Exploring non-linear correlators on AGP

Armin Khamoshi*

Department of Physics and Astronomy, Rice University, Houston, TX 77005-1892

Guo P. Chen

Department of Chemistry, Rice University, Houston, TX 77005-1892

Thomas M. Henderson and Gustavo E. Scuseria

Department of Physics and Astronomy, Rice University, Houston, TX 77005-1892

(Dated: December 7, 2020)

Single-reference methods such as Hartree-Fock-based coupled cluster theory are well known for their accuracy and efficiency for weakly correlated systems. For strongly correlated systems, more sophisticated methods are needed. Recent studies have revealed the potential of the antisymmetrized geminal power (AGP) as an excellent initial reference for the strong correlation problem. While these studies improved on AGP by linear correlators, we explore some non-linear exponential ansätze in this paper. We investigate two approaches in particular. Similar to Phys. Rev. B 91, 041114(R) (2015), we show that the similarity transformed Hamiltonian with a Hilbert-space Jastrow operator is summable to all orders and can be solved over AGP by projecting Schrödinger’s equation. The second approach is based on approximating the unitary pair-hopper ansatz recently proposed for application on a quantum computer. We report benchmark numerical calculations against the ground state of the pairing Hamiltonian for both of these approaches.

I. INTRODUCTION

Many-body methods in electronic structure theory frequently start from a mean-field reference. Commonly, a single-reference Slater determinant obtained by Hartree-Fock (HF) calculations is used for post-HF methods that make correction to this mean-field [1]. Such methods include configuration interaction (CI) and coupled cluster theory (CC). For weakly correlated systems, the single-reference CC is often regarded as the gold standard and calculations are routinely performed using CC with single and double excitations (CCSD) very accurately at an affordable cost of $O(M^6)$ where $M$ is the size of the system [1–3]. Unfortunately, the same cannot be said for the strong correlation problem. When many-body interactions are strong, more than one Slater determinant becomes important and methods based on a single HF reference often break down [4–6]. Although multireference methods exist, they have certain limitations and are not always straightforward to use [7–10]. As such, a new way of thinking and more sophisticated methods are desirable.

A useful way of organizing the Hilbert space in the context of strong correlation is by seniority [11, 12]. Loosely speaking, seniority ($\Omega$) is the number of unpaired fermions in some appropriate pairing scheme. For example, a two-body Hamiltonian can generally be written as [13]

$$H = H^{\Omega=0} + H^{\Omega=2} + H^{\Omega=4}$$

where the superscripts indicate the portion of the Hamiltonian that couples those determinants whose seniority differs by 0, 2, etc. Numerical evidence has shown that the low-seniority portions of the Hilbert space can recover a very large part of the correlation energy and have the correct qualitative behavior [12, 14]. The seniority-zero subspace can be solved exactly by the doubly occupied configuration interaction (DOCI) but it scales combinatorially with system size [15–17]. Remarkably, CC restricted to paired double excitation (pCCD) [18], also known as AP1roG [19], is often nearly exact in this subspace with a very low computational cost for realistic molecular Hamiltonians [20–22]. However, this correspondence is not universal and it has been shown that pCCD and its extensions break down for some problems including the attractive pairing Hamiltonian [5, 13, 23–25],

$$H = \sum_p \epsilon_p N_p - G \sum_{pq} P^\dagger_p P^\dagger_q,$$  

where $\epsilon_p$ is the single-particle energy level, $G$ tunes the strength of the pairwise, infinite-range interactions, and

$$P_{pq}^\dagger = c_{p}^\dagger c_{q}^\dagger,$$

$$N_p = c_{p}^\dagger c_p,$$

such that $c_{p}^\dagger$ is a fermion in spin-orbital $p$, and $p$ is its “paired” companion. Generally, in the seniority-zero subspace, all operators can be organized as a linear combination of terms like $P_{p}^\dagger \ldots N_{q} \ldots P_{r}^\dagger$ [26], where the nilpotent operators $P_{p}^\dagger$ and $P_{p}$ together with $N_{p}$ are gen-

* Correspondence email address: armin.khamoshi@rice.edu
erators of a \( su(2) \) Lie algebra

\[
\begin{align}
\{P_p, P_q^\dagger\} &= \delta_{pq} (1 - N_p), \\
\{N_p, P_q^\dagger\} &= 2\delta_{pq} P_q^\dagger.
\end{align}
\]

The breakdown of CC in the pairing Hamiltonian can be understood in terms of the spontaneous breaking of number symmetry, which occurs at some critical value of \( G, G_c > 0 \). As a result, the mean-field solution gives rise to the number-broken Bardeen–Cooper–Schrieffer wave-function (BCS) \[27\] for \( G > G_c \), and a number-conserving Slater determinant for \( G < G_c \), including all \( G < 0 \) where the interaction is repulsive. Indeed, it is in the attractive regime where CC as well as a whole host of other conventional methods struggle \[25\]. It is noteworthy that the pairing Hamiltonian is exactly solvable \[23, 28\].

Recently, we proposed to use the AGP wavefunction as the initial reference for this and potentially more general problems \[25, 26, 29–32\]. The main advantage of AGP for the seniority-zero sector is that on the one hand, it puts coefficients on each Slater determinant in the DOCI while still conserving seniority, and on the other hand, AGP is computationally straightforward. Put differently, a seniority-zero wavefunction has natural orbitals defined by the pairing scheme chosen, and the natural orbitals of AGP, HF, pCCD, and DOCI are all the same, differing only in occupation. The AGP wavefunction can then be expanded in its seniority-zero natural orbital determinants and can occupy all \( \binom{M}{N} \) such determinants, where \( N \) is the number of fermion pairs. It thus encompasses all possible configurations in the seniority-zero space, albeit with inexact coefficients. However, unlike DOCI, AGP can be optimized at mean-field cost \( O(M^3) \) \[33\]. It also encompasses the HF state, so it has a much richer structure as an initial reference. Moreover, we have shown recently that the reduced density matrices (RDM) over AGP can be computed efficiently as all higher order RDMs are decomposable in terms of lower order ones. \[26\] These realizations permit us to envision post-AGP methods—analogous to post-HF—to capture the remaining error and provide accurate ansätze in both weak and strong correlation. A case can be made for using AGP in the seniority non-zero spaces where pairs are broken, but we concentrate on the seniority-zero case at this stage.

