




Speeding up *ab initio* diffusion Monte Carlo simulations by a smart lattice regularization

Kousuke Nakano ^{1,*}, Ryo Maezono ^{2,3} and Sandro Sorella ^{1,†}

¹*International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy*

²*School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Asahidai 1-1, Nomi, Ishikawa 923-1292, Japan*

³*Computational Engineering Applications Unit, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*



(Received 2 September 2019; revised manuscript received 26 November 2019; accepted 13 March 2020; published 7 April 2020)

One of the most significant drawbacks of the all-electron *ab initio* diffusion Monte Carlo (DMC) is that its computational cost drastically increases with the atomic number (Z), which typically scales with Z^{-6} . In this study, we introduce a very efficient implementation of the lattice regularized diffusion Monte Carlo (LRDMC), where the conventional time discretization is replaced by its lattice space counterpart. This scheme enables us to conveniently adopt a small lattice space in the vicinity of nuclei, and a large one in the valence region, by which a considerable speedup is achieved, especially for large atomic number Z . Indeed, the computational performances of the improved LRDMC can be theoretically established based on the Thomas-Fermi model for heavy atoms, implying the optimal Z^{-5} scaling for all-electron DMC calculations. This improvement enables us to apply the DMC technique even for superheavy elements ($Z \geq 104$), such as oganesson ($Z = 118$), which has the highest atomic number of all synthesized elements so far.

DOI: [10.1103/PhysRevB.101.155106](https://doi.org/10.1103/PhysRevB.101.155106)

I. INTRODUCTION

In recent years, the grand challenge in materials modeling is to provide extremely accurate reference energetics often well beyond the standard benchmark provided by the density functional theory (DFT) that notoriously is not enough predictive in several materials of both scientific [1,2] and technological interests [3,4]. This is also particularly important in view of existing progress in machine learning algorithms to define accurate classical force field potentials with reference data as unbiased as possible [5–8]. For such problems, explicitly correlated wave-function-based approaches are necessary [9–13], such as the ones used in quantum chemistry and the ones relying on statistical approaches that are known under the generic name “quantum Monte Carlo” (QMC) [14]. In practice, for electronic systems containing more than a handful of atoms, QMC remains the only possible wave function based reference method, partly because of its favorable scaling with system size and the fact that it can be used efficiently on massively parallel supercomputers. One of the most powerful QMC techniques is based on a systematic ground-state projection of a carefully determined trial state [15], using the so-called diffusion Monte Carlo (DMC) with the fixed node approximation (FN). This choice represents a good compromise between accuracy and efficiency because FN is necessary for avoiding the well-known sign problem, and gives the best (i.e., the lowest energy) variational state with the same sign of the trial function. Despite this, FN remains a highly expensive computational tool, especially for systems containing nuclei with large atomic number Z .

Many sophisticated pseudopotentials have been developed and intensively used so far [16–21] in order to avoid an almost prohibitive computational cost. However, they are usually determined within other schemes and require further approximations (e.g., the locality [14]) that spoil the consistency of the method and often sacrifice the variational principle. We remark that several pseudopotentials used in QMC (e.g., the so-called BFD ones [22,23]) are based on the Hartree-Fock (HF) approximation that completely misses the correlation energy. Their use can be therefore justified only empirically and does not guarantee any consistency, namely that the FN energy differences are consistent with or without pseudopotentials.

All-electron calculations do not suffer from the above problems, but they are rarely applied for atoms of large atomic number Z in QMC due to the expensive computational cost. The major drawback is that, in the electronic wave function, the core and the valence regions are characterized by very different length scales. Therefore, within the most straightforward QMC algorithm, the smallest scale ($\sim Z^{-1}$) should be adopted for the proposed random displacement of the electrons, in order to avoid significant biases. How to treat different length (time) scales in a system is one of the central issues in computational physics, e.g., in molecular dynamics (RESPA [24,25]) and *ab initio* calculations. In the *ab initio* community, the Voronoi scheme [26], Becke fuzzy cells integration [27], and the multigrid technique [28] have been applied to accelerate the computation with an accurate quadrature in the vicinity of nuclei. Similarly, in the QMC community, Umrigar *et al.* have devised an accelerated Metropolis algorithm for the variational Monte Carlo (VMC) [29,30], by which electrons in the vicinity of nuclei are displaced with a step much shorter than the one used in the valence region by employing spherical polar coordinates. They also

*kousuke_1123@icloud.com

†sorella@sissa.it

devised a method for the diffusion Monte Carlo (DMC) [31] by reducing the velocity in the vicinity of nuclei to prevent from overshooting electrons. Despite that this improves the accuracy by a sizable amount, the major drawback of the conventional DMC is that the imaginary time step has to remain necessarily the same both for the valence and the core region [31] because it is difficult to decompose the propagation in imaginary time into the sum of multiple operators acting on different length scales. In this work, we will provide a solution to the different time (length) scale problem in DMC that is extremely relevant for its computational efficiency.

Lattice regularized diffusion Monte Carlo (LRDMC) can straightforwardly handle different length scales of a wave function [15,32], so that electrons in the vicinity of the nuclei and those in the valence region can be appropriately diffused. Henceforth this remedy is referred to as “double-grid LRDMC,” and “single-grid LRDMC” refers to the simpler version that adopts only a single lattice space as introduced in Ref. [32]. Although the double-grid LRDMC was claimed to solve the above drawback, no meaningful speedup has been achieved so far. Indeed, the original double-grid LRDMC has been used only for a very limited number of applications, specifically for light elements, such as carbon [32] and sodium [33] because, as shown later on, the original formulation leads to significant and uncontrolled errors for heavy elements. In the following, we introduce a strategy to handle different length and time scales in electronic ground-state calculations by having in mind the Thomas-Fermi theory, that is quite general and allows in particular a much more powerful double-grid LRDMC. After this appropriate treatment, this method becomes applicable to any element without introducing significant bias, while definitely accelerating the computation especially for large atomic number Z .

II. BOOSTING THE DOUBLE-GRID LRDMC

In LRDMC, the original continuous Hamiltonian is regularized by an approximate one H^a such that $H^a \rightarrow H$ for $a \rightarrow 0$, where a is the lattice mesh size used to discretize the continuous space [15,32]. Indeed, the kinetic part is approximated by a finite difference form:

$$\Delta_i \approx \Delta_i^{a,a'} \equiv \Delta_i^{a,p} + \Delta_i^{a',1-p}, \quad (1)$$

where $\Delta_i^{a,p}$ and $\Delta_i^{a',1-p}$ are discretized Laplacians by a small lattice space (a) and a large one (a'), respectively. The function $p(\vec{r})$, defining $\Delta_i^{a,p}$ and $\Delta_i^{a',1-p}$, parametrizes the probability to use the smaller and therefore more accurate lattice space (a) when an electron is close to a heavy nucleus. In the original work, $p(r)$ was chosen to be a simple Padé function [32]:

$$p(\vec{r}) = (1 + r_c^2 |\vec{r} - \vec{R}_c|^2)^{-1}; \quad (2)$$

instead a Gaussian-type function is employed in this work:

$$p(\vec{r}) = \exp\left(-\frac{|\vec{r} - \vec{R}_c|^2}{2r_c^2}\right), \quad (3)$$

where \vec{R}_c is the position of the nucleus closest to the electron in \vec{r} , and r_c is an important parameter determining the electrons treated with the smaller lattice space a (henceforth

referred to as core electrons), in other words, the ones inside the sphere of radius r_c (see Appendix A).

The key parameters of the double-grid LRDMC are a'/a and r_c . A smaller r_c and a larger ratio a'/a accelerate the double-grid scheme as compared with the corresponding single-grid one, whereas the bias (i.e., the difference between the single-grid and the double-grid LRDMC energies at the same a) is correspondingly increased. Therefore, a proper determination of the two parameters is essential to balance accuracy and efficiency of the double-grid LRDMC. a'/a was originally parametrized as

$$\frac{a'}{a} = \sqrt{\frac{Z^2}{4} + 1}, \quad (4)$$

with the Padé function $p(\vec{r})$ in Eq. (2) and $r_c = \frac{1}{2}Z$ [32], where Z is the atomic number considered. However, as it is shown in the following, the above choice is not suitable for large atomic number (Z).

In the following, we briefly describe a strategy to speed up a LRDMC calculation. First, we discuss how to properly determine a'/a . Since the computational cost of LRDMC is proportional to the inverse square of the lattice spaces (a and a'), the acceleration of the double-grid vs single-grid LRDMC (denoted as speedup) can be analytically estimated in terms of a'/a and the average number of electrons in the core/valence regions, according to the following relation:

$$\text{speedup}^{-1} = \frac{N_{\text{core}}(r_c)}{Z} + \frac{N_{\text{valence}}(r_c)}{Z} \left(\frac{a'}{a}\right)^{-2}, \quad (5)$$

where $N_{\text{core}}(r_c)$ and $N_{\text{valence}}(r_c) = Z - N_{\text{core}}(r_c)$ are the average numbers of electrons that are diffused with the smaller (a) and the larger (a') lattice spaces, respectively [34]. On physical grounds, the average numbers of core and valence electrons satisfy the inequality $N_{\text{core}} \ll N_{\text{valence}}$. On the other hand, the systematic error of the double-grid scheme referred to the corresponding single-grid one at the same a (denoted as bias) cannot be analytically estimated. This is because it is a very complicated function of a , a' , r_c , and Z . It is, however, possible to estimate an appropriate value according to the following consideration: (i) If we chose a too large value of a'/a , most of the computational time would be spent for the core electrons, and we could certainly decrease the bias by a smaller a' without affecting much the efficiency. (ii) On the other hand, if we chose a'/a (> 1) too close to 1, the bias would be minimal (i.e., equal to the single-grid LRDMC), but the speedup could be substantially increased by a larger a' without affecting the bias. Thus, we determine a'/a in a way that the speedup becomes an appropriate percentage of the maximum one [e.g., $\text{speedup} = \frac{1}{2} \left(\frac{N_{\text{core}}(r_c)}{Z}\right)^{-1}$], yielding

$$\frac{a'}{a} = \sqrt{\frac{lN_{\text{valence}}(r_c)}{N_{\text{core}}(r_c)}} \equiv \sqrt{\frac{l[Z - N_{\text{core}}(r_c)]}{N_{\text{core}}(r_c)}}, \quad (6)$$

where l is an efficiency parameter independent of Z . We found that $l = 2.0$ is an optimal value (see Appendix B in detail). The determination of a'/a [Eq. (6)] is an excellent compromise to balance acceleration and accuracy.

