Computing physical properties with quantum Monte Carlo methods with statistical fluctuations independent of system size

Roland Assaraf*

Laboratoire de Chimie Théorique, CNRS-UMR 7616, Université Pierre et Marie Curie Paris VI, Case 137, 4, place Jussieu 75252 PARIS Cedex 05, France

(Received 28 July 2014; published 24 December 2014)

We show that the recently proposed correlated sampling without reweighting procedure extends the locality (asymptotic independence of the system size) of a physical property to the statistical fluctuations of its estimator. This makes the approach potentially vastly more efficient for computing space-localized properties in large systems compared with standard correlated methods. A proof is given for a large collection of noninteracting fragments. Calculations on hydrogen chains suggest that this behavior holds not only for systems displaying short-range correlations, but also for systems with long-range correlations.

DOI: 10.1103/PhysRevE.90.063317 PACS number(s): 05.10.Ln, 31.15.−p, 02.70.Ss, 05.45.Xt

Quantum Monte Carlo (QMC) methods are stochastic techniques used to solve the Schrödinger equation. Like any Monte Carlo method, they rely on statistical sampling to compute large-dimensional integrals. This enables better scaling (both in CPU time and memory requirements) as a function of system size than deterministic methods. Further, these methods can be easily and efficiently implemented on parallel computers, making them an excellent choice for solving many-body problems. Many applications can be found in nuclear physics [1,2], condensed matter physics [3,4], or quantum chemistry [5].

In practice, accurate (small systematic error) and precise (small statistical uncertainty) total ground-state energies can be computed with these methods [6,7]. Ground-state energies are key ingredients to compute the physical properties of interest (dissociation energies, electron affinity, forces, etc.), because the latter can be expressed as energy differences $E_\lambda - E_0$, where E_0 and E_λ are the ground-state energies of two closely related Hamiltonians H_0 and H_λ,

$$H_\lambda = H_0 + \lambda O, \tag{1}$$

where λ is a small perturbation parameter. In addition, $E_\lambda - E_0$ provides an estimate of the expectation value of the observable O in a first-order expansion in λ (Hellmann–Feynman theorem).

In practice, for fermionic systems we compute approximations to ground-state energies. Starting from this point, E_λ and E_0 will also stand for approximations (e.g., variational or fixed-node approximations) of the exact energies of H_λ and H_0, respectively. If E_λ and E_0 are accurate (small systematic errors), then $E_\lambda - E_0$ is an accurate estimate of the energy difference. However, since $E_\lambda - E_0$ is small, the issue is then to compute $E_\lambda - E_0$ with a sufficiently small statistical uncertainty. As we discuss now, this is not always easy.

There are typically three strategies to compute exactly energy differences (or observables) in QMC. (i) The difference $E_\lambda - E_0$ can be directly and simply obtained from independent calculations of E_λ and E_0. (ii) The most popular techniques are based on the idea of correlated sampling with reweighting [8,9]: the algorithm builds a single sample to compute both E_λ and E_0. Weighting factors are introduced in the expectation values to avoid any systematic error (see below). Different techniques can be classified in this category, like zero-variance zero-bias improved estimators in variational Monte Carlo [10]. The forward walking method [11] can also be reformulated by using this point of view [12]. (iii) A third strategy (to compute the expectation value of an observable) makes use of the Hellmann–Feynman theorem, sampling (in principle) the exact ground-state distribution ϕ_0^2 and averaging O on that sample (no weighting factors are introduced since we sample the right distribution). This is done by using QMC methods based on explicit path integral sampling, like reptation Monte Carlo or path integral ground-state Monte Carlo [13,14].

Beyond obtaining $E_\lambda - E_0$ with no (additional) systematic error, an efficient energy difference method should fulfill the following three conditions: First, the statistical uncertainty has to be proportional to λ. This is necessary to obtain a finite statistical uncertainty on the energy derivative. This feature is satisfied by (ii) and (iii) but not by (i). Second, it should obey the zero-variance principle, i.e., in the limit of exact wave function and possibly its parameter derivatives, the variance should vanish. This principle provides flexibility to reduce statistical fluctuations. This is of major importance when, for example, O has infinite fluctuations (as is the case for forces on nuclei, or the histogram estimator for the density in the limit of vanishingly small bin size). The zero-variance principle is satisfied by (i) and (ii) but not by (iii). These two criteria are well documented in the literature. We emphasize here that they should be completed by a third one: the locality property for the statistical fluctuations. Many important observables or properties are local—they depend on a few degrees of freedom. For example, the force on a nucleus is generally not influenced by a distant noninteracting fragment. One would like to use a method such that the statistical uncertainty of the estimators of local properties does not depend on distant degrees of freedom. This condition is satisfied by (iii) but not by (i) and (ii) [10].

