Doctorate Dissertation
博士論文

Geminal theory for strongly correlated few-body systems
(少数多体強相関系に対するジェミナル理論)

A Dissertation Submitted for Degree of Doctor of Philosophy
December 2018
平成 30 年 12 月博士(理学)申請

Department of Physics, Graduate School of Science,
The University of Tokyo
東京大学大学院理学系研究科物理学専攻

Airi Kawasaki
川﨑　愛理

Abstract

To calculate strongly correlated systems accurately, one needs to treat strong and weak correlations simultaneously. Popular methods, such as Hartree-Fock theory or density functional theory, can treat the weak correlation but cannot treat the strong correlation. On the other hand, the full-configuration-interaction (full-CI) method can treat the strong correlation well but it needs high calculation cost. In this context, I focus on the wave function theory for the electron pair. In condensed matter physics, the electron pair is an important concept for superconductivity. Also in the field of chemistry, the concept of electron pair has been used to represent a chemical bond since a long time ago. The electron pair is called geminal in chemistry and many calculation methods using geminal were developed. The method of expressing different chemical bonds using different geminals is called antisymmetrized-product-of-geminals (APG) theory. Although the idea of APG was proposed long ago, there are few APG studies because of its high calculation cost and complexity in the calculation.

In this thesis, I overcome the computational difficulty of APG using a tensor decomposition method developed in mathematics and make the variational determination of APG tractable numerically. This makes it possible to analyze the APG wave function without introducing additional approximations to the geminals and thus to understand the inherent advantage and disadvantage in describing strongly correlated few-body systems. This understanding helps me to develop a method to incorporate the electron correlation beyond the original APG. This novel method is based on a polynomial extension of APG; note that the original APG has a monomial form. With the polynomial extension, I succeed in relating geminals to the valence bond even in the strong correlation regime. I recognize the polynomial extension as an introduction of the "resonance" effect into the APG-based valence bond theory and thus is a natural improvement.

I also develop variations of the APG calculation. I develop a simplified geminal method, which is different from the ones developed previously in the literatures. I also develop a geminal method specialized for a strongly correlated impurity system embedded in a weakly correlated medium.

The present study analyzes geminal theories comparatively from (a) the simplest one,

called antisymmetrized geminal powers (AGP), which is a mean-field theory of geminals, (b) APG and (c) its extension, bridging thereby the HF to full-CI via electron-pair theories of different levels. I believe that the present work has made clearer the property of geminal theory and would stimulate further sophistication of the geminal-based valence bond theory.

Acknowledgements

I am deeply grateful to my supervisor, Professor Osamu Sugino. I enjoyed studying with him for five years. Without him I did not choose an academic career. I appreciate his helpful advice. I also thank our group members.

I would like to thank Professor Paul W. Ayers for giving me some advice about my study and introduced many research groups to me. I also thank Professor Gustavo E. Scuseria for discussing geminal researches.

I would like to offer my special thanks to the Madoka Mochida at the Public Relation Office of the Institute for Solid State Physics. Most of this thesis was written in the PR office. I was so thankful that she provided me a comfortable place. I was inspired by seeing her working very hard. I was so glad we were able to spend time together. I enjoyed having lunch with her. I appreciate her caring about my health condition. When I was down, she encouraged me. Her words always cheer me up. I could not write this thesis without her. I cannot thank her enough. I was happy to be able to meet her.

Finally, special thanks to a singer, Mai Kuraki. Her music always gives me the energy to go on.

Contents

Chapter 1

Introduction

Determination of the ground state of many-body systems is one of the most basic and important problems in many fields of science such as condensed matter physics, quantum chemistry and nuclear physics. Density functional theory (DFT) [1] is established as a de facto standard theory for solids, liquids, molecules, atoms and nuclear matters although the application has so far been limited to moderately correlated systems. This limitation is due to the fact that popular DFT methods such as the Kohn-Sham (KS) method [2] are based on the single Slater determinant that is used for a non-interacting reference system under the influence of an effective KS potential. The reference system, however, cannot be properly connected to the target system when the correlation becomes significant, at least, within the current level of the theory. In this context, hybrid methods like DFT+U [3] and DFT+DMFT [4] were developed to augment the description of strong correlation effect when the correlation occurs locally within an atom of a solid, although legitimate approaches such as the quantum Monte Carlo method [5] or the field theoretical approaches [6] are more favorable ultimately. It is noteworthy that similar hybrid methods such as the density-matrix-embedding theory (DMET) [7, 8] were developed for chemical systems, where an accurate wave function like the one described using the configuration interaction (CI) is embedded in a wave function of a non-interacting system. Considering the success of DMET in describing some of the chemical systems and its potential applicability to more general materials where local correlation is important, it is important to advance the hybrid method for strongly correlated few-body systems embedded in a moderately interacting medium. As a step towards this long-range goal, I develop in this thesis a method for strongly correlated few-body systems and then try to embed it into a mean-field theory.

