

Near-exact treatment of seniority-zero ground and excited states with a Richardson-Gaudin mean-field

Charles-Émile Fecteau, Samuel Cloutier, Jean-David Moisset, Jérémy Boulay

*Département de chimie, Université Laval, Québec, Québec, G1V 0A6,
Canada*

Patrick Bultinck

Department of Chemistry, Ghent University, Gent, Belgium

Alexandre Faribault

*Université de Lorraine, CNRS, LPCT, F-54000 Nancy,
France*

Paul A. Johnson^{a)}

*Département de chimie, Université Laval, 1045 avenue de la Médecine, bureau 1220,
Québec, Québec, G1V 0A6, Canada*

(Dated: 20 April 2022)

Eigenvectors of the reduced Bardeen-Cooper-Schrieffer (BCS) Hamiltonian, Richardson-Gaudin (RG) states, are used as a variational wavefunction Ansatz for strongly-correlated electronic systems. These states are geminal products whose coefficients are solutions of non-linear equations. Previous results showed an un-physical apparent avoided crossing in ground state dissociation curves for hydrogen chains. In this contribution it is shown that each seniority-zero state of the molecular Coulomb Hamiltonian corresponds directly to an RG state. However, the seniority-zero ground state does not correspond to the ground state of a reduced BCS Hamiltonian. The difficulty is in choosing the correct RG state. The systems studied showed a clear choice and we expect it should always be possible to reason physically which state to choose.

^{a)}Electronic mail: paul.johnson@chm.ulaval.ca

I. INTRODUCTION

Most systems in quantum chemistry are weakly-correlated and thus well-described in the orbital picture. Qualitatively, the wavefunction is one Slater determinant, with small contributions from single- and double-excitations. This picture is well-understood as a mean-field of electrons, so methods built from such a mean-field, in particular Kohn-Sham density functional theory (DFT) and coupled-cluster (CC) with singles and doubles, perform quite well.¹

This picture fails for strongly-correlated systems. The physical wavefunction has important contributions from many Slater determinants, and all must be included to arrive at a qualitatively-correct result. If there are a few important Slater determinants, then complete active space self-consistent field (CASSCF) or complete active space configuration interaction (CASCI) are sufficient to solve the problem. As the number of important Slater determinants grows, these methods become much more expensive and the correct Slater determinants to include are not always obvious. Methods for strong correlation continue to be developed²⁻¹² and here we offer an alternative based on the concept of electron pairs.

Strongly-correlated systems are broadly-defined as any system that is not weakly-correlated. To make measurable progress it is necessary to subdivide these systems into different types based on what they *are*, rather than what they are not. It has been shown^{13,14} that classifying systems based on their number of unpaired electrons, their *seniority*, is a productive way to subdivide strongly-correlated systems. The first step is thus to solve the problem for systems with no unpaired electrons, otherwise stated, seniority-zero systems. Our goal is to show definitively that seniority-zero systems are effectively treated as weakly-correlated pairs of electrons.

The idea of using electron-pairs, or geminals, is an old idea.¹⁵⁻²⁰ The most general seniority-zero geminal wavefunction, the antisymmetrized product of interacting geminals (APIG),²¹⁻²⁴ is intractable for numerical computation, both variationally and by projection. Many of APIG's specific cases are however feasible, though each have drawbacks. The antisymmetrized geminal power (AGP)²⁵⁻²⁸ has a long history in quantum chemistry, and is under current study.²⁹⁻³⁵ Variationally it is feasible, though unfortunately AGP is not size-consistent using a restricted orbital basis. This prevents it from properly describing molecular dissociations. Neuscamman has shown that size-consistency may be restored at the cost of

introducing Jastrow factors.^{36–38} The antisymmetrized product of strongly-orthogonal geminals (APSG)^{39,40} is variationally feasible and dissociates molecules correctly. It is, however, difficult to assign the orbitals correctly into disjoint spaces. The antisymmetrized product of 1-reference orbital geminals (AP1roG),⁴¹ or equivalently pair-coupled-cluster doubles (pCCD),⁴² is at present the method to beat: it treats the ground state of repulsive seniority-zero systems quite well^{43–51} and scales like $\mathcal{O}(N^4)$, with N the number of orbitals. For attractive interactions, AP1roG / pCCD has been shown to fail catastrophically.⁴⁵ Systems which do not preserve seniority are also a problem⁵² but would be for any seniority-zero geminal. Ultimately, AP1roG / pCCD is a state-specific method that must be solved by projection.

The final specific cases of APIG are the eigenvectors of the reduced Bardeen-Cooper-Schrieffer (BCS)^{53–55} Hamiltonian, which we refer to as Richardson-Gaudin (RG)^{56–59} states. These states are geminal product wavefunctions whose geminal coefficients are solutions of a set of non-linear equations. Scalar products and reduced density matrices (RDM) of RG states are easily evaluated, and thus variational optimizations with RG states are feasible. Recently,⁶⁰ we performed variational calculations for hydrogen chain dissociations using RG *ground* states, i.e. ground states of a reduced BCS Hamiltonian, as a wavefunction Ansatz. We observed un-physical behaviour which was quite disappointing. To judge whether this was the fault of the RG wavefunction form, we performed variational calculations with RG states that were not eigenvectors of a reduced BCS Hamiltonian, so-called off-shell RG states.⁶¹ Off-shell RG states have much more variational freedom, but as they are not eigenvectors of a physical model, they are not computationally tractable. Our results showed that for H_4 and H_6 the off-shell RG results matched APIG to at least 0.01 milliHartree. For H_8 , we were forced to relax our convergence criterion, but the agreement remained quite clear. These results show that the form of the RG wavefunction was not the problem, but perhaps the specific RG state used. In this contribution, we look at many RG states and easily find those optimal for Hydrogen chains, as well as N_2 , which are prototypical examples of strong-correlation in Chemistry. The agreement with the best seniority-zero wavefunction possible, doubly-occupied configuration interaction (DOCI),^{62–64} is excellent everywhere though deviates linearly in the weakly-correlated limit.

In section II we very briefly outline the structure of RG states, the approach for numerically obtaining them, and the energy functional to be evaluated. In section III variational

calculations are performed for molecular systems with RG states. In particular, for H_4 a reduced BCS Hamiltonian is optimized for each possible RG state. The correct RG state is found, and with that reduced BCS Hamiltonian, all the RG states are computed and found to match each of the molecular DOCI states. For H_6 and H_8 similar results are obtained, and for N_2 the optimal RG state matches quite well with DOCI.

II. THEORY

In this section we briefly outline the necessary properties of RG states. In particular, we summarise a numerical procedure for obtaining them and expressions for their reduced density matrix (RDM) elements. This leads to an energy functional for the molecular Coulomb Hamiltonian, whose variables are the parameters defining the reduced BCS Hamiltonian.

A. Reduced BCS Hamiltonian and RG states

RG states are geminal product wavefunctions with a particular structure. They describe M pairs among N spatial orbitals built from the pair representation of $su(2)$

$$S_i^+ = a_{i\uparrow}^\dagger a_{i\downarrow}^\dagger, \quad S_i^- = a_{i\downarrow} a_{i\uparrow}, \quad S_i^z = \frac{1}{2} \left(a_{i\uparrow}^\dagger a_{i\uparrow} + a_{i\downarrow}^\dagger a_{i\downarrow} - 1 \right) \quad (1)$$

where $a_{i\uparrow}^\dagger$ ($a_{i\downarrow}$) creates (removes) an up-spin (down-spin) electron in spatial orbital i . The pair operators satisfy the structure

$$[S_i^+, S_j^-] = 2\delta_{ij} S_i^z \quad (2a)$$

$$[S_i^z, S_j^\pm] = \pm\delta_{ij} S_i^\pm. \quad (2b)$$

Rather than S_i^z , it is productive to use the number operator

$$\hat{n}_i = 2S_i^z + 1 = a_{i\uparrow}^\dagger a_{i\uparrow} + a_{i\downarrow}^\dagger a_{i\downarrow} \quad (3)$$

which counts the number of electrons in spatial orbital i .

The reduced BCS Hamiltonian

$$\hat{H}_{BCS} = \frac{1}{2} \sum_{i=1}^N \varepsilon_i \hat{n}_i - \frac{g}{2} \sum_{i,j=1}^N S_i^+ S_j^-, \quad (4)$$

expresses competition between aufbau filling of the spatial orbitals and an isotropic pair scattering between them. It is defined by its single-particle energies $\{\varepsilon\}$ and the pairing strength g . The RG states

$$|\{u\}\rangle = S^+(u_1)S^+(u_2)\dots S^+(u_M)|\theta\rangle \quad (5)$$

are built from the geminals,

$$S^+(u) = \sum_{i=1}^N \frac{S_i^+}{u - \varepsilon_i}, \quad (6)$$

which create pairs de-localized over the spatial orbitals. RG states are the eigenvectors of the reduced BCS Hamiltonian (4) provided that the set of complex numbers $\{u\}$, the *rapidities*, are solutions of Richardson's equations

$$\frac{2}{g} + \sum_{i=1}^N \frac{1}{u_a - \varepsilon_i} + \sum_{b(\neq a)=1}^M \frac{2}{u_b - u_a} = 0. \quad (7)$$

This construction is a particular example of the algebraic Bethe Ansatz (ABA).⁶⁵⁻⁶⁷ An introduction in terms of electrons is presented in refs.^{68,69} Our approach is to use RG states as a variational Ansatz. To do so, we optimize the parameters $\{\varepsilon\}$ and g defining the reduced BCS Hamiltonian. The rapidities are not free. They must be obtained from Richardson's equations.

Many algorithms exist to solve Richardson's equations,⁷⁰⁻⁷⁴ though most struggle near critical points, where two rapidities are exactly equal to one of the single-particle energies ε . The fastest and most robust approach^{75,76} involves changing to so-called *eigenvalue-based variables* (EBV)

$$U_i = \sum_{a=1}^M \frac{g}{\varepsilon_i - u_a} \quad (8)$$

which satisfy the coupled non-linear equations

$$U_i^2 - 2U_i - g \sum_{k(\neq i)=1}^N \frac{U_k - U_i}{\varepsilon_k - \varepsilon_i} = 0, \quad \forall i = 1, \dots, N. \quad (9)$$

These equations are much easier to solve numerically, as there are no longer the divergences in the denominators for the rapidities. Taking the sum of Richardson's equations, one can see that the sum of the EBV is a constant, in particular it is twice the number of pairs

$$\sum_{i=1}^N U_i = 2M. \quad (10)$$

It is essential to enforce (10) as an additional constraint so that there are $N + 1$ equations defining the N variables $\{U\}$. Otherwise, the solutions *will* cross into different sectors of M .