Calculations of local properties with the methods (ii) and (iii) get less precise when the system size N gets larger. This is because expectation values saturate as a function of N while the statistical uncertainty σ has the same behavior as the energy: $\sigma \sim \sqrt{N/M}$ where M is the number of independent Monte Carlo configurations. The relative error

*assaraf@lct.jussieu.fr

is then proportional to $\sqrt{N}$ for a given value of M. Extensive observables do not have this problem. The relative error is proportional to $1/\sqrt{N}$ for a given size M of the sample. Nevertheless, when methods (ii) and (iii) are employed, large statistical fluctuations may also come from irrelevant degrees of freedom: e.g., core electrons have little influence on a dipole moment or a polarizability, but a large contribution to σ. Note that local properties carry more detailed information than their nonlocal counterparts—the dipole moment of a finite system is the sum of the dipole moments of its fragments, a one-body observable can be computed from the one-body density, and so on. This justifies why the local property of the statistical fluctuations is an important efficiency criterion. So far, locality has been exploited in QMC to reduce the computational cost of the sampling [15], but, not the statistical fluctuations.

Recently, a method has been proposed to compute small energy differences: the correlated sampling method with no reweighting [16]. It was introduced mainly to compute the derivative of the diffusion Monte Carlo energy (one of the best estimates one can have in QMC) with no bias and finite statistical fluctuations. We show here that this method obeys the locality property for the statistical uncertainty. This makes the method the first one to fulfill all three conditions we enumerated. We present a proof for a set of noninteracting systems and show numerically that this property extends to interacting systems by using simulations on chains of hydrogen atoms.

We now give a sketch of the correlated sampling method with and without reweighting to put them in perspective. In QMC, E_λ, the ground-state energy of the Hamiltonian H_λ, is computed as the expectation value of a random variable e_λ over some known probability measure $\pi_\lambda(\mathbf{R})$,

$$E_\lambda = \langle e_\lambda(\mathbf{R})\rangle_{\pi_\lambda}. \tag{2}$$

The nature of E_λ, π_λ, e_λ, and $\mathbf{R}$ might depend on the variant of the QMC method. For example, in the context of the variational Monte Carlo method (VMC), E_λ is the variational energy, i.e., the (quantum) expectation value of H_λ on a variational wave function ψ_λ, $\pi_\lambda \equiv \psi_\lambda^2$, $e_\lambda \equiv H_\lambda\psi_\lambda/\psi_\lambda$ (the so-called local energy), and $\mathbf{R}$ usually stands for the $3N$ coordinates of the N particles. In other methods, $\mathbf{R}$ might stand for a Feynman–Kac path [13,17] or a determinant of molecular orbitals [7,18–20].

Correlated sampling methods with reweighting (CSRs) typically compute $E_\lambda - E_0$ as

$$\begin{aligned} E_\lambda - E_0 &= \langle e_\lambda(\mathbf{R})w_\lambda(\mathbf{R}) - e_0(\mathbf{R})\rangle_{\pi_0} \\ &= \langle e_\lambda - e_0\rangle_{\pi_0} + \mathrm{cov}(e_\lambda, w_\lambda), \end{aligned} \tag{3}$$

where the weight w_λ is

$$w_\lambda(\mathbf{R}) \equiv \frac{\frac{\pi_\lambda(\mathbf{R})}{\pi_0(\mathbf{R})}}{\left\langle\frac{\pi_\lambda}{\pi_0}\right\rangle_{\pi_0}}. \tag{4}$$

The nonlocality of the fluctuations comes from the second term of expression (3) [10]. This is because the weight w_λ is multiplied by the local energy whose variance grows with N.

