

RESEARCH ARTICLE | MARCH 27 2026

## Fast evaluation of unbiased atomic forces in *ab initio* variational Monte Carlo via the Lagrangian technique

Special Collection: [2025 JCP Emerging Investigators Special Collection](#)

Kousuke Nakano ; Stefano Battaglia ; Jürg Hutter



*J. Chem. Phys.* 164, 124113 (2026)

<https://doi.org/10.1063/5.0311160>



View  
Online



Export  
Citation

### Articles You May Be Interested In

Diffusion Monte Carlo method for barrier heights of multiple proton exchanges and complexation energies in small water, ammonia, and hydrogen fluoride clusters

*J. Chem. Phys.* (February 2024)

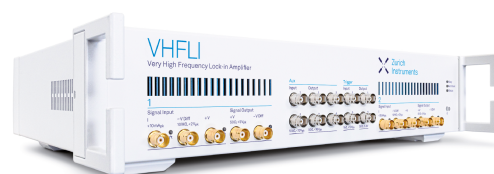
Correlation consistent effective core potentials for late 3d transition metals adapted for plane wave calculations

*J. Chem. Phys.* (November 2022)

A new generation of effective core potentials: Selected lanthanides and heavy elements II

*J. Chem. Phys.* (September 2025)

02 April 2026 00:30:04



Zurich  
Instruments

### Freedom to Innovate.

The New VHFLI 200 MHz Lock-in Amplifier.

Orchestrate pulses, triggers, and acquisition as the hub of your experiment.  
Discover more – run every signal analysis tool, simultaneously.

Order now

# Fast evaluation of unbiased atomic forces in *ab initio* variational Monte Carlo via the Lagrangian technique

Cite as: J. Chem. Phys. **164**, 124113 (2026); doi: [10.1063/5.0311160](https://doi.org/10.1063/5.0311160)

Submitted: 7 November 2025 • Accepted: 11 March 2026 •

Published Online: 27 March 2026 • Publisher Error Corrected: 30 March 2026



View Online



Export Citation



CrossMark

Kousuke Nakano,<sup>1,a)</sup>  Stefano Battaglia,<sup>2,b)</sup>  and Jürg Hutter<sup>2</sup> 

## AFFILIATIONS

<sup>1</sup>Center for Basic Research on Materials (CBRM), National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

<sup>2</sup>Department of Chemistry, University of Zurich (UZH), Winterthurerstrasse 190, 8057 Zurich, Switzerland

**Note:** This paper is a part of the 2025 JCP Emerging Investigators Special Collection.

<sup>a)</sup>Author to whom correspondence should be addressed: [kousuke\\_1123@icloud.com](mailto:kousuke_1123@icloud.com)

<sup>b)</sup>[me@stefanobattaglia.ch](mailto:me@stefanobattaglia.ch). Current address: Microsoft Research, AI for Science.

## ABSTRACT

*Ab initio* quantum Monte Carlo (QMC) methods are state-of-the-art electronic structure calculations based on highly parallelizable stochastic frameworks for accurate solutions of the many-body Schrödinger equation, suitable for modern many-core supercomputer architectures. Despite its potential, one of the major drawbacks that still hinders QMC applications, especially when targeting dynamical properties of large systems or extensive datasets, is the lack of an affordable method to compute atomic forces that are consistent with the corresponding potential energy surfaces (PESs), also known as *unbiased* atomic forces. Recently, one of the authors in the present paper proposed a way to obtain unbiased forces with the Jastrow-correlated Slater determinant *Ansatz*, where the determinant part is frozen to the values obtained by a mean-field method, such as density functional theory [K. Nakano, M. Casula, and G. Tenti, Phys. Rev. B **109**, 205151 (2024)]. However, the proposed method has a significant drawback for its applications: for a system with  $N$  nuclei, one requires  $6N$  additional density functional theory (DFT) calculations to get unbiased forces, which is not negligible as the system size increases. This paper presents a way to replace the  $6N$  DFT calculations with a single coupled-perturbed Kohn–Sham calculation, following the so-called *Lagrangian technique* established in quantum chemistry. This improves the computational cost and scalability of the method. We also demonstrate that the developed unbiased variational Monte Carlo (VMC) force calculation improves not only the consistency with PESs but also its accuracy, by investigating three molecules from the rMD17 benchmark set, and comparing the unbiased VMC forces with those obtained by the coupled-cluster singles and doubles with perturbative triples [CCSD(T)] calculations. We found that the bare VMC forces are biased from the CCSD(T) ones, while the unbiased ones give values closer to those of the CCSD(T) ones. Our benchmark test also reveals that the unbiased VMC forces yield very consistent values with hybrid and meta generalized gradient approximations (e.g.,  $\omega$ B97X-D3BJ and  $\omega$ B97M-D3BJ), but do not necessarily yield values that are very close to those of CCSD(T). Our finding paves the way to generate machine learning interatomic potentials based on VMC forces more efficiently and accurately.

© 2026 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0311160>

## I. INTRODUCTION

*Ab initio* quantum Monte Carlo (QMC) methods have been studied as accurate and highly parallelizable electronic structure methods that can leverage the capabilities of current and near-future exascale supercomputers.<sup>1</sup> While density functional theory (DFT) is

the most widely used *ab initio* method, it sometimes faces significant challenges in accurately describing peculiar systems, such as strongly correlated and dispersion-dominated materials. QMC methods are attracting attention as promising alternatives to address these limitations. Despite the maturity of energy calculations by QMC, the atomic force calculation (i.e., first derivatives of the total energy with

respect to atomic positions) has remained an open challenge, limiting the broader applicability of QMC. There are two main real-space QMC frameworks: variational Monte Carlo (VMC) and fixed-node diffusion Monte Carlo (FN-DMC).<sup>2</sup> In this study, we focus on VMC because the calculation of forces within the FN-DMC framework<sup>3</sup> is much more complicated. The FN-DMC force calculation is challenging because the FN wavefunction is not an eigenstate for the force estimator unlike for the Hamiltonian (cf. therefore, forces are computed using the so-called mixed-estimator) and the derivative of the FN wavefunction is hard to compute (cf. it is typically replaced with the trial WF).<sup>4–9</sup> Notice that other QMC frameworks, such as auxiliary-field quantum Monte Carlo (AFQMC) approach, can also compute atomic forces and pressures.<sup>10</sup> The methodological development for force evaluation in VMC has a long history, dating back to the 1980s. An *intrinsic* difficulty in computing atomic forces by VMC is the so-called *infinite-variance problem*: when the force is evaluated by Monte Carlo sampling, the variance of the force estimator diverges.<sup>11</sup> Over the past decades, this difficulty has been addressed by several complementary strategies. One of the earliest attempts was the space-warp coordinate transformation (SWCT), originally introduced by Umrigar.<sup>12</sup> This was generalized by Filippi and Umrigar within a correlated sampling framework,<sup>13</sup> and then, its differential formulations were developed by Assaraf and Caffarel<sup>14</sup> and Sorella and Capriotti.<sup>15</sup> Although these advances cure the divergence problem associated with the electron–ion Coulomb singularity, the other divergence arising from the nodal surface also had to be addressed.<sup>16</sup> Successful remedies are reweighting of the sampling distribution<sup>16,17</sup> and the tail regression estimator.<sup>18–20</sup> Related to the divergence problem, it was later revealed that one must also consider linear dependence of localized Gaussian basis sets (i.e., ill-conditioned overlap matrices) in periodic boundary condition calculations.<sup>21</sup> From an implementation point of view, the application of the adjoint algorithmic differentiation to atomic force evaluation<sup>15</sup> was a major milestone. It allows one to compute atomic forces without hand-coded analytic derivatives, dramatically reducing human effort in the derivative implementation.<sup>15</sup> Subsequently, a fast scheme for derivatives of multi-determinant wave functions was also formulated,<sup>22</sup> which has recently facilitated the generation of high-quality training data for machine learning force fields.<sup>23</sup> Very recently, neural network *Ansätze* combined with the VMC approach have been proposed to compute accurate atomic forces and pressures.<sup>24–26</sup> These developments have brought VMC force evaluations closer to practical utility.

Nevertheless, force calculations in VMC have been limited to small molecules,<sup>5,8,9,23,27–31</sup> clusters,<sup>32,33</sup> and crystals with light elements, such as solid/liquid hydrogen<sup>34–41</sup> and diamond.<sup>21</sup> This limitation likely comes from a *practical* issue underlying VMC-based force evaluations, which we will describe hereafter: Let  $\mathbf{R}_\alpha$  be the atomic position of the nucleus  $\alpha$ . The atomic force acting on  $\alpha$  is defined as the negative gradient of the energy with respect to  $\mathbf{R}_\alpha$ ,

$$\mathbf{F}_\alpha^{\text{VMC}} = -\frac{dE}{d\mathbf{R}_\alpha} = -\left\langle \frac{\partial}{\partial \mathbf{R}_\alpha} E_L \right\rangle \quad (1a)$$

$$-2 \left\langle (E_L - E) \frac{\partial \log \Psi}{\partial \mathbf{R}_\alpha} \right\rangle \quad (1b)$$

$$-\sum_{i=1}^{N_p} \frac{\partial E}{\partial p_i} \frac{dp_i}{d\mathbf{R}_\alpha}, \quad (1c)$$

where  $\Psi$  is an optimized wavefunction,  $\langle A \rangle$  indicates the quantum average of the local operator  $A$  over the sampling of  $|\Psi|^2$ ,  $E_L$  is the so-called local energy ( $E_L \equiv \hat{H}\Psi/\Psi$ ), with  $E \equiv \langle E_L \rangle$ , and  $\{p_1, \dots, p_{N_p}\}$  is the set of  $N_p$  variational parameters included in the  $\Psi$  *Ansatz*. Equations (1a)–(1c) are called the Hellmann–Feynman (HF), Pulay, and non-variational (NV) terms, respectively. When the wavefunction includes all relevant variational parameters and is fully optimized, the NV term vanishes.<sup>30,42</sup> However, especially for large systems or extensive datasets, optimization of all wavefunction variational parameters is often infeasible in practice. Therefore, the so-called Jastrow–Slater determinant (JSD) approach is typically employed, where the Slater determinant (SD) *obtained from mean-field methods* such as DFT is kept fixed, and only the Jastrow factor is variationally optimized. Historically, even in this case, the NV term was often neglected on the assumption that its contribution was small and forces were computed using only the HF and Pulay terms. However, recent studies have demonstrated that ignoring the NV term can lead to significant biases.<sup>30,31,42</sup> The bias arising from the NV term is referred to as the self-consistency error (SCE).<sup>30</sup> In 2021, Tiihonen *et al.*<sup>30</sup> reported that the SCE becomes more severe for heavier elements. Follow-up studies confirmed that the SCE for liquid hydrogen is on the order of 1 GPa,<sup>43</sup> and it reaches up to 5 GPa for cubic boron nitride (c-BN).<sup>42</sup> This implies that applying VMC to realistic systems beyond model compounds should address the SCE issue; otherwise, it becomes a critical bottleneck in VMC applications.

Nakano *et al.* (one of the authors of this present paper) recently formulated a method to eliminate the SCE.<sup>42</sup> The idea is conceptually straightforward: a direct evaluation of the NV term using a hybrid VMC-DFT approach. The partial derivatives  $\partial E/\partial p_i$  are computed via standard VMC, while the wavefunction response term  $dp_i/d\mathbf{R}_\alpha$  is evaluated using DFT by using the finite difference method (FDM). Using this scheme, unbiased forces have been successfully obtained for various systems, such as H<sub>2</sub>, Cl<sub>2</sub>, c-BN,<sup>42</sup> and liquid hydrogen.<sup>39</sup> Although this methodology successfully demonstrated a proof-of-concept for SCE elimination,<sup>39,42</sup> a significant practical limitation still remains. The problem is that *6N additional DFT calculations are required* to compute unbiased forces on all  $N$  atoms.<sup>42</sup> While this is acceptable for small systems, it becomes computationally prohibitive when targeting large systems or preparing extensive datasets for machine-learning potential (MLP) training.

This paper presents a method to eliminate the *6N* additional DFT calculations when correcting the NV term to get unbiased VMC forces. The key idea is the following: the bias arising from the NV term is not only a specific problem in VMC calculations but also in wavefunction theory (WFT) calculations when non-variational wavefunctions, such as those used in perturbative methods, are used to compute forces. One well-established approach in WFT is the so-called *Lagrangian technique*.<sup>44,45</sup> In this work, we extend the Lagrangian technique to the VMC framework, enabling the unbiased atomic force and pressure calculations for both molecular and periodic systems by a single-shot VMC and a single-shot coupled-perturbed Hartree–Fock (CPHF) or coupled-perturbed Kohn–Sham (CPKS) calculation. This eliminates the need for *6N* DFT evaluations and improves the feasibility of the method. Furthermore, we demonstrate that the developed unbiased force evaluation improves not only the consistency but also the accuracy of the VMC forces with the Jastrow–Slater determinant *Ansatz*. We benchmarked our

method on three molecules, ethanol, malonaldehyde, and benzene, taken from the rMD17 benchmark set,<sup>46,47</sup> and compare biased and unbiased VMC forces to those obtained with the coupled-cluster singles and doubles with perturbative triples [CCSD(T)]<sup>48</sup> method, known as the gold-standard in quantum chemistry calculations.

This paper is organized as follows. In Sec. II, we describe the details of the Lagrangian technique for the VMC framework. In Sec. III, we show the validation of our formalism and implementations in CP2K<sup>49</sup> and TurboRVB<sup>50</sup> software suites. In Sec. IV, we report benchmark results of the unbiased VMC force calculations for the three molecules taken from the rMD17 benchmark set. In Sec. V, we conclude our work and discuss the current limitations and future studies of the developed method.

## II. METHODS: LAGRANGIAN TECHNIQUE EXTENDED TO VMC

The key idea of the Lagrangian technique<sup>44,45</sup> is to construct a function, the *Lagrangian*, which has the same energy as the VMC one, but that incorporates all the constraints imposed during the optimization of the wavefunction. By making the Lagrangian stationary with respect to all its parameters, we then obtain a fully variational expression for which we can compute the nuclear gradient directly. The VMC energy depends on a set of Jastrow variational parameters,  $\{p_i\}_{i=1}^{N_f}$ , collected in the vector  $\mathbf{p}$ , and the occupied molecular orbital (MO) coefficients of the Slater determinant (SD) denoted by the tensor  $\mathbf{C}$ , with elements  $C_{\mu i}^{\sigma}$ . The indices  $\mu, \nu, \lambda$ , and  $\delta$  label atomic orbitals (AOs); the indices  $i, j, k$ , and  $l$  label occupied MOs; and  $\sigma, \tau$  label the spin. Sums over these indices automatically imply complete sums over the total number of occupied orbitals  $N_{\text{MO}}$ , atomic orbitals  $N_{\text{AO}}$ , and the number of spin channels  $N_{\text{spin}}$ , respectively. The overlap matrix between AOs or MOs (depending on its subscripts) is denoted by  $S$ . In the following, we always label the atomic and molecular orbitals as subscripts and the spin index as a superscript, where possible.

Let us start by introducing the VMC Lagrangian,

$$L_{\text{VMC}}(\mathbf{p}, \mathbf{C}, \mathbf{Z}, \mathbf{W}) = E_{\text{VMC}}(\mathbf{p}, \mathbf{C}) + \sum_{\lambda k \tau} Z_{\lambda k}^{\tau} \frac{\partial L_{\text{SCF}}(\mathbf{C})}{\partial C_{\lambda k}^{\tau}} - \sum_{kl\tau} W_{kl}^{\tau} (S_{kl}^{\tau} - \delta_{kl}). \quad (2)$$

The first term is the VMC energy; the second term corresponds to the stationary conditions of the underlying SCF reference calculation; and the third term enforces the orthonormality of the occupied MOs. The tensors  $\mathbf{Z}$  and  $\mathbf{W}$  contain the Lagrange multipliers. We do not give the explicit form of  $L_{\text{SCF}}$  here because only its derivative (defined later) is needed. In the most general case, all boldfaced quantities represent tensors of rank 3, with dimension  $N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}}$  for  $\mathbf{C}$  and  $\mathbf{Z}$  and dimension  $N_{\text{MO}} \times N_{\text{MO}} \times N_{\text{spin}}$  for  $\mathbf{W}$ . In case of restricted calculations, where the MO coefficients are the same for both spin channels,  $N_{\text{spin}} = 1$ , and the tensors reduce to matrices. Several equivalent formulations of the SCF term of the Lagrangian exist. In this work, we closely follow the atomic orbital formulation used for the implementation of density-corrected DFT gradients in CP2K.<sup>51</sup> This results in the Roothaan–Hall equations for the occupied MOs,

$$\frac{\partial L_{\text{SCF}}}{\partial C_{\lambda k}^{\tau}} = \sum_{\nu} F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau}, \quad (3)$$

where  $F_{\lambda \nu}^{\tau}$  is the Fock/Kohn–Sham matrix and  $\epsilon_k^{\tau}$  is the Kohn–Sham eigenvalue of  $k$ -th MO with the spin  $\tau$ . The explicit form of the Fock/Kohn–Sham matrix is introduced later. When the Lagrangian is stationary with respect to all its parameters, we can rewrite  $\mathbf{F}_{\alpha}^{\text{VMC}}$  with  $L_{\text{VMC}}$  as

$$\mathbf{F}_{\alpha}^{\text{VMC}} = -\frac{dE_{\text{VMC}}}{d\mathbf{R}_{\alpha}} = -\frac{dL_{\text{VMC}}}{d\mathbf{R}_{\alpha}} = -\frac{\partial L_{\text{VMC}}}{\partial \mathbf{R}_{\alpha}}. \quad (4)$$

To make the Lagrangian stationary with respect to all its parameters, we have to ensure the following set of equations is satisfied:

$$\frac{\partial L_{\text{VMC}}}{\partial p_i} = \frac{\partial E_{\text{VMC}}}{\partial p_i} = 0, \quad (5a)$$

$$\frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} = \frac{\partial E_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} + \sum_{\lambda k \tau} Z_{\lambda k}^{\tau} \frac{\partial^2 L_{\text{SCF}}}{\partial C_{\mu i}^{\sigma} \partial C_{\lambda k}^{\tau}} - \sum_{kl\tau} W_{kl}^{\tau} \frac{\partial S_{kl}^{\tau}}{\partial C_{\mu i}^{\sigma}} = 0, \quad (5b)$$

$$\frac{\partial L_{\text{VMC}}}{\partial Z_{\lambda k}^{\tau}} = \sum_{\nu} F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau} = 0, \quad (5c)$$

$$\frac{\partial L_{\text{VMC}}}{\partial W_{kl}^{\tau}} = S_{kl}^{\tau} - \delta_{kl} = 0. \quad (5d)$$

Equation (5a) is the variational condition on the VMC energy with respect to the Jastrow parameters. For an optimized Jastrow wave function, this condition is automatically fulfilled. Equation (5b) is zero by virtue of the Lagrange multipliers; that is, we *choose* the elements  $Z_{\lambda k}^{\tau}$  and  $W_{kl}^{\tau}$  such that  $\frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} = 0$  for  $\forall \mu, i, \sigma$ . Equations (5c) and (5d) are automatically fulfilled if we use canonical orbitals from the underlying SCF calculation.