Recent studies in our group have developed CI-based methods over AGP and numerical calculations show excellent accuracy when applied to the pairing Hamiltonian. These methods include orthogonal excitations \[25, 29\] and those based on the generators of the Lie algebra Eq. (3) \[31\]. It has been shown that these models are formally equivalent and can be viewed as “genuine replacement” methods \[31\]. Encouraged by the accuracy of linear models, we would like to explore non-linear correlator ansätze over AGP. Non-linear models in principle can approximate the DOCI coefficient differently and it remains to be seen whether they provide an advantage over the linear case. Developing a CC theory over AGP is not trivial and remains an open problem \[25\].

In this paper, we explore two exponential correlators over AGP which we build from number-conserving combinations of the generators of the Lie algebra, Eq. (4). The first is the exponential of Jastrow \[34\] operators on AGP (JAGP). Jastrow-Gutzwiller correlators are ubiquitous in physics and chemistry and are widely used in Monte Carlo calculations. \[34–45\] It was shown in Ref. \[46, 47\] that the the similarity transformation of any fermionic Hamiltonian, \( e^{-JH}e^J \), where \( J \) is a Hilbert-space Jastrow operator, is analytically summable to all orders. Inspired by that, we extend the formalism to the pairing \( su(2) \) algebra Eq. (4) and show that we can efficiently solve the projected Schrödinger equation over AGP. In a way, this could be viewed as a form of generalized coupled cluster theory, but instead of the traditional particle-hole excitations, correlation is mediated by number operators that create excitations on AGP—a possibility that does not occur for single Slater determinants as they are eigenfunctions of \( J \).

The second ansatz is the unitary pair-hopper that we proposed recently for applications in quantum computers \[32\]. Unitary ansätze can be implemented efficiently on a quantum computer, \[48\] but some approximations must be made in order to implement them on a classical computer. Here, we approximate the unitary transformation of the Hamiltonian using the Baker-Campbell-Hausdorff (BCH) formula by truncating it at the 4th commutator. Affordable computational scaling is made possible by the reconstruction formulae \[26\] with which we can eliminate the bottleneck of computing and storing higher rank density matrices from our calculations. As an alternative to the brute-force expansion of the BCH formula, we apply the canonical transformation theory as formulated in Ref. \[49–53\] to recursively sum the transformed Hamiltonian. We report benchmark calculations of all methods against the ground state energy of the pairing Hamiltonian.

The organization of this paper is as follows. To familiarize the reader with our notation, in Sec. II we review some basic properties of AGP and its density matrices. In Sec. III, we present our formalism for solving similarity transformed JAGP as well as its application to the pairing Hamiltonian. Sec. IV discusses the truncated commutator expansion and canonical transformation with the unitary pair-hoppers ansatz over the pairing Hamiltonian. Lastly, some discussion and concluding remarks are provided in Sec. V.

II. BACKGROUND

The AGP wavefunction with \( N \) pairs corresponds to a condensate where all \( 2N \) fermions are placed in the same spinoidal. Mathematically it can be written as \[54, 55\]

\[
|\text{AGP}\rangle = \frac{1}{N!} (\Gamma^\dagger)^N |\text{vac}\rangle, \tag{5}
\]

\[\Gamma = \sum_{p=1}^N \gamma p = \sum_{p=1}^N \sum_{q=1}^N \gamma pq, \tag{4e}\]
that for a generic many-body operator \( \hat{\Gamma} \) in the natural orbital basis \([26]\). To see this, first note that computed efficiently at polynomial cost and are very sparse.

Despite being a linear combination of a combinatorial number of determinants, reduced density matrices of seniority non-conserving operators are identically zero.

Therefore, without loss of generality, we obtain the geminal operator in the natural orbital basis of the geminal as

\[
\Gamma^\dagger = \sum_{pq} \eta_{pq} c_p^\dagger c_q^\dagger.
\]

Here, \( p, q \) denote spin-orbitals and \( \eta_{pq} \) is a skew-symmetric matrix known as the \textit{geminal coefficient}. To make a connection with Eq. (3) and for mathematical convenience, we perform an orbital rotation that brings the matrix of geminal coefficients into a block-diagonal form \([56]\), that is

\[
\eta = \bigoplus_{p=1}^M \left( \begin{array}{cc} 0 & \eta_p \\ -\eta_p & 0 \end{array} \right).
\]

Therefore, we can organize all nonzero reduced density matrices in ascending rank as follows

\[
Z^{p(1)}_p = \langle \text{AGP} | N_p \text{AGP} \rangle, \quad Z^{p(2)}_p = \langle \text{AGP} | P_p^\dagger P_p | \text{AGP} \rangle, \quad Z^{p(2)}_{pq} = \langle \text{AGP} | N_p N_q | \text{AGP} \rangle, \quad Z^{p(1)}_{pq} = \langle \text{AGP} | P_p^\dagger N_q N_p | \text{AGP} \rangle, \quad Z^{p(3)}_{pq} = \langle \text{AGP} | N_p N_q N_r | \text{AGP} \rangle, \quad \ldots
\]

where the first integer in the superscript indicates the number of \( N_p \) operators in the middle and the second integer is the rank of the density matrix. When indices of an RDM are different, we refer to it as an irreducible density matrix. This is because \( N_p = N_p^2/2 = 2P_p^\dagger P_p \) together with \( P_p^\dagger P_p = P_p^\dagger N_p \) in the seniority-zero space imply that RDMs with repeated indices are either zero or can be written in terms of lower rank density matrices.

Every matrix element of an AGP RDM, regardless of the rank, can be computed directly at \( \mathcal{O}(M^2) \) cost using a method based on elementary symmetric polynomials (ESP) \([26]\). Even more efficiently, one can take advantage of the reconstruction formulae to express any irreducible RDM as a linear combination of lower rank ones, which can be performed all the way down to occupation numbers and geminal coefficients \([26]\). This allows us to compute any density matrix of rank 2 or higher with \( \mathcal{O}(1) \) cost per element. Finally, we can take advantage of the equivalence between AGP and the number-projected BCS wavefunction \([11, 57]\) to extract AGP density matrices via grid integrations of BCS transition density matrices, at \( \mathcal{O}(N_{\text{grid}}) \) cost per element, where \( N_{\text{grid}} \) is the size of the numerical quadrature for gauge integration \([33]\). While less efficient than the reconstruction formulae, this final approach can be very efficient when used to contract density matrices with other factorized quantities. In the subsequent sections, we use these methods to compute all density matrices in our calculations.