The derivative of expression (3) provides the expectation value of the observable O. The derivative of the first term has the same value as the Hellmann–Feynman (HF) estimate. The derivative of the second term is called the Pulay correction. The latter is zero for the exact ground state, or if the wave function has been optimized in a sufficiently large parameter space. However, this is frequently not the case. Although robust and efficient methods have been recently developed for optimizing all the parameters to minimize the VMC energy [21,22], these methods rely on evaluating derivatives of the variational energy, using again the expression (3), and are less precise when N is larger. One way to reduce the fluctuations is to use the space-warp transformation [9,23]. When a nucleus is displaced, the electrons move in such a way that those close to the nucleus move almost rigidly with the nucleus and those far away move very little. Besides increasing the correlation between e_λ and e_0, this reduces the fluctuations in the weights in Eq. (3). Although smaller, the statistical fluctuations still increase with system size.

We now discuss the method that does cure this problem. The correlated sampling method with no reweighting (CSNR) [16] obtains $E_\lambda - E_0$ as follows:

$$E_\lambda - E_0 = \langle e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0)\rangle. \tag{5}$$

First, $\mathbf{R}_\lambda$ and $\mathbf{R}_0$ sample π_λ and π_0, respectively. Second, the pair $(\mathbf{R}_0,\mathbf{R}_\lambda)$ appearing in Eq. (5) is correlated in such a way that $\langle(\mathbf{R}_\lambda - \mathbf{R}_0)^2\rangle \propto \lambda^2$ (with a finite prefactor). These two ingredients ensure, respectively, that there is no bias, and that the variance of $e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0)$ is of order λ^2 (for a finite precision on the derivative). These two properties of the CSNR method are shared by the CSR method, but we will demonstrate that the CSNR has the additional desirable property that the statistical error of local observables rapidly saturates as a function of system size.

The configurations $\mathbf{R}(t)$ and $\mathbf{R}_\lambda(t)$ are produced by two (close) stochastic processes having a stability property versus chaos, or a so-called synchronization behavior: Two trajectories having different initial conditions but sharing the same sequence of random numbers coalesce sufficiently fast in time. For example, the overdamped Langevin process, described by the dynamics

$$\mathbf{R}_\lambda(t+dt) = \mathbf{R}_\lambda(t) + \mathbf{b}_\lambda dt + d\mathbf{W}_\lambda, \tag{6}$$

where $\mathbf{b}_\lambda = \nabla\ln(\pi_\lambda)$ and $d\mathbf{W}_\lambda$ is a Wiener process, sometimes has this stability property. With this choice, the CSNR method performs two overdamped Langevin processes corresponding to $\lambda = 0$ and $\lambda \neq 0$, but with $d\mathbf{W}_\lambda = d\mathbf{W}_0$ (same sequence of random numbers).

Locality property. We now prove that the statistical fluctuations of the estimator $e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0)$ in Eq. (5) has the locality property under the strong hypothesis of a fully separable system and a strictly local perturbation. This means that a physical configuration $\mathbf{R}$ can be split into two sets $\mathbf{R} = (\mathbf{R}^l,\mathbf{R}^u)$ such that the degrees of freedom $\mathbf{R}^l$ and $\mathbf{R}^u$ are not coupled by the Hamiltonian. Mathematically, the latter is a direct sum

$$H_\lambda = H_\lambda^l + H^u, \tag{7}$$

where H_λ^l (H^u) act only on the degrees of freedom $\mathbf{R}^l$ ($\mathbf{R}^u$). $\mathbf{R}^l$ and $\mathbf{R}^u$ represent coordinates of electrons belonging to two noninteracting (distant) molecules, $\mathbf{R}^l = (\mathbf{r}^1,\ldots,\mathbf{r}^L)$ and $\mathbf{R}^u = (\mathbf{r}^{L+1},\ldots,\mathbf{r}^N)$. The strict locality of the physical property is manifest in that only H_λ^l is parametrized by λ, not H^u. For such a system made of two independent fragments, the stochastic dynamics followed by $\mathbf{R}^l$ and $\mathbf{R}^u$ are required

to be independent: the steady state of the sampling process is then a product:

$$\pi_\lambda(\mathbf{R}) = \pi^l_\lambda(\mathbf{R}^l)\pi^u(\mathbf{R}^u). \tag{8}$$

Suppose, for example, that $\pi_\lambda(\mathbf{R})$ is obtained from the familiar overdamped Langevin process. Given Eq. (8), the drift is

$$\mathbf{b}_\lambda(\mathbf{R}) = \mathbf{b}^l_\lambda(\mathbf{R}^l) + \mathbf{b}^u(\mathbf{R}^u), \tag{9}$$

implying that the process indeed produces independent samples for $\mathbf{R}^l$ and $\mathbf{R}^u$.

