

RESEARCH ARTICLE | SEPTEMBER 22 2026

Improving the accuracy of fixed-node diffusion Monte Carlo for non-covalent interactions

Special Collection: [2026 JCP Emerging Investigators Special Collection](#)

Kousuke Nakano ; Benjamin X. Shi ; Dario Alfè ; Andrea Zen 



J. Chem. Phys. 165, 124102 (2026)

<https://doi.org/10.1063/5.0348824>



View
Online



Export
Citation

Articles You May Be Interested In

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

J. Chem. Phys. (March 2026)

TURBORVB: A many-body toolkit for *ab initio* electronic simulations by quantum Monte Carlo

J. Chem. Phys. (May 2020)

Understanding intermolecular interactions of large systems in ground state and excited state by using density functional based tight binding methods

J. Chem. Phys. (May 2021)



The Journal of Chemical Physics

Special Topics Open for Submissions

[Learn More](#)

Improving the accuracy of fixed-node diffusion Monte Carlo for non-covalent interactions

Cite as: J. Chem. Phys. 165, 124102 (2026); doi: 10.1063/5.0348824

Submitted: 18 June 2026 • Accepted: 14 August 2026 •

Published Online: 22 September 2026



Kousuke Nakano,^{1,a)} Benjamin X. Shi,² Dario Alfè,^{3,4,5} and Andrea Zen^{3,4,b)}

AFFILIATIONS

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

²Initiative for Computational Catalysis, Flatiron Institute, 160 5th Avenue, New York, New York 10010, USA

³Dipartimento di Fisica Ettore Pancini, Università di Napoli Federico II, Monte S. Angelo, I-80126 Napoli, Italy

⁴Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom

⁵Thomas Young Centre and London Centre for Nanotechnology, 17-19 Gordon Street, London WC1H 0AH, United Kingdom

Note: This paper is part of the 2026 JCP Emerging Investigators Special Collection.

^{a)}Author to whom correspondence should be addressed: kousuke_1123@icloud.com

^{b)}andrea.zen@unina.it

ABSTRACT

Diffusion quantum Monte Carlo (DMC) and coupled cluster theory [CCSD(T)] are widely used benchmark methods for noncovalent interactions (NCIs). However, recent studies have reported notable discrepancies for several hydrogen-bonded and dispersion-dominated systems, raising questions about the accuracy of the approximations underlying each approach. In DMC, the dominant residual error is expected to stem from the fixed-node (FN) approximation, where the nodal surface is typically taken from a single Slater determinant (SD) derived from density functional theory or Hartree–Fock calculations. In this work, we assess the impact of nodal-surface optimization on DMC predictions for 12 compounds spanning diverse NCIs. In particular, we compare binding energies obtained with a recently proposed antisymmetrized geminal power Ansatz parameterized with natural orbitals (AGPn) and with the conventional SD ansatz. AGPn is a generalization of the SD Ansatz and, when variationally optimized at the DMC level, yields lower total energies. We find that for hydrogen-bonded systems, AGPn improves the agreement with CCSD(T), whereas its effect on dispersion-dominated systems is negligible within the statistical uncertainty. The obtained mean absolute deviations between CCSD(T) and FN-SD-DMC binding energies are 0.43(2) and 0.21(2) kcal/mol for the hydrogen-bonded and dispersion-dominated systems, respectively, and those between CCSD(T) and FN-AGPn-DMC binding energies are 0.18(2) and 0.28(3) kcal/mol for the hydrogen-bonded and dispersion-dominated systems, respectively. These results suggest that the reported discrepancies between DMC and CCSD(T) for hydrogen bonds originate primarily from the mean-field nodal surface of the SD ansatz, while the origin of the discrepancies for dispersion-dominated systems remains an open question.

© 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/10.1063/5.0348824>

I. INTRODUCTION

The central task of *ab initio* electronic structure calculations is to solve the many-body Schrödinger equation. Several methods have been developed to solve this equation, including wavefunction methods (quantum chemistry calculations), density functional theory (DFT), and quantum Monte Carlo (QMC). On the one hand, coupled cluster theory with single, double, and perturbative triple excitations [CCSD(T)]¹ has been regarded as a reference method

within the wavefunction methods. On the other hand, the fixed-node diffusion quantum Monte Carlo (FN-DMC)² has also been regarded as a reference method, especially in materials science and condensed matter physics. Typically, FN-DMC uses a Jastrow factor as the guiding function, multiplying a single Slater determinant (SD). Thereafter, we call this approach FN-SD-DMC. For small molecular systems, FN-SD-DMC and CCSD(T) typically yield consistent values, e.g., for intermolecular binding energies. However, in 2021, Al-Hamdani *et al.* reported significant discrepancies between

CCSD(T) and FN-SD-DMC calculations when evaluating the binding energies of large molecules bound by van der Waals forces, such as C₆₀ buckyball inside a [6]-cycloparaphenyleneacetylene ring (C₆₀@[6]CPPA).³ Fishman *et al.* reported that such discrepancy is also observed in π -stacking complexes, such as acene and alkadiene dimers.⁴ Furthermore, in 2025, more binding energies were benchmarked on the S66 dataset⁵ of medium-sized dimers. It was reported that discrepancies between CCSD(T) and FN-SD-DMC were observed not only for large molecules bound by van der Waals forces but also for some molecules bound by hydrogen bonds. Since neither CCSD(T) nor FN-SD-DMC is exact, and both rely on approximations, it remains difficult to unambiguously determine which approach is more accurate for NCIs.

On the coupled cluster (CC) side, considerable effort has been devoted to investigating how binding energies change as higher-order perturbative corrections are included^{6–11} or as higher-order screening effects are treated differently,¹² such as by CCSDT, CCSDT(Q), and CCSD(cT).¹³ The details of these studies are discussed later in Sec. III. On the FN-DMC side, by contrast, the main line of investigation has been to assess how the *nodal surface* of the trial wave function affects binding energies. The present work follows this direction by investigating how the FN-DMC binding energies change when the nodal surface of the trial wave function is improved.

Several choices have to be made when setting up FN-SD-DMC simulations, such as deciding how to initialize the Slater determinant; which functional form for the Jastrow factor, basis set,¹⁴ and pseudopotentials to employ (all-electron calculations in FN-DMC are possible, but explicitly treating core electrons entails a severe computational cost, which scales roughly as Z^6 , where Z is the largest atomic number in the system¹⁵); how to deal with the non-local terms in the pseudopotentials;^{16–18} which scheme to adopt to stabilize the simulation;^{19,20} how to eliminate the finite time step and finite-population biases; and how to optimize the variational parameters in the Jastrow factor and control the optimization uncertainty. Reproducibility across different QMC codes is reached when a consistent setup is adopted and the infinitesimal time step limit is carefully assessed.²¹ The choice in SD initialization and the pseudopotential localization affects the total energies of FN-SD-DMC, but energy differences are typically minimally affected by these choices, indicating a strong cancellation of these biases, especially on evaluations of interaction energies on NCIs. The systems for which FN-SD-DMC and CCSD(T) disagree are no exception: FN-SD-DMC simulations adopting different setups have been performed^{3,5,22} and the dependence of the interaction energy on the setup is minimal and is much smaller than the difference with the CCSD(T) evaluations. Thus, if there is a bias on FN-SD-DMC evaluations, it would come from the fixed-node approximation (i.e., the nodal surface).

The fixed-node approximation is the least systematically controllable approximation in FN-DMC, arising from the so-called Fermion sign problem.^{2,23,24} In FN-SD-DMC, the nodes are defined by the single Slater determinant, as the Jastrow factor is positive. Typically, in FN-SD-DMC, the SD is provided by a preceding mean-field calculation, such as Hartree–Fock (HF) or DFT calculations. In this situation, what can be done from the FN-DMC side is to investigate whether the binding energy obtained by using a

better wavefunction—beyond a mean-field nodal surface—remains unchanged, or approaches the CCSD(T) value. This makes it possible to assess whether the discrepancy between CCSD(T) and FN-SD-DMC originates from the mean-field nodal surfaces used in FN-SD-DMC.