1.1 Wave-function theories

Among the wave-function theories, configuration interaction (CI) is the most versatile and accurate one, but the application has been severely restricted even for few-body systems owing to the number of configurations that grows explosively with the number of electrons. Historically, there has been much effort to reduce the number of configurations from a perturbative consideration. For example, the configurations were limited by taking only few-particle excitations from a single Slater determinant or from multi-determinants in the early stage, and later, the configurations were extended to include the linked clusters of the excitations in the coupled-cluster (CC) theory [9, 10]. Those methods are, however, successful only for moderately correlated system where the single Slater determinant is a good starting point. To approach the strongly correlated systems in the full-CI level, the density-matrix renormalization-group (DMRG) [11] was adapted to few-body molecular systems [12, 13]. Although the quantum chemical DMRG was shown to work for a number of molecules, it is successful only when localized molecular orbitals (MO) can be arranged one-dimensionally, reflecting the general property of DMRG. In this context, I pay attention to a different wave-function theory called a geminal theory, which can compactly represent the wave function.

1.2 Valence bond theory

As a step for the explanation of the geminal theory, let me briefly follow the history of the theory of chemical bond. Already in 1910's, Lewis introduced the concept of the chemical bond, which is formed when two electrons overlap. This concept was subsequently formulated in the valence bond theory in terms of the molecular orbitals (MOs), or the hybridized atomic orbitals such as the sp_2 and sp_3 hybridization, which accommodate up to two electrons. With the MOs, the electronic structure is characterized by a network of the valence bonds, or more correctly, as the resonance of possible networks. For example, the electronic structure of a benzene molecule is captured as the resonance of two Kekulé structures consisting of single and double bonds. The molecular structures and stability, in addition, can be explained by the repulsive interaction of the valence bonds in the valence shell electron pair repulsion (VSPER) theory [14]. Although the valence bond theory has so far been successfully developed for moderately correlating systems such as aromatic molecules, extension of the theory toward more strongly correlating systems has been rarely studied as far as I know. Considering the potential of the valence bond theory in explaining few-body systems, I will focus on the possibility of the extension of the valence bond theory for the strongly correlated systems.

1.3 Geminal theory

Two electrons constituting a valence bond may be most generally described by the geminal defined as

$$g\left(r_1, r_2\right) = \sum_{ij} \epsilon_{ij} \varphi_i\left(r_1\right) \varphi_j\left(r_2\right), \tag{1.1}$$

where φ's are the molecular orbitals (MOs) that are not necessarily orthogonal to each other. The geminal is antisymmetric with respect to the exchange of space-spin coordinate of an electron, $r_1 \leftrightarrow r_2$, because of the antisymmetric matrix ϵ. Equation (1.1) may be alternatively written using the creation operator $\hat{a}_i^\dagger$ as

$$\hat{g} \equiv \sum_{ij} \epsilon_{ij} \hat{a}_i^\dagger \hat{a}_j^\dagger. \tag{1.2}$$

The geminal thus defined plays a central role in this thesis.

This definition was first used by Coleman in his geminal wave function theory [15], where the many-body wave function is represented as an antisymmetric product of the geminals:

$$\psi^{\mathrm{AGP}}\left(r_1, r_2, \cdots\right) = \hat{A}\left[g\left(r_1, r_2\right) g\left(r_3, r_4\right) \cdots\right], \tag{1.3}$$

or

$$|\psi^{\mathrm{AGP}}\rangle = \hat{g}^{n/2}|0\rangle, \tag{1.4}$$

where $\hat{A}$ is an operator to fully antisymmetrize the geminals and n is the number of electrons. The wave function thus constructed is called as antisymmetrized geminal powers (AGP), which I will detail in Chapter 2. In short, the AGP wave function is an extension of the Hartree-Fock (HF) wave function in that AGP is a mean-field theory of an electron pair while HF is a mean-field theory of an individual electron. The intra-pair correlation is automatically considered in AGP, while it is not the case in HF. The AGP wave function is related to the BCS wave function [16] in that the former is obtained by fixing the number of electrons in the latter, indicating a formal similarity between AGP and BCS in spite of their conceptual difference.