### III. JASTROW-AGP

#### A. Analytic properties

Let the \( k \)-body, or rank-\( k \), Hilbert-space Jastrow operator \([58]\) be written as

\[
J_k = \frac{1}{2k} \sum_{p_1 < \ldots < p_k} \alpha_{p_1 \ldots p_k} N_{p_1} N_{p_2} \ldots N_{p_k},
\]

where the sums run over all orbitals, \( \alpha \) is a symmetric tensor and is invariant under the exchange of its indices, and \( 1/2^k \) is introduced for convenience. Importantly, since \( N_p^2 = 2N_p \) in the seniority-zero space, we impose that \( \alpha \) is zero if any two indices are the same and eliminate these terms from the sum in Eq. (13).
From the $k$-body Jastrow operator we can define the $k$-body Jastrow-AGP ($J_k$AGP) wavefunction as

$$|J_k\text{AGP}\rangle = e^{J_k}|\text{AGP}\rangle$$

As we prove in Appendix A, we do not need to include lower body Jastrow operators in this $J_k\text{AGP}$ wavefunction, as they are contained inside $J_k$. Thus, for example, we may write

$$e^{J_1+J_2}|\text{AGP}\rangle = e^{J_2}|\text{AGP}\rangle.$$  \hspace{1cm} (15)

where $J'_2$ is another two-body Jastrow operator.

Slater determinants built from the same orbitals used to define the Hilbert-space Jastrow operators are eigenfunctions of $J_k$, that is

$$J_k|\Phi_{\mu}\rangle = J_{k,\mu}|\Phi_{\mu}\rangle,$$  \hspace{1cm} (16a)

$$J_{k,\mu} = \frac{1}{2k} \sum_{p_1 < \ldots < p_k} \alpha_{p_1 \ldots p_k} N_{p_1,\mu} \ldots N_{p_k,\mu}$$

$$= \sum_{i_1 < \ldots < i_k} \alpha_{i_1 \ldots i_k},$$  \hspace{1cm} (16b)

where $|\Phi_{\mu}\rangle$ is a determinant, $N_{p,\mu}$ is the occupation number of level $p$ in determinant $\mu$, and the indices $i_1 \ldots i_k$ run over the orbitals occupied in $|\Phi_{\mu}\rangle$. From this, it follows that

$$e^{J_k} \sum_{\mu} c_{\mu} |\Phi_{\mu}\rangle = \sum_{\mu} e^{J_{k,\mu}} c_{\mu} |\Phi_{\mu}\rangle.$$  \hspace{1cm} (17)

For a complex $\alpha$, the magnitude of the coefficient of a given determinant is modified by the Hermitian part of $J_k$, i.e. the real part of $\alpha$. The anti-Hermitian part of $J_k$ (the imaginary part of $\alpha$) only adjusts the phase of each coefficient. It is thus apparent that for an $N$-body initial wavefunction $|\Psi\rangle$ in which all determinants $|\Phi_{\mu}\rangle$ have non-zero coefficients, $\exp(J_N)|\Psi\rangle$ can be made an exact eigenstate of a chosen Hamiltonian provided that $J_N$ is non-Hermitian.

AGP is, in general, such an initial wavefunction, so $\exp(J_N)$ acting on it can be made exact for the seniority zero space. For any $k \leq N$, it follows from Eq. (16) and Eq. (17) that

$$e^{J_k}|\text{AGP}\rangle = \sum_{p_1 < \ldots < p_N} \exp \left( \sum_{1 \leq j_1 < \ldots < j_k \leq N} \alpha_{p_{j_1} \ldots p_{j_k}} \eta_{p_{j_1} \ldots p_{j_k}} P_{p_{j_1}} \ldots P_{p_{j_k}} \right) |\text{AGP}\rangle.$$  \hspace{1cm} (18)

In contrast, we can see how the linear wavefunction $|\Psi\rangle = J_k|\text{AGP}\rangle$, which we refer to as $J_k\text{CI AGP}$ [31], modifies AGP as follows

$$|\Psi\rangle = \sum_{p_1 < \ldots < p_N} \left( \sum_{1 \leq j_1 < \ldots < j_k \leq N} \alpha_{p_{j_1} \ldots p_{j_k}} \right) \eta_{p_{j_1} \ldots p_{j_k}} P_{p_{j_1}} \ldots P_{p_{j_k}} |\text{AGP}\rangle.$$  \hspace{1cm} (19)

Thus, the way in which the linear and non-linear $J_k$ correlators factorize the DOCI coefficients can be easily deduced from Eq. (18) and Eq. (19). Together with Eq. (10), we can readily see that

$$\begin{align*}
D_{p_1 \ldots p_N}^{(\text{Linear})} &\approx \sum_{1 \leq j_1 < \ldots < j_k \leq N} \alpha_{p_{j_1} \ldots p_{j_k}} \eta_{p_{j_1} \ldots p_{j_k}} \eta_{p_{j_k+1} \ldots p_{N}} , \hspace{1cm} (20a) \\
D_{p_1 \ldots p_N}^{(\text{Non-linear})} &\approx \prod_{1 \leq j_1 < \ldots < j_k \leq N} e^{\alpha_{p_{j_1} \ldots p_{j_k}}} \eta_{p_{j_1} \ldots p_{j_k}} \eta_{p_{j_k+1} \ldots p_{N}} . \hspace{1cm} (20b)
\end{align*}$$

Indeed, the two factorizations are quite different. An implication can be made, for example, about the sign of each coefficient; if $\alpha$ is purely real, the DOCI coefficient in Eq. (20a) can be negative only if some $\eta$’s are negative, whereas the same restriction is not implied in (20b).

Having shown analytically how $J_k$ acts on AGP, we concern ourselves with optimizing the energy over JAGP in the subsequent sections.

