Optimization of ground and excited state wavefunctions and van der Waals clusters

M. P. Nightingale and Vilen Melik-Alaverdian

Department of Physics, University of Rhode Island, Kingston RI 02881, USA

(October 24, 2018)

A quantum Monte Carlo method is introduced to optimize excited state trial wavefunctions. The method is applied in a correlation function Monte Carlo calculation to compute ground and excited state energies of bosonic van der Waals clusters of up to seven particles. The calculations are performed using trial wavefunctions with general three-body correlations.

PACS codes: 03.65.02, 02.50.N, 02.70.L, 36.40.M, 34.30

Weakly-bound clusters display strong anharmonicity, and this makes solving the Schrödinger equation for such systems a formidable computational challenge. The discrete variable representation method (DVR) [1] has been applied with success to compute the energies of vibrational states for systems with up to six degrees of freedom. Its computational complexity scales exponentially with dimension, a problem that Monte Carlo methods can avoid. Indeed, correlation function Monte Carlo calculations are applicable to higher dimensional systems, although in practice they are restricted to a smaller number of excited states.

The accuracy of Monte Carlo projection methods can be improved dramatically by employing approximate eigenfunctions. In fact, without good initial approximations one rarely obtains results of sufficient accuracy. In this Letter, we discuss a systematic and efficient method to construct approximate eigenfunctions by optimization of many-parameter trial functions. We then use these functions in a correlation function Monte Carlo calculation. We expect that also POITSE calculations can be improved by the same means. A variant of the method described here was applied previously to study critical point interactions.

In that case, for $k = 1, \ldots, n$,

$$
\psi^{(k)}(R) = \sum_i \beta_i(R) d_i^{(k)}
$$

is an eigenvector of $H$ with an eigenvalue $\tilde{E}_k$ equal to the exact energy $E_k$, if $d_i^{(k)}$ is a right eigenvector of $E$ with eigenvalue $\tilde{E}_k$. We rewrite Eq. (1) in matrix form

$$
B' = BE,
$$

where $B_{\sigma i} = \hat{\beta}_i(R_{\sigma})$ and $B'_{\sigma i} = \hat{\beta}'_i(R_{\sigma})$, with $\hat{\beta}_i = \beta_i/\psi_g$ and $\hat{\beta}'_i = H \beta_i/\psi_g$.

In practice, the subspace spanned by the basis functions is not invariant, so that for given matrices $B$ and $B'$ Eq. (2) is an overdetermined set of equations for the unknown matrix $E$. If one multiplies through from the left by $B'^T$, the transpose of $B$, one obtains by inversion

$$
E = (B'^T B)^{-1} (B'^T B') \equiv N^{-1}H.
$$

As readily verified, this is the least-squares solution of Eq. (3); note that the rows of the matrices $B$ and $B'$ are weighted by the guiding function so that the elements of the matrices $N$ and $H$ approach the standard quantum mechanical overlap integrals and matrix elements in the limit of an infinite Monte Carlo sample. Eqs. (2) and (4) are usually derived from stationarity (in the linear parameters) of the average energy. If the latter is estimated by a finite-sample average, requiring stationarity of this estimate yields Eq. (4) with $H$ replaced by its symmetrized analog. Since the exact quantum mechanical expression is indeed symmetrical, one might be
inclined to use the symmetrized $H$. However, only the non-symmetric expression $B^T B'$ in Eq. (3) satisfies the zero-variance principle of yielding exact results independent of the configuration sample if the basis functions span an invariant subspace of the Hamiltonian. As in the ideal case, Eq. (2) determines the linearly optimized trial functions, but now one has $E_k \lesssim \hat{E}_k$, an inequality which for a finite Monte Carlo sample may be violated because of statistical noise.

