Boosting the accuracy and speed of quantum Monte Carlo: Size consistency and time step

Andrea Zen, Sandro Sorella, Michael J. Gillan, Angelos Michaelides, and Dario Alfè

Phys. Rev. B **93**, 241118 — Published 29 June 2016

DOI: 10.1103/PhysRevB.93.241118

Boosting the accuracy and speed of quantum Monte Carlo: size-consistency and time-step

Andrea Zen[1,2,3], Sandro Sorella[4], Michael J. Gillan[1,2,3], Angelos Michaelides[1,2,3], and Dario Alfè[1,2,3,5*]
[1] *London Centre for Nanotechnology, Gordon St., London WC1H 0AH, UK* [2] *Thomas Young Centre, University College London, London WC1H 0AH, UK* [3] *Dept. of Physics and Astronomy, University College London, London WC1E 6BT, UK* [4] *International School for Advanced Studies (SISSA), Via Beirut 2-4, 34014 Trieste, Italy and INFM Democritos National Simulation Center, Trieste, Italy* [5] *Dept. of Earth Sciences, University College London, London WC1E 6BT, UK*
(Dated:)

Diffusion Monte Carlo (DMC) simulations for fermions are becoming the standard for providing high quality reference data in systems that are too large to be investigated via quantum chemical approaches. DMC with the fixed-node approximation relies on modifications of the Green function to avoid singularities near the nodal surface of the trial wavefunction. Here we show that these modifications affect the DMC energies in a way that is not size-consistent, resulting in large time-step errors. Building on the modifications of Umrigar *et al.* and DePasquale *et al.* we propose a simple Green function modification that restores size-consistency to large values of the time-step, which substantially reduces time-step errors. The new algorithm also yields remarkable speedups of up to two orders of magnitude in the calculation of molecule-molecule binding energies and crystal cohesive energies, thus extending the horizons of what is possible with DMC.

The determination of accurate reference energetics for solids is one of the grand challenges of materials modelling. Reliable reference data is needed to make accurate predictions about any number of phenomena, such as phase stability, adsorption on surfaces and crystal polymorph prediction. Very often density functional theory (DFT) provides sufficient accuracy for this and as such has been immensely successful in furthering our understanding of materials [1, 2]. However, there are many materials and materials related problems for which DFT does not deliver the desired accuracy [3]. For such problems explicitly correlated wave-function based approaches are needed, such as the approaches of quantum chemistry, quantum Monte Carlo (QMC), and combinations thereof [4–15]. In practice for condensed phase systems with more than a handful of atoms in the unit cell QMC remains the only feasible reference method, partly because of its favorable scaling with system size and the fact that it can be used efficiently on massively parallel supercomputers. Indeed QMC, mostly within the diffusion Monte Carlo (DMC) approach, is increasingly used to provide benchmark data for solids and to tackle interesting materials science problems that have been beyond the reach of DFT [16–29]. DMC is also proving increasingly useful in exposing and helping to explain problems with DFT and as such in helping to further the development of DFT.

DMC is in principle an exact technique to solve the imaginary time dependent Schrödinger equation. The discretization of time in practical implementations introduces a time-step (τ) error, the computational cost of which is proportional to $1/\tau$. Recently Gillan *et al.* [21] showed that for CH_4-H_2O clusters current implementations of DMC appear to be non size-consistent, i.e. the total energy of a system of M non-interacting molecules is not proportional to M. Here we show that this is a general problem, we identify its source, and propose a simple modification that solves it. Moreover, we observe that the time-step error in binding energy evaluations is mostly due to this size-consistency issue. Our proposed method also leads to remarkable speedups, by significantly increasing the accuracy of large τ DMC evaluations [30].