Many variational Ansatzes beyond the single Slater-determinant have been proposed. Examples are multideterminant ansatz,^{25–28} including selected-CI/CIPSI expansions,^{29–34} antisymmetrized geminal power (AGP) Ansatz,³⁵ backflow,^{36–38} Pfaffian,^{38–40} and more recently, neural-network-based Ansatzes.^{41–44} They have been applied to binding energy calculations of various molecules.^{22,29,33,34,40,44–49} A pioneering work in the context of the present study is the FN-DMC calculation with the backflow Ansatz (denoted as FN-BF-DMC in this study) for AcOH–AcOH and water-peptide by Lambie *et al.*⁴⁹ The study revealed that the consistency between CCSD(T) and FN-BF-DMC is much better than that between CCSD(T) and FN-SD-DMC; thus, the reported discrepancy between CCSD(T) and FN-SD-DMC is attributed to the mean-field nodal surface used in FN-SD-DMC. However, in any ansatz, the drawback is the inevitable trade-off between flexibility and computational cost. More flexible Ansatzes can yield lower variational energies, but they generally require a higher computational cost. For the evaluation of interaction energies, however, obtaining lower total energies is not by itself sufficient: the Ansatz must describe the complex and the isolated fragments with comparable accuracy, so that size-extensivity and size-consistency are preserved, and the optimization uncertainty must be kept under control, ideally below the statistical error. This increases the computational cost and makes systematic studies of diverse compounds more challenging. In fact, Lambie *et al.*⁴⁹ stated that their FN-BF-DMC calculations used 2.9×10^6 and 1.5×10^6 core-hours for AcOH–AcOH and water-peptide on an AMD EPYC cluster machine, respectively, while their FN-SD-DMC counterparts cost 0.42 and 0.27×10^6 core-hours, respectively. This may limit applications of larger compounds.

We recently developed a DMC workflow along this line based on AGP Ansatz with natural orbitals (denoted as FN-AGPn-DMC in this study),⁵⁰ where only a relatively small number of pairing coefficients of the geminal pairing function are optimized. This is obtained by writing the pairing function in terms of natural orbitals constructed from a beyond HF/DFT calculation and keeping them frozen in the optimization while optimizing their weights. Since AGPn is a generalization of a single Slater determinant, it provides a variationally more flexible Ansatz than the SD form and can, therefore, improve the fixed-node energy when properly optimized. Moreover, the computational cost of FN-AGPn-DMC has the same scaling as that of FN-SD-DMC.⁵⁰ This enables one to study large systems with better nodal surfaces than mean-field ones at an affordable computational cost.

This paper reports DMC binding-energy calculations using the FN-AGPn-DMC approach for a diverse set of non-covalent interaction (NCI) compounds, together with comparisons against FN-SD-DMC and CCSD(T) evaluations. We selected 12 NCI compounds for which discrepancies between FN-SD-DMC and CCSD(T) have been reported. The 12 compounds are taken from the S22,⁵¹ S66,⁵² and A24⁵³ datasets and are divided into two primary categories: hydrogen-bonded and dispersion-dominated systems (π -stacking,

pure London dispersion, and mixed interactions). Our calculations show that for hydrogen-bonded systems, FN-AGPn-DMC yields binding energies closer to CCSD(T) than FN-SD-DMC does. For dispersion-dominated systems, FN-AGPn-DMC instead yields binding energies consistent with those obtained with FN-SD-DMC despite improving the wavefunction and systematically lowering the total energies. Thus, our study suggests that the discrepancies between FN-SD-DMC and CCSD(T) for hydrogen-bonded systems originate primarily from inaccuracies in FN-SD-DMC, in particular from the nodal surface of a single Slater determinant. The same conclusion cannot be straightforwardly generalized to dispersion-dominated systems, for which FN-SD-DMC and FN-AGPn-DMC interaction energies agree within the statistical uncertainty. Therefore, the origin of the discrepancies between DMC and CCSD(T) for dispersion-dominated systems remains an open question.

II. COMPUTATIONAL METHOD

In this work, we employed our recently developed method for variational optimization of the nodal surface with the FN-DMC framework. Typically, in DMC methods, the nodal surface is deterministically given by the single Slater determinant constructed from a mean-field calculation, such as HF or DFT, and kept frozen at the DMC level. We denote it FN-SD-DMC in this work.

Our developed scheme⁵⁰ aims to obtain a better wavefunction by variational optimization of the nodal surface at the DMC level. The choice is the most prominent feature of our scheme. To balance the accuracy and feasibility, we used an antisymmetrized geminal power (AGP) wavefunction parameterized with natural orbitals constructed from an MP2 calculation, denoted as AGPn ansatz. The details are described in Ref. 50. We briefly recall why the AGP Ansatz can be regarded as a generalization of the SD ansatz.

The AGPn Ansatz can be understood as a compact multideterminant extension of the single Slater determinant. For a closed-shell system, the AGP part is written as

$$|\Phi_{\text{AGPn}}\rangle \propto \left(\sum_{p=1}^n \lambda_p a_{p\uparrow}^\dagger a_{p\downarrow}^\dagger \right)^{N/2} |0\rangle, \quad (1)$$

where $a_{p\sigma}^\dagger$ creates an electron with spin σ in the p th MP2 natural orbital, λ_p is the corresponding pairing coefficient, and n is the number of natural orbitals included in the pairing function. The reference single Slater determinant is obtained by doubly occupying the occupied natural orbitals,

$$|\Phi_{\text{SD}}\rangle = \prod_i^{\text{occ}} a_{i\uparrow}^\dagger a_{i\downarrow}^\dagger |0\rangle. \quad (2)$$

By expanding Eq. (1) around this reference determinant, the AGPn wavefunction can be expressed as

$$|\Phi_{\text{AGPn}}\rangle \propto \left[1 + \sum_i^{\text{occ}} \sum_a^{\text{virt}} \frac{\lambda_a}{\lambda_i} P_a^\dagger P_i + \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{virt}} \frac{\lambda_a \lambda_b}{\lambda_i \lambda_j} P_a^\dagger P_b^\dagger P_j P_i + \dots \right] |\Phi_{\text{SD}}\rangle, \quad (3)$$

where $P_p^\dagger = a_{p\uparrow}^\dagger a_{p\downarrow}^\dagger$ and $P_p = a_{p\downarrow} a_{p\uparrow}$. Equation (3) shows that the AGPn ansatz contains the SD determinant as the leading term,

followed by pair excitations in which an occupied natural orbital is emptied and a virtual natural orbital is doubly occupied. Therefore, the SD Ansatz is recovered when the pairing coefficients associated with the virtual natural orbitals are set to zero, whereas AGPn systematically extends the nodal surface by including selected occupied-to-virtual pair excitations.

In the workflow, only a restricted subset of the natural-orbital pairing coefficients was optimized, thereby reducing the number of variational parameters. Regarding how many weights to optimize, we targeted natural orbitals with MP2 calculated occupation ratios >0.001 for optimization (this is a hyperparameter). We confirmed that the threshold of 0.001 yields a converged FN-AGPn-DMC energy relative to the occupation threshold (i.e., the number of considered virtual orbitals) for the water dimer.

Figure 1(a) shows the relation between the final FN-AGPn-DMC energy with respect to the occupation threshold chosen for variational weight parameters in the pairing function in the water dimer. We tested 0.06, 0.02, 0.01, 0.001, 0.0001, and 0.00001 (the corresponding numbers of considered virtual orbitals are 5, 8, 18, 38, 85, and 139, respectively) and also repeated FN-DMC optimization three times with different seeds and hyperparameters. The figure shows that the chosen parameter 0.001 yields a converged energy, and the optimization uncertainty is within the statistical errors. Figure 1(b) schematically illustrates how the AGPn Ansatz improves the nodal surface relative to the SD Ansatz by mixing occupied and virtual MP2 natural orbitals.

The AGPn was combined with a conventional Jastrow factor. The Jastrow function in this study is composed of an inhomogeneous one-body, exponential-type two-body, and atomic orbital-based three-body terms.⁵⁴ The Jastrow factor was optimized only with VMC gradients and it was held fixed during optimization with (approximate) FN gradients.⁵⁵ The weights for the natural orbitals were optimized with the FN gradients. The parameters were optimized using the stochastic reconfiguration method⁵⁶ implemented in TurboRVB⁵⁴ with an adaptive learning rate.⁴⁶

We selected 12 compounds as the benchmark set for this study, taken from the S22, S66, and A24 datasets. They are the acetic acid dimer (AcOH–AcOH, No. 20 in S66), the formic acid dimer (HCOOH–HCOOH, No. 3 in S22), the dimer of water and peptide (Water–Peptide, No. 4 in S66), the dimer of methanol and pyridine (MeOH–pyridine, No. 19 in S66), the water dimer (water–water, No. 1 in S66), the methylamine dimer (MeNH₂–MeNH₂, No. 10 in S66), the dimer of peptide and pentane (peptide–pentane, No. 46 in S66), the dimer of uracil and cyclopentane (uracil–cyclopentane, No. 42 in S66), the dimer of pentane and acetamide (pentane–AcNH₂, No. 62 in S66), the benzene dimer (benzene–benzene, No. 24 in S66), the dimer of ethene and ethyne (ethene–ethyne, No. 22 in A24), and the ethyne dimer (ethyne–ethyne, No. 24 in A24). The compounds are divided into four categories: hydrogen bonds (water–water, HCOOH–HCOOH, AcOH–AcOH, MeNH₂–MeNH₂, water–peptide, and MeOH–pyridine), π -stacking (benzene–benzene), pure London dispersion (uracil–cyclopentane, ethene–ethyne, ethyne–ethyne, and peptide–pentane), and mixed-interactions (pentane–AcNH₂). In the main text, we categorize them into two primary groups, hydrogen-bonds and dispersion-dominated compounds (π -stacking, pure London dispersion, and mixed-interactions). The target compounds are all dimers