When $g = 0$, the reduced BCS Hamiltonian simplifies to aufbau filling, and the equations (9) decouple producing the clear solution

$$U_i (U_i - 2) = 0, \quad (11)$$

so that each EBV is either zero or two. Since the sum of the EBV is $2M$, M of the EBV are two, and the remaining $N - M$ are zero. The complete set of $\binom{N}{M}$ solutions are obtained from the possible choices of assigning the values of zero or two to the EBV, and the states are closed-shell Slater determinants. By convention^{75–78} the RG states are labelled as a string of 1's and 0's based on the choice of *pair*-occupations at $g = 0$: a doubly-occupied spatial orbital is labelled “1” while empty spatial orbitals are labelled “0” (ordered energetically). We refer to such a choice of 1's and 0's as a *bitstring* or simply a state. The equations (9) are solved by choosing a solution at $g = 0$ and evolving g slowly to its final value. Remarkably, each state evolves continuously, so it is unambiguous to label an RG state at arbitrary g based on its occupations at $g = 0$. To be clear, the bitstrings refer specifically to the closed-shell Slater determinant solutions at $g = 0$. At non-zero g , the RG states are *not* Slater determinants but the bitstrings still label the RG states uniquely. Individual RG states can, and do, cross though the ground state *of the reduced BCS Hamiltonian* is always the state labelled $1 \dots 10 \dots 0$, while the highest excited state is always $0 \dots 01 \dots 1$. For example, a 4-site system with 2 pairs has the states 1100 (which is the ground state at all values of g), 1010, 1001, 0110, 0101, and 0011 (which is the highest energy state at all values of g). The labelling of the states suggests an aufbau principle, though as curves cross, the ordering of states does not respect the principle at all couplings. As we have shown,⁷⁹ the dominant couplings between two RG states occur for “single-pair” excitations, with double-pair excitations having non-zero contributions. Past doubles, the couplings go to zero quite quickly. A similar conclusion has been drawn in ref.⁷⁸

B. Solving Richardson's equations

We will briefly summarize the approach^{75,76} while highlighting details that we have found to be important. The solver has two components: a solver for the EBV equations (9) and a

root-finding procedure to extract the rapidities $\{u\}$. Unfortunately, the expressions for RDM elements require the rapidities, so substantial effort will be devoted to obtaining them. Scalar products and 1-RDM expressions are known in terms of the EBV,⁸⁰ but 2-RDM expressions at present are not.

Solving the EBV equations begins by choosing a bitstring and assigning the corresponding solution at $g = 0$ based on (11). With a set of EBV, a step δg is taken by first expanding the EBV as a Taylor series in g to a particular order, then refining the solution with a Newton-Raphson procedure. Of course the solution could be evolved to the final g by only solving with Newton-Raphson at each step, but the Taylor series allows one to take much larger steps while maintaining the accuracy. The terms in the Taylor series are easily calculated. They are solutions of sets of over-determined linear equations who all share the same matrix. In particular, the first-derivatives of the EBV with respect to g , subject to the constraints (10) are the solutions of the linear equations

$$A \frac{\partial \mathbf{U}}{\partial g} = \mathbf{r}_1 \quad (12)$$

where the $(N + 1) \times N$ matrix A has elements

$$A_{ij} = \begin{cases} 2U_i - 2 + \sum_{k(\neq i)=1}^N \frac{g}{\varepsilon_k - \varepsilon_i}, & i = j \\ -\frac{g}{\varepsilon_j - \varepsilon_i}, & i \neq j \end{cases}, \quad 1 \leq i, j \leq N, \quad (13)$$

$$A_{N+1,j} = 1$$

and the elements of the first-order right-hand-side (RHS) are

$$(\mathbf{r}_1)_i = \sum_{k(\neq i)=1}^N \frac{U_k - U_i}{\varepsilon_k - \varepsilon_i}, \quad 1 \leq i \leq N \quad (14)$$

$$(\mathbf{r}_1)_{N+1} = 0.$$

It can be shown by induction that the p th derivatives $\frac{\partial^p \mathbf{U}}{\partial g^p}$ satisfy the set of linear equations

$$A \frac{\partial^p \mathbf{U}}{\partial g^p} = \mathbf{r}_p \quad (15)$$

with the elements of the p th RHS being

$$(\mathbf{r}_p)_i = p \sum_{k(\neq i)=1}^N \frac{\frac{\partial^{p-1} U_k}{\partial g^{p-1}} - \frac{\partial^{p-1} U_i}{\partial g^{p-1}}}{\varepsilon_k - \varepsilon_i} - \sum_{m=1}^{p-1} \binom{p}{m} \frac{\partial^m U_i}{\partial g} \frac{\partial^{p-m} U_i}{\partial g^{p-m}}, \quad 1 \leq i \leq N \quad (16)$$

$$(\mathbf{r}_p)_{N+1} = 0.$$

Each set of linear equations is solved by QR-factorizing the matrix $A = QR$, and solving the linear equations

$$R \frac{\partial^p \mathbf{U}}{\partial g^p} = Q^\dagger \mathbf{r}_p \quad (17)$$

by backward substitution. This gives the least-squares solution to each set of linear equations, though physically we know there is always an “exact” solution. Since the matrix A is common, only one QR-factorization is required to compute the Taylor series. How many terms are required? While it may not be optimal, we have seen that 4th-order is more than sufficient.

With a Taylor-series approximation to the EBV at $g + \delta g$, a Newton-Raphson solution quickly converges the EBV to higher accuracy. The Jacobian of the EBV equations is the same matrix A (13), and is once-again factored as $A = QR$, to iteratively solve the solutions of equations

$$R(\mathbf{U}_{n+1} - \mathbf{U}_n) = -Q^\dagger \mathbf{f}(\mathbf{U}_n) \quad (18)$$

for the change $\mathbf{U}_{n+1} - \mathbf{U}_n$. Unfortunately, the Jacobian, and hence its QR-factorization, must be computed at each iteration. The RHS is the EBV equations (9) along with the sum rule (10) evaluated at the n th iteration

$$f_i(\mathbf{U}_n) = \begin{cases} U_{i,n}^2 - 2U_{i,n} - g \sum_{k(\neq i)=1}^N \frac{U_{k,n} - U_{i,n}}{\varepsilon_k - \varepsilon_i}, & 1 \leq i \leq N \\ \sum_{k=1}^N U_{k,n} - 2M, & i = N + 1 \end{cases} \quad (19)$$

Thus, each step from $g = 0$ to g consists of a Taylor series approximation followed by a Newton-Raphson solution. This solution works quite well, and its computational cost scales with the number of steps required, which is in itself a problem. To ensure that the solutions are correctly and unambiguously followed, the initial step from $g = 0$ must be smaller than the smallest difference in single particle energies $\Delta\varepsilon$. If there are two ε that are close in energy, this means a *very large* number of steps will be required. We have adopted a dynamic step size algorithm to minimize the number of steps: a step δg is attempted, but is rejected if either the terms computed in the Taylor series grow in size or if after the Newton-Raphson solution the change in the EBV with respect to the previous step is too large. (While we have not performed exhaustive optimization, a 25% change in the norm of the EBV is judged to be too large. Tighter criteria do not appear to provide any

benefit.) A rejected step is reattempted with $\frac{1}{2}\delta g$. With the dynamic step-size procedure, the number of steps appears to scale logarithmically with g : the growth in the number of steps required for a half-filled 40-site reduced BCS Hamiltonians with equally-spaced levels is shown in Figure 1. In section III, it will be pertinent to employ the state with alternating occupations 1010...10 which we will call the Néel state. In the pair representation of $\text{su}(2)$, this state (at zero coupling) corresponds to alternating pair-empty occupations of the spatial orbitals. In a $\frac{1}{2}$ -spin representation of $\text{su}(2)$ this state is the up-down Néel ordering.

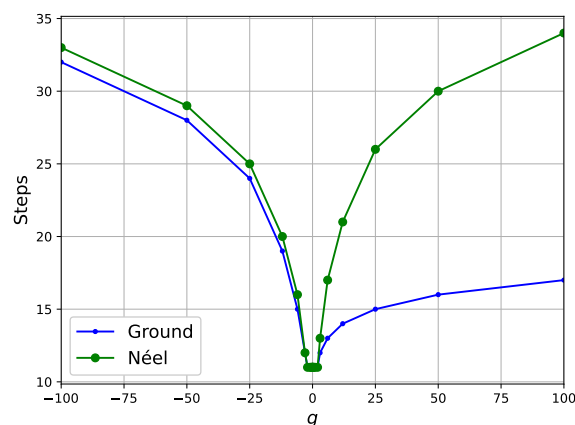


FIG. 1: Number of steps required to solve EBV equations, for the ground-state and the Néel state, with dynamic step-size algorithm for a reduced BCS Hamiltonian of 40 equally spaced single-particle energies and 20 pairs.

With a solution for the EBV, the rapidities are extracted by the procedure of ref.⁷⁶ Consider a polynomial $P(z)$, whose roots are the rapidities

$$P(z) = \prod_{a=1}^M (z - u_a). \quad (20)$$

It is straightforward to show that the logarithmic derivative of $P(z)$ evaluated at the single-particle energies gives the EBV scaled by g

$$\frac{1}{g}U_i = \frac{P'(\varepsilon_i)}{P(\varepsilon_i)}. \quad (21)$$

Further, Richardson's equations (7) are equivalent to a 2nd-order ordinary differential equation (ODE) for $P(z)$:

$$P''(z) - F(z)P'(z) + G(z)P(z) = 0, \quad (22)$$

with

$$F(z) = \frac{2}{g} + \sum_{i=1}^N \frac{1}{(z - \varepsilon_i)} \quad (23)$$

$$G(z) = \frac{1}{g} \sum_{i=1}^N \frac{U_i}{(z - \varepsilon_i)}. \quad (24)$$

Now it should be clear why the rapidities are special: they are the roots of an orthogonal polynomial. To obtain them, the polynomial $P(z)$ is Lagrange-interpolated on a grid of $M+1$ points $\{z_a\}$ which reduces the ODE (22) to a set of linear equations for the weights of $P(z)$ in the Lagrange basis. This yields a representation of $P(z)$ which solves (22), whose roots are extracted one at a time by Laguerre's method with deflation. This procedure works very well provided that the chosen grid is a reasonable guess for the rapidities. This means that an initial grid chosen at $g = 0$ can be evolved by solving for the rapidities at intermediate steps in the EBV solution. It is not necessary to solve for the rapidities at each step. The grid evolves well enough by solving at every other step of the EBV solution. In refs.^{76,78} the authors point out that if the rapidities are only required at a large g , then a physical argument can be made to estimate the grid-points, eliminating intermediate calculations. Rather than Lagrange interpolation, we considered using Chebyshev polynomials as for real-valued functions Chebyshev interpolation seems to be the unambiguously correct approach.⁸¹ We abandoned this idea rather quickly as the rapidities generally occur in complex-conjugate pairs, for which Chebyshev interpolation is much less efficient.

