Variational Wave Functions and Their Overlap with the Ground State

Christophe Mora[1,2] and Xavier Waintal[3]

[1]Laboratoire Pierre Aigrain, Département de Physique, Ecole Normale Supérieure, 24 rue Lhomond, F-75005 Paris, France
[2]Institut für Theoretische Physik, Heinrich-Heine Universität, D-40225 Düsseldorf
[3]Nanoelectronics Group, Service de Physique de l'État Condensé, CEA Saclay, 91191 Gif-sur-Yvette cedex, France

(Received 27 February 2007; published 20 July 2007)

An intrinsic measure of the quality of a variational wave function is given by its overlap with the ground state of the system. We derive a general formula to compute this overlap when quantum dynamics in imaginary time is accessible. The overlap is simply related to the area under the $E(\tau)$ curve, i.e., the energy as a function of imaginary time. This has important applications to, for example, quantum Monte Carlo simulations where the overlap becomes as a simple by-product of routine simulations. As a result, we find that the practical definition of a good variational wave function for quantum Monte Carlo simulations, i.e., fast convergence to the ground state, is equivalent to a good overlap with the actual ground state of the system.

DOI: 10.1103/PhysRevLett.99.030403 PACS numbers: 05.30.Jp, 03.75.Hh, 05.10.Ln

Variational wave functions (VWF) are very valuable tools to study interacting quantum systems. Examples are numerous and can be found in many fields of physics, like ^{4}He [1], high T_c superconductors [2], or the fractional quantum Hall effect [3]. Given a variational wave function Ψ_V, it is not difficult to sample $|\Psi_V|^2$ from which one obtains the expectation value of many physical observables. It remains difficult, however, to determine to which degree Ψ_V is a good approximation of the true ground state Ψ_0 of the system. A first answer to this question lies in the variational theorem itself: as the variational energy is always higher than the ground-state energy, one attempts to get the VWF with the lowest energy. The drawback of this criterion is that VWFs with similar variational energies can convey very different physics as the energy is very sensitive to the short range part of the VWF, i.e., when two particles are close to each other, but rather weakly to the long range part. This is unfortunate as other physical observables may depend crucially on the VWF. Another (standard) possibility is to pick the VWF that has the lowest variance of its energy. The variance criterion has some advantages over the energy [4] as it is a more absolute criteria: a zero variance means that the VWF is an eigenstate of the system (but not necessarily the ground state).

In this Letter, we concentrate on a more intrinsic criterion, namely, maximizing the overlap Ω:

$$\Omega \equiv \frac{|\langle\Psi_V|\Psi_0\rangle|^2}{\langle\Psi_0|\Psi_0\rangle\langle\Psi_V|\Psi_V\rangle} \tag{1}$$

of Ψ_V with the actual ground state Ψ_0 of the system. A direct calculation of Ω would require the complete knowledge of the ground state Ψ_0, which is extremely computationally demanding. The main result of this Letter, namely, Eqs. (3) and (4), is that Ω can be related to the energy, upon projection of the initial VWF in imaginary time. Hence, Ω can be simply obtained as the by-product of quantum Monte Carlo simulations. In this Letter, we focus on this particular method to determine Ω. We would like to emphasize, however, that our main result together with inequalities (6), (7), and (11) are general and can be applied to other methods.

The development of zero temperature quantum Monte Carlo (QMC) techniques (diffusive or Green function Monte Carlo algorithms) relies on the availability of good VWFs. Although those techniques allow one to access ground-state properties, VWFs play a crucial role for two reasons. First, QMC uses a VWF to properly sample the Hilbert space through "importance sampling." An homogeneous sampling of Hilbert space would mean spending a very large amount of time sampling regions of the Hilbert space that have almost no contribution to Ψ_0. Second, even though QMC simulations can give interesting information on the system they do not necessarily explain the physical mechanisms involved. Valuable insight can often be obtained just by looking at how the VWF is constructed. For instance, the short and long range behavior of a Jastrow type VWF, or more dramatically the analytical structure of the Laughlin wave function in the fractional quantum Hall effect, conveys in itself useful information. With QMC comes an empirical criterion for what is a good VWF: it should allow the simulation to converge from Ψ_V to Ψ_0 as fast as possible. We shall see that this criterion actually coincides with maximizing the overlap Ω of Ψ_V with the actual ground state Ψ_0.