Regarding the estimator of the energy, the latter should be the sum of two independent contributions

$$e_\lambda(\mathbf{R}) = e^l_\lambda(\mathbf{R}^l) + e^u(\mathbf{R}^u). \tag{10}$$

This condition is fulfilled when e_λ is the local energy (VMC method, reptation method), and the variational function $\psi_\lambda(\mathbf{R})$ obeys the same separation property as the exact ground state, i.e., $\psi_\lambda(\mathbf{R}) = \psi^l_\lambda(\mathbf{R}_l)\psi^u(\mathbf{R}^u)$. Since the stochastic process does not couple the variables $\mathbf{R}^l$ and $\mathbf{R}^u$, the dynamics on the variables $\mathbf{R}^l$ for an overdamped Langevin process is

$$\mathbf{R}^l_\lambda(t+dt) = \mathbf{R}^l_\lambda(t) + \mathbf{b}^l_\lambda\big[\mathbf{R}^l_\lambda(t)\big]dt + d\mathbf{W}^l, \tag{11}$$

$$\mathbf{R}^l_0(t+dt) = \mathbf{R}^l_0(t) + \mathbf{b}^l_0\big[\mathbf{R}^l_0(t)\big]dt + d\mathbf{W}^l. \tag{12}$$

In a similar way, the dynamics on the variables $\mathbf{R}^u$ is

$$\mathbf{R}^u_\lambda(t+dt) = \mathbf{R}^u_\lambda(t) + \mathbf{b}^u\big[\mathbf{R}^u_\lambda(t)\big]dt + d\mathbf{W}^u, \tag{13}$$

$$\mathbf{R}^u_0(t+dt) = \mathbf{R}^u_0(t) + \mathbf{b}^u\big[\mathbf{R}^u_0(t)\big]dt + d\mathbf{W}^u. \tag{14}$$

Hence the stochastic processes (13) and (14) are the same because $\mathbf{b}^u$ does not depend on λ. Besides, they share a common Wiener process. Because of the stability property with respect to chaos, there is a finite time beyond which $\mathbf{R}^u_\lambda(t) = \mathbf{R}^u_0(t)$ to any level of accuracy. Using this last identity and expression (10), the estimator of the energy difference appearing in the brackets of Eq. (5) is

$$e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0) = e^l_\lambda\big(\mathbf{R}^l_\lambda\big) - e^l_0\big(\mathbf{R}^l_0\big), \tag{15}$$

where of course the configurations arising in the identity (15) are evaluated at the same time t. In conclusion, the energy difference is an expectation value of a random variable $e_\lambda - e_0$ depending only on the coordinates $(\mathbf{R}^l, \mathbf{R}^l_\lambda)$ and not on $\mathbf{R}^u$. As a direct consequence, the fluctuations of $e_\lambda - e_0$ depend also on $(\mathbf{R}^l, \mathbf{R}^l_\lambda)$ and not on $\mathbf{R}^u$. This ends the proof. Note that the wave function ψ_λ should in general be symmetrized (bosons) or antisymmetrized (fermions). However, this constraint does not modify the proof when all the particles of the two separate subsystems (represented by the two Hamiltonians H^l_λ and H^u) are localized in different nonoverlapping regions of space. This is because the symmetrization or the antisymmetrization of ψ_λ can still be written as a product. For example, for fermions, the antisymmetrization of the product $\psi_\lambda(\mathbf{R}) = \psi^l_\lambda(\mathbf{R}^l)\psi^u(\mathbf{R}^u)$ can be written as

$$A\psi_\lambda(\mathbf{R}) \propto \big[A\psi^l_\lambda(\mathbf{R}^l)\big][A\psi^u(\mathbf{R}^u)]. \tag{16}$$

The identity (16) holds because all terms corresponding to exchanges of two particles lying in the two separate regions of space are zero.