The solution for $\mathcal{E}$ as written in Eq. (1) is numerically unstable since the matrix $N$ is ill conditioned because of near-linear dependence of the $\beta_k$. The solution to this problem is to use a singular value decomposition to obtain a numerically regularized inverse $B^{-1}$. In terms of the latter, one finds from Eq. (3)

$$\mathcal{E} = B^{-1} B'.$$

With the linear variational parameters optimized for fixed non-linear parameters in the elementary basis functions, we optimize --following Umrigar et al. the non-linear parameters by minimization of

$$\chi^2 = \frac{\sum_{\sigma} \left( \tilde{\psi}_k^{(l)}(R_{\sigma}) - \tilde{E}_k \tilde{\psi}_k^{(l)}(R_{\sigma}) \right)^2}{\sum_{\sigma} \tilde{\psi}_k^{(l)}(R_{\sigma})^2},$$

the variance of the local energy of the wavefunction given in Eq. (2); again, the guiding function is incorporated via $\tilde{\psi}_k^{(l)} = \psi_k^{(l)}/\tilde{\psi}_k^{(l)}$ and $\tilde{\psi}_k^{(l)} = H\tilde{\psi}_k^{(l)}/\tilde{\psi}_k^{(l)}$.

We now address the choice of the guiding function $\tilde{\psi}_k$. To obtain acceptable statistical errors, the sample has to have sufficient overlap with the desired excited states. In our case, this can be accomplished with $\tilde{\psi}_k^{(l)} = \tilde{\psi}_k^{(1)}$ with a power $p$ in the range $2 \lesssim p \lesssim 3$, while the ground state wavefunction $\tilde{\psi}_k^{(1)}$ is obtained after a few initial iterations.

The elementary basis functions are the final ingredient of the computation. They are defined as functions of all interatomic distances $r_{\sigma\tau}$ and scaled variables $\tilde{r}_{\sigma\tau} = f(r_{\sigma\tau})$. Here, $f$ maps the interatomic distances monotonically onto the interval $(-1, 1)$ such that most of the variation occurs where the wavefunction differs most from zero; the explicit form of $f$ is not important for the current discussion.

The elementary basis functions used for energy level $k$ have non-linear variational parameters $a_{j}^{(k)}$, and are of the form

$$\beta_l(R) = \sum_j a_j^{(k)} s_j(R) - \sum_{\sigma<\tau} \left( \kappa_{k} r_{\sigma\tau} + \frac{\sqrt{m}}{5 r_{\sigma\tau}} \right).$$

The prefactor polynomial $s_i$ has bosonic symmetry, and contains general three-body correlations, since all polynomials symmetric in $x, y$, and $z$ can be written as polynomials in the three invariants $I_i = x^i + y^i + z^i$, with $l = 1, 2, 3$ and v.v.

The number of elementary basis functions is limited by a bound on the total degree $\sum_l m_{l\mu}$; the polynomials $s_i$ in the exponent are of the same form as those in the prefactor, and their number in Eq. (6) is limited similarly.

The constant $\kappa_k$ is determined self-consistently so that the wavefunction has the correct exponential decay in the limit that a single atom goes to infinity. Assuming $-\kappa$ is plausible for the small clusters studied here -- that the energy of a cluster is roughly proportional to the number of atom pairs, we find

$$\kappa_k = \frac{2}{N_c - 1} \sqrt{-m \tilde{E}_k / N_c}.$$
the analog of Eq. (4) is used to compute eigenvalues, rather than Eq. (3). The singular value decomposition (SVD) cannot be used for the correlation function Monte Carlo because there are too many configurations $R_{0}$ to store the required matrices $B$ and $B'$. However, since the optimized basis functions are few in number and roughly orthonormal—at least for small projection times $t$—the SVD is not essential in this case.

Before we present estimates of the excited state energies, we discuss the sources of error of this method. In addition to the statistical errors, there are two systematic errors. For any finite projection time $t$ and in the limit of vanishing statistical errors, the energies computed by this method are upper bounds to the exact energies. In practice, since the statistical errors increase with projection time, one should choose the smallest projection time such that the projection and statistical errors are of the same order of magnitude. To pinpoint that time one has to distinguish real trends from false ones due to correlated noise. This is always tricky, but a troublesome detail is that at that point the results tend to have a non-Gaussian distribution, which makes it difficult to produce error bars with a well defined statistical meaning. In addition, there is the time-step error, which arises because of the imaginary-time evolution operator $\exp(-tH)$ has to be evaluated as the limit $\tau \to 0$ of $[\exp(-tH) + O(\tau^{2})]^{1/\tau}$, but this error is much easier to control.