The polynomial $P(z)$ is written in the basis of Lagrange polynomials with the grid $\{z_a\}$ as

$$P(z) = l(z) \sum_{a=1}^{M+1} \frac{w_a}{z - z_a} \quad (25)$$

with $\{w\}$ the weights, and

$$l(z) = \prod_{a=1}^{M+1} (z - z_a). \quad (26)$$

The first derivative of $P(z)$ is easily evaluated

$$P'(z) = l'(z) \sum_{a=1}^{M+1} \frac{w_a}{z - z_a} - l(z) \sum_{a=1}^{M+1} \frac{w_a}{(z - z_a)^2}, \quad (27)$$

and simplifies when divided by $l(z)$

$$\frac{P'(z)}{l(z)} = \sum_{a=1}^{M+1} \sum_{b(\neq a)=1}^{M+1} \frac{w_a}{(z - z_a)(z - z_b)}. \quad (28)$$

Likewise, the second derivative scaled by $l(z)$ is simple

$$\frac{P''(z)}{l(z)} = \sum_{a=1}^{M+1} \sum_{b(\neq a)=1}^{M+1} \sum_{c(\neq a,b)=1}^{M+1} \frac{w_a}{(z - z_a)(z - z_b)(z - z_c)}. \quad (29)$$

Substituting these expressions into the ODE (22), and evaluating at the grid-points gives a set of linear equations for the weights $\{w\}$

$$\sum_{b(\neq a)=1}^{M+1} \sum_{c(\neq a,b)=1}^{M+1} \frac{w_a + 2w_b}{(z_a - z_b)(z_a - z_c)} - F(z_a) \left(\frac{w_a + w_b}{z_a - z_b} \right) + G(z_a)w_a = 0, \quad 1 \leq a \leq M + 1, \quad (30)$$

along with a normalization on the weights

$$\sum_{a=1}^{M+1} w_a = 1. \quad (31)$$

Taken together this is a system of over-determined linear equations

$$B\mathbf{w} = \mathbf{b} \quad (32)$$

for the weights which is again solved by QR factorization, with $B = QR$ as

$$R\mathbf{w} = Q^\dagger \mathbf{b}. \quad (33)$$

The $(M + 2) \times (M + 1)$ matrix B has elements

$$B_{ab} = \begin{cases} \sum_{c(\neq a)=1}^{M+1} \sum_{d(\neq a)}^{M+1} \frac{1}{(z_a - z_c)(z_a - z_d)} - \sum_{c(\neq a)=1}^{M+1} \frac{F(z_a)}{z_a - z_c} + G(z_a), & a = b \\ \sum_{c(\neq a,b)=1}^{M+1} \frac{2}{(z_a - z_b)(z_a - z_c)} - \frac{F(z_a)}{z_a - z_b}, & a \neq b \end{cases}, \quad (34)$$

in the first $M + 1$ rows, while the last row is

$$B_{M+2,b} = 1. \quad (35)$$

The RHS $\mathbf{b}$ is $M + 1$ zeros followed by a single entry 1. Solving the linear equations (33) gives a representation of the polynomial $P(z)$ which satisfies the ODE (22). A single root α

may be found with Laguerre's method, and the polynomial is deflated by removing the grid point z_α closest to the root. In the Lagrange polynomial basis, this amounts to removing the weight w_α corresponding to the root, and causes the other weights to be modified

$$w_b \rightarrow w_b \frac{z_b - z_\alpha}{z_b - \alpha}. \quad (36)$$

The new weights represent a polynomial of the original rank minus one, and the procedure may be repeated until all roots are found. Provided the choice of grid is "close enough" to the roots, this method works quite well.

The grid consists of $M+1$ points. At $g=0$, we assign M of them to be the single-particle energies of the occupied levels (corresponding to which EBV are non-zero) plus some small random noise, while the last grid-point is chosen as 10 times the largest single-particle energy. At each iteration the grid is updated to the rapidities from the previous iteration, unless the difference between a grid point and a previous rapidity is less than a threshold. As g increases, many rapidities will stabilize and this threshold is imposed to prevent the matrix B from having zero denominators. The final grid-point is updated to be 10 times the absolute value of the largest rapidity.

Finally, with a set of rapidities at the desired g , a Newton-Raphson routine applied directly to Richardson's equations (7) can be used to refine the rapidities to higher precision. The whole construction could thus be viewed as a way of pre-conditioning a Newton-Raphson solver at the final g . With this approach we were able to obtain rapidities to double precision.

C. Energy functional

We want to solve the Coulomb Hamiltonian for molecules

$$\hat{H}_C = \sum_{i,j=1}^N h_{ij} \sum_{\sigma} a_{i\sigma}^{\dagger} a_{j\sigma} + \frac{1}{2} \sum_{i,j,k,l=1}^N V_{ijkl} \sum_{\sigma\tau} a_{i\sigma}^{\dagger} a_{j\tau}^{\dagger} a_{l\tau} a_{k\sigma} \quad (37)$$

in which σ and τ represent spin labels, and the 1- and 2-electron integrals are computed in a basis $\{\phi\}$

$$h_{ij} = \int d\mathbf{r} \phi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right) \phi_j(\mathbf{r}) \quad (38)$$

$$V_{ijkl} = \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\phi_i^*(\mathbf{r}_1) \phi_j^*(\mathbf{r}_2) \phi_k(\mathbf{r}_1) \phi_l(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (39)$$

RG states have zero seniority, meaning that the energy of (37) computed with an RG state $|\{u\}\rangle$

$$E[\{\varepsilon\}, g] = 2 \sum_{i=1}^N h_{ii} \gamma_i + \sum_{i=1}^N \sum_{j(\neq i)=1}^N (2V_{ijij} - V_{ijji}) D_{ij} + \sum_{i,j=1}^N V_{iijj} P_{ij} \quad (40)$$

depends only on the 1-RDM elements γ_i along with the non-zero elements of the 2-RDM: the diagonal-correlation function D_{ij} and the pair-correlation function P_{ij} . Explicitly, these elements are

$$\gamma_i = \frac{1}{2} \frac{\langle \{u\} | \hat{n}_i | \{u\} \rangle}{\langle \{u\} | \{u\} \rangle} \quad (41a)$$

$$D_{ij} = \frac{1}{4} \frac{\langle \{u\} | \hat{n}_i \hat{n}_j | \{u\} \rangle}{\langle \{u\} | \{u\} \rangle} \quad (41b)$$

$$P_{ij} = \frac{\langle \{u\} | S_i^+ S_j^- | \{u\} \rangle}{\langle \{u\} | \{u\} \rangle}. \quad (41c)$$

It is important to note that as written, D_{ii} and P_{ii} correspond to the same element of the 2-RDM. As a convention, we assign this element to $P_{ii} = \gamma_i$ while taking $D_{ii} = 0$. The energy (40) will be minimized with respect to the single-particle energies $\{\varepsilon\}$ and the pairing strength g . Again, a particular choice of $\{\varepsilon\}$ and g dictates the values of the rapidities $\{u\}$ through Richardson's equations (7).

The RDM elements for RG states have been written many times.^{77,78,82–84} RDM elements of any order can be evaluated with the derivatives of the rapidities with respect to the single-particle energies, which are obtained as solutions of the linear equations

$$G \frac{\partial \mathbf{u}}{\partial \varepsilon_k} = b_k \quad (42)$$

with G the Gaudin matrix

$$G_{ab} = \begin{cases} \sum_{i=1}^N \frac{1}{(u_a - \varepsilon_i)^2} - \sum_{c(\neq a)=1}^M \frac{2}{(u_a - u_c)^2}, & a = b \\ \frac{2}{(u_a - u_b)^2}, & a \neq b. \end{cases} \quad (43)$$

In particular, the normalized 1-RDM elements are

$$\gamma_k = \sum_{a=1}^M \frac{\partial u_a}{\partial \varepsilon_k} \quad (44)$$

while the 2-RDM elements are, for $i \neq j$,

$$D_{ij} = \sum_{a=1}^M \sum_{b=a+1}^M \frac{(u_a - \varepsilon_i)(u_b - \varepsilon_j) + (u_a - \varepsilon_j)(u_b - \varepsilon_i)}{(\varepsilon_i - \varepsilon_j)(u_b - u_a)} \left(\frac{\partial u_a}{\partial \varepsilon_i} \frac{\partial u_b}{\partial \varepsilon_j} - \frac{\partial u_a}{\partial \varepsilon_j} \frac{\partial u_b}{\partial \varepsilon_i} \right) \quad (45)$$

$$P_{ij} = \sum_{a=1}^M \frac{(u_a - \varepsilon_i)}{(u_a - \varepsilon_j)} \frac{\partial u_a}{\partial \varepsilon_i} - 2 \sum_{a=1}^M \sum_{b=a+1}^M \frac{(u_b - \varepsilon_i)(u_a - \varepsilon_i)}{(\varepsilon_i - \varepsilon_j)(u_b - u_a)} \left(\frac{\partial u_a}{\partial \varepsilon_i} \frac{\partial u_b}{\partial \varepsilon_j} - \frac{\partial u_a}{\partial \varepsilon_j} \frac{\partial u_b}{\partial \varepsilon_i} \right). \quad (46)$$

Thus to compute the 2-RDM elements one must solve the N sets of linear equations (42). As G is square, it is cheapest to use LU-factorization.⁸⁵ Each set of linear equations can be solved with cost $\mathcal{O}(M^2)$, so that the N sets have a cost of $\mathcal{O}(NM^2)$. The scaling bottleneck comes from evaluating the 2-RDM: there are $\mathcal{O}(N^2)$ elements, and each requires a summation over M^2 objects. The energy (40) can thus be evaluated with $\mathcal{O}(N^2M^2)$ scaling.

D. Procedure

As there is certainly chance for confusion, it is essential to state explicitly what will be done. The goal is to find which RG bitstring is best for molecular systems.