Next, we discuss a strategy to properly determine r_c . Since here we are interested in the asymptotic behavior of the

TABLE I. LRDMC energies of He, Be, Ne, Ar, Kr, and Xe atoms obtained by the single- and double-grid schemes at $a = (3.5Z)^{-1}$.

Element	Lattice space		Single grid Energy (Ha)	Double grid (this work)		Double grid (previous)	
	$a \equiv (\alpha Z)^{-1}$	α		Energy (Ha)	Bias (mHa) ^a	Energy (Ha)	Bias (mHa) ^a
He ($Z = 2$)	0.142 857	3.50	-2.903 732 1(62)	-2.903 759 7(64)	0.0(0.0)	-2.903 739 8(63)	0.0(0.0)
Be ($Z = 4$)	0.071 429	3.50	-14.667 247(31)	-14.667 294(32)	0.0(0.0)	-14.667 316(31)	0.1(0.0)
Ne ($Z = 10$)	0.028 571	3.50	-128.926 26(13)	-128.927 19(14)	0.9(0.2)	-128.927 45(14)	1.2(0.2)
Ar ($Z = 18$)	0.015 873	3.50	-527.495 42(18)	-527.497 56(19)	2.1(0.3)	-527.505 17(20)	9.7(0.3)
Kr ($Z = 36$)	0.007 937	3.50	-2753.771 51(78)	-2753.773 54(74)	2.0(1.1)	-2753.830 69(84)	59.2(1.1)
Xe ($Z = 54$)	0.005 291	3.50	-7234.8320(13)	-7234.838 45(88)	6.4(1.6)	-7235.0409(15)	208.9(2.0)

^aThe difference in total energy between the single- and double-grid LRDMCs.

algorithmic accuracy and efficiency for large Z , it is convenient to adopt the Thomas-Fermi approximation [35], according to which $N_{\text{core}}(r_c)$ for $Z^{1/3}r_c \ll 1$ is given by (see Appendix A in detail)

$$N_{\text{core}}(r_c) \propto (Zr_c)^{3/2}. \quad (7)$$

At this point, it is important to consider that the bias depends on the two lengths, namely, the value of r_c (the bias is minimal for $r_c \rightarrow \infty$) and the value of a' (the bias is minimal if $a' = a \simeq 1/Z$). Now, these two contributions are expected to be of the same order if we take $a' \simeq r_c$ because we can assume that for $r > r_c$, far from the core region, the wave function is smooth and the Laplacian can be discretized with a lattice space $a' \lesssim r_c$. This represents the most balanced choice, providing a good compromise between efficiency (smaller r_c and larger a') and accuracy (the other way around). With the above condition, by substituting the Thomas-Fermi expression of Eq. (7) in Eq. (6) for $a \propto 1/Z$, we obtain

$$r_c \gtrsim a' \simeq \frac{1}{Z} \sqrt{\frac{Z}{(Zr_c)^{3/2}}}, \quad (8)$$

yielding $r_c \propto Z^{-\theta}$ with $\theta = 5/7$. Therefore, our choice in the following is $r_c(Z) = \beta Z^{-5/7}$, where β is a Z -independent prefactor. Although the above discussion based on the Thomas-Fermi model is exact only for $Z \rightarrow \infty$, our VMC calculations show that the scaling [i.e., $N_{\text{core}}(r_c) \propto Z^{3/7}$] is empirically correct even for small Z (see Appendix A). The prefactor β should be small enough so that the scaling is valid in a wide range of Z values, even outside the asymptotic power-law regime. Therefore, $\beta = 0.75$ is employed in this study (see Appendix A).

As a summary, we determine $r_c(Z)$ according to $r_c(Z) = \beta Z^{-5/7}$, with $\beta = 0.75$. Then, a corresponding appropriate a'/a is determined according to Eq. (6). In this study, $N_{\text{core}}(r_c)$ and $N_{\text{valence}}(r_c)$ are estimated by the Slater's effective model [36] with the exponents that Clementi *et al.* proposed based on their HF calculations [37,38]. Since the computational cost of the all-electron single-grid DMC has turned out to scale with $Z^{5.5-6.5}$ [39-41] and the single-grid LRDMC similarly behaves, it is obviously very important to accelerate the double-grid LRDMC for heavy elements. In the following, we assume that the unbiased $a \rightarrow 0$ fixed node estimate can be obtained by a low-order polynomial fit of several energy calculations corresponding to different $a \gtrsim \frac{1}{Z}$. Since both terms in Eq. (5) are proportional to $Z^{-4/7}$ [$N_{\text{core}}(r_c) \propto Z^{3/7}$

and $d'/a \propto Z^{2/7}$], it is expected that the parametrization improves the complexity of the single-grid LRDMC by $\sim Z^{4/7} \simeq Z^{0.57}$, that represents a remarkable achievement especially for large Z .

III. PRACTICAL TEST OF THE DEVELOPED PARAMETRIZATION

In Table I, we show the LRDMC energies of He ($Z = 2$), Be ($Z = 4$), Ne ($Z = 10$), Ar ($Z = 18$), Kr ($Z = 36$), and Xe ($Z = 54$) atoms for $a = (3.5Z)^{-1}$ obtained by the single-grid (standard), the previous, and the double-grid parametrizations. These results indicate that the double-grid LRDMC energies obtained with the previous parametrization are significantly biased, especially for large atomic number Z . On the other hand, our parametrization suppresses these large biases, and the obtained LRDMC energies are essentially consistent with the single-grid ones for all Z , implying that the scaling law derived by means of the Thomas-Fermi model ($r_c \propto Z^{-5/7}$ and $d'/a \propto Z^{2/7}$) is in very good agreement with the numerical simulation. Thus, our double-grid LRDMC accelerates the computation without introducing significant biases, no matter how large is Z .

In practice, it is important to evaluate the actual computational time required to obtain a given absolute error in the total energy, as a function of the atomic number Z . We measured the computational times for He, Be, Ne, Ar, Kr, and Xe, wherein $a = (3.5Z)^{-1}$ is employed [42]. This is consistent with the typical setting of the time step in the standard DMC ($\tau \propto Z^{-2}$ [41]). Figure 1 shows that our parametrization accelerates the single-grid LRDMC calculations by $\times 1.1$, $\times 1.6$, $\times 2.8$, $\times 4.0$, $\times 5.8$, and $\times 7.5$ for He, Be, Ne, Ar, Kr, and Xe, respectively [43]. Our practical test shows that the single-grid LRDMC scales with $Z^{5.54}$, which is already slightly better than the previous report for the standard DMC algorithm ($Z^{5.97}$ with $\tau \propto Z^{-2}$ [41], where Umrigar's improvement [31] was employed). The double-grid parametrization improves the scaling of the LRDMC from $Z^{5.54}$ (single grid) to $Z^{4.95}$ (double grid) though it does not introduce significant bias, as described in the previous paragraph. The improvement of the scaling ($Z^{0.59}$) is consistent with our expectation ($Z^{4/7} \simeq Z^{0.57}$). To our best knowledge, $Z^{4.95}$ is the best scaling for the all-electron FN calculations so far.

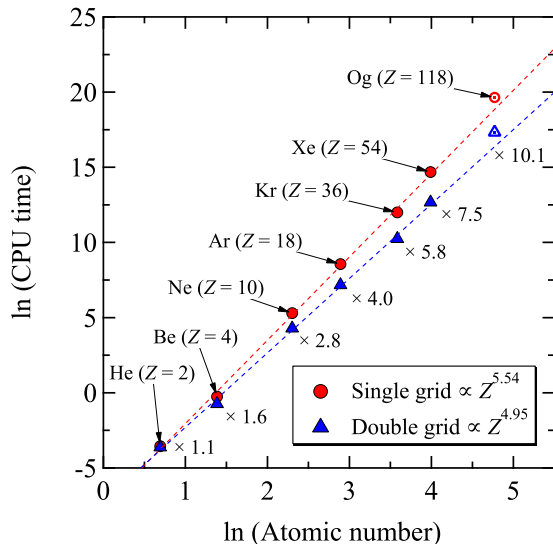


FIG. 1. Computational costs (CPU times) of the single- and double-grid LRDMC, measured at $a = (3.5Z)^{-1}$. The figure shows the logarithm of the CPU times required to obtain an absolute error bar (1.0 mHa) in the total energy. The CPU times were measured in seconds. The speedups are indicated below the blue triangles. The scalings described here were obtained without including the Og data.

