[16] may not be well converged with respect to the number of random walkers, while getting a truly converged DMC energies for a system as complex as water hexamer would probably require excessively too many random walkers. To add to the controversy, the most recent publication[13] of two of the authors from ref. 16 find the prism to be the most stable isomer on the same WHBB PES, yet, the estimated energy difference is still much smaller than that from the SCP estimate. For the HBB2-pol PES[18], which is identical to WHBB, except for the three-body interactions that were fit at CCSD(T) level, ref. 13 predicts the cage isomer to be the most stable structure. Also note that the RE-PIMD free energy differences could not be converged below 30K (or perhaps even higher) and that the T=0K result (the ground state) was estimated by harmonic extrapolation from mid-temperature range, a procedure that is at least questionable. In any case, at this point we cannot make a definite verdict on the actual reason for the discrepancy between the SCP and RE-PIMD predictions, but we hope the results reported in this paper would shed some light on this issue. For the present study, we find it instructive to consider three PES's: q-TIP4P/F[19], WHBB[17] and HBB2-pol[18].

To this end it is worth noting that the rotational spectroscopy experiments find the cage to be more stable[20], but unfortunately at this point a direct comparison of experimental findings with calculations cannot be made very meaningful. With four-body and higher contributions responsible for 5% of the total energy, the error of truncating the many body expansion at three-body terms (involved in the construction of either WHBB of HBB2-pol PES's) is more than $500\,\mathrm{cm}^{-1}$, much larger than the energy difference in question.

II. METHODS

A. Self Consistent Phonons

Consider an N-body system described by the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2}\nabla^{\mathrm{T}}\mathbf{M}^{-1}\nabla + V(\mathbf{r}), \tag{1}$$

where $V(\mathbf{r})$ defines the PES, $\mathbf{M} = \mathrm{diag}\{m_i\}$ is the mass matrix, and $\mathbf{r} \in \mathbb{R}^{3N}$, the coordinate vector.

It is assumed that the system is trapped in a local minimum and is at equilibrium at temperature T. We then define a local harmonic approximation

$$\hat{H}_h(T) = -\frac{\hbar^2}{2}\nabla^{\mathrm{T}}\mathbf{M}^{-1}\nabla + \frac{1}{2}(\mathbf{r}-\mathbf{q})^{\mathrm{T}}\mathbf{K}\,(\mathbf{r}-\mathbf{q}) + V_0\ , \tag{2}$$

where V_0 is the minimum of the effective harmonic potential, $\mathbf{q}$ its center, and $\mathbf{K}$, the Hessian. The Gibbs-Bogolyubov inequality

$$F \leq F_{\text{trial}} := F_h + \langle V \rangle_h - \langle V_h \rangle_h \tag{3}$$

provides an upper limit F_{trial} for the free energy F of the system when the ensemble average $\langle \cdot \rangle_h = \text{Tr}[e^{-\beta \hat{H}_h} \cdot]$ is more conveniently taken with respect to the approximation $\hat{H}_h$ instead of the true $\hat{H}$. F_h denotes the free-energy of the harmonic system and V_h is the harmonic potential. The ensemble average with respect to the reference harmonic potential can be written as

$$\begin{aligned} \langle f(\hat{\mathbf{r}}) \rangle_h &:= \text{Tr}[e^{-\beta \hat{H}_h} f(\hat{\mathbf{r}})] \\ &= \|2\pi \mathbf{D}\|^{-1/2} \int d\mathbf{x}\, e^{-\frac{1}{2}\mathbf{x}^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{x}} f(\mathbf{q}+\mathbf{x}), \end{aligned} \tag{4}$$

with the displacement-displacement correlation matrix

$$\mathbf{D} = \mathbf{M}^{-\frac{1}{2}}\, d\,(\mathbf{\Omega})\, \mathbf{M}^{-\frac{1}{2}} \tag{5}$$

and the auxiliary function

$$d(\omega) := (\hbar/2)\omega^{-1} \coth\left(\beta \hbar \omega / 2\right). \tag{6}$$

The frequency matrix $\mathbf{\Omega}$ is defined via the mass-scaled effective Hessian:

$$\mathbf{\Omega}^2 = \tilde{\mathbf{K}} = \mathbf{M}^{-\frac{1}{2}} \mathbf{K} \mathbf{M}^{-\frac{1}{2}}. \tag{7}$$

The trial free-energy to be minimized takes thus the form

$$F_{\text{trial}} = k_{\mathrm{B}} T \log \left\| 2 \sinh \left(\frac{\hbar \beta}{2} \tilde{\mathbf{K}}^{1/2} \right) \right\| + \langle V \rangle_h - \frac{1}{2} \text{Tr}[\mathbf{K}\mathbf{D}]. \tag{8}$$

Minimizing F_{trial} with respect to $\mathbf{q}$ and $\mathbf{K}$ leads to the following system of coupled nonlinear equations[2]:

$$\langle \nabla V \rangle_h = 0 \tag{9a}$$

$$\left\langle \nabla \nabla^{\mathrm{T}} V \right\rangle_h = \mathbf{K}. \tag{9b}$$

Taking the $T \to 0$ limit in (3) we obtain a variational principle for the ground-state energy E_0:

$$E_0 \leq E_0^{\text{SCP}} := \langle 0 | \hat{H}_h | 0 \rangle + \langle 0 | V - V_h | 0 \rangle \tag{10a}$$

$$:= \langle 0 | \hat{H} | 0 \rangle, \tag{10b}$$

where $|0\rangle$ is the ground state of $\hat{H}_h$. SCP becomes equivalent to the Ritz variational principle because the SCP equations (9) minimize E_0^{SCP} with respect to $\mathbf{q}$ and $\mathbf{K}$. $|0\rangle$ provides thus the SCP approximation for the ground state of $\hat{H}$. With $\delta V = V - V_h$ one obtains further

$$E_0^{\text{SCP}} = \frac{1}{2} \sum_m \hbar \omega_m + \langle 0 | \delta V | 0 \rangle. \tag{11}$$

We will show in Section II D that in the SCP basis approach, only sufficiently large basis can further optimize E_0^{SCP}.

B. Quasi Monte Carlo

Unless the potential energy V has a special form, as exploited in Ref. 6, the expectation values (4) required by Eq. (9) cannot be evaluated analytically. A possible Monte-Carlo approach would generate a sequence $\mathbf{x}_s \in \mathbb{R}^{3N}$, $s = 1 \ldots S$ obeying the multivariate normal distribution

$$p_{\mathbf{D}}(\mathbf{x}) = \|2\pi \mathbf{D}\|^{-1/2} e^{-\frac{1}{2}\mathbf{x}^{\mathrm{T}} \mathbf{D}^{-1} \mathbf{x}} \tag{12}$$

and the average in Eq. (4) becomes

$$\langle f \rangle_h = \frac{1}{S} \sum_{s=1}^{S} f(\mathbf{x}_s). \tag{13}$$

$p_{\mathbf{D}}(\mathbf{x})$ could then be sampled by transforming a sequence uniformly distributed on $[0,1]^{3N}$ using the inverse cumulative distribution function (CDF) for the normal distribution. Random or pseudo-random sequences show a tendency of point clustering, leaving areas of the integration domain uncovered, which causes the typical $\mathcal{O}(S^{-1/2})$ convergence. At this rate, the approach would require millions of potential energy evaluations, which is prohibitive, especially for an *ab-initio* PES.