In practice, the time step dt of the overdamped Langevin dynamics can be small but finite. This is at the origin of a finite time-step error which can be hardly avoided in diffusion Monte Carlo but can be suppressed in variational Monte Carlo by using the acceptance-rejection method [24]. A difficulty arises with such method: one walker can be accepted (e.g., $\mathbf{R}_\lambda$) and the other can be rejected (e.g., $\mathbf{R}_0$). This introduces undesirable weights in the expectation values, and the expression (5) has to be modified accordingly. The fluctuations of the weights can be arbitrarily small by choosing smaller time steps but still increase with system size. Here, in our applications, we avoid these weights by modifying the acceptance-rejection method as follows: whenever one walker should be accepted and not the other, we reject both walkers. This introduces a systematic error, which can be easily controlled (by using smaller time steps). In addition, it obviously fulfills the local property (unlike the statistical fluctuations coming from the weights).

The only hypotheses we made are the stability with respect to chaos of the processes, and the independence of the variables $\mathbf{R}^l_\lambda(t)$ and $\mathbf{R}^u(t)$. Besides the overdamped Langevin process, any stochastic process fulfilling these two conditions will produce statistical fluctuations obeying this local property.

Numerical results. In practice, $\mathbf{R}^l(t)$ and $\mathbf{R}^u(t)$ might be weakly correlated rather than exactly independent. Also, the perturbation may not be strictly local. As a test bed we choose linear chains of hydrogen atoms, since large sizes can be handled easily, and the popular overdamped Langevin process, with the usual guiding functions such as a restricted Hartree–Fock (RHF) solution (with or without a Jastrow factor), displays numerically the required synchronization behavior [16]. We use the same geometries and observable as Assaraf *et al.* [10]. (i) The first geometry minimizes the Hartree–Fock energy; it consists of H_2 molecules (interatomic distances $\sim$1.4 au) separated by $\sim$6.5 au. This system is nonmetallic, displaying short-range correlations. (ii) The second geometry consists of equally separated hydrogen nuclei (1.4 au). It is known to have metallic properties, i.e., long-range correlations between electrons. We compute the derivative of the energy with respect to the first nuclear position along the axis of the chain (force), using finite differences: The secondary system (parameter $\lambda \neq 0$) corresponds to the first atom displaced by $\lambda = 0.0001$ au. For easy checks, results are presented in the VMC framework with, unless specified otherwise, a monodeterminantal variational wave function obtained from a self-consistent field calculation (SCF).

Expectation values. Expectation values of the force are shown in Fig. 1. The Hellmann–Feynman (HF) force should agree with the derivative of the variational energy on a RHF solution but, as shown in the figure, the (HF) force obviously has a large bias on its SCF finite-basis approximation. The Pulay correction is not small and must be taken into account.

The CSNR estimator (5) turns out to be in perfect agreement with the SCF force. This is consistent with the theory. This estimator takes into account the Pulay correction but, as we see now, it does not have the drawback of a direct calculation of this contribution.

Locality of the statistical fluctuations. We report the probability density function (pdf) of the random variable $e_\lambda(\mathbf{R}_\lambda) - e_0(\mathbf{R}_0)$ for the metallic chain in Fig. 2. As a function of N, the pdf appears to have almost converged when $N = 24$ and fully converged when $N = 48$. This behavior is of course

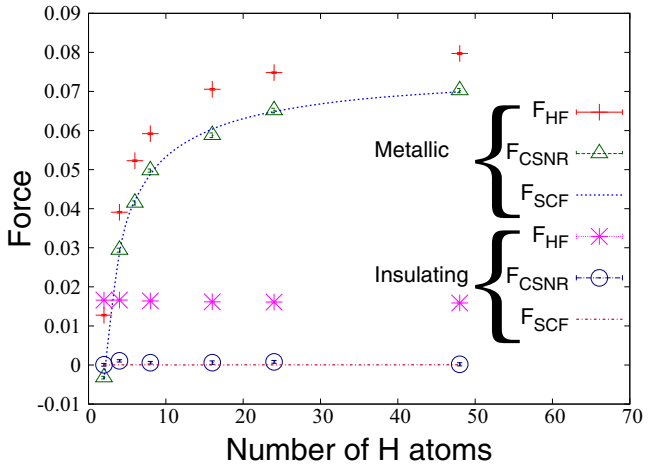

FIG. 1. (Color online) Forces computed from SCF, the correlated sampling with no reweighting (CSNR) estimator, and the Hellmann–Feynman (HF) estimates for insulating and metallic chains. The CSNR force agrees with the SCF force, but the HF force does not—the Pulay term is not small.