At first sight, it seems that Eq. (5b) is not sufficient to determine both  $\mathbf{Z}$  and  $\mathbf{W}$ . However, only certain sub-blocks are different from zero, such that by projecting the stationary conditions onto occupied and virtual orbital spaces, we are able to obtain enough equations for their determination. To this end, let us introduce the projection operator  $\mathbf{Q}$ , with elements,

$$Q_{\mu\nu}^{\sigma} = \delta_{\mu\nu} - \sum_{\lambda k} C_{\mu k}^{\sigma} C_{\lambda k}^{\sigma} S_{\lambda\nu}. \quad (6)$$

The Lagrange multipliers  $Z_{\lambda k}^{\tau}$  can be best understood as expansion coefficients of the orbital response due to the correlation energy introduced by the Jastrow factor. Hence, only the occupied-virtual blocks contribute to the response, while the occupied-occupied block is assumed to be zero,

$$Z_{ij}^{\sigma} = \sum_{\mu\nu} Z_{\mu i}^{\sigma} S_{\mu\nu} C_{\nu j}^{\sigma} = 0, \quad (7)$$

as unitary rotations among occupied orbitals leave the energy unchanged. It is convenient at this stage to introduce the explicit form of the Fock/Kohn–Sham matrix,

$$F_{\mu\nu}^{\sigma}[\mathbf{P}] = h_{\mu\nu} + \sum_{\lambda\delta\tau} P_{\lambda\delta}^{\tau} [(\mu\nu|\lambda\delta) - c_x \delta_{\sigma\tau} (\mu\lambda|\nu\delta)] + V_{\mu\nu}^{\text{xc},\sigma}, \quad (8)$$

where  $(\mu\nu|\lambda\delta)$  are two-electron integrals over primitive atomic orbitals,

$$(\mu\nu|\lambda\delta) = \iint \psi_\mu(\mathbf{r})\psi_\nu(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_\lambda(\mathbf{r}')\psi_\delta(\mathbf{r}') d\mathbf{r}d\mathbf{r}'; \quad (9)$$

and  $h_{\mu\nu}$  values represent the kinetic energy and nuclear electron attraction integrals. This form is completely general and accommodates Hartree–Fock, pure semilocal functionals, and hybrid functionals using the exact exchange parameter  $c_x$  and the exchange–correlation potential  $V_{\mu\nu}^{\text{xc},\sigma}$ . The tensor  $\mathbf{P}$  contains the spin-density matrices in AO basis, with elements,

$$P_{\mu\nu}^\sigma = \sum_i C_{\mu i}^\sigma C_{\nu i}^\sigma. \quad (10)$$

The stationary conditions for the MO coefficients are given by equating to zero the derivatives of  $L_{\text{VMC}}$  with respect to  $C_{\mu i}^\sigma$ ,

$$\begin{aligned} \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^\sigma} &= \frac{\partial E_{\text{VMC}}}{\partial C_{\mu i}^\sigma} + \sum_{\lambda k \tau} Z_{\lambda k}^\tau \sum_\nu \frac{\partial}{\partial C_{\mu i}^\sigma} (F_{\lambda\nu}^\tau C_{\nu k}^\tau - S_{\lambda\nu} C_{\nu k}^\tau \epsilon_k^\tau) \\ &- \sum_{kl \tau} W_{kl}^\tau \frac{\partial S_{kl}^\tau}{\partial C_{\mu i}^\sigma} = 0. \end{aligned} \quad (11)$$

To arrive at working equations to determine the multipliers, we project Eq. (11) onto the occupied orbital space using  $\mathbf{C}$  and onto the unoccupied orbital space using  $\mathbf{Q}$ ,

$$\sum_\mu \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^\sigma} Q_{\mu\nu}^\sigma = 0 \rightarrow Z_{\nu i}^\sigma \quad \forall \nu, i, \sigma, \quad (12a)$$

$$\sum_\mu \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^\sigma} C_{\mu j}^\sigma = 0 \rightarrow W_{ij}^\sigma \quad \forall i, j, \sigma. \quad (12b)$$

In the following, we report only the final expressions obtained from Eqs. (12a) and (12b); their explicit derivation is given in Appendix B. The projection onto the virtual space, Eq. (12a), results in a linear system of equations  $\mathbf{AZ} = -\mathbf{B}$ , known as the Z-vector equations, from which the Lagrange multipliers  $\mathbf{Z}$  can be determined,

$$\sum_\lambda Z_{\lambda i}^\sigma (F_{\lambda\nu}^\sigma - S_{\lambda\nu} \epsilon_i^\sigma) + \sum_\mu H_{\mu i}^\sigma[\mathbf{P}] Q_{\mu\nu}^\sigma = -\sum_\mu B_{\mu i}^\sigma Q_{\mu\nu}^\sigma. \quad (13)$$

The elements of the tensor  $\mathbf{B}$  are the derivatives of the VMC energy with respect to the MO coefficients,

$$B_{\mu i}^\sigma = \frac{\partial E_{\text{VMC}}}{\partial C_{\mu i}^\sigma}, \quad (14)$$

while the tensor  $\mathbf{A}$  contains quantities related to the SCF stationary conditions and the MOs orthonormalization constraints; hence, it does not depend on the VMC calculation. The elements  $H_{\mu i}^\sigma[\mathbf{P}]$  are derivatives of the Fock matrix with respect to the MO coefficients,

$$H_{\mu i}^\sigma[\mathbf{P}] = \sum_{\lambda\nu\tau} \bar{P}_{\lambda\nu}^\tau (2(\lambda\nu|\mu i_\sigma) - c_x \delta_{\tau\sigma} [(\lambda\mu|\nu i_\sigma) + (\lambda i_\sigma|\nu\mu)]) + 2f_{\lambda\nu\tau,\mu i_\sigma}^{\text{xc}}; \quad (15)$$

where  $f_{\lambda\nu\tau,\mu i_\sigma}^{\text{xc}}$  is the exchange–correlation kernel, with an index contracted with a molecular orbital, and the tensor  $\bar{\mathbf{P}}$  is a symmetrized

form of the response density matrix (for increased numerical stability), given by

$$\bar{P}_{\mu\nu}^\sigma = \frac{1}{2} \sum_i (Z_{\mu i}^\sigma C_{\nu i}^\sigma + C_{\mu i}^\sigma Z_{\nu i}^\sigma). \quad (16)$$

Once the response tensor  $\mathbf{Z}$  is determined by solving Eq. (13), we can also determine the second set of Lagrange multipliers  $\mathbf{W}$ , which are given by

$$W_{ij}^\sigma = \frac{1}{2} H_{ij}^\sigma[\mathbf{P}] + \frac{1}{2} \sum_\mu B_{\mu i}^\sigma C_{\mu j}^\sigma. \quad (17)$$

Finally, the atomic force is given by the negative gradient of the Lagrangian with respect to the nuclear positions, reading

$$\frac{\partial L_{\text{VMC}}}{\partial \mathbf{R}_\alpha} = \frac{\partial E_{\text{VMC}}}{\partial \mathbf{R}_\alpha} - \sum_{\mu\nu\sigma} \Lambda_{\mu\nu}^\sigma \frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_\alpha} + \sum_{\mu\nu j \sigma} Z_{\mu j}^\sigma \frac{\partial F_{\mu\nu}^\sigma}{\partial \mathbf{R}_\alpha} C_{\nu j}^\sigma, \quad (18)$$

with the intermediate quantity  $\Lambda_{\mu\nu}^\sigma$  defined as

$$\Lambda_{\mu\nu}^\sigma = \sum_{ij} C_{\mu i}^\sigma W_{ij}^\sigma C_{\nu j}^\sigma + \frac{1}{2} \sum_i \epsilon_i^\sigma (Z_{\mu i}^\sigma C_{\nu i}^\sigma + C_{\mu i}^\sigma Z_{\nu i}^\sigma), \quad (19)$$

embodying the conditions of orthonormality between the response MOs and the so-called energy-weighted response density matrix.

An important aspect of the above-mentioned Lagrangian approach in the context of VMC is that the elements of the right-hand side tensor  $\mathbf{B}$  of Eq. (13) (i.e., the energy derivative with respect to the MO coefficients) come with statistical noise. It is important to know how this noise propagates from the VMC energy derivative to the response vectors and, ultimately, to the forces. There are several ways to estimate the errors: the easiest way is to use a resampling method, such as the Jackknife approach. As an outcome of a VMC calculation, we can reblock samples of  $B_{\mu i}^\sigma$  and estimate the error bars of the corrections by iteratively solving the Z-vector equations. Another way to estimate the errors is based on the analytical error propagation of  $B_{\mu i}^\sigma$ , for which the variance–covariance tensor of  $B_{\mu i}^\sigma$  is required. In this paper, we employed the former approach, i.e., the error bar estimations were made using the Jackknife method. Further details about the VMC energy and gradients calculations, as well as the derivation of the analytical error propagation, are shown in Appendix A.

### III. VALIDATION OF THE METHOD

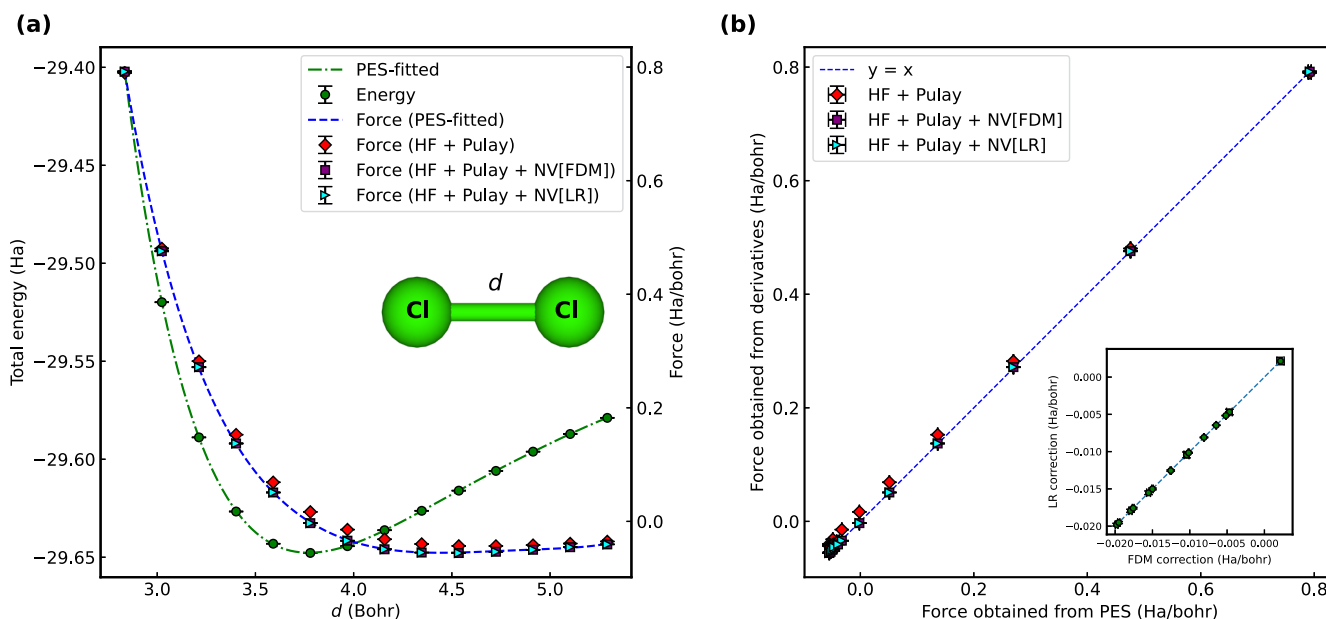
In this section, we validate the Lagrangian technique described above, which we have implemented by interfacing the linear response module of CP2K<sup>49</sup> with TurboRVB<sup>50</sup> and exchanging the data through the TREX-IO library.<sup>52</sup> The validation means demonstrating the following two points: (i) forces and pressures obtained with the Lagrangian technique agree with numerical derivatives of the potential energy surface (PES) (i.e., the results are unbiased) and (ii) these forces and pressures are consistent with those computed by using the finite-difference method (FDM) developed in Ref. 42. The validation was performed for a Cl<sub>2</sub> molecule, crystalline c-BN, and acetamide (CH<sub>3</sub>CONH<sub>2</sub> ≡ AcNH<sub>2</sub>) dimer. For Cl<sub>2</sub>, we evaluated forces; for c-BN, we evaluated both forces and pressure; for AcNH<sub>2</sub> dimer, we evaluated forces along the O–H bonds acting on each

fragment and total forces; for  $\text{Cl}_2$ , we used the ccECP pseudopotentials<sup>53</sup> together with their accompanying cc-pVDZ basis sets; and for c-BN, we used the ccECP<sup>54</sup> pseudopotentials combined with DZVP-MOLOPT-SR-GTH basis sets<sup>55,56</sup> to avoid spuriously large force error bars by Gaussian-type orbitals (GTOs) with large exponents.<sup>21</sup> The cBN calculations employed a conventional  $1 \times 1 \times 1$  supercell with a  $\Gamma$ -point. For the  $\text{AcNH}_2$  dimer, we used the ccECP pseudopotentials<sup>53</sup> together with their accompanying cc-pVDZ basis sets. To estimate the statistical error of the linear-response (LR) correction, we used the jackknife approach:<sup>57</sup> we repeatedly applied the LR procedure to samples of  $\partial E/\partial P_{\mu\nu}^\dagger$  taken from TurboRVB and computed the jackknife mean and variance. Figures 1–3 summarize the results for the  $\text{Cl}_2$  forces, for the c-BN forces and pressure, and for  $\text{AcNH}_2$  dimer forces, addressing points (i) and (ii). As shown, the forces and pressures obtained by the Lagrangian technique are consistent with the numerical derivatives of the potential energy surface and match the values obtained with the FDM.<sup>42</sup> These results confirm that the Lagrangian technique works as expected. The inset of Fig. 3 also demonstrates that the total forces obtained by the Lagrangian technique are always zero, as they should be. Notice that we do not compare the obtained PESs and EOS with the experimental bond lengths and volume since the used basis sets, the supercell size, and the  $k$ -point ( $\Gamma$  only) are suitable only for numerically checking (i) and (ii).

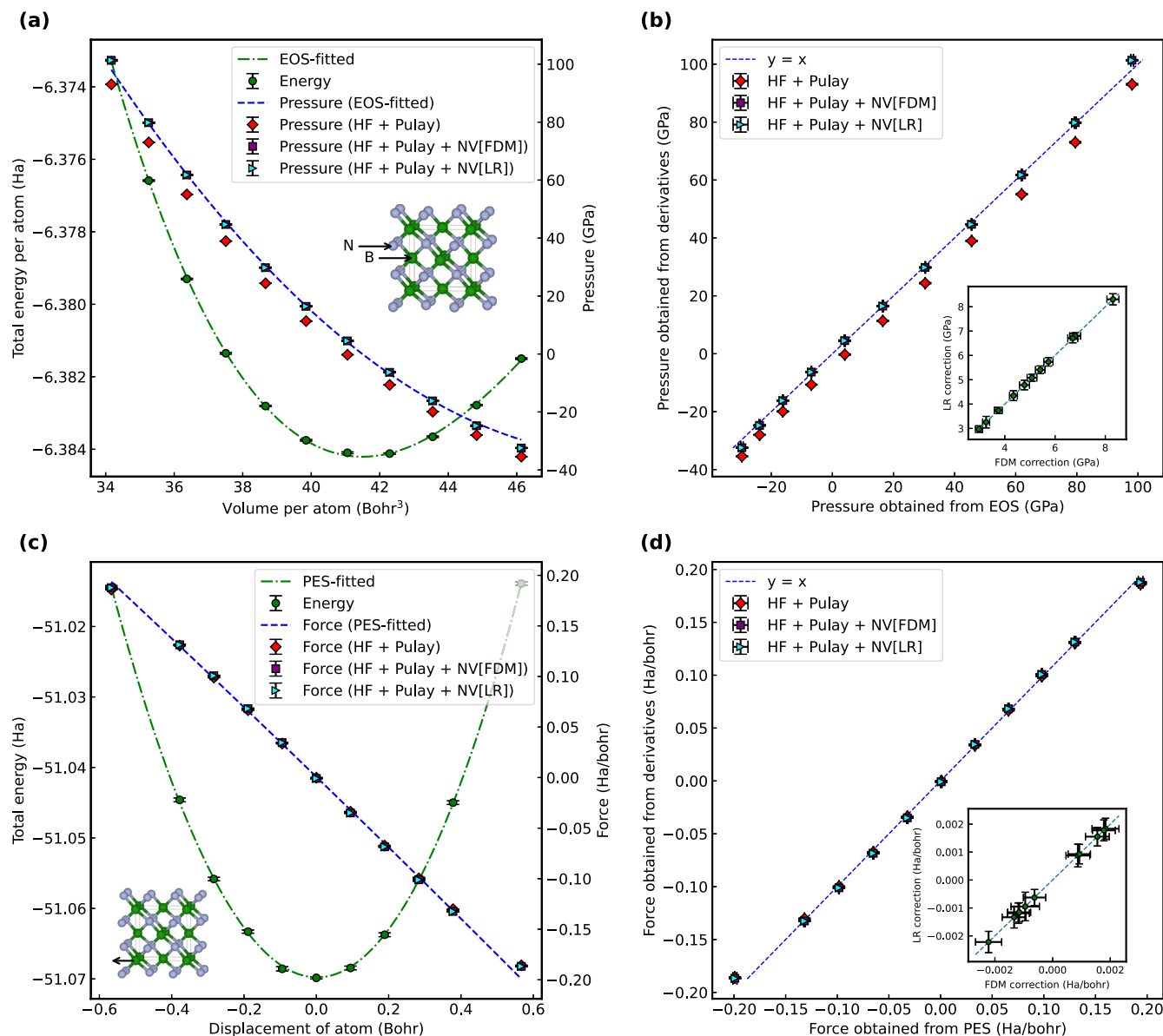
We next discuss the computational cost of the Lagrangian technique. The primary advantage of the Lagrangian technique is efficiency: unlike the FDM approach,<sup>42</sup> which requires  $6N$  separate DFT calculations to obtain the forces, the Lagrangian technique

needs only a single-shot CPHF/CPKS calculation. On the one hand, the FDM approach results in an  $O(N^4)$  asymptotic scaling because it requires  $6N$  times DFT calculations and the computational complexity of DFT is typically  $O(N^3)$ . On the other hand, a CPHF/CPKS calculation shows the same asymptotic scaling as a single-shot HF/KS-SCF calculation does, implying that the asymptotic scaling of the Lagrangian technique remains  $O(N^3)$ . Therefore, the Lagrangian approach is more affordable as the system size increases. We note that the dominant computational cost for atomic force calculations in total arises from the VMC step, not from the LR calculation step. In this work, the number of Markov chain steps for the final energy and force estimates was adjusted so that the error bars on the total energy were 0.1, 0.2, and 0.1 mHa for the  $\text{Cl}_2$ , cBN, and  $\text{AcNH}_2$  dimers, respectively. In such cases, the feasibility of unbiased VMC force calculations is determined by the cost of the corresponding biased VMC force calculations.

