Efficient and robust quantum Monte Carlo estimate of the total and spin electron densities at nuclei

P. Håkansson[a)] and Massimo Mella[b)]
School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

The computational performance of two different variational quantum Monte Carlo estimators for both the electron and spin densities on top of nuclei are tested on a set of atomic systems containing also third-row species. Complications due to an unbounded variance present for both estimators are circumvented using appropriate sampling strategies. Our extension of a recently proposed estimator [Phys. Rev. A **69**, 022701 (2004)] to deal with heavy fermionic systems appears to provide improved computational efficiency, at least an order of magnitude, with respect to alternative literature approaches for our test set. Given the importance of an adequate sampling of the core region in computing the electron density at a nucleus, a further reduction in the overall simulation cost is obtained by employing accelerated sampling algorithms. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2982930]

I. INTRODUCTION

Quantum Monte Carlo (QMC) is a set of powerful tools, useful for solving the Schrödinger equation for atoms and molecules.[1] Requiring the least computational resources, variational Monte Carlo (VMC) represents the "workhorse" of this family. Armed with the recently acquired capability of optimizing a correlated trial wave function (TWF) Ψ_T by minimizing its energy,[2–5] VMC has been shown to provide accurate estimates for the ground state properties of electronic systems.[5–9]

Not being constrained to the use of Gaussian atomic basis sets, VMC may also represent a valuable tool in computing nondifferential properties such as electron and spin densities on top of nuclei. The latter quantities are, respectively, defined as a function of the expectation values $\langle \delta(r_{i\mu}) \rangle$ and $\langle \delta_\sigma(r_{i\mu}) \rangle$, where δ is the Dirac delta, $r_{i\mu}$ is the distance between the ith electron of spin σ (either α or β) and the nucleus μ. In practical applications, accurate values for $\langle \delta(r_\mu) \rangle$ and $\langle \delta_\alpha(r_\mu) \rangle - \langle \delta_\beta(r_\mu) \rangle$ are needed to compute the isotropic Fermi contact term of the electron spin resonance (ESR) hyperfine coupling in systems containing heavy atoms.[10] In this context, VMC may serve either as a standalone technique or as a prerequisite for the use of diffusion Monte Carlo (DMC) when highly accurate results are needed. In this latter case, VMC results are combined with DMC computed weights to obtain true ground state properties, although within the framework of the fixed node approximation.[11]

When using QMC to compute the properties of electronic systems, a good level of accuracy may be expected thanks to the fact that electron correlation is to a large extent included. In dealing with electron-nucleus coalescence prob-

ability, however, one must also tackle the difficulty of estimating this infinitely local property with a method that provides only a statistical representation of the wave function (or, for what matters, of the electron density ρ). Thus, initial attempts at estimating $\langle \delta(r_\mu) \rangle$ were carried out by enclosing the nucleus of interest (μ) within spheres of decreasing radius, computing the probability density of finding an electron within each sphere, and finally extrapolating this set of results to zero radius.[12] While practical, this approach is, however, flawed by an inherently unbounded variance. A more sophisticated variant of this approach exploited the weak limit properties of a set of Gaussian functions as a possible way to improve the precision of $\langle \delta(r_{i\mu}) \rangle$, but the variance was shown to be unbounded for this method also.[13]

Due to the importance of estimating the Dirac delta expectation value, several researchers have sought more efficient and better behaving QMC estimators addressing the problem with the infinite variance. In this respect, a possible way forward is represented by the use of the differential identity $4\pi\delta(R) = -\nabla^2(1/|R|)$, the latter leading to a global estimator for $\langle \delta(r_\mu) \rangle$ that is easily implemented using precomputed quantities. Such an estimator, however, still suffers from an unbounded variance due to the presence of a term proportional to $1/|R|^2$, a difficulty that can be circumvented by sampling the alternative distribution $\Psi_T^2(\Sigma_i 1/r_{i\mu}^2)$.[14,15] This, however, assures a bounded variance only for the nucleus μ. More recently, Assaraf *et al.*[16] introduced the use of an auxiliary function $f(r_{i\mu})$ to regularize the differential identity without modifying the sampling distribution. In their proposal, one computes the contribution to the density on top of the nucleus μ due to the ith electron using $4\pi\delta(r_{i\mu}) = -(\nabla_i^2[f(r_{i\mu})]/r_{i\mu})$, where f satisfies the condition $\lim_{r_{i\mu} \to 0} f(r_{i\mu}) = 1$. With a suitable choice of f, the variance of the differential estimator becomes bounded apparently eliminating the necessity of using alternative sampling distributions.

A completely different approach to the task of comput-

---

a)Current address: Computational Engineering and Design Group, School of Engineering Sciences, University of Southampton, University Road, Southampton SO17 1BJ, United Kingdom.
b)Electronic mail: mellam@cardiff.ac.uk.

129, 124101-1

ing $\langle\delta(r_{i\mu})\rangle$ was proposed by Alexander et al.[17] In the latter, M independent electronic configurations are generated from an analytically normalizable distribution $g(\mathbf{R})$ and used to estimate the normalization of Ψ_T by means of

$$(N_T^2)^{-1} = M^{-1}\sum_i^M \psi_T^2(\mathbf{R}_i)/g(\mathbf{R}_i). \tag{1}$$

The Dirac delta expectation value is subsequently computed by using the normalized Ψ_T and sampling a subset of the configuration space with one of the electrons sitting on top of a nucleus. Although original, this method was found to have serious shortcomings for many electron atoms, these being mainly due to the lack of importance sampling. The latter was introduced in a later work,[18] with the sampling of the region near a nodal surface being de-emphasized by selecting a single configuration from a small subset of points obtained from a Hartree–Fock wave function. The insufficient sampling at short electron-electron r_{ij} and electron-nucleus distances was subsequently addressed in Ref. 19 by introducing a weight that approaches infinity for $r_{ij}, r_{i\mu} \rightarrow 0$. Despite all these improvements, this technique has never been applied to heavy atoms and appears to be useful only for first row atoms and ions.

An alternative scheme for evaluating the expectation value of the Dirac delta operator, one which is, in principle, exact as the approaches discussed above, was introduced by Chiesa et al.[20] to compute the electron-positron annihilation probability. This method, based on an analytical transformation of the original expectation value that does not require the use of a differential identity, makes also $\langle\delta(r_{ij})\rangle$ suitable for computation using a straightforward VMC simulation. Bearing in mind the relevance of $\langle\delta(r_{i\mu})\rangle$ for the calculation of the hyperfine coupling constants in open shell systems, the aim of this work is to extend the "local" VMC estimator introduced in Ref. 20 for a possible application in the field and to test its computational efficiency. In doing this, our long term goal would be to build a robust approach capable of tackling first-row transition metal compounds free of the difficulties introduced by the use of Gaussian basis sets and spin contaminated wave functions.

Due to the good performance shown in computing the electron density of systems containing light atoms, the renormalized "differential" estimator introduced in Ref. 16 could also represent a suitable approach to the same task. Thus, a numerical comparison between the two estimators has been carried out in this work, aiming also to gain a better understanding of the computational effort involved in estimating $\langle\delta(r_\mu)\rangle$ and $\langle\delta_\sigma(r_\mu)\rangle$ by means of VMC.

During the preliminary test stage, however, it was discovered that a straightforward implementation of both estimators is plagued with variance-related problems due to the presence of fermionic nodal surfaces in the core electron region. The numerical characterization of this difficulty (described in Secs. II and IV) allowed us to suggest the use of alternative sampling distributions to circumvent the problem and to meaningfully compare the renormalized differential estimator[16] with the one proposed in this work, demonstrating a better efficiency for the latter.

The outline of the manuscript is as follows. In Sec. II, the approaches presented in Refs. 16 and 20 to estimate $\langle\delta\rangle$ are revised. We also discuss the origin of the difficulty that forces one to use sampling distributions different from the traditionally employed Ψ_T^2. Section III presents our numerical results, starting with a discussion on the robustness and performance of the straightforward implementation of both estimators while using the standard Metropolis Monte Carlo as a sampling scheme. The results obtained using some alternative sampling distributions and algorithms are discussed in Sec. IV. Finally, Sec. V contains our conclusions and the outlook for future work.