When different geminal types are used for different electrons as

$$\psi^{\mathrm{APG}}\left(r_1, r_2, \cdots\right) = \hat{A}\left[g_1\left(r_1, r_2\right) g_2\left(r_3, r_4\right) \cdots\right], \tag{1.5}$$

or as

$$|\psi^{\mathrm{APG}}\rangle = \prod_{m=1}^{n/2} \hat{g}[m]|0\rangle, \tag{1.6}$$

where the different geminals are labeled by index m. The wave function is called antisymmetrized product of geminals (APG) [17]. APG is an extension of AGP in that the former can describe, to some extent, the correlation between geminals missing in the latter. Importantly, geminals in APG can be related to the chemical bond. Suppose, for example, that the first geminal is a spin-singlet combination of an sp_3 hybridized MO; then the geminal can be assigned as an sp_3 bond. The geminals localized in a bond center may represent a covalent bond, and those localized in a back bond region may represent a lone pair. Although the existing APG theory is not comparable to full-CI both in accuracy and versatility, the theory has a potential not only to properly redefine the chemical bond in the moderate correlation regime but also to extend the chemical bond in the strong correlation scheme. Considering the present level of APG, however, it is currently important to level up the APG and bridge the gap existing between APG and the full-CI.

1.4 Previous APG theory

Although the APG theory was proposed long time ago, the study of APG has not yet been proceeded very well. This is because of the known complexity in the variational calculation and the large computational cost that increases exponentially with the number of electrons [18]. In this context, rather than advancing the APG-based theory toward higher accuracy, much effort has been made for simplifying the APG calculation. For example, the APG wave function has been simplified by (a) restricting geminals to be strongly orthogonal to each other in the antisymmetrized product of strongly orthogonal geminal (APSG) scheme [19, 20, 21] and by (b) using the same pairing scheme for all geminals, or using the same set of MOs throughout, in the antisymmetrized product of interacting geminals (APIG) [22]. In this context, I begin by advancing the APG theory so that the restrictions can be removed. For this purpose, I will introduce in this thesis a transformation called Waring decomposition to convert an APG wave function to a linear combination of the AGP ones, for which an analytical form for the total energy is available. Indeed, the method for the linear combination of AGP, which I call as the AGP-CI, has been developed and tested as detailed in the next paragraph. Therefore, by introducing the Waring decomposition, it is possible to variationally obtain the wave function and the total energy. This indicates that the potential-energy curve and the atomic force, which are crucially important quantities in chemistry, can also be calculated although I do not go into this direction in this thesis. The Waring decomposition is thus the most important step in this thesis.

1.5 APG versus AGP

The formula for the AGP-CI was derived by Onishi and Yoshida [23] and was applied to small molecules and a four-site Hubbard model by Uemura *et al.* [24], as will be detailed in Chapter 2. The research of Uemura *et al.* is based on the fact that the many-body wave function can be more efficiently expanded by AGPs than by Slater determinants. In other words, the CI-coefficient tensor $A_{i_1,i_2,\cdots,i_n}$ of an n electron system defined in

$$\Psi\left(r_1, r_2, \cdots\right) = \sum_{i_1,i_2,\cdots} A_{i_1,i_2,\cdots,i_n} \varphi_{i_1}\left(r_1\right) \varphi_{i_2}\left(r_2\right) \cdots \tag{1.7}$$

can be conveniently expanded as products of an anti-symmetric matrix F as

$$A_{i_1,i_2,\cdots,i_n} = \sum_{r=1}^{R} \sum_{\sigma \in S_{n/2}} \mathrm{sgn}(\sigma) F^r_{\sigma(i_1)\sigma(i_2)} F^r_{\sigma(i_3)\sigma(i_4)} \cdots \tag{1.8}$$

with

$$\sum_{i_1,i_2} F^r_{i_1 i_2} \varphi^r_{i_1}\left(r_1\right) \varphi^r_{i_2}\left(r_2\right) \tag{1.9}$$

forming a geminal $g_r\left(r_1, r_2\right)$. The many-body wave function can thus be expanded by R pieces of different AGPs as

$$\sum_{r=1}^{R} \mathrm{AGP}^r\left(r_1, r_2, \cdots\right) \tag{1.10}$$

with

$$\mathrm{AGP}^r\left(r_1, r_2, \cdots\right) \equiv \sum_{\sigma \in S_{n/2}} \mathrm{sgn}(\sigma) F^r_{\sigma(i_1)\sigma(i_2)} F^r_{\sigma(i_3)\sigma(i_4)} \cdots . \tag{1.11}$$