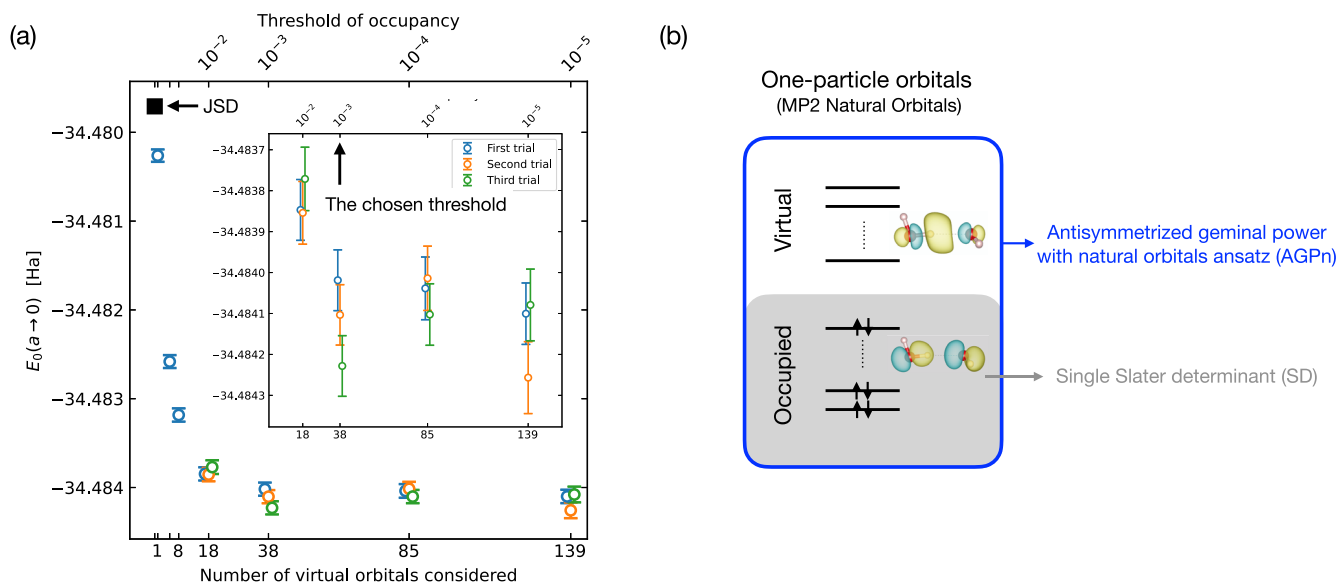


FIG. 1. (a) FN-AGPn-DMC total energies of the water–water dimer, obtained using different occupation thresholds for choosing the number of optimized variational weight parameters. The tested occupation thresholds, 0.06, 0.02, 0.01, 0.001 (used for production runs), 0.0001, and 0.000 01, correspond to 5, 8, 18, 38, 85, and 139 virtual orbitals, respectively. The FN-SD-DMC energy is also plotted as a reference. The inset in panel (a): The optimizations were repeated three times with different seeds. The error bars represent 1σ . (b) A schematic figure representing how the SD and AGPn Ansatzes are constructed from one-body orbitals.

composed of two fragments. For each compound, we calculated either three or four configurations: Fragment 1, Fragment 2 (if fragments are not identical), Complex, and Far-away complex (where the two fragments are sufficiently separated; used for checking the size-consistency).

For DMC calculations, we employed the correlation consistent effective core potential (ECP)⁵⁷ and accompanying aug-cc-pVTZ basis sets.⁵⁷ Our recent study¹⁴ has revealed that even in FN-DMC calculations, the choice of basis function matters for binding energies of the 0.1 kcal/mol order due to the basis set incompleteness error (BSIE). The aug-cc-pVTZ basis set sufficiently reduces the BSIE.¹⁴ For DFT and MP2 calculations for the trial wave function generations, we employed PySCF.^{58–60} The LDA-PZ functional was employed for the baseline DFT calculations, as it produces good orbitals for FN-SD-DMC, typically better than orbitals from HF or from other DFT exchange–correlation (XC) functionals.^{61–64} However, we note that the choice of XC typically has little or negligible effect on the binding energies, especially for NCIs, because of the FN-bias cancellation when considering energy differences. This is observed for the acetic acid dimer in Ref. 5, for benzene dimer and other dimers in Ref. 3, and in several other systems in Refs. 61, 63, and 65–69.

We exported the TREFIO⁷⁰ files from PySCF and converted them to the TurboRVB wavefunction format using the converter implemented in TurboGenius.⁷¹ Subsequently, we performed the aforementioned FN-DMC calculations. For DMC calculations, we utilized the lattice regularized diffusion Monte Carlo method (LRDMC)⁷² implemented in TurboRVB. LRDMC realizes the projection as standard DMC does, and it has been confirmed that controlling all controllable errors yields identical results.^{21,73} In

LRDMC calculations, the controllable error is the real-space grid, a . This error can be eliminated by extrapolating $a \rightarrow 0$. In this study, we adopted values of $a = 0.1, 0.15, 0.2$, and 0.3 Bohr and fitted the total energy using the quadratic function $E(a) = E_0 + k_1 a^2 + k_2 a^4$. For the total energy, we report the extrapolated value, $E(a \rightarrow 0)$. For binding energies,

$$E_b = E(\text{Complex}) - E(\text{Monomer 1}) - E(\text{Monomer 2})$$

is computed. We found that the dependence of E_b on a is flat within the error bars for $a \leq 0.2$ for all compounds. Therefore, the value at $a = 0.2$ is reported. For size-consistency calculations,

$$E_s = E(\text{far-away}) - E(\text{Monomer 1}) - E(\text{Monomer 2})$$

is computed. We found that the dependence of E_s on a is also flat within the error bars for $a \leq 0.2$ for all compounds. Therefore, the value at $a = 0.2$ is reported. In this study, since ECPs are used, the treatment of the so-called non-local term is crucial to satisfy the variational principle in DMC calculations. Regarding this point, LRDMC corresponds to the implementation of the T-move method in standard DMC,^{17,72} thus ensuring the variational principle. Furthermore, we combined this with the Determinant Locality Approach (DLA).¹⁸ This method with the SD ansatz, when starting from the same DFT calculation, yields the same result in the limit $a \rightarrow 0$, i.e., independent of the Jastrow factor parameterization (referred to as the DTM method²¹). All calculations in this study were automated and performed using TurboWorkflows.⁷¹