Ref. 7 adapted a quasi-MC approach that reduced the required number of potential energy evaluations by several orders of magnitude, thus, making it applicable to *ab-initio* PES. Quasi-MC uses quasi-random, or, more precisely, deterministic low-discrepancy sequences, that have much better local distributions that, in particular, do not suffer from clustering. Several low-discrepancy sequences exist[21–25], but the Sobol sequence has shown so far better performance[26,27] and is thus used here. For a d-dimensional integral, quasi-MC integration achieves a convergence rate of $\mathcal{O}(S^{-1}(\log S)^d)$, sometimes even $\mathcal{O}(S^{-1})$ for certain types of integrands[28].

C. Solving the SCP Equations

In solving Eqs. (9) for $\mathbf{q}$ and $\mathbf{K}$, Calvo *et al*[4] used a mixed approach, a Newton-Raphson-like iteration for $\mathbf{q}$ and self-consistent iteration for $\mathbf{K}$. This method worked well for aromatic hydrocarbon molecules, such as pyrene, achieving convergence in as few as 10 iterations, but it did not work for water hexamer[7], because of highly ill-conditioned Hessian matrix $\mathbf{K}$ (and, therefore, ill-conditioned $\mathbf{D}$). Thus for such cases a much more conservative iterative procedure has to be employed, a procedure that would generally require many more iterations. For even larger water clusters very low frequency modes may emerge, thus leading to an extremely stiff non-linear system that may be problematic to solve in a self-consistent fashion. Interestingly, as shown in ref. 6, in the $T \to 0$ limit, i.e. when the density matrix is defined solely by the ground state, the SCP

method becomes equivalent to the Variational Gaussian Wavepacket method[29–31], in which case the self-consistent equations (9) can be solved in a deterministic fashion by propagating a system of coupled ordinary differential equations in "imaginary time" τ. In the limit $\tau \to \infty$ this procedure always converges to a stationary solution of Eq. (9).

D. Properties of the SCP ground state

As established previously[7], the ground state, $|0\rangle$, of the SCP harmonic Hamiltonian $\hat{H}_h$ coincides with the SCP approximation for the ground state of the true Hamiltonian $\hat{H}$, which may be obtained by variationally optimizing (in the Ritz sense) a Gaussian wavepacket. We will now consider other eigen states of $\hat{H}_h$ and investigate the properties of $\hat{H}$ evaluated in this basis.

We split the system Hamiltonian $\hat{H}$ into the SCP Hamiltonian and the difference potential δV:

$$\hat{H} = \hat{H}_h + V - V_h = \hat{H}_h + \delta V. \tag{14}$$

Let U be the unitary matrix that diagonalizes the mass scaled Hessian (7)

$$\mathrm{diag}(\omega_k^2) = U^{\mathrm{T}} \tilde{\mathbf{K}} U. \tag{15}$$

Following ref. 7 we restrict the Hamiltonian to the $(3N-6)$-dimensional subspace, the Eckart subspace, that excludes the translations of the whole cluster and its rotations around the center of mass. For the converged SCP solution the excluded six degrees of freedom correspond to six zero eigenvalues of $\tilde{\mathbf{K}}$. The normal mode transform in the Eckart subspace is then given by

$$\mathbf{r} = \mathbf{M}^{-\frac{1}{2}} U \boldsymbol{\alpha}^{-1} \mathbf{y} + \mathbf{q}, \tag{16}$$

with the frequency-scaled normal mode coordinate vector $\mathbf{y} = (y_1, ..., y_{3N-6})^{\mathrm{T}}$, diagonal matrix $\boldsymbol{\alpha} := \mathrm{diag}(\alpha_1, \ldots, \alpha_{3N-6})$ and $\alpha_n = \sqrt{\omega_n/\hbar}$.

The eigen-basis of $\hat{H}_h$ is then a direct product of the $(3N-6)$ harmonic eigen-states:

$$\Psi_{v_1,...,v_{N-6}}(\mathbf{r}) = \|\boldsymbol{\alpha}\|^{\frac{1}{2}} \|\mathbf{M}\|^{\frac{1}{4}} \prod_{n=1}^{3N-6} \phi_{v_n}(\alpha_n U_n^{\mathrm{T}} \mathbf{M}^{\frac{1}{2}} (\mathbf{r}-\mathbf{q})), \tag{17}$$

where U_n is the n-th column of U, that is, the unit vector of the n-th eigenmode and $\phi_v(y)$ are the harmonic eigenfunctions,

$$\phi_v(y) = \frac{1}{\sqrt{2^v \sqrt{\pi} v!}} \, e^{-y^2/2} H_v(y). \tag{18}$$

From now on we will use a more convenient notation to define a harmonic eigenstate,

$$|n_1^{v_1}, n_2^{v_2} ...\rangle, \tag{19}$$

which includes the indices, $n_1,\ n_2, ...$, of only those modes that are excited and $v_1 > 0,\ v_2 > 0, ...$ then indicate their excitations. For the ground state we will use $|0\rangle$. To further simplify the notations we will drop the excitation indices for singly excited modes, e.g., $|n\rangle \equiv |n^1\rangle$, $|n,m\rangle \equiv |n^1, m^1\rangle$, etc.

Let us now partition the basis according to the maximum number of quanta carried by a basis function. We thus define the singly-excited basis $\mathcal{B}_h^{(1)}$, which includes the ground state $|0\rangle$ and all the singly-excited states $|n\rangle$. The doubly-excited basis $\mathcal{B}_h^{(2)}$ includes $\mathcal{B}_h^{(1)}$ and all the doubly-excited states of the two kinds $|n,m\rangle$ and $|n^2\rangle$, and so on.

Diagonalizing $\hat{H}$ in the $\mathcal{B}_h^{(1)}$ basis can either improve the accuracy of the SCP ground-state $|0\rangle$ or leave it unchanged; diagonalizing it in $\mathcal{B}_h^{(2)}$ can either improve it further or leave it unchanged, and so on.