IV. APPLICATION TO SUPERHEAVY ELEMENTS

We are now able to apply the LRDMC and its extrapolation (i.e., $a \rightarrow 0$) to superheavy elements ($Z \geq 104$) within a reasonable computational time. In this study, we deal with Og ($Z = 118$) that has the largest atomic number of all the known elements [44]. First, we evaluated the CPU time required to obtain an absolute error bar (1.0 mHa) in the total energy for Og. We found that, even when considering also Og, the expected polynomial scaling of the double-grid LRDMC remains, as shown in Fig. 1. The double-grid parametrization allows us to speed up LRDMC for Og by more than an order of magnitude (\sim ten times), so that the calculation is now possible and timely. Notice that, by including the Og data in the extrapolations, we obtained $Z^{5.66}$ and $Z^{5.10}$ scaling for the single- and double-grid LRDMC, respectively. They are a little worse than those without including the Og data (i.e., $Z^{5.54}$ and $Z^{4.95}$). This slight increase in the exponents might be related to the issue of polynomial scaling performances for large number of electrons, within the conventional branching algorithm adopted here [45]. These problems can probably be solved by adopting a more efficient path-integral approach [46]. Table II shows the *extrapolated* ($a \rightarrow 0$) nonrelativistic LRDMC energies of various elements including Og. The details of the extrapolations are shown in Appendixes C and D. The result shows that the correlation energies retrieved at the LRDMC level scales with $E_c \propto Z^{1.26}$, which is close to the exact one, $E_c \propto Z^{1.35}$ (up to Xe, see Appendix D). This is consistent with the previous study [41] that shows the DMC correlation energy scales with $E_c \propto Z^{1.26}$ up to Xe. Our result implies that the DMC scheme can recover a satisfactory correlation energy even when Z significantly increases. This is a great advantage of the DMC method compared with

TABLE II. Nonrelativistic LRDMC energies of He, Be, Ne, Ar, Kr, Xe, and Og obtained by the double-grid schemes with extrapolation ($a \rightarrow 0$). The Hartree-Fock and the corresponding correlation energies (E_c) retrieved by the LRDMC calculations are also shown. The unit of energy is Hartree.

Element	HF ^a	LRDMC	E_c
He ($Z = 2$)	-2.861 680	-2.903 720 0(53)	0.04
Be ($Z = 4$)	-14.573 02	-14.667 261(25)	0.09
Ne ($Z = 10$)	-128.547 10	-128.926 10(11)	0.38
Ar ($Z = 18$)	-526.817 51	-527.496 04(28)	0.68
Kr ($Z = 36$)	-2752.054 98	-2753.771 05(55)	1.72
Xe ($Z = 54$)	-7232.138 36	-7234.8350(10)	2.70
Og ($Z = 118$)	-46 324.355 82	-46 331.4380(53)	7.08

^aSee Ref. [47].

the quantum chemistry ones, because, even within the gold-standard CCSD(T) scheme, the correlation energy is rapidly going down when Z increases (e.g., $\sim 20\%$ for Kr [48]). The implementation of the relativistic effect [49–52] will make the double-grid LRDMC much more useful for studying electronic structures of large atomic number elements in the near future.

V. SUMMARY

In this study, we have introduced a powerful strategy to deal with the different time and length scales in electronic systems that has allowed us to establish an improved double-grid lattice regularized diffusion Monte Carlo (LRDMC) with careful balance between speedup and accuracy, yielding unprecedented computations even for large atomic number Z . The speedup of the LRDMC is predicted theoretically within the standard Thomas-Fermi model for atoms with large atomic number, and the calculation is indeed accelerated in practice by a large amount, especially for large atomic number Z . As a result, the computational scaling is improved from $Z^{\sim 5.5}$ (the single-grid LRDMC) to $Z^{\sim 5}$ (the double-grid LRDMC). The double-grid LRDMC has been applied to the largest superheavy element with more than an order of magnitude speedup. This technique can be used for polyatomic systems more efficiently (see Appendix E). Our solution is not based on a particular property of the DMC algorithm, rather on the general and accepted Thomas-Fermi theory in condensed matter physics. This implies that other electronic structure packages could also take advantage of the proposed approach for dealing with different length scales.

ACKNOWLEDGMENTS

The computations in this work have been mainly performed using the facilities of the Research Center for Advanced Computing Infrastructure at the Japan Advanced Institute of Science and Technology (JAIST). The authors appreciate helpful comments by A. Zen. K.N. is grateful for a financial support from the Simons Foundation. R.M. is grateful for financial support from MEXT-KAKENHI (Grants No. 19H04692 and No. 16KK0097), from FLAG-SHIP2020 (Projects No. hp190169 and No. hp190167 at

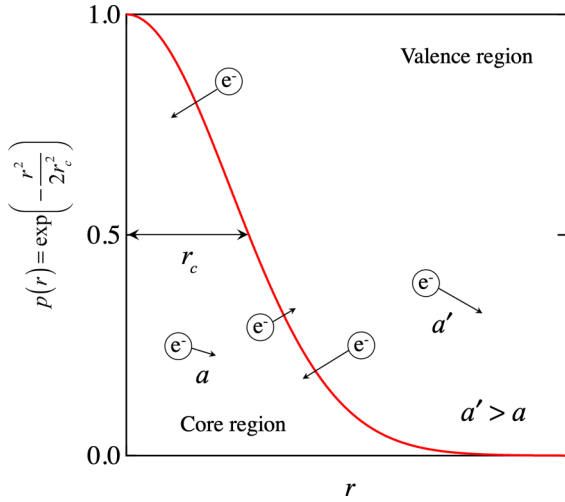


FIG. 2. A two-dimensional schematic picture of the double-grid LRDMC algorithm. Each electron at \vec{r} is displaced by a shorter (a) or a longer (a') lattice space according to the probability of $p(\vec{r})$ or $1 - p(\vec{r})$, where $0 \leq p(\vec{r}) \leq 1$. Since $p(\vec{r})$ is a function decaying suddenly far from nuclei (e.g., Gaussian function), most electrons in the vicinity of nuclei are displaced by a , and those far from nuclei are by a' .

K-computer), from Toyota Motor Corporation, from I-O DATA Foundation, from the Air Force Office of Scientific Research (AFOSR-AOARD/FA2386-17-1-4049; FA2386-19-1-4015), and from JSPS Bilateral Joint Projects (with India DST). S.S. acknowledges financial support by PRIN Project No. 2017BZPKSZ and computational resources by PRACE Project No. 2019204934.

APPENDIX A: THE SCALING OF THE NUMBER OF ELECTRONS WITHIN r_c OBTAINED BY THE THOMAS-FERMI MODEL

In the double-grid LRDMC, the kinetic part is approximated by a finite difference form:

$$\Delta_i \approx \Delta_i^{a,a'} \equiv \Delta_i^{a,p} + \Delta_i^{a',1-p}, \quad (\text{A1})$$

where $\Delta_i^{a,p}$ and $\Delta_i^{a',1-p}$ are discretized Laplacians by a small lattice space (a) and a large one (a'), respectively [32]. The function $p(\vec{r})$, defining $\Delta_i^{a,p}$ and $\Delta_i^{a',1-p}$, parametrizes the probability to use the smaller (a) when an electron is close to an heavy nucleus, and a large one (a') far from a nucleus. In this work, a Gaussian-type function was employed for $p(r)$:

$$p(\vec{r}) = \exp\left(-\frac{|\vec{r} - \vec{R}_c|^2}{2r_c^2}\right), \quad (\text{A2})$$

where \vec{R}_c is the position of the nucleus closest to the electron in \vec{r} , and r_c determines the sphere in which the electrons are treated with the smaller lattice space a (Fig. 2).

According to the Thomas-Fermi model [35], the electron density of the atomic number Z can be represented by

$$\rho(r) = Z^2 f\left(\frac{Z^{1/3}r}{b}\right), \quad (\text{A3})$$

where b is the constant value $b = (9\pi^2/128)^{1/3}$, $f(x)$ is

$$f(x) = \frac{32}{9\pi^3} \left(\frac{\chi(x)}{x}\right)^{3/2}, \quad (\text{A4})$$

and $\chi(x)$ is the universal function independent of Z . Therefore, when the Gaussian-type function [Eq. (A2)] is employed for $p(r)$, the number of electrons within r_c is defined by

$$N_{\text{core}}(r_c) = \int_0^\infty dr 4\pi r^2 Z^2 f\left(\frac{Z^{1/3}r}{b}\right) \exp\left(-\frac{r^2}{2r_c^2}\right), \quad (\text{A5})$$

where we assume $\vec{R}_c = 0$ in Eq. (A2). Equation (A5) can be rewritten using $Z^{1/3}r/b = x$:

$$N_{\text{core}}(r_c) = Z \int_0^\infty dx x^{1/2} \chi(x)^{3/2} \exp(-\xi x^2), \quad (\text{A6})$$

where $\xi = (bZ^{-1/3}/\sqrt{2}r_c)^2$. $\chi(x)$ can be approximated by the following polynominal expression at small x region ($x \ll 1$) [35]:

$$\chi(x) = 1 - Ax + \dots, \quad (\text{A7})$$

where A is the constant value $A = 1.8858$. If ξ is large enough (i.e., $Z^{1/3}r_c/b \ll 1$), only the small x region contributes to the integral and the high-order terms can be neglected. Therefore, the above equation can be approximated by

$$N_{\text{core}}(r_c) \simeq Z \int_0^\infty dx x^{1/2} \exp(-\xi x^2). \quad (\text{A8})$$