- First, the molecular Coulomb Hamiltonian (37) is constructed. The 1- and 2-electron integrals are in the orbitals optimized for the DOCI ground state.
- Next, a bitstring is chosen which represents a particular RG state.
- For the chosen bitstring, the energy functional (40) is minimized with respect to the parameters $\{\varepsilon\}$ and g . This defines a reduced BCS Hamiltonian whose particular state labelled by the chosen bitstring is the variational approximation to the ground state of the Coulomb Hamiltonian.
- This procedure is repeated for many bitstrings to establish which best minimizes the energy of the Coulomb Hamiltonian.

It is clear that this procedure is in general intractable as there are $\binom{N}{M}$ bitstrings. As will be seen in III the choice of bitstring can be reasoned physically.

III. NUMERICAL RESULTS

A. Preliminaries

Dissociation curves were computed for symmetric hydrogen chains as well as molecular nitrogen in the minimal basis STO-6G. In a given set of orbitals, RG is strictly a variational approximation to DOCI. The best possible case for RG is orbital-optimized (OO)-DOCI, and thus all presented data are in the OO-DOCI orbitals originally computed for ref.⁶⁰ using a multi-configurational self-consistent field (MCSCF) in the complete doubly-occupied space with GAMESS (US).⁸⁶ It is to be emphasized that all the DOCI curves are computed with the orbitals optimized for the DOCI ground state. DOCI excited states, particularly near equilibrium, benefit from state specific orbital-optimization^{87,88} while enforcing orthogonality to the ground state. Full configuration interaction (FCI) results, also from ref.⁶⁰ computed with psi4,^{89,90} demonstrate that for the hydrogen chain dissociations OO-DOCI is close to exact while for molecular nitrogen it is not. Restricted Hartree-Fock (RHF) results were computed with Gaussian 16.⁹¹

The variational optimization has room for improvement. At present the variables $\{\varepsilon\}$ and g are pre-conditioned with the covariance matrix adaptation evolution strategy (CMA-ES)⁹² and optimized with the Nelder-Mead simplex algorithm.⁹³ This is a rather brute force approach which is acceptable as the present results are proof-of-principle.

It is necessary to impose a consistency check on the RDM expressions. The energy of the reduced BCS Hamiltonian (4) can be computed in several different manners. It is the sum of the rapidities

$$E_{BCS} = \sum_{a=1}^M u_a, \quad (47)$$

it has an expression in terms of the EBV

$$E_{BCS} = \frac{g}{2}M(M - N - 1) + \frac{1}{2} \sum_{i=1}^N \varepsilon_i U_i, \quad (48)$$

and it can be evaluated directly

$$E_{BCS} = \sum_{i=1}^N \varepsilon_i \gamma_i - \frac{g}{2} \sum_{i,j=1}^N P_{ij}. \quad (49)$$

The conditions (47) and (48) are verified in the solver for Richardson's equations. At each energy evaluation in the variational optimization, the reduced BCS energy is computed with both (47) and (49). If the difference is larger than 1×10^{-6} Hartree the point is rejected. This loss in precision occurs when two single-particle energies come too close to one another, and could be avoided using the degenerate version of the EBV equations solver.⁷⁶ This will be done in a future contribution as at present the results converge quite well to the correct answer.

B. H_4 Dissociation

In ref.⁶⁰ we observed that the RG ground state does not describe hydrogen chain dissociation correctly at intermediate distances. Near the minimum, the treatment is acceptable, and at dissociation the treatment is correct, but in the middle there is an unphysical local maximum. Similar behaviour for H_8 was observed in ref.¹³ We have recently shown⁶¹ that this was not a failure of the RG wavefunction form, but must be due to the choice of using only RG ground states. We therefore look at all possible RG states to choose the best. For H_4 we consider the $\binom{4}{2}$ possible states and variationally optimize parameters for each. The results are presented in figure 2. As we choose a bitstring (state) and optimize the

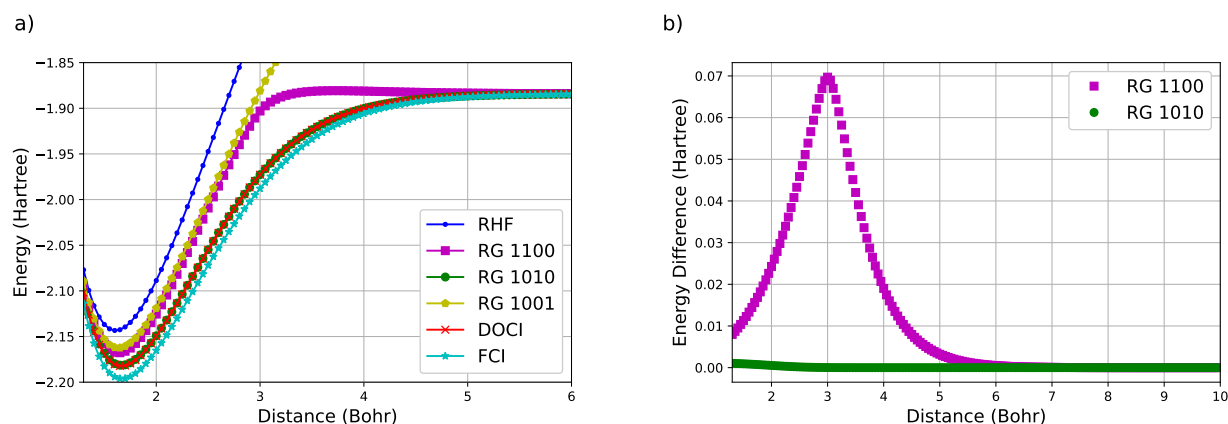


FIG. 2: (a) Bond dissociation curves for H_4 . (b) Deviations from DOCI for the 1100 and the 1010 RG states. All results computed with the STO-6G basis set in the basis of OO-DOCI orbitals.

parameters $\{\varepsilon\}$ and g , not all bitstrings are unique: the ground state of a reduced BCS

Hamiltonian with a negative (repulsive) interaction is equivalent to the highest excited state of a different reduced BCS Hamiltonian with a positive (attractive) interaction. As a result, there are only three distinct choices since with this equivalence $1100 = 0011$, $1010 = 0101$, and $1001 = 0110$. Variational calculations were performed for all six choices, but as the results are doubly-degenerate, only one set are plotted.

It is clear that the 1010 state is qualitatively correct, while the others are not. The 1001 state reproduces the AGP results of ref.⁶¹ while the 1100 state corresponds to the H_8 result observed in ref.¹³ Energy differences with respect to DOCI are presented in figure 2b. We expected that the optimal choice of state would differ at the minimum and at dissociation, but this does not appear to be the case: the 1010 state is optimal everywhere. Can the choice 1010 be reasoned physically? To do so, we must look at the parameters defining the optimal reduced BCS Hamiltonian for the state 1010, which are shown in figure 3. The

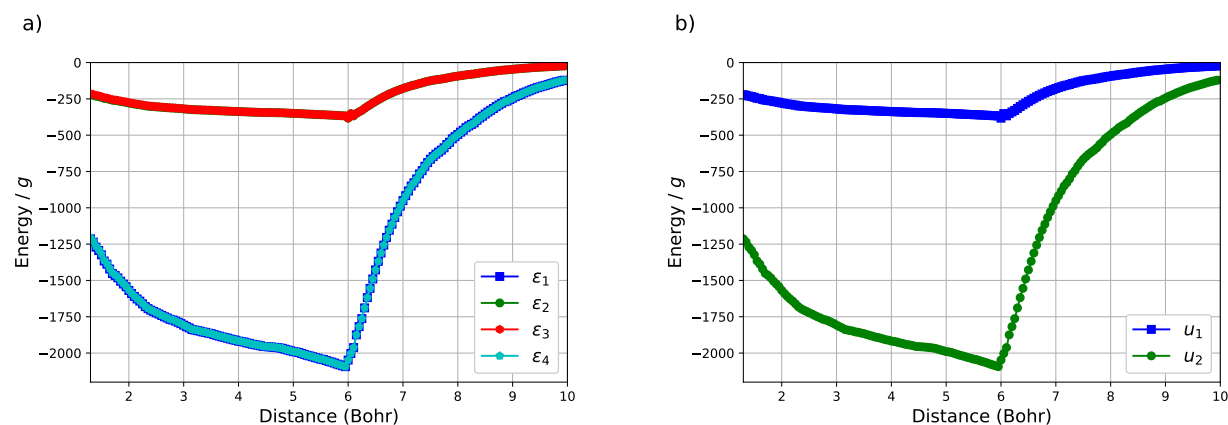


FIG. 3: (a) Optimal reduced BCS parameters $\{\epsilon\}$ scaled by $|g|$ for the 1010 RG state. (b) Rapidities of the 1010 RG state.

single-particle energies $\{\epsilon\}$ scaled by the absolute value of the pairing interaction are plotted in figure 3a. The interaction g is always repulsive. The single-particle energies are separated into two groups, with both members of each group being near-degenerate, though distinct. The rapidities corresponding to the 1010 state of this Hamiltonian are shown in figure 3b. The rapidities are each trapped between the near-degenerate single-particle energies, i.e. $\epsilon_1 < u_2 < \epsilon_4$ and $\epsilon_2 < u_1 < \epsilon_3$. The parameters break the system into two sub-systems and the bitstring 1010 puts 1 rapidity in each sub-system. The individual subsystems correspond to H_2 dissociations, which the RG state 10 treats exactly.⁶⁰ The OO-DOCI orbitals, shown

in ref.⁹⁴, correspond to individual H_2 systems. There is a cusp in the optimal parameters near $r = 6.0$ bohr, though this is not reflected in the energy of the 1010 state applied to the Coulomb Hamiltonian for H_4 . In figure 2a) the axis is cut near $r = 6.0$ in order to better show behaviour of the other curves near the minimum. The 1010 result is smooth to $r = 10.0$ bohr. The non-zero RDM elements for the 1010 state are shown in figure 4. All

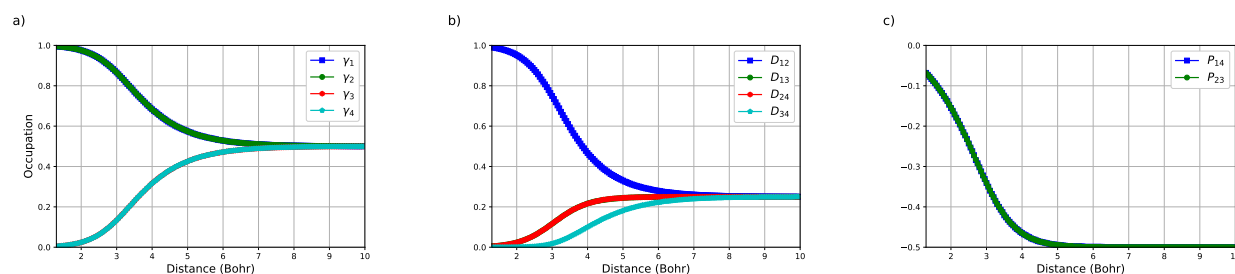


FIG. 4: (a) 1-RDM elements of the 1010 RG state. (b) Non-zero D_{ij} elements of the 1010 RG state. (c) Non-zero P_{ij} elements of the 1010 RG state.

are smooth all the way to dissociation.