### B. Energy optimization with JAGP

Consider $k = 1$ for which the $J_1\text{AGP}$ wavefunction is just an AGP,

$$e^{J_1}|\text{AGP}\rangle = \sum_{p_1 < \ldots < p_N} e^{\sum_{j} \alpha_{p_j} \left( \prod_{k} \eta_{p_k} P_{p_k} \right) \langle \text{vac} \rangle}.$$  \hspace{1cm} (21a)

$$= \sum_{p_1 < \ldots < p_N} \left( \prod_{k} \eta_{p_k} e^{\alpha_{p_k} P_{p_k}} \right) \langle \text{vac} \rangle .$$  \hspace{1cm} (21b)

Evidently, the action of $J_1$ just changes the geminal coefficients $\eta_{p_k}$ to $e^{\alpha_{p_k}}$. The $k = 1$ J1AGP can optimize an AGP, but it cannot correlate it; we need $k > 1$ to go beyond mean-field. Unfortunately, except for $k = 1$, the norm and density matrices over $e^{J_k}|\text{AGP}\rangle$ are not trivial to compute, and aside from quantum Monte Carlo methods, [36, 38, 42, 43, 59, 60] we are not aware of an efficient non-stochastic algorithm to compute the overlaps. This for examples makes the variational optimization of JAGP energy

$$E_{\text{var.-JAGP}} = \frac{\langle \text{AGP}|e^{J}H|\text{AGP}\rangle}{\langle \text{AGP}|e^{J}e^{J}|\text{AGP}\rangle} ,$$  \hspace{1cm} (22)

out of reach for the existing non-stochastic methods.

However, a case can be made for the similarity transformed Hamiltonian using $J_2$, that is

$$E_{\text{ST-J2AGP}} = \langle \text{AGP}|e^{-J_2}H e^{J_2}|\text{AGP}\rangle .$$  \hspace{1cm} (23)

The two-body Jastrow correlator has a long history in electronic structure methods and has shown to be capable of recovering a significant portion of the correlation energy. [37, 42, 44] Previous work in our group has shown that the similarity transformation of fermionic Hamiltonians using a two-body Jastrow operator is summable to
all orders and the amplitudes can be solved by left projection via components of $J_2$ acting on the HF reference [46, 47]. Following the same line of reasoning, we extend that formalism to the pairing $su(2)$ algebra Eq. (4) and show how to obtain the working equations for AGP in the next section.

The main idea relies on the fact that the similarity transformation of the Hamiltonian reduces the rank of $J_2$ by one, thus the remaining one-body Jastrow operator can be subsequently absorbed into the reference. The same idea can be applied to the unitary transformation to begin by working through the case in which we use a Slater determinant in AGP and do not alter their that break time-reversal symmetry. For other Hamiltonians, the mere change in phase is insufficient to make the wavefunction exact.

C. Similarity transformation with JAGP

In this section, we extend the results of Ref. [46, 47] to the AGP wavefunction and our $su(2)$ algebra. For pedagogical reasons, we keep this section self-contained. Suppose one had a generic Jastrow operator $J$. In a manner similar to coupled cluster theory, we intend to solve for the ground state energy by similarity transforming the Hamiltonian,

$$
\hat{H} = e^{-\tilde{J}} \hat{H} e^{\tilde{J}},
$$

(25a)

$$
\tilde{E} = \langle \text{AGP} | \hat{H} | \text{AGP} \rangle.
$$

(25b)

Similarity transformation does not change the spectrum of $\hat{H}$, but $\tilde{H}$ is no longer Hermitian and $\tilde{E}$ is not variational. Therefore we obtain the residual equations by left projecting the Schrödinger equation to get

$$
\langle N_p, \ldots N_p \rangle \hat{H} - \tilde{E} \langle N_p, \ldots N_p \rangle = 0 \quad \forall p_1 < \ldots < p_k
$$

(26)

where the expectation values are taken with respect to AGP. In traditional coupled cluster theory using particle-hole excitations, one expresses the similarity transformation by the Baker-Campbell-Hausdorff (BCH) expansion whose series naturally truncates at the 4th commutator for a two-body Hamiltonian. In contrast, the transformation using the Jastrow operator does not truncate, but it is summable to all orders for any Hamiltonian.

Our ultimate goal is to use a two- or higher-body Jastrow operators to similarity transform the Hamiltonian and provide correlation. However, it will prove useful to begin by working through the case in which we use a one-body operator, for which the algebra is much simpler. Therefore, consider the $J_1$ operator

$$
J_1 = \frac{1}{2} \sum_p \alpha_p N_p.
$$

(27)

Using Eq. (4), it is easy to show that

$$
\begin{align*}
& e^{-J_1} P_p^\dagger e^{J_1} = e^{-\alpha_p} P_p^\dagger, \\
& e^{-J_1} P_p e^{J_1} = e^{\alpha_p} P_p, \\
& e^{-J_1} N_p e^{J_1} = N_p.
\end{align*}
$$

(28a)

Thus, the similarity-transformed Hamiltonian simply adjusts the coefficients of the $P_1$ and $P$ operators.

Alternatively, we can use Eq. (21) to absorb $\exp(J_1)$ into AGP. Although we have seen how this can be done, we will extract the same result in an different fashion. Notice by Eq. (5) that

$$
e^{J_1} |\text{AGP}\rangle = \frac{1}{N!} \left(e^{J_1} \Gamma^\dagger e^{-J_1}\right)^N |\text{vac}\rangle,
$$

(29)

where we used the fact that $\exp(J_1)|\text{vac}\rangle = |\text{vac}\rangle$. From the algebra, Eq. (4), it follows

$$
\begin{align*}
e^{J_1} \Gamma e^{-J_1} &= \sum_p \eta_p e^{\alpha_p} P_p^\dagger, \\
e^{J_1} \Gamma e^{-J_1} &= \sum_p \eta_p e^{-\alpha_p} P_p
\end{align*}
$$

(30a)

(30b)

from which it is clear that $e^{J_1} |\text{AGP}\rangle$ leads to $\eta_p \rightarrow \eta_p e^{\alpha_p}$ for all geminal coefficients in the ket, and $|\text{AGP}\rangle e^{-J_1}$ results in $\eta_p \rightarrow \eta_p e^{-\alpha_p}$ in the bra.