We performed a benchmark test for the LR calculation using CP2K for cBN systems with  $N = k \times k \times k \times 8$  atoms in a simulation cell. We measured the timing of DFT and LR calculations on a dual-socket AMD EPYC 7742 64-Core Processor for  $k$  ranging from 1 to 8. The DZVP basis (13 bsf/atom) is employed with the ccECP pseudopotential. Table I summarizes the timings of the calculations, and Fig. 4 plots the scaling analysis. The scaling analysis reveals that, in the region we tested, DFT and LR exhibit  $N^{-3}$  and  $N^{-2}$  behavior, respectively, which are  $N$  order of magnitude smaller than the expected asymptotic scalings. This is probably because the investigated  $N$  values are too small to see the expected asymptotic scaling. However, it demonstrates that the prefactor and scaling of a



**FIG. 1.** Validation of the Lagrangian technique for  $\text{Cl}_2$ . (a) Potential energy surface (PES) as a function of bond length (green circles) and its fit (green dotted line). The reference force on the right-hand Cl atom ( $\text{Cl}_2$  oriented along the  $x$  axis), obtained from the derivative of the fitted PES, is indicated by the blue dashed line. The red diamonds, purple squares, and cyan triangles denote forces evaluated with the HF + Pulay terms, HF + Pulay terms + NV correction (via FDM), and HF + Pulay terms + NV correction (via LR), respectively. (b) Comparison between the PES-derived (reference) forces and those obtained from the estimators. Inset: direct comparison of the NV corrections via FDM and LR.

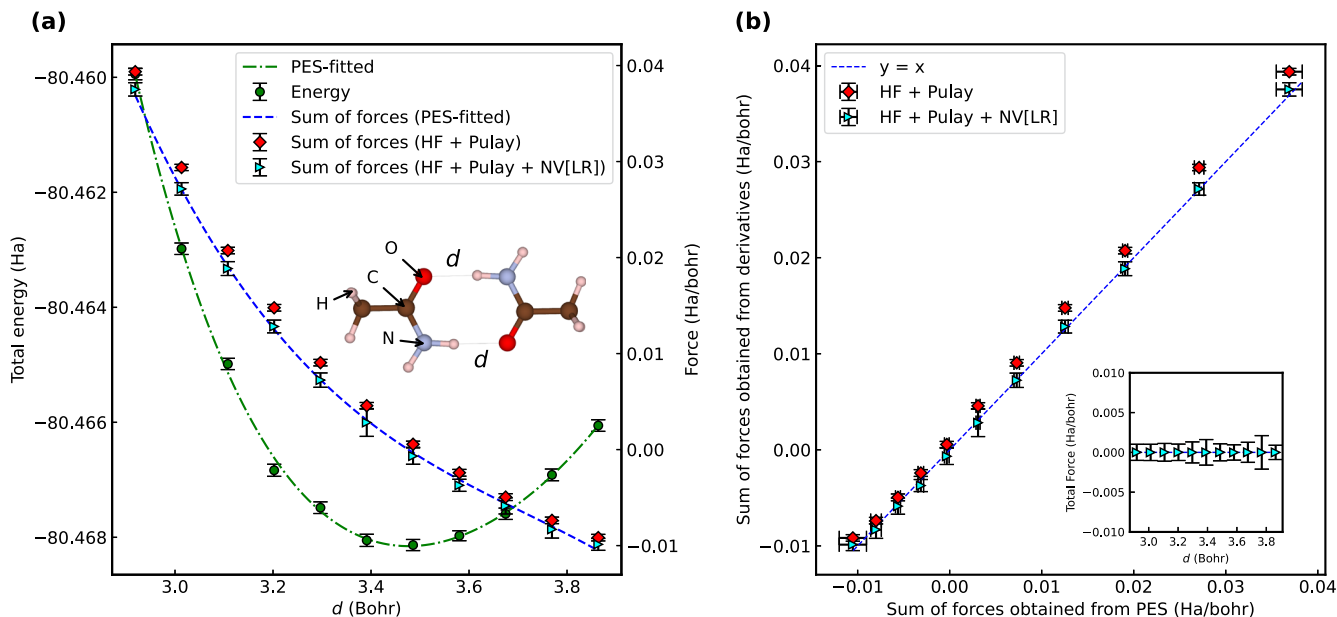


**FIG. 2.** Validation of the Lagrangian technique for cBN. (a) Equation of state (EOS) as a function of volume (green circles) and its fit (green dotted line). The reference pressure, obtained from the derivative of the fitted EOS, is indicated by the blue dashed line. The red diamonds, purple squares, and cyan triangles denote pressures evaluated with the HF + Pulay terms, HF + Pulay terms + NV correction (via FDM), and HF + Pulay terms + NV correction (via LR), respectively. (b) Comparison between the EOS-derived (reference) pressures and those obtained from the estimators. Inset: direct comparison of the NV corrections via FDM and LR. (c) PES as a function of a displacement of the B atom located at the origin (0, 0, 0) of the unit cell in the x direction (green circles) and its fit (green dotted line). The reference force on the B atom, obtained from the derivative of the fitted PES, is indicated by the blue dashed line. The red diamonds, purple squares, and cyan triangles denote forces evaluated with the HF + Pulay terms, HF + Pulay terms + NV correction (via FDM), and HF + Pulay terms + NV correction (via LR), respectively. (d) Comparison between the PES-derived (reference) forces and those obtained from the estimators. Inset: direct comparison of the NV corrections via FDM and LR.

single-shot LR are better than those of  $6N$ -times DFT calculations. The systematic test shows that our method is faster than the FDM method relying on the  $6N$ -times DFT calculations.

From a technical point of view, the exchange of wavefunctions between CP2K<sup>49</sup> and TurboRVB<sup>50</sup> required careful consideration

to deal with differences in conventions, which is a common challenge among electronic structure packages. Examples are the use of spherical-harmonic vs Cartesian Gaussian type orbitals (GTOs), normalization constants, and their application stage (pre- vs post-contraction and shell-wise vs AO-level), and related notational



**FIG. 3.** Validation of the Lagrangian technique for AcNH<sub>2</sub> dimer. (a) Potential energy surface (PES) as a function of O–H bond length (green circles) denoted as  $d$  and its fit (green dotted line). The reference force (i.e., the sum of forces acting on the right fragment along the two O–H bonds, where the two O–H bonds are aligned along the  $x$  axis), obtained from the derivative of the fitted PES, is indicated by the blue dashed line. The red diamonds and cyan triangles denote the sum of forces in the  $x$  direction evaluated with the HF + Pulay terms and HF + Pulay terms + NV correction (via LR), respectively. (b) Comparison between the PES-derived (reference) forces and those obtained from the estimator, HF + Pulay terms + NV correction (via LR). Inset: The total forces of the system along the O–H bonds (i.e., in the  $x$  direction), which should be zero, as a function of O–H bond length.

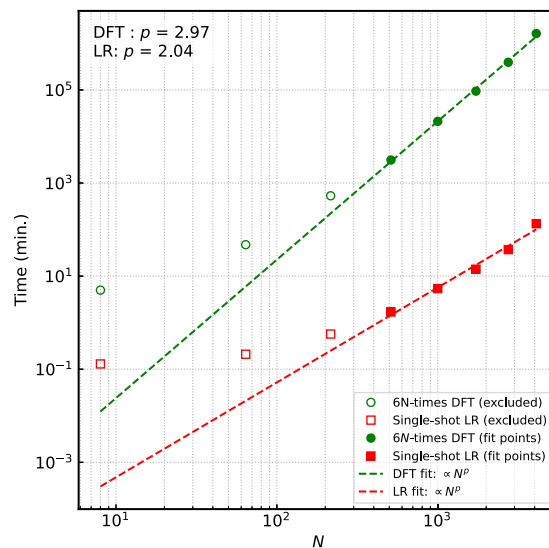
**TABLE I.** Comparison of timings of 6N-times DFT calculations and single-shot LR calculation for cBN systems with  $N = k \times k \times k \times 8$  atoms. The DZVP basis (13 bsf/atom) is employed with a ccECP pseudo-potential. The timings were measured in minutes on a dual-socket AMD EPYC 7742 64-Core Processor.

$k$	$N$	6N-times DFT	Single-shot LR
1	8	4.99	0.13
2	64	47.60	0.21
3	216	441.33	0.57
4	512	3 130.68	1.70
5	1000	21 095.40	5.39
6	1728	94 109.82	14.11
7	2744	392 623.87	37.12
8	4096	1 628 106.34	133.98

choices. In this work, we adopted the TREX-IO<sup>52</sup> format, a flexible library that standardizes conventions for exchanging wavefunctions, structures, and pseudopotentials among electronic structure packages. TREX-IO effectively absorbed inter-code convention mismatches and played a key role in robust data exchange between CP2K and TurboRVB in this work.

#### IV. APPLICATION TO ORGANIC MOLECULES IN RMD17 DATASET

Now that we have established an unbiased VMC force evaluation based on the Lagrangian technique, a systematic assessment of



**FIG. 4.** Scaling analysis of DFT and LR calculations with respect to the number of atoms ( $N$ ) in the cBN simulation cell.

VMC force accuracy is feasible across a wide range of compounds. VMC and DMC results are increasingly being used as a training dataset for MLPs.<sup>23,35,37–41,43</sup> In this study, we selected ethanol, malonaldehyde, and benzene from the rMD17 dataset<sup>46,47</sup> and

computed (i) unbiased VMC forces established in this work; (ii) VMC forces with fully optimized wavefunctions; (iii) Hartree–Fock (HF), second-order Møller–Plesset perturbation theory (MP2), CCSD, and DFT forces; and (iv) CCSD(T) forces as references. We took 100 structures from molecular dynamics trajectories distributed as an rMD17 dataset.<sup>46,47</sup> In a previous study that constructed MLPs with CCSD(T) Ref. 58, the Dunning cc-pVTZ<sup>59</sup> was used for ethanol as well as cc-pVDZ<sup>59</sup> for malonaldehyde and benzene to generate a training dataset for MLPs. Here, we used cc-pVQZ for the three molecules for the CCSD(T) calculations to obtain forces closer to the complete basis-set limit. For comparison, we also carried out HF, MP2,<sup>60</sup> and CCSD<sup>61</sup> with the cc-pVQZ basis and DFT calculations with the def2-QZVPPD<sup>62</sup> basis set using several exchange–correlation functionals with the D3BJ dispersion correction.<sup>63,64</sup> In particular, we chose SVWN,<sup>65</sup> Perdew–Burke–Ernzerhof (PBE),<sup>66</sup> PBE + D3BJ, PBE0<sup>67</sup>+D3BJ,  $\omega$ B97X + D3BJ,<sup>68,69</sup> and  $\omega$ B97M + D3BJ.<sup>69,70</sup> In the case of MP2, CCSD, and CCSD(T) calculations, the 1s shells were kept frozen. All of these calculations were performed with Psi4 software suite version 1.9.1.<sup>71</sup> We notice that, for each of ethanol, malonaldehyde, and benzene, the first structure was also computed with Molpro version 2025.3<sup>72–74</sup> and PySCF version 2.10.0<sup>75,76</sup> to double-check the Psi4 results. We confirmed that Psi4, PySCF, and Molpro yielded consistent energies and forces for all three molecules. SVWN and PBE were selected as computationally inexpensive exchange–correlation functionals. PBE0 was included to allow comparison with a previous VMC/DMC benchmark study,<sup>23</sup> which combined Tkatchenko–Scheffler (TS)<sup>77</sup> and the many-body dispersion (MBD)<sup>78</sup> corrections with PBE and PBE0, respectively. Since the Psi4 software suite does not implement TS or MBD, we instead employed D3BJ dispersion correction in this work. The  $\omega$ B97 family of functionals was chosen because it is often used in popular MLP datasets (e.g., SPICE<sup>79</sup>). For the VMC calculations, we employed ccECP pseudopotentials<sup>54</sup> together with the accompanied cc-pVTZ basis set. Following the previous VMC/DMC benchmark study by Sloatman *et al.*,<sup>23</sup> we also benchmarked ccECP-based CCSD against all-electron CCSD using PySCF<sup>75,76</sup> and confirmed that the force errors for the three molecules are <1 kcal/mol/Å in root-mean-square errors (RMSEs). In the VMC calculations for the unbiased forces, the Slater determinant part was fixed to orbitals obtained from DFT (LDA-PZ) and only the Jastrow part was variationally optimized using stochastic reconfiguration<sup>80</sup> with an optimal step size. In the VMC calculations with fully optimized wavefunction, both the molecular orbital coefficients in the Slater determinant and the Jastrow part were variationally optimized. The Jastrow factor comprised a two-body term satisfying the electron–electron cusp<sup>50</sup> and an inhomogeneous one-body term and a three-body term parameterized using localized atomic orbitals.<sup>50</sup> For the inhomogeneous one-body and three-body Jastrow basis sets, we used uncontracted primitive GTOs with [3s1p] for H and [4s2p1d] for C, N, and O; the corresponding Gaussian exponents were optimized for each configuration during the optimization procedure. Although matrix elements coupling different atomic pairs (often referred to as four-body Jastrow terms<sup>50</sup>) are frequently set to zero in practice,<sup>50</sup> we retained them here because they have an impact on the forces (see Appendix A 6). The total numbers of optimized variational parameters were 3422, 5069, and 8140 for ethanol, malonaldehyde, and benzene, respectively. The trial wavefunctions were generated

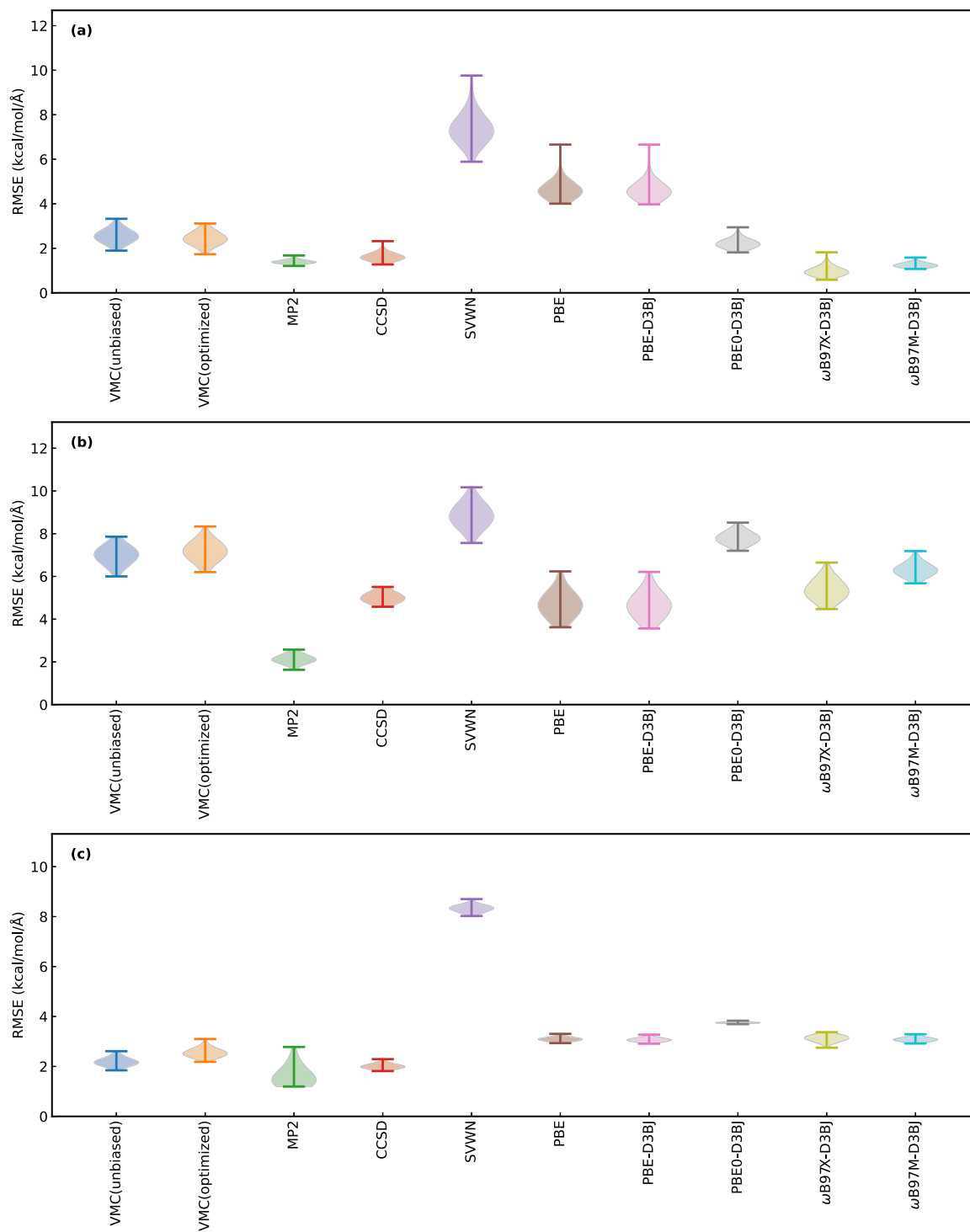
with CP2K,<sup>49</sup> VMC simulations were performed with TurboRVB,<sup>50</sup> and the force-bias correction based on the Lagrangian technique was applied via CP2K.<sup>49</sup>

Table II reports the RMSE of all force components (i.e.,  $N_{\text{atoms}} \times 3$  directions  $\times$  100 structures) relative to CCSD(T) for each molecule. Figure 5 shows, for each molecule, the violin plot of the distribution of RMSE in 100 configurations for selected methods. We first emphasize that the benefit of the Lagrangian technique is evident from Table II. Comparing the CCSD(T) reference with biased VMC forces (i.e., without the Lagrangian-technique correction), the RMSEs are 2.924(9), 9.56(1), and 3.19(1) kcal mol<sup>-1</sup> Å<sup>-1</sup> for ethanol, malonaldehyde, and benzene, respectively. In contrast, the unbiased VMC forces developed in this study yield RMSEs of 2.549(9), 7.00(1), and 2.18(1) kcal mol<sup>-1</sup> Å<sup>-1</sup> for ethanol, malonaldehyde, and benzene, respectively. They suggest that when accurate forces are sought with only partially optimized wavefunctions, our correction scheme is essential. The VMC forces obtained with the fully optimized wavefunction yield RMSEs of 2.455(9), 7.19(1), and 2.54(1) kcal mol<sup>-1</sup> Å<sup>-1</sup> for ethanol, malonaldehyde, and benzene, respectively. They are very close to the unbiased VMC ones, demonstrating that, *even without optimizing the determinant part*, correcting the NV term can achieve force accuracies comparable with those obtained with fully optimized wavefunctions, thereby reducing optimization effort.