A review of DMC can be found elsewhere [4, 31], and is summarized in the SI [32]. To understand the size-consistency issue we recall the main ideas of the method and how it is applied in practice. Consider the Schrödinger equation in imaginary time for a system including N particles with the *fixed-node* constraint, i.e. with the solution $\Phi(\mathbf{R},t)$, where $\mathbf{R}$ is the electronic configuration and t the time, forced to have the same nodal surface of some guiding function $\psi_G(\mathbf{R})$ (the $3N-1$ hyper-surface where $\psi_G=0$). This is achieved, within the importance sampling scheme, by introducing the *mixed distribution* $f(\mathbf{R},t)=\psi_G(\mathbf{R})\Phi(\mathbf{R},t)$, which satisfies the equation:

$$-\frac{\partial f}{\partial t}=-\frac{1}{2}\nabla^2 f+\nabla\cdot[\mathbf{V}f]-Sf\,. \qquad (1)$$

Here we have omitted the functional dependence of the terms and $\mathbf{V}(\mathbf{R})\equiv\nabla\log|\psi_G(\mathbf{R})|$ is usually called the *drift velocity*, $S(\mathbf{R})\equiv E_T-E_L(\mathbf{R})$ is the *branching* term, E_L is the local energy, and E_T is an energy shift. The three terms on the right hand side of Eq. 1 are responsible for diffusion, drift and branching processes, respectively. Eq. 1 can be rewritten in integral form:

$$f(\mathbf{R},t+t_0)=\int G(\mathbf{R},\mathbf{R}';t)f(\mathbf{R}',t_0)d\mathbf{R}' \qquad (2)$$

where $G(\mathbf{R},\mathbf{R}';t)$ is the Green function for the importance sampling. The DMC method is a stochastic realization of Eq. 2, in which a series of *walkers* initially distributed as some $f(\mathbf{R},0)=\sum_i \delta(\mathbf{r}-\mathbf{r_i})$ is propagated ahead in time through a branching-drift-diffusion process [32]. In the limit $t\to\infty$ the walkers end up distributed as $\psi_G(\mathbf{R})\phi(\mathbf{R})$, with $\phi(\mathbf{R})$ the ground state of the Hamiltonian subject to the fixed-node constraint.

A practical implementation of Eq. 2 faces a problem: $E_L(\mathbf{R})$ and $\mathbf{V}$ diverge at the nodes of ψ_G as the inverse of the distance between the nodal surface and $\mathbf{R}$. As $\tau\to 0$ these two singularities are not an issue because the drift term prevents the walkers from approaching the nodal surface. However, for finite τ, walkers can end up close to the nodal surface with catastrophic consequences. A practical solution to this problem is to introduce limits to the drift velocity and to the local energy. Umrigar, Nightingale and Runge [31] (UNR) proposed to replace $\mathbf{V}=(\mathbf{v}_1,\ldots,\mathbf{v}_N)$ with $\bar{\mathbf{V}}=(\bar{\mathbf{v}}_1,\ldots,\bar{\mathbf{v}}_N)$, defined as:

$$\bar{\mathbf{v}}_i=\frac{-1+\sqrt{1+2av_i^2\tau}}{av_i^2\tau}\mathbf{v}_i;\quad \mathbf{v}_i=\nabla_i\log|\psi_G(\mathbf{R})|\,, \quad (3)$$

with a an adjustable parameter between 0 and 1. This expression provides a rough approximation to the average velocity over a time-step, which has the effect of limiting the drift distance [31]. The branching factor $S(\mathbf{R})$ is replaced with:

$$\bar{S}(\mathbf{R})=[E_T-E_{\rm best}]+[E_{\rm best}-E_L(\mathbf{R})]\frac{\bar{V}}{V}, \quad (4)$$

where $E_{\rm best}$ is the best estimate of the energy, $V=\|\mathbf{V}\|$ and $\bar{V}=\|\bar{\mathbf{V}}\|$. This limiting procedure is elegant and minimises instabilities because the divergences of $E_L(\mathbf{R})$ at the nodes are cancelled by divergences in V. As a result it is now standard in most DMC simulations. However, this limiting procedure is an approximation of the Green function which renders DMC size-inconsistent, see discussion in Sec. I.D of SI [32]. The issue disappears for $\tau\to 0$, where $\bar{V}/V\to 1$, but for $\tau>0$ the total energy is not proportional to the size of the system. Since the main application area of DMC is the calculation of medium to large systems for which relatively small energy differences are computed but very small τ cannot be afforded, this issue threatens the usefulness of DMC in material science.