The idea to maximize the overlap as a variational criterion was put forward 20 years ago [5–7]. It was shown that for a certain class of VWF, merely of the Jastrow type and its extensions, maximizing Ω amounts to choosing the Jastrow function that reproduces exactly the (mixed estimator of the) static structure function. Hence the practical criterion was to optimize the static structure function obtained from variational Monte Carlo simulations by comparison with that coming out of QMC simulations. Here we

 © 2007 The American Physical Society

choose a different route and show that Ω can be computed directly as a simple by-product of a QMC run. In QMC techniques one starts with an initial wave function Ψ_V and then integrates in a stochastic way the imaginary time Schrödinger equation $\partial_\tau \Psi = -H\Psi$, the formal solution being $\Psi(\tau) = e^{-\tau H}\Psi_V$. For a large τ, $\Psi(\tau)$ actually converges toward Ψ_0. In practical simulations, one usually computes the energy

$$E(\tau) = \frac{\langle \Psi_V | H | \Psi(\tau)\rangle}{\langle \Psi_V | \Psi(\tau)\rangle} = \frac{\langle \Psi_V | H e^{-\tau H} | \Psi_V\rangle}{\langle \Psi_V | e^{-\tau H} | \Psi_V\rangle} \tag{2}$$

until $E(\tau)$ has converged to its asymptotic value, the exact ground-state energy E_0. An example of a typical $E(\tau)$ curve is shown in Fig. 1 for the system discussed at the end of this Letter. By introducing κ as

$$\Omega \equiv e^{-\kappa}, \tag{3}$$

the principal result of this Letter is that κ is simply given by the shaded area in Fig. 1. Or more precisely,

$$\kappa = \int_0^\infty d\tau [E(\tau) - E_0]. \tag{4}$$

It is therefore straightforward to obtain κ from a QMC simulation.

Equation (4) calls for a few comments. (i) As mentioned above, Eq. (4) combines a theoretical criterion, maximizing the overlap between Ψ_V and Ψ_0, and a practical criterion that is fast convergence to E_0. Conversely, a wave function that has a low variational energy but converges very slowly toward E_0 has a poor overlap with the actual ground state. (ii) κ is not a substitute for the variational energy (or variance) as it requires a full QMC simulation. (iii) In contrast with energy and energy variance, Ω is dimensionless and hence has an absolute meaning. For example, $\Omega = 0.97$ means that the VWF captures 97% of the ground state. (iv) In some instances, one is interested in the thermodynamic $N \to \infty$ limit, where N is the number of particles in the system. It is easy to be convinced that the generic behavior of Ω is an exponential decrease with N. This can be seen, for instance, in a noninteracting Bose-Einstein condensate, where any error in the one particle wave function appears to the power N in the many-body wave function. More generally, the energy is usually an extensive quantity and hence κ scales linearly with N for large N. In that case, the correct measure of the accuracy of a VWF is κ/N so that $\Omega = (0.99)^N$, for example, means that the VWF captures "per particle" 99% of the ground state. κ/N usually shows weak finite size effect upon increasing N. (v) In practice, we have found that the accuracy of κ is similar to that obtained for other physical quantities like density or density-density correlations but slightly less precise than that achieved for E_0. (Relative precisions of 10^{-4} or better are routinely obtained for the latter.)

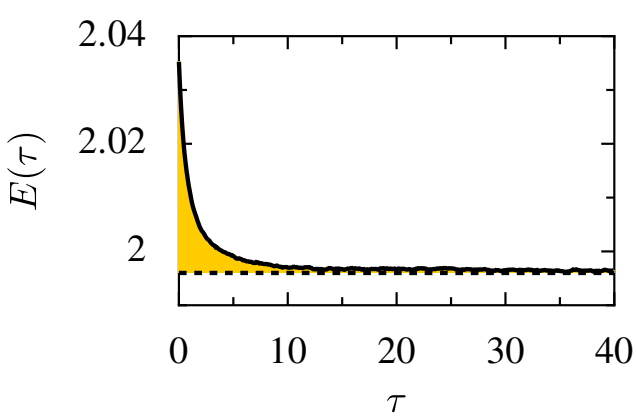

FIG. 1 (color online). Typical curve for the energy $E(\tau)$ as a function of the imaginary time τ. The thick line is a typical trace (with high precision) while the dashed line corresponds to the asymptotic value E_0. The logarithm of the overlap Ω is simply given by the shaded area between the two curves (see text). The data have been obtained for the model of dipolar bosons, defined Eq. (8), with very high precision.