Since the integral can be replaced by the gamma function,

$$\int_0^\infty dx x^{2s-1} \exp(-\xi x^2) = \frac{1}{2} \Gamma(s) \xi^{-s}, \quad (\text{A9})$$

the number of electrons within r_c can be represented as:

$$N_{\text{core}}(r_c) \simeq \frac{1}{2} Z \Gamma\left(\frac{3}{4}\right) \xi^{-3/4} \equiv \frac{1}{2} \left(\frac{b}{\sqrt{2}}\right)^{-3/2} \Gamma\left(\frac{3}{4}\right) (Zr_c)^{3/2}. \quad (\text{A10})$$

By substituting r_c with $\beta Z^{-\theta}$, we finally get the relation

$$N_{\text{core}}(r_c) \simeq \frac{1}{2} (b\sqrt{2}\beta)^{-3/2} \Gamma\left(\frac{3}{4}\right) Z^{(3/2)(1-\theta)} \propto Z^{(3/2)(1-\theta)}, \quad (\text{A11})$$

and for $\theta = 5/7$:

$$N_{\text{core}}(r_c) \propto Z^{3/7}. \quad (\text{A12})$$

Equation (A12) is valid only when the inequality $\xi \gg 1$ is satisfied. This depends on the prefactor β as well as the atomic number Z . In practice, the prefactor β should be small enough so that $N_{\text{core}}(r_c) \propto Z^{3/7}$ is valid in a wide range of Z values, even outside the asymptotic power-law regime. Figure 3 shows the plot of $N_{\text{core}}(r_c)$ divided by $Z^{3/7}$ obtained by VMC calculations vs β for Ne, Ar, Kr, and Xe atoms. This figure shows that $\beta = 0.75$ is small enough to satisfy the above scaling.

The electron densities obtained by the VMC calculations were validated by comparison with the experimental atomic scattering factors (ASFs), as shown in Fig. 4. ASFs can be readily calculated using the electron densities obtained by the

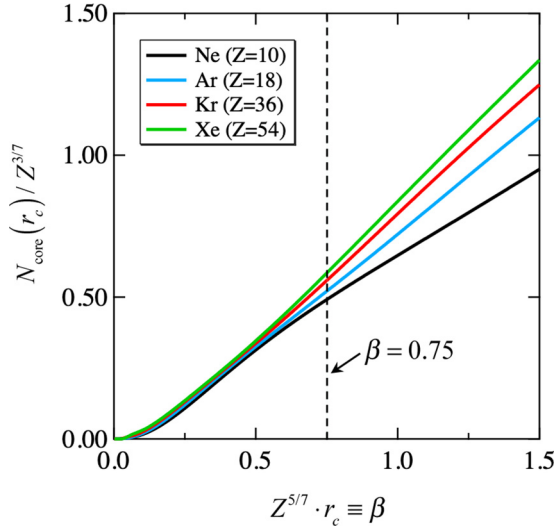


FIG. 3. The plot of $N_{\text{core}}(r_c)/Z^{3/7}$ obtained by VMC calculations v.s β , showing that the scaling ($N_{\text{core}}(r_c) \propto Z^{3/7}$) is satisfied at small β region. $\beta = 0.75$ is employed in the present work.

ab initio VMC calculations according to the following relation [53]:

$$\text{ASF} = \int_0^\infty 4\pi r^2 \rho(r) \frac{\sin kr}{kr} dr, \quad (\text{A13})$$

where $k = 4\pi \sin \theta / \lambda$, 2θ is the scattering angle and λ is the wavelength. TurboRVB enables us to calculate a radial distribution function as well as an electron density from a many-body wave function.

APPENDIX B: ADJUSTMENT OF THE EFFICIENCY PARAMETER l IN EQ. (6) IN THE MAIN TEXT

We compared the single-grid and the double-grid LRDMC energies by changing the efficiency parameter l in Eq. (6) in

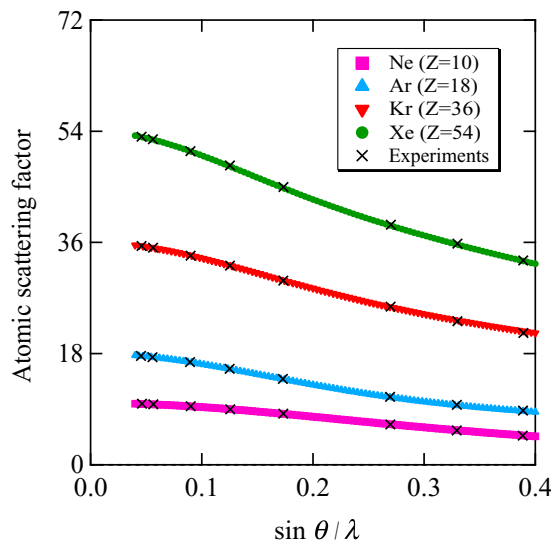


FIG. 4. Atomic scattering factors obtained by the VMC calculations, and those obtained by scattering x-ray measurements [53].

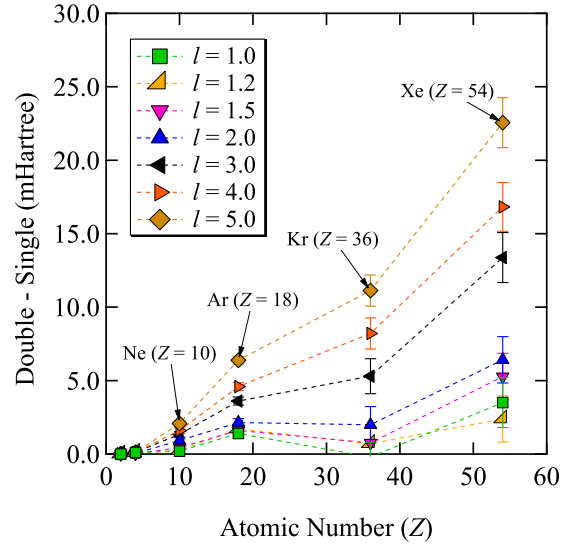


FIG. 5. The energy differences between the single-grid and the double-grid LRDMC for He, Ne, Ar, Kr, and Xe with various l in Eq. (6) of the main text.

the main text in order to find a optimal value. The result is shown in Fig. 5. Since the biases significantly increase as l becomes larger than 3.0, we have concluded that $l = 2.0$ is an optimal value to balance acceleration and accuracy.

APPENDIX C: A REMEDY FOR SMOOTH EXTRAPOLATIONS

In order to improve the quality of the energy extrapolation for $a \rightarrow 0$, it is important to increase r_c as a increases. This is because if r_c is fixed (i.e., a'/a is also fixed according to Eq. (6) in the main text), $a' \simeq r_c$ is no longer satisfied in a large a region, which introduces a large bias by the larger lattice space a' especially in the vicinity of the border between the core and valence regions. A simple parametrization to solve this problem is $r_c(a, Z) = r_c(Z)f(a)$, where $f(a)$ is an arbitrary function satisfying $f(0) = 1$ and $f(\infty) = \text{const}$. In this study, a simple polynomial function,

$$f(a) = \frac{\kappa(Z \cdot a)^2 + 1}{(Z \cdot a)^2 + 1} \equiv \frac{\kappa \cdot \alpha^{-2} + 1}{\alpha^{-2} + 1}, \quad (\text{C1})$$

is employed, where κ is a prefactor, and $a = (\alpha Z)^{-1}$. Therefore, $r_c(a, Z)$ can be parametrized as

$$r_c(a, Z) = r_c(Z)f(a) \equiv \frac{\beta(\kappa\alpha^{-2} + 1)}{\alpha^{-2} + 1} Z^{-5/7}. \quad (\text{C2})$$

Equation (6) in the main text indicates that $N_{\text{core}}(r_c)$ should be smaller than $N_{\text{valence}}(r_c)$ for any a and Z , otherwise the double-grid LRDMC becomes useless (i.e., $a'/a < 1$) in a certain case. According to the Thomas-Fermi theory [35], $N_{\text{core}}(r_c)$ becomes equal to $N_{\text{valence}}(r_c)$ at $r_c = 1.33Z^{-1/3}$. Therefore, the following inequality should be satisfied for all Z and a so that a'/a becomes larger than 1.0 even when l in Eq. (6) is set to 1.0,

$$r_c(a, Z) < 1.33Z^{-1/3} \quad (\forall a \in a > 0, \forall Z \in Z \geq 3). \quad (\text{C3})$$

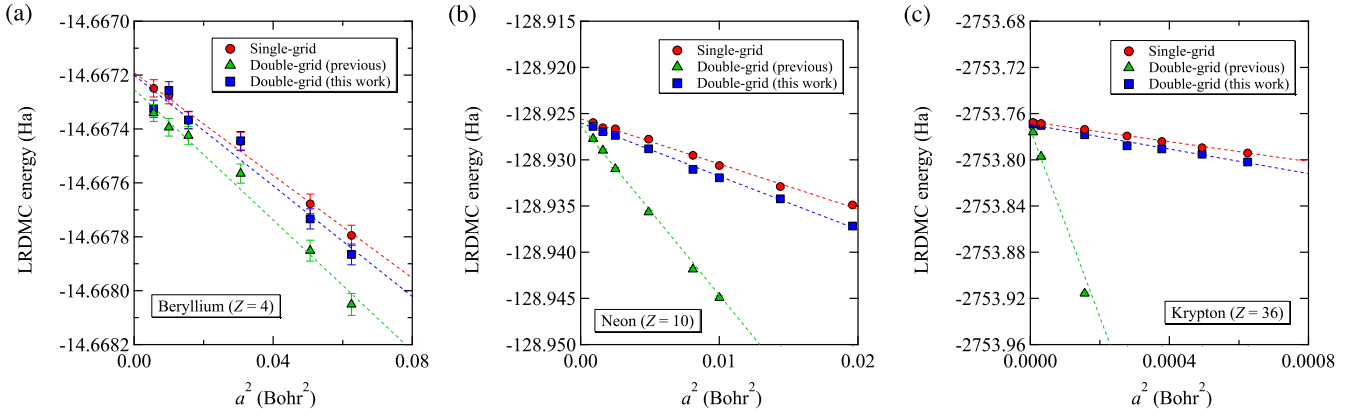


FIG. 6. LRDMC energies of (a) Be ($Z = 4$), (b) Ne ($Z = 10$), and (c) Kr ($Z = 36$) atoms obtained by the single-grid (red circles), the previous (green triangles), and the (blue squares) double-grid algorithms. The broken lines represent extrapolations with a quadratic function $E(a^2) = E_0 + ka^2$.