In the above, S_m is the symmetric group of degree m, sgn is the signature of the permutation, R is the anti-symmetric rank of the tensor $A_{i_1,i_2,\cdots,i_n}$, and the expansion Eq. (1.8) is the one called as the anti-symmetric tensor decomposition. By optimizing the molecular orbitals, φ^r_i, one can expect to reduce the rank R, although a method of finding the lower limit of R has not been developed. Uemura *et al.* showed that the value R is less than ten for the small systems that they investigated. The AGP-CI calculation scales as $O\left(n^5 R^2\right)$ and the lower limit of R would possibly scale exponentially with n when estimating with the general theory of tensor decomposition [25].

By improving the algorithms, AGP-CI may be sophisticated further and would provide a more efficient way for the variational calculation. However, I will develop a method for APG in this thesis instead of improving the AGP-CI. This is because, as discussed above, the APG theory has provided an intuitive chemical picture on the molecular systems

and is expected to be the better method for strongly correlated few-body systems. The difference between the AGP-based and APG-based approaches will be more striking with increasing degrees of freedom (DoF) where interpretation of the results may be difficult because of the number of DoF. With the importance of the APG-based method in mind, I will develop an algorithm for the APG together with APG-CI, although I will call the latter the polynomial APG as will be explained below.

1.6 APG calculation

I will use, as a benchmark system, a one-dimensional Hubbard model with the on-site Coulomb interaction U being up to 10 times larger than the transfer interaction t. Below, I will rescale the energy so that $t = 1$ for simplicity. This system will correspond to a one-dimensionally linked hydrogen atoms; the system will also correspond to a hypothetical carbon ring with one of the two π orbitals passivated by a hydrogen atom, leaving thereby a system consisting of pz orbitals only. Note that, in the latter case, the system will correspond to a benzene molecule when the number of sites is taken to be six.

I will show, by performing variational calculation of the Hubbard model, that APG yields the value of the total energy significantly close to that of full-CI compared with AGP and HF, but the APG wave function incorrectly breaks the spin symmetry exhibiting a spin wave. The broken symmetry is apparent both in the density profile and in the pair correlation function, but is more apparent in the wave function. When the geminals constituting the APG wave function are transformed into canonical orbitals, as is commonly done in the analysis of the BCS wave function, the amplitude of the canonical orbitals are found to be unnaturally biased to a few sites and do not behave like a chemical bond as expected to do in a benzene molecule. This biased behavior of the geminals, derived by using only one APG in the variational calculation, seems to arise from the neglect of the resonance. Indeed, by using multiple APGs, one can recover a plausible behavior as the bond. Although the result that the resonance is essentially important may sound too trivial, I will arrive at this conclusion via a naïve trial-and-error investigation of possible extensions of the APG theory. I will show that the variationally superior wave function behaves more favorably as a resonating valence bond.

1.7 Technical aspect of APG calculation

The APG calculation scales as $O\left(n^5 \times (2^n n!)^2\right)$ since $R \sim 2^n n!$ without the resonance. In addition, the gradient-based variational calculation such as the conjugate-gradient (CG) method often slows down by the existence of a saddle point as will be detailed more in 3.8. That is, the variational calculation is made difficult by the non-convex character, which

is presumably due to the antisymmetric character of the CI-coefficient tensor. Because of this, I start from typically 100 initial conditions in performing the variational calculation instead of taking a Hessian-based algorithm, which is an alternative and steady approach. Along this strategy for the calculation, I did the calculation up to a 12-electron system although improvement in the algorithm will allow me to handle larger systems. In this context, I will try, in sections 3.5 and 4.6, to speed up the calculation by restricting the degrees of freedom (DoF) for the geminal. The strategy of the restriction is based on the observation that the variationally determined geminals have a peculiar structure; they consist of a major components having the dominant eigenvalue, several minor components with much less dominant eigenvalues, and others with negligible eigenvalues. Here, the dominance has been judged from the absolute value of the eigenvalue. Making use of the aforementioned structure of the geminals, I use only the major component for the variation in a simplified calculation, in other words, I prepare antisymmetric matrices of rank-n for the variational calculation. I will call the method a low-rank approximation. I will show that the total energy does not so sensitively increase by using the low-rank restriction, suggesting that the restriction is a reasonable choice for the simplification. I will further test the restriction using the polynomial APG.