Proof of Eq. (4).—The proof is straightforward. We introduce the pseudopartition function $Z(\tau) \equiv \langle \Psi_V | \Psi(\tau)\rangle = \langle \Psi_V | e^{-\tau H} | \Psi_V\rangle$, which, in analogy to the finite temperature partition function, is related to the energy through $E(\tau) = -\partial_\tau \log Z(\tau)$. Defining $\bar{\Psi}_0$ as the ground state normalized to unity, $\Psi(\tau)$ converges towards $\Psi(\tau) \to \sqrt{Z(2\tau)}\bar{\Psi}_0$ for large τ. Using the definition of Ω and $Z(\tau)$ one obtains

$$\Omega = \lim_{\tau\to\infty} \frac{[Z(\tau)]^2}{Z(0)Z(2\tau)}. \tag{5}$$

Further, from $E(\tau) = -\partial_\tau \log Z(\tau)$ one finds $\log[Z(\tau)/Z(0)] = -\int_0^\tau E(\tau)d\tau$ and $\log[Z(2\tau)/Z(\tau)] = -\int_\tau^{2\tau} E(\tau)d\tau$. In the latter, for τ large enough, $E(\tau) \sim E_0$ such that $Z(2\tau)/Z(\tau) \sim \exp(-E_0\tau)$. Collecting terms together in Eq. (5), we arrive at Eq. (4).

Link with mixed estimators.—One drawback of QMC calculations is that when the quantum average of an observable $\hat{A}$ is measured, the mixed estimate $A_{\rm MX} = \langle \Psi_V | \hat{A} | \Psi_0\rangle / \langle \Psi_V | \Psi_0\rangle$ naturally emerges. For some observables it is possible to actually calculate the correct quantity $A_0 = \langle \Psi_0 | \hat{A} | \Psi_0\rangle / \langle \Psi_0 | \Psi_0\rangle$ using, for instance, forward walking techniques [8]. However, this concerns mainly local quantities and does not apply to nonlocal ones. For the latter, one has to rely on the extrapolation formula, $A_0 \approx 2A_{\rm MX} - A_V$, where $A_V = \langle \Psi_V | \hat{A} | \Psi_V\rangle / \langle \Psi_V | \Psi_V\rangle$. This formula [9] is only valid when Ψ_0 is close to Ψ_V, otherwise it is uncontrolled. If $\hat{A}$ is definite positive, it takes the form $\hat{A} = d^\dagger d$ and the overlap allows one to obtain the following lower bound for A_0,

$$A_0 \geq \frac{[A_{\rm MX}]^2}{A_V}\Omega. \tag{6}$$

A particularly interesting example is the condensate frac-

tion of a Bose condensate where the operator $d^\dagger$ creates a particle in the $k = 0$ state. Unfortunately, this lower bound is only useful for small systems like atoms or molecules, or for extremely good VWFs, since the exponentially decreasing nature of the overlap quickly makes it meaningless. To prove Eq. (6), one simply writes Schwartz inequality, $|\langle C|B\rangle|^2 \leq \langle C|C\rangle\langle B|B\rangle$ with $|C\rangle = d|\Psi_0\rangle$ and $|B\rangle = d|\Psi_V\rangle$.