Thus, we obtain $\kappa < 2.69$ for $\beta = 0.75$. $\kappa = 2.5$ is employed here. The modified algorithm determines $r_c(a, Z)$ using Z and a according to Eq. (C2), then, a corresponding proper a'/a is calculated by the obtained $r_c(a, Z)$ according to Eq. (6) in the main text, wherein $N_{\text{core}}(r_c)$ and $N_{\text{valence}}(r_c)$ are estimated by the Slater's effective models [36] with the exponents that Clementi *et al.* proposed based on HF calculations [37,38].

In Fig. 6 and Table III, we show the LRDMC energies of Be ($Z = 4$), Ne ($Z = 10$), and Kr ($Z = 36$) atoms obtained by the single-grid, the previous, and the double-grid algorithms with the above remedy ($l = 1.0$, $\beta = 0.75$, $\kappa = 2.5$). The LRDMC energies were extrapolated with a quadratic function $E(a^2) = E_0 + ka^2$, as shown in Fig. 6. We remark that in Fig. 6 the LRDMC energies obtained by the previous parametrization are significantly biased in the large a region especially in Kr

TABLE III. LRDMC energies of Be, Ne, and Kr atoms obtained by the single- and double-grid schemes. These LRDMC calculations were performed by the double-grid scheme improved for smooth extrapolations ($a \rightarrow 0$) described in Appendix C ($l = 1.0$, $\beta = 0.75$, $\kappa = 2.5$). The obtained LRDMC energies were extrapolated with a quadratic function $E(a^2) = E_0 + ka^2$, as shown in Fig. 6.

Element	Lattice space		Single grid Energy (Ha)	Double grid (this work)				
	$a \equiv (\alpha Z)^{-1}$	α		Energy (Ha)	Bias (mHa) ^a	Acceleration ^b	a'/a	$r_c(a, Z)$
Be ($Z = 4$)	$a \rightarrow 0$	$\alpha \rightarrow \infty$	-14.667 191(22)	-14.667 200(22)	0.0(0)			
	0.08	3.33	-14.667 249(32)	-14.667 325(32)	0.1(0)	$\times 2.0$	1.733	0.313
	0.10	2.50	-14.667 274(32)	-14.667 257(32)	0.0(0)	$\times 1.9$	1.654	0.336
	0.13	2.00	-14.667 368(32)	-14.667 367(32)	0.0(0)	$\times 1.8$	1.580	0.362
	0.18	1.43	-14.667 446(34)	-14.667 444(34)	0.0(0)	$\times 1.6$	1.458	0.416
	0.23	1.11	-14.667 678(36)	-14.667 734(38)	0.1(1)	$\times 1.5$	1.372	0.466
	0.25	1.00	-14.667 795(38)	-14.667 866(38)	0.1(1)	$\times 1.4$	1.340	0.488
Ne ($Z = 10$)	$a \rightarrow 0$	$\alpha \rightarrow \infty$	-128.925 559(85)	-128.925 996(87)	0.2(2)			
	0.03	3.33	-128.925 97(14)	-128.926 38(14)	0.4(2)	$\times 3.5$	2.449	0.163
	0.04	2.50	-128.926 53(14)	-128.926 95(14)	0.4(2)	$\times 3.3$	2.352	0.175
	0.05	2.00	-128.926 65(14)	-128.927 36(14)	0.7(2)	$\times 3.1$	2.255	0.188
	0.07	1.43	-128.927 76(15)	-128.928 81(15)	1.0(2)	$\times 2.7$	2.080	0.216
	0.09	1.11	-128.929 52(15)	-128.931 04(16)	1.5(2)	$\times 2.5$	1.942	0.242
	0.10	1.00	-128.930 63(15)	-128.931 94(16)	1.3(2)	$\times 2.4$	1.887	0.253
	0.12	0.83	-128.932 89(17)	-128.934 22(17)	1.3(2)	$\times 2.2$	1.797	0.273
	0.14	0.71	-128.934 89(19)	-128.937 16(21)	2.3(4)	$\times 2.1$	1.731	0.289
Kr ($Z = 36$)	$a \rightarrow 0$	$\alpha \rightarrow \infty$	-2753.767 13(41)	-2753.769 45(46)	2.3(0.6)			
	0.00278	10.00	-2753.767 70(76)	-2753.768 91(65)	1.2(1.0)	$\times 7.0$	3.641	0.059
	0.00556	5.00	-2753.768 56(60)	-2753.770 25(64)	1.7(0.9)	$\times 6.7$	3.541	0.061
	0.01250	2.22	-2753.773 69(58)	-2753.778 37(86)	4.7(1.0)	$\times 5.5$	3.159	0.073
	0.01667	1.67	-2753.779 37(75)	-2753.787 84(69)	8.5(1.0)	$\times 4.8$	2.936	0.081
	0.01944	1.43	-2753.784 12(73)	-2753.790 63(73)	6.5(1.0)	$\times 4.5$	2.809	0.087
	0.02222	1.25	-2753.789 51(77)	-2753.795 14(73)	5.6(1.1)	$\times 4.2$	2.701	0.092
	0.02500	1.11	-2753.794 12(82)	-2753.801 90(85)	7.8(1.2)	$\times 4.0$	2.609	0.097

^aThe difference in total energy between the single- and double-grid algorithms.

^bThe accelerations were not measured by actual CPU times but the number of off-diagonal moves per a given time step t_{bra} .

TABLE IV. The nonrelativistic ground-state energies of He, Be, Ne, Ar, Kr, Xe, and Og. The LRDMC calculations shown here were performed by the double-grid scheme improved for smooth extrapolations ($a \rightarrow 0$) described in Appendix C ($l = 2.0$, $\beta = 0.75$, $\kappa = 2.5$). We extrapolated the LRDMC energies with a quadratic function $E(a^2) = E_0 + ka^2$, as shown in Figs. 7 and 8.

Element	Method	Total energy (Ha)	Correlation (%)
He ($Z = 2$) ^a	HF ^h	-2.861 680	0
	VMC-JSD ⁱ	-2.903 527(9)	99.53(2)
	DMC ⁱ	-2.903 719(2)	99.99(0)
	VMC-JSD	-2.903 351 8(79)	99.11(2)
	VMC-JAGP	-2.903 443 1(70)	99.33(2)
	LRDMC	-2.903 720 0(53)	99.99(1)
	Exact ^j	-2.903 724	100
Be ($Z = 4$) ^b	HF ^h	-14.573 02	0
	VMC-AGP ^k	-14.665 04(4)	97.54(0)
	DMC ^k	-14.667 26(1)	99.89(0)
	VMC-JAGP	-14.665 828(42)	98.38(0)
	LRDMC	-14.667 261(25)	99.89(0)
	Exact ^l	-14.667 36	100
Ne ($Z = 10$) ^c	HF ^h	-128.547 10	0
	VMC-JSD ⁱ	-128.891(5)	88(1)
	DMC ⁱ	-128.9231(1)	95.94(3)
	VMC-JSD	-128.898 03(12)	89.55(12)
	VMC-JAGP	-128.903 54(44)	90.95(11)
	LRDMC	-128.926 10(11)	96.71(3)
	Exact ^l	-128.939	100
Ar ($Z = 18$) ^d	HF ^h	-526.817 51	0
	VMC-JSD ⁱ	-527.3817(2)	77.02(3)
	DMC ⁱ	-527.4840(2)	90.99(3)
	VMC-JSD	-527.419 37(25)	82.17(5)
	VMC-JAGP	-527.431 64(37)	83.84(5)
	LRDMC	-527.496 04(28)	92.63(4)
	Exact ⁿ	-527.55	100
Kr ($Z = 36$) ^e	HF ^h	-2752.054 98	0
	VMC-JSD ⁱ	-2753.2436(6)	57.28(3)
	DMC ⁱ	-2753.7427(6)	81.34(3)
	VMC-JSD	-2753.614 20(46)	75.14(2)
	VMC-JAGP	-2753.628 41(44)	75.83(2)
	LRDMC	-2753.771 05(55)	82.70(3)
	Exact ^o	-2754.13	100
Xe ($Z = 54$) ^f	HF ^h	-7232.138 36	0
	VMC-JSD ⁱ	-7233.700(2)	45.51(6)
	DMC ⁱ	-7234.785(1)	77.12(3)
	VMC-JSD	-7234.507 30(87)	69.03(3)
	VMC-JAGP	-7234.564 94(82)	70.71(2)
	LRDMC	-7234.8350(10)	78.58(3)
	Exact ^o	-7235.57	100

($Z = 36$). On the other hand, our parametrization suppresses these significant biases, and the obtained LRDMC energies in the small a region ($a \rightarrow 0$) are essentially unbiased for all Z . In Table III, it is evident that $r_c(a, Z)$ increases as a increases, and a'/a decreases as a increases, by which the condition $a' \simeq r_c$ is satisfied for any a and Z . In this way, unnecessary large biases (i.e., not saving computational time) are suppressed, and smooth extrapolations are achieved. Notice that Table III indicates that this modified algorithm implies smaller

TABLE IV. (Continued.)