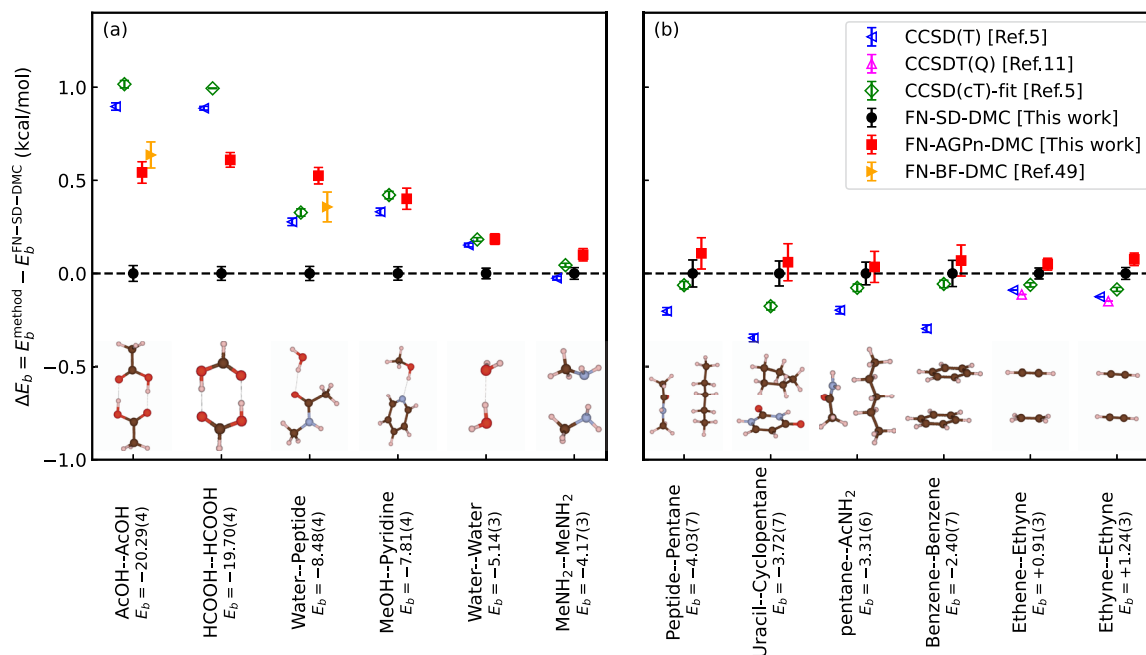


FIG. 2. Difference in binding energies computed by FN-SD-DMC, FN-AGPn-DMC, FN-BF-DMC, CCSD(T), CCSDT(Q), and CCSD(cT)-fit for (a) hydrogen bonds and (b) dispersion-dominated bonds. The CCSD(T), CCSDT(Q), CCSD(cT)-fit (except for HCOOH–HCOOH, ethene–ethyne, and ethyne–ethyne, which were computed in this work using the same setting in Ref. 5), and FN-BF-DMC values are taken from Refs. 5, 11, and 49, respectively. E_b in the x labels refers to the binding energies (kcal/mol) obtained by FN-SD-DMC in this work. The error bars represent 1σ .

TABLE I. Binding energies (kcal/mol). The numbers in the parentheses represent 1σ error bars. $\Delta E_b^x \equiv E_b^x - E_b^{\text{FN-SD-DMC}}$, where $x \in \{\text{FN-AGPn-DMC}, \text{CCSD(T)}\}$.

System	FN-SD-DMC	FN-AGPn-DMC	CCSD(T)	$\Delta E_b^{\text{FN-AGPn-DMC}}$	$\Delta E_b^{\text{CCSD(T)}}$
AcOH–AcOH	–20.29(4)	–19.74(6)	–19.39(2) ⁵	0.54(7)	0.90(5)
HCOOH–HCOOH	–19.70(4)	–19.09(4)	–18.81(1) ⁷⁴	0.61(5)	0.89(4)
Water–peptide	–8.48(4)	–7.95(4)	–8.20(2) ⁵	0.52(6)	0.28(4)
MeOH–pyridine	–7.81(4)	–7.41(6)	–7.48(2) ⁵	0.40(7)	0.33(4)
Water–water	–5.14(3)	–4.96(3)	–4.99(1) ⁵	0.18(4)	0.15(3)
MeNH ₂ –MeNH ₂	–4.17(3)	–4.07(3)	–4.20(1) ⁵	0.10(5)	–0.03(3)
Peptide–pentane	–4.03(7)	–3.92(8)	–4.23(2) ⁵	0.1(1)	–0.20(8)
Uracil–cyclopentane	–3.72(7)	–3.7(1)	–4.07(2) ⁵	0.1(1)	–0.35(7)
Pentane–AcNH ₂	–3.31(6)	–3.28(8)	–3.51(2) ⁵	0.0(1)	–0.20(6)
Benzene–benzene	–2.40(7)	–2.33(8)	–2.70(2) ⁵	0.1(1)	–0.30(7)
Ethene–ethyne	0.91(3)	0.96(3)	0.821 ⁵³	0.05(4)	–0.09(3)
Ethyne–ethyne	1.24(3)	1.32(3)	1.115 ⁵³	0.08(5)	–0.12(3)

III. RESULTS AND DISCUSSION

The key result of this work is summarized in Fig. 2 and Table I. Figure 2 shows the deviation of the binding energy obtained by FN-AGPn-DMC from those obtained by FN-SD-DMC and Figs. 2(a) and 2(b) show those for the hydrogen and dispersion bonds, respectively. Table I summarizes the binding energy values corresponding to Fig. 2. For comparison, Fig. 2 also plots the deviations of the binding energies obtained by using other methods, CCSD(T), CCSDT(Q), CCSD(cT)-fit, and FN-BF-DMC, which will be discussed later. Figure 3 summarizes the mean absolute deviation

(MAD) of binding energies obtained by CCSD(T), CCSD(cT)-fit, FN-SD-DMC, and FN-AGPn-DMC. On the one hand, it is clear from Fig. 2(a) that, for binding energies for hydrogen bonds, FN-AGPn-DMC weakens the binding energies, resulting in values closer to those of CCSD(T), specifically for formic acid dimer and acetic acid dimer. While the correction overshoots for water–peptide, the consistency between CCSD(T) and FN-AGPn-DMC (0.18(2) kcal/mol in MAD) is better than CCSD(T) and FN-SD-DMC (0.43(2) kcal/mol in MAD), as shown in Fig. 3. The result implies that for hydrogen bonds, the reported deviation between FN-SD-DMC and CCSD(T) has arisen from the nodal surfaces obtained

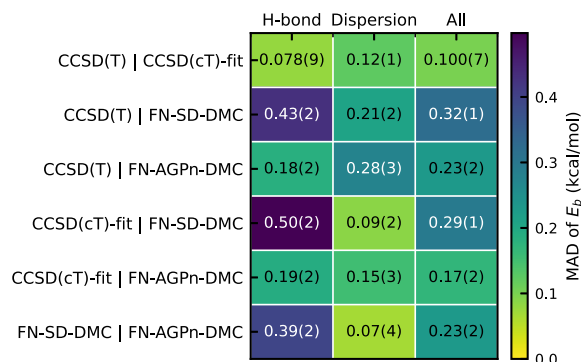


FIG. 3. Pairwise MADs of binding energies obtained by CCSD(T), CCSD(cT)-fit, FN-SD-DMC, and FN-AGPn-DMC for hydrogen bonds (H-bond), dispersion-dominated bonds (dispersion), and all of them (All). The unit is kcal/mol.

from a mean-field approach. The implication agrees with the recent work⁴⁹ for AcOH-AcOH and water-peptide. On the other hand, as shown in Fig. 2(b), for the dispersion-dominated bonds, FN-AGPn-DMC shows negligible effect on binding energy with respect to FN-SD-DMC. This suggests that for these systems, the nodal-surface improvement achieved by the present AGPn Ansatz largely cancels between the complex and the fragments, leaving the interaction energies essentially unchanged within the statistical uncertainty. Our calculation implies that the reported discrepancies between FN-DMC and CCSD(T) for those bonds might not be attributable to a mean-field nodal surface in FN-DMC, suggesting that other factors could be involved. This is discussed later.

From a practical perspective, it is also important to examine the computational cost of the FN-AGPn-DMC workflows. As described in Ref. 50, the number of parameters optimized in the FN-AGPn-DMC workflow is limited to the weights associated with the pairing function; thus, it linearly scales with the system size and remains on the order of 10^2 in practice. For example, even for uracil-cyclopentane (dimer), which is the compound with the largest number of valence electrons in this study, the number of optimized parameters is only 183. Furthermore, AGPn itself is a single determinant of the pairing function. Consequently, the computational cost of FN-AGPn-DMC is not significantly different from that of FN-SD-DMC. For example, for uracil-cyclopentane (dimer), the computational cost of FN-AGPn-DMC was about two times that of FN-SD-DMC when achieving the same error bar in total energy (c.f., the computational cost of FN-BF-DMC is ~ 5 – 7 times that of FN-SD-DMC⁴⁹). Our FN-SD-DMC calculations consume about 1.2×10^6 core-hours on the Fugaku supercomputer to achieve a 1σ error bar of 0.07 kcal/mol in the binding energy of uracil-cyclopentane, whereas the FN-AGPn-DMC counterparts consume about 1.1×10^6 core-hours to achieve a 1σ error bar of 0.1 kcal/mol.

Several checks were performed to verify the validity of our calculations; the results are summarized in the Appendix. The total energies for all monomers and complexes obtained by FN-AGPn-DMC are lower than those obtained by FN-SD-DMC. The

variational principle guarantees that our FN-AGPn-DMC calculations give a better nodal surface than the FN-SD-DMC calculations do. We also confirmed that the underbindings by FN-AGPn-DMC are not due to deterioration in size consistency. CCSD(T) is size-consistent,^{1,75} whereas, in QMC, wavefunction optimization can degrade size-consistency,^{46,76} depending on the Ansatz choice. Improving the wavefunctions does not necessarily lead to better binding energies because the wavefunctions for the dimer may differ from those for the fragments, leading to artificial under- or over-binding. In this work, we verified that the sum of the energies of the fragments constituting the complex is consistent with the total energy of a far-away complex within the error bars, where the two fragments are sufficiently separated. For all complexes in this study, size consistency is satisfied within the error bars, suggesting that the underbinding is intrinsic.