We have chosen the 1010 state, and optimized the parameters $\{\varepsilon\}$ and g to minimize the energy of the Coulomb Hamiltonian. The optimized parameters define a reduced BCS Hamiltonian, for which all the eigenvectors are known as the distinct solutions of Richardson's equations. Each solution was computed and their corresponding Coulomb Hamiltonian energies are plotted in figure 5. It should be highlighted that the DOCI states are in the orbitals optimized for the DOCI ground state, and are labelled by their energetic ordering. The 3rd and 5th DOCI excited states have an avoided crossing near $r = 2.45$ bohr. As can be seen, the RG states correspond *directly* to the DOCI states. The difference in energy between the RG states and the DOCI states are plotted in figure 6. The 1010 state is essentially the same as the DOCI ground state everywhere except near the minimum. There, the system is weakly-correlated, the RG states are more or less Slater determinants, and a short perturbative expansion would yield the correct energy. It is remarkable how well the RG states match the DOCI states. They do not capture the avoided crossing near $r = 2.45$, but this would be asking too much. The RG states are computed from the parameters defining the optimal ground state mean-field. Avoided crossings are "correlation" effects: to capture them, we would need to diagonalize the Coulomb Hamiltonian matrix in the basis of RG states. There is also deviation from the DOCI states around $r = 6.0$ bohr, where the cusp

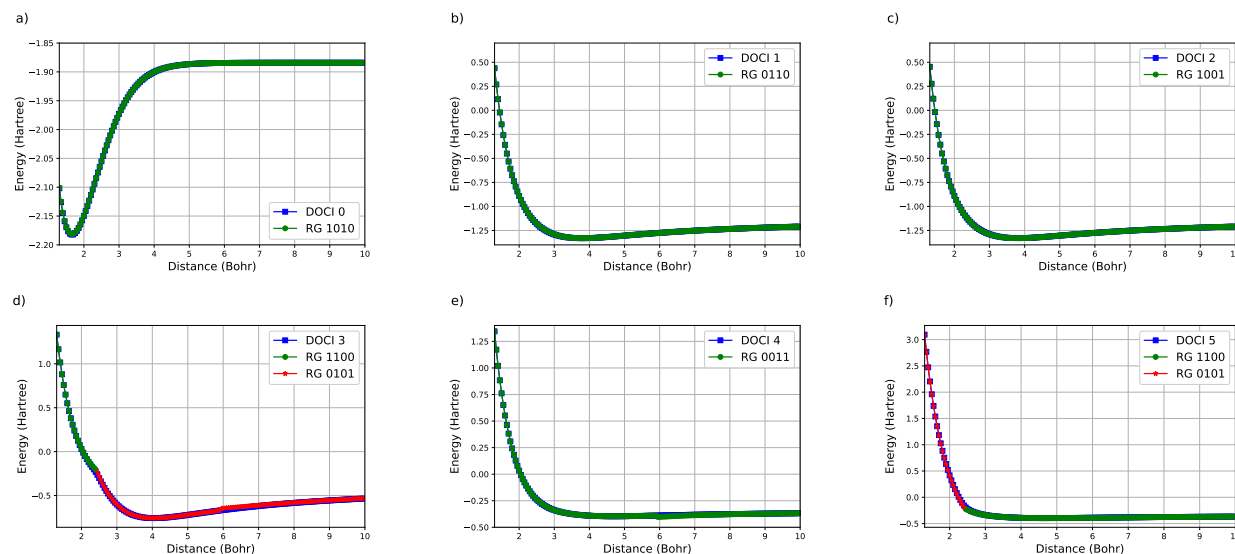


FIG. 5: RG states computed with reduced BCS parameters defining the Hamiltonian corresponding to the optimal 1010 solution and DOCI states. All results computed in the STO-6G basis set in the basis of OO-DOCI orbitals. For (a)-(c) and (e) the RG states correspond directly to DOCI states. For (d) and (f) the DOCI states have an avoided crossing, and thus switch RG states.

in the optimal reduced BCS Hamiltonian occurs. Our hypothesis was that RG states would be a better basis for DOCI than Slater determinants, but these results tend to suggest that the RG states *are* the DOCI states, except near particular points. We would have been satisfied with the conclusion that the 1010 state is essentially the DOCI ground state, but the matching of all states is strongly suggestive that the problem has been correctly understood. 1010 always represents the ground state, but the energetic ordering of the excited RG states changes over the course of the dissociation. Near the minimum, excitations are ordered by their rank in the sense that the “double” excitation 0101 is the highest in energy. At dissociation, the ordering is different, though we will postpone the discussion to the results for H_6 as it will be more clear.

C. H_6 Dissociation

For H_6 we proceeded directly to using the 101010 state, presented in figure 7. It is clear that the 101010 RG state is very close to the DOCI ground state everywhere except the

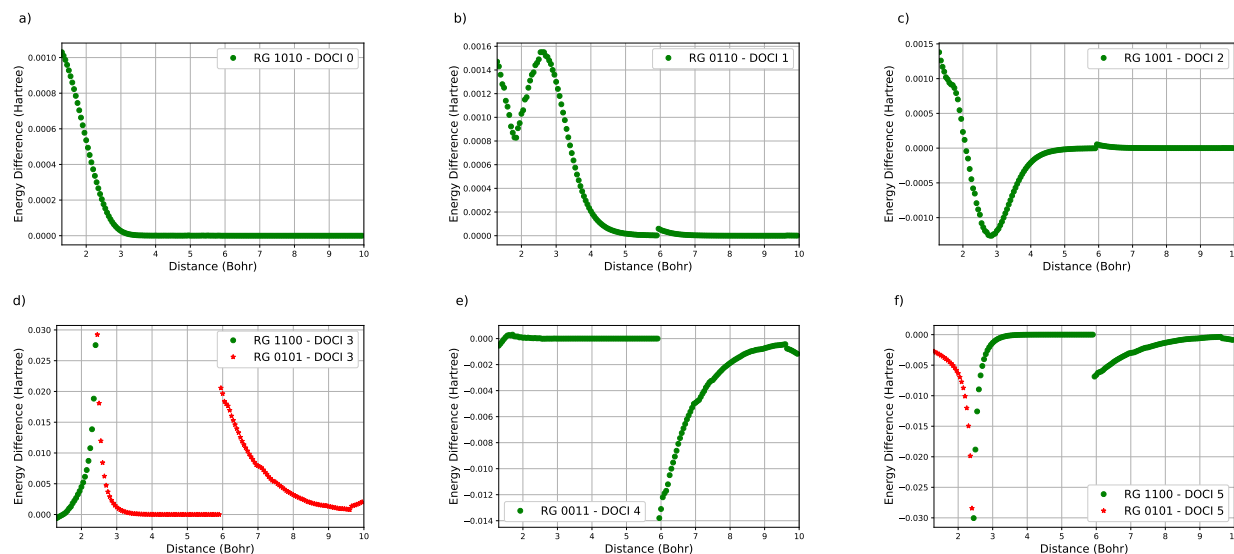


FIG. 6: Energy differences between RG states computed with reduced BCS parameters defining the Hamiltonian corresponding to the optimal 1010 solution and DOCI states. All results computed in the STO-6G basis set in the basis of OO-DOCI orbitals. For (a)-(c) and (e) the RG states correspond directly to DOCI states. For (d) and (f) the DOCI states have an avoided crossing, and thus switch RG states.

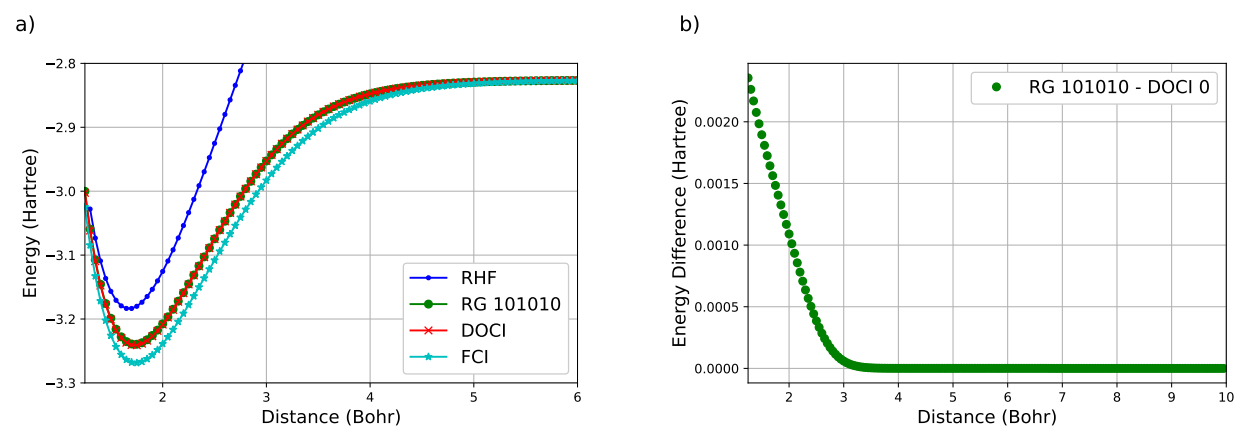


FIG. 7: (a) Bond dissociation curves for H_6 . (b) Energy difference between the 101010 RG state and the DOCI ground state. All results computed with the STO-6G basis set in the basis of OO-DOCI orbitals.

minimum where a short expansion would fix the problem completely. With the parameters defining the optimal Hamiltonian for the 101010 state, we computed all the RG states and compared them with the DOCI states. The DOCI spectrum is a little more complicated, to

say the least, though after resolving many crossings and avoided crossings, the RG states again seem to match the DOCI states very well. The results are plotted in appendix A. As for H_4 , there is a cusp in the optimal reduced BCS Hamiltonian parameters though it occurs near $r = 6.80$ bohr. This does not affect the smoothness of the 101010 approximation to the DOCI ground state, but is visible in other states.