Using Eqs. (17), (18) and the SCP condition (9) it follows that

$$\begin{aligned} \langle 0|\delta V|k\rangle &= \frac{1}{\alpha_k \sqrt{2}} \langle \partial_{u_k} \delta V \rangle && (20) \\ &= \frac{1}{\alpha_k \sqrt{2}} U_k^{\mathrm{T}} \mathbf{M}^{-\frac{1}{2}} \langle \nabla V \rangle = 0. && (21) \end{aligned}$$

Furthermore, for $n \neq k$:

$$\begin{aligned} \langle n|\delta V|k\rangle &= \frac{1}{2\alpha_n \alpha_k} \langle \partial_{u_n} \partial_{u_k} \delta V \rangle \\ &= \frac{1}{2\alpha_n \alpha_k} U_n^{\mathrm{T}} \mathbf{M}^{-\frac{1}{2}} \left[\langle \nabla \nabla^{\mathrm{T}} V \rangle - \mathbf{K} \right] \mathbf{M}^{-\frac{1}{2}} U_k \\ &= 0, \end{aligned} \tag{22}$$

and

$$\begin{aligned} \langle n|\delta V|n\rangle &= \langle 0|\delta V|0\rangle + \frac{1}{2\alpha_n^2} \langle \partial_{u_n}^2 \delta V \rangle \\ &= \langle 0|\delta V|0\rangle . \end{aligned} \tag{23}$$

In other words, $\hat{H}$ is diagonal in the singly-excited basis:

$$\langle n|\hat{H}|k\rangle = \delta_{nk}(E_0 + \hbar\omega_n). \tag{24}$$

For the matrix elements involving the doubly-excited states we obtain additionally:

$$\langle 0|\delta V|k^2\rangle = \frac{1}{\sqrt{2}}[\langle k|\delta V|k\rangle - \langle 0|\delta V|0\rangle] = 0 \tag{25a}$$

$$\langle 0|\delta V|k,l\rangle = \langle k|\delta V|l\rangle = 0. \tag{25b}$$

To summarize the above findings, the "perturbation" δV does not couple the ground state to either the singly- or doubly-excited states. Thus, diagonalizing the Hamiltonian with respect to either $\mathcal{B}_h^{(1)}$ or $\mathcal{B}_h^{(2)}$ leaves the SCP ground state $|0\rangle$ or its energy E_0 unchanged. Therefore, in order to optimize the SCP ground state the Hamiltonian must be diagonalized in at least triply-excited basis $\mathcal{B}_h^{(3)}$.

E. Computation of the potential energy matrix elements

Our goal is to evaluate the following integrals:

$$\langle n_1^{v_1}, n_2^{v_2}, \ldots | \delta V | {m_1}^{w_1}, {m_2}^{w_2}, \ldots \rangle = \int d^{3N-6}\mathbf{y}\ p(\mathbf{y}) \times \frac{H_{n_1}(y_{n_1})}{\sqrt{2^{v_1} v_1!}} \frac{H_{n_2}(y_{n_2})}{\sqrt{2^{v_2} v_2!}} \cdots \frac{H_{m_1}(y_{m_1})}{\sqrt{2^{w_1} w_1!}} \frac{H_{m_2}(y_{m_2})}{\sqrt{2^{w_2} w_2!}} \cdots \times \delta V(\mathbf{M}^{-\frac{1}{2}} U \boldsymbol{\alpha}^{-1} \mathbf{y}) \quad (26)$$

where we defined the normal distribution

$$p(\mathbf{y}) := \pi^{-\frac{3N-6}{2}} e^{-\mathbf{y}^{\mathrm{T}}\mathbf{y}}. \quad (27)$$

The bad news is that the number of matrix elements to be computed proliferates rapidly with the system size, e.g., as $\mathcal{O}(N^6)$ if triple excitations are included, and for the water hexamer with $(3N-6) = 48$ vibrational degrees of freedom, this requires evaluation of 216,829,900 individual integrals. Quite clearly, for a large system, such as water hexamer, unless due to some miracle a significant fraction of the matrix elements could be neglected, we cannot hope for including higher order excitations. In fact, we will rather demonstrate that omitting essentially any significant fraction of the matrix elements will lead to substantial loss of accuracy. However, the good news is that the triple excitation basis may already provide a highly accurate result.

Another good news is that, due to the special structure of the integrand in Eq. (26) all such integrals can be estimated simultaneously within the quasi-MC approach, Eq. (13) of Section II B, using the energies evaluated on a single grid of size S resulted from a quasi-random Sobol sequence $\{\mathbf{y}_s\}$ $(s = 1, ..., S)$ distributed normally according to $p(\mathbf{y}_s)$. Partial integration in Eq. (26) yields several equivalent expressions involving lower degree polynomials and derivatives of δV (see ref. 7 for details). The variance of the quasi-MC integrand is thus reduced and shorter quasi-random sequences will suffice. Taking the average of all these equivalent forms further reduces the variance. Computation of potential energy gradients is usually inexpensive, and most *ab-initio* packages can provide them with no extra cost. Yet, higher derivatives come at a considerable additional cost which has to be weighed against the savings in the total number of energy evaluations. Our tests indicate that just using the gradients reduces the required length of the quasi-random sequence by at least an order of magnitude.

Tracking the expressions for so many matrix elements, especially when taking the advantage of the partial integration, may be very tedious and confusing. In the Appendix we provide a Mathematica notebook that produces such expressions by first finding all equivalent forms for the required matrix elements via partial integration followed by their averaging. The expressions are then exported as C source code.

High-order matrix elements involve high-order polynomials that can reveal high-order correlations in the quasi-random sequence. Some matrix elements may show oscillatory behavior with periods given by powers of two. One solution is to use sequences of sufficiently large size, which is exactly $S = 2^L$. Another solution is to use **scrambled quasi-random sequences**, e.g., that proposed by Matoušek[32] and available in `MATLAB` as `'MatousekAffineOwen'`. Other scrambling techniques can reduce the variance of the quasi-MC even further[33].

F. Potential energy surfaces for H_2O

q-TIP4P/F[19] is an empirical, four-site water model parametrized to reproduce the liquid structure, diffusion coefficient, and IR absorption frequencies in quantum, rather than classical simulations[19]. It is flexible and anharmonic in the intramolecular OH-stretch vibrations. In the next section we show that SCP is highly accurate for an isolated OH-stretch, and by including triple excitations we can recover the ground-state within less than $0.2\,\mathrm{cm}^{-1}$.

The WHBB[16] and the HBB2-pol[18] employ a "stratified" many-body expansion of the interaction energy:

$$V(1, \ldots, M) = \sum_i V^{1b}(i) + \sum_{i<j} V^{2b}(i,j) + \sum_{i<j<k} V^{3b}(i,j,k) + \ldots + V^{Mb}(i, \ldots, M). \quad (28)$$

where M stands for the total number of water molecules in the system. The one-body (or 1b) potential $V^{1b}(i) = V(i) - V_{eq}(i)$ accounts for deformations of the individual molecules and is simply the PES of an isolated molecule. The 2b interaction energy is defined as

$$V^{2b}(1,2) = V(1,2) - V(1) - V(2), \quad (29)$$

the 3b interactions as

$$V^{3b}(1,2,3) = V(1,2,3) - \sum_{i<j}^{3} V^{2b}(i,j) + \sum_{i}^{3} V(i), \quad (30)$$

and so on. The expansion was found to converge quickly for long-range weak interactions, in particular, also for water clusters[34]. For example, 4b and higher-order terms represent less than 5% of the total energy of the water hexamer and also 21-mer water cluster[35].