To quantify the size-consistency problem consider two systems A and B with energies E_A and E_B, and define $E_{A,B}^{\rm separated}$ as the energy of the system with A and B at large enough distance from each other to have zero interaction. The quantity $E_s=E_{A,B}^{\rm separated}-(E_A+E_B)$ is therefore expected to be equal to zero and if it is not it measures the size-consistency error. To compute the binding energy of the system where A and B are interacting and have a total energy $E_{A,B}^{\rm bonded}$ it is useful to define $E_b=E_{A,B}^{\rm bonded}-(E_A+E_B)$ and $E_{bs}=E_{A,B}^{\rm bonded}-E_{A,B}^{\rm separated}$.

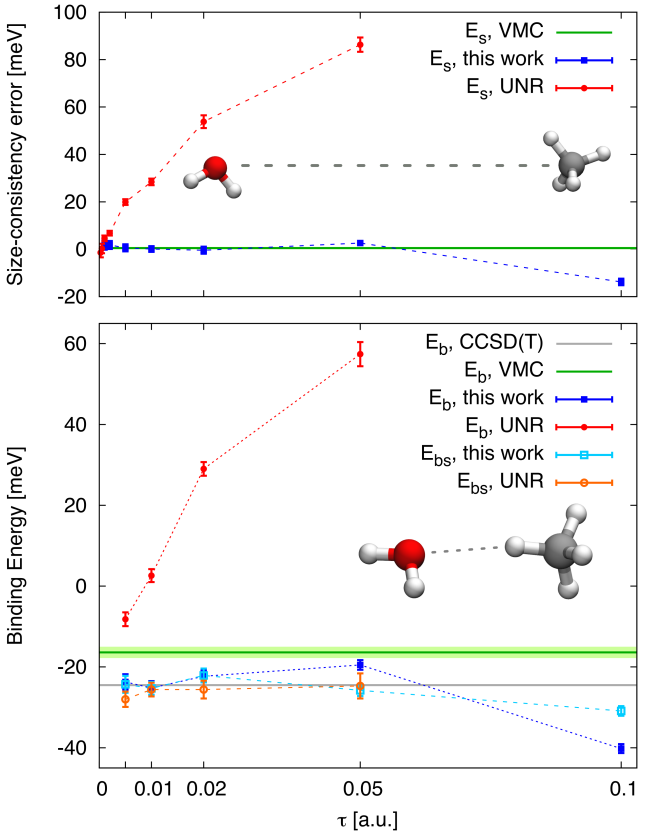

FIG. 1. (Top) Size-consistency error E_s (see text) and (bottom) binding energy [using two different definitions, E_b and E_{bs} (see text)] for the CH_4-H_2O system. Results from the limited branching term given by Eq. 4 (UNR) or the approach introduced here Eqs. 5,6 (this work) are reported. VMC and CCSD(T) [21] evaluations are also shown. Error bars are one standard deviation. The insets show the structures of the complexes which have the molecules at large (top) and near the equilibrium (bottom) separation.

Here E_b may be affected by a size-consistency problem, whereas E_{bs} is not. To illustrate the problem we have selected three representative examples with a broad range of interaction strengths, involving both isolated and periodic systems.

DMC simulations were carried out with the CASINO code [33]. We used Dirac-Fock pseudopotentials [34, 35] with the locality approximation [36]. The trial wavefunctions were of the Slater-Jastrow type with single Slater determinants and the single particle orbitals obtained from DFT-LDA plane-wave calculations performed with PWSCF [37] and re-expanded in terms of B-splines [38].

Our first example is a system formed by a CH_4 (A) and a H_2O (B) molecule. $E_{A,B}^{\rm separated}$ is obtained for a C-O distance of 11.44 Å. On the basis of CCSD(T) calculations we know that the residual interaction energy is <0.1 meV, negligible for our purposes. E_s is zero also for variational Monte Carlo (VMC), showing that the trial wavefunction of the dimer $\psi_{\rm CH_4,H_2O}^{\rm separated}$ is effectively factor-

ized: $\psi^{\rm separated}_{CH_4,H_2O} = \psi_{CH_4} \otimes \psi_{H_2O}$.