Link with the energy gap.—Another application of the calculation of the overlap is that it gives access to an upper bound value for the gap of the system, i.e., to information on the excitation spectrum. Indeed, it is straightforward to show that the difference Δ between the first excited state and E_0 obeys: [10]

$$\Delta \leq \frac{E_V - E_0}{1 - \Omega}. \tag{7}$$

Again, such an upper bound value is usually useless in the thermodynamic limit ($N \to \infty$) where there usually is a continuum of excitations. It can be used, however, for molecular or atomic systems. In this case, one should look for a wave function close to the first excited state. Such a VWF has a rather low energy but also a low overlap with the ground state. In practice, the VWF, being close to an eigenstate, has a low variance, and as the variance is $-\partial E/\partial\tau|_{\tau=0}$, the $E(\tau)$ curve converges slowly to the ground-state energy, hence resulting in a small overlap. The quality of this upper bound value depends on the projection of the VWF on the excited states other than the first one and the inequality becomes an equality when the VWF has projections only on the ground state and first excited state.

Application.—To illustrate the utility of the above discussion, we now turn to a specific example where the calculation of the overlap can be particularly useful. The system discussed below is close to a first order liquid-solid transition. On one hand, the chosen VWF is close to the liquid state so that optimizing the variance or the energy leads closer to the liquid state. On the other hand, the true ground state is the crystal as indicated by the calculation of the overlap. More precisely, we consider a system of N bosons in two dimensions with a repulsive dipolar interaction. This system is subject to current research as it is a candidate for a realizing a crystal in ultracold atom experiments [11,12]. A detailed study of the model will be presented elsewhere [13], and we focus here on the relevance of Ω. The scaled Hamiltonian takes the form

$$H = -\frac{1}{r_s}\sum_{i=1}^{N}\nabla_i^2 + 2\sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|^3}, \tag{8}$$

where r_S controls the relative strength of the dipolar interaction over the kinetic one. This model shows a first order transition at $r_S = r_S^* \approx 27$: for $r_S \leq r_S^*$ the system is in a Bose-Einstein phase while for $r_S \geq r_S^*$ the system crystallizes into a triangular lattice. Note that this crystal-Bose-Einstein transition has been discussed recently in Refs. [11,12]. We use a VWF of Bijl-Jastrow form:

$$\Psi_V(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \prod_{i=1}^{N}\Phi_1(\mathbf{r}_i)\prod_{j<k}\Phi_2(|\mathbf{r}_j - \mathbf{r}_k|), \tag{9}$$

where the Jastrow part includes two-body correlations, $\Phi_2(r) = \exp(-2\sqrt{\frac{r_s}{r}}e^{-r/A})$. The one-body part allows to break translational symmetry and interpolates from a condensatelike VWF to a crystal-like VWF. Noting Δy, the distance between lattice sites of the crystal along the y axis, we define the vector $\mathbf{q}_1 = (0, 2\pi/\Delta y)$ in the reciprocal lattice. Vectors $\mathbf{q}_{2/3}$ are obtained by rotating $\mathbf{q}_1$ by an angle of $2\pi/3$ and $-2\pi/3$. We choose a one-body trial function

$$\Phi_1(\mathbf{r}) = \prod_{i=1,2,3}[1 + \alpha\cos(\mathbf{q}_i \cdot \mathbf{r})], \tag{10}$$

whose maxima reproduce the triangular lattice expected for the crystal. This function is well suited to describe quantum melting since it interpolates between a flat liquid (BEC)-type pattern for $\alpha = 0$ to a triangular crystal form for $\alpha \neq 0$. Setting $r_S = 28.6$, i.e., slightly above the crystalline transition, we investigate our VWF as a function of the translational symmetry breaking parameter α. The results are presented in Fig. 2. The details of our algorithm can be found in Refs. [13,14]. In particular, we used the Green function Monte Carlo technique on a spatial grid with a filling factor $\nu = 1/56$ particles per site.