The energetic ordering of the RG states is different near the minimum and at dissociation, though both follow clear patterns. Near the minimum, the RG states closely resemble Slater determinants, and thus the usual rules apply with the clear distinction being that the 101010 state is the minimum. Generally, the next lowest in energy are the single pair excitations, corresponding to the exchange of a 1 and a 0 in the bitstring. Double-pair excitations are generally next, followed by the triple-excitation 010101 at the top of the spectrum. At dissociation the rules are different. The state 101010 at dissociation is essentially 3 distinct subsystems, each of which has a state 10. It is useful to separate excitations into two types. First, there are *swaps* for which a subsystem switches 10 to 01. Second, there are *transfers* which correspond more or less to a transfer of a pair from one subsystem to another: one 10 becomes 00 while another becomes 11. With these ideas, the energetic ordering of the states is clearer. The lowest excited states are 011010, 100110, and 101001. They are degenerate and are “single-swaps” of the ground state 101010. The next states are the double-swaps 011001, 010110, and 100101 which are nearly degenerate (they differ by less than 2 milli-Hartree). Next are the 6 single-transfer states 100011, 101100, 001011, 111000, 001110, and 110010 which occur in 3 near-degenerate pairs. The triple-swap 010101 is next, followed finally by the single-swap-single-transfer states 010011, 011100, 110100, 000111, 001101, and 110001 which also occur in three near-degenerate pairs.

D. H_8 Dissociation

For H_8 the results are more or less the same. The ground state is very-well described by the RG state 10101010 as shown in figure 8. We did not plot the $\binom{8}{4} = 70$ states to compare with the DOCI states, but at dissociation the energetic ordering follows the same pattern as for H_6 . The lowest excited states are the degenerate single-swaps, followed by double-swaps, the single-transfers, the triple-swaps, and the single-swap-single-transfer states. Next are the quadruple swap 01010101, the double-swap-single-transfer states followed finally by

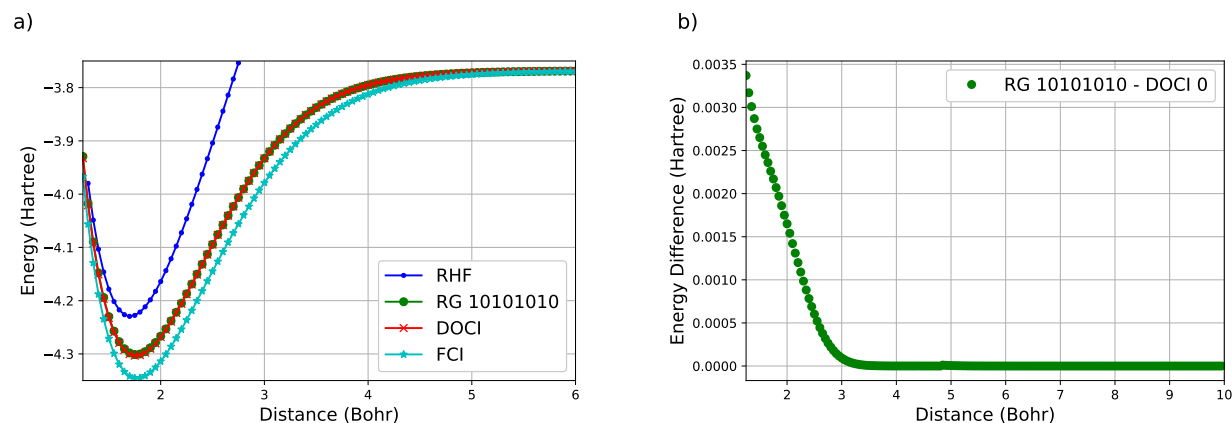


FIG. 8: (a) Bond dissociation curves for H_8 . (b) Energy difference between the 10101010 RG state and the DOCI ground state. All results computed with the STO-6G basis set in the basis of OO-DOCI orbitals.

the double-transfer states. The two double-transfer states 00111100 and 11000011 occur in the middle of the double-swap-single-transfer states. These are the only exceptions to the ordering pattern.

E. N_2 Dissociation

Finally, we computed results for the dissociation of N_2 . It was not obvious which bitstring would lead to the best RG treatment of the DOCI ground state, so several were computed and are plotted in figure 9. It is clear that the bitstring 1111101010 is close to the DOCI ground state, with a deviation that follows the same pattern seen for the H-chains: they agree quite well everywhere except near the minimum. This bitstring is easily reasoned physically: the core is always filled, hence the string 1111, while the valence orbitals are half-filled at dissociation, hence 101010. It is interesting that most of the bitstrings we tried are correct at dissociation. The ordering of the excited states is less clear.

The numerical results herein presented clearly demonstrate that an RG mean-field is very close to the DOCI ground state, and the corresponding RG excited states closely match the DOCI excited states, provided that the correct RG mean-field is chosen. On its face, choosing the right state is an NP-complete problem, though our results should suggest that physical reasoning should lead directly to the correct choice, or at least narrow the choice down enough that a greedy algorithm or dynamic programming would find the correct state. It is

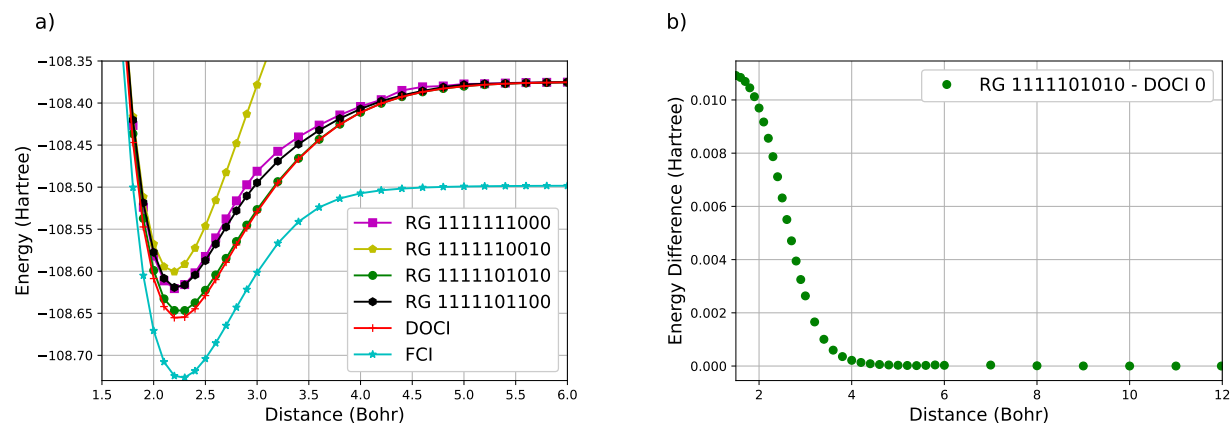


FIG. 9: (a) Bond dissociation curves for N₂. (b) Energy difference between the 111101010 RG state and the DOCI ground state. All results computed with the STO-6G basis set in the basis of OO-DOCI orbitals.

clear enough that for any system at half-filling, the correct choice is the Néel state 1010...10. Based on the results for N₂, we suggest a hypothesis based on splitting the spatial orbital basis into core, valence and virtual orbitals. The core is always full, and the virtuals are always empty, so will correspond to 11..11 and 00..00 respectively. The valence orbitals will be partially filled, and each pair of partial fillings correspond to 10. For N₂ in a minimal basis, the valence is half-filled, so the choice 101010 is obvious and unambiguous. For systems with valence orbitals less than half-filled it is less clear which bitstring is optimal. Further calculations are obviously necessary to confirm or refine this hypothesis, which we have begun and will report in a subsequent contribution. Choosing the correct bitstring is not the same as choosing an active space. The bitstring is *not* tied to specific orbitals since the single-particle energies must be sorted for the Richardson's equations solver. Thus, to pick the correct bitstring, we do not need to know which specific orbitals are partially occupied, only where they would occur in an energetic ordering.

IV. CONCLUSION

Variational calculations for bond-dissociation processes were performed using RG states as a mean-field. It is evident that a particular choice of RG state is very close to the DOCI ground state. Remarkably, the other RG states match the DOCI excited states quite well. Our hypothesis was that RG states are a better basis for DOCI than Slater

determinants, though it seems that the states are near-identical. For hydrogen chains, the Néel state, corresponding to the bitstring 1010..10 is always the optimal state. For N_2 , the state 1111101010 was found to be optimal. Both choices can be reasoned physically. Our hypothesis for choosing the correct state is based on separating the orbital spectrum into core, valence and virtual spaces. This does *not* require identifying which specific orbitals are in which space, only the number that are in each space.

The optimal variational RG curves are nearly identical to DOCI everywhere except near the minimum, where the correct result is weakly-correlated electrons. There, the RG states have a weak pairing strength g , and are thus quite close to Slater determinants. A short perturbative expansion, or CI diagonalization, will account for the missing energy.

With this we consider the principle proven: a single RG state effectively reproduces DOCI for Coulomb systems. This approach is a physical solution to a physical problem. While AP1roG / pCCD is currently the method to beat, the advantages of RG states are clear: AP1roG / pCCD is state-specific and must be solved by projection whereas the RG mean-field for the physical ground state gives a complete set of states which approximate the seniority-zero excited states quite well. Both methods have $\mathcal{O}(N^4)$ scaling in a given basis, or $\mathcal{O}(N^5)$ with orbital-optimization. Obviously non-zero seniority sectors will need to be treated separately. In the near term this may be accomplished in any number of ways by introducing non-zero seniorities post-hoc. In the longer term, this should be done correctly at the mean-field level which is substantially more difficult.

V. ACKNOWLEDGEMENTS

P.A.J. was supported by NSERC and Compute Canada. C.-É.F. is grateful for funding from the Vanier Canada Graduate Scholarships. S. C. was supported by NSERC.

VI. DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Appendix A: H_6 RG Excited States

The DOCI states for the symmetric H_6 dissociation have *many* crossings and avoided crossings. It is possible to decompose each DOCI state in terms of RG states (as was done in figure 5) but the figures become unreadable as up to 8 RG states contribute to individual DOCI states. Therefore, we have manually switched the DOCI curves to match the individual RG states, and the labelling corresponds to the energetic ordering at $r = 1.25$ bohr.