In WHBB, the 1b potential is a spectroscopically accurate monomer potential, and the 2b and 3b potentials are permutationally invariant fits using tens of thousands of CCSD(T)/aug-cc-pVTZ and MP2/aug-cc-pVTZ *ab-initio* data points. Semiempirical 4b and higher-order contributions are also included, as described in the TTM3-F potentials[36].

HBB2-pol tries to both improve the accuracy and reduce the computational cost of the 3b terms in WHBB, which is the most time consuming part of the energy evaluation in the latter. CCSD(T) level electronic energies

are now used as reference not only for the 2b fits, but for the 3b ones as well. On the other hand, HBB2-pol exploits the fact that most of the 3b interaction energy in water arises from induction and makes the ansatz:

$$V^{3b}_{\rm HBB2\text{-}pol} = V^{3b}_{\rm ind} + V^{3b}_{\rm poly}. \quad (31)$$

The induction scheme of TTM4-F is used for the induction term $V^{3b}_{\rm ind}$, and $V^{3b}_{\rm poly}$ is a short range correction, which is fit in terms of low order polynomials. This leads to large savings in computational time, since in WHBB the full V^{3b} term was fit to high, fifth- and sixth-order, polynomials[37].

III. NUMERICAL EXPERIMENTS

We examine here the improvement in the ground-state energy in ever increasing basis of the SCP Hamiltonian, $\mathcal{B}^{(1)}_h$, $\mathcal{B}^{(2)}_h$,$\mathcal{B}^{(3)}_h$, etc., in which the system Hamiltonian is diagonalized. We begin with a simple but instructive case of a 1D Morse oscillator and then consider a very challenging case of water hexamer.

A. The 1D Morse oscillator

The Morse potential reads

$$U_{\rm Morse}(r) = D_e(1 - e^{-\alpha(r-r_0)})^2. \quad (32)$$

The corresponding Schrödinger equation is analytically soluble with eigenenergies given by

$$E_{\rm Morse}(v) = \hbar\omega_e(v+1/2) - \frac{[\hbar\omega_e(v+1/2)]^2}{4D_e}, \quad (33)$$

where $\omega_e = \alpha\sqrt{2D_e/m}$ and $0 \le v \le 2D_e/\hbar\omega_e - 1$. The level of anharmonicity is controlled by parameter α.

We solve the SCP equations to find the optimal harmonic approximation and construct the harmonic SCP basis. We then compute the ground state of the Hamiltonian matrix and analyze its accuracy as we increase the number of basis functions. We also consider the standard harmonic approximation (HA) emerging from the Taylor expansion of the potential at its minimum and the corresponding basis.

We first consider both the strongly quantum and strongly anharmonic Morse potential, $D_e = r_e = 1$ and $\alpha = 1/\sqrt{2}$, that supports only two bound states. Figure 1 shows the ground state energy estimate as a function of the basis size, comparing the SCP and HA bases; the inset shows the relative error with respect to the exact value of the ground state energy E_0. Basis size = 1 corresponds to the pure SCP result, which for this strongly anharmonic case is accurate to about 7 %. Basis size = 2 corresponds to the diagonalization of the Hamiltonian using singly-excited basis, and so on. The SCP energies are identical up to basis size = 3 (double excitations), thus

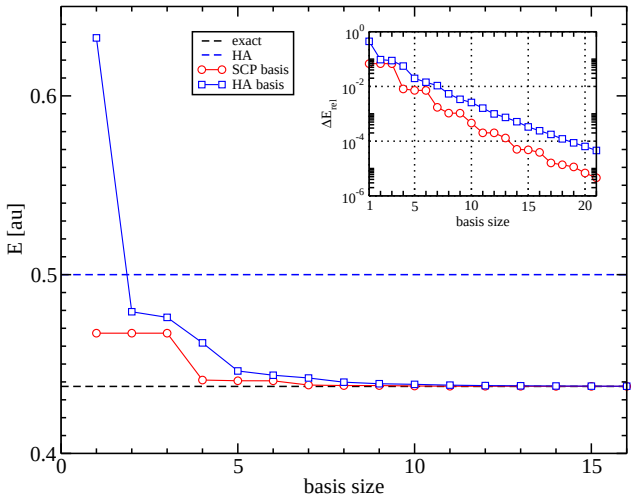

FIG. 1. Ground state energy of a strongly anharmonic Morse oscillator ($D_e = r_e = 1$ and $\alpha = 1/\sqrt{2}$) obtained from diagonalization of the Hamiltonian in the SCP basis (red) or the HA basis (blue), as a function of the basis size. Dashed black: exact E_0 (Eq. (33)); dashed blue: pure HA result $E_0^{\rm HA} = \hbar\omega_e/2$.

confirming the properties derived in Section II D: basis functions carrying at least triple excitations are required to further improve the SCP ground state. For the 1D case this translates to a minimum of four basis functions, which provide the accuracy of $\sim 1\%$. In order to obtain a comparable accuracy using the HA basis at least seven basis functions are required.

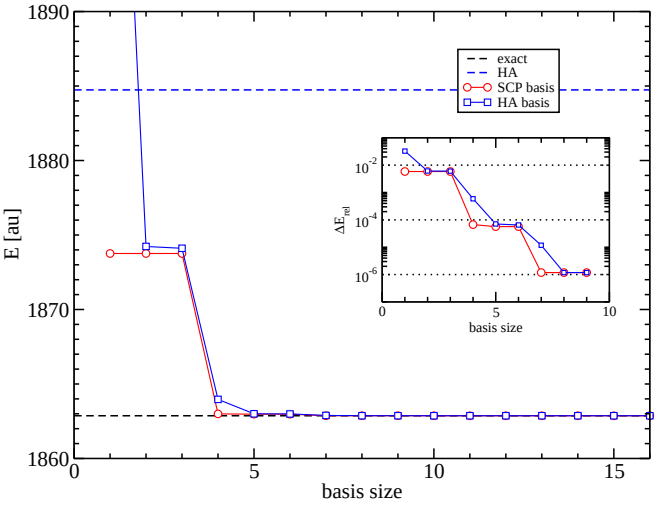

FIG. 2. Same as Figure (1), but for Morse parameters used to describe the OH-stretch in the q-TIP4P/F water model[19]: $D_e = 116.09$ kcal mol^{-1}, $r_0 = 0.9419$ Å and $\alpha = 2.287$ Å^{-1}.

Figure 2 illustrates a dominantly harmonic case, corresponding to the OH-stretch pair potential in the q-TIP4P/F water model[19], that is $D_e = 116.09$ kcal mol^{-1}, $r_0 = 0.9419$ Å and $\alpha = 2.287$ Å^{-1}. The HA estimates improve substantially, but the SCP result is consistently

much better. With 4 basis functions, $|E_0^{(\mathrm{SCP})} - E_0| \leq 0.13\,\mathrm{cm}^{-1}$, that is $\Delta E_{\mathrm{rel}} \leq 10^{-4}$.