In Fig. 1 (top) we plot E_s computed with DMC as a function of τ. For $\tau \to 0$, $E_s \to 0$ as expected. However, at a typical time-step $\tau = 0.005$ a.u. [21] the error is already ∼20 meV, which is about the same size as the binding energy of the dimer near the equilibrium distance, and it grows to over 80 meV at $\tau = 0.05$ a.u.. In Fig. 1 (bottom) we show the binding energy of the molecule for a configuration near the equilibrium distance [39]. As expected from the large size-consistency problem highlighted above, the binding energy computed with E_b is wrong, and has a strong time-step dependence. Extrapolating to zero time-step using the whole $0.005 \leq \tau \leq 0.05$ range yields $E_b = 11 \pm 7$ meV. Using only the range $0.005 \leq \tau \leq 0.02$ a value of $E_b = 21 \pm 2$ meV is obtained, which is close to the benchmark energy $E_b = 24.5$ meV, obtained with coupled cluster with singles, doubles and perturbative triples (CCSD(T) and a large basis set) [21]. By contrast, E_{bs} is effectively time-step independent up to $\tau = 0.05$, is in better agreement with the reference value, and removes the need for uncertain and arbitrary extrapolations. The UNR limiting procedure is too unstable above $\tau = 0.05$ and even at $\tau = 0.05$ we have not been able to obtain a very small statistical error due to instabilities in the simulations [32].

Although one could envisage always using definitions analogous to E_{bs} to compute binding energies, it is much more desirable to be able to use E_b instead, particularly when one is concerned with the binding energy of more than just a dimer [40].

To address this size-consistency issue we propose a new limiting procedure. As proven in Sec. I.D of the SI [32], the UNR limit for the drift term, Eq. 3, does not affect size-consistency, thus we only need to modify the branching term. Our method is based on the idea that any modifications to the Green function should be as insensitive as possible to the size of the system. Inspired by the prescriptions of DePasquale *et al.* [41], in which the local energy entering the branching factor is limited by a cutoff $E_{\rm cut}$, a modified branching factor is defined as:

$$\begin{aligned} \bar{S}(\mathbf{R}) &= E_T - \bar{E}_L(\mathbf{R}); \\ \bar{E}_L(\mathbf{R}) &= E_{\rm best} + {\rm sign}[E_L(\mathbf{R}) - E_{\rm best}] \times \\ &\quad \min\{E_{\rm cut}, |E_L(\mathbf{R}) - E_{\rm best}|\}; \end{aligned} \tag{5}$$

In the original [41] recipe $E_{\rm cut} = 2/\sqrt{\tau}$. This has the consequence that for larger systems a larger fraction of the distribution of the branching factor is modified, leading again to a size-consistency issue. Here we propose:

$$E_{\rm cut} = \alpha\sqrt{N/\tau}, \tag{6}$$

where N is the number of electrons in the system. Since the variance of the system is proportional to N, this ensures that the proportion of the distribution of the branching factor modified by the cutoff is similar for systems with different values of N [42]. As with the original approach [41], the exact Green function is restored in the limit $\tau \to 0$. The parameter α is an arbitrary constant to be conveniently chosen. For large enough values of α (and/or small values of τ) the Green function becomes exact, but then singularities reappear. For small values of α (and/or large values of τ) the bias in the DMC energy becomes large. We have found that a good compromise is obtained by setting $\alpha = 0.2$. The results obtained with this newly proposed scheme are displayed in Fig. 1, showing that the bias in the DMC energy is now size-consistent up to very large values of τ. The new scheme also reduces the time-step error on the absolute energies [32].

If the composite system is made of non-identical subsystems (like our water-methane system) then the method becomes less accurate at large τ, mainly because of the different widths of the S distributions. In particular, the cutoff at $\tau = 0.1$ a.u. corresponds to $E_{\rm cut}$ of around 3.5 σ, 2.7 σ and 3.0 σ for CH_4, H_2O and CH_4-H_2O, respectively, where σ indicates the corresponding standard deviation of the VMC local energy [43]. With such small cutoff energies, the percentage of the respective distributions that are cut are different enough to affect the bias of the local energy in a non size-consistent way, which is why the error reappears at large values of τ.

Binding energies computed with the new method are displayed in the bottom panel of Fig. 1, showing that E_{bs} has the same accuracy as that computed with the UNR branching factor, but now also E_b is accurate. The new method is stable also for τ =0.1 a.u., although at this very large value of the time-step the binding energy starts to show non negligible errors. Note that in order to obtain a sufficiently high accuracy on E_b with the UNR branching factor, without relying on extrapolations, we would need to reduce the time-step to at least $\tau \sim 0.0005$ a.u., which is two orders of magnitude smaller than what is required with our newly proposed method.