Now, let us discuss the origins of the discrepancy between FN-SD-DMC and CCSD(T) binding energies for the non-covalent bonds. We claim that, for the hydrogen bonds, the reported discrepancies are likely due to the employed DFT nodal surface because FN-AGPn-DMC shows significant change in the binding energies and gives binding energies closer to the CCSD(T) values, especially for the hydrogen-bonding systems of formic acid dimer and acetic acid dimer. Our claim is consistent with many recent studies. For instance, Lambie *et al.* reported the binding energies of acetic acid dimer and water-peptide⁴⁹ by FN-BF-DMC. They revealed that the obtained binding energies are underbound with respect to FN-SD-DMC values. Their FN-BF-DMC binding energies are consistent with our FN-AGPn-DMC ones within the error bars, as shown in Fig. 2. Semidalas *et al.*¹⁰ found that, for hydrogen bonds, CCSD(T) gives consistent binding energies with more accurate CCSDT and CCSDT(Q) calculations. They also argue that the discrepancies reported for the hydrogen-bond systems between FN-SD-DMC and CCSD(T) are very hard to rationalize with remaining flaws of CCSD(T).¹⁰ The present results are consistent with this interpretation.

The remaining discrepancies (~ 0.25 kcal/mol) between CCSD(T) and FN-AGPn-DMC may still reflect residual fixed-node errors beyond the flexibility of the present AGPn ansatz. Indeed, after the controllable DMC biases have been removed, further improvements from the DMC side can only come from improving the nodal surface. The present workflow achieves this within a constrained AGPn parameterization, in which the orbitals entering the pairing function are fixed to MP2 natural orbitals and only a restricted subset of pairing coefficients is optimized. A natural next step is, therefore, to relax this constraint by adopting more flexible nodal parameterizations, such as multi-determinant trial wave functions or neural-network-based Ansatzes.^{41–44} For interaction energies, however, this increased flexibility must be balanced against the need to preserve size consistency and reliable error cancellation between the complex and its fragments.

The situation for the dispersion-dominated bonds is different. Our calculations reveal that FN-AGPn-DMC has negligible effect on the binding energy for those bonds compared with FN-SD-DMC; thus, the discrepancy between FN-SD-DMC and CCSD(T) is not explained by the DFT nodal surface, and other sources should contribute to it. The origin of the discrepancy in those bonds is under intensive debate in the community. One of the pioneering studies

was done by Schäfer *et al.* reporting changes in binding energies for nine π -stacking compounds using CCSD(cT).¹² They demonstrated that for the nine compounds, CCSD(cT) restores excellent agreement for interaction energies with FN-SD-DMC ones. Based on the way to estimate CCSD(cT) binding energies proposed by Schäfer *et al.* [denoted as CCSD(cT)-fit], Shi *et al.*⁵ reported the CCSD(cT)-fit binding energies for the S66 compounds. The trend is the same as Schäfer's work. Indeed, for π stacking and other dispersion-dominated compounds, CCSD(cT)-fit gives different binding energies than CCSD(T); thus, CCSD(cT) and CCSD(cT)-fit binding energies are close to FN-SD-DMC ones. The reported and computed CCSD(cT)-fit values are plotted in Fig. 2. Very interestingly, for the tested dispersion bonds, CCSD(cT)-fit gives binding energies more consistent with our FN-SD-DMC and FN-AGPn-DMC ones, implying that only CCSD(T) gives overbindings for those molecules. However, we note that although CCSD(cT) has recently been proposed as an improved alternative to CCSD(T) for non-covalent binding energies, its use as a reference remains under intensive debate because CCSD(cT) is an approximation to CCSDT, and CCSDT is not always accurate.^{6–10} Our work reveals that the discrepancies in the binding energies of the dispersion-dominated interactions might not be explained by the DFT nodal surface used in FN-SD-DMC calculations, but the origin of the discrepancies remains an open question.

IV. CONCLUDING REMARKS

In conclusion, we computed the binding energies of 12 compounds with non-covalent bonds, for which the binding energies are reportedly inconsistent between CCSD(T) and FN-SD-DMC calculations, by the FN-DMC with nodal surfaces given by the antisymmetrized geminal power Ansatz with natural orbitals, denoted as FN-AGPn-DMC. The AGPn Ansatz yields a better nodal surface than the conventional single Slater determinant Ansatz (SD); thus, FN-AGPn-DMC gives more accurate energies than FN-SD-DMC does, allowing one to discuss whether the discrepancy arises from poor nodal surfaces in FN-SD-DMC. On the one hand, our calculations reveal that for hydrogen bonds, FN-AGPn-DMC yields weaker binding than FN-SD-DMC, leading to greater consistency with CCSD(T). It indicates that the discrepancy previously reported in the binding energy calculations between CCSD(T) and FN-SD-DMC is due to the nodal surface given by a mean-field approach, such as density functional theory (DFT) calculation. On the other hand, our results show that for dispersion-dominated bonds, FN-AGPn-DMC yields binding energies consistent with those of FN-SD-DMC, thereby not improving the consistency with CCSD(T). The origin of the discrepancies for dispersion-dominated bonds remains an open question. One possible source is the limited nodal surface flexibility of the constrained AGPn parameterization employed in the present study, in which the orbitals in the pairing function are fixed to MP2 natural orbitals and only a restricted subset of the pairing coefficients is optimized. A natural next step is, therefore, to go beyond this constrained parameterization and explore more flexible nodal parameterizations, such as multi-determinant trial wave functions and neural-network-based Ansatzes. Further investigation on the CC side would also be important. As discussed in this work, for the dispersion-dominated systems studied here, different perturbative corrections yield noticeably

different binding energies, implying that reference CC values are not yet established for many NCIs. Complementary progress on both the DMC and CC sides will, therefore, be essential for clarifying the origin of the remaining discrepancies.

ACKNOWLEDGMENTS

K.N. is grateful for computational resources from the Numerical Materials Simulator at the National Institute for Materials Science (NIMS). K.N. is grateful for computational resources from the supercomputer Fugaku provided by RIKEN through the HPCI System Research Projects (Project IDs: hp240033 and hp250031). K.N. acknowledges financial support from MEXT Leading Initiative for Excellent Young Researchers (Grant No. JPMXS0320220025), from Iketani Science and Technology Foundation (Grant No. 0361248-A), and from JSPS KAKENHI [Grant-in-Aid for Scientific Research(C), Grant No. 25K00216]. D.A. and A.Z. acknowledge support from Leverhulme Trust (Grant No. RPG-2020-038). D.A. and A.Z. also acknowledge support from the European Union under the Next generation EU (Project Nos. 20222FXZ33 and P2022MC742). The authors acknowledge the use of the ARCHER UK National Supercomputing Service (<https://www.archer2.ac.uk>). The molecular structures were depicted using VESTA.⁷⁷ The Flatiron Institute is a division of the Simons Foundation. The authors thank Professor Martin Suhm for his suggestion of the formic acid dimer as an interesting application of our FN-AGPn-DMC approach and Dr. Pablo López Ríos (Max Planck Institute for Solid State Research) for his suggestion of the convergence and optimization uncertainty checks.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Kousuke Nakano: Conceptualization (equal); Data curation (lead); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (lead); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). **Benjamin X. Shi:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (equal); Validation (equal); Writing – review & editing (equal). **Dario Alfè:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Andrea Zen:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (equal); Methodology (equal); Project administration (lead); Supervision (lead); Visualization (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The QMC kernel used in this work, TurboRVB, is available from its GitHub repository, (<https://github.com/sissaschool/turborvb>). TurboRVB is supported by “Project for advancement of

software usability in materials science”⁷⁸ in FY2025 by the Institute for Solid State Physics (ISSP) at the University of Tokyo. The geometries, wavefunctions, pseudo-potentials, other input, and output files are available in NIMS Materials Data Repository (MDR) [<https://doi.org/10.48505/nims.6500>]. The geometries can also be obtained from the Benchmark Energy and Geometry Database (BEGDB).⁷⁹

APPENDIX: SUPPORTING DATA

Table II presents the results of the LRDMC calculations using the frozen DFT orbitals in this study and those of a previous study using DMC with the frozen DFT orbitals. Although different software and implementations were used in each study, the binding energies agree within the statistical errors. This confirms the validity of our calculation results. Furthermore, we thoroughly checked size-consistency. In particular, we verified whether the sum of the energies of the fragments constituting the complex (i.e., Monomer 1 and Monomer 2) is consistent with the total energy of a far-away complex, where the two fragments are sufficiently separated (10 Å was employed in this study) and treated as a single molecule. This is because failure to check the property could lead to situations where

TABLE II. Binding energies (kcal/mol) obtained by FN-SD-DMC calculations. The calculation settings (basis-set, effective core potential, XC for trial WF, locality approximation, and software package) used in this work, Refs. 5, 14, and 80, are (aug-cc-pVTZ,⁵⁷ ccECP,⁵⁷ LDA-PZ, T-move with DLA,¹⁸ and TurboRVB⁵⁴), (plane-wave, eCEPP,⁸¹ LDA-PZ, DLA, and CASINO⁸²), (aug-VTZ,⁸³ BFD-ECP,⁸³ LDA-PZ, T-move,¹⁷ and QWalk⁸⁴), and (aug-cc-pV6Z,⁵⁷ ccECP,⁵⁷ LDA-PZ, T-move with DLA,¹⁸ and TurboRVB⁵⁴), respectively. The numbers in parentheses represent 1 σ error bars.