For ethanol, Sloatman *et al.*<sup>23</sup> performed an extensive benchmark for a set of 200 configurations and reported an RMSE of 3.055 kcal mol<sup>-1</sup> Å<sup>-1</sup> for VMC forces based on a fully optimized Jastrow–Slater wavefunction. Our ethanol RMSEs, 2.549(9) kcal mol<sup>-1</sup> Å<sup>-1</sup> with the frozen DFT orbitals and 2.455(9) kcal mol<sup>-1</sup> Å<sup>-1</sup> with the fully optimized Jastrow–Slater wavefunction, are better than their values. This improvement is probably due to the much larger Jastrow factor used in this study. We note that our ethanol configurations are not identical to theirs; thus, a strict one-to-one comparison is not possible. If one pursues the force accuracy of  $\sim 1$  kcal mol<sup>-1</sup> Å<sup>-1</sup> relative to CCSD(T), DMC calculations would be required, as revealed by Sloatman *et al.*<sup>23</sup> While the present Lagrangian technique is applicable to DMC force calculations, in

TABLE II. RMSE of atomic forces of all the configurations for the selected methods relative to CCSD(T). The unit is kcal/mol/Å.

Method	Ethanol	Malonaldehyde	Benzene
VMC(biased)	2.924(9)	9.56(1)	3.19(1)
VMC(unbiased)	2.549(9)	7.00(1)	2.18(1)
VMC(optimized)	2.455(9)	7.19(1)	2.54(1)
HF	9.45	23.44	5.57
MP2	1.39	2.13	1.64
CCSD	1.62	5.01	2.00
SVWN	7.37	8.86	8.32
PBE	4.63	4.72	3.09
PBE-D3BJ	4.60	4.68	3.06
PBE0-D3BJ	2.21	7.80	3.75
$\omega$ B97X-D3BJ	0.98	5.40	3.13
$\omega$ B97M-D3BJ	1.24	6.31	3.08



**FIG. 5.** Violin plots showing, for each configuration of (a) ethanol, (b) malonaldehyde, and (c) benzene, the RMSE of atomic forces for the selected methods relative to CCSD(T). The unit is kcal/mol/Å.

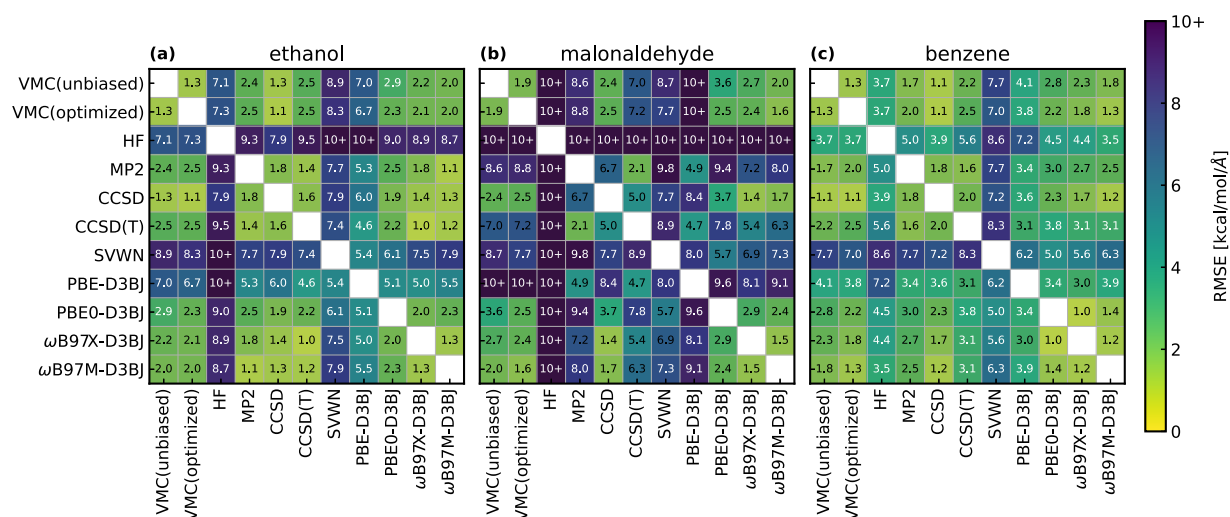
principle, a full DMC implementation is beyond the scope of this paper and is left for future work.

Table II also lists the HF, MP2, CCSD, and other DFT results for all three molecules. For ethanol,  $\omega$ B97X-D3BJ and  $\omega$ B97M-D3BJ yield  $\sim 1$  kcal mol $^{-1}$  Å $^{-1}$  RMSEs relative to CCSD(T). Slootman *et al.*<sup>23</sup> compared VMC with dispersion-corrected PBE-TS and PBE0-MB, finding  $\text{RMSE}(\text{PBE0-MB}) < \text{RMSE}(\text{VMC}) < \text{RMSE}(\text{PBE-TS})$ , that is, VMC performed between a generalized gradient approximation (GGA) and a hybrid GGA. Our results reproduce this trend. Interestingly, this trend is not universal across the molecules. For benzene, the VMC-unbiased and VMC-optimized RMSEs are 2.18(1) and 2.54(1) kcal mol $^{-1}$  Å $^{-1}$ , respectively, whereas  $\omega$ B97X-D3BJ and  $\omega$ B97M-D3BJ are worse than the VMC ones, exceeding 3 kcal mol $^{-1}$  Å $^{-1}$ . For malonaldehyde, both the VMC-unbiased and VMC-optimized RMSEs yield large RMSEs, 7.00(1) and 7.19(1) kcal mol $^{-1}$  Å $^{-1}$ , respectively. Not only the VMCs but also PBE0-D3BJ and  $\omega$ B97 functional family give large RMSEs in order of  $\sim 5$ – $8$  kcal mol $^{-1}$  Å $^{-1}$ . For all three compounds, HF or SVWN shows the worst RMSE among the tested methods, while MP2 shows a very good performance. These results demonstrate that our unbiased VMC forces appear advantageous over biased VMC ones in terms of accuracy. However, VMC is outperformed by MP2 for all three compounds. Although VMC offers a more favorable scaling with the number of electrons [typically,  $O(N^3)$ <sup>81</sup> and  $O(N^{4-5})$  for VMC and MP2, respectively] and is naturally suitable for parallelization, the large prefactor in practical VMC implementations complicates any conclusion about superior wall-time efficiency (e.g., VMC with  $\sim 0.3$  kcal/mol/Å error bars in the force components is roughly a factor of 100 heavier than MP2 for ethanol).

To further analyze the discrepancy between VMC and CCSD(T) for malonaldehyde, we computed the pairwise RMSE between the methods we have tested in this study and plotted these

values with the color bar shown in Fig. 6. This figure reveals that the consistency between methods differs significantly for ethanol, malonaldehyde, and benzene. For ethanol and benzene, most methods generally show good agreement, except for HF and SVWN. However, for malonaldehyde, many method pairs yield RMSE exceeding 10, indicating significant discrepancy among them. Furthermore, the consistency between CCSD and CCSD(T) also varies considerably among ethanol, malonaldehyde, and benzene. While the RMSE between CCSD and CCSD(T) is small for ethanol and benzene, it becomes very large for malonaldehyde. Interestingly, (unbiased and optimized) VMC, PBE0-D3BJ,  $\omega$ B97X-D3BJ, and  $\omega$ B97M-D3BJ show good agreement in RMSEs for all three compounds. These results suggest that, for malonaldehyde, VMC does not provide physically incorrect forces; rather, the parenthesis T correction in the CCSD(T) calculation may have a peculiar influence on the atomic forces.

Although the reason for the significant deviation in malonaldehyde is not yet fully understood, one possibility is that malonaldehyde may have a significant multi-reference character. For CCSD and CCSD(T) calculations, the so-called D1 and T1 diagnostics<sup>82</sup> have been proposed as measures to estimate the degree of the multi-reference character. Empirically, it has been advised to pay attention to the reliability of CCSD and CCSD(T) calculations when D1 exceeds 0.05 or T1 exceeds 0.02.<sup>82</sup> The D1 and T1 diagnostics taken from the computed 100 structures for ethanol, malonaldehyde, and benzene are D1 values of 0.023 (Max:0.032 and Min:0.020), 0.051 (Max:0.063 and Min:0.043), and 0.030 (Max:0.031 and Min:0.029), and T1 values of 0.0097 (Max:0.0115 and Min:0.0093), 0.0162 (Max:0.0179 and Min:0.0147), and 0.0106 (Max:0.0109 and Min:0.0103). The D1 value for malonaldehyde shows an average exceeding the threshold (0.05), raising suspicion of a significant multi-reference character. To further investigate this discrepancy, quantum chemistry calculations would require approaches such as



**FIG. 6.** Pairwise RMSEs of atomic forces (kcal/mol/Å) between all pairs of the benchmarked methods for the three molecules: (a) ethanol, (b) malonaldehyde, and (c) benzene. The color scale is shared among the plots and is saturated at 10 kcal/mol/Å (i.e., values  $\geq 10$  are mapped to the same color). The numbers indicate the corresponding RMSEs for each pair.

complete active space self-consistent field (CASSCF), while QMC calculations would require VMC or DMC calculations using a multi-determinant *Ansatz*, as reported by Stoolman *et al.*<sup>23</sup> A broader benchmark, e.g., covering more chemically varied systems and including DMC benchmark, is also a promising future work.

## V. CONCLUDING REMARKS AND FUTURE PERSPECTIVES

In this paper, we introduce an application of the Lagrangian technique in variational Monte Carlo (VMC) calculations to get atomic forces and pressures consistent with potential energy surfaces (also known as unbiased forces and pressures). The previous study to get unbiased forces and pressures with the Jastrow–Slater determinant *Ansatz* requires  $6N$  times additional density functional theory (DFT) calculations ( $N$  refers to the number of nuclei in a system),<sup>42</sup> which hinders the applications of unbiased VMC force/pressure calculations in large molecules/crystal or in generations of training data for machine-learning potentials (MLPs). The formulation that we report in this paper replaces the additional  $6N$ -times DFT calculations with a single coupled-perturbed Hartree–Fock or coupled-perturbed Kohn–Sham calculation. This improves the feasibility of the VMC force calculations with the conventional Jastrow–Slater determinant *Ansatz*. We focus on the Jastrow–Slater determinant *Ansatz* with GTO basis sets in this study as a simple and practical case. However, the Lagrangian technique is versatile for any other functional forms, such as Jastrow multi-determinant *Ansatz*, and any other basis sets, such as plane-waves, with non-variational parameters. The difference from this study is to derive the corresponding Lagrangian, depending on the employed *Ansatz* and basis set. As an application of our developed method, we demonstrate that the unbiased force evaluation improves not only the consistency but also the accuracy of the VMC forces. We selected three molecules from the rMD17 benchmark set,<sup>46,47</sup> ethanol, malonaldehyde, and benzene, and compared biased VMC forces, unbiased VMC forces, and forces obtained by coupled-cluster singles and doubles with perturbative triples [CCSD(T)]<sup>48</sup> calculations. We found that the biased VMC forces are far from the CCSD(T) ones, while the unbiased ones give values close to those of the CCSD(T) ones. The unbiased force evaluation described here, for instance, plays a crucial role in generating high-quality training data for MLPs.

## ACKNOWLEDGMENTS

K.N. acknowledges computational resources from the Numerical Materials Simulator at National Institute for Materials Science (NIMS), from the Research Institute for Information Technology at Kyushu University under the category of General Projects on the supercomputer GENKAI, and from the supercomputer Fugaku provided by RIKEN through the HPCI System Research Projects (Project No. hp250031). K.N. acknowledges financial support from the MEXT Leading Initiative for Excellent Young Researchers (Grant No. JPMXS0320220025), the Murata Science and Education Foundation (Grant No. M24AN006), the Japan Science and Technology Agency (JST), PRESTO (Grant No. JPMJPR24J9), and the JSPS KAKENHI [Grant-in-Aid for Scientific Research (C), Grant No. 25K00216]. This research was also supported by the NCCR MARVEL, a National Center of Competence in Research, funded

by the Swiss National Science Foundation (Grant No. 205602). K.N. acknowledges Michele Casula (CNRS) for fruitful discussion about VMC force calculations. K.N. acknowledges the valuable comments on the method that the author developed before from Dr. Emmanuel Giner (CNRS) and Dr. David Ceperley (UIUC) in a conference. We thank Dr. Guglielmo Mazzola (UZH) for organizing a seminar delivered by one of the authors K.N. at UZH. The seminar and subsequent discussions led to this collaboration. The molecular and crystal structures were depicted using VESTA.<sup>83</sup> We acknowledge Grammarly in its use to help us check English grammar.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Kousuke Nakano:** Conceptualization (equal); Data curation (lead); Funding acquisition (equal); Investigation (lead); Methodology (equal); Project administration (equal); Resources (lead); Software (equal); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review & editing (equal). **Stefano Battaglia:** Conceptualization (equal); Data curation (supporting); Funding acquisition (equal); Investigation (supporting); Methodology (lead); Resources (equal); Software (equal); Validation (supporting); Visualization (supporting); Writing – original draft (equal); Writing – review & editing (equal). **Jürg Hutter:** Conceptualization (lead); Funding acquisition (equal); Methodology (equal); Project administration (equal); Resources (equal); Software (equal); Supervision (lead); Writing – original draft (supporting); Writing – review & editing (equal).

## DATA AVAILABILITY

The *ab initio* DFT and QMC packages used in this work, CP2K and TurboRVB, are available from their GitHub repositories <https://github.com/cp2k/cp2k> and <https://github.com/sissaschool/turborvb>, respectively. The data to reproduce the rMD17 benchmark test are available from NIMS Materials Data Repository (MDR) (<https://doi.org/10.48505/nims.6164>).

## APPENDIX A: DETAILS OF THE VARIATIONAL MONTE CARLO CALCULATIONS

### 1. Variational Monte Carlo methods

The expectation value of the energy for a given wavefunction,  $\Psi$ , is evaluated as

$$\langle E \rangle = \frac{\int d\mathbf{x} \Psi^2(\mathbf{x}) \cdot \hat{H}\Psi(\mathbf{x})/\Psi(\mathbf{x})}{\int d\mathbf{x} \Psi^2(\mathbf{x})} = \int d\mathbf{x} E_L(\mathbf{x}) \pi(\mathbf{x}), \quad (\text{A1})$$

where  $\mathbf{x} = (\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)$  refers to the  $N$  electron coordinates and their spins and

$$E_L(\mathbf{x}) \equiv \frac{\hat{H}\Psi(\mathbf{x})}{\Psi(\mathbf{x})} \quad \text{and} \quad \pi(\mathbf{x}) \equiv \frac{\Psi^2(\mathbf{x})}{\int d\mathbf{x}' \Psi^2(\mathbf{x}')}$$

are the local energy and the probability distribution, respectively. The  $3N$ -dimensional integration can be stochastically evaluated by generating a set  $\{\mathbf{x}_i\}$  according to the probability distribution  $\pi(\mathbf{x})$  generated using the Markov chain Monte Carlo (MCMC) technique and by averaging the obtained local energies  $E_L(\mathbf{x}_i)$ ,

$$E_{\text{MCMC}} = \langle E_L(\mathbf{x}) \rangle_{\pi(\mathbf{x})} \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{x}_i). \quad (\text{A2})$$

$E_{\text{MCMC}}$  is always associated with an error since it is evaluated by a stochastic sampling. We note that the error bar is inversely proportional to the square root of the number of samplings.