II. METHODS

As mentioned in Sec. I, the aim of this work is to provide efficient strategies for computing the electron density on top of heavy nuclei with VMC. We therefore begin extending the local estimator approach introduced in Ref. 20 and revising the theory of the renormalized differential estimator presented by Assaraf et al.[16]

As suggested in Ref. 20, one may write the expectation value $\langle\delta(r_\mu)\rangle$ of the Dirac delta measuring the density of an N electron system on top of nucleus μ as

$$\langle\delta(r_\mu)\rangle_{\Psi^2}$$
$$= \frac{\sum_{i=1}^N \int |\Phi^i(r_1, \ldots, r_{i-1}, r_{i+1}, r_N)|^2 dr_1, \ldots, dr_{i-1}dr_{i+1}dr_N}{\int |\Psi(r_1, \ldots, r_N)|^2 dr_1, \ldots, dr_N}, \tag{2}$$

with

$$\Phi^i(r_1, \ldots, r_{i-1}, r_{i+1}, \ldots, r_N)$$
$$= \int \delta(r_\mu - r_i)\Psi(r_1, \ldots, r_N)dr_i. \tag{3}$$

Here, $\Psi(r_1, \ldots, r_N)$ is the ground state molecular wave function. The electron spin density takes the usual form $\sum_i^{N\alpha}\langle\delta_\alpha(r_{i\mu})\rangle_{\Psi^2} - \sum_j^{N\beta}\langle\delta_\beta(r_{j\mu})\rangle_{\Psi^2}$ where N_α and N_β are the number of spin up and down electrons, respectively. Equation (2) is made suitable for VMC estimation restoring an identical dimensionality for the denominator and numerator multiplying each term of the latter by

$$\mathcal{N}^2 \int \xi(r_i)^2 dr_i = 1, \tag{4}$$

where $\xi(r_i)$ is a square integrable auxiliary function whose analytical form should be chosen to reduce the variance of the MC estimator. Introducing the TWF $\Psi_T(r_1, \ldots, r_N)$ approximating $\Psi(r_1, \ldots, r_N)$, Eq. (2) can be rewritten as

$$\langle\delta(r_\mu)\rangle_{\Psi^2} = \mathcal{N}^2\sum_{i=1}^N \frac{\int \left|\frac{\Phi^i_T\xi}{\Psi_T}\right|^2 |\Psi_T|^2 dr_1,\ldots,dr_N}{\int |\Psi_T|^2 dr_1,\ldots,dr_N}$$

$$\approx \frac{\mathcal{N}^2}{M}\sum_{j=1}^M\sum_{i=1}^N \left|\frac{\Phi^i_T\xi}{\Psi_T}\right|^2, \qquad (5)$$

where M configurations, sampled according to Ψ_T^2, are used for the estimate. Equation (5), the main result of this section and henceforth referred to as "local estimator," is straightforwardly implemented in any VMC code by moving, in turn, each electron on top of nucleus μ to estimate the various Φ^i_T. As a linguistic note, we mention here that the adjective local for the estimator defined in Eq. (5) is meant to indicate the absence of terms depending on first or higher order derivatives of the trial wave function.

As an alternative to the use of Eq. (5), one could employ the renormalized differential estimator by Assaraf *et al.*[16] This was originally introduced with the broader aim of improving the performance of computing the electron density at a generic position r_g in space. This differential estimator reads

$$\langle\delta(r_g)\rangle = -\frac{1}{4\pi}\sum_{i=1}^N$$

$$\times \left\langle \frac{1}{|r_i-r_g|}\frac{\nabla_i^2[f(r_i,r_g)\Psi_T^2(r_1,\ldots,r_N)]}{\Psi_T^2(r_1,\ldots,r_N)}\right\rangle_{\Psi_T^2}, \quad (6)$$

where $f(r_i,r_g)$ is a smooth auxiliary function fulfilling $\lim_{r_i\to r_g} f(r_i,r_g)=1$ and Ψ_T is a trial wave function approximating Ψ. In this work, we closely follow the suggestion provided in Ref. 16 and use the specialized form

$$f(r_i,r_\mu) = 1 + 2Z_A|r_i-r_\mu| \qquad (7)$$

to reduce the statistical error in the nuclear cusp region.

Although the implementation of both Eqs. (5) and (6) in a VMC code is fairly straightforward, we need to stress that difficulties are going to be encountered simulating fermionic systems due to the presence of large and unbounded fluctuations in regions surrounding a nodal surface. Even though these fluctuations have no direct effect on the expectation value itself, they violate the hypothesis of the central limit theorem, so that it becomes impossible to estimate the statistical uncertainty associated with $\langle\delta\rangle$. Perhaps, the simplest way to appreciate the origin of this problem is to recognize that the nodal surfaces of Ψ_T^2 do not coincide with the ones of $\nabla^2[f(r_i,r_\mu)\Psi_T^2(r_1,\ldots,r_N)]$ and $(\Phi^i_T\xi)^2$. Computing the variance of both estimators with respect to the sampling density Ψ_T^2, one easily realizes that the second order divergence on a node becomes fourth order and hence not integrable (in the Appendix, we exemplify this problem using a determinantal model for the ^3S state of He).

A possible way of circumventing this problem would be to modify the conventional walker density $p(R) = \psi_T^2/\int|\psi_T|^2 dR$ using

$$p(R) = \frac{\chi(R)w(R)}{\int \chi(R)w(R)dR}. \qquad (8)$$

Here, the weighting function $w(R)=|\psi(R)_T|^2/\chi(R)$ corrects for the sampling bias introduced by using a shifted distribution of the kind

$$\chi = [D(r_1,\ldots,r_N)^2 + cE(r_1,\ldots,r_N)]\exp(J)^2 \qquad (9)$$

rather than Ψ_T^2. In the above equation, $c>0$ is a constant, D is the determinant part of Ψ_T, E is a semipositive symmetric function of the electronic coordinates, and $\exp(J)$ is an explicitly correlated factor. With this substitution, the general form of the two estimators becomes

$$\langle\hat{O}\rangle = \frac{\int \hat{O}w\chi dr_1,\ldots,dr_N}{\int w\chi dr_1,\ldots,dr_N} = \frac{\langle Ow\rangle_\chi}{\langle w\rangle_\chi} \equiv \frac{\mathcal{D}}{\mathcal{W}}, \qquad (10)$$

with $O=\mathcal{N}^2\sum_{i=1}^N|(\Phi^i_T\xi)/\Psi_T|^2$ [cf. Eq. (5)] and $O=-(1/4\pi)\sum_{i=1}^N 1/|r_{i\mu}|(\nabla_i^2[f\Psi_T^2]/\Psi_T^2)$ [cf. Eq. (6)] for the local and differential estimators, respectively. The advantage of using Eq. (10) comes from the fact that it has a bounded variance var(δ) provided that E does not go to zero on the node faster than a first order function. In this work, we propose the use of the following analytical forms for the distribution χ:

$$\chi_{abs} = [D(r_1,\ldots,r_N)^2 + c|D(r_1,\ldots,r_N)|]\exp(J)^2, \qquad (11)$$

$$\chi_f = \left\{ D(r_1,\ldots,r_N)^2 + c\prod_{i=1}^N [1+\exp(\gamma(r_c-r_i))]^{-1}\right\}$$
$$\times\exp(J)^2, \qquad (12)$$

$$\chi_{abs,f} = \left\{ D(r_1,\ldots,r_N)^2 + c|D(r_1,\ldots,r_N)| \right.$$
$$\left. \times\prod_{i=1}^N [1+\exp(\gamma(r_c-r_i))]^{-1}\right\}\exp(J)^2. \qquad (13)$$

Notice that both χ_{abs} and $\chi_{abs,f}$ go to zero linearly with respect to the magnitude of a vector perpendicular to the nodal surface, whereas χ_f includes the strictly positive Fermi function.

When using the χ distributions in Eqs. (11)–(13) to guide the sampling, a population bias is introduced in the estimate provided by Eq. (10).[21] Different from what happens in DMC simulations, such bias is not due to the correlation between configuration introduced by controlling the population size but rather by the fact that $\mathcal{D}$ and $\mathcal{W}$ must be independently estimated before computing the ratio that provides one with the desired expectation value. When only a limited number of samples are available (e.g., when computing the expectation value over a single simulation block), the statistical error of the denominator $\mathcal{W}$ translates into a systematic error of the average over the block mean values.[21] We found it sufficient, however, to compute the VMC results using 1000 walkers and a large number of steps (≥ 1000) per block to eliminate this population bias.