Element	Method	Total energy (Ha)	Correlation (%)
Og ($Z = 118$) ^g	HF ^h	-46 324.355 82	0
	VMC-JAGP	-46 330.7920(19)	66.92(5)
	LRDMC	-46 331.4380(53)	71.01(16)
	Exact ^p	-46 334.33	100

^aOur modified cc-pVDZ basis is composed of $3s1p$ ($Z \leq 5.77$) and $2s1p$ ($Z \leq 1.275$) for the determinant and Jastrow part, respectively.

^bOur modified cc-pVDZ basis is composed of $7s4p1d$ ($Z \leq 100.5$) and $5s3p1d$ ($Z \leq 9.17$) for the determinant and Jastrow part, respectively.

^cOur modified cc-pVDZ basis is composed of $6s4p1d$ ($Z \leq 173.5$) and $3s3p1d$ ($Z \leq 7.81$) for the determinant and Jastrow part, respectively.

^dOur modified cc-pVDZ basis is composed of $8s8p1d$ ($Z \leq 459.7$) and $6s6p1d$ ($Z \leq 64.69$) for the determinant and Jastrow part, respectively.

^eOur modified cc-pVDZ basis is composed of $11s11p6d$ ($Z \leq 6582.01$) and $7s8p4d$ ($Z \leq 129.00$) for the determinant and Jastrow part, respectively.

^fOur modified ADZP basis is composed of $15s16p9d2f$ ($Z \leq 19789.22$) and $11s12p7d$ ($Z \leq 335.98$) for the determinant and Jastrow part, respectively.

^gOur modified cc-pVDZ basis is composed of $20s20p16d10f$ ($Z \leq 73546.30$) and $18s16p12d7f$ ($Z \leq 15476.90$) for the determinant and Jastrow part, respectively.

^hSee Ref. [47].

ⁱSee Ref. [41].

^jSee Ref. [61].

^kSee Ref. [55].

^lSee Ref. [61].

^mSee Ref. [62].

ⁿSee Ref. [63].

^oSee Ref. [64].

^pThis value was estimated by the extrapolation with the correlation energies of He-Xe.

speedups as a increases. This is because $r_c(a, Z)$ [Eq. (C2)] becomes slightly larger than the original $r_c(Z)$ due to $f(a)$ as a increases. However, the effect is negligible in practice because small a calculations are much more important than large ones (i.e., the computational cost is proportional to a^{-2}). Thus, within the remedy, the double-grid algorithm achieves both acceleration and smooth extrapolation.

APPENDIX D: THE DETAILS OF VMC AND LRDMC CALCULATIONS

The variational (VMC) and lattice regularized quantum Monte Carlo (LRDMC) calculations for He, Be, Ne, Ar, Kr, Xe, Og, and C_6H_6 were performed using TurboRVB [54]. In the VMC calculations, we employed the Jastrow Slater (JSD) or the Antisymmetrized Geminal Power (JAGP) [55] ansatz. The ansatz is composed of a Jastrow and an antisymmetric part ($\Psi = J\Phi_{AS}$). The singlet antisymmetric part is denoted as the antisymmetrized geminal power (AGP) that reads

$$\Psi_{AGP}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \hat{A}[\Phi(\mathbf{r}_1^\uparrow, \mathbf{r}_1^\downarrow)\Phi(\mathbf{r}_2^\uparrow, \mathbf{r}_2^\downarrow)\dots\Phi(\mathbf{r}_{N/2}^\uparrow, \mathbf{r}_{N/2}^\downarrow)], \quad (D1)$$

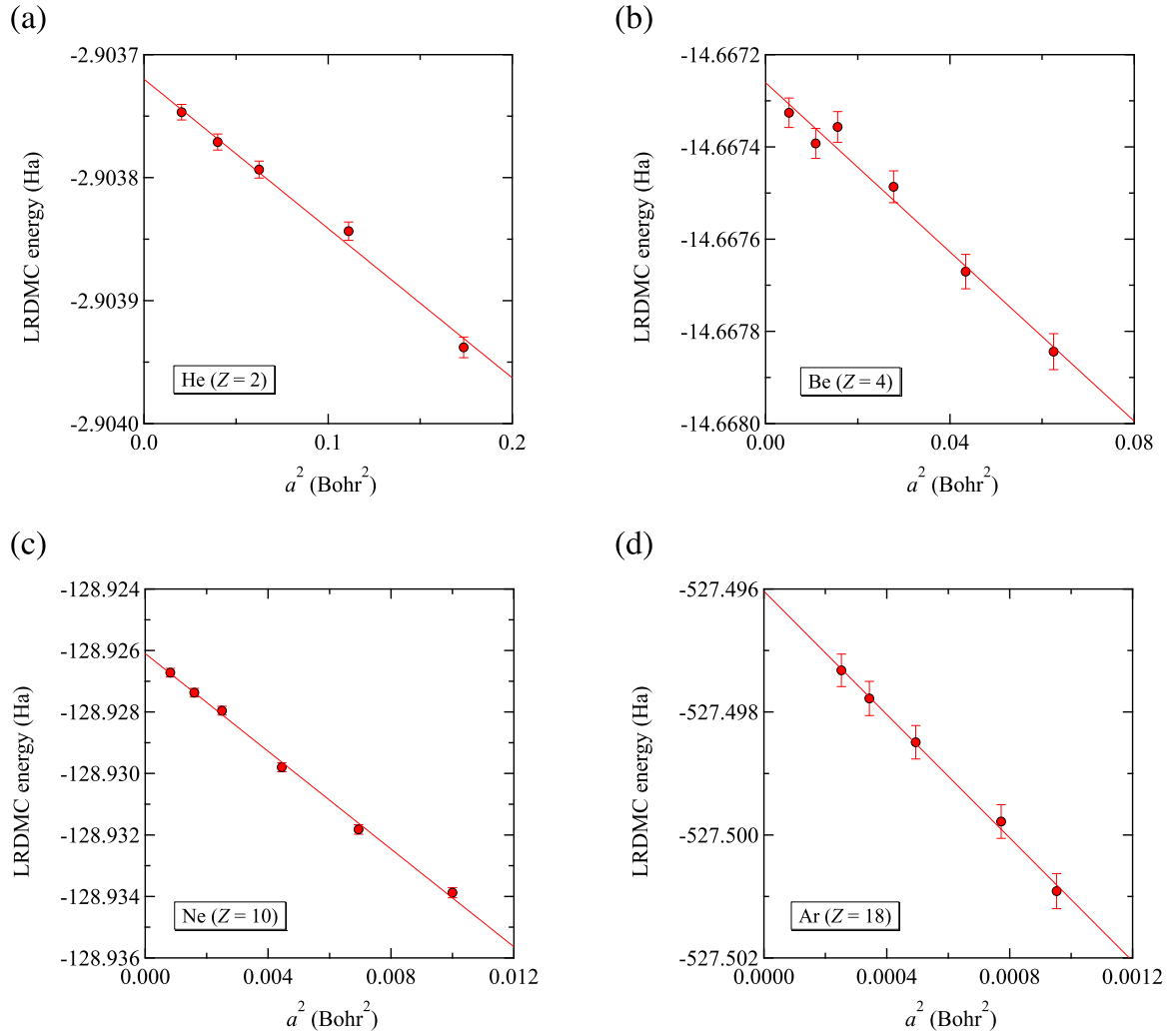


FIG. 7. LRDMC energies of (a) He ($Z = 2$), (b) Be ($Z = 4$), (c) Ne ($Z = 10$), and (d) Ar ($Z = 18$) atoms obtained by the double-grid scheme improved for a smooth extrapolation ($a \rightarrow 0$) described in Appendix C ($l = 2.0$, $\beta = 0.75$, $\kappa = 2.5$). The solid lines represent extrapolations with a quadratic function $E(a^2) = E_0 + ka^2$.

where \hat{A} is the antisymmetrization operator, and $\Phi(\mathbf{r}^\uparrow, \mathbf{r}^\downarrow)$ is called the pairing function. The spatial part of the geminal function is expanded over the Gaussian-type atomic orbitals:

$$\Phi_{\text{AGP}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{l,m,a,b} f_{\{a,l\},\{b,m\}} \psi_{a,l}(\mathbf{r}_i) \psi_{b,m}(\mathbf{r}_j), \quad (\text{D2})$$

where $\psi_{a,l}$ and $\psi_{b,m}$ are primitive Gaussian atomic orbitals, their indices l and m indicate different orbitals centered on atoms a and b , and i and j are coordinates of spin-up and -down electrons, respectively. When the JAGP is expanded over p molecular orbitals where p is equal to half of the total number of electrons ($N/2$), the JAGP coincides with the JSD [15,56]. The Jastrow term is composed of one-body, two-body, and three-/four-body factors ($J = J_1 J_2 J_{3/4}$). The one-body and two-body factors are used to fulfill the electron-ion and electron-electron cusp conditions, respectively, and the three-/four-body factor is employed to consider the further electron-electron correlation. The one-body Jastrow factor

reads

$$J_1(\mathbf{r}_1, \dots, \mathbf{r}_N) = \exp \left(\sum_{i,l,l} g_{l,l} \chi_{l,l}(\mathbf{r}_i) \right) \cdot \prod_i \tilde{J}_1(\mathbf{r}_i), \quad (\text{D3})$$

$$\tilde{J}_1(\mathbf{r}) = \exp \left(\sum_I -(2Z_I)^{3/4} u(2Z_I^{1/4} |\mathbf{r} - \mathbf{R}_I|) \right), \quad (\text{D4})$$

where \mathbf{r}_i are the electron positions, \mathbf{R}_I are the atomic positions with corresponding atomic number Z_I , l runs over atomic orbitals $\chi_{l,l}^I$ centered on the atom I , and $u(r)$ contains a variational parameter b :

$$u(r) = \frac{b}{2} (1 - e^{-r/b}). \quad (\text{D5})$$

Notice that the homogeneous one-body Jastrow part [$\tilde{J}_1(\mathbf{r})$] is also used to modify the single-particle orbital in DFT calculations to fulfill the electron-ion cusp condition explicitly