not trivial since the metallic chain is far from the simple separable model discussed above. The observable being the Coulomb force, it is not strictly localized either (it decays like the square inverse of the distance to the first atom). For the insulating chains, the pdfs are numerically the same for all sizes considered (N ranging from 2 to 48) and are not reported here. We conclude that any estimate of the statistical uncertainty will not depend asymptotically on N, for these systems. In other words, the locality property for the statistical uncertainty extends to systems composed of interacting fragments with non-strictly-local perturbations.

Remark about the pdf. The pdf reported in Fig. 2 has heavy tails: it asymptotically behaves as $1/|x|^{2.5}$, implying a infinite variance. This is because the cusp conditions are not fulfilled by the SCF trial function. This behavior disappears when a

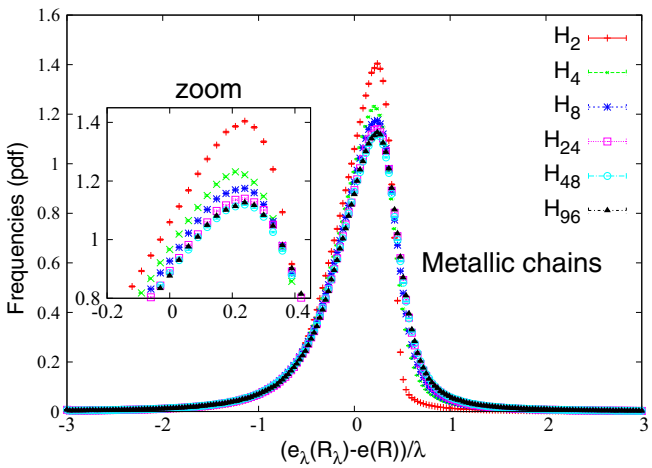

FIG. 2. (Color online) Probability density function (pdf) of the estimator of the energy derivative in the CSNR method. Even for these metallic chains, the pdf rapidly converges with chain length.