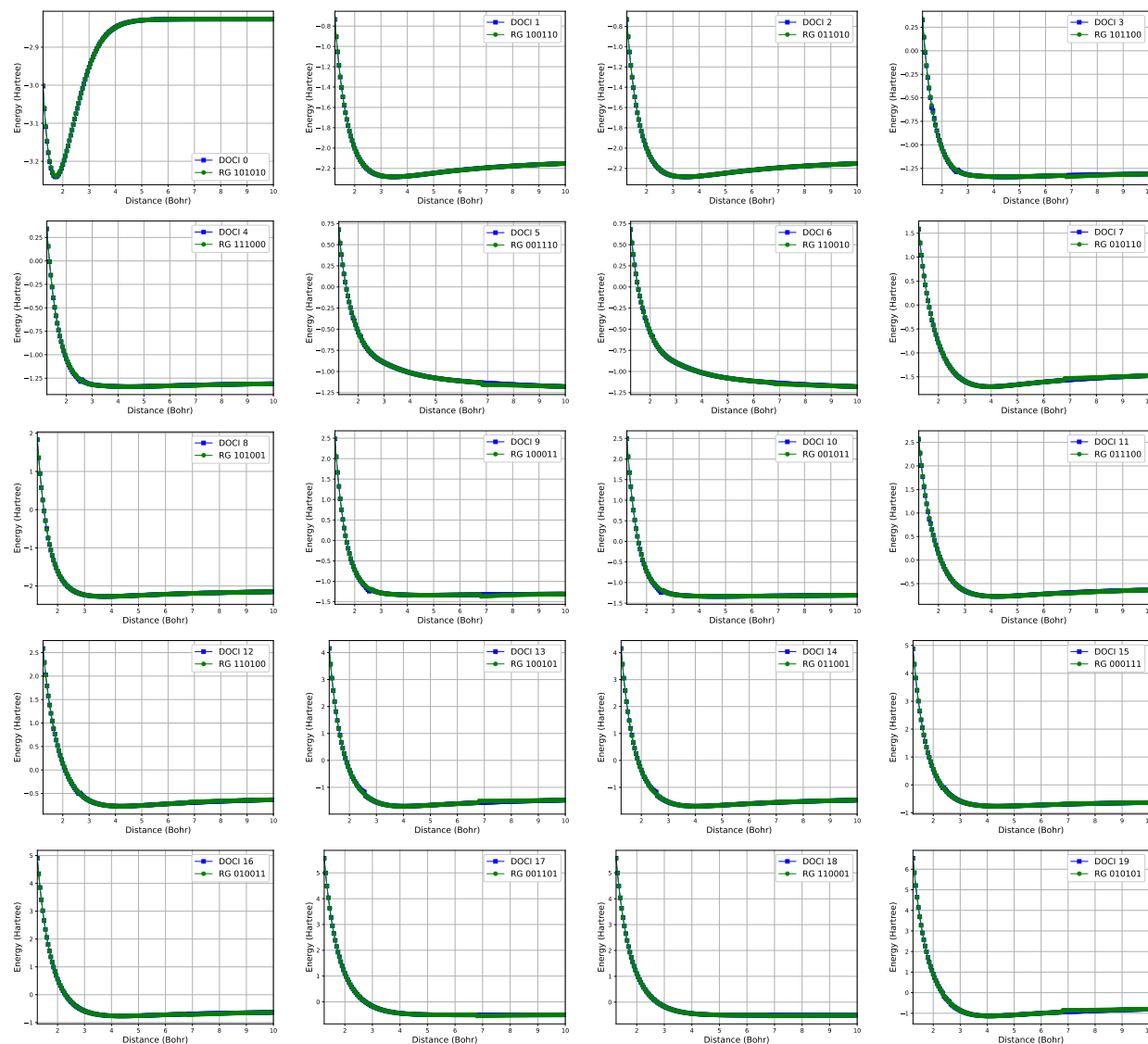


FIG. 10: (a)-(t) RG states computed with reduced BCS parameters defining the Hamiltonian corresponding to the optimal 101010 solution and DOCI states. All results computed in the STO-6G basis set in the basis of OO-DOCI orbitals.

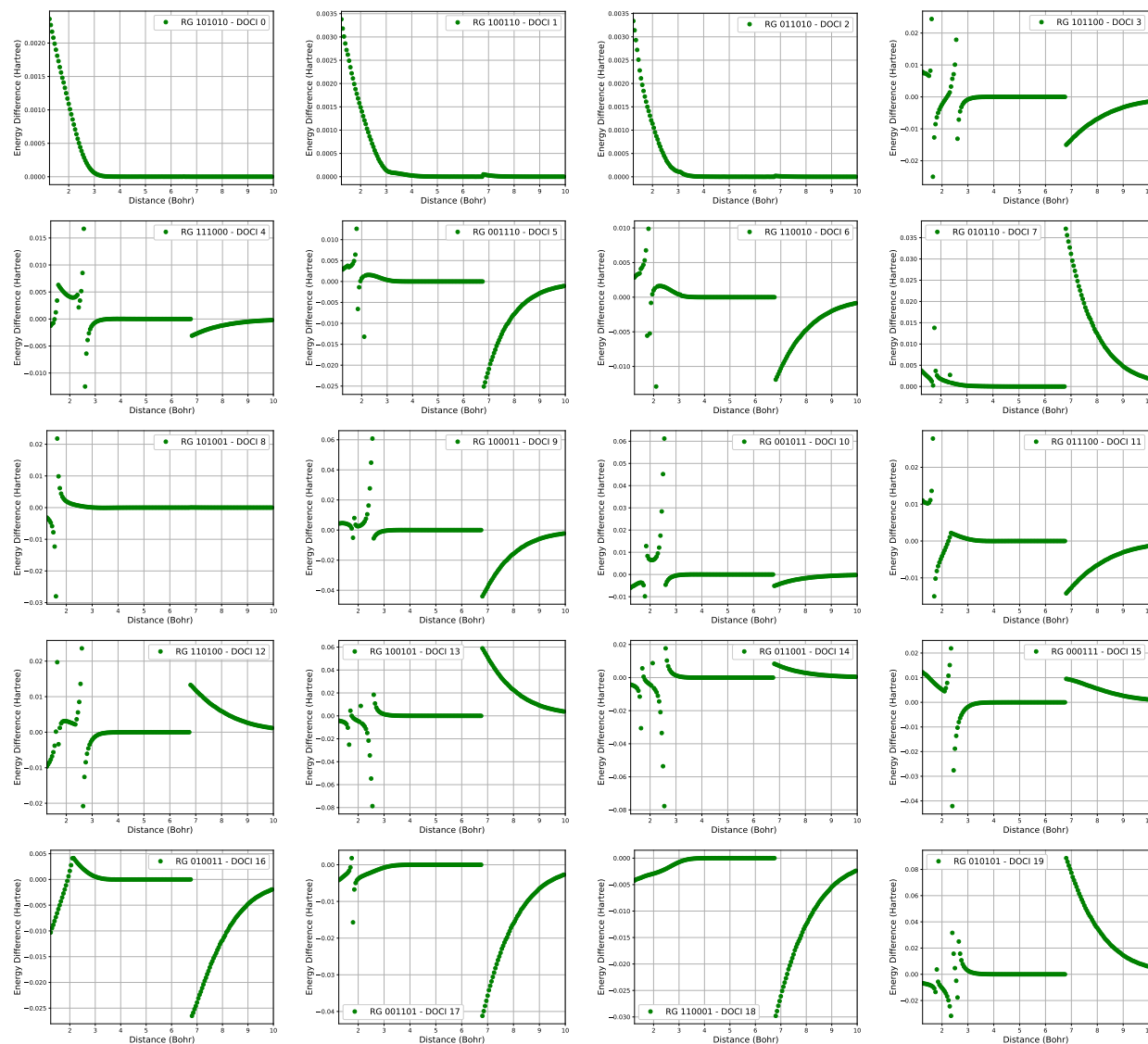


FIG. 11: (a)-(t) Energy differences between RG states computed with reduced BCS parameters defining the Hamiltonian corresponding to the optimal 101010 solution and DOCI states. All results computed in the STO-6G basis set in the basis of OO-DOCI orbitals.

REFERENCES

- ¹T. Helgaker, P. Jørgenson, and J. Olsen. *Molecular Electronic-Structure Theory*. Wiley & Sons, West Sussex, 2000.
- ²B. Huron, J. P. Malrieu, and P. Rancurel. *The Journal of Chemical Physics*, 58:5745, 1973.

- ³G. K.-L. Chan and M. Head-Gordon. *The Journal of Chemical Physics*, 116:4462, 2002.
- ⁴G. K.-L. Chan. *The Journal of Chemical Physics*, 120:3172, 2004.
- ⁵A. J. W. Thom and A. Alavi. *The Journal of Chemical Physics*, 123:204106, 2005.
- ⁶G. H. Booth and Alavi A. *The Journal of Chemical Physics*, 132:174104, 2010.
- ⁷G. K.-L. Chan and S. Sharma. *Annual Review of Physical Chemistry*, 62:465, 2011.
- ⁸G. H. Booth, A. Gruneis, G. Kresse, and A. Alavi. *Nature*, 493:431, 2013.
- ⁹S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar. *Journal of Chemical Theory and Computation*, 13:1595, 2017.
- ¹⁰A. A. Holmes, C. J. Umrigar, and S. Sharma. *The Journal of Chemical Physics*, 147:164111, 2017.
- ¹¹J. Li, M. Otten, A. A. Holmes, S. Sharma, and C. J. Umrigar. *The Journal of Chemical Physics*, 149:214110, 2018.
- ¹²Y. Yao and C. J. Umrigar. *Journal of Chemical Theory and Computation*, 17:4183, 2021.
- ¹³L. Bytautas, T. M. Henderson, C. A. Jimenez-Hoyos, J. K. Ellis, and G. E. Scuseria. *The Journal of Chemical Physics*, 135:044119, 2011.
- ¹⁴L. Bytautas, G. E. Scuseria, and K. Ruedenberg. *The Journal of Chemical Physics*, 143:094105, 2015.
- ¹⁵V. Fock. *Doklady Akademii Nauk SSSR*, 73:735, 1950.
- ¹⁶R. McWeeny. *Proceedings of the Royal Society of London*, A253:242, 1959.
- ¹⁷R. McWeeny. *Reviews of Modern Physics*, 32:335, 1960.
- ¹⁸R. McWeeny and B. Sutcliffe. *Proceedings of the Royal Society of London*, A273:103, 1963.
- ¹⁹V. A. Nicely and J. F. Harrison. *The Journal of Chemical Physics*, 54:4363, 1971.
- ²⁰W. F. Siems and R. D. Poshusta. *The Journal of Chemical Physics*, 64:6, 1976.
- ²¹D. M. Silver. *The Journal of Chemical Physics*, 50:5108, 1969.
- ²²D. M. Silver. *The Journal of Chemical Physics*, 52:299, 1970.
- ²³D. M. Silver, E. L. Mehler, and K. Ruedenberg. *The Journal of Chemical Physics*, 52:1174, 1970.
- ²⁴D. M. Silver, K. Ruedenberg, and E. L. Mehler. *The Journal of Chemical Physics*, 52:1206, 1970.
- ²⁵A. J. Coleman. *Journal of Mathematical Physics*, 6:1425, 1965.
- ²⁶J. V. Ortiz, B. Weiner, and Y. Öhrn. *International Journal of Quantum Chemistry*, 15:113, 1981.