Suppose that if  $E_0$  is the ground state energy, the *variational theorem* guarantees that the expectation value of  $E$  never becomes smaller than  $E_0$ . In other words, to approach the ground state, one can optimize a given WF by introducing a set of parameters  $\mathbf{p} \equiv (\dots, p_i, \dots)$  to the WF  $\Psi(\mathbf{x}, \mathbf{p})$ ,

$$E_{\text{VMC}}(\mathbf{p}) = \int d\mathbf{x} E_L(\mathbf{x}, \mathbf{p}) \pi(\mathbf{x}, \mathbf{p}) \geq E_0. \quad (\text{A3})$$

This framework is called variational Monte Carlo (VMC). VMC is conceptually very simple, but the optimization of a many-body WF is a complex numerical task due to several reasons, such as the presence of several local minima in the energy landscape and statistical errors associated with MCMC evaluation. Great improvements in this field have been achieved in the QMC community. For instance, in TurboRVB, the automatic differentiation<sup>15</sup> allows the stable and efficient calculation of the energy derivatives with variational parameters and offers an efficient optimization method, the so-called stochastic reconfiguration method (also known as natural gradient optimization).<sup>80,84</sup>

## 2. Wavefunction Ansatz

TurboRVB employs the Jastrow antisymmetrized geminal power (JAGP)<sup>85</sup> Ansatz. The Ansatz is composed of a Jastrow and an antisymmetric part ( $\Psi = J \cdot \Psi_{\text{AGP}}$ ). The antisymmetric part reads

$$\Psi_{\text{AGP}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \det \left[ g(\mathbf{r}_1^\uparrow, \mathbf{r}_1^\downarrow) g(\mathbf{r}_2^\uparrow, \mathbf{r}_2^\downarrow) \cdots g(\mathbf{r}_{N/2}^\uparrow, \mathbf{r}_{N/2}^\downarrow) \right], \quad (\text{A4})$$

where  $g(\mathbf{r}^\uparrow, \mathbf{r}^\downarrow)$  is called the pairing function. Although, for the sake of clarity, an unpolarized system with an even number  $N$  of electrons is considered here, it can be generalized to describe a polarized system with unpaired electrons.<sup>50</sup> The pairing function,  $g(\mathbf{r}_i^\uparrow, \mathbf{r}_j^\downarrow)$ , is expanded over Gaussian-type atomic orbitals (GTOs),

$$g(\mathbf{r}_i^\uparrow, \mathbf{r}_j^\downarrow) = \sum_{\mu\nu} P_{\mu\nu}^{\uparrow\downarrow} \psi_\mu(\mathbf{r}_i^\uparrow) \psi_\nu(\mathbf{r}_j^\downarrow), \quad (\text{A5})$$

where  $\psi(\mathbf{r})$  are primitive or contracted GTOs. Notice that in Eq. (A5), the expansion coefficients of the pairing functions are labeled with  $P_{\mu\nu}^{\sigma\tau}$  to highlight their similarity with the density matrix of Eq. (10). Depending on the symmetry of  $P_{\mu\nu}^{\uparrow\downarrow}$ , the pairing function is dubbed as AGP with spin-singlet pairs (AGPs), AGP with spin-singlet pairs for triplet correlations (AGPu), and the combination of AGPs and AGPu,<sup>50</sup> which correspond to a symmetric, skew-symmetric, and general  $P_{\mu\nu}^{\uparrow\downarrow}$ , respectively.<sup>86</sup> The basis sets employed in this study are described in the main text. When the JAGP is

expanded over  $p = N_{\text{el}}/2$  molecular orbitals (i.e., only the occupied molecular orbitals are considered), the JAGP coincides with the Jastrow–Slater determinant (JSD) Ansatz.<sup>50</sup> In this way, we composed the JSD ansatz, where the SD part was taken from a DFT calculation. Indeed, the coefficients in the pairing functions (i.e., the variational parameters in the SD part) are obtained by CP2K and fixed during the VMC optimization step. Only the coefficients in the Jastrow factor are optimized at the VMC level.

The Jastrow term in this study is composed of inhomogeneous one-body, two-body, and three-body factors ( $J = J_1 J_2 J_3$ ).<sup>50</sup> We employed a spin-independent Jastrow parameterization in this study. The two-body factor is used to satisfy the electron–electron cusp conditions, and the inhomogeneous one-body and three-/four-body factors are adopted to take the further electron–electron correlation into consideration. The one-body Jastrow factor added to satisfy the electron–nucleus cusp condition is not used in this study because effective core potentials remove the Coulomb singularity. The inhomogeneous one-body Jastrow part is parameterized as

$$J_1(\mathbf{r}_1 \sigma_1, \dots, \mathbf{r}_N \sigma_N) = \sum_i \sum_\mu \xi_\mu \psi_\mu^{\text{Jas}}(\mathbf{r}_i), \quad (\text{A6})$$

where  $\mathbf{r}_i$  represents the electron positions,  $\xi_\mu$  represents the variational parameters, and  $\mu$  is the GTO index. The two-body Jastrow factor is defined by a function as

$$J_2(\mathbf{r}_1, \dots, \mathbf{r}_N) = \exp \left( \sum_{i < j} v(|\mathbf{r}_i - \mathbf{r}_j|) \right). \quad (\text{A7})$$

$v(r)$  is a Pade-type function,

$$v(r) = \frac{1}{2} r \cdot (1 + F \cdot r)^{-1}, \quad (\text{A8})$$

or the exponential-type function,

$$v(r) = \frac{1}{2F} (1 - e^{-F \cdot r}), \quad (\text{A9})$$

where  $F$  is a variational parameter. The three-body Jastrow factors are

$$J_3(\mathbf{r}_1, \dots, \mathbf{r}_N) = \exp \left( \sum_{i < j} \Xi_{\text{Jas}}(\mathbf{r}_i, \mathbf{r}_j) \right) \quad (\text{A10})$$

and

$$\Xi_{\text{Jas}}(\mathbf{r}_i, \mathbf{r}_j) = \sum_{\mu, \nu} \xi_{\mu, \nu} \psi_\mu^{\text{Jas}}(\mathbf{r}_i) \psi_\nu^{\text{Jas}}(\mathbf{r}_j), \quad (\text{A11})$$

where  $\xi_{\mu, \nu}$  are the variational parameters and the indices  $\mu$  and  $\nu$  refer to different orbitals. The basis sets employed in this study are described in the main text. The Jastrow factors are optimized using the stochastic reconfiguration method.

## 3. Computing derivatives

The elements of the right-hand side tensor  $\mathbf{B}$  of Eq. (13) can be efficiently computed from a VMC sampling,

$$B_{\mu i}^\sigma = -2 \langle E_L(\mathbf{x}) [ (O_{\mu i}^\sigma(\mathbf{x}) - \bar{O}_{\mu i}^\sigma) ] \rangle, \quad (\text{A12})$$

where we made apparent the dependence of the local operators on the total electronic coordinate  $\mathbf{x}$ , sampled by VMC, to distinguish them from constant values. In Eq. (A12),  $O_{\mu i}^{\sigma}(\mathbf{x}) = \partial \ln \Psi(\mathbf{x}) / \partial C_{\mu i}^{\sigma}$  and  $\bar{O}_{\mu i}^{\sigma} \sim \langle O_{\mu i}^{\sigma}(\mathbf{x}) \rangle$ . We note that  $O_{\mu i}^{\sigma}(\mathbf{x})$  can be efficiently computed in a VMC calculation using the adjoint algorithmic differentiation,<sup>15</sup> and the divergences of the generalized forces can be cured by reweighting methods.<sup>16,17</sup> It is extremely important that the variational force is evaluated in a covariance form of random variables to reduce its fluctuations.<sup>87,88</sup> In practice, TurboRVB always computes the derivatives of  $E$  with respect to the AGP matrix elements  $P_{\mu\nu}^{\uparrow\downarrow}$  in Eq. (A5),  $\partial E / \partial P_{\mu\nu}^{\uparrow\downarrow}$ . In other words, this is the AO representation of the pairing function. Thus, the chain rule should be applied to obtain  $\partial E / \partial C_{\mu i}^{\sigma}$  from  $\partial E / \partial P_{\mu\nu}^{\uparrow\downarrow}$ . Molecular spin orbitals are defined as

$$\phi_k^{\sigma}(\mathbf{r}) = \sum_{\mu} C_{\mu k}^{\sigma} \psi_{\mu}(\mathbf{r}). \quad (\text{A13})$$

Enforcing JSD  $\leftrightarrow$  JAGP equivalence (idempotent geminal) via the occupied-orbital expansion,

$$g(\mathbf{r}_i^{\uparrow}, \mathbf{r}_j^{\downarrow}) = \sum_{k \in \text{occ}} \phi_k^{\uparrow}(\mathbf{r}_i^{\uparrow}) \phi_k^{\downarrow}(\mathbf{r}_j^{\downarrow}), \quad (\text{A14})$$

implies the geminal coefficients,

$$P_{\mu\nu}^{\uparrow\downarrow} = \sum_{k \in \text{occ}} C_{\mu k}^{\uparrow} C_{\nu k}^{\downarrow} \quad (\text{A15})$$

and

$$P_{\mu\nu}^{\downarrow\uparrow} = \sum_{k \in \text{occ}} C_{\mu k}^{\downarrow} C_{\nu k}^{\uparrow}. \quad (\text{A16})$$

We define  $\tilde{B}_{\mu\nu}^{\uparrow\downarrow} := \partial E / \partial P_{\mu\nu}^{\uparrow\downarrow}$  and  $\tilde{B}_{\mu\nu}^{\downarrow\uparrow} := \partial E / \partial P_{\mu\nu}^{\downarrow\uparrow}$ . By the chain rule,

$$\begin{aligned} B_{\mu i}^{\sigma} &:= \frac{\partial E}{\partial C_{\mu i}^{\sigma}} = \sum_{\lambda\nu} \frac{\partial E}{\partial P_{\lambda\nu}^{\uparrow\downarrow}} \frac{\partial P_{\lambda\nu}^{\uparrow\downarrow}}{\partial C_{\mu i}^{\sigma}} + \sum_{\lambda\nu} \frac{\partial E}{\partial P_{\lambda\nu}^{\downarrow\uparrow}} \frac{\partial P_{\lambda\nu}^{\downarrow\uparrow}}{\partial C_{\mu i}^{\sigma}} \\ &= \begin{cases} \sum_{\nu} \tilde{B}_{\mu\nu}^{\uparrow\downarrow} C_{\nu i}^{\downarrow} + \sum_{\lambda} \tilde{B}_{\lambda\mu}^{\uparrow\downarrow} C_{\lambda i}^{\downarrow} & (\sigma = \uparrow) \\ \sum_{\lambda} \tilde{B}_{\lambda\mu}^{\downarrow\uparrow} C_{\lambda i}^{\uparrow} + \sum_{\nu} \tilde{B}_{\mu\nu}^{\downarrow\uparrow} C_{\nu i}^{\uparrow} & (\sigma = \downarrow). \end{cases} \quad (\text{A17}) \end{aligned}$$

For spin-restricted cases,  $\tilde{B}_{\mu\nu}^{\uparrow\downarrow} = \tilde{B}_{\mu\nu}^{\downarrow\uparrow} := \tilde{B}_{\mu\nu}$ ,  $C_{\nu i}^{\uparrow} = C_{\nu i}^{\downarrow} := C_{\nu i}$ , and  $\tilde{B}_{\mu\nu} = \tilde{B}_{\nu\mu}$  because of the symmetry between spins. Thus,  $B_{\mu i}^{\uparrow} = B_{\mu i}^{\downarrow} = \sum_{\nu} 2\tilde{B}_{\mu\nu} C_{\nu i}$ . Notice that we compute the derivatives of the energy with respect to the geminal matrix elements in this study simply because TurboRVB internally holds these values. If one uses a QMC package holding the derivatives of the energy with respect to molecular orbitals, one can just use them for the right-hand side tensor  $\mathbf{B}$  of Eq. (13).

The derivatives of the energy with respect to the atomic positions, i.e., the nuclear forces, with the so-called space warp coordinate transformation (SWCT)<sup>12</sup> are computed as<sup>15</sup>

$$\mathbf{F}_{\alpha}^{\text{VMC}} = -\frac{dE}{d\mathbf{R}_{\alpha}} = -\left\langle \frac{\partial E_L}{\partial \mathbf{R}_{\alpha}} + \sum_i \omega(\mathbf{R}_{\alpha}, \mathbf{r}_i) \frac{\partial E_L}{\partial \mathbf{r}_i} \right\rangle \quad (\text{A18})$$

$$\begin{aligned} &-2\left\langle (E_L - E) \left( \frac{\partial \log \Psi}{\partial \mathbf{R}_{\alpha}} + \sum_i \left[ \omega(\mathbf{R}_{\alpha}, \mathbf{r}_i) \frac{\partial \log \Psi}{\partial \mathbf{r}_i} \right. \right. \right. \\ &\left. \left. \left. + \frac{1}{2} \frac{\partial \omega(\mathbf{R}_{\alpha}, \mathbf{r}_i)}{\partial \mathbf{r}_i} \right] \right) \right\rangle, \quad (\text{A19}) \end{aligned}$$

where

$$\omega_{\alpha}(\mathbf{r}) = \frac{\kappa(|\mathbf{r} - \mathbf{R}_{\alpha}|)}{\sum_{\beta=1}^M \kappa(|\mathbf{r} - \mathbf{R}_{\beta}|)}, \quad (\text{A20})$$

with  $\kappa(r) = 1/r^4$ .<sup>12</sup> In order to evaluate these differential expressions,  $6N_e + 6N_a$  components have to be evaluated, namely,  $\left\langle \frac{\partial E_L}{\partial \mathbf{r}_i} \right\rangle$ ,  $\left\langle \frac{\partial \log \Psi}{\partial \mathbf{r}_i} \right\rangle$ ,  $\left\langle \frac{\partial E_L}{\partial \mathbf{R}_{\alpha}} \right\rangle$ , and  $\left\langle \frac{\partial \log \Psi}{\partial \mathbf{R}_{\alpha}} \right\rangle$ .<sup>15</sup> These values are efficiently computed in TurboRVB by using the automatic differentiation technique.

When computing the above-mentioned derivatives, the so-called reweighting method is essential to mitigate the heavy-tailed distribution problem mentioned in Sec. I. In the MCMC, the probability distribution can also differ from  $\pi(\mathbf{x})$ . Indeed, one can use an arbitrary probability distribution function  $\pi'(\mathbf{x}) = \Psi_G^2(\mathbf{x}) / \int d\mathbf{x} \Psi_G^2(\mathbf{x})$ , and estimate a generic local observable  $O(\mathbf{x})$  either by using  $\bar{O}_{\text{MCMC}} = \langle O(\mathbf{x}) \rangle_{\pi(\mathbf{x})}$  or

$$\bar{O}_{\text{MCMC}} = \frac{\langle O(\mathbf{x}') \mathcal{W}(\mathbf{x}') \rangle_{\pi'(\mathbf{x}')}}{\langle \mathcal{W}(\mathbf{x}') \rangle_{\pi'(\mathbf{x}')}} \approx \frac{\sum_{i=1}^{M'} O(\mathbf{x}'_i) \mathcal{W}(\mathbf{x}'_i)}{\sum_{i=1}^{M'} \mathcal{W}(\mathbf{x}'_i)}, \quad (\text{A21})$$

where  $\mathcal{W}(\mathbf{x}) = |\Psi(\mathbf{x}) / \Psi_G(\mathbf{x})|^2$  and the points  $\mathbf{x}'_i$  are distributed according to  $\pi'$ . This *reweighting* scheme is very important when evaluating energy gradients with finite variances. TurboRVB implements the so-called Attaccalite and Sorella (AS) reweighting scheme<sup>16</sup> to suppress the divergences of derivatives in the vicinity of the nodal surface.

#### 4. Error propagation

Important for error propagation is the fact that the elements  $B_{\mu i}^{\sigma}$  come with an error bar  $\delta B_{\mu i}^{\sigma}$ , whereby these elements are also correlated with each other. In the following, it is most convenient to consider the tensor  $\mathbf{B}$  as a vector with a single compound index  $B_{\mu i}^{\sigma} = B_{(\mu i \sigma)} = B_I$ , and the elements of  $\mathbf{A}$  in Eq. (13),  $A_{\mu i, \nu j}^{\sigma\tau} = A_{(\mu i \sigma), (\nu j \tau)} = A_{I, J}$ , as forming a matrix of dimension  $(N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}}) \times (N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}})$ . Let us denote the covariance matrix (tensor) of  $X$  as  $\text{Cov}(\mathbf{X})$ . With the element-wise notation, Eq. (13) is written as

$$\sum_{\nu j \tau} A_{\mu i, \nu j}^{\sigma\tau} Z_{\nu j}^{\tau} = -B_{\mu i}^{\sigma}, \quad (\text{A22})$$

$$\sum_{\mu i \sigma} (A^{-1})_{\lambda k, \mu i}^{\eta\sigma} \sum_{\nu j \tau} A_{\mu i, \nu j}^{\sigma\tau} Z_{\nu j}^{\tau} = -\sum_{\mu i \sigma} (A^{-1})_{\lambda k, \mu i}^{\eta\sigma} B_{\mu i}^{\sigma}, \quad (\text{A23})$$

$$\sum_{\mu\sigma} \sum_{\nu\tau} (A^{-1})_{\lambda k, \mu\sigma}^{\eta\sigma} A_{\mu\sigma, \nu\tau}^{\sigma\tau} Z_{\nu\tau}^{\tau} = - \sum_{\mu\sigma} (A^{-1})_{\lambda k, \mu\sigma}^{\eta\sigma} B_{\mu\sigma}^{\sigma} \quad (\text{A24})$$

$$\sum_{\nu\tau} \delta_{\nu\lambda} \delta_{jk} \delta_{\eta\tau} Z_{\nu\tau}^{\tau} = - \sum_{\mu\sigma} (A^{-1})_{\lambda k, \mu\sigma}^{\eta\sigma} B_{\mu\sigma}^{\sigma} \quad (\text{A25})$$

$$Z_{\lambda k}^{\eta} = - \sum_{\mu\sigma} (A^{-1})_{\lambda k, \mu\sigma}^{\eta\sigma} B_{\mu\sigma}^{\sigma} \quad (\text{A26})$$

In other words, we have

$$Z_{\mu\sigma}^{\sigma} = - \sum_{\nu\tau} G_{\mu\sigma, \nu\tau}^{\sigma\tau} B_{\nu\tau}^{\tau} \quad (\text{A27})$$

where  $G$  is an analog of the matrix inverse

$$\sum_{\mu\sigma} G_{\lambda k, \mu\sigma}^{\eta\sigma} A_{\mu\sigma, \nu\tau}^{\sigma\tau} = \delta_{\nu\lambda} \delta_{jk} \delta_{\eta\tau}, \quad \sum_{\mu\sigma} A_{\lambda k, \mu\sigma}^{\eta\sigma} G_{\mu\sigma, \nu\tau}^{\sigma\tau} = \delta_{\nu\lambda} \delta_{jk} \delta_{\eta\tau}. \quad (\text{A28})$$

Therefore,  $\text{Cov}(\mathbf{Z})$  reads

$$\text{Cov}(\mathbf{Z})_{\mu\sigma, \nu\tau}^{\sigma\tau} = \sum_{\lambda k, \delta l, \zeta} G_{\mu\sigma, \lambda k}^{\sigma\eta} \text{Cov}(\mathbf{B})_{\lambda k, \delta l}^{\eta\zeta} (G^T)_{\delta l, \nu\tau}^{\zeta\tau} \quad (\text{A29})$$