Next we present results for excited state energies for clusters with up to seven atoms. First, we computed energies for trimers of Ne, Ar, Kr, and Xe ($m = 7.092 \times 10^{-3}$, $6.9635 \times 10^{-4}$, $1.9128 \times 10^{-4}$, and $7.8508 \times 10^{-5}$). Since our variational functions contain general three-body correlations, the accuracy of the wavefunctions and energies for the trimers can be improved without apparent limit other than the machine precision. During optimization of the wavefunctions for the trimers, we typically start with the ground state wavefunction which has prefactor degree of 5 or 6. For the trimers we chose not to vary the polynomial coefficients in the exponent and simply used the fixed terms required by the boundary conditions. The quality of the wavefunctions may be improved by varying polynomial coefficients in the exponent, and for larger clusters it becomes important to include such polynomials.

For the optimization we used samples consisting of 4000 configurations, and we gradually increased the prefactor degree to improve the quality of the trial functions. For Ne trimers we performed diffusion Monte Carlo calculations using optimized wavefunctions with prefactor degrees up to 14. Although, in principle, for trimers nothing should preclude further improvement, the observed changes were statistically insignificant in the 12 to 14 degree range. Table I contains results for degree 12.

There was no statistically significant difference between time steps $\tau = 0.4$, 0.2, and 0.1, and thus no noticeable time-step error. In the diffusion Monte Carlo calculations we used 1.3 million Monte Carlo steps. For Ar, Kr, and Xe trimers we found that the quality of the wavefunctions does not improve beyond the prefactor degree 10. The results in Table I for the three more massive noble gas atoms were obtained using trial wavefunctions with such polynomials. Convergence with respect to the time-step was established by comparing $\tau = 0.8$, 0.6, and 0.4. The number of Monte Carlo steps is the same as for Ne. Except for the energy of the fifth level of Ne, which is 0.009 too high, the results in Table I agree with, and in some cases improve upon, those of Leitner et al. [3].

In Table II we present results for the energies of the first five levels of Ar clusters of sizes 4 through 7. Our method allows one to go beyond 7 atom clusters, but as one can see from Table II the statistical error increase with system size. To obtain more accurate results for larger clusters it would probably be helpful to include higher order correlations in the wavefunction. In the calculations for 4 through 7 atom clusters we used a 10 degree prefactor and an exponent of degree three. Again, 1.3 million step diffusion Monte Carlo results were compared for $\tau = 0.8$, 0.6, and 0.4.

As to the performance of our method as the mass $m$ decreases and the atoms become more weakly bound, we find that both the optimization and the projection methods begin to fail, because the elementary basis functions lack the required variational freedom. This breakdown is illustrated in Fig. 1, which contains three energy levels as a function of mass for a four atom cluster. The results are plotted using variables chosen so that there is linear dependence both

**TABLE I.** Energy levels $E_k$ of the rare gas trimers; the errors are estimated to be a few units in the least significant digit.

| $k$ | Ne$_3$ | Ar$_3$ | Kr$_3$ | Xe$_3$ |
|-----|--------|--------|--------|--------|
| 1   | -1.719569 | -2.55328943 | -2.76055534 | -2.84524150 |
| 2   | -1.22283 | -2.2501855 | -2.5812399 | -2.7249558 |
| 3   | -1.1420  | -2.126361  | -2.5069468  | -2.6750648 |
| 4   | -1.038   | -1.99643   | -2.412444   | -2.608615  |
| 5   | -0.890   | -1.9467    | -2.387973   | -2.592226  |

**TABLE II.** Energy levels of Ar clusters of up to seven atoms; the errors are estimated to be a few units in the least significant digit.