- ²⁷C. R. Sarma, J. Paldus, and Y. Öhrn. *International Journal of Quantum Chemistry*, 36:35, 1989.
- ²⁸A. J. Coleman. *International Journal of Quantum Chemistry*, 63:23, 1997.
- ²⁹T. M. Henderson and G. E. Scuseria. *The Journal of Chemical Physics*, 151:051101, 2019.
- ³⁰A. Khamoshi, T. M. Henderson, and G. E. Scuseria. *The Journal of Chemical Physics*, 151:184103, 2019.
- ³¹T. M. Henderson and G. E. Scuseria. *The Journal of Chemical Physics*, 153:084111, 2020.
- ³²A. Khamoshi, F. A. Evangelista, and G. E. Scuseria. *Quantum Science and Technology*, 6:014004, 2020.
- ³³R. Dutta, T. M. Henderson, and G. E. Scuseria. *Journal of Chemical Theory and Computation*, 16:6358, 2020.
- ³⁴A. Khamoshi, G. P. Chen, T. M. Henderson, and G. E. Scuseria. *The Journal of Chemical Physics*, 154:074113, 2021.
- ³⁵R. Dutta, T. M. Henderson, and G. E. Scuseria. *The Journal of Chemical Physics*, 154:114112, 2021.
- ³⁶E. Neuscamman. *Physical Review Letters*, 109:203001, 2012.
- ³⁷E. Neuscamman. *The Journal of Chemical Physics*, 139:194105, 2013.
- ³⁸E. Neuscamman. *Molecular Physics*, 114:577, 2016.
- ³⁹A. C. Hurley, J. Lennard-Jones, and J. A. Pople. *Proceedings of the Royal Society*, A220:446, 1953.
- ⁴⁰W. Kutzelnigg. *The Journal of Chemical Physics*, 40:3640, 1964.
- ⁴¹P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck. *Journal of Chemical Theory and Computation*, 9:1394, 2013.
- ⁴²T. Stein, T. M. Henderson, and G. E. Scuseria. *The Journal of Chemical Physics*, 140:214113, 2014.
- ⁴³P. A. Limacher, T. D. Kim, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck. *Molecular Physics*, 112:853, 2014.
- ⁴⁴P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck. *Physical Chemistry Chemical Physics*, 16:501, 2014.
- ⁴⁵T. M. Henderson, G. E. Scuseria, J. Dukelsky, A. Signoracci, and T. Duguet. *Physical Review C*, 89:054305, 2014.

- ⁴⁶T. M. Henderson, I. W. Bulik, T. Stein, and G. E. Scuseria. *The Journal of Chemical Physics*, 141:244104, 2014.
- ⁴⁷K. Boguslawski, P. Tecmer, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck. *Physical Review B*, 98:201106(R), 2014.
- ⁴⁸K. Boguslawski, P. Tecmer, P. Bultinck, S. De Baerdemacker, D. Van Neck, and P. W. Ayers. *Journal of Chemical Theory and Computation*, 10:4873, 2014.
- ⁴⁹K. Boguslawski, P. Tecmer, P. A. Limacher, P. A. Johnson, P. W. Ayers, P. Bultinck, S. De Baerdemacker, and D. Van Neck. *Journal of Chemical Theory and Computation*, 140:214114, 2014.
- ⁵⁰P. Tecmer, K. Boguslawski, P. A. Johnson, M. Chan, T. Verstraelen, and P. W. Ayers. *Journal of Physical Chemistry*, A118:9058, 2014.
- ⁵¹K. Boguslawski and P. W. Ayers. *Journal of Chemical Theory and Computation*, 11:5252, 2015.
- ⁵²J. M. Wahlen-Strothmann, T. M. Henderson, and G. E. Scuseria. *Molecular Physics*, 116:186, 2018.
- ⁵³J. Bardeen, L. N. Cooper, and J. R. Schrieffer. *Physical Review*, 106:162, 1957.
- ⁵⁴J. Bardeen, L. N. Cooper, and J. R. Schrieffer. *Physical Review*, 108:1175, 1957.
- ⁵⁵J. R. Schrieffer. *Theory of Superconductivity*. CRC Press, Boca Raton, 1964.
- ⁵⁶R. W. Richardson. *Physics Letters*, 3:277, 1963.
- ⁵⁷R. W. Richardson and N. Sherman. *Nuclear Physics*, 52:221, 1964.
- ⁵⁸R. W. Richardson. *Journal of Mathematical Physics*, 6:1034, 1965.
- ⁵⁹M. Gaudin. *Journal de Physique*, 37:1087, 1976.
- ⁶⁰P. A. Johnson, C.-É. Fecteau, F. Berthiaume, S. Cloutier, L. Carrier, M. Gratton, P. Bultinck, S. De Baerdemacker, D. Van Neck, P. Limacher, and P. W. Ayers. *The Journal of Chemical Physics*, 153:104110, 2020.
- ⁶¹J.-D. Moisset, C.-É. Fecteau, and P. A. Johnson. *arXiv*, page 2202.09401, 2022.
- ⁶²F. Weinhold and E. B. Wilson. *The Journal of Chemical Physics*, 46:2752, 1967.
- ⁶³F. Weinhold and E. B. Wilson. *The Journal of Chemical Physics*, 47:2298, 1967.
- ⁶⁴D. B. Cook. *Molecular Physics*, 30:733, 1975.
- ⁶⁵H. Bethe. *Zeitschrift für Physik*, 71:205, 1931.
- ⁶⁶L. D. Faddeev, E. K. Sklyanin, and L. A. Takhtadzhian. *Theoretical and Mathematical Physics*, 40:688, 1980.

- ⁶⁷V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin. *Quantum Inverse Scattering Method and Correlation Functions*. Cambridge University Press, Cambridge, 1993.
- ⁶⁸L. Carrier, C.-É. Fecteau, and P. A. Johnson. *International Journal of Quantum Chemistry*, 120:e26255, 2020.
- ⁶⁹J.-D. Moisset, L. Carrier, and P. A. Johnson. *arXiv*, page 2109.05144, 2021.
- ⁷⁰S. Rombouts, D. Van Neck, and J. Dukelsky. *Physical Review C*, 69:061303(R), 2004.
- ⁷¹X. Guan, K. D. Launey, M. Xie, L. Bao, F. Pan, and J. P. Draayer. *Physical Review C*, 86:024313, 2012.
- ⁷²W. V. Pogosov. *Journal of Physics: Condensed Matter*, 24:075701, 2012.
- ⁷³S. De Baerdemacker. *Physical Review C*, 86:044332, 2012.
- ⁷⁴P. W. Claeys, S. De Baerdemacker, M. Van Raemdonck, and D. Van Neck. *Physical Review B*, 91:155102, 2015.
- ⁷⁵A. Faribault, O. El Araby, C. Sträter, and V. Gritsev. *Physical Review B*, 83:235124, 2011.
- ⁷⁶O. El Araby, V. Gritsev, and A. Faribault. *Physical Review B*, 85:115130, 2012.
- ⁷⁷A. Faribault, P. Calabrese, and J.-S. Caux. *Physical Review B*, 77:064503, 2008.
- ⁷⁸A. Faribault, P. Calabrese, and J.-S. Caux. *Physical Review B*, 81:174507, 2010.
- ⁷⁹P. A. Johnson, F. Fortin, S. Cloutier, and C.-É. Fecteau. *The Journal of Chemical Physics*, 154:124125, 2021.
- ⁸⁰P. W. Claeys, D. Van Neck, and S. De Baerdemacker. *SciPost Physics*, 3:028, 2017.
- ⁸¹L. N. Trefethen. *Approximation Theory and Approximation Practice, Extended Edition*. SIAM, 2019.
- ⁸²L. Amico and A. Osterloh. *Physical Review Letters*, 88:127003, 2002.
- ⁸³G. Gorohovsky and E. Bettelheim. *Physical Review B*, 84:224503, 2011.
- ⁸⁴C.-É. Fecteau, H. Fortin, S. Cloutier, and P. A. Johnson. *The Journal of Chemical Physics*, 153:164117, 2020.
- ⁸⁵L. N. Trefethen and D. Bau III. *Numerical linear algebra*, volume 50. Siam, 1997.
- ⁸⁶G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez, T. Harville, S. Irle, J. Ivanic, K. Kowalski, S. S. Leang, H. Li, W. Li, J. J. Lutz, I. Magoulas, J. Mato, V. Mironov, H. Nakata, B. Q. Pham, P. Piecuch, D. Poole, S. R. Pruitt, A. P. Rendell, L. B. Roskop, K. Ruedenberg, T. Sattasathuchana, M. W. Schmidt, J. Shen, L. Slipchenko, M. Sosonkina, V. Sundriyal, A. Tiwari, J. L. Galvez Vallejo, B. Westheimer, M. Włoch, P. Xu, F. Zahariev, and M. S.

- Gordon. *The Journal of Chemical Physics*, 152:154102, 2020.
- ⁸⁷A. Marie, F. Kossoski, and P.-F. Loos. *The Journal of Chemical Physics*, 155:104105, 2021.
- ⁸⁸F. Kossoski, A. Marie, A. Scemama, M. Caffarel, and P.-F. Loos. *Journal of Chemical Theory and Computation*, 17:4756, 2021.
- ⁸⁹C. D. Sherill and H. F. Schaefer III. *Advances in Quantum Chemistry*, 34:143, 1999.
- ⁹⁰R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherill. *Journal of Chemical Theory and Computation*, 13:3185, 2017.
- ⁹¹J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. M. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian 16 revision c.01. *Gaussian Inc. Wallingford CT*, 2016.
- ⁹²N. Hansen and A. Ostermeier. *Evolutionary Computation*, 9:159, 2001.
- ⁹³J. A. Nelder and R. Mead. *Computer Journal*, 7:308, 1965.
- ⁹⁴W. Poelmans. *Variational determination of the two-particle density matrix: The case of doubly-occupied space*. PhD thesis, Ghent University, 2015.