The second system we examined is the buckyball catcher, the C_{60}-$C_{60}H_{28}$ ($A-B$) complex. This is an example of a whole class of supramolecular systems which is generally out of reach of the most accurate quantum chemistry methods and so at present DMC is the prime candidate for examining such systems. For the calculation of $E^{\rm separated}_{A,B}$ we considered the system with the two fragments separated by 10 Å. The residual interaction energy at this distance is $\simeq$ 10 meV [44], which is again negligible compared to the energies involved. The new limiting procedure results in very good cancellation of time-step error and it is size-consistent up to at least $\tau = 0.05$ a.u.. The UNR branching factor causes a slightly larger time-step dependence of both E_b and E_{bs}, and the top panel of Fig. 2 highlights once again the size-

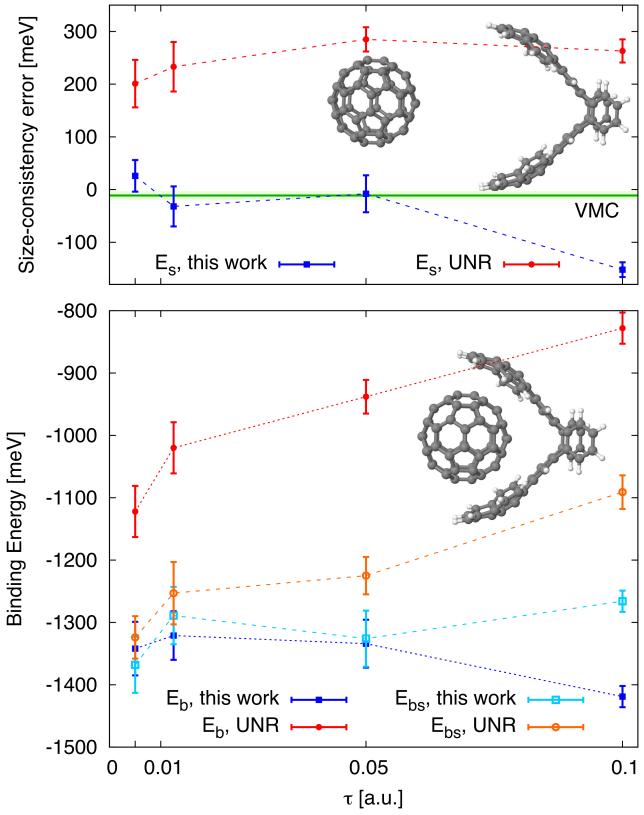

FIG. 2. Same as Fig. 1 but in this case for the C_{60}-$C_{60}H_{28}$ system.

consistency problem. Incidentally, the binding energy of this complex reported in [45] was computed using UNR and E_b, therefore it had a size-consistency error of ~ 0.2 eV. Note that in this case any sensible extrapolation to zero time-step would result in a large size-consistency error, and therefore to obtain accurate results we should use $\tau \sim 0.0005$ a.u., if not even smaller, which is over two orders of magnitude more expensive and out of reach even on the biggest supercomputers currently available.

Our third and final test was performed on a square lattice ice system, a H-bonded 2D-periodic system which has been the subject of recent theoretical [46, 47] and experimental [48] studies. The simulation cell comprises 64 water molecules. In Fig. 3 we show the cohesive energy as a function of time-step. The cohesive energy computed with the new limiting procedure is independent of time-step up to at least $\tau = 0.05$ a.u., while that computed with the UNR branching factor has errors even at the shortest time-step that we could afford ($\tau = 0.002$ a.u.). The non-linear trend of the UNR curve makes any $\tau \to 0$ extrapolation unreliable, unless simulations with $\tau < 0.001$ a.u. could be afforded. Given the size of this system this makes such calculations prohibitively expensive. Remarkably, the new method does not require any uncertain time-step extrapolations and yields a speedup of around two orders of magnitude.