System	This work	Previous studies
AcOH–AcOH	−20.29(4)	−20.30(8) ⁵
HCOOH–HCOOH	−19.70(4)	−19.6(1) ⁸⁰
Water–peptide	−8.48(4)	−8.58(7) ⁵
MeOH–pyridine	−7.81(4)	−7.89(7) ⁵
Water–water	−5.14(3)	−5.16(3) ⁵
MeNH ₂ –MeNH ₂	−4.17(3)	−4.20(6) ⁵
Peptide–pentane	−4.03(7)	−3.82(8) ⁵
Uracil–cyclopentane	−3.72(7)	−3.59(9) ⁵
Pentane–AcNH ₂	−3.31(6)	−3.08(7) ⁵
Benzene–benzene	−2.40(7)	−2.33(7) ⁵
Ethene–ethyne	0.91(3)	1.04(9) ¹⁴
Ethyne–ethyne	1.24(3)	1.32(8) ¹⁴

TABLE III. Total energies (Ha), binding energy [$E_b = E_{\text{Complex}} - (E_{\text{Monomer 1}} + E_{\text{Monomer 2}})$] (kcal/mol), and size-consistency check [$E_s = E_{\text{Far-away}} - (E_{\text{Monomer 1}} + E_{\text{Monomer 2}})$] (kcal/mol). The numbers in parentheses represent 1 σ statistical uncertainty. The total energies are extrapolated values ($a \rightarrow 0$), while E_b and E_s are computed with $a = 0.2$.

System	Ansatz	$E_{\text{Monomer 1}}$	$E_{\text{Monomer 2}}$	E_{Complex}	$E_{\text{Far-away}}$	E_b	E_s
AcOH–AcOH	FN-SD-DMC	−45.820 54(7)	...	−91.673 6(2)	−91.641 3(2)	−20.29(4)	−0.05(4)
	FN-AGPn-DMC	−45.824 94(7)	...	−91.681 5(2)	−91.650 2(3)	−19.74(6)	−0.09(6)
HCOOH–HCOOH	FN-SD-DMC	−38.930 99(5)	...	−77.893 5(1)	−77.861 9(1)	−19.70(4)	−0.02(4)
	FN-AGPn-DMC	−38.935 53(5)	...	−77.901 4(2)	−77.871 2(1)	−19.09(4)	−0.11(4)
Water–peptide	FN-SD-DMC	−17.235 65(3)	−47.182 84(9)	−64.432 0(1)	−64.418 6(1)	−8.48(4)	−0.10(4)
	FN-AGPn-DMC	−17.237 93(3)	−47.186 4(1)	−64.437 2(2)	−64.424 7(1)	−7.95(4)	0.01(4)
MeOH–pyridine	FN-SD-DMC	−24.090 33(3)	−41.252 71(8)	−65.355 5(2)	−65.343 0(2)	−7.81(4)	−0.06(4)
	FN-AGPn-DMC	−24.092 77(3)	−41.257 1(1)	−65.361 7(2)	−65.349 9(2)	−7.41(6)	−0.01(6)
Water–water	FN-SD-DMC	−17.235 77(4)	−17.235 81(3)	−34.479 70(5)	−34.471 48(6)	−5.14(3)	−0.07(3)
	FN-AGPn-DMC	−17.238 02(3)	−17.238 16(3)	−34.484 06(5)	−34.476 20(4)	−4.96(3)	−0.04(3)
MeNH ₂ –MeNH ₂	FN-SD-DMC	−18.587 62(3)	−18.587 54(3)	−37.182 00(8)	−37.175 2(1)	−4.17(3)	−0.05(3)
	FN-AGPn-DMC	−18.589 37(3)	−18.589 33(4)	−37.185 4(1)	−37.178 7(1)	−4.07(3)	0.02(3)
Peptide–pentane	FN-SD-DMC	−35.587 0(1)	−47.182 93(9)	−82.775 8(4)	−82.769 6(4)	−4.03(7)	0.07(8)
	FN-AGPn-DMC	−35.588 8(1)	−47.186 5(1)	−82.782 1(6)	−82.774 4(6)	−3.92(8)	−0.08(8)
Uracil–cyclopentane	FN-SD-DMC	−77.007 5(1)	−34.375 06(8)	−111.389 1(4)	−111.382 8(4)	−3.72(7)	−0.1(1)
	FN-AGPn-DMC	−77.013 4(2)	−34.376 80(9)	−111.396 0(5)	−111.389 9(5)	−3.7(1)	−0.1(1)
Pentane–AcNH ₂	FN-SD-DMC	−40.313 88(9)	−35.586 9(1)	−75.906 1(4)	−75.901 1(4)	−3.31(6)	−0.15(6)
	FN-AGPn-DMC	−40.317 5(1)	−35.589 1(2)	−75.909 7(5)	−75.906 3(6)	−3.28(8)	0.07(7)
Benzene–benzene	FN-SD-DMC	−37.619 2(2)	...	−75.242 0(5)	−75.238 7(5)	−2.40(7)	−0.06(7)
	FN-AGPn-DMC	−37.622 5(3)	...	−75.248 8(6)	−75.246 0(6)	−2.33(8)	−0.07(8)
Ethene–ethyne	FN-SD-DMC	−13.711 58(4)	−12.456 40(3)	−26.166 38(5)	−26.167 96(6)	0.91(3)	−0.04(3)
	FN-AGPn-DMC	−13.713 82(4)	−12.459 54(3)	−26.171 67(7)	−26.173 28(6)	0.96(3)	−0.01(3)
Ethyne–ethyne	FN-SD-DMC	−12.456 36(4)	...	−24.910 77(5)	−24.912 70(5)	1.24(3)	0.01(3)
	FN-AGPn-DMC	−12.459 47(3)	...	−24.916 88(6)	−24.918 91(6)	1.32(3)	0.02(3)

the binding energy is underestimated due to the artifact. In particular, if the complex is optimized to a lower energy solely because it has more atoms than the individual fragments, the apparent binding energy might be underestimated. In this study, as shown in Table III and Fig. 4, we calculated the total energy of several far-away complexes and confirmed that the total energy is consistent with the sum of two fragments within the statistical errors. We note that the binding energies and size-consistency checks were computed with

$a = 0.2$. The other important check we have done is the total energy comparison between FN-SD-DMC and FN-AGPn-DMC (Fig. 5). Table III lists the total energies ($a \rightarrow 0$) obtained for each fragment and complex. These results show that for all compounds, the AGPn Ansatzes yield sufficiently lower energies compared to the SD with the frozen DFT orbitals. This indicates that FN-AGPn-DMC provides a better nodal surface than FN-SD-DMC (from the variational principle).

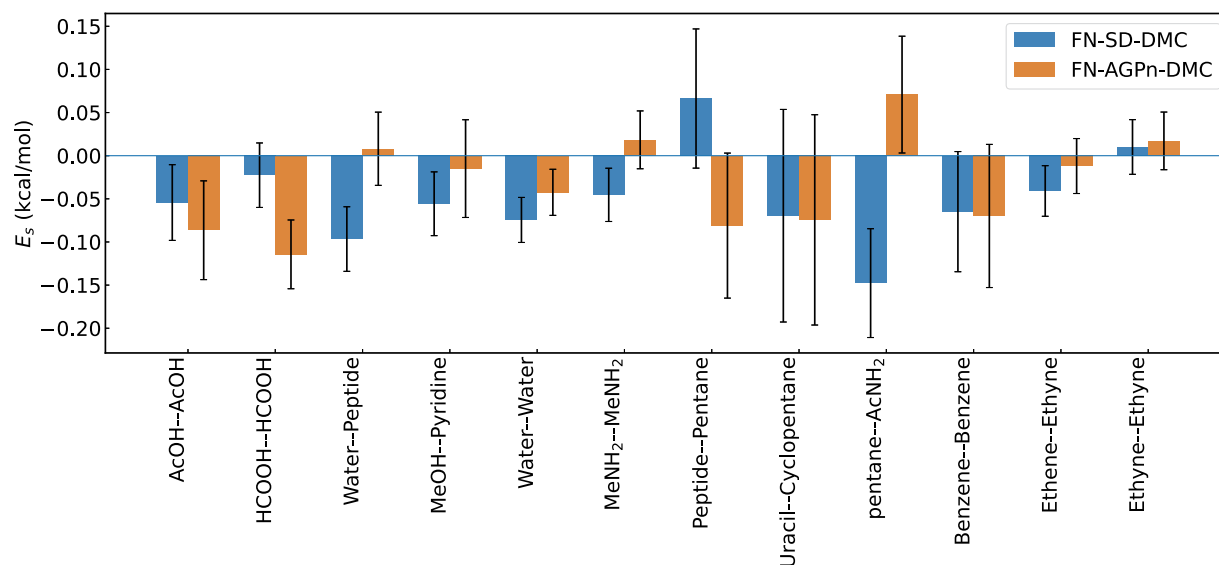


FIG. 4. Results for the size-consistency check for FN-SD-DMC and FN-AGPn-DMC. $E_s = E_{\text{Far-away}} - (E_{\text{Monomer 1}} + E_{\text{Monomer 2}})$ (kcal/mol). Error bars represent 1σ .