We now proceed to derive the equations for similarity transformation with $J_2$. Define $J_1(p) = \sum_q \alpha_{pq} N_p/4$ and recall $\alpha_{pp} = 0$, then as in Eq. (28) we obtain

$$
\begin{align*}
e^{-J_2} P_p^\dagger e^{J_2} &= e^{-J_1(p)} P_p^\dagger e^{J_1(p)}, \\
e^{-J_2} P_p e^{J_2} &= e^{J_1(p)} P_p e^{J_1(p)}, \\
e^{-J_2} N_p e^{J_2} &= N_p
\end{align*}
$$

(31a)

(31b)

(31c)

where we used $|J_1(p)\rangle P_p^\dagger = 0$ to symmetrically distribute the exponentials in the right hand side of Eq. (31) on either side of each operator. In fact, using these equations, we can transform any many-body operator of the form $P_p^\dagger \ldots N_{q_n} \ldots P_r$... in a symmetric way. In general, it can be shown that

$$
\begin{align*} 
& e^{-J_2} P_p^\dagger \ldots P_p^\dagger N_{q_1} \ldots N_{q_n} P_{r_1} \ldots P_{r_n} e^{J_2} = \\
& e^{\sum_{i=1}^n J_1(r_i) - J_1(p_i)} P_p^\dagger \ldots P_p^\dagger e^{\sum_{i=1}^n J_1(r_i) - J_1(p_i)}
\end{align*}
$$

(32)

To obtain a closed-form expression for the energy and the residual equations, we need to further absorb each $J_1(p)$ into AGP. Following the same steps as in Eq. (30), we can see that

$$
\begin{align*}
e^{J_1(q) - J_1(p)} \Gamma^\dagger e^{J_1(p) - J_1(q)} &= \\
= \sum \left(\eta_p e^{(\alpha_q - \alpha_p)/2}\right) P_p^\dagger = \tilde{\Gamma}^\dagger.
\end{align*}
$$

(33)
This together with Eq. (32) show that we can evaluate the expectation value of any similarity transformed operator of the form $P_p^1\ldots N_q\ldots P_p\ldots$ with $J_2$ by scaling the geminal coefficients appropriately.

The same basic steps follow for even higher-body Jastrow operators: the initial similarity-transformation of the Hamiltonian leads to modified matrix elements along with exponentials of Jastrow operators whose ranks are reduced by one. However, only $J_2$ permits easy evaluation of the remaining expectation values.

For unitary JAGP, as in Eq. (24), one follows the same steps as above but replaces $\alpha \rightarrow i\alpha$.

**D. Application to the pairing Hamiltonian**

We now show how to apply the equations above to similarly transform the pairing Eq. (2). It follows form Eq. (28) and Eq. (32) that

$$e^{-J_2}He^{J_2} = \sum_p \epsilon_p N_p - G \sum_{pq} (e^{J_1(q)-J_1(p)}P_p^1P_q^1e^{J_1(q)-J_1(p)}).$$

For every $P_p^1P_q^1$, define $\tilde{\text{AGP}}_{pq} = e^{J_1(q)-J_1(p)}\text{AGP}$, and note that $\tilde{\text{AGP}}_{pq}$ is an AGP whose geminal coefficients have been rescaled by $\tilde{\eta}_r = \eta_r e^{(\alpha_r - \alpha_r^2)/2}$ for all orbitals $r$ from Eq. (33). As such, we can obtain the energy,

$$\tilde{E} = \langle e^{-J_2}He^{J_2} \rangle = \sum_p \epsilon_p \langle \text{AGP}|N_p|\text{AGP}\rangle - G \sum_{pq} \langle \tilde{\text{AGP}}_{pq}|P_p^1P_q^1|\tilde{\text{AGP}}_{pq}\rangle.$$  

Following similar steps as above, it is straightforward to get closed-form expressions for the residual equations, Eq. (26), as well. We simply use

$$\langle N_r N_s e^{-J_2}H e^{J_2} \rangle = \sum_p \epsilon_p \langle \text{AGP}|N_p N_r N_s|\text{AGP}\rangle - G \sum_{pq} \langle \tilde{\text{AGP}}_{pq}|N_r N_s P_p^1 P_q^1|\tilde{\text{AGP}}_{pq}\rangle$$

which we can then be normal ordered in terms of $P_p^1\ldots N_q\ldots P_p\ldots$.

As outlined in Sec. II, it is a property of AGP-based RDMs that $Z_{r_{pq}}^{(1,1)} = \langle \text{AGP}|P_r|\tilde{\text{AGP}}_{pq}\rangle$ is sufficient to construct all higher-body density matrices, where the cost of building $Z_{r_{pq}}^{(1,1)}$ is $O(M^3)$. Of course the same principle applies to RDMs using $\tilde{\text{AGP}}_{pq}$, except that $\tilde{Z}_{r_{pq}}^{(1,1)} = \langle \tilde{\text{AGP}}_{pq}|N_r|\tilde{\text{AGP}}_{pq}\rangle$ needs to be evaluated for every $pq$ in the Hamiltonian, as demonstrated for Eq. (35). The cost of building $\tilde{Z}_{r_{pq}}^{(1,1)}$ for all $pq$ is $O(M^5)$ and is the leading cost of evaluating both the energy and the residuals. Without the reconstruction formulae, the cost would be $O(M^6)$.

In Fig. 1, we show the energy error and the fraction of correlation energy captured by ST-J2AGP. The correlation energy is computed as the difference with HF and the right is the total energy error.

**FIG. 1.** Comparison of energy obtained using ST-J2AGP at 12 levels, half-filled with J2-CI and var-J2AGP. The left figure is the correlation energy as the difference with HF and the right is the total energy error.
this J2CI-AGP has $O(M^6)$ scaling, or $O(M^8)$ without reconstruction formulae.

For larger systems, the over-correlation of ST-J2AGP near the recoupling regime worsens while the linear variational theory remains well-behaved. Thus, although ST-J2AGP has a lower cost than other AGP-based methods developed recently in our group, neither var-J2AGP nor ST-J2AGP yield as accurate energies for the pairing Hamiltonian. While correlators build from particle-hole excitations generally benefit from being put in an exponential operator, it is far from clear that the same is true of correlators built from Hilbert space Jastrow operators. Thus in the next section we discuss an exponential ansatz that we build from different generators.

IV. UNITARY PAIR-HOPPER CORRELATOR

An alternative way of creating an exponential ansatz based on two-body, number preserving operators with the generators of the algebra is to use a linear combination of terms like $P^*_p P_q$. Unlike the Jastrow operators discussed in Sec. III, a unitary or similarity transformation using $P^*_p P_q$ operators is not easily summable nor does it truncate naturally the way it does in CC theory. This is a direct consequence of the commutation relations Eq. (4). As a result, some form of approximation must be made. Recently, we proposed a unitary ansatz that we call unitary pair-hopper for an implementation on a quantum computer [32]. Benchmark calculations against the ground state energy of the pairing-Hamiltonian show highly accurate results.

In this section, we seek approximation schemes to this model that can be implemented affordably on a classical computer. In particular, we implement two approaches: First, we use the BCH formula and truncate the commutator expansion at some high order. Second, we apply the canonical transformation theory as studied extensively in Ref. [49–53] to this problem later in this section.