These Lagrange multipliers enter the formula for the force only linearly and they get contracted with deterministic quantities only; see Eq. (18). So, once we obtain  $\text{Cov}(\mathbf{Z})$ , it is trivial to propagate the error to the forces. Indeed, in Eq. (18), the force correction term,  $F^c$ , can be written as

$$F^c = \text{const.} + \sum_{\mu\sigma} R_{\mu\sigma}^{\sigma} Z_{\mu\sigma}^{\sigma}, \quad (\text{A30})$$

where  $\text{const.}$  aggregates all terms that are *not* multiplied by  $Z_{\mu\sigma}^{\sigma}$  and

$$R_{\mu\sigma}^{\sigma} = \epsilon_{i\sigma} \sum_{\nu} \frac{\partial S_{\mu\nu}}{\partial R} C_{\nu i\sigma} + \sum_{\nu} \frac{\partial F_{\mu\nu\sigma}}{\partial R} C_{\nu i\sigma} - \sum_{jk} \sum_{\lambda\nu\tau} I_{\mu\sigma, jk}^{\sigma\tau} C_{\lambda j}^{\tau} \frac{\partial S_{\lambda\nu}}{\partial R} C_{\nu k}^{\tau}, \quad (\text{A31})$$

with  $I_{\mu\sigma, jk}^{\sigma\tau}$  containing the two-electron integrals entering in the definition of  $W_{ij}^{\sigma}$  through the Fock matrix derivative with respect to the MO coefficients, i.e., Eq. (15). Thus, the variance of  $F^c$  is

$$\text{Var}(F^c) = \sum_{\mu\sigma} \sum_{\nu\tau} R_{\mu\sigma}^{\sigma} \text{Cov}(\mathbf{Z})_{\mu\sigma, \nu\tau}^{\sigma\tau} R_{\nu\tau}^{\tau}. \quad (\text{A32})$$

By plugging  $\text{Cov}(\mathbf{B})$  into the above-mentioned equation, we obtain

$$\text{Var}(F^c) = \sum_{\mu\sigma} \sum_{\lambda k} \sum_{\delta l} \sum_{\nu\tau} R_{\mu\sigma}^{\sigma} G_{\mu\sigma, \lambda k}^{\sigma\eta} \text{Cov}(\mathbf{B})_{\lambda k, \delta l}^{\eta\zeta} (G^T)_{\delta l, \nu\tau}^{\zeta\tau} R_{\nu\tau}^{\tau}. \quad (\text{A33})$$

Dropping the element-wise notation, we get

$$\text{Var}(F^c) = \bar{R}^T \mathbf{G} \text{Cov}(\mathbf{B}) \mathbf{G}^T \bar{R}. \quad (\text{A34})$$

The computation of Eq. (A34) involves the contraction of matrices of dimension  $(N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}}) \times (N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}})$ , which results in an overall sixth power with respect to system size. This will limit the maximum system size for which we

will be able to compute the error estimation on the forces. The computation of the  $\mathbf{G}$  tensor itself also scales with the sixth power in system size. A possible way to reduce the computational cost is to perform a low-rank decomposition of the  $\mathbf{B}$  tensor,

$$\mathbf{B} = \mathbf{b} \mathbf{b}^T, \quad (\text{A35})$$

where  $\mathbf{b} \in \mathcal{R}^{(N_{\text{AO}} \times N_{\text{MO}} \times N_{\text{spin}}) \times N_{\text{b}}}$ , where  $N_{\text{b}}$  is very small. Then, for each of the  $N_{\text{b}}$  vectors,  $\mathbf{b}_k$ , we have that  $\mathbf{b}_k^T \mathbf{G}$  is the same as solving  $\mathbf{A} \mathbf{Z} = \mathbf{B}$ , and we can then use a slightly modified version of the force evaluation subroutine to compute  $\mathbf{Z} \bar{R}$ .

## 5. Computational details

The DFT calculations (including the LR calculations) and the VMC calculations in this work were performed on cluster machines operated by the National Institute for Materials Science (NIMS), on the Genkai supercomputer at Kyushu University, and on the Fugaku supercomputer at RIKEN. For the VMC optimization for  $\text{Cl}_2$  dimer, cBN, and  $\text{AcNH}_2$  dimer test cases reported in Sec. III, we adjusted the number of independent MCMC chains and the per-chain MCMC sample count so that the error bar of the per-iteration VMC energy estimate was 5.0 mHa. The wavefunction optimization was then run for 300 iterations. As described in the main text, only the Jastrow factor was optimized variationally. Denoting by  $\vec{f}$ , the vector of the parameter gradients  $\equiv \{\dots, -\frac{\partial E}{\partial p_i}, \dots\}$ , where  $p_i$  are the variational parameters, we monitored  $\text{devmax} \equiv \max_i (|f_i| / |\text{std } f_i|)$  and verified for all parameters converged in a statistical sense ( $\text{devmax} \leq 4$ ). The optimization rate and regularization parameters were set to  $\text{tpar} = 0.35$  and  $\text{parr} = 0.001$ , respectively. See Ref. 50 for the implementation detail in TurboRVB. After optimization, we performed fresh MCMC simulations to compute the total energy, atomic forces, and the parameter gradients. The energies/forces and parameter gradients were evaluated in separate MCMC runs. For the energy and forces runs, the number of MCMC steps was adjusted so that the error bars on the total energy estimate were 0.1, 0.2, and 0.1 mHa for the  $\text{Cl}_2$ , cBN, and  $\text{AcNH}_2$  dimer, respectively; the resulting error bars on the force components were of comparable magnitude. For the parameter-gradients runs, after equilibration, the total pre-blocking MCMC sample count was set to  $N_{\text{sample}} = n_{\text{chain}} \times n_{\text{samples\_per\_chain}} = 120 \times 30\,000 = 3\,600\,000$ . For the error bar estimations of the force corrections, we set the number of blocks to 50 per chain. For ethanol, malonaldehyde, and benzene calculations reported in Sec. IV, we adjusted the MCMC sample count such that the error bar of the per-iteration VMC energy estimate was 5.0 mHa and performed 100 iterations of VMC optimization. As in Sec. III, only the Jastrow factor was optimized and  $\text{devmax}$  was confirmed for all parameters. We used  $\text{tpar} = 0.35$  and  $\text{parr} = 0.001$ . After optimization, energies, forces, and the parameter gradients were recomputed by MCMC in separate runs. For the energy and force runs, the number of MCMC steps was tuned so that the error bar of the energy estimate was 0.1 mHa, leading to force uncertainties of a similar magnitude. For the parameter-gradient runs, the total number of the pre-blocking MCMC steps was set to  $N_{\text{sample}} = n_{\text{chain}} \times n_{\text{samples\_per\_chain}} = 128 \times 50\,000 = 6\,400\,000$ . For the calculations in Sec. IV, we did not evaluate the error bar of the force-correction term.

## 6. Parameter study for the rMD17 benchmark test

VMC results are strongly dependent on the chosen *Ansatz*, such as the basis set in the determinant part, the exchange–correlation functional used to prepare the determinant (if it is not optimized at the VMC level), and the functional form of the Jastrow part and its employed basis set. Therefore, we did a benchmark test for many combinations of them. In particular, we tested cc-pVTZ and cc-pVQZ basis sets accompanied with ccECP for the determinant part, HF calculation and LDA-PZ and B3LYP DFT calculations to prepare the determinant part, the Padé and exponential type functional forms for the two-body Jastrow part, including and not-including the so-called four-body Jastrow parameters (i.e., the pairs of different nuclei in the Jastrow matrix elements) with a small (H:3s, C:3s1p, O:3s1p), medium (H:3s1p, C:4s2p1d, O:4s2p1d), or large (H:4s2p1d, C:5s3p2d, O:5s3p2d) Jastrow basis set. In this benchmark test set, we used only 20 out of 100 structures.

As metrics for assessing the quality of the results, we prepared two values. The first one is the VMC energy reported as the value relative to the result for label 10. For each configuration  $i$  in the test set, we define the energy error of the label  $x$  relative to the reference (label 10) as

$$\Delta E_i^{(\text{label } x)} = E_i^{(\text{label } x)} - E_i^{(\text{label } 10)}, \quad (\text{A36})$$

where  $E_i^{(\text{label } x)}$  is the energy computed with the parameter set (label  $x$ ) and  $E_i^{(\text{label } 10)}$  is the reference energy. The mean error ( $\text{ME}_E$ ) is then given by

$$\text{ME}_E^{(\text{label } x)} = \frac{1}{N} \sum_{i=1}^N \Delta E_i^{(\text{label } x)}, \quad (\text{A37})$$

where  $N$  is the number of configurations ( $N = 20$  in this study). With this notation,  $\text{ME}_E > 0$  ( $\text{ME}_E < 0$ ) indicates that label  $x$  gives worse (better) WF than the reference one on average according to the variational principle. The second one, which is a more direct metric for

this study, is the root-mean-square error (RMSE) of all the force components with respect to the reference all-electron CCSD(T) ones computed with the cc-pVQZ basis set. For each configuration  $i$  and force component  $\alpha$ , we define the force error of a parameter set (label  $x$ ) relative to the reference [all-electron CCSD(T)] as

$$\Delta F_{i,\alpha}^{(\text{label } x)} = F_{i,\alpha}^{(\text{label } x)} - F_{i,\alpha}^{(\text{CCSD(T)})}, \quad (\text{A38})$$

where  $F_{i,\alpha}^{(\text{label } x)}$  is the force component  $\alpha$  computed with the parameter set (label  $x$ ) and  $F_{i,\alpha}^{(\text{CCSD(T)})}$  is the reference force. The RMSE is computed as

$$\text{RMSE}_F^{(\text{label } x)} = \sqrt{\frac{1}{NM} \sum_{i=1}^N \sum_{\alpha=1}^M [\Delta F_{i,\alpha}^{(\text{label } x)}]^2}, \quad (\text{A39})$$

where  $N$  is the number of configurations and  $M$  is the number of force components ( $M = 3 \times$  the number of atoms in the system). The obtained results summarized in Tables III–V.

We found, as expected, that the VMC energy decreases as the Jastrow factor becomes more flexible. For example, comparisons between labels (1 and 2), (4 and 5), (7 and 8), or (9 and 10) indicate that including the four-body term reduces  $\text{ME}_E$ . However, comparisons such as labels (2 and 7) or (5 and 9) show that labels 7 and 9 yield smaller  $\text{ME}_E$  than labels 2 and 5, respectively, despite having fewer variational parameters. This suggests that using only the three-body term with a larger Jastrow basis is a more efficient way to gain the correlation energy than including a four-body term with a smaller Jastrow basis. For ethanol and malonaldehyde, we also tested an even larger Jastrow basis (label 13). Although some of the 20 configurations exhibited smaller  $\text{ME}_E$  than the medium-sized Jastrow basis set (label 8), the average  $\text{ME}_E$  became positive. We interpret this as indicating that the optimization becomes less robust for this oversized Jastrow parameterization, probably due to local minima. Therefore, in the main text, we adopt the medium-sized Jastrow basis set. In contrast to  $\text{ME}_E$ , the RMSE for force ( $\text{RMSE}_F$ )

**TABLE III.** Parameter study for ethanol. For each label, the determinant basis set, ECP, XC, the Jastrow parameters (2-body and 3-/4-body), the Jastrow basis set, and the total number of variational parameters are listed, together with the mean error in energy  $\text{ME}_E$  (Ha) and the root-mean-square error in corrected force  $\text{RMSE}_F$  (kcal/mol/Å).

Label	Det. basis	ECP	XC	Jastrow 2b	Jastrow 3b/4b	Jas. basis	Parameter	$\text{ME}_E$ (Ha)	$\text{RMSE}_F$ (kcal/mol/Å)
1	cc-pVTZ	ccECP	B3LYP	Padé	3b	Small	148	+0.4749(19)	3.451(17)
2	cc-pVTZ	ccECP	B3LYP	Padé	3b/4b	Small	715	+0.2542(19)	2.976(17)
3	cc-pVTZ	ccECP	B3LYP	Exp.	3b/4b	Small	715	+0.3480(19)	3.267(16)
4	cc-pVTZ	ccECP	LDA-PZ	Padé	3b	Small	148	+0.4443(19)	3.198(16)
5	cc-pVTZ	ccECP	LDA-PZ	Padé	3b/4b	Small	715	+0.2296(19)	2.809(17)
6	cc-pVTZ	ccECP	HF	Padé	3b/4b	Small	715	+0.4843(18)	3.711(17)
7	cc-pVTZ	ccECP	B3LYP	Padé	3b	Medium	587	+0.0832(18)	2.515(17)
8	cc-pVTZ	ccECP	B3LYP	Padé	3b/4b	Medium	3422	+0.0387(18)	2.741(18)
9	cc-pVTZ	ccECP	LDA-PZ	Padé	3b	Medium	587	+0.0664(19)	2.344(17)
10	cc-pVTZ	ccECP	LDA-PZ	Padé	3b/4b	Medium	3422	+0.0(0)	2.469(19)
11	cc-pVTZ	ccECP	HF	Padé	3b/4b	Medium	3422	+0.2224(19)	3.480(20)
12	cc-pVQZ	ccECP	B3LYP	Padé	3b/4b	Small	715	+0.1811(18)	3.087(16)
13	cc-pVTZ	ccECP	B3LYP	Padé	3b/4b	Large	13 394	+0.0578(19)	3.313(27)

**TABLE IV.** Parameter study for malonaldehyde. For each label, the determinant basis set, ECP, XC, the Jastrow parameters (2-body and 3-/4-body), the Jastrow basis set, and the total number of variational parameters are listed, together with the mean error in energy  $ME_E$  (Ha) and the root-mean-square error in corrected force  $RMSE_F$  (kcal/mol/Å).

Label	Det. Basis	ECP	XC	Jastrow 2b	Jastrow 3b/4b	Jas. Basis	Parameter	$ME_E$ (Ha)	$RMSE_F$ (kcal/mol/Å)
1	cc-pVTZ	ccECP	B3LYP	Pade	3b	Small	184	+0.6652(27)	6.953(24)
2	cc-pVTZ	ccECP	B3LYP	Pade	3b/4b	Small	958	+0.3877(27)	7.012(25)
3	cc-pVTZ	ccECP	B3LYP	Exp.	3b/4b	Small	958	+0.5266(28)	7.525(25)
4	cc-pVTZ	ccECP	LDA-PZ	Pade	3b	Small	184	+0.6275(28)	6.953(26)
5	cc-pVTZ	ccECP	LDA-PZ	Pade	3b/4b	Small	958	+0.3922(27)	6.974(25)
6	cc-pVTZ	ccECP	HF	Pade	3b/4b	Small	958	+0.8491(28)	9.025(28)
7	cc-pVTZ	ccECP	B3LYP	Pade	3b	Medium	803	+0.0893(27)	6.886(27)
8	cc-pVTZ	ccECP	B3LYP	Pade	3b/4b	Medium	5069	+0.0195(28)	7.213(28)
9	cc-pVTZ	ccECP	LDA-PZ	Pade	3b	Medium	803	+0.1125(27)	6.939(27)
10	cc-pVTZ	ccECP	LDA-PZ	Pade	3b/4b	Medium	5069	+0.0(0)	6.967(30)
11	cc-pVTZ	ccECP	HF	Pade	3b/4b	Medium	5069	+0.4373(28)	8.937(30)
12	cc-pVQZ	ccECP	B3LYP	Pade	3b/4b	Small	958	+0.2624(29)	7.946(28)
13	cc-pVTZ	ccECP	B3LYP	Pade	3b/4b	Large	16 499	+0.1427(28)	7.742(29)

**TABLE V.** Parameter study for benzene. For each label, the determinant basis set, ECP, XC, the Jastrow parameters (2-body and 3-/4-body), the Jastrow basis set, and the total number of variational parameters are listed, together with the mean error in energy  $ME_E$  (Ha) and the root-mean-square error in corrected force  $RMSE_F$  (kcal/mol/Å).

Label	Det. Basis	ECP	XC	Jastrow 2b	Jastrow 3b/4b	Jas. Basis	Parameter	$ME_E$ (Ha)	$RMSE_F$ (kcal/mol/Å)
1	cc-pVTZ	ccECP	B3LYP	Pade	3b	Small	225	+0.6395(27)	2.854(18)
2	cc-pVTZ	ccECP	B3LYP	Pade	3b/4b	Small	1548	+0.4278(27)	2.494(19)
3	cc-pVTZ	ccECP	B3LYP	Exp.	3b/4b	Small	1548	+0.4589(27)	2.456(17)
4	cc-pVTZ	ccECP	LDA-PZ	Pade	3b	Small	225	+0.5619(27)	2.735(18)
5	cc-pVTZ	ccECP	LDA-PZ	Pade	3b/4b	Small	1548	+0.3625(27)	2.518(19)
6	cc-pVTZ	ccECP	HF	Pade	3b/4b	Small	1548	+0.7913(27)	2.493(19)
7	cc-pVTZ	ccECP	B3LYP	Pade	3b	Medium	985	+0.1601(26)	2.280(18)
8	cc-pVTZ	ccECP	B3LYP	Pade	3b/4b	Medium	8140	+0.0542(27)	2.448(19)
9	cc-pVTZ	ccECP	LDA-PZ	Pade	3b	Medium	985	+0.1288(26)	2.544(22)
10	cc-pVTZ	ccECP	LDA-PZ	Pade	3b/4b	Medium	8140	+0.0(0)	2.091(21)
11	cc-pVTZ	ccECP	HF	Pade	3b/4b	Medium	8140	+0.3252(26)	2.480(18)

does not exhibit a monotonic improvement upon increasing the Jastrow flexibility. For instance, on focusing on labels 9 and 10 for the benzene molecule, the inclusion of the 4-body Jastrow improves from 2.544(22) kcal/mol/Å (label 9) to 2.091(21) kcal/mol/Å (label 10), while for ethanol and malonaldehyde, it does not show a similar improvement. This should reflect the fact that, unlike the total energy, forces are differential quantities (i.e., energy differences with respect to atomic displacements) and improving the wave function does not necessarily improve forces. On the other hand, there are cases where the choice of the determinant part clearly matters. For example, HF-based calculations (labels 6 and 11) yield substantially worse  $RMSE_F$  than LDA-based calculations (e.g., labels 5 and 10, respectively) for all three molecules. Regarding the basis set for the determinant, we also tested cc-pVQZ for ethanol and malonaldehyde (label 13), but we did not observe improvement in  $RMSE_F$ ; therefore, we adopt cc-pVTZ in the main text. We did not find a single parameter set simultaneously yielding the best  $ME_E$  and  $RMSE_F$  for all three compounds. Therefore, we adopt label 10 in the main

text because it gives the smallest  $ME_E$  for all three compounds and also provides good  $RMSE_F$  values.