1.8 AGP embedding scheme

Finally, I will formulate an embedding method called four-body correlation embedding method. Therein, I assume a system consisting of a single site of strong on-site Coulomb interaction ($U/t = 10$) surrounded by a few tens of sites with negligible Coulomb interaction. I will do a variational calculation using an AGP-type trial wave function of the form:

$$\exp\left[\sum_{pqrs} G_{pqrs}\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_r^\dagger\hat{a}_S^\dagger\right]|\psi^{\mathrm{AGP}}\rangle, \tag{1.12}$$

where $\hat{a}_p^\dagger$ is the creation operator of the p th MO. The summation over the sites is restricted so that only important components of the antisymmetric tensor G can be used for the variation: Here, the two of the four indices of G are restricted to be at the strongly correlated site (while other indices are unrestricted) to consider correlation of geminals around that site. The number of electrons is fixed using a technique described below. By performing the variational calculation with the help of the Waring decomposition, I will show that the resulting wave function is quite close to the exact solution, showing the promise of such an embedding method.

1.9 Summary of Introduction

I will detail my thesis work in the following chapters along with the line briefly introduced in this chapter. The motivation of this work is to develop a wave function theory for strongly correlated few-body systems that are either isolated or embedded in a weakly interacting environment. The defects in semiconductor or catalysts on a hydrocarbon material will be the future target. Considering that existing DFT methods or wave function methods are not very suitable for the purpose, I develop a rather unconventional wave function method based on the antisymmetrized product of geminals (APG). One may hope to relate the geminals constituting APG to the valence bond and the linear combination of APG to the resonance, so that I expect the APG-based method to yield a means to understand the electronic structure in terms of the bond even in the strongly correlated regime. I step forward in this direction by enabling the variational calculation of the APG wave function; the Waring decomposition plays a crucial role in the computational method although other improvements are also done to enable the calculation. The single APG calculation is superior to AGP and HF in the variational sense, but symmetry breaking occurs incorrectly possibly because of the inherent structure of the trial APG wave function. By taking into account the linear combination of APG, one can overcome the problem; in my experience, the determinant polynomial form for the linear combination exhibits superior property as the variational wave function and a plausible behavior as the chemical bond. Since the APG calculation has been quite time-consuming, I introduce a low-rank approximation to the geminals without significant degradation of the calculated result. I suggest such an approximate scheme is promising and that, with the future improvements of the algorithm, APG-based calculation would provide reasonably accurate and intuitively recognizable results that may advance our understanding on the strongly correlated few-body systems.

The hierarchy of geminal theories is shown in FIG. 1.1. The APG without restricted degrees of freedom is the most elaborate method among the APG families like the APSG and APIG.

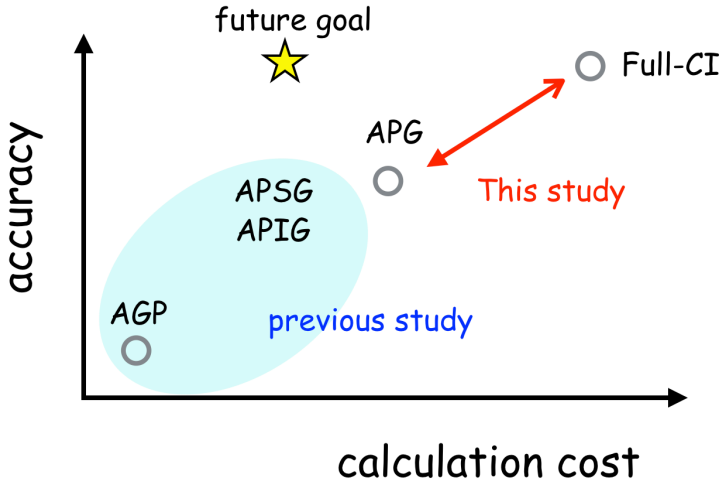

Figure 1.1: The hierarchy of geminal theories.