TABLE I. Average electron and spin densities ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) for first, second, and third row atoms obtained using the differential estimator and sampling Ψ_T^2. τ is the time step, err($\cdot$) is the standard error associated with a mean value, and dev($\langle\delta\rangle$)$=(\langle\delta\rangle-\delta_{HF})/$err($\delta$) is the relative statistical deviation from the exact restricted Hartree–Fock density.

Metropolis results	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	dev(δ)
He	0.0800	3.5970	1.07×10^{-3}	−0.000 004	1.1×10^{-3}	−0.24
Li	0.0200	13.8336	4.16×10^{-3}	0.1676	4.2×10^{-3}	−0.16
C	0.0100	127.5010	5.81×10^{-2}	0.0515	5.7×10^{-2}	−0.9
Ne	0.0050	620.1368	5.53×10^{-1}	0.3378	5.4×10^{-1}	−0.16
Ar	0.001	3845.5	2.1×10^{0}	1.8	2.2×10^{0}	2.5
K	0.0005	4547.2	2.5×10^{0}	6.6	2.7×10^{0}	3.0
Langevin results						
Ar	0.001	3845	3.38×10^{0}	−1.8000	3.06×10^{0}	1.6
K	0.0005	4541	6.46×10^{0}	−1.8295	3.06×10^{0}	0.22

III. RESULTS

In this section, the performances of Eqs. (5) and (6) are tested, computing the electron and electron spin nuclear densities for several first, second, and third row atoms. To simplify the presentation of our results, we begin considering the calculation of restricted Hartree–Fock (RHF) densities[22] in some detail. Exploiting the fact that RHF densities are known analytically, this approach allows us to assess the presence of possible unexpected problems in the Monte Carlo simulations. Unless otherwise mentioned, the latter were carried out with a standard implementation of the Metropolis Monte Carlo algorithm, which attempts single-electron displacements with uniform probability within a cubic box centered on the initial Cartesian position of the electron itself. Results for correlated electrons are presented at a later stage.

A. Testing the performances of the estimators with Ψ_T^2 sampling

Let us start by considering the Metropolis Monte Carlo results obtained using RHF wave functions for He, Li, C, Ne, Ar, and K provided in Tables I and II for Eqs. (6) and (5), respectively. All data have been obtained using simulations sampling an identical number of configurations. In the calculations using the local estimator [Eq. (5)], we opted to use the auxiliary function $\xi(r)=\exp(-Z_{eff}r)$, with Z_{eff} being chosen by comparing the standard errors obtained using several equispaced Z_{eff} values centered around the atomic number Z during a single simulation. We shall return to discuss alternative forms for auxiliary function below.

From Tables I and II we note, in particular, that both estimators provide accurate results, the difference between the exact HF density (δ_{HF}) and the simulation result (δ) being well within the associated statistical error err(δ) as indicated by the relative deviation dev(δ)$=(\delta-\delta_{HF})/$err(δ). The electron density and the spin density are sampled from the same Markov chain and are seen to have similar statistical errors, a result independent of the estimator used. Indeed, sampling the spin density $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$ from the same Markov chain is necessary in order to keep the statistical error low. Despite these similarities, the advantage provided by the local estimator [Eq. (5)] is clearly apparent, particularly for the nodeless ground state of He. For this system, Eq. (5) provides a standard error 29 times lower than Eq. (6) when using simulations of identical length.

For the remaining atoms, the standard error obtained with the local estimator is roughly an order of magnitude smaller than the one obtained using the differential counterpart.

Two details are perhaps worth mentioning at this stage. First, there is the fact that we have made some attempts at optimizing the Z_A value used in Eq. (7) to minimize the variance of Eq. (6) and to provide one with a fairer comparison between the two estimators. Our attempts, though, suggested that using the atomic number as Z_A is an almost optimal strategy and that only minimal improvements are

TABLE II. Average electron and spin densities ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) for first, second, and third row atoms obtained using the local estimator and sampling Ψ_T^2 with the Metropolis algorithm. τ is the time step, err($\cdot$) is the standard error associated with a mean value, Z_{eff} is the parameter of the auxiliary function ξ, and dev($\langle\delta\rangle$)$=(\langle\delta\rangle-\delta_{HF})/$err($\delta$) is the relative statistical deviation from the exact restricted Hartree–Fock density.

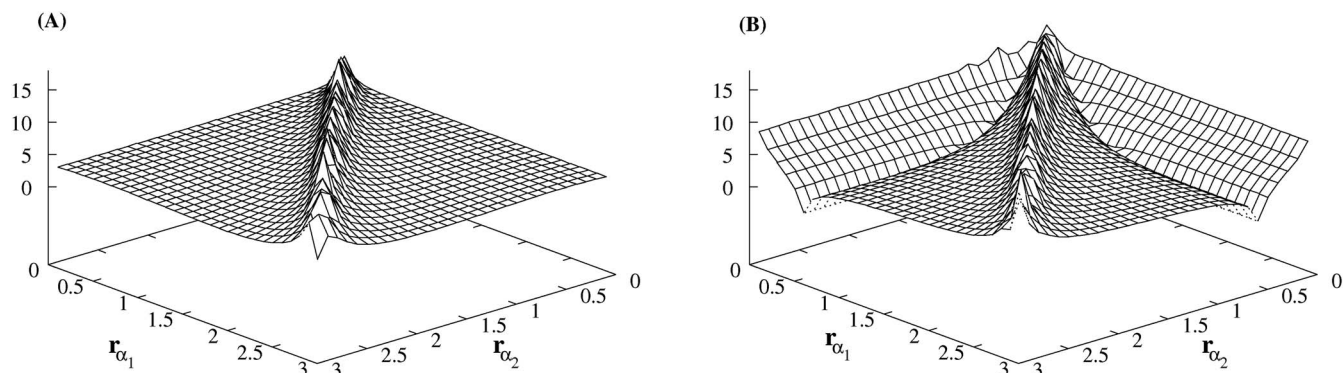|	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	dev(δ)	Z_{eff}
He	0.4	3.597 21	2.7×10^{-5}	−0.000 007	2.8×10^{-5}	−1.16	1.6
Li	0.08	13.834	3.2×10^{-4}	0.167	3.1×10^{-4}	−1.32	2.69
C	0.04	127.562	8.7×10^{-3}	−0.026	9.5×10^{-3}	1.02	6.88
Ne	0.004	620.23	4.9×10^{-2}	0.017	5.1×10^{-2}	0.12	12.14
Ar	0.004	3839.1	3.5×10^{-1}	0.2	3.4×10^{-1}	−3.30	23.7
K	0.0020	4540.0	7.2×10^{-1}	1.17	7.4×10^{-1}	0.48	26.33

FIG. 1. Logarithm of the local variance $\log_{10}[\text{var}(\delta_{\alpha_1}+\delta_{\alpha_2})]$ as a function of the distance (in bohr) from the nucleus for identical spin electrons in the Li atom. Results obtained using the Metropolis algorithm. Panel A: Local estimator [Eq. (5) with $Z_{\text{eff}}=2.69$]; panel B: Differential estimator [Eq. (6)].

obtained by means of a complete optimization. Second, we found simulations employing the differential estimator to be slightly faster than the ones using the local estimator (e.g., 20% faster for C when a HF plus two-body correlation functions is used), due to the need to compute a few additional terms in the latter. Despite such increase in the computational cost, Eq. (5) still appears to provide one with a more efficient algorithm when Ψ_T^2 is sampled.