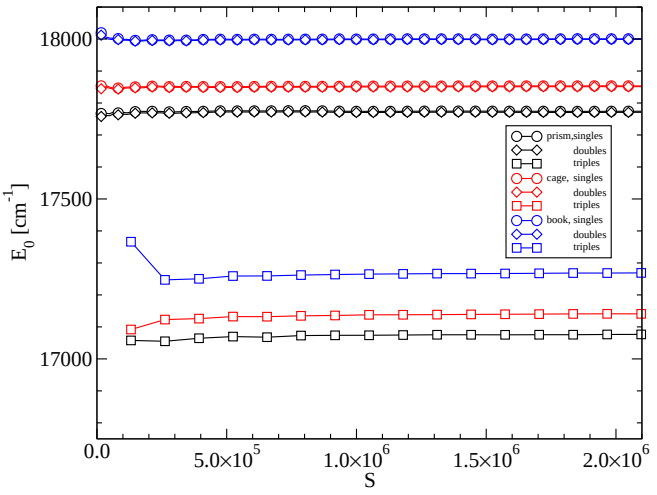

FIG. 3. Ground state energies of the HBB2-pol water hexamer for cage, prism and book isomers as a function of the length, S, of the quasi-random Sobol sequence.

B. The water hexamer

Here we compute the ground-state energies of the cage, prism and book isomers of water hexamer for three different PES's, the q-TIP4P/F[19], WHBB[16] and HBB2-pol[18]. Note that depending on the PES the energies of these isomers could be very close, and even obtaining the correct energy ordering may be nontrivial.

For each isomer the calculation is done in two stages. First, a relatively well-converged SCP result is obtained for the parameters of the SCP Hamiltonian, $\mathbf{K}$ and $\mathbf{q}$, essentially following the protocol of ref. 7. Note that here high accuracy is not required as the SCP parameters only define a basis in which the Hamiltonian is to be diagonalized. All the potential energy matrix elements are then evaluated in the SCP basis simultaneously according to the quasi-MC approach described in Section II E, and the lowest energy of the Hamiltonian matrix is computed. These values are denoted by $E_0^{(m)}$ according to the notation, $\mathcal{B}_h^{(m)}$, of the basis set. In addition, we define E_0^{SCP}, which is the value obtained by the SCP method. Fig. 3 shows $E_0^{(m)}$ for each of the three isomers of the HBB2-pol water hexamer as a function of S, the length of the quasi-random Sobol sequence used to compute the matrix elements. According to the properties of the SCP approach discussed in Section II D, the converged results should be identical with the SCP result using either $\mathcal{B}_h^{(1)}$ or $\mathcal{B}_h^{(2)}$ bases, i.e., the fully converged energies should satisfy:

$$E_0^{(2)} = E_0^{(1)} = E_0^{\mathrm{SCP}} \quad \text{(converged SCP)} \qquad (34)$$

However, if the numerical solution of the SCP equations (9) is insufficiently accurate, the $\mathcal{B}_h^{(2)}$ basis will show an improvement over the $\mathcal{B}_h^{(1)}$ basis, due to the variational nature of the method,

$$E_0^{(2)} \leq E_0^{(1)} \neq E_0^{\mathrm{SCP}} \quad \text{(non-converged SCP)} \qquad (35)$$

with the energy differences serving as a measure for convergence. In fact, the corresponding energy differences that we do observe are of the order of $1\,\mathrm{cm}^{-1}$.

With the $\mathcal{B}_h^{(3)}$ basis the matrix elements involve higher-order polynomials that are more difficult to converge. This explains why in this case a satisfactory convergence for $E_0^{(3)}$ values is achieved only using $S \sim 5 \times 10^5$ energy evaluations, which is by about an order of magnitude more than what is needed to obtain converged results for either of E_0^{SCP}, $E_0^{(1)}$ or $E_0^{(2)}$.

TABLE I. Ground state energies of the book and cage isomers relative to that of the prism computed by pure SCP, $\Delta E_0^{\mathrm{SCP}}$, and within $\mathcal{B}_h^{(3)}$ basis, $\Delta E_0^{(3)}$, for three different PES's.

	q-TIP4P/F		WHBB		HBB2-pol	
	$\Delta E_0^{\mathrm{SCP}}$	$\Delta E_0^{(3)}$	$\Delta E_0^{\mathrm{SCP}}$	$\Delta E_0^{(3)}$	$\Delta E_0^{\mathrm{SCP}}$	$\Delta E_0^{(3)}$
cage	−137.52	−141.2	107.91	102.21	79.26	64.3
book	190.73	190.5	570.38	563.22	226.62	192.4

q-TIP4P/F and WHBB energies converge in a similar fashion, with an improvement $\delta E_0 \sim -700\,\mathrm{cm}^{-1}$ when all triple excitations are included in the basis. More surprisingly though, in spite of the fairly large energy correction for all structures on all PES's, the *relative* energies of the isomers,

$$\Delta E_0^{(m)} := E_0^{(m)} - E_0^{(m)}(\mathrm{prism}),$$

are hardly affected. Those are shown in Table I for the cage-prism and book-prism and for the three different PES's. For each case we report the two most interesting values, $\Delta E_0^{\mathrm{SCP}}$ and $\Delta E_0^{(3)}$. In spite of the considerable computational effort involved in a $\mathcal{B}_h^{(3)}$ calculation (216,829,900 individual integrals have been converged) q-TIP4P/F and WHBB show no significant change, i.e. less than $8\,\mathrm{cm}^{-1}$! HBB2-pol shows a somewhat larger change, $\sim 15\,\mathrm{cm}^{-1}$ for cage-prism and $\sim 34\,\mathrm{cm}^{-1}$ for book-prism, but it is still very small when compared to the absolute energy changes, $\delta E_0 = E_0^{(3)} - E_0^{\mathrm{SCP}}$, due to the inclusion of triple excitations. This suggests that the higher order anharmonic corrections that are included with the triple excitations are simultaneous for different structures, i.e., they are weakly affected by the structural differences of different isomers.