## APPENDIX B: DERIVATION OF LAGRANGE MULTIPLIERS EQUATIONS

To arrive at the algebraic equations that determine the Lagrange multipliers, we need to explicitly evaluate the derivatives of the VMC Lagrangian with respect to the MO coefficients and project them onto the occupied and virtual orbital manifolds. We start by computing the derivatives appearing in Eq. (11), which we repeat here for convenience,

$$\frac{\partial L_{VMC}}{\partial C_{\mu i}^{\sigma}} = \frac{\partial E_{VMC}}{\partial C_{\mu i}^{\sigma}} + \sum_{\lambda k} Z_{\lambda k}^{\tau} \sum_{\nu} \frac{\partial}{\partial C_{\mu i}^{\sigma}} (F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau}) - \sum_{kl} W_{kl}^{\tau} \frac{\partial S_{kl}^{\tau}}{\partial C_{\mu i}^{\sigma}}. \quad (B1)$$

The first term on the right-hand side is the derivative of the VMC energy with respect to the MO coefficients, which we simply label as  $B_{\mu i}^{\sigma}$  as in the main text. How these derivatives are obtained from the VMC calculation is discussed in [Appendix A 3](#).

The second term in Eq. (B1) corresponds to the derivative of the Roothaan–Hall equations, reading

$$\sum_{\lambda k \tau} Z_{\lambda k}^{\tau} \sum_{\nu} \frac{\partial}{\partial C_{\mu i}^{\sigma}} (F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau}). \quad (\text{B2})$$

It is more convenient to work out the action of the derivative operator on the Fock and overlap components separately. By the chain rule, the first part is given by

$$\frac{\partial}{\partial C_{\mu i}^{\sigma}} F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} = \frac{\partial F_{\lambda \nu}^{\tau}}{\partial C_{\mu i}^{\sigma}} C_{\nu k}^{\tau} + F_{\lambda \nu}^{\tau} \frac{\partial C_{\nu k}^{\tau}}{\partial C_{\mu i}^{\sigma}}, \quad (\text{B3})$$

where the derivative of the Fock matrix with respect to the MO coefficients reads

$$\begin{aligned} \frac{\partial F_{\lambda \nu}^{\tau}}{\partial C_{\mu i}^{\sigma}} &= \frac{\partial}{\partial C_{\mu i}^{\sigma}} \left( h_{\lambda \nu} + \sum_{\gamma \delta} P_{\gamma \delta}^{\eta} [(\lambda \nu | \gamma \delta) - c_x \delta_{\tau \eta} (\lambda \gamma | \nu \delta)] + V_{\lambda \nu}^{\text{xc}, \tau} \right) \\ &= \sum_{\gamma \delta \eta} \frac{\partial P_{\gamma \delta}^{\eta}}{\partial C_{\mu i}^{\sigma}} [(\lambda \nu | \gamma \delta) - c_x \delta_{\tau \eta} (\lambda \gamma | \nu \delta)] + \frac{\partial V_{\lambda \nu}^{\text{xc}, \tau}}{\partial C_{\mu i}^{\sigma}} \\ &= \sum_{\gamma \delta \eta} \left( \delta_{\sigma \eta} [\delta_{\mu \gamma} C_{\delta i}^{\eta} + \delta_{\mu \delta} C_{\gamma i}^{\eta}] \right) [(\lambda \nu | \gamma \delta) - c_x \delta_{\tau \eta} (\lambda \gamma | \nu \delta)] + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} \\ &= \sum_{\gamma \delta \eta} \delta_{\sigma \eta} \delta_{\mu \gamma} C_{\delta i}^{\eta} (\lambda \nu | \gamma \delta) + \sum_{\gamma \delta \eta} \delta_{\sigma \eta} \delta_{\mu \delta} C_{\gamma i}^{\eta} (\lambda \nu | \gamma \delta) \\ &\quad - \sum_{\gamma \delta \eta} c_x \delta_{\sigma \eta} \delta_{\tau \eta} \delta_{\mu \gamma} C_{\delta i}^{\eta} (\lambda \gamma | \nu \delta) - \sum_{\gamma \delta \eta} c_x \delta_{\sigma \eta} \delta_{\tau \eta} \delta_{\mu \delta} C_{\gamma i}^{\eta} (\lambda \gamma | \nu \delta) + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} \end{aligned}$$

$$\begin{aligned} &\sum_{\lambda k \tau} Z_{\lambda k}^{\tau} \sum_{\nu} \frac{\partial (F_{\lambda \nu}^{\tau} C_{\nu k}^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau})}{\partial C_{\mu i}^{\sigma}} \\ &= \sum_{\lambda k \tau} Z_{\lambda k}^{\tau} \sum_{\nu} \left[ C_{\nu k}^{\tau} (2(\lambda \nu | \mu i \sigma) - c_x \delta_{\tau \sigma} [(\lambda \mu | \nu i \sigma) + (\lambda i \sigma | \nu \mu)]) + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} + \delta_{\tau \sigma} \delta_{\nu \mu} \delta_{i k} F_{\lambda \nu}^{\tau} - \delta_{\tau \sigma} \delta_{\mu \nu} \delta_{i k} S_{\lambda \nu} \epsilon_k^{\tau} - S_{\lambda \nu} C_{\nu k}^{\tau} \frac{\partial \epsilon_k^{\tau}}{\partial C_{\mu i}^{\sigma}} \right] \\ &= \sum_{\lambda \nu k \tau} Z_{\lambda k}^{\tau} C_{\nu k}^{\tau} (2(\lambda \nu | \mu i \sigma) - c_x \delta_{\tau \sigma} [(\lambda \mu | \nu i \sigma) + (\lambda i \sigma | \nu \mu)]) + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} + \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda \mu}^{\sigma} - S_{\lambda \mu} \epsilon_i^{\sigma}) - \underbrace{\sum_{k \tau} \sum_{\lambda \nu} Z_{\lambda k}^{\tau} S_{\lambda \nu} C_{\nu k}^{\tau} \frac{\partial \epsilon_k^{\tau}}{\partial C_{\mu i}^{\sigma}}}_{=0} \\ &= \sum_{\lambda \nu \tau} \tilde{P}_{\lambda \nu}^{\tau} (2(\lambda \nu | \mu i \sigma) - c_x \delta_{\tau \sigma} [(\lambda \mu | \nu i \sigma) + (\lambda i \sigma | \nu \mu)]) + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} + \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda \mu}^{\sigma} - S_{\lambda \mu} \epsilon_i^{\sigma}) = H_{\mu i}^{\sigma} [\tilde{\mathbf{P}}] + \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda \mu}^{\sigma} - S_{\lambda \mu} \epsilon_i^{\sigma}), \quad (\text{B9}) \end{aligned}$$

where we have identified the intermediate tensor  $\mathbf{H}[\tilde{\mathbf{P}}]$  as the derivative of the Fock matrix with respect to the MO coefficients [see Eq. (15) in the main text], traced with the asymmetric response spin-density tensor  $\tilde{\mathbf{P}}$ . Notice that contrary to the main text, here we

$$\begin{aligned} &= \sum_{\delta} C_{\delta i}^{\sigma} (\lambda \nu | \mu \delta) + \sum_{\gamma} C_{\gamma i}^{\sigma} (\lambda \nu | \gamma \mu) \\ &\quad - \sum_{\delta} c_x \delta_{\tau \sigma} C_{\delta i}^{\sigma} (\lambda \mu | \nu \delta) - \sum_{\gamma} c_x \delta_{\tau \sigma} C_{\gamma i}^{\sigma} (\lambda \gamma | \nu \mu) + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} \\ &= 2(\lambda \nu | \mu i \sigma) - c_x \delta_{\tau \sigma} [(\lambda \mu | \nu i \sigma) + (\lambda i \sigma | \nu \mu)] + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}}. \quad (\text{B4}) \end{aligned}$$

Here, we have introduced the exchange–correlation kernel  $f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}}$ , defined as

$$f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}} = \iint \frac{\delta^2 E^{\text{xc}}}{\delta \rho_{\tau}(\mathbf{r}) \delta \rho_{\sigma}(\mathbf{r}')} \Omega_{\lambda \nu}(\mathbf{r}) \Omega_{\mu i}(\mathbf{r}') \mathbf{r} \mathbf{r}' d\mathbf{r} d\mathbf{r}', \quad (\text{B5})$$

where  $\Omega$  denotes an orbital-pair function,  $\Omega_{\lambda \nu}(\mathbf{r}) = \psi_{\lambda}(\mathbf{r}) \psi_{\nu}(\mathbf{r})$  (for AOs) and  $\Omega_{\mu i}(\mathbf{r}) = \phi_{\mu \sigma}(\mathbf{r}) \phi_{i \sigma}(\mathbf{r})$  (for MOs). The second term of Eq. (B3) is trivial,

$$F_{\lambda \nu}^{\tau} \frac{\partial C_{\nu k}^{\tau}}{\partial C_{\mu i}^{\sigma}} = F_{\lambda \nu}^{\tau} \delta_{\tau \sigma} \delta_{\nu \mu} \delta_{i k}. \quad (\text{B6})$$

Inserting Eqs. (B4) and (B6) into Eq. (B3) results in

$$\begin{aligned} \frac{\partial F_{\lambda \nu}^{\tau}}{\partial C_{\mu i}^{\sigma}} C_{\nu k}^{\tau} + F_{\lambda \nu}^{\tau} \frac{\partial C_{\nu k}^{\tau}}{\partial C_{\mu i}^{\sigma}} &= [2(\lambda \nu | \mu i \sigma) - c_x \delta_{\tau \sigma} [(\lambda \mu | \nu i \sigma) + (\lambda i \sigma | \nu \mu)] \\ &\quad + 2f_{\lambda \nu \tau, \mu i \sigma}^{\text{xc}}] C_{\nu k}^{\tau} + F_{\lambda \nu}^{\tau} \delta_{\tau \sigma} \delta_{\nu \mu} \delta_{i k}. \quad (\text{B7}) \end{aligned}$$

The second term of Eq. (B2) is given by

$$\begin{aligned} \frac{\partial}{\partial C_{\mu i}^{\sigma}} S_{\lambda \nu} C_{\nu k}^{\tau} \epsilon_k^{\tau} &= S_{\lambda \nu} \frac{\partial C_{\nu k}^{\tau}}{\partial C_{\mu i}^{\sigma}} \epsilon_k^{\tau} + S_{\lambda \nu} C_{\nu k}^{\tau} \frac{\partial \epsilon_k^{\tau}}{\partial C_{\mu i}^{\sigma}} \\ &= \delta_{\tau \sigma} \delta_{\mu \nu} \delta_{i k} S_{\lambda \nu} \epsilon_k^{\tau} + S_{\lambda \nu} C_{\nu k}^{\tau} \frac{\partial \epsilon_k^{\tau}}{\partial C_{\mu i}^{\sigma}}. \quad (\text{B8}) \end{aligned}$$

Plugging Eqs. (B7) and (B8) into Eq. (B2) finally results in

derive all the equations with the asymmetric response spin-density tensor, although, in practice, we use a symmetrized version as it is numerically more stable.

At last, the third term in Eq. (B1) is

$$\begin{aligned} \sum_{kl\tau} W_{kl}^{\tau} \frac{\partial S_{kl}^{\tau}}{\partial C_{\mu i}^{\sigma}} &= \sum_{kl\tau} W_{kl}^{\tau} \left( \frac{\partial}{\partial C_{\mu i}^{\sigma}} \sum_{\lambda\nu} C_{\lambda k}^{\tau} S_{\lambda\nu} C_{\nu l}^{\tau} \right) = \sum_{kl\tau} W_{kl}^{\tau} \\ &\times \left( \sum_{\lambda\nu} \delta_{\lambda\mu} \delta_{ik} \delta_{\tau\sigma} S_{\lambda\nu} C_{\nu l}^{\tau} + \sum_{\lambda\nu} \delta_{\mu\nu} \delta_{il} \delta_{\tau\sigma} S_{\lambda\nu} C_{\lambda k}^{\tau} \right) \\ &= \sum_{kl\tau} \delta_{ik} \delta_{\tau\sigma} W_{kl}^{\tau} \sum_{\nu} S_{\mu\nu} C_{\nu l}^{\tau} + \sum_{kl\tau} \delta_{il} \delta_{\tau\sigma} W_{kl}^{\tau} \sum_{\lambda} S_{\lambda\mu} C_{\lambda k}^{\tau} \\ &= \sum_{lv} W_{il}^{\sigma} S_{\mu\nu} C_{\nu l}^{\sigma} + \sum_{kl} W_{ki}^{\sigma} S_{\lambda\mu} C_{\lambda k}^{\sigma} = 2 \sum_{j\nu} W_{ij}^{\sigma} S_{\mu\nu} C_{\nu j}^{\sigma}, \end{aligned} \quad (\text{B10})$$

where in the last step, we simply relabeled the indices and used the fact that both  $\mathbf{W}$  and  $\mathbf{S}$  are symmetric.

To determine the Lagrange multipliers, we equate Eq. (B1) to zero and project it onto the occupied and virtual MO manifolds [see Eqs. (12a) and (12b) of the main text], solving for  $\mathbf{W}$  and  $\mathbf{Z}$ , respectively. Projection onto the virtual manifold results in a linear system of equations, known as the Z-vector equations, while  $\mathbf{W}$  drops out. Let us start by explicitly writing out Eq. (12a),

$$\begin{aligned} \sum_{\mu} \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} Q_{\mu\nu}^{\sigma} &= \underbrace{\sum_{\mu} B_{\mu i}^{\sigma} Q_{\mu\nu}^{\sigma}}_{\text{term I}} + \underbrace{\sum_{\mu} \left[ H_{\mu i}^{\sigma}[\tilde{\mathbf{P}}] + \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma}) \right]}_{\text{term II}} Q_{\mu\nu}^{\sigma} \\ &- \underbrace{\sum_{\mu} 2 \sum_{j\lambda} W_{ij}^{\sigma} S_{\mu\lambda} C_{\lambda j}^{\sigma} Q_{\mu\nu}^{\sigma}}_{\text{term III}} = 0, \end{aligned} \quad (\text{B11})$$

where we recall here the definition of  $Q_{\mu\nu}^{\sigma}$ ,

$$Q_{\mu\nu}^{\sigma} = \delta_{\mu\nu} - \sum_{\lambda k} C_{\mu k}^{\sigma} C_{\lambda k}^{\sigma} S_{\lambda\nu}. \quad (\text{B12})$$

Term I in Eq. (B1) remains untouched. Term III vanishes

$$\begin{aligned} \sum_{\mu} 2 \sum_{j\lambda} W_{ij}^{\sigma} S_{\mu\lambda} C_{\lambda j}^{\sigma} Q_{\mu\nu}^{\sigma} &= 2 \sum_{\mu j\lambda} W_{ij}^{\sigma} S_{\mu\lambda} C_{\lambda j}^{\sigma} \delta_{\mu\nu} \\ &- 2 \sum_{\mu j\lambda} \sum_{\delta k} W_{ij}^{\sigma} S_{\mu\lambda} C_{\lambda j}^{\sigma} C_{\mu k}^{\sigma} C_{\delta k}^{\sigma} S_{\delta\nu} \\ &= 2 \sum_{j\lambda} W_{ij}^{\sigma} S_{\nu\lambda} C_{\lambda j}^{\sigma} - 2 \sum_{j k \delta} \underbrace{W_{ij}^{\sigma} S_{k j}^{\sigma} C_{\delta k}^{\sigma}}_{=\delta_{kj}} S_{\delta\nu} \\ &= 2 \sum_{j\lambda} W_{ij}^{\sigma} S_{\nu\lambda} C_{\lambda j}^{\sigma} - 2 \sum_{j\delta} W_{ij}^{\sigma} S_{\delta\nu} C_{\delta j}^{\sigma} = 0, \end{aligned} \quad (\text{B13})$$

where we have used the fact that the overlap matrix is symmetric. Finally, we look at term II, which is the most complicated one,

$$\begin{aligned} \sum_{\mu} \left[ H_{\mu i}^{\sigma}[\tilde{\mathbf{P}}] + \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma}) \right] Q_{\mu\nu}^{\sigma} \\ = \underbrace{\sum_{\mu} H_{\mu i}^{\sigma}[\tilde{\mathbf{P}}]}_{\text{term II.A}} Q_{\mu\nu}^{\sigma} + \underbrace{\sum_{\mu\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma})}_{\text{term II.B}} Q_{\mu\nu}^{\sigma}. \end{aligned} \quad (\text{B14})$$

Term II.A does not simplify, while term II.B simplifies owing to the Roothaan–Hall equation, which vanishes when the Lagrangian is stationary with respect to the MO coefficients,

$$\begin{aligned} \sum_{\mu\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma}) Q_{\mu\nu}^{\sigma} &= \sum_{\mu\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma}) \delta_{\mu\nu} - \sum_{\mu\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\mu}^{\sigma} - S_{\lambda\mu} \epsilon_i^{\sigma}) \sum_{\delta k} C_{\mu k}^{\sigma} C_{\delta k}^{\sigma} S_{\delta\nu} \\ &= \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\nu}^{\sigma} - S_{\lambda\nu} \epsilon_i^{\sigma}) - \sum_{\lambda\delta k} Z_{\lambda i}^{\sigma} \sum_{\mu} (F_{\lambda\mu}^{\sigma} C_{\mu k}^{\sigma} - S_{\lambda\mu} C_{\mu k}^{\sigma} \epsilon_i^{\sigma}) C_{\delta k}^{\sigma} S_{\delta\nu} \\ &= \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\nu}^{\sigma} - S_{\lambda\nu} \epsilon_i^{\sigma}) - \sum_{\delta k} \sum_{\lambda\mu} Z_{\lambda i}^{\sigma} (S_{\lambda\mu} C_{\mu k}^{\sigma} \epsilon_k^{\sigma} - S_{\lambda\mu} C_{\mu k}^{\sigma} \epsilon_i^{\sigma}) C_{\delta k}^{\sigma} S_{\delta\nu} \\ &= \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\nu}^{\sigma} - S_{\lambda\nu} \epsilon_i^{\sigma}) - \sum_{\delta k} \left[ \underbrace{\sum_{\lambda\mu} Z_{\lambda i}^{\sigma} S_{\lambda\mu} C_{\mu k}^{\sigma}}_{=0} \right] \epsilon_k^{\sigma} C_{\delta k}^{\sigma} S_{\delta\nu} \\ &- \sum_{\delta k} \left[ \underbrace{\sum_{\lambda\mu} Z_{\lambda i}^{\sigma} S_{\lambda\mu} C_{\mu k}^{\sigma}}_{=0} \right] \epsilon_i^{\sigma} C_{\delta k}^{\sigma} S_{\delta\nu} = \sum_{\lambda} Z_{\lambda i}^{\sigma} (F_{\lambda\nu}^{\sigma} - S_{\lambda\nu} \epsilon_i^{\sigma}). \end{aligned} \quad (\text{B15})$$

Putting together Eqs. (B14) and (B15) results in

$$\begin{aligned} \sum_{\mu} H_{\mu i}^{\sigma}[\hat{\mathbf{P}}]Q_{\mu\nu}^{\sigma} + \sum_{\mu\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\mu}^{\sigma} - S_{\lambda\mu}\epsilon_i^{\sigma})Q_{\mu\nu}^{\sigma} \\ = \sum_{\mu} H_{\mu i}^{\sigma}[\hat{\mathbf{P}}]Q_{\mu\nu}^{\sigma} + \sum_{\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\nu}^{\sigma} - S_{\lambda\nu}\epsilon_i^{\sigma}). \end{aligned} \quad (\text{B16})$$

Finally, inserting Eqs. (B13) and (B16) into Eq. (B1) yields

$$\begin{aligned} \sum_{\mu} \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} Q_{\mu\nu}^{\sigma} = \sum_{\mu} B_{\mu i}^{\sigma} Q_{\mu\nu}^{\sigma} + \sum_{\mu} H_{\mu i}^{\sigma}[\hat{\mathbf{P}}]Q_{\mu\nu}^{\sigma} \\ + \sum_{\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\nu}^{\sigma} - S_{\lambda\nu}\epsilon_i^{\sigma}) = 0 \end{aligned} \quad (\text{B17})$$

$$\sum_{\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\nu}^{\sigma} - S_{\lambda\nu}\epsilon_i^{\sigma}) + \sum_{\mu} H_{\mu i}^{\sigma}[\hat{\mathbf{P}}]Q_{\mu\nu}^{\sigma} = -\sum_{\mu} B_{\mu i}^{\sigma} Q_{\mu\nu}^{\sigma},$$

which corresponds to Eq. (13) of the main text.