A. Truncated unitary pair-hopper

Recall the definition of the anti-Hermitian pair-hopper operator [32]

$$\mathcal{T} = \sum_{p < q} \tau_{pq} (P^*_p P_q - P^*_q P_p),$$

where $\tau_{pq}$ is the amplitude and is antisymmetric. $\mathcal{T}$ is derived from

$$K_{pq} - K^*_p K_q ^* \propto P^*_p P_q - P^*_q P_p,$$

where $K_{pq}|\text{AGP}\rangle = 0$ is the two-body killer of AGP [29]. Its exact expression is

$$K_{pq} = \eta^2_{pq} P^*_p P_q + \eta^2_{pq} P^*_q P_p + \frac{1}{2} \eta_p \eta_q (N_p N_q - N_p - N_q).$$

Notice that $e^{\mathcal{T}|\text{HF}\rangle}$ is equivalent to paired unitary CC with generalized indices (UGCC).

Define the unitary pair-hopper ansatz as $e^{\mathcal{T}|\text{AGP}\rangle}$ which we denote as uP-AGP. We want to solve for the amplitudes by minimizing the energy, that is

$$E(\tau) = \langle \text{AGP}|e^{-\mathcal{T}} H e^{\mathcal{T}}|\text{AGP}\rangle$$

$$\tau^* = \arg \min_{\tau_{pq} \in \mathbb{R}} E(\tau).$$

However, since the BCH expansion of the unitary transformed Hamiltonian does not truncate naturally, we choose to truncate the expansion at the highest order that gives us an affordable scaling, i.e. $O(M^6)$. Again, using the reconstruction formulae, the expected value of all two-body or higher density matrices can be accessed at $O(1)$ cost per element. This makes the asymptotic scaling of the commutator expansion to be dominated by the many-body contractions. For example, for a two-body Hamiltonian, the first commutator, $\{[H, \mathcal{T}]\}$ contains three-body contractions; the second commutator, $\{[H, \mathcal{T}], \mathcal{T}\}$ contains four-body contractions, and so on. By truncating the expansion at the 4th commutator, we obtain a scaling of $O(M^6)$ for the energy. In other words, we can state

$$E(\tau) \approx \frac{\langle H \rangle}{O(M^2)} + \frac{\langle \{H, \mathcal{T}\} \rangle}{O(M^4)} + \frac{1}{2!} \frac{\langle \{\{H, \mathcal{T}\}, \mathcal{T}\} \rangle}{O(M^6)}$$

$$+ \frac{1}{3!} \frac{\langle \{\{\{H, \mathcal{T}\}, \mathcal{T}\}, \mathcal{T}\} \rangle}{O(M^8)} + \frac{1}{4!} \frac{\langle \{\{\{H, \mathcal{T}\}, \mathcal{T}\}, \mathcal{T}, \mathcal{T}\} \rangle}{O(M^{10})},$$

such that addition of every term in the expansion multiplies the leading scaling by a factor of $M$. Obtaining the equations above analytically is cumbersome but not prohibitive. We use our home-built algebraic manipulator software, drudge [61], to generate this and the analytic gradient of the energy. As the result, we can compute both $E(\tau)$ and $\nabla E(\tau)$ at $O(M^6)$ cost albeit with a relatively large prefactor.

We minimize the energy as in Eq. (40) using the limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) algorithm [62] where we provide the gradients using our analytical expressions. We could alternatively find the solutions to $\nabla E(\tau) = 0$ using non-linear root-finding algorithms, but we find that the former is typically more robust when an accurate initial guess is provided.

Because truncating the commutator expansion yields a non-variational energy expression, we must take care with the minimization of the energy. This becomes particularly important in larger systems for which the numerical algorithms may go below the exact energy and eventually blow up. Fortunately, this failure mode is easy to identify as it is accompanied by extremely large amplitudes in $\tau$. To avoid this problem, we introduce an $L_2$
In contrast, our AGP based methods are well behaved in all correlation regimes.

The results for a larger system is show in Fig. 3 where we compare the accuracy of the truncated uP-AGP with ST-J_2AGP and P-CI. Although all of these methods have the same number of variational parameters, they clearly differ considerably in their accuracy. The observation that uP-AGP could be more accurate than P-CI in principle is a testament to its non-linear nature. Previously, it was shown that CI methods based on the generators of the algebra, namely J_2CI, P-CI, and K-CI, yield identical results [31]. The other advantage of the non-linear model is that it is agnostic to the linear dependencies of the ansatz. This is true for both ST-J_2AGP and the unitary pair-hopper. Although removing the linear dependencies by diagonalizing the metric can yield faster convergence, our numerical results show very little difference in the final energy.

The downside of the truncated uP-AGP is that it has a relatively large prefactor which is a result of the brute-force expansion of the BCH formula. In the next section we apply canonical transformation theory as another means for approximating the unitary transformed Hamiltonian.

\[ e^{-A}He^A = H + [H, A] + \frac{1}{2!}[[H, A], A] + \ldots \]  

### B. AGP-based canonical transformation

In canonical transformation (CT) theory, as formulated originally by Yanai and Chan [49, 50], the commutator expansion of a unitary transformed Hamiltonian,

\[ e^{-A}He^A = H + [H, A] + \frac{1}{2!}[[H, A], A] + \ldots \]  

where $\lambda > 0$ is the regularization parameter and is sufficiently small to allow a robust convergence. In effect, the regularization penalizes the energy optimization in proportion to the magnitude of the amplitudes, thereby preventing it from blowing up. Although this compromises the accuracy of final energy, we gain a smoother convergence. To find $m$ sufficiently small to allow a robust convergence. In effect, $\lambda > m$. We repeat this process until we can no longer converge or the change in energy is negligible. The convergence threshold of the gradient is set to $10^{-8}$ in our calculations.

In Fig. 2 we plot the results for the pairing Hamiltonian in a small system in which we can compare the accuracy of our truncated commutator expansion with that of exact uP-AGP which we obtain from a FCI code. As we can see in the plot, the agreement between the exact and truncated uP-AGP is excellent on the attractive side while at the same time being slightly more accurate than the linear wavefunction $\mathcal{T}|\text{AGP}$ denoted as P-CI AGP. On the repulsive side, the results for the truncated uP-AGP is comparable to P-CI and becomes more accurate only at sufficiently small $G/G_c$.