In an attempt at characterizing and rationalizing the different performances of the two estimators, we decided to study their local variance $\text{var}(\delta_{\alpha_1}+\delta_{\alpha_2})=\langle|\delta(r_{\alpha_1})+\delta(r_{\alpha_2})|^2\rangle -\langle\delta(r_{\alpha_1})+\delta(r_{\alpha_2})\rangle^2$ as a function of the distance r_{α_i} from which the ith electron is moved on top of the nucleus. In the previous equation, $\delta(r_{\alpha_i})$ indicates one of the two estimators defined in Eqs. (5) and (6). Figure 1 shows the results obtained for the Li atom when $\text{var}(\delta_{\alpha_1}+\delta_{\alpha_2})$ is plotted as a two-dimensional function of the distance between the each identical spin electrons and the nucleus. We notice a large variance in the region where $r_{\alpha_1} \approx r_{\alpha_2}$ (i.e., the location of the nodal plane of the RHF wave function) in both cases. However, Eq. (6) also presents a large variance at short distances from the nucleus ($r \rightarrow 0$, panel B), whereas the local estimator behaves more smoothly in the same region (panel A). Considering the weighting provided by Ψ_T^2 in computing the expectation values, one arrives at the conclusion that the range of electron-nucleus distance $0 \leq r \leq 0.5$ bohr provides a substantial contribution to the variance of the differential

estimator in the case of Li. Similar conclusions can also be reached plotting $\text{var}(\delta)$ for Li as a function of a single electron-nucleus distance (panel A, Fig. 2). Also, notice that the histograms present a fairly rough shape despite the long simulations employed to collect the data, an outcome that may also be due to the fluctuations of the variance close to the nodal surface.

Although somewhat different in shape, a similar trend is observed also for Ne (panel B, Fig. 2), with the highest value of $\text{var}(\delta)$ being located in the range of distances $0 \leq r \leq 1$ bohr. In this case, however, the fluctuations of the differential estimator in the valence region are expected to contribute to the total variance due to the higher valence electron density than in the case of lithium. Interestingly, a vastly different behavior of $\text{var}(\delta)$ is displayed by the local estimator for Ne, with fluctuations in the core regions being reduced by roughly two orders of magnitude and being negligible in the valence. This latter effect is likely to be due to the short range exponential form of the auxiliary function ξ, for which the value $Z_{\text{eff}}=12.5$ was used in order to minimize the variance of the estimator.

Taken together, the results shown in Figs. 1 and 2 suggest that the lower standard error obtained using the local estimator with respect to the current implementation of the differential one is likely to be due to reduced fluctuations in both the core and the valence region. Nevertheless, the large fluctuations close to the node found for both Eqs. (5) and (6)
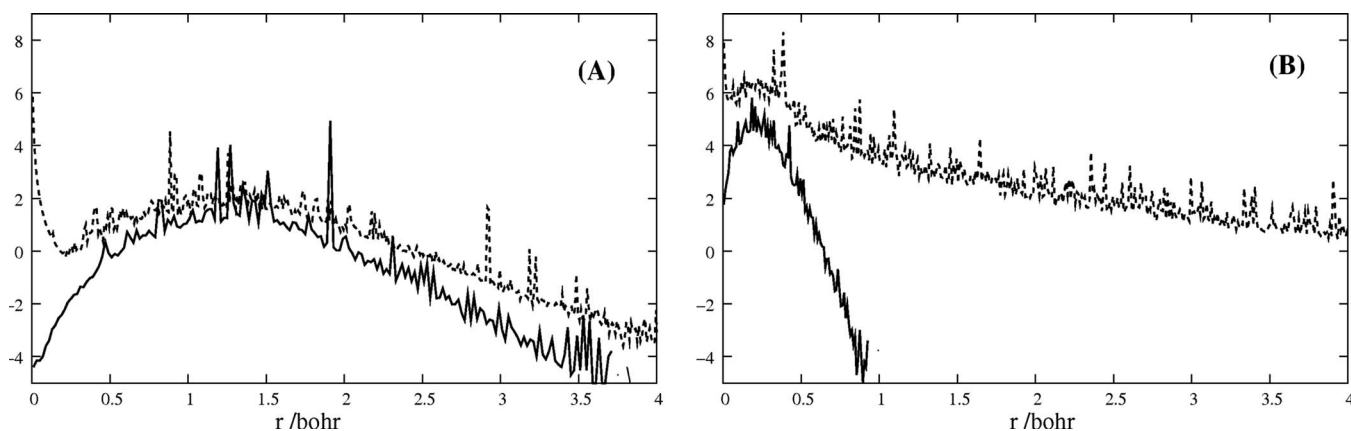

FIG. 2. Logarithm of the local variance $\log_{10}[\text{var}(\delta)]$ as a function of the electron-nucleus distance (in bohr) obtained using the Metropolis algorithm. Panel A: Li atom; panel B: Ne atom. The solid and dashed lines correspond to the local [Eq. (5)] and differential [Eq. (6)] estimators, respectively.
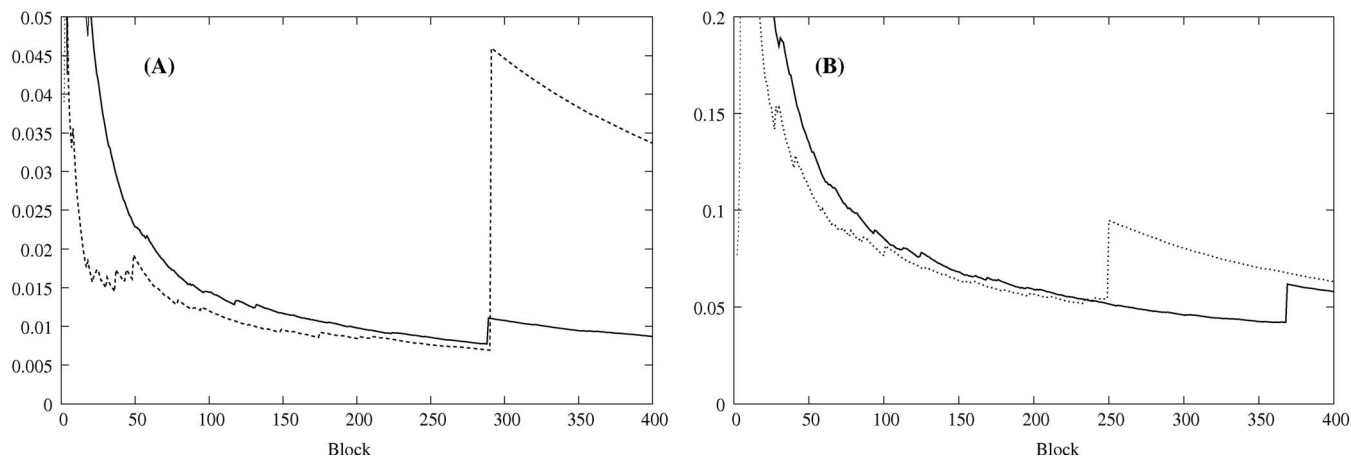
FIG. 3. Evolution of the standard error associated with $\langle \delta \rangle$ for carbon as a function of the number of blocks (2×10^6 configurations each) in a simulation sampling Ψ_T^2. Panel A: Local estimator [Eq. (5)]; panel B: Differential estimator [Eq. (6)]. Solid line: RHF wave function; dashed line: Explicitly correlated wave function. Notice the different ordinate scales in the two plots.

(see Fig. 1) are an important source of concerns and must be addressed to obtain robust estimates. As a matter of fact, one can easily prove that both Eqs. (6) and (5) have an unbounded variance as discussed in Sec. I; the effect of the latter, however, can be hidden by the finite simulation length and by the fact that the density is vanishingly small around a nodal surface. This point is exemplified in Fig. 3, which shows the behavior of the standard error as a function of the number of blocks when using Ψ_T^2 as sampling density. Clearly, the sudden "jumps" (roughly a factor of 2–5) in the standard error highlighted in Fig. 3 are rare but catastrophic events that spoil the robustness of the MC estimate. This is particularly apparent for explicitly correlated wave functions and for the local estimator due to its lower standard error. Needless to say, one would hope that both estimators may enjoy a higher statistical efficiency for heavier systems once the problem of unbound variance is eliminated.

B. Improved sampling densities

In order to address the complications due to the unbound variance and to obtain reliable results and error estimates, we

tested the performance of the sampling distributions provided in Eqs. (11)–(13). These distributions, however, require few parameters ($\{c\}$ or $\{c, r_c, \gamma\}$) to be chosen/optimized in order to obtain the highest possible statistical efficiency from the VMC simulations. In our attempt at selecting those parameters, we devised a different strategy for each distribution.