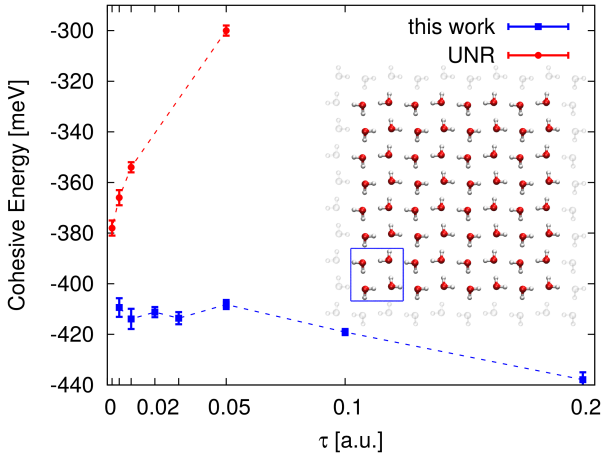

FIG. 3. Cohesive energy of a two-dimensional periodic square ice system with the UNR and current branching terms. The inset of the structure shows the simulated 64 molecule supercell as colored molecules, and the primitive unit cell in the blue square.

In summary, we have proposed a procedure that reduces DMC time-step errors by a large factor, and restores size-consistency. The method is based on the UMR scheme with an alternative branching factor. The modification is straightforward to implement, requiring just a change to a single line of code. We have demonstrated the new method on a CH_4-H_2O dimer, the C_{60}-$C_{60}H_{28}$ supramolecular system and 2-dimensional ice. Besides solving the size-consistency problem, speedups of two orders of magnitude are obtained (see Fig. 4 in [32]) and the need for time-step extrapolations is removed. The improvement appears particularly promising for investigations on molecular materials and to discriminate between crystal polymorphs. Moreover, the recent emergence of QMC-based molecular dynamics [24–26], which until now has only been affordable within VMC, could now be in reach with the more accurate fixed-node DMC approach.

AZ and AM's work has been sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-12-1-2099 and by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 616121 (HeteroIce project). AM is also supported by the Royal Society through a Wolfson Research merit Award. SS acknowledges CINECA for the use of computational facilities, under IscrB_SUMCHAL grant. Calculations were performed on the U.K. national service ARCHER, the UK's national high-performance computing service, which is funded by the Office of Science and Technology through EPSRC's High End Computing Programme. This research also used resources of the Oak Ridge Leadership Computing Facility located in the Oak

Ridge National Laboratory, which is supported by the Office of Science of the Department of Energy under Contract No. DE-AC05-00OR22725. We thank Cyrus Umrigar for useful discussions and Jan Hermann for providing the estimated residual binding energy of the $C_{60} - C_{60}H_{28}$(shifted) complex.

* d.alfe@ucl.ac.uk

[1] J. Hafner, C. Wolverton, and G. Ceder, MRS Bulletin **31**, 659 (2011).

[2] J. Neugebauer and T. Hickel, Wiley Interdisciplinary Reviews: Computational Molecular Science **3**, 438 (2013).

[3] A. J. Cohen, P. Mori-Sanchez, and W. Yang, Chem. Rev. **112**, 289 (2012).

[4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[5] C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, in *Rev. Comput. Chem.* (John Wiley & Sons, Inc., 2007) pp. 1–82.

[6] R. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[7] G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002).

[8] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[9] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, Nature **493**, 365 (2013).

[10] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).

[11] M. Casula, C. Filippi, and S. Sorella, Phys. Rev. Lett. **95**, 100201 (2005).

[12] M. Casula, S. Moroni, S. Sorella, and C. Filippi, J. Chem. Phys. **132**, 154113 (2010).

[13] J. Harl and G. Kresse, Phys. Rev. Lett. **103**, 056401 (2009).

[14] L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, Nat. Mater. **9**, 741 (2010).

[15] X. Ren, P. Rinke, C. Joas, and M. Scheffler, J. Mater. Sci. **47**, 7447 (2012).

[16] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).

[17] M. A. Morales, J. R. Gergely, J. McMinis, J. M. McMahon, J. Kim, and D. M. Ceperley, J. Chem. Theory Comput. **10**, 2355 (2014).

[18] S. J. Cox, M. D. Towler, D. Alfè, and A. Michaelides, J. Chem. Phys. **140**, 174703 (2014).