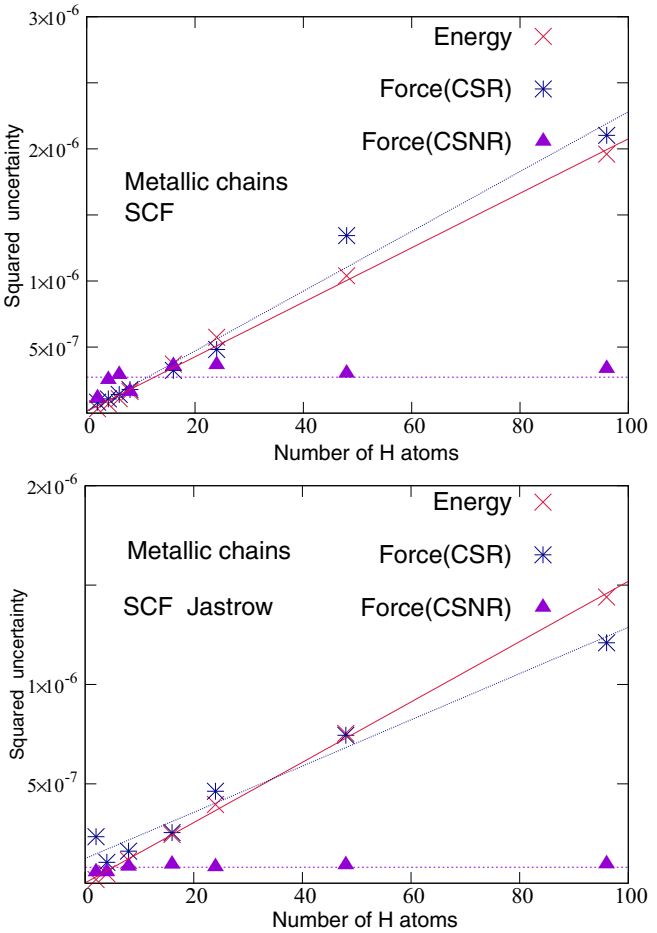

FIG. 3. (Color online) Squared uncertainties, σ^2, in the variational energy and forces. Estimates are made from simulations with calculation times of 4000 au and with 400 walkers. (top) SCF wave function. (bottom) SCF function times a minimal Jastrow factor. The straight lines are linear fits. The value of σ^2 for the energy and CSR forces increases linearly with chain length whereas that for the CSNR forces rapidly saturates with chain length.

correlation factor (Jastrow) is introduced:

$$\psi(\mathbf{R}) = \psi_{SCF}(\mathbf{R}) \exp\left(\sum_{1\geqslant i<j\leqslant N} \alpha_{ij} \frac{r_{ij}}{1+r_{ij}}\right),$$

where r_{ij} is the distance between the two electrons i, j, and $\alpha_{ij} = 0.5$ (0.25) if electrons i and j have different (the same) spins. With this variational function, we found that the pdf asymptotically behaves as $1/|x|^4$, and the estimator has a finite variance. Besides, with a Jastrow factor, the pdf saturates as a function of N in the same way.

For the sake of comparison, we report statistical uncertainties squared, σ^2, in Fig. 3. For the CSNR estimator, σ^2 saturates as a function of N, as expected. This has to be compared with the CSR estimator (3) for which $\sigma^2(N)$ has an asymptotically linear growth, similar to the energy estimator. When the Jastrow factor is included, the CSNR estimator is already better than the CSR estimator for $N = 2$ atoms. For $N = 96$ the gain is about a factor of 10.

Locality of the dynamics. We have seen in the proof that the locality of the statistical uncertainty in the CSNR method came from the fact that the coordinates of the particles which are unaffected by the perturbation coalesce after a given time $\mathbf{R}^u_\lambda(t) = \mathbf{R}^u_0(t)$ [Eqs. (14) and (13)]. This was proved for a fully separable system and a strictly local perturbation. We now address how this property extends when these conditions are not fulfilled. For that purpose, we extract from the configuration $\mathbf{R}_0(t)$ the particles i which belong to a given region V of the space. We then introduce a (local) quadratic distance (LQD), between these particles and those in $\mathbf{R}_\lambda(t)$ having the same index:

$$\mathrm{LQD} = \frac{1}{\lambda^2}\left\langle \sum_{i/\mathbf{r}^i_0 \in V} \left|\mathbf{r}^i_\lambda - \mathbf{r}^i_0\right|^2 \right\rangle. \tag{17}$$

This quantity is a measure of the influence of the perturbation on the stochastic dynamics in the region V. Indeed for the fully separable model with a strictly localized perturbation, the LQD is exactly zero if V is included in the region described by H^u. The regions of space we consider are spherical shells centered on the first nucleus. The LQDs are reported in Fig. 4 for the metallic and the insulating chains as a function of the (inner) radius of the shell. First, the LQDs appear to converge as a function of N for both systems. Second, it decays exponentially for the insulating chain and algebraically for the metallic chains ($\propto 1/r^{3(\pm 0.1)}$) [25]. These are important results because it suggests that the stability property against chaos is itself a local property. This means that if the dynamics is stable for isolated fragments (a small molecule, a crystal cell, . . .), it is likely to be stable even when these fragments are weakly correlated. The overdamped Langevin process was found to be unstable for molecules with larger atoms, such as alkanes or lithium clusters, but the typical time when the instability occurs seems to be independent of the size (1 au for alkanes and 10 au for lithium clusters). Statistical fluctuations on the force calculation appear to be independent of the system size during this transient stable regime. This suggests that the locality property of the statistical uncertainty holds for many systems, as soon as the dynamics is nonchaotic. Developing stable stochastic dynamics with respect to chaos in VMC and other variants of QMC methods appears to be a promising path to solve the long-standing problem of accuracy and precision for properties, since the CSNR approach fulfills at the same time the three necessary conditions for efficiency enumerated in the introduction.