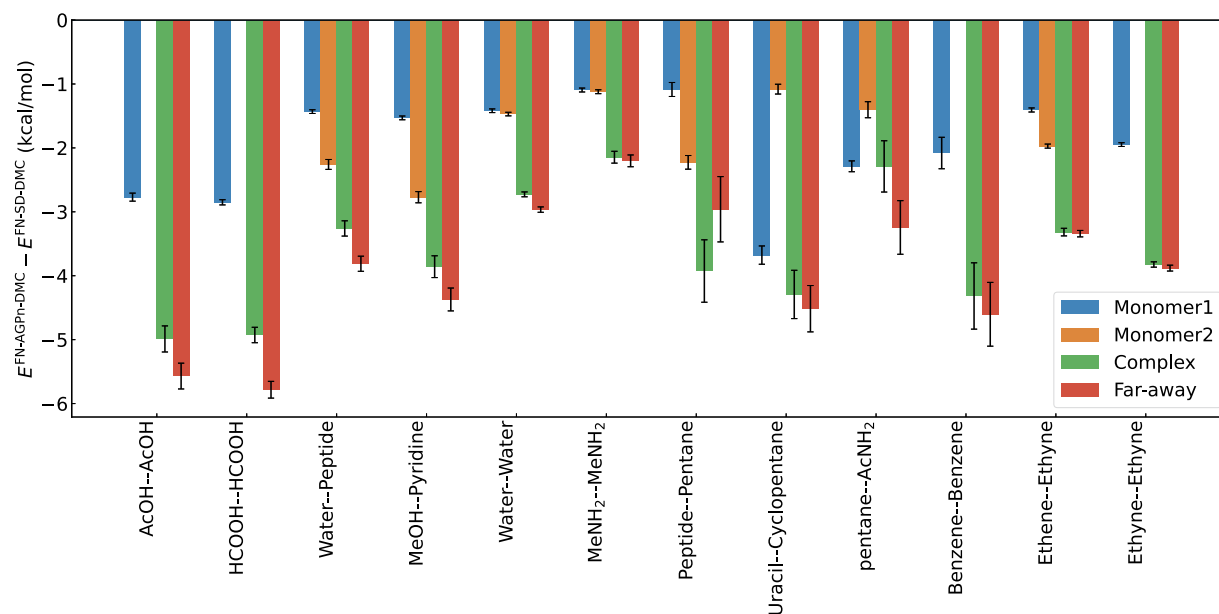


FIG. 5. Differences in total energies between FN-SD-DMC and FN-AGPn-DMC for Monomer1, Monomer2 (if it is not identical to Monomer1), Complex, and Far-away complex. The FN-AGPn-DMC yields lower energies than FN-SD-DMC does for all cases, indicating that the obtained nodal surfaces are improved for all cases.