For benchmarking purposes, we also plot HF-based UGCC and UCC in Fig. 2. The results are obtained from a FCI code. As expected UCC works well in the repulsive regime, but it breaks down as $G$ approaches $G_c$. On the other hand, UGCC is remarkably more accurate than UCC, but it too eventually breaks down for large $G/G_c$. In contrast, our AGP based methods are well behaved in all correlation regimes.
where $A$ is an anti-Hermitian operator, is recursively summed by systematically replacing all higher-body operators obtained from each commutator by an approximate, lower-body operator. For example, if $A$ and $H$ are each two-body operators, one approximates $[H, A] \rightarrow [H, A]_{(1,2)}$, wherein the three-body operator resulting from $[H, A]$ is approximated in terms of one- and two-body operators. This process yields an effective Hamiltonian more and more diagonal by successively approximating higher order excitations. It can be viewed as a kind of renormalization group approach and is rooted in an earlier work by White [64] as well as flow renormalization approach by Wegner [65] and Glazek and Wilson [66]. In this section, however, we very closely follow the formalism presented in Ref. [49–53] by Yanai, Chan, and Neuscamman.

The chief difference between our approach and those of earlier papers is in the operator decomposition or downfolding scheme. Instead of invoking the cumulant decomposition of RDMs [67] as a means to get lower-body operators, we use AGP and our reconstruction formulae. Just as in cumulant decomposition, the reconstruction formulæ are statements about the RDMs, but we can extended them to operators as an approximation.

For example, consider the three-body irreducible density matrix $Z_{pq}^{(3)}$. From the reconstruction formulæ [26], it follows that

$$\langle N_p N_q N_r \rangle = \lambda_{qp} \lambda_{rp} \langle N_p \rangle + \lambda_{pq} \lambda_{qr} \langle N_q \rangle + \lambda_{pr} \lambda_{qr} \langle N_r \rangle,$$

(47)

where $\lambda_{pq} = 2\eta_{pq}^2/(\eta^2_{pq} - \eta^2_p)$. For CT, we propose to approximate $N_p N_q N_r$ by removing the expectation values to get

$$N_p N_q N_r \rightarrow \lambda_{qp} \lambda_{rp} N_p + \lambda_{pq} \lambda_{qr} N_q + \lambda_{pr} \lambda_{qr} N_r.$$  

(48)

Note that since the reconstruction formulæ apply only to irreducible RDMs, we limit our downfolding to irreducible operators, i.e. $P_p \ldots, N_q \ldots P_r \ldots$ where all indices are different. Indeed, all reducible operators (RDMs) can be written as a sum of irreducible ones.

To carry out the CT procedure, we take the unitary pair-hopper $e^{T}$ ansatz to be the generator of our transformation, and we seek an effective Hamiltonian of the

![Graph showing accuracy of the truncated unitary pair-hopper as compared with P-CI AGP and ST-J at 20 sites, half-filled. The left figure shows the correlation energy as the difference with HF and the right is the total energy error.](image)

FIG. 3. Accuracy of the truncated unitary pair-hopper as compared with P-CI AGP and ST-J at 20 sites, half-filled. The left figure shows the correlation energy as the difference with HF and the right is the total energy error.
form

$$H = \sum_p h_p N_p + \sum_{p \neq q} w_{pq} N_p N_q + \sum_{pq} v_{pq} P_p^\dagger P_q,$$  \hspace{1cm} (49)

where $w_{pq}$ and $v_{pq}$ are symmetric matrices. Note that Eq. (49) is the general form of a two-body Hamiltonian in the seniority-zero space [13]. In order to get a relatively more accurate approximation in CT, we would like to delay the downfolding as much possible. To this end, following Ref. [52], we split $\bar{H}$ such that $\bar{H}_1$ and $\bar{H}_2$ are the one- and two-body parts of the effective Hamiltonian respectively, and the downfolding is delayed until four-body operators appear. As such, one obtains

$$\bar{H}_1 = H_1 + [H_1, T] + \frac{1}{2!}[[H_1, T], T]_{(1,2)} + \frac{1}{3!}[[[H_1, T], T], T]_{(1,2)} + \ldots$$  \hspace{1cm} (50a)

$$\bar{H}_2 = H_2 + [H_2, T]_{(1,2)} + \frac{1}{2!}[[H_2, T], T]_{(1,2)} + \frac{1}{3!}[[[H_2, T], T], T]_{(1,2)} + \ldots$$  \hspace{1cm} (50b)

Similarly, to get accurate residual equations, we split $R_{pq} = R_{1,pq} + R_{2,pq}$ and evaluate the residuals associated with the one- and two-body parts recursively

$$R_{1,pq} = \langle [H_1, T_{pq}] \rangle + \frac{1}{2!}[[H_1, T_{pq}], T_{pq}]_{(1,2)} + \ldots,$$  \hspace{1cm} (51a)

$$R_{2,pq} = \langle [H_2, T_{pq}]_{(1,2)} \rangle + \frac{1}{2!}[[H_2, T_{pq}], T_{pq}]_{(1,2)} + \ldots$$  \hspace{1cm} (51b)

In both Eq. (50) and Eq. (51) we retain all two-body terms that naturally appear in the commutator expansion and absorb their coefficients into $w_{pq}$ and $v_{pq}$ in Eq. (49). All higher-body operators are downfolded directly into one-body with the intent of making the Hamiltonian more diagonal. The cost of evaluating the effective Hamiltonian and the residuals is $O(M^4)$.

Having the effective Hamiltonian and the residual equations, one can solve the amplitudes using a nonlinear root-finding algorithm. In so doing, we can provide an approximation to the Jacobian matrix as follows

$$J_{pq,rs} = \langle[[\bar{H}, T_{rs}], T_{pq}]\rangle.$$  \hspace{1cm} (52)

The cost of building the Jacobian is also $O(M^4)$ by using the reconstruction formula.

We apply the CT as described above to the pairing Hamiltonian Eq. (2). As encountered in other implementations of CT, we find that the optimization over the amplitude equations is ill-conditioned [51–53]. These numerical difficulties are attributed to near-zero eigenvalues of the Jacobian and are said to be similar to the intruder states in the second order perturbation theory [52]. Several techniques have been proposed to mitigate these difficulties [52, 53]. In our implementation, we find that shifting of the amplitude equations [53]

$$\bar{R}_{\mu} = R_{\mu} + \lambda \tau_{\mu},$$  \hspace{1cm} (53a)

$$\Rightarrow J_{\mu \nu} = J_{\mu \nu} + \lambda \delta_{\mu \nu},$$  \hspace{1cm} (53b)

where $\mu, \nu$ are flattened indices, to be the most effective. This is analogous to the regularization introduced in Eq. (42). The shifting parameter $\lambda > 0$ can be chosen ad hoc [53].