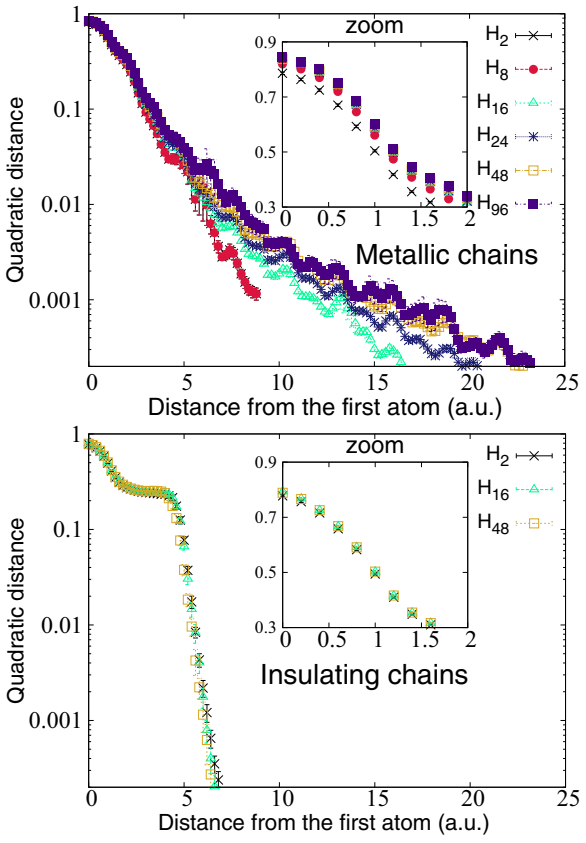

FIG. 4. (Color online) Quadratic distance between two electrons of $\mathbf{R}(t)$ and $\mathbf{R}_\lambda(t)$ having the same index versus distance to first (displaced) atom.

We warmly thank Julien Toulouse for reading the paper and Cyrus Umrigar for useful discussions and his highly significant contribution in improving the text.

[1] S. C. Pieper and R. B. Wiringa, Annu. Rev. Nucl. Part. Sci. **51**, 53 (2001).
[2] M. A. Hohensee, H. Müller, and R. B. Wiringa, Phys. Rev. Lett. **111**, 151102 (2013).
[3] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[4] M. Capello, F. Becca, S. Yunoki, and S. Sorella, Phys. Rev. B **73**, 245116 (2006).
[5] B. M. Austin, D. Y. Zubarev, and W. A. Lester, Chem. Rev. (Washington, DC, U.S.) **112**, 263 (2012).
[6] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, J. Chem. Phys. **136**, 124116 (2012).
[7] G. H. Booth and A. Alavi, J. Chem. Phys. **132**, 174104 (2010).
[8] J. Vrbik and S. M. Rothstein, J. Chem. Phys. **96**, 2071 (1992).
[9] C. Filippi and C. J. Umrigar, Phys. Rev. B **61**, R16291 (2000).
[10] R. Assaraf and D. Domin, Phys. Rev. E **89**, 033304 (2014).
[11] K. S. Liu, M. K. Kalos, and G. V. Chester, Phys. Rev. A **10**, 303 (1974).
[12] R. Assaraf and M. Caffarel, J. Chem. Phys. **119**, 10536 (2003).
[13] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).
[14] A. Sarsa, K. E. Schmidt, and W. R. Magro, J. Chem. Phys. **113**, 1366 (2000).
[15] A. J. Williamson, R. Q. Hood, and J. C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).

[16] R. Assaraf, M. Caffarel, and A. C. Kollias, Phys. Rev. Lett. **106**, 150601 (2011).
[17] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
[18] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003).
[19] F. R. Petruzielo, A. A. Holmes, H. J. Changlani, M. P. Nightingale, and C. J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012).
[20] R. Baer, M. Head-Gordon, and D. Neuhauser, J. Chem. Phys. **109**, 6219 (1998).
[21] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).
[22] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).
[23] C. J. Umrigar, Int. J. Quant. Chem. **36**, 217 (1989).
[24] W. Hastings, Biometrika **57**, 97 (1970).
[25] We notice small oscillations on top of the algebraic behavior: the minima are found to be close to the nuclei. This is consistent with a weaker influence of the perturbation in attractive regions of space.