[19] A. Benali, L. Shulenburger, N. A. Romero, J. Kim, and O. A. von Lilienfeld, J. Chem. Theory Comput. **10**, 3417 (2014).

[20] Y. S. Al-Hamdani, M. Ma, D. Alfè, O. A. von Lilienfeld, and A. Michaelides, J. Chem. Phys. **142**, 181101 (2015).

[21] M. J. Gillan, D. Alfè, and F. R. Manby, J. Chem. Phys. **143**, 102812 (2015).

[22] Y. Virgus, W. Purwanto, H. Krakauer, and S. Zhang, Phys. Rev. B **86**, 241406 (2012).

[23] M. Morales, R. Clay, C. Pierleoni, and D. Ceperley, Entropy **16**, 287 (2014).

[24] G. Mazzola, S. Yunoki, and S. Sorella, Nat. Commun. **5**, 3487 (2014).

[25] G. Mazzola and S. Sorella, Phys. Rev. Lett. **114**, 105701 (2015).

[26] A. Zen, Y. Luo, G. Mazzola, L. Guidoni, and S. Sorella, J. Chem. Phys. **142**, 144111 (2015).

[27] J. Chen, X. Ren, X.-Z. Li, D. Alfè, and E. Wang, J. Chem. Phys. **141**, 024501 (2014).

[28] L. K. Wagner, Int. J. Quantum Chem. **114**, 94 (2013).

[29] L. K. Wagner and P. Abbamonte, Phys. Rev. B **90**, 125129 (2014).

[30] We note that other QMC approaches, such as the variational Monte Carlo (VMC) or the lattice regularized diffusion Monte Carlo (LRDMC) [11] methods, do not suffer from these problems. This has been shown in [12], where the effect of the cutoff in the local energy on the size-consistency issue was carefully considered also for LRDMC. In this paper, however, we are concerned with the much more widely used DMC.

[31] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).

[32] See supplementary information.

[33] R. J. Needs, M. D. Towler, N. D. Drummond, and P. L. Rios, J. Phys.: Condens. Matter **22**, 023201 (2010).

[34] J. R. Trail and R. J. Needs, J. Chem. Phys. **122**, 014112 (2005).

[35] J. R. Trail and R. J. Needs, J. Chem. Phys. **122**, 174109 (2005).

[36] L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[37] S. Baroni, A. Dal Corso, S. de Gironcoli, and P. Giannozzi, `http://www.pwscf.org`.

[38] D. Alfè and M. J. Gillan, Phys. Rev. B **70**, 161101 (2004).

[39] Note that this is not the water-methane dimer equilibrium configuration, but just a configuration in which the C-O distance is near the equilibrium value. The coordinates of the atoms in this configuration are reported in the SI.

[40] For example, in the case of a cluster formed by a large number of molecules the construction of the system with all molecules far enough away from each other could be difficult, or even impossible, and alternative correction schemes would be required [21].

[41] M. F. DePasquale, S. M. Rothstein, and J. Vrbik, J. Chem. Phys. **89**, 3629 (1988).

[42] Note that, given $f(S_A)$ the distribution of the branching factor S_A of some system A, the distribution $f(M; S_A)$ of a system containing M non-interacting copies of A does not have, in general, the same form. This is because the central limit theorem implies that $f(M; S_A)$ becomes Gaussian for large enough M, but in general $f(S_A)$ is not Gaussian. Thus the distribution cannot be modifierd in a way that is exactly size-consistent and our proposed method is therefore only approximate.

[43] The standard deviation σ_{DMC} of the DMC distributions will, in general, be different from the σ of the VMC distributions, but the same arguments would apply.

[44] J. Hermann, Private communication.

[45] A. Tkatchenko, D. Alfè, and K. S. Kim, J. Chem. Theory Comp. **8**, 4317 (2012).

[46] J. Chen, G. Schusteritsch, C. J. Pickard, C. G. Salzmann, and A. Michaelides, Phys. Rev. Lett. **116**, 025501 (2016).

[47] F. Corsetti, P. Matthews, and E. Artacho, Sci. Rep. **6**, 18651 (2016).

[48] G. Algara-Siller, O. Lehtinen, F. C. Wang, R. R. Nair,

U. Kaiser, H. A. Wu, A. K. Geim, and I. V. Grigorieva,
Nature **519**, 443 (2015).