| $k$ | Ar$_4$ | Ar$_5$ | Ar$_6$ | Ar$_7$ |
|-----|--------|--------|--------|--------|
| 1   | -5.11811 | -7.7851 | -10.8879 | -14.191 |
| 2   | -4.785  | -7.567 | -10.561 | -13.969 |
| 3   | -4.674  | -7.501 | -10.51  | -13.80  |
| 4   | -4.530  | -7.39  | -10.46  | -13.74  |
| 5   | -4.39   | -7.36  | -10.35  | -13.71  |
for large masses and for energies close to zero. As the energy of the levels approaches zero, the scatter in the data points increases, and ultimately the method fails to produce reliable results. Again, the use of trial wave-functions with four-body correlations is likely to make it possible to continue to smaller masses.

This research was supported by the (US) National Science Foundation (NSF) through Grant DMR-9725080. It is our pleasure to thank David Freeman and Cyrus Umrigar for valuable discussions and suggestions for improvements of the manuscript.

[1] D.H. Zhang, Q. Wu, J.Z.H. Zhang, M. von Dirke, and Z. Bačić, J. Chem. Phys. 102, 2315 (1995).
[2] D.M. Ceperley and B. Bernu, J. Chem. Phys. 89, 6316 (1988).
[3] B. Bernu, D.M. Ceperley, and W.A. Lester, Jr., J. Chem. Phys. 93, 552 (1990); W.R. Brown, W.A. Glauser, and W.A. Lester, Jr., J. Chem. Phys. 103, 9721 (1995).
[4] D. Blume, M. Lewerenz, P. Niyaz, and K.B. Whaley, Phys. Rev. E 55, 3664 (1997). D. Blume and K. B. Whaley, J. Chem. Phys. 112, 2218 (2000) and references therein.
[5] M.P. Nightingale and H.W.J. Blöte, Phys. Rev. Lett. 80, 1007 (1998).
[6] M.P. Nightingale and H.W.J. Blöte, Phys. Rev. B 62, 1089 (2000).
[7] J. de Boer, Physica, 14, 139 (1948).
[8] J.K.L. MacDonald, Phys. Rev. 43, 830 (1933).
[9] G.H. Golub and C.F. van Loan, Matrix computations (Second Edition), (Johns Hopkins University Press, 1989) Chapter 5.5.
[10] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes, (Cambridge University Press, Cambridge 1992), Section 2.6.
[11] We ignored all singular values $\sigma_k$ with $\sigma_k < 10^3 \sigma_1 \epsilon_{\text{mach}}$, where $\epsilon_{\text{mach}}$ is the double precision machine accuracy.
[12] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. 60, 1719 (1988); C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, in Computer Simulation Studies in Condensed Matter Physics, Recent Developments, edited by D.P. Landau K.K. Mon and H.B. Schüttler, Springer Proceedings in Physics (Springer, Berlin, 1988).
[13] Andrei Mushinskii and M.P. Nightingale, J. Chem. Phys. 101, 8831, (1994).
[14] M. Meierovich, A. Mushinskii, and M.P. Nightingale, J. Chem. Phys. 105, 6498 (1996).
[15] D.M. Leitner, J.D. Doll, and R.M. Whitnell, J. Chem. Phys. 94, 6644 (1991).
[16] V. Melik-Alaverdian and M.P. Nightingale, Int. J. Mod. Phys. C 10, 1409, (2000).
[17] H. Hetherington, Phys. Rev. A 30, 2713 (1984).
[18] C.J. Umrigar, M.P. Nightingale, and K.J. Runge, J. Chem. Phys. 99, 2865 (1993) and references therein; we used a single, weighted walker, and an accept reject step with $\tau_{\text{eff}} = \tau (0)$ for accepted (rejected) moves.
[19] M. Caffarel and P. Claverie, J. Chem. Phys. 88, 1088 (1988); ibid. p. 1100.
[20] M. Foulkes, L. Mitas, R. Needs and G. Rajagopal, Rev. Mod. Phys. , 73, 33 (2001).