Tables II and III try to break down the overall triple-excitation correction, δE_0, into its most important contributions. Table II lists absolute values and Table III, the corresponding relative significance with respect to

TABLE II. Energy lowering δE_0 [cm^{-1}] while gradually expanding $\mathcal{B}_h^{(2)}$ up to $\mathcal{B}_h^{(3)}$ by including the types of mode couplings indicated in the left most column. The following notations describing the couplings are used:

$\|n^3\rangle$:	all triple excitations within isolated modes.
s-l:	all triple excitations each combining a stretch with a libration mode.
s-l(local):	triple excitations coupling only stretch modes with libration modes that both have a large amplitude on the same molecule.
s-s:	all triple excitations each coupling any two stretch modes with each other.
l-l:	all triple excitations each coupling any two libration modes with each other.
any-s-l:	all triple excitations each coupling any mode with one stretch and one libration mode.

	q-TIP4P/F			WHBB			HBB2-pol		
$\mathcal{B}_h^{(2)}$	prism	cage	book	prism	cage	book	prism	cage	book
$\|n^3\rangle$	−53.53	−73.93	−66.87	−92.79	−73.06	−38.20	−108.38	−97.98	−81.65
$\|n^3\rangle$, s-l(local)	−341.12	−299.68	−206.18	−393.31	−410.74	−307.67	−343.42	−307.86	−294.63
$\|n^3\rangle$, s-l(local), s-s, l-l	−395.05	−338.49	−273.42	−443.49	−475.79	−402.08	−398.90	−369.40	−361.91
$\|n^3\rangle$, s-l	−545.01	−569.92	−565.52	−572.46	−565.43	−514.04	−575.71	−581.45	−584.83
$\|n^3\rangle$, s-l, s-s, l-l	−593.30	−603.76	−622.80	−617.91	−625.47	−598.18	−625.31	−635.04	−642.96
$\|n^3\rangle$, any-s-l,s-s, l-l	−667.50	−672.22	−692.32	−687.71	−694.02	−665.44	−695.18	−703.80	−714.46
$\mathcal{B}_h^{(3)}$	−707.96	−711.70	−732.67	−724.23	−730.72	−699.61	−731.35	−740.07	−751.05

TABLE III. Same as Table II but relative to $\delta E_0 = E_0^{(3)} - E_0^{\rm SCP}$.

	q-TIP4P/F			WHBB			HBB2-pol		
$\mathcal{B}_h^{(2)}$	prism	cage	book	prism	cage	book	prism	cage	book
$\|n^3\rangle$	7.56	10.39	9.13	12.81	10.00	5.46	14.82	13.24	10.87
$\|n^3\rangle$, s-l(local)	48.18	42.11	28.14	54.31	56.21	43.98	46.96	41.60	39.23
$\|n^3\rangle$, s-l(local), s-s, l-l	**55.80**	**47.56**	**37.32**	**61.24**	**65.11**	**57.47**	**54.54**	**49.91**	**48.19**
$\|n^3\rangle$, s-l	76.98	80.08	77.19	79.04	77.38	73.47	78.72	78.57	77.87
$\|n^3\rangle$, s-l, s-s, l-l	**83.80**	**84.83**	**85.00**	**85.32**	**85.60**	**85.50**	**85.50**	**85.81**	**85.61**
$\|n^3\rangle$, any-s-l, s-s, l-l	94.29	94.45	94.49	94.96	94.98	95.12	95.06	95.10	95.13
$\mathcal{B}_h^{(3)}$	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

the overall shift. Inclusion of the triple-excitation states of the type $|m^3\rangle$ were very successful in improving the ground-state energy for an isolated OH-stretch, as shown in Fig. 2, but for the hexamer they only contribute 5-15% of the total δE_0. The largest contributions come from the s-l couplings, i.e., the triple excitations combining stretch (i.e., the OH-stretch) and libration modes (i.e., frustrated rotation modes of water monomers) with no other modes involved. Consider a stretch SCP normal mode and a libration mode. If they both have a large component within the same molecule, then we label the corresponding coupling with s-l(local). We observe that local stretch-libration couplings alone can account for almost half of the energy lowering, but the contribution is not uniform across all structures and PES's. For book on the q-TIP4P/F surface it is only about 28%. Adding all s-s and l-l couplings does not make the contribution more uniform. However, going beyond local and considering all possible s-l couplings, including the s-s and l-l ones, yields 83-86% of the total lowering for all structures and all PES's. Also, allowing for couplings that combine a stretch, a libration and any other mode yields 95% of the overall δE_0. All other couplings, including those between bending modes and frustrated translations are than responsible for the remaining 5%.

We do not have a good explanation for why the ground state energy estimate is lowered by almost the same amount for all structures and PES's when the basis is enlarged to triple excitations. Investigations of other isomers with a radically different structure and symmetry, such as the ring hexamer, could perhaps break the rule.

The analysis presented in Tables II and III suggests that within a variational basis set approach any reduced dimensional treatment of water clusters, or any approximation that retains only certain "important" coupling terms and drops all the other "unimportant" terms, would introduce significant errors in the final result that would be too large, e.g., for a correct assessment of which structural isomer is energetically more stable. This speculation is of course correct unless a massive error cancellation takes place. Irrelevant to how appealing the latter idea is, but all our attempts to identify the groups of

modes and coupling terms, for which such cancellations can occur, failed so far.

IV. CONCLUSIONS AND CRITICAL REMARKS.

The SCP method is an efficient way to include the anharmonic effects in the estimation of the free energy of a large molecular system and, in particular, to estimate its vibrational ground state (zero temperature). In this paper we explored the possibility of further improvement of the SCP ground state by diagonalizing the system Hamiltonian in the SCP harmonic basis. The latter can be constructed from the optimal harmonic approximation that is a byproduct of the SCP method. We have demonstrated that the SCP basis is much more efficient than the standard harmonic basis that emerges from the normal mode analysis of the system PES at its minimum. For a multidimensional system, the main numerical challenge, the computation of the potential energy matrix elements in a harmonic basis, is here circumcised by quasi-MC method that reduces the otherwise very difficult numerical task to a practically tractable problem, which can even be solved for a general *ab-initio* PES without making any assumptions or approximations. The applicability and practicality of the method was demonstrated on a challenging problem, the ground state calculation of isomers of water hexamer.

While the ground state energy estimated by pure SCP is already quite accurate, its further improvement cannot be obtained within the SCP basis including only excitations up to the second order. The good news though is that including the triple excitations lead to a substantial improvement of the SCP result. For example, for a 1D Morse potential that mimics the OH-stretch potential in water the ground state energy is practically exact within the triply-excited basis. This turns out to be very important as for a multidimensional system the total basis size increases rapidly with the dimensionality, so that for a system as large as water hexamer the triply-excited basis essentially hits its practical limit. The lowering of the ground state energy due to the inclusion of triple excitations for this case was $\delta E_0 = E_0^{(3)} - E_0^{\mathrm{SCP}} \approx -700\,\mathrm{cm}^{-1}$, which turned out to be nearly the same across all structures and PES's, a fact which also kept the relative energies of the various isomers almost constant, with the notable exception of the HBB2-pol PES. While these results are certainly very encouraging, further numerical tests are required, in particular, investigations on structurally more different isomers, such as the ring hexamer, as well as different PES's could reveal whether the almost constant energy lowering is a special property of water systems, or a property of particular structures of water clusters, or an artifact of the method.