In the simplest case provided by χ_{abs} [Eq. (11)], the parameter of the distribution was roughly optimized by means of relatively short simulations (each sampling 4×10^7 configurations) with values for c a factor of 5 apart from each other. For the long production runs, we selected the value of c that gave the lowest standard error for $\langle \delta \rangle$. As for χ_f, one should note that the Fermi function makes it strictly nonzero, thus eliminating the presence of nodes. However, the parameters must be chosen so that its value is high enough around a node to remedy the problem while still obtaining a low variance. To this end, we found that the parameters $\{\gamma, r_c\}$, i.e., respectively, the parameters for the decay and extent of the "flat" region of $\{\exp[\gamma(r_c - r_i)]\}^{-1}$, could be chosen from the radial electron density ρ obtained sampling Ψ_T^2. This is
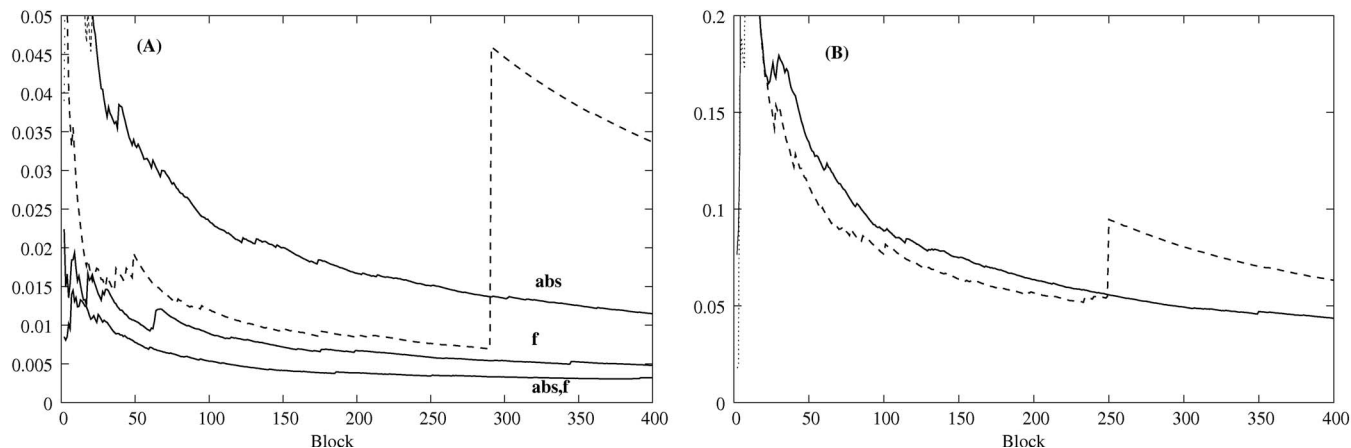

FIG. 4. Evolution of the standard error associated with $\langle \delta \rangle$ for carbon as a function of the number of simulation blocks (2×10^6 configurations each) for an explicitly correlated Ψ_T. Panel A: Local estimator [Eq. (5)]; panel B: Differential estimator [Eq. (6)]. The dashed lines represent the results obtained with a direct sampling of Ψ_T^2. The solid lines in panel A have been obtained using χ_{abs} (top), χ_f (middle), and $\chi_{abs,f}$ (bottom). The solid line in (B) represents the results obtained using $\chi_{abs,f}$.

TABLE III. Average electron and spin densities for first row atoms ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) obtained using the local estimator and sampling χ_{abs} with the Metropolis algorithm. τ is the time step, err($\cdot$) is the standard error associated with a mean value, Z_{eff} is the parameter of the auxiliary function ξ, is the relative statistical deviation from the exact restricted Hartree–Fock density, c the parameter in Eq. (11), and $\mathcal{W}$ is the average configuration weight.

	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	dev(δ)	Z_{eff}	c	$\mathcal{W}$
Li	0.08	13.8349	5.3×10^{-4}	0.168	4.6×10^{-4}	−1.3	2.58	0.1	0.02
C	0.04	127.5371	1.2×10^{-4}	−0.005	1.2×10^{-4}	1.38	5.75	0.01	0.01
Ne	0.038	619.9105	1.1×10^{-1}	−0.03	1.19×10^{-1}	2.92	11.5	0.0002	0.14

done by matching the decay of the Fermi function (γ) to the exponentially decaying tail of ρ and determining the onset of the exponential tail region (r_c) by visual inspection. A value of $c=1\times10^{-7}$ was found adequate to fully eliminate the nodal problem and was kept identical for all atoms. We found, however, that this way of choosing the parameters for χ_f generates a sampling distribution requiring a large number of MC steps to equilibrate for atoms heavier than Ne; to correct for this deficiency, a slightly higher γ and a shorter r_c were used.

As for $\chi_{abs,f}$ [Eq. (13)], this combines the properties of χ_{abs} and χ_f and shows no problem with the equilibration of the distribution for large atoms. In an attempt at obtaining the best possible performance out of this analytical form, we tried to optimize its parameters by minimizing the standard deviation σ_δ^2 using a fixed set of walkers distributed according to Ψ_T^2. The form of the cost function used in this approach reads

$$\sigma_\delta^2 = \frac{\sigma_{\mathcal{D}}^2}{\mathcal{W}^2} + \frac{\mathcal{D}^2\sigma_{\mathcal{D}}^2}{\mathcal{W}^4} - 2\frac{\mathcal{D}\sigma_{\mathcal{D}\mathcal{W}}^2}{\mathcal{W}^3}, \qquad (14)$$

where $\mathcal{D}$ and $\mathcal{W}$ are defined in Eq. (10) and σ_χ^2 and $\sigma_{\chi\mathcal{Y}}^2$ are their variance and covariance, respectively. A full optimization of the parameters for $\chi_{abs,f}$ turned out to be unfruitful, with unphysical values obtained at the end of the minimization procedure. We found, however, that optimizing c with Eq. (14) and choosing $\{\gamma, r_c\}$ as indicated above for χ_f is indeed a useful approach.

Panel (A) of Fig. 4 displays the behavior of the standard error err($\langle\delta\rangle$) as a function the number of blocks during a simulation obtained using the local estimator for the carbon atom and the three distributions discussed above. Also shown, there are results from a direct sampling of Ψ_T^2 (dashed line). From this figure, one notices that the weighted sampling eliminates the pathological "jumps" in the standard error, a result confirmed in several long realizations. A similar outcome was obtained also for the differential estimator

[Panel (B), Fig. 4], where the results for a simulation using $\chi_{abs,f}$ are shown. Notice, however, that the differential estimator still presents a substantially larger statistical error than the local estimator despite the improvements introduced by the different sampling densities. Quantitatively, the results shown in Fig. 4 indicate the efficiency of Eq. (5) as being 100 times higher than the one of Eq. (6) when using $\chi_{abs,f}$ for the carbon atom. We therefore feel that this comparison highlights a substantially better performance for the local estimator in computing the electron density on top of nuclei and, consequently, the following discussion will focus mainly on the performance of Eq. (5).

Tables III–V provide the final quantitative results obtained using Eqs. (11)–(13) as sampling densities, the local estimator with $\xi(r)=\exp(-Z_{eff}r)$, and the Metropolis algorithm for a larger set of atoms. Quantitatively, the largest standard error is obtained using χ_{abs}, which also presents an unfavorable scaling of err($\langle\delta\rangle$) with respect to the nuclear charge (compare Tables II and III). This finding is likely to be due to a slower decay of χ_{abs} as a function of r than Ψ_T^2 and the associated broadening of the electron density. For this reason, we limited our tests on χ_{abs} only to the Li, C, and Ne atoms. Conversely, the results obtained with χ_f (Table IV) have standard errors similar to the ones obtained during a "fortunate" realization of the direct Ψ_T^2 sampling, suggesting therefore that it is free from the rapid increase in standard error presented by χ_{abs}. As for $\chi_{abs,f}$, the latter clearly provides us with the best statistical precision out of the three modified sampling densities χ. Indeed a comparison between Tables V and II indicates that the weighted sampling approach is of comparable accuracy to a fortunate realization of the direct Ψ_T^2 sampling (i.e., when no sudden jumps are witnessed during a simulation).

In an attempt at further improving the performance of the local estimator, we also investigated the effect of more general forms for the auxiliary function ξ, whose only restriction is that it must be a three-dimensional normalizable

TABLE IV. Average electron and spin densities for first, second, and third row atoms ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) obtained using the local estimator and sampling χ_f with the Metropolis algorithm. τ is the time step, err($\cdot$) is the standard error associated with a mean value, Z_{eff} is the parameter of the auxiliary function ξ, dev($\langle\delta\rangle$)=($\langle\delta\rangle-\delta_{HF}$)/err($\delta$) is the relative statistical deviation from the exact restricted Hartree–Fock density, (c,γ,r_c) the parameters in Eq. (12), and $\mathcal{W}$ is the average configuration weight.