The results for 40 levels, 20 pairs is shown in Fig. 4. CT with the unitary pair-hopper on AGP (CT-uP-AGP) is considerably less expensive than the truncated uP-AGP as well as the linear P-CI AGP. However, the approximation made to higher order excitations compromises the accuracy of CT-uP-AGP. Systematic improvements

FIG. 4. Accuracy of canonical transformation with the unitary pair-hoppers, CT-uP-AGP, and that of P-CI AGP. The left figure is the correlation energy as the difference with HF and the right is the total energy error.
can be made to CT-nP-AGP by delaying the downfolding even further, but we leave these improvements for future work.

V. CONCLUSIONS

We explore several non-linear exponential ansätze on AGP in this paper. First, we study Hilbert-space Jastrow correlators on AGP. We show analytically that while a non-Hermitian JAGP exponential ansatz can in principle be made exact, practical computational limitations with non-stochastic methods compel us to make certain approximations. To this end, we extend the formalism of Ref. [46] to AGP and the \( su(2) \) algebra, and show how to similarity transform any given Hamiltonian in the seniority-zero space and solve for the amplitudes in a CC-style manner for AGP. Benchmark calculations of this ST-J2AGP on the ground state of the pairing Hamiltonian show significant improvement on AGP but the resulting energies are less accurate compared to their CI counter-part, J2-Cl AGP. The variational ansatz, vari-J2AGP, is generally more accurate than the similarity transformed one, but it is out of reach for non-stochastic methods to the best of our knowledge. The cost of ST-J2AGP is \( O(M^6) \).

In the second half of the paper, we sought an approximation to the unitary pair-hopper ansatz. Owing to its non-linear nature, exact variational calculations show uP-AGP can be considerably more accurate than the linear P-Cl AGP. However, due to its non truncating commutator expansion, some approximation must be made to implement it efficiently on a classical computer. For an approximation to uP-AGP, we expand \( \exp(-T H e^T) \) using the well-known BCH formula up to the highest affordable computational scaling \( O(M^6) \). Numerical results indicate improvement over P-Cl AGP on the attractive regime of the pairing Hamiltonian.

The downside of truncated uP-AGP is that it has a relatively large prefactor in its asymptotic scaling. To resolve this, we implement a slight variation of canonical transformation theory [52] wherein we carry out the operator decomposition from the reconstruction formulæ [26] instead of the commonly used cumulant decomposition [67]. Our CT-uP-AGP is significantly less expensive than either truncated uP-AGP or P-Cl AGP, but the approximation made to the higher order excitations noticeably compromises its accuracy.

Considering the pros and cons of each method, it is far from clear whether exponential ansätze for post-AGP methods are necessarily more advantageous than linear correlators. While there are non-linear ansätze that are more accurate than their linear counter-parts, practical considerations for implementing them could limit their scope. So far, the linear ansätze have shown remarkable accuracy, they are straightforward to implement and seem rather resilient to system size, at least in the pairing Hamiltonian.

Lastly, because the interactions are infinite-range in the pairing Hamiltonian, we intentionally avoid discussing size-extensivity in this paper. Analysis regarding size-extensivity of post-AGP methods will be reported in due time.

VI. ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation under Grant No. CHE-1762320. G.E.S. is a Welch Foundation Chair (Grant No. c-0036). A.K is thankful to Yiheng Qiu for useful discussions regarding Jastrow AGP and to Gaurav Harsha for assistance with the drudge software. We thank Rishab Dutta for letting us use his JCI-AGP code.

Appendix A: Proof for J1 CI Manifold Containing Lower-Rank JCI Manifolds over AGP

The \( J_k \) CI-AGP manifold contains the \( J_{k-1} \) CI-AGP manifold and thus all lower-rank JCI-AGP manifolds. This has been observed numerically in Ref. [31]. Here we provide a proof for the \( k = 2 \) case, which can be readily generalized to higher ranks. Along the same lines, \( J_k \) AGP does not need to include lower-rank Jastrow operators, which is shown as a corollary.

For an AGP with \( N \) pairs of electrons

\[
J_1|\text{AGP}\rangle = \frac{1}{2N} \sum_p N_p J_1|\text{AGP}\rangle
\]

\[
= \frac{1}{4N} \sum_{pq} \alpha_p N_p N_q |\text{AGP}\rangle
\]

\[
= \frac{1}{4N} \sum_{p<q} (\alpha_p + \alpha_q) N_p N_q |\text{AGP}\rangle + \frac{1}{4N} \sum_p \alpha_p N_p^2 |\text{AGP}\rangle
\]

\[
= \frac{1}{4N} \sum_{p<q} (\alpha_p + \alpha_q) N_p N_q |\text{AGP}\rangle + \frac{1}{2N} \sum_p \alpha_p N_p |\text{AGP}\rangle.
\]

It follows that

\[
J_1|\text{AGP}\rangle = \frac{1}{4} \sum_{p<q} \frac{\alpha_p + \alpha_q}{N-1} N_p N_q |\text{AGP}\rangle.
\]

Namely, \( J_1|\text{AGP}\rangle \) can be expressed as a linear combination of states in \( J_2 \) CI.

From Eq. (A1), we can show Eq. (15) by defining

\[
\alpha'_{pq} = \frac{\alpha_p + \alpha_q}{N-1} + \alpha_{pq}
\]

and

\[
J_2' = \frac{1}{4} \sum_{p<q} \alpha'_{pq} N_p N_q.
\]

Generalization to higher-rank cases is straightforward.
many-body theory development, Thesis (2018).

[62] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, SIAM J. Sci. Comput. 16, 1190 (1995).

[63] J. Nocedal and S. J. Wright, Numerical Optimization (Springer Science & Business Media, New York, NY, 2006).

[64] S. R. White, J. Chem. Phys. 117, 7472 (2002).

[65] F. Wegner, Ann. Phys. 506, 77 (1994).

[66] S. D. Glazek and K. G. Wilson, Phys. Rev. D 49, 4214 (1994).

[67] W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. 110, 2800 (1999).