The very good agreement of the SCP energies with those from the triple-excitation basis suggest that SCP alone is very good at capturing energy differences, at least for water systems. At the same time, we cannot rule out the opposite possibility, i.e., that a localized (harmonic) basis is not appropriate in capturing the inherent curvature of some low frequency modes, thus not being able to correctly describe the structural differences in water systems, and, in addition, the apparent cancellation of systematic errors is an artifact of the method. This pessimistic scenario could then explain the disagreement of our results with those of refs. 13 and 16.

Finally, the approach presented here is in principle valid for computation of the excited vibrational states as well: one only needs to increase the harmonic basis by including sufficiently high excitations to obtain converged results for the corresponding excited states. However, for the present case of water hexamer further increase of the basis is unfeasible, while the triple excitations are insufficient. In addition, within the triply-excited basis we cannot even improve the SCP result for the fundamental frequencies[7], simply because excited states are not improved consistently with the improvement of the ground state, thus leading to worsening of the fundamental frequency estimates rather than their improvement.

ACKNOWLEDGEMENTS

We would like to thank Sandra Brown for useful discussions and for her tolerance during the discussions she witnessed. Joel Bowman and Yimin Wang are acknowledged for sending us the code of their WHBB PES, and Francesco Paesani and Volodymyr Babin, for the source code of their HBB2-pol PES. I.G. and S.J. were supported by the National Science Foundation (NSF) Grant No. DMS-1101578, and V.A.M., by NSF Grant No. CHE-1152845.

Appendix A: Expressions for the potential energy matrix elements

This appendix includes Mathematica code that will compute expressions for all matrix elements mentioned in this work. The notebook automatically performs partial integrations wherever possible in order to take advantage of the PES gradient. If partial integration can be performed with respect to more than one integration variable, then an average of all possible expressions is taken in order to reduce the variance. The notebook also generates the corresponding C source code.

```
Needs["SymbolicC`"];

Phi[y_, n_]:=1/Sqrt[2^n Factorial[n]] Pi^(-1/4) Exp[-y^2/2] HermiteH[n, y];
```

```
PartialIntegration[expr_, x_] :=
    If[ expr===(expr/.x->0),
        ,
        FullSimplify[D[FullSimplify[(expr-Coefficient[expr,x, 0])/(2 x)] f[x],x]
            + Coefficient[expr,x, 0] f[x] ]
    ]

MakeCArrays[expr_]:=
    (
        (expr/.xx_^2->sqr[xx]) /.
            {qm->q[[m]], qn->q[[n]], qp->q[[p]], qk->q[[k]], ql->q[[l]], qj->q[[j]] }
    ) /. {f'[_[[n_]]]->Vq[[n]], f[_[[n_]]]->V}

MakeCFunction[name_, args_, body_] := CFunction[
                    double, name, {int, #}&/@ args,
                    CReturn[CExpression[MakeCArrays[FullSimplify[Mean[body]]] ]]
                    ]

MakeFunction[braNames_, braStates_, ketNames_, ketStates_]:=(
    bra:= Pick[Transpose[{braNames, braStates}],#>0&/@braStates];
    ket:= Pick[Transpose[{ketNames, ketStates}],#>0&/@ketStates];

    funcName := StringJoin[
            (ToString[StringForm["````_", #1, #2 ] ] & @@@bra)
            ~Join~ {"f"} ~Join~
            (ToString[StringForm["_````", #1, #2 ] ] & @@@ket)
        ];

    funcArgs := Transpose[ bra ~Join~ ket][[1]] ;
    states :=Transpose[ bra ~Join~ ket][[2]] ;
    integrationVariables := Symbol[ToString[StringForm["q``", #]]]& /@ funcArgs;

    funcBody := PartialIntegration[
                    Times @@ ( Phi[#1, #2]/Phi[#1, 0]&
                                @@@ Transpose[{integrationVariables, states}]),
                    #]& /@ DeleteDuplicates[integrationVariables];

    {funcName, DeleteDuplicates[funcArgs], funcBody}
)


Excitations={   {0, 0, 0}, {1, 0, 0}, {2, 0, 0}, {3, 0, 0},
                {1, 1, 0}, {1, 2, 0}, {2, 1, 0}, {1, 1, 1} };

mnpklj=Join @@
    Table[
        Table[
            MakeFunction[{m, n, p}, Excitations[[mi]],
                            {k, l, j}, Excitations[[ki]] ],
            {ki,mi,Length[Excitations]} ],
        {mi, 1, Length[Excitations]} ];

mnpmlj=Join @@  Table[
                    Table[
                        MakeFunction[{m, n, p}, Excitations[[mi]],
                                        {m, l, j}, Excitations[[ki]] ],
                        {ki,mi,Length[Excitations]} ],
                    {mi, 2, Length[Excitations]} ];

mnpmnj=Join @@  Table[
                    Table[
                        MakeFunction[{m, n, p}, Excitations[[mi]],
                                        {m, n, j}, Excitations[[ki]] ],
                        {ki,mi,Length[Excitations]} ],
                    {mi, 5, Length[Excitations]} ];

mnpmnp=Join @@  Table[
                    Table[
                        MakeFunction[{m, n, p}, Excitations[[mi]],
```

```
                                {m, n, p}, Excitations[[ki]] ],
                        {ki,mi,Length[Excitations]} ],
                {mi, 8, Length[Excitations]} ];

ToCCodeString[MakeCFunction@@@Join[mnpklj, mnpmlj, mnpmnj, mnpmnp]]
```


[1]T. R. Koehler, "Theory of the self-consistent harmonic approximation with application to solid neon," Phys. Rev. Lett., **17**, 89–91 (1966).

[2]N. S. Gillis, N. R. Werthamer, and T. R. Koehler, "Properties of crystalline argon and neon in the self-consistent phonon approximation," Phys. Rev., **165**, 951–959 (1968).

[3]J. Cao and G. A. Voth, "Modeling physical systems by effective harmonic oscillators: The optimized quadratic approximation," The Journal of Chemical Physics, **102**, 3337–3348 (1995).

[4]F. Calvo, P. Parneix, and N.-T. Van-Oanh, "Finite-temperature infrared spectroscopy of polycyclic aromatic hydrocarbon molecules. ii. principal mode analysis and self-consistent phonons," The Journal of Chemical Physics, **133**, 074303 (2010).

[5]I. Georgescu, J. Deckman, L. J. Fredrickson, and V. A. Mandelshtam, "Thermal gaussian molecular dynamics for quantum dynamics simulations of many-body systems: Application to liquid para-hydrogen," The Journal of Chemical Physics, **134**, 174109 (2011).

[6]I. Georgescu and V. A. Mandelshtam, "Self-consistent phonons revisited. i. the role of thermal versus quantum fluctuations on structural transitions in large lennard-jones clusters," The Journal of Chemical Physics, **137**, 144106 (2012).

[7]S. E. Brown, I. Georgescu, and V. A. Mandelshtam, "Self-consistent phonons revisited. ii. a general and efficient method for computing free energies and vibrational spectra of molecules and clusters," The Journal of Chemical Physics, **138**, 044317 (2013).