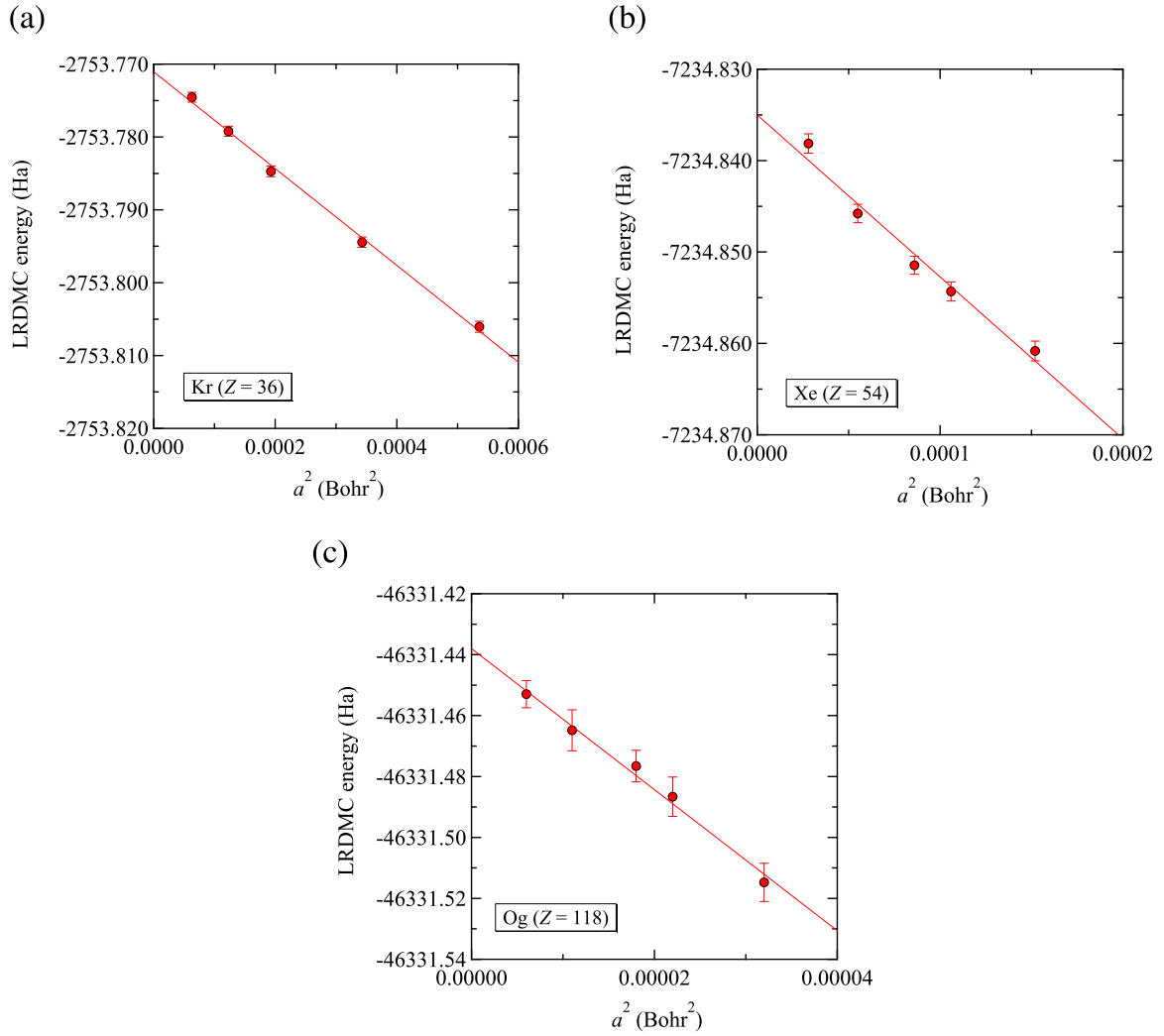


FIG. 8. LRDMC energies of (a) Kr ($Z = 36$), (b) Xe ($Z = 54$), and (c) Og ($Z = 118$) atoms obtained by the double-grid scheme improved for a smooth extrapolation ($a \rightarrow 0$) described in Appendix C ($l = 2.0$, $\beta = 0.75$, $\kappa = 2.5$). The solid lines represent extrapolations with a quadratic function $E(a^2) = E_0 + ka^2$.

[33]. The two-body Jastrow factor is defined as

$$J_2(\mathbf{r}_1, \dots, \mathbf{r}_N) = \exp\left(\sum_{i<j} v(|\mathbf{r}_i - \mathbf{r}_j|)\right), \quad (\text{D6})$$

where $v(r)$ is

$$v(r) = \frac{1}{2}r(1 - Fr)^{-1} \quad (\text{D7})$$

and F is a variational parameter. The three-body Jastrow factor is

$$J_{3/4}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \exp\left(\sum_{i<j} \Phi_{\text{Jas}}(\mathbf{r}_i, \mathbf{r}_j)\right), \quad (\text{D8})$$

and

$$\Phi_{\text{Jas}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{l,m,a,b} g_{a,l,m,b} \chi_{a,l}^{\text{Jas}}(\mathbf{r}_i) \chi_{b,m}^{\text{Jas}}(\mathbf{r}_j), \quad (\text{D9})$$

TABLE V. LRDMC energies of the benzene molecule (C_6H_6) obtained by the single and double-grid schemes at $a = (3.5Z_{\text{max}})^{-1}$.

Molecule	Lattice space		Single grid	Double grid (this work)		
	$a \equiv (\alpha Z_{\text{max}})^{-1}$	α	Energy (Ha)	Energy (Ha)	Bias (mHa) ^a	Acceleration ^b
C_6H_6	0.047 619	3.50	-232.192 58(55)	-232.194 24(56)	1.7(8)	$\times 1.9$

^aThe difference in total energy between the single- and double-grid algorithms.

^bThe acceleration of actual CPU time required for a fixed reference error in the total energy.

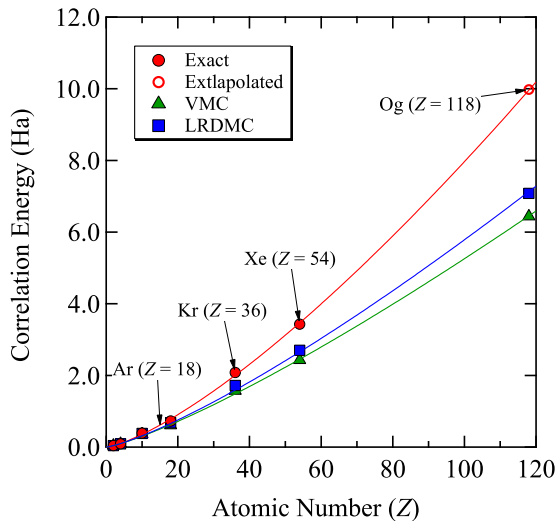


FIG. 9. The correlation energies E_c as a function of atomic number Z , giving $E_c \propto Z^{1.35}$, $E_c \propto Z^{1.24}$, and $E_c \propto Z^{1.26}$ for the exact values, VMC, and LRDMC calculations, respectively. Notice that the exact correlation energy of Og (the red blank circle) was estimated by the extrapolation with those of He-Xe.

where the indices l and m again indicate different orbitals centered on corresponding atoms a and b . In practice, the coefficients of the three-/four-body Jastrow factor become zero in atom calculations for $a \neq b$ because only one nucleus exists in a system. In the present study, we employed the cc-pVDZ (He, Be, Ne, Ar, Kr, and C_6H_6) and ADZP (Xe) basis sets taken from the Basis Set Exchange [57], and the nonrelativistic cc-pVDZ basis (Og) taken from the recent work [58], both for the determinant and the Jastrow parts. The variational JSD and JAGP wave functions were optimized using the stochastic configuration in combination with the linear method [59,60], by which all variational parameters in the Jastrow and the determinant parts including the exponents were optimized. In LRDMC calculations, the optimized JAGP wave functions were used for the guiding functions. Table IV shows the summary of our VMC and LRDMC calculations.

We employed the improved double-grid LRDMC for the extrapolations (Appendix C) with $l = 2.0$, $\beta = 0.75$, and $\kappa = 2.5$; on the other hand, we applied the simple double-grid LRDMC, namely $l = 2.0$, $\beta = 0.75$, and $\kappa = 1.0$ for the single-point calculations at $a = (3.5Z)^{-1}$ (Table I in the main text). The LRDMC energies were extrapolated with a quadratic function $E(a^2) = E_0 + ka^2$, as shown in Figs. 7 and 8. The exact correlation energies and those obtained by our VMC and LRDMC calculations are plotted in Fig. 9. Thanks to our careful optimizations, our VMC-JSD energies are lower than the previous results, especially when Z becomes larger. Remarkably, the JAGP ansatz further improves the variational energies. Our JAGP ansatz also improves the LRDMC energies (i.e., the nodal surfaces) as well as the variational energies.