REFERENCES

- ¹K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989).
- ²W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- ³Y. S. Al-Hamdani, P. R. Nagy, A. Zen, D. Barton, M. Kállay, J. G. Brandenburg, and A. Tkatchenko, *Nat. Commun.* **12**, 3927 (2021).
- ⁴V. Fishman, M. Lesiuk, J. M. L. Martin, and A. Daniel Boese, *J. Chem. Theory Comput.* **21**, 2311 (2025).
- ⁵B. X. Shi, F. Della Pia, Y. S. Al-Hamdani, A. Michaelides, D. Alfè, and A. Zen, *J. Chem. Phys.* **162**, 144107 (2025).
- ⁶J. Řezáč, L. Šimová, and P. Hobza, *J. Chem. Theory Comput.* **9**, 364 (2013).
- ⁷L. Demovičová, P. Hobza, and J. Řezáč, *Phys. Chem. Chem. Phys.* **16**, 19115 (2014).
- ⁸K. U. Lao, *J. Chem. Phys.* **161**, 234103 (2024).
- ⁹S. Lambie, D. Kats, D. Usvyat, and A. Alavi, *J. Chem. Phys.* **162**, 114112 (2025).
- ¹⁰E. Semidalas, A. D. Boese, and J. M. L. Martin, *Chem. Phys. Lett.* **863**, 141874 (2025).
- ¹¹S. Lambie, C. Rickert, D. Usvyat, A. Alavi, and D. Kats, *J. Chem. Phys.* **163**, 111101 (2025).
- ¹²T. Schäfer, A. Irmeler, A. Gallo, and A. Grüneis, *Nat. Commun.* **16**, 9108 (2025).
- ¹³N. Masios, A. Irmeler, T. Schäfer, and A. Grüneis, *Phys. Rev. Lett.* **131**, 186401 (2023).
- ¹⁴K. Nakano, B. X. Shi, D. Alfè, and A. Zen, *J. Chem. Theory Comput.* **21**, 4426 (2025).
- ¹⁵D. M. Ceperley, *J. Stat. Phys.* **43**, 815 (1986).
- ¹⁶L. Mitáš, E. L. Shirley, and D. M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991).
- ¹⁷M. Casula, *Phys. Rev. B* **74**, 161102 (2006).
- ¹⁸A. Zen, J. G. Brandenburg, A. Michaelides, and D. Alfè, *J. Chem. Phys.* **151**, 134105 (2019).
- ¹⁹C. J. Umrigar, M. P. Nightingale, and K. J. Runge, *J. Chem. Phys.* **99**, 2865 (1993).
- ²⁰A. Zen, S. Sorella, M. J. Gillan, A. Michaelides, and D. Alfè, *Phys. Rev. B* **93**, 241118 (2016).
- ²¹F. Della Pia, B. X. Shi, Y. S. Al-Hamdani, D. Alfè, T. A. Anderson, M. Barborini, A. Benali, M. Casula, N. D. Drummond, M. Dubecký, C. Filippi, P. R. C. Kent, J. T. Krogel, P. López Ríos, A. Lüchow, Y. Luo, A. Michaelides, L. Mitás, K. Nakano, R. J. Needs, M. C. Per, A. Scemama, J. Schultze, R. Shinde, E. Sliotman, 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).
- ²²A. Benali, H. Shin, and O. Heinonen, *J. Chem. Phys.* **153**, 194113 (2020).
- ²³J. B. Anderson, *J. Chem. Phys.* **65**, 4121 (1976).
- ²⁴A. Annarelli, D. Alfè, and A. Zen, *J. Chem. Phys.* **161**, 241501 (2024).
- ²⁵C. Filippi and C. J. Umrigar, *J. Chem. Phys.* **105**, 213 (1996).
- ²⁶C. J. Umrigar and C. Filippi, *Phys. Rev. Lett.* **94**, 150201 (2005).
- ²⁷J. Toulouse and C. J. Umrigar, *J. Chem. Phys.* **128**, 174101 (2008).
- ²⁸M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria, *J. Chem. Theory Comput.* **8**, 2181 (2012).
- ²⁹E. Giner, A. Scemama, and M. Caffarel, *J. Chem. Phys.* **142**, 044115 (2015).
- ³⁰M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, *J. Chem. Phys.* **144**, 151103 (2016).
- ³¹A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, *J. Comput. Chem.* **37**, 1866 (2016).
- ³²M. Dash, S. Moroni, A. Scemama, and C. Filippi, *J. Chem. Theory Comput.* **14**, 4176 (2018).
- ³³A. Scemama, Y. Garniron, M. Caffarel, and P.-F. Loos, *J. Chem. Theory Comput.* **14**, 1395 (2018).
- ³⁴A. Scemama, E. Giner, A. Benali, and P.-F. Loos, *J. Chem. Phys.* **153**, 174107 (2020).
- ³⁵M. Casula and S. Sorella, *J. Chem. Phys.* **119**, 6500 (2003).
- ³⁶N. D. Drummond, P. L. Ríos, A. Ma, J. R. Trail, G. G. Spink, M. D. Towler, and R. J. Needs, *J. Chem. Phys.* **124**, 224104 (2006).
- ³⁷P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. E* **74**, 066701 (2006).
- ³⁸M. Bajdich, L. Mitás, L. K. Wagner, and K. E. Schmidt, *Phys. Rev. B* **77**, 115112 (2008).
- ³⁹M. Bajdich, L. Mitás, G. Drobny, L. K. Wagner, and K. E. Schmidt, *Phys. Rev. Lett.* **96**, 130201 (2006).
- ⁴⁰C. Genovese, T. Shirakawa, K. Nakano, and S. Sorella, *J. Chem. Theory Comput.* **16**, 6114 (2020).
- ⁴¹J. Hermann, Z. Schätzle, and F. Noé, *Nat. Chem.* **12**, 891 (2020).
- ⁴²D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, *Phys. Rev. Res.* **2**, 033429 (2020).
- ⁴³X. Li, Z. Li, and J. Chen, *Nat. Commun.* **13**, 7895 (2022).
- ⁴⁴W. Ren, W. Fu, X. Wu, and J. Chen, *Nat. Commun.* **14**, 1860 (2023).
- ⁴⁵I. G. Gurtubay and R. J. Needs, *J. Chem. Phys.* **127**, 124306 (2007).
- ⁴⁶S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).
- ⁴⁷S. Azadi and R. E. Cohen, *J. Chem. Phys.* **143**, 104301 (2015).
- ⁴⁸K. Nakano, R. Maezono, and S. Sorella, *J. Chem. Theory Comput.* **15**, 4044 (2019).
- ⁴⁹S. Lambie, P. L. Ríos, D. Kats, and A. Alavi, *J. Chem. Theory Comput.* **22**, 7693 (2026).
- ⁵⁰K. Nakano, S. Sorella, D. Alfè, and A. Zen, *J. Chem. Theory Comput.* **20**, 4591 (2024).
- ⁵¹P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006).
- ⁵²J. Řezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011).
- ⁵³J. Řezáč and P. Hobza, *J. Chem. Theory Comput.* **9**, 2151 (2013).
- ⁵⁴K. 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).
- ⁵⁵J. McFarland and E. Manousakis, *Phys. Rev. A* **105**, 032815 (2022).
- ⁵⁶S. Sorella, *Phys. Rev. Lett.* **80**, 4558 (1998).
- ⁵⁷M. C. Bennett, C. A. Melton, A. Annaberdiyev, G. Wang, L. Shulenburger, and L. Mitás, *J. Chem. Phys.* **147**, 224106 (2017).
- ⁵⁸Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma *et al.*, *WIREs Comput. Mol. Sci.* **8**, e1340 (2018).
- ⁵⁹Q. 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).
- ⁶⁰Q. Sun, M. R. Hermes, X. Wu, H. Zhai, X. Zhang, A. M. Ahmed, J. J. Aúcar, O. J. Backhouse, S. Banerjee, P. Bao, N. A. Bogdanov, K. Bystrom, F. Chapoton, N.-Y. Chen, I. Y. Chernyshov, H. S. Clifford, S. Cohen-Janes, Z.-H. Cui, Y. D. Damour, N. Dattani, L. B. Dittmer, S. Ehlert, J. J. Eriksen, F. A. Evangelista, S. A. Ewing, A. Farahvash, K. Focke, Y. Gao, K. E. Gasperich, N. Gillispie, J. Greiner, M. R. Hennefarth, J. Hermann, C. Hillenbrand, J. Huhtasalo, B. Ibrahim, B. Jangid, A. N. Javaremi, A. J. Jenkins, Y. Jin, D. S. King, D. P. Kooi, J. S. Kurian, H. R. Larsson, B. T. G. Lau, S. Lee, S. Lehtola, C. Li, H. Li, J. Li, R. Li, S. Li, A. O. Lykhin, A. Mahajan, N. Mauger, P. del Mazo-Sevillano, J. Moussa, K. Nakano, V. A. Neufeld, L. Peng, H. Q. Pham, P. Pinski, P. Pokhilko, Z. Pu, Y. Qian, S. J. Quinon, W. T. Schulze, T. R. Scott, A. Seal, J. D. Serna, J. E. T. Smith, K. E. Smyser, T. Stahl, C. Sun, K. J. Sung, E. Trushin, S. Upadhyay, E. A. Vo, T. Vogels, S. Wang, T. Wang, X. Wang, X. Wang, Y. Wang, M. Williamson, J. Yang, H.-Z. Ye, C.-N. Yeh, H. Yu, J. Yu, V. W. Yu, C. Zhang, D. Zhang, Y. Zhang, Z. Zhao, Z. Zhou, A. J. Zhu, T. Zhu, T. C. Berkelbach, L. Gagliardi, S. Sharma, A. Sokolov, and G. K.-L. Chan, "The Python simulations of chemistry framework: 10 years of an open-source quantum chemistry project," *J. Chem. Phys.* **165**, 102502 (2026).
- ⁶¹J. Ma, D. Alfè, A. Michaelides, and E. Wang, *J. Chem. Phys.* **130**, 154303 (2009).
- ⁶²M. C. Per, K. A. Walker, and S. P. Russo, *J. Chem. Theory Comput.* **8**, 2255 (2012).

- ⁶³Y. Kadioglu, J. A. Santana, H. D. Özaydin, F. Ersan, O. Ü. Aktürk, E. Aktürk, and F. A. Reboredo, *J. Chem. Phys.* **148**, 214706 (2018).
- ⁶⁴M. C. Per, E. K. Fletcher, and D. M. Cleland, *J. Chem. Phys.* **150**, 184101 (2019).
- ⁶⁵N. A. Benedek, I. K. Snook, M. D. Towler, and R. J. Needs, *J. Chem. Phys.* **125**, 104302 (2006).
- ⁶⁶M. Dubecký, R. Derian, P. Jurečka, L. Mitás, P. Hobza, and M. Otyepka, *Phys. Chem. Chem. Phys.* **16**, 20915 (2014).
- ⁶⁷Y. S. Al-Hamdani, D. Alfè, O. A. von Lilienfeld, and A. Michaelides, *J. Chem. Phys.* **141**, 18C530 (2014).
- ⁶⁸K. Krongchon, B. Busemeyer, and L. K. Wagner, *J. Chem. Phys.* **146**, 124129 (2017).
- ⁶⁹Y. S. Al-Hamdani, A. Zen, and D. Alfè, *J. Chem. Phys.* **159**, 204708 (2023).
- ⁷⁰E. 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).
- ⁷¹K. Nakano, O. Kohulák, A. Raghav, M. Casula, and S. Sorella, *J. Chem. Phys.* **159**, 224801 (2023).
- ⁷²M. Casula, C. Filippi, and S. Sorella, *Phys. Rev. Lett.* **95**, 100201 (2005).
- ⁷³K. Nakano, S. Sorella, and M. Casula, *J. Chem. Phys.* **163**, 194117 (2025).
- ⁷⁴T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, *J. Chem. Phys.* **132**, 144104 (2010).
- ⁷⁵R. J. Bartlett, *J. Phys. Chem.* **93**, 1697 (1989).
- ⁷⁶E. Neuscamman, *Phys. Rev. Lett.* **109**, 203001 (2012).
- ⁷⁷K. Momma and F. Izumi, *J. Appl. Crystallogr.* **44**, 1272 (2011).
- ⁷⁸K. Yoshimi, Y. Motoyama, T. Aoyama, M. Kawamura, and N. Kawashima, *Sci. Technol. Adv. Mater.: Methods* **5**, 2564055 (2025).
- ⁷⁹J. Řezáč, P. Jurečka, K. E. Riley, J. Černý, H. Valdes, K. Pluháčková, K. Berka, T. Řezáč, M. Pitoňák, J. Vondrášek, and P. Hobza, *Collect. Czech. Chem. Commun.* **73**, 1261 (2008).
- ⁸⁰M. Dubecký, *J. Chem. Theory Comput.* **13**, 3626 (2017).
- ⁸¹J. R. Trail and R. J. Needs, *J. Chem. Phys.* **146**, 204107 (2017).
- ⁸²R. J. Needs, M. D. Towler, N. D. Drummond, P. López Ríos, and J. R. Trail, *J. Chem. Phys.* **152**, 154106 (2020).
- ⁸³M. Burkatzki, C. Filippi, and M. Dolg, *J. Chem. Phys.* **126**, 234105 (2007).
- ⁸⁴L. K. Wagner, M. Bajdich, and L. Mitás, *J. Comput. Phys.* **228**, 3390 (2009).