[8]S. Carter, J. M. Bowman, and N. C. Handy, "Multimode calculations of rovibrational energies of c2h4 and c2d4," Molecular Physics, **110**, 775–781 (2012), http://www.tandfonline.com/doi/pdf/10.1080/00268976.2012.669504.

[9]S. Carter, S. J. Culik, and J. M. Bowman, "Vibrational self-consistent field method for many-mode systems: A new approach and application to the vibrations of co adsorbed on cu(100)," The Journal of Chemical Physics, **107**, 10458–10469 (1997).

[10]H.-D. Meyer, U. Manthe, and L. Cederbaum, "The multi-configurational time-dependent hartree approach," Chemical Physics Letters, **165**, 73 – 78 (1990), ISSN 0009-2614.

[11]M. H. Beck, A. Jäckle, G. A. Worth, and H. D. Meyer, "The multiconfiguration time-dependent hartree (mctdh) method: a highly efficient algorithm for propagating wavepackets," Physics Reports, **324**, 1 – 105 (2000), ISSN 0370-1573.

[12]H.-D. Meyer and G. A. Worth, "Quantum molecular dynamics: propagating wavepackets and density operators using the multiconfiguration time-dependent hartree method," Theoretical Chemistry Accounts, **109**, 251–267 (2003).

[13]V. Babin and F. Paesani, "The curious case of the water hexamer: Cage vs. prism," Chemical Physics Letters, **580**, 1 – 8 (2013), ISSN 0009-2614.

[14]D. M. Bates and G. S. Tschumper, "Ccsd(t) complete basis set limit relative energies for low-lying water hexamer structures," The Journal of Physical Chemistry A, **113**, 3555–3559 (2009), pMID: 19354314, http://pubs.acs.org/doi/pdf/10.1021/jp8105919.

[15]M. Losada and S. Leutwyler, "Water hexamer clusters: Structures, energies, and predicted mid-infrared spectra," The Journal of Chemical Physics, **117**, 2003–2016 (2002).

[16]Y. Wang, V. Babin, J. M. Bowman, and F. Paesani, "The water hexamer: Cage, prism, or both. full dimensional quantum simulations say both," Journal of the American Chemical Society, **134**, 11116–11119 (2012), http://pubs.acs.org/doi/pdf/10.1021/ja304528m.

[17]Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, and J. M. Bowman, "Flexible, ab initio potential, and dipole moment surfaces for water. i. tests and applications for clusters up to the 22-mer," The Journal of Chemical Physics, **134**, 094509 (2011).

[18]G. R. Medders, V. Babin, and F. Paesani, "A critical assessment of two-body and three-body interactions in water," Journal of Chemical Theory and Computation, **9**, 1103–1114 (2013), http://pubs.acs.org/doi/pdf/10.1021/ct300913g.

[19]S. Habershon, T. E. Markland, and D. E. Manolopoulos, "Competing quantum effects in the dynamics of a flexible water model," J. Comp. Phys., **131**, 024501 (2009).

[20]C. Pérez, M. T. Muckle, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel, and B. H. Pate, "Structures of cage, prism, and book isomers of water hexamer from broadband rotational spectroscopy," Science, **336**, 897–901 (2012), http://www.sciencemag.org/content/336/6083/897.full.pdf.

[21]J. Halton, "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," Numerische Mathematik, **2**, 84–90 (1960), ISSN 0029-599X.

[22]I. Sobol, "On the distribution of points in a cube and the approximate evaluation of integrals," USSR Computational Mathematics and Mathematical Physics, **7**, 86–112 (1967), ISSN 0041-5553.

[23]I. Sobol, "Uniformly distributed sequences with an additional uniform property," USSR Computational Mathematics and Mathematical Physics, **16**, 236–242 (1976), ISSN 0041-5553.

[24]H. Faure, "Discrépance de suites associées à un système de numération (en dimension s)," Acta Arithmetica, **41**, 337–351 (1982).

[25]H. Niederreiter, "Point sets and sequences with small discrepancy," Monatshefte für Mathematik, **104**, 273–337 (1987), ISSN 0026-9255, 10.1007/BF01294651.

[26]W. Morokoff and R. Caflisch, "Quasi-random sequences and their discrepancies," SIAM Journal on Scientific Computing, **15**, 1251–1279 (1994), http://epubs.siam.org/doi/pdf/10.1137/0915077.

[27]W. J. Morokoff and R. E. Caflisch, "Quasi-monte carlo integration," Journal of Computational Physics, **122**, 218–230 (1995), ISSN 0021-9991.

[28]R. Schürer, "A comparison between (quasi-)monte carlo and cubature rule based methods for solving high-dimensional integration problems," Math. Comput. Simul., **62**, 509–517 (2003), ISSN 0378-4754.

[29]P. Frantsuzov, A. Neumaier, and V. A. Mandelshtam, "Gaussian resolutions for equilibrium density matrices," Chemical Physics Letters, **381**, 117 – 122 (2003), ISSN 0009-2614.

[30]P. A. Frantsuzov and V. A. Mandelshtam, "Quantum statistical mechanics with gaussians: Equilibrium properties of van der waals clusters," J. Comp. Phys., **121**, 9247–9256 (2004).

[31]J. Deckman and V. A. Mandelshtam, "Effects of quantum delocalization on structural changes in lennard-jones clusters†," J. Phys. Chem. A, **113**, 7394–7402 (2009).

[32]J. Matoušek, "On thel2-discrepancy for anchored boxes," Journal of Complexity, **14**, 527 – 556 (1998), ISSN 0885-064X.

[33]J. Baldeaux, J. Dick, J. Greslehner, and F. Pillichshammer, "Construction algorithms for higher order polynomial lattice rules," Journal of Complexity, **27**, 281 – 299 (2011), ISSN 0885-064X, ¡ce:title¿Dagstuhl 2009¡/ce:title¿.

[34]A. Hermann, R. P. Krawczyk, M. Lein, P. Schwerdtfeger, I. P. Hamilton, and J. J. P. Stewart, "Convergence of the many-body expansion of interaction potentials: From van der waals to covalent and metallic systems," Phys. Rev. A, **76**, 013202 (2007).

[35]J. Cui, H. Liu, and K. D. Jordan, "Theoretical characterization of the (h2o)21 cluster: Application of an n-body decomposition

procedure," The Journal of Physical Chemistry B, **110**, 18872–18878 (2006).

[36] G. S. Fanourgakis and S. S. Xantheas, "Development of transferable interaction potentials for water. v. extension of the flexible, polarizable, thole-type model potential (ttm3-f, v. 3.0) to describe the vibrational spectra of water clusters and liquid water," The Journal of Chemical Physics, **128**, 074506 (2008).

[37] U. Gora, R. Podeszwa, W. Cencek, and K. Szalewicz, "Interaction energies of large clusters from many-body expansion," The Journal of Chemical Physics, **135**, 224102 (2011).