	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	dev(δ)	Z_{eff}	c	γ	r_c	$\mathcal{W}$
Li	0.08	13.8341	4.88×10^{-4}	0.16 730	5.9×10^{-4}	−0.2673	2.69	1.0×10^{-7}	1.11	4.5	0.12
C	0.04	127.553	5.34×10^{-3}	−0.0096	5.38×10^{-3}	0.0250	5.88	1.0×10^{-7}	2.04	1.5	0.46
Ne	0.004	620.031	8.98×10^{-2}	0.15	9.6×10^{-2}	−2.1425	11.3	1.0×10^{-7}	2.87	1.3	0.13
Ar	0.004	3839.53	3.9×10^{-1}	0.47	4.09×10^{-1}	−1.9311	23.7	1.0×10^{-7}	2.3	0.7	0.996
K	0.002	4538.71	5.7×10^{-1}	−0.13	5.97×10^{-1}	−1.6996	28.33	1.0×10^{-7}	2.77	1.1	0.58

TABLE V. Average electron and spin densities for first, second, and third row atoms ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) obtained using the local estimator and sampling $\chi_{abs,f}$ with the Metropolis algorithm. τ is the time step, err($\cdot$) is the standard error associated with a mean value, Z_{eff} is the parameter of the auxiliary function ξ, dev($\langle\delta\rangle$)=($\langle\delta\rangle-\delta_{HF}$)/err($\delta$) is the relative statistical deviation from the exact restricted Hartree–Fock density, (c,γ,r_c) the parameters in Eq. (13), and $\mathcal{W}$ is the average configuration weight.

	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	dev(δ)	Z_{eff}	c	γ	r_c	$\mathcal{W}$
Li	0.08	13.8343	1.70×10^{-4}	0.1672	1.65×10^{-4}	0.04	2.69	0.01	1.11	4.50	0.37
C	0.04	127.5578	3.92×10^{-3}	−0.004 29	3.91×10^{-3}	1.21	5.875	0.01	2.036 80	1.5	0.83
Ne	0.004	620.2153	3.58×10^{-2}	−0.000 744	3.73×10^{-2}	−0.22	10.97	0.1	2.872 18	1.3	0.67
Ar	0.004	3839.5391	3.22×10^{-1}	0.09	3.33×10^{-1}	−2.29	23.7	0.016	2.296 92	2.0	0.71
K	0.002	4538.6754	4.44×10^{-1}	0.41	4.23×10^{-1}	−2.29	27	0.1	0.842 55	2.0	0.73

function [see Eq. (4)]. As a preliminary test, a sum of two exponentials fitted to the radial electron density was used for He obtaining $\langle\delta\rangle=3.597\ 253\pm9\times10^{-6}$, i.e., factor of 2 lower standard error than the single exponent auxiliary function and 86 times lower than the differential estimator. Unfortunately, the biexponential form did not provide us with any improvement over the single exponential for larger systems with nodal surfaces, this outcome prompting us to test even more flexible auxiliary functions. Among those tested, we mention only $\xi(r)=a\exp(-Z_{eff}r)+(1-a)F(r)$, where $F(r)$ is the Fermi function. The latter form mimics the $\chi_{abs,f}$ single electron density for first row atoms significantly better than a single exponential, but it provided no reduction in the standard error associated with $\langle\delta\rangle$ for Li. Similar results were obtained for other model ξ's, suggesting therefore that the use of the density as weighting function is not likely to give any improvement. In the end, the good results obtained using the single exponential auxiliary function and the simplifying matter of a single parameter optimization strongly motivates its general use in computing $\langle\delta\rangle$.

To conclude this section, we report in Table VI the results obtained using explicitly correlated TWFs built as a product between the seven term total symmetric correlation factor suggested by Schmidt and Moskowitz[23] and a RHF function, $\chi_{abs,f}$ as distribution and the local estimator. In these cases, the parameters of the correlation factor were optimized by minimizing the variance of the local energy over a fixed sample of configurations. Notice that the analytical form chosen for the correlated TWFs prevents us from comparing our data for $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$ either with experiments or results obtained using more accurate model wave functions. Given this fact and that the aim of this work is only to develop efficient estimators for the Dirac delta expectation value, we restrict our comments, mentioning only that no substantial changes in the computational procedure were needed and that we obtained standard errors similar to the ones obtained using a simpler determinant function.

IV. OPTIMAL MARKOV CHAINS

In VMC simulations, configurations sampling the appropriate target density are usually generated by means of a Markov chain. Although convenient in terms of code complexity and flexibility, the usage of Markov chains produces samples that are sequentially correlated especially in the core region.[24–27] One should therefore strive to reduce the correlation time for the specific observable of interest (e.g., the density), finding optimally efficient Markov chains by correctly choosing the sampling algorithm and the step size used to propagate the configurations in space. The aim of this section is thus to provide a short discussion of our efforts toward an appropriate selection of those two variables.

Let us start by mentioning that we found it convenient to use the weighted mean square electron displacement $\langle\Delta r^2(\tau)\rangle\exp(-2Z_{eff}r)$, a function of the electron-nucleus distance (r), as an indicator of the overall sampling efficiency and to select an almost optimal step size τ. $\langle\Delta r^2(\tau)\rangle\exp(-2Z_{eff}r)$ reflects the average walker displacement in the core region, of special importance for the local estimator, and it is an inexpensive way of selecting τ that scores as well as the correlation function for δ. The results shown in Fig. 5, all obtained using the Metropolis algorithm, give an indication of the longest and shortest weighted mean square displacement obtained for neon over a broad range of step lengths when sampling either $\chi_{abs,f}$ or Ψ_T^2. Notice that the value of $\sqrt{\langle\Delta r^2(\tau)\rangle\exp(-2Z_{eff}r)}$ close to the nucleus is at least an order of magnitude smaller than τ for the Metropolis algorithm, suggesting that a substantial amount of rejection is indeed taking place. Interestingly, it appears that the weighted mean square displacement obtained using $\chi_{abs,f}$ is

TABLE VI. Average electron and spin densities for first, second, and third row atoms ($\langle\delta\rangle$ and $\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$, a.u.) obtained using the local estimator and sampling $\chi_{abs,f}$ with the Metropolis algorithm. τ is the time step, Z_{eff} is the parameter of the auxiliary function ξ, err($\cdot$) is the standard error associated with a mean value, (c,γ,r_c) the parameters in Eq. (13), and $\mathcal{W}$ is the average configuration weight. Explicitly correlated Ψ_T.

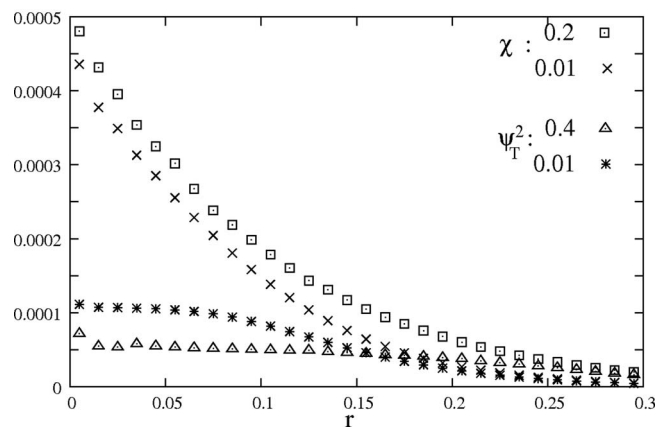|	τ	$\langle\delta\rangle$	err(δ)	$\langle\delta_\alpha\rangle-\langle\delta_\beta\rangle$	err($\delta_\alpha-\delta_\beta$)	Z_{eff}	c	γ	r_c	$\mathcal{W}$
Li	0.5	13.7836	2.3×10^{-4}	0.1656	3.3×10^{-4}	3.8	0.3	2.11	4.5	0.02
C	0.08	127.6793	2.6×10^{-3}	−0.1110	3.2×10^{-3}	6.0	2.1	2.04	1.5	0.02
Ne	0.2	613.71	3.0×10^{-2}	−0.0047	3.0×10^{-2}	10.14	3.9	2.87	1.5	0.01
Ar	0.08	3724.6753	8.0×10^{-1}	−1.87	8.9×10^{-1}	19.2	1.6	2.3	2.0	0.02
K	0.04	4395.1288	7.2×10^{-1}	0.35	7.5×10^{-1}	20.3	1.6	0.84	2.0	0.12