To determine the  $\mathbf{W}$  Lagrange multipliers, we project the derivative of the Lagrangian onto the occupied orbitals and equate to zero, cf. Eq. (12b). This equation is identical to Eq. (B11), with the only difference being the projector,

$$\begin{aligned} \sum_{\mu} \frac{\partial L_{\text{VMC}}}{\partial C_{\mu i}^{\sigma}} C_{\mu j}^{\sigma} = \sum_{\mu} B_{\mu i}^{\sigma} C_{\mu j}^{\sigma} + \sum_{\mu} \left[ H_{\mu i}^{\sigma}[\hat{\mathbf{P}}] + \sum_{\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\mu}^{\sigma} - S_{\lambda\mu}\epsilon_i^{\sigma}) \right] C_{\mu j}^{\sigma} \\ - \sum_{\mu} 2 \sum_{k\lambda} W_{ik}^{\sigma} S_{\mu\lambda} C_{\lambda k}^{\sigma} C_{\mu j}^{\sigma} = 0. \end{aligned} \quad (\text{B18})$$

No simplifications are possible for the first term, while the third term is simply

$$\sum_{\mu} 2 \sum_{k\lambda} W_{ik}^{\sigma} S_{\mu\lambda} C_{\lambda k}^{\sigma} C_{\mu j}^{\sigma} = 2 \sum_k \underbrace{W_{ik}^{\sigma} \sum_{\mu\lambda} C_{\mu j}^{\sigma} S_{\mu\lambda} C_{\lambda k}^{\sigma}}_{=\delta_{jk}} = 2W_{ij}^{\sigma}. \quad (\text{B19})$$

The middle term simplifies to

$$\begin{aligned} \sum_{\mu} \left[ H_{\mu i}^{\sigma}[\hat{\mathbf{P}}] + \sum_{\lambda} Z_{\lambda i}^{\sigma}(F_{\lambda\mu}^{\sigma} - S_{\lambda\mu}\epsilon_i^{\sigma}) \right] C_{\mu j}^{\sigma} \\ = H_{ji}^{\sigma}[\hat{\mathbf{P}}] + \sum_{\mu\lambda} Z_{\lambda i}^{\sigma} F_{\lambda\mu}^{\sigma} C_{\mu j}^{\sigma} - \underbrace{\sum_{\mu\lambda} Z_{\lambda i}^{\sigma} S_{\lambda\mu} C_{\mu j}^{\sigma} \epsilon_i^{\sigma}}_{=0}, \end{aligned} \quad (\text{B20})$$

$$= H_{ij}^{\sigma}[\hat{\mathbf{P}}] + \underbrace{\sum_{\mu\lambda} Z_{\lambda i}^{\sigma} S_{\lambda\mu} C_{\mu j}^{\sigma} \epsilon_j^{\sigma}}_{=0}, \quad (\text{B21})$$

$$= H_{ij}^{\sigma}[\hat{\mathbf{P}}], \quad (\text{B22})$$

using the symmetry of  $\mathbf{H}$  and Eq. (7). With these results, we can rearrange Eq. (B18) to solve for  $\mathbf{W}$ , yielding

$$W_{ij}^{\sigma} = \frac{1}{2} H_{ij}^{\sigma}[\hat{\mathbf{P}}] + \frac{1}{2} \sum_{\mu} B_{\mu i}^{\sigma} C_{\mu j}^{\sigma} \quad (\text{B23})$$

that is reported as Eq. (17) in the main text.

## REFERENCES

- R. Shinde, C. Filippi, A. Scemama, and W. Jalby, *Nat. Rev. Phys.* **7**, 378–387 (2025).
- W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- F. Della Pia, B. X. Shi, Y. S. Al-Hamdani, D. Alf e, T. A. Anderson, M. Barborini, A. Benali, M. Casula, N. D. Drummond, M. Dubeck y, C. Filippi, P. R. C. Kent, J. T. Krogel, P. L opez R os, A. L uchow, Y. Luo, A. Michaelides, L. Mitas, K. Nakano, R. J. Needs, M. C. Per, A. Scemama, J. Schultze, R. Shinde, E. Sloatman, S. Sorella, A. Tkatchenko, M. Towler, C. J. Umrigar, L. K. Wagner, W. A. Wheeler, H. Zhou, and A. Zen, *J. Chem. Phys.* **163**, 104110 (2025).
- P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester, Jr., *Int. J. Quantum Chem.* **29**, 589 (1986).
- S. Chiesa, D. M. Ceperley, and S. Zhang, *Phys. Rev. Lett.* **94**, 036404 (2005).
- A. Badinski and R. J. Needs, *Phys. Rev. B* **78**, 035134 (2008).
- R. Assaraf, M. Caffarel, and A. C. Kollias, *Phys. Rev. Lett.* **106**, 150601 (2011).
- S. Moroni, S. Saccani, and C. Filippi, *J. Chem. Theory Comput.* **10**, 4823 (2014).
- J. Van Rhijn, C. Filippi, S. De Palo, and S. Moroni, *J. Chem. Theory Comput.* **18**, 118 (2021).
- S. Chen and S. Zhang, *Phys. Rev. B* **107**, 195150 (2023).
- R. Assaraf and M. Caffarel, *J. Chem. Phys.* **113**, 4028 (2000).
- C. J. Umrigar, *Int. J. Quantum Chem.* **36**, 217 (1989).
- C. Filippi and C. J. Umrigar, *Phys. Rev. B* **61**, R16291 (2000).
- R. Assaraf and M. Caffarel, *J. Chem. Phys.* **119**, 10536 (2003).
- S. Sorella and L. Capriotti, *J. Chem. Phys.* **133**, 234111 (2010).
- C. Attaccalite and S. Sorella, *Phys. Rev. Lett.* **100**, 114501 (2008).
- S. Pathak and L. K. Wagner, *AIP Adv.* **10**, 085213 (2020).
- J. R. Trail, *Phys. Rev. E* **77**, 016703 (2008).
- J. R. Trail, *Phys. Rev. E* **77**, 016704 (2008).
- P. L. R os and G. J. Conduit, *Phys. Rev. E* **99**, 063312 (2019).
- K. Nakano, T. Morresi, M. Casula, R. Maezono, and S. Sorella, *Phys. Rev. B* **103**, L121110 (2021).
- C. Filippi, R. Assaraf, and S. Moroni, *J. Chem. Phys.* **144**, 194105 (2016).
- E. Sloatman, I. Poltavsky, R. Shinde, J. Cocomello, S. Moroni, A. Tkatchenko, and C. Filippi, *J. Chem. Theory Comput.* **20**, 6020 (2024).
- Y. Qian, W. Fu, W. Ren, and J. Chen, *J. Chem. Phys.* **157**, 164104 (2022).
- Y. Qian, X. Li, and J. Chen, *Faraday Discuss.* **254**, 529 (2024).
- J. Lai, B. Kan, Y. Wu, Q. Fu, H. Shang, Z. Li, and J. Yang, *J. Chem. Theory Comput.* **20**, 9478 (2024).
- A. Zen, Y. Luo, S. Sorella, and L. Guidoni, *J. Chem. Theory Comput.* **9**, 4332 (2013).
- Y. Luo, A. Zen, and S. Sorella, *J. Chem. Phys.* **141**, 194112 (2014).
- Y. F. Liu, B. Andrews, and G. J. Conduit, *J. Chem. Phys.* **150**, 034104 (2019).
- J. Tiihonen, R. C. Clay III, and J. T. Krogel, *J. Chem. Phys.* **154**, 204111 (2021).
- K. Nakano, A. Raghav, and S. Sorella, *J. Chem. Phys.* **156**, 034101 (2022).
- F. Mouhat, S. Sorella, R. Vuilleumier, A. M. Saitta, and M. Casula, *J. Chem. Theory Comput.* **13**, 2400 (2017).
- F. Mouhat, M. Peria, T. Morresi, R. Vuilleumier, A. M. Saitta, and M. Casula, *Nat. Commun.* **14**, 6930 (2023).
- G. Mazzola, R. Helled, and S. Sorella, *Phys. Rev. Lett.* **120**, 025701 (2018).
- A. Tirelli, G. Tenti, K. Nakano, and S. Sorella, *Phys. Rev. B* **106**, L041105 (2022).
- K. K. Ly and D. M. Ceperley, *J. Chem. Phys.* **156**, 044108 (2022).
- H. Niu, Y. Yang, S. Jensen, M. Holzmann, C. Pierleoni, and D. M. Ceperley, *Phys. Rev. Lett.* **130**, 076102 (2023).
- D. M. Ceperley, S. Jensen, Y. Yang, H. Niu, C. Pierleoni, and M. Holzmann, *Electron. Struct.* **6**, 015011 (2024).
- G. Tenti, K. Nakano, and M. Casula, *J. Chem. Theory Comput.* **21**, 9335 (2025).
- K. K. Ly and D. M. Ceperley, *Phys. Rev. B* **111**, 104102 (2025).
- S. Goswami, S. Jensen, Y. Yang, M. Holzmann, C. Pierleoni, and D. M. Ceperley, *J. Chem. Phys.* **162**, 054118 (2025).
- K. Nakano, M. Casula, and G. Tenti, *Phys. Rev. B* **109**, 205151 (2024).
- G. Tenti, K. Nakano, A. Tirelli, S. Sorella, and M. Casula, *Phys. Rev. B* **110**, L041107 (2024).

- <sup>44</sup>T. Helgaker and P. Jørgensen, *Theor. Chim. Acta* **75**, 111 (1989).
- <sup>45</sup>F. Jensen, "Lagrangian techniques," in *Introduction to Computational Chemistry*, 3rd ed. (John Wiley & Sons, 2017), Chap. 11.5.
- <sup>46</sup>A. S. Christensen and O. A. von Lilienfeld, *Mach. Learn.: Sci. Technol.* **1**, 045018 (2020).
- <sup>47</sup>A. S. Christensen and O. A. von Lilienfeld (2020). "rMD17," Figshare. <https://doi.org/10.6084/m9.figshare.12672038>
- <sup>48</sup>K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- <sup>49</sup>T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter, *J. Chem. Phys.* **152**, 194103 (2020).
- <sup>50</sup>K. Nakano, C. Attaccalite, M. Barborini, L. Capriotti, M. Casula, E. Coccia, M. Dagrada, C. Genovese, Y. Luo, G. Mazzola, A. Zen, and S. Sorella, *J. Chem. Phys.* **152**, 204121 (2020).
- <sup>51</sup>F. Belleflamme and J. Hutter, *Phys. Chem. Chem. Phys.* **25**, 20817 (2023).
- <sup>52</sup>E. Posenitskiy, V. G. Chilkuri, A. Ammar, M. Hapka, K. Pernal, R. Shinde, E. J. Landinez Borda, C. Filippi, K. Nakano, O. Kohulák, S. Sorella, P. de Oliveira Castro, W. Jalby, P. L. Ríos, A. Alavi, and A. Scemama, *J. Chem. Phys.* **158**, 174801 (2023).
- <sup>53</sup>M. C. Bennett, G. Wang, A. Annaberdiyev, C. A. Melton, L. Shulenburg, and L. Mitas, *J. Chem. Phys.* **149**, 104108 (2018).
- <sup>54</sup>M. C. Bennett, C. A. Melton, A. Annaberdiyev, G. Wang, L. Shulenburg, and L. Mitas, *J. Chem. Phys.* **147**, 224106 (2017).
- <sup>55</sup>S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- <sup>56</sup>J. VandeVondele and J. Hutter, *J. Chem. Phys.* **127**, 114105 (2007).
- <sup>57</sup>F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017).
- <sup>58</sup>S. Chmiela, H. E. Sauceda, K.-R. Müller, and A. Tkatchenko, *Nat. Commun.* **9**, 3887 (2018).
- <sup>59</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>60</sup>C. Møller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- <sup>61</sup>G. D. Purvis III and R. J. Bartlett, *J. Chem. Phys.* **76**, 1910 (1982).
- <sup>62</sup>F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- <sup>63</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
- <sup>64</sup>S. Grimme, S. Ehrlich, and L. Goerigk, *J. Comput. Chem.* **32**, 1456 (2011).
- <sup>65</sup>S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- <sup>66</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- <sup>67</sup>C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- <sup>68</sup>N. Mardirossian and M. Head-Gordon, *Phys. Chem. Chem. Phys.* **16**, 9904 (2014).
- <sup>69</sup>A. Najibi and L. Goerigk, *J. Chem. Theory Comput.* **14**, 5725 (2018).
- <sup>70</sup>N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.* **144**, 214110 (2016).
- <sup>71</sup>D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish *et al.*, *J. Chem. Phys.* **152**, 184108 (2020).
- <sup>72</sup>H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 242 (2012).
- <sup>73</sup>H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, T. F. Miller III, A. Mitrushchenkov, K. A. Peterson, I. Polyak, G. Rauhut, and M. Sibaev, *J. Chem. Phys.* **152**, 144107 (2020).
- <sup>74</sup>H.-J. Werner, P. J. Knowles, P. Celani, W. Györfy, A. Hesselmann, D. Kats, G. Knizia, A. Köhn, T. Korona, D. Kreplin, R. Lindh, Q. Ma, F. R. Manby, A. Mitrushchenkov, G. Rauhut, M. Schütz, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardtsson, A. Berning, J. A. Black, P. J. Bygrave, R. Cimraglia, D. L. Cooper, D. Coughtrie, M. J. O'Neill, P. Palmieri, D. Peng, K. A. Peterson, K. Pflüger, F. Eckert, S. Erfort, E. Goll, C. Hampel, G. Hetzer, J. G. Hill, M. Hodges, T. Hrenar, G. Jansen, C. Köppl, C. Kollmar, S. J. R. Lee, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, B. Mussard, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. A. Peterson, K. Pflüger, R. Pitzer, I. Polyak, M. Reiher, J. O. Richardson, J. B. Robinson, B. Schröder, M. Schwilk, T. Shiozaki, M. Sibaev, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, J. Toulouse, M. Wang, M. Welborn, and B. Ziegler, Molpro, 2025.3, a package of *ab initio* programs.
- <sup>75</sup>Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1340 (2018).
- <sup>76</sup>Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, *J. Chem. Phys.* **153**, 024109 (2020).
- <sup>77</sup>A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.* **102**, 073005 (2009).
- <sup>78</sup>A. Tkatchenko, R. A. DiStasio, R. Car, and M. Scheffler, *Phys. Rev. Lett.* **108**, 236402 (2012).
- <sup>79</sup>P. Eastman, P. K. Behara, D. L. Dotson, R. Galvelis, J. E. Herr, J. T. Horton, Y. Mao, J. D. Chodera, B. P. Pritchard, Y. Wang, G. De Fabritiis, and T. E. Markland, *Sci. Data* **10**, 11 (2023).
- <sup>80</sup>S. Sorella, *Phys. Rev. Lett.* **80**, 4558 (1998).
- <sup>81</sup>K. P. Esler, J. Kim, D. M. Ceperley, W. Purwanto, E. J. Walter, H. Krakauer, S. Zhang, P. R. C. Kent, R. G. Hennig, C. Umrigar, M. Bajdich, J. Kolorenč, L. Mitas, and A. Srinivasan, *J. Phys.: Conf. Ser.* **125**, 012057 (2008).
- <sup>82</sup>M. Mörchen, L. Freitag, and M. Reiher, *J. Chem. Phys.* **153**, 244113 (2020).
- <sup>83</sup>K. Momma and F. Izumi, *J. Appl. Crystallogr.* **44**, 1272 (2011).
- <sup>84</sup>S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).
- <sup>85</sup>M. Casula and S. Sorella, *J. Chem. Phys.* **119**, 6500 (2003).
- <sup>86</sup>Any square matrix can be uniquely decomposed into a sum of symmetric and skew-symmetric matrices.
- <sup>87</sup>S. Sorella, *Phys. Rev. B* **71**, 241103(R) (2005).
- <sup>88</sup>C. J. Umrigar and C. Filippi, *Phys. Rev. Lett.* **94**, 150201 (2005).