APPENDIX E: AN APPLICATION OF THE DOUBLE-GRID ALGORITHM TO A POLYATOMIC SYSTEM

The double-grid algorithm can also be applied to polyatomic systems such as molecules and crystals. For a polyatomic system, the smallest length scale is determined by the heaviest atom in the system with $Z = Z_{\max}$. Therefore, in this case, we can change the definition of R_c in Eq. (3) slightly, by considering only the distances of the electrons with the heaviest atoms. In this way, when electrons are close to the lighter elements, they always move with the larger lattice space a' , without introducing a sizable bias. Conversely, for r_c , one can adopt the value calculated with a single reference heavy atom, as we have done in this work. It is clear, therefore, that a more significant speedup can be achieved by using Eq. (5), especially when the number of heavy atoms in the system is very small (e.g., transition-metal porphyrin complexes, metallofullerenes). As the first step to large systems, we considered the benzene molecule ($Z_{\max} = 6$). Table V shows that the bias of the double-grid LRDMC is as small as in the atomic cases while the computational time is accelerated by $\times 1.9$, significantly larger than the one $\simeq 1.5$ estimated from Fig. 1, demonstrating that the double-grid algorithm is already advantageous for polyatomic systems, even without too heavy nuclei and too many light ones.

- [1] B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen, *Nature (London)* **518**, 179 (2015).
- [2] S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, and A. Kis, *Nat. Rev. Mater.* **2**, 17033 (2017).
- [3] G. G. Naumis, S. Barraza-Lopez, M. Oliva-Leyva, and H. Terrones, *Rep. Prog. Phys.* **80**, 096501 (2017).
- [4] S. Sorella, K. Seki, O. O. Brovko, T. Shirakawa, S. Miyakoshi, S. Yunoki, and E. Tosatti, *Phys. Rev. Lett.* **121**, 066402 (2018).
- [5] J. Behler and M. Parrinello, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [6] J. Schmidt, J. Shi, P. Borlido, L. Chen, S. Botti, and M. A. Marques, *Chem. Mater.* **29**, 5090 (2017).
- [7] Y. Li, H. Li, F. C. Pickard IV, B. Narayanan, F. G. Sen, M. K. Chan, S. K. Sankaranarayanan, B. R. Brooks, and B. Roux, *J. Chem. Theory Comput.* **13**, 4492 (2017).
- [8] R. Kobayashi, D. Giofré, T. Junge, M. Ceriotti, and W. A. Curtin, *Phys. Rev. Mater.* **1**, 053604 (2017).
- [9] S. Zhang, J. Carlson, and J. E. Gubernatis, *Phys. Rev. Lett.* **74**, 3652 (1995).
- [10] G. H. Booth, A. Grüneis, G. Kresse, and A. Alavi, *Nature (London)* **493**, 365 (2013).
- [11] A. A. Holmes, H. J. Changlani, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 1561 (2016).
- [12] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, *J. Chem. Theory Comput.* **12**, 3674 (2016).
- [13] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [14] W. Foulkes, L. Mitás, R. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [15] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, 2017).

- [16] J. R. Trail and R. J. Needs, *J. Chem. Phys.* **142**, 064110 (2015).
- [17] J. T. Krogel, J. A. Santana, and F. A. Reboredo, *Phys. Rev. B* **93**, 075143 (2016).
- [18] J. R. Trail and R. J. Needs, *J. Chem. Phys.* **146**, 204107 (2017).
- [19] M. C. Bennett, C. A. Melton, A. Annaberdiyev, G. Wang, L. Shulenburg, and L. Mitas, *J. Chem. Phys.* **147**, 224106 (2017).
- [20] M. C. Bennett, G. Wang, A. Annaberdiyev, C. A. Melton, L. Shulenburg, and L. Mitas, *J. Chem. Phys.* **149**, 104108 (2018).
- [21] A. Annaberdiyev, G. Wang, C. A. Melton, M. Chandler Bennett, L. Shulenburg, and L. Mitas, *J. Chem. Phys.* **149**, 134108 (2018).
- [22] M. Burkatzki, C. Filippi, and M. Dolg, *J. Chem. Phys.* **126**, 234105 (2007).
- [23] M. Burkatzki, C. Filippi, and M. Dolg, *J. Chem. Phys.* **129**, 164115 (2008).
- [24] M. E. Tuckerman, G. J. Martyna, and B. J. Berne, *J. Chem. Phys.* **93**, 1287 (1990).
- [25] M. Tuckerman, B. J. Berne, and G. J. Martyna, *J. Chem. Phys.* **97**, 1990 (1992).
- [26] G. Te Velde and E. Baerends, *J. Comput. Phys.* **99**, 84 (1992).
- [27] A. D. Becke, *J. Chem. Phys.* **88**, 2547 (1988).
- [28] T. L. Beck, *Rev. Mod. Phys.* **72**, 1041 (2000).
- [29] C. J. Umrigar, *Phys. Rev. Lett.* **71**, 408 (1993).
- [30] M. Stedman, W. Foulkes, and M. Nekovee, *J. Chem. Phys.* **109**, 2630 (1998).
- [31] C. Umrigar, M. Nightingale, and K. Runge, *J. Chem. Phys.* **99**, 2865 (1993).
- [32] M. Casula, C. Filippi, and S. Sorella, *Phys. Rev. Lett.* **95**, 100201 (2005).
- [33] K. Nakano, R. Maezono, and S. Sorella, *J. Chem. Theory Comput.* **15**, 4044 (2019).
- [34] The computational cost discussed here is not an actual CPU time but the number of off-diagonal moves per a given time step t_{bra} . Actual CPU time is discussed later.
- [35] L. D. Landau and E. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Pergamon, London, 1958).
- [36] J. C. Slater, *Phys. Rev.* **36**, 57 (1930).
- [37] E. Clementi and D. L. Raimondi, *J. Chem. Phys.* **38**, 2686 (1963).
- [38] E. Clementi, D. L. Raimondi, and W. P. Reinhardt, *J. Chem. Phys.* **47**, 1300 (1967).
- [39] D. Ceperley, *J. Stat. Phys.* **43**, 815 (1986).
- [40] B. L. Hammond, P. J. Reynolds, and W. A. Lester, Jr., *J. Chem. Phys.* **87**, 1130 (1987).
- [41] A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. E* **71**, 066704 (2005).
- [42] This benchmark was performed using 8 Intel Xeon E5-2680v2-2.8 GHz CPUs (i.e., 320 cores) installed on a SGI cluster.
- [43] The accelerations of actual CPU times are smaller than those of acceptance ratios. This is because the double-grid LRDMC consumes more CPU times when computing the discretized Laplacians and potentials.
- [44] S. A. Giuliani, Z. Matheson, W. Nazarewicz, E. Olsen, P.-G. Reinhard, J. Sadhukhan, B. Schuettrumpf, N. Schunck, and P. Schwerdtfeger, *Rev. Mod. Phys.* **91**, 011001 (2019).
- [45] N. Nemeč, *Phys. Rev. B* **81**, 035119 (2010).
- [46] M. Boninsegni and S. Moroni, *Phys. Rev. E* **86**, 056712 (2012).
- [47] S. L. Saito, *At. Data Nucl. Data Tables* **95**, 836 (2009).
- [48] R. D. Johnson, III Computational Chemistry Comparison and Benchmark Database, Release 19, National Institute of Standards and Technology, <http://cccbdb.nist.gov> (accessed 7 November 2019).
- [49] Y. Nakatsuka, T. Nakajima, M. Nakata, and K. Hirao, *J. Chem. Phys.* **132**, 054102 (2010).
- [50] Y. Nakatsuka and T. Nakajima, *J. Chem. Phys.* **137**, 154103 (2012).
- [51] C. A. Melton, M. Zhu, S. Guo, A. Ambrosetti, F. Pederiva, and L. Mitas, *Phys. Rev. A* **93**, 042502 (2016).
- [52] C. A. Melton, M. C. Bennett, and L. Mitas, *J. Chem. Phys.* **144**, 244113 (2016).
- [53] D. R. Chipman and L. D. Jennings, *Phys. Rev.* **132**, 728 (1963).
- [54] S. Sorella, TurboRVB: Quantum Monte Carlo Software for Electronic Structure Calculations, <https://people.sissa.it/~sorella/web> (accessed 30 August 2019).
- [55] M. Casula and S. Sorella, *J. Chem. Phys.* **119**, 6500 (2003).
- [56] M. Marchi, S. Azadi, M. Casula, and S. Sorella, *J. Chem. Phys.* **131**, 154116 (2009).
- [57] B. P. Pritchard, D. Altarawy, B. T. Didier, T. D. Gibson, and T. L. Windus, *J. Chem. Inf. Model.* **59**, 4814 (2019).
- [58] P. Jerabek, O. R. Smits, J. Mewes, K. A. Peterson, and P. Schwerdtfeger, *J. Phys. Chem. A* **123**, 4201 (2019).
- [59] S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).
- [60] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Phys. Rev. Lett.* **98**, 110201 (2007).
- [61] C. L. Pekeris, *Phys. Rev.* **112**, 1649 (1958).
- [62] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, *Phys. Rev. A* **44**, 7071 (1991).
- [63] S. J. Chakravorty and E. R. Davidson, *J. Phys. Chem.* **100**, 6167 (1996).
- [64] E. Clementi and D. Hofmann, *J. Mol. Struct.: THEOCHEM* **330**, 17 (1995).