FIG. 5. Weighted mean square displacement $\langle \Delta r^2(\tau) \rangle \exp(-2Z_{\text{eff}}r)$ (bohr^2) as a function of the electron-nucleus distance r (bohr). Shown are the highest and lowest weighted mean square displacements obtained with the Metropolis algorithm over the range of step lengths $0.01 \leq \tau \leq 0.4$ for $\chi = \chi_{\text{abs},f}$ ($\square, \times$) and $\chi = \Psi_T^2$ ($\triangle, *$). Notice the higher $\langle \Delta r^2 \rangle \exp(-2Z_{\text{eff}}r)$ values obtained using the alternative sampling distributions $\chi_{\text{abs},f}$.

substantially larger (up to five times) than the same quantity obtained sampling Ψ_T^2. This could be understood by recalling that $\chi_{\text{abs},f}$ is written as a linear combination between D^2 and $|D|$, which is expected to generate a smoother behavior of the sampling density. When using $\chi_{\text{abs},f}$ as sampling distribution for Ne, the largest weighted mean square displacement over the important range of r is obtained with $\tau = 0.2$, a fact that translates into a 25% lower standard error than for the $\tau = 0.01$ case. Considering the analytical form [Eq. (14)] of the statistical error associated with the weighted sampling [Eq. (10)], one would expect that a larger variance should be obtained than in the case where Ψ_T^2 sampling is carried out due to weight fluctuations and the presence of a positive covariance. Figure 5 shows, however, that $\chi_{\text{abs},f}$ provides one with a more efficient sampling and an explanation for the lower standard error discussed earlier. As for the direct sampling of Ψ_T^2, $\tau = 0.01$ appear to produce longer displacements and smaller standard errors than obtained using larger time steps, in good agreement with what was previously obtained.[27] It is important, however, to remember that using only the simulation's standard error to gauge the best step length with Ψ_T^2 is an inherently flawed procedure due to the possibility of events such as the ones depicted in Fig. 3.

Turning to the issue of selecting an appropriate sampling algorithm, it is common lore that using a single electron move implementation of the Langevin algorithm[11] provides a more precise mean value of the local energy than the equivalent Metropolis scheme, and that electron density histograms obtained with the two methods closely overlap. In this work, the Langevin algorithm was implemented in the standard way[11] for single electron moves, with a sequence of drift (based on the quantum velocity $\nabla_i \ln \Psi_T^2$), Gaussian diffusion, and acceptance/rejection steps. Similarly to what is usually reported in the literature, we found that the Langevin algorithm provides one with a 40% higher efficiency while using both Eqs. (5) and (6) and directly sampling Ψ_T^2 for the ground state of He. However, we also found that biased results may appear for systems whose wave function contains nodal surfaces. For instance, $\langle \delta \rangle$ obtained using Eq. (5) was found to deviate substantially from the exact RHF density in the case of carbon and lithium (6 and 16 times the standard error, respectively). Similar results were obtained using the differential estimator for carbon. This, however, is different from the unbounded variance problem discussed previously for the Metropolis algorithm, which produces unbiased Ψ_T^2 densities. An explanation for these results is obtained by considering the diagonal part of the two-body density matrix around a nodal surface. Figure 6 shows the histogram obtained with a long simulation for the two electrons of identical spin in Li. From this, one can easily notice that the Langevin algorithm [panel (B)] generated persistent configurations around the node, therefore spoiling the quadratic shape of the two-particle density matrix. Conversely, the Metropolis distribution appears smooth and displays the expected quadratic behavior around the node [panel (A)]. The situation is largely improved sampling $\chi_{\text{abs},f}$ with the Langevin algorithm, the numerical results for $\langle \delta \rangle$ obtained with Eq. (5) differing from the exact RHF value by less than three times the standard error. This finding is most likely connected to a weaker divergence of the quantum force $\nabla \ln \chi_{\text{abs},f}$ due to the presence of $|D|$. err($\langle \delta \rangle$), however, turned out to be substantially larger than the one obtained with the Metropolis algorithm. In this respect, a possible avenue of improvement may be represented by the improved Langevin algorithm introduced in Refs. 28 and 26. Given the
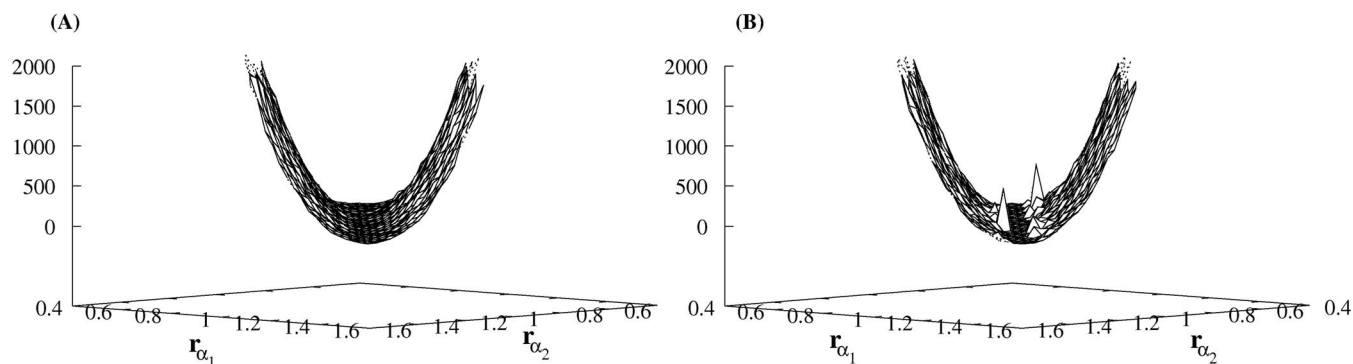

FIG. 6. Histogram of the two-body diagonal density matrix for identical spin electrons in Li as a function of their distance from the nucleus (bohr). Panels A and B show the results obtained using the Metropolis and Langevin algorithms, respectively. In panel B, notice the relatively high peaks due to the presence of persisting configurations around the nodal region ($r_{\alpha_1} \sim r_{\alpha_2}$) that are absent in the distribution generated with the Metropolis scheme.

overall satisfactory performance provided by the Metropolis algorithm, we decided, however, not to pursue this avenue of exploration further.

Concluding this section, we also mention the possibility of using "accelerated sampling" algorithms[24,26,27,29] to increase the efficiency in computing $\langle \delta \rangle$. In this work, we implemented the delayed rejection scheme proposed by Bressanini et al.[27] for the Metropolis algorithm, an approach where an initial electron displacement attempted with a step length suitable for the valence region is followed by a second shorter one upon rejection of the first movement. With the appropriate normalization of the acceptance probability, the second displacement is more likely to be accepted in the core region, thus producing a more efficient Markov chain. Indeed, we found that the delayed rejection method led to an overall increase in the average accepted displacement, with the larger improvements being obtained in the nuclear region for both $\chi_{abs,f}$ and Ψ_T^2. A twofold improvement in efficiency was obtained using $\chi_{abs,f}$ selecting the ratio between long and short step lengths in accord with the suggestion made in Ref. 27. In our view, this illustrates that accelerated algorithms may be a way to improve efficiency when sampling $\chi_{abs,f}$. On the other hand, no consistent reduction in standard error was seen when sampling Ψ_T^2 with the delayed rejection scheme, a finding suggesting again the presence of problems generated by the nodal surface.

V. CONCLUSIONS

Starting from earlier work on the interparticle contact density in small systems,[20] we introduced a local estimator [Eq. (5)] for the electron and spin density on top of heavy nuclei suitable for electronic structure QMC simulations. Extensive numerical simulations carried out on the latter and with the scheme proposed in Ref. 16 highlighted numerical issues related to the presence of an unbounded variance. This, due to a second order divergence of the integrand in both Eqs. (5) and (6) on the nodes of Ψ_T, was circumvented with a weighted sampling approach based on alternative sampling densities χ [Eqs. (11)–(13)]. Using the sampling density in Eq. (13), the local estimator [Eq. (5)] features a smaller computational cost, at least an order of magnitude for heavy atoms, than its differential counterpart [Eq. (6)]. Using appropriate accelerated sampling algorithms for the core region (e.g., the delayed rejection scheme tested in this work) the total cost of evaluating $\langle \delta \rangle$ with Eq. (5) is further reduced. According to these results, the local estimator appears therefore as the method of choice when computing contact electron and spin densities, quantities that bear strong relevance in the interpretation of electron paramagnetic resonance spectra. As usual, further improvements in efficiency may be expected using sampling algorithms that reduce even more the serial correlation in the VMC Markov chain provided that persistent configurations are completely eliminated (for example, see Ref. 28).