The system is clearly in a crystal state, as demonstrated by computing the static structure factor, and hence a good variational ansatz is expected for $\alpha \neq 0$. Figure 2 shows, however, that both the variational energy and variance indicate a minimum for $\alpha = 0$. On the other hand, κ

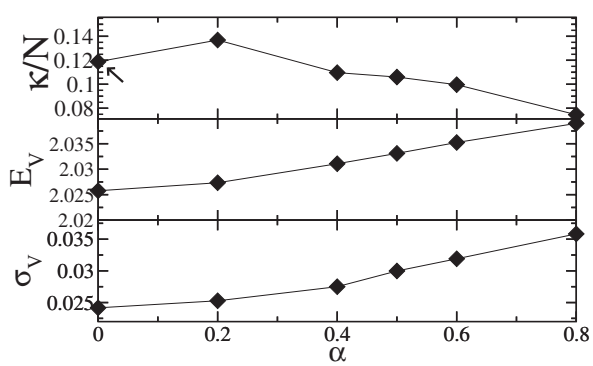

FIG. 2. Upper panel: parameter κ/N measuring the overlap between the guiding function and the actual ground state of the system as a function of the symmetry breaking parameter α. The arrow indicates that the point at $\kappa(\alpha = 0)$ is lower than the correct value, a small $\alpha \neq 0$ being needed for the simulation to fully converge to the ground state. Middle panel: variational energy E_V as a function of α. Lower panel: square root of the quantum variance σ_V^2 of the VWF as a function of the symmetry breaking parameter α. All data are taken for 32 bosons in a 32×56 grid at $r_S = 28.6$.

decreases with α indicating that the phase with $\alpha \neq 0$ is the correct one. This is an extreme case where the traditional criterion on the energy and variance actually gives a false answer even qualitatively while the overlap indicates the correct answer.

Conclusion.—It is interesting to note that both the standard criteria used to characterize a VWF (variational energy and variance), as well as the overlap discussed in the present Letter, are all different aspects of the energy-imaginary time $E(\tau)$ curve: the variational energy $E_V = E(\tau = 0)$, the variational variance $\sigma_V^2 = -\partial E/\partial\tau|_{\tau=0}$, and the overlap is the integral of $E(\tau)$. Hence the different criteria give informations on different characteristics of $E(\tau)$. While no general statement can be made, in many instances the VWFs used are a fair description of the ground state. In these cases $E(\tau)$ looks typically like the curve shown in Fig. 1 and, in particular, has a positive curvature. Assuming a positive curvature, we find that the various criteria are related to each other as

$$\Omega \leq \exp[-(E_V - E_0)^2/(2\sigma_V^2)], \tag{11}$$

obtained using the fact that $E(\tau)$ lies above its tangent at $\tau = 0$. Although the above inequality is not completely general, it gives a rough estimate of the overlap in many practical cases. One can note that strictly speaking, the above estimate decreases when the variance decreases. It is therefore imperative that the variance and variational energy are optimized simultaneously.

To conclude this Letter, we have shown that from the data given by usual QMC calculations, it is possible to extract relevant information on the quality of the variational wave function used. We stress that our main result, Eq. (4), applies generally to any projection technique in imaginary time and not only to QMC.

We thank O. Parcollet for fruitful discussions and S. Dhilon for many useful comments. C. M also acknowledges support from the DFG-SFB Transregio 12.

[1] B. K. Clark and D. M. Ceperley, Phys. Rev. Lett. **96**, 105302 (2006).
[2] P. W. Anderson, Science **235**, 1196 (1987).
[3] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
[4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[5] L. Reatto, Phys. Rev. B **26**, 130 (1982).
[6] G. L. Masserini and L. Reatto, Phys. Rev. B **30**, R5367 (1984).
[7] G. L. Masserini and L. Reatto, Phys. Rev. B **35**, 6756 (1987).
[8] M. Calandra Buonaura and S. Sorella, Phys. Rev. B **57**, 11 446 (1998).
[9] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Comput. Phys. **96**, 258 (1991).
[10] C. Eckert, Phys. Rev. **36**, 878 (1930).
[11] G. E. Astrakharchik, J. Boronat, I. L. Kurbakov, and Yu. E. Lozovik, Phys. Rev. Lett. **98**, 060405 (2007).
[12] H. P. Büchler, E. Dernler, M. Lukin, A. Micheli, N. Prokof'ev, G. Pupillo, and P. Zoller, Phys. Rev. Lett. **98**, 060404 (2007).
[13] C. Mora, O. Parcollet, and X. Waintal, arXiv:condmat/0703620.
[14] X. Waintal, Phys. Rev. B **73**, 075417 (2006).