As mentioned in Sec. I, an alternative to the two estimators benchmarked in this work would be represented by the method proposed by Alexander and Coldwell.[19] Unfortunately, a direct comparison of the three methods is hampered by the different sampling techniques. Notwithstanding this hurdle, we notice that the magnitude of the statistical error quoted in Ref. 19 is similar to the one obtained using Eq. (6) when first row atoms are studied and a comparable number of independent samples used. In turn, this suggests that the local estimator should perform better in terms of overall efficiency. Besides, one would expect the sampling of independent configurations with the method in Ref. 19 to become highly problematic for third row species due to the large number of electron-electron pairs to account for.

Although not attempted in this work, the local estimator presented in Eq. (5) may be extended to deal with electron-electron coalescence probabilities following the suggestions provided in Ref. 20. At the current stage, however, the differential estimator proposed by Toulouse et al.[30] appears to provide one with a sufficient statistical precision to cope with all the foreseeable tasks, and it is therefore deemed as not necessary. Instead, a more pressing issue is represented by the spin contamination that plagues the calculation of accurate spin densities for open shell systems containing heavy atoms. We expect this to become especially important when using nonsymmetric correlation factors,[31] the latter being used to satisfy the different cusp conditions for identical and different spin electron pairs. To solve this issue, we are currently trying to reduce the spin contamination in VMC trial wave functions using a projection technique. In principle, the impact of spin contamination on the accuracy of $\langle \delta_\alpha \rangle - \langle \delta_\beta \rangle$ may be reduced by exploiting DMC simulations to project out, although partially, the contaminant states. Even though it appears straightforward to extend the DMC scheme in Ref. 20 to the case of electron-nuclear coalescence probabilities, it is our belief that employing spin projected trial wave functions to guide DMC simulations would lead to more accurate results for $\langle \delta \rangle$ and $\langle \delta_\alpha \rangle - \langle \delta_\beta \rangle$. Testing this idea is left as a future avenue of exploration.

ACKNOWLEDGMENTS

APPENDIX: UNBOUNDED VARIANCE OF $\langle \delta \rangle$

To exemplify the problem with the unbounded variance of $\langle \delta \rangle$ due to the presence of fermionic nodal surfaces, we use the simple case of the lowest ^3S state for the He atom. A determinant model for the wave function of this state reads

$$\Psi_T(r_1, r_2) = \mathcal{M}[\phi_1(r_1)\phi_2(r_2) - \phi_2(r_1)\phi_1(r_2)], \quad (A1)$$

where $\mathcal{M}$ is the normalization constant and the ϕ_i's are spherical orbitals. Despite the fact that the above function has a node for $r_1 = r_2$, $\langle \delta \rangle$ is well defined as can easily be shown by inserting Ψ_T in the expression for the expectation value. To show that the variance $(var(\delta))$ of the density estimator presented in Eq. (5) is indeed unbounded, let us start by noticing that a term of the form

$$\int dr_1 dr_2 \left| \frac{\Psi_T(0,r_2)\xi(r_1)}{\Psi_T(r_1,r_2)} \right|^4 |\Psi_T(r_1,r_2)|^2$$

$$= \int dr_1 dr_2 \frac{|\Psi_T(0,r_2)\xi(r_1)|^4}{|\Psi_T(r_1,r_2)|^2} \tag{A2}$$

is present in the expression for $\mathrm{var}(\delta)$. From the equation above, one easily recognizes that the numerator is always positive apart from when $r_2=0$, whereas the denominator has a second order zero as the original trial function on the manifold $r_1=r_2$. In principle, this would not be a problem if it was possible to find a coordinate transformation with a Jacobian contrasting the second order divergence on this manifold. A typical application of this idea in the context of VMC simulations is provided by the calculation of $\langle 1/r_{i\mu}\rangle_{\Psi_T^2}$. In this case, the variance of $\langle 1/r_{i\mu}\rangle_{\Psi_T^2}$ is bound, thanks to the fact that the divergence in the integrand of $\langle 1/r_{i\mu}^2\rangle_{\Psi_T^2}$ is eliminated by the factor $r_{i\mu}^2 dr_{i\mu}$ in the spherical coordinates Jacobian. For the ^3S state of helium, it would be natural to use the Hylleraas coordinate set $(s,t,u)=(r_1+r_2,r_1-r_2,r_{12})$, due to which the nodal surface assumes the simple form $t=0$. The volume element in Hylleraas coordinates is proportional to $(s^2-t^2)u\,ds\,dt\,du$, which unfortunately, contains the t-independent term s^2u conserving the second order divergence in the integrand. A similar conclusion is easily reached for the differential estimator in Eq. (6).

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[2] X. Lin, H. Zhang, and A. M. Rappe, J. Chem. Phys. **112**, 2650 (2000).
[3] J. Toulouse and C. Umrigar, J. Chem. Phys. **126**, 084102 (2007).
[4] A. Scemama and C. Filippi, Phys. Rev. B **73**, 241101 (2006).
[5] J. Toulouse and C. Umrigar, J. Chem. Phys. **128**, 174101 (2008).
[6] F. J. Gálvez, E. Buendía, and A. Sarsa, J. Chem. Phys. **115**, 1166 (2001).
[7] F. J. Gálvez, E. Buendía, and A. Sarsa, J. Chem. Phys. **122**, 154307 (2005).
[8] M. W. Lee, M. Mella, and A. M. Rappe, J. Chem. Phys. **122**, 244103 (2005).
[9] M. D. Brown, J. R. Trail, P. López Ríos, and R. J. Needs, J. Chem. Phys. **126**, 224110 (2007).
[10] M. Munzarova and M. Kaupp, J. Phys. Chem. A **103**, 9966 (1999).
[11] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in ab initio Quantum Chemistry* (World Scientific, Singapore, 1994).
[12] N. Jiang and D. M. Schrader, J. Chem. Phys. **109**, 9430 (1998).
[13] M. Mella, D. Bressanini, and G. Morosi, J. Chem. Phys. **111**, 108 (1999).
[14] J. Vrbik, M. F. DePasquale, and S. M. Rothstein, J. Chem. Phys. **88**, 3784 (1988).
[15] P. Langfelder, S. M. Rothstein, and J. Vrbik, J. Chem. Phys. **107**, 8525 (1997).
[16] R. Assaraf, M. Caffarel, and A. Scemama, Phys. Rev. E **75**, 035701 (2007).
[17] S. A. Alexander, R. L. Coldwell, and J. D. Morgan, J. Chem. Phys. **97**, 8407 (1992).
[18] S. A. Alexander and R. L. Coldwell, J. Chem. Phys. **103**, 2572 (1995).
[19] S. A. Alexander and R. L. Coldwell, J. Mol. Struct. **487**, 67 (1999).
[20] S. Chiesa, M. Mella, and G. Morosi, Phys. Rev. A **69**, 022701 (2004).
[21] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Academic, New York, 1986).
[22] E. Clementi and C. Roetti, At. Data Nucl. Data Tables **14**, 177 (1974).
[23] K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).
[24] C. J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993).
[25] Z. Sun, M. M. Soto, and W. A. Lester, Jr., J. Chem. Phys. **100**, 1278 (1994).
[26] M. Mella, J. B. Anderson, and A. Lüchow, Chem. Phys. Lett. **265**, 467 (1997).
[27] D. Bressanini, G. Morosi, S. Tarasco, and A. Mira, J. Chem. Phys. **121**, 3446 (2004).
[28] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).
[29] M. L. Stedman, W. M. C. Foulkes, and M. Nekovee, J. Chem. Phys. **109**, 2630 (1998).
[30] J. Toulouse, R. Assaraf, and C. J. Umrigar, J. Chem. Phys. **126**, 244122 (2007).
[31] C.-J. Huang, C. Filippi, and C. J. Umrigar, J. Chem. Phys. **108**, 8838 (1998).