Chapter 2

Review

First, let me review the geminal theory found in the literatures. The theory has been developed almost independently in different communities. Therefore, I will introduce them separately. The contents explained in this chapter have overlap with those shown in Chapters 1 and 3 but I do so for the benefit of readers unfamiliar to the geminal theories.

2.1 AGP theory in chemical physics community

Coleman did a pioneering work on AGP in 1960s [15], and after long intervention, AGP was paid attention by the condensed-matter community in 1980s as an extension of the BCS theory. Goscinski [26] derived a total energy expression for an N-particle system by neglecting those terms scaling to N as $O\left(N^{-\alpha}\right)$, with $\alpha \geq 2$, and the resulting mathematical structure was discussed. Importantly, the author recognized that the eigenvalue of the geminals, say g_i, exhibits a rapidly decreasing series when sorted by the magnitude, and this fact was used to simplify the calculation. This approach is similar to the low-rank approximation of the geminal as discussed below in this thesis.

Weiner and Goscinski [27] provided a total-energy expression also for the generalized AGP (GAGP), where AGP wave functions are "embedded" in a Slater determinant that is assumed to be strongly orthogonal to the geminals

$$|\psi\rangle = \hat{g}^{n/2}|\psi^{\mathrm{HF}}\rangle, \tag{2.1}$$

where "strongly orthogonal" condition means

$$\int dr g\left(r, r'\right) \varphi_i(r) = 0 \tag{2.2}$$

for all MOs $\varphi_i(r)$ constituting the Slater determinant. Intuitively, the strong orthogonal condition is satisfied when, for example, the geminal is localized in a region where the MOs have small amplitude.

In 2000s, Mazziotti developed the geminal functional theory (GFT) as an extension of DFT by using the geminals in AGP as the fundamental parameter that replaces the density used in DFT [28]. This is based on the KEC theorem stating that any first order density matrix with evenly degenerate eigenvalues can be derived from an N-particle AGP wave function [29]. Here, KEC was named after the researchers Kumar, Erdahl, and Coleman, who contributed to deriving the early version of the theorem [15]. Mazziotti derived an approximate GFT functional and calculated the total energy of small molecules and atoms; the obtained correlation energy was found to be superior to HF and AGP calculations. Compared with the full-CI calculation, the error in the correlation energy was typically $\pm 10\%$; note that GFT is not variational and the total energy can be above or below that of full-CI.

In 2011, Scuseria *et al.* applied AGP to small molecules and compared the result with the one obtained using HF. The AGP trial wave function was prepared from the BCS one by the particle number projection method, known in nuclear physics community [Note that one of the most general formulations for the projection method was given by Hara and Iwasaki [30]]. The AGP wave function with a given number of electrons N is thus

$$|\psi^{\mathrm{AGP}}\rangle = \int_0^{2\pi} \frac{d\theta}{2\pi} \exp(\mathrm{i}\theta(\hat{N} - N))|\psi^{\mathrm{BCS}}\rangle, \tag{2.3}$$

where

$$|\psi^{\mathrm{BCS}}\rangle = \exp\left(\frac{1}{2}\sum_{ab} F_{ab}\hat{a}_a^\dagger \hat{a}_b^\dagger\right)|0\rangle \tag{2.4}$$

and

$$\hat{N} = \sum_i \hat{a}_i^\dagger \hat{a}_i \tag{2.5}$$

is the number operator. Projecting the AGP wave function into an eigenstate of spin, point group symmetry, and time reversal symmetry, the parameters for the geminal F_{ab} were varied with the help of the Onishi-Yoshida formula [23]. There was sizable improvement over HF, but still there is appreciable deviation from the full-CI result.

2.2 APG theory in physical chemistry community

In 1959, Shull [31] called "geminal" the two-electron wave function describing an electron pair and characterized it as an important building block of the full many-body wave function. Afterwards, the property of the geminal was studied in 1950s and 1960s. Hurley *et al.* [19] developed a molecular-orbital (MO) theory of geminal, and further, explained the molecular structure in terms of the repulsive field between the geminals. Parr *et al.*

[32] extended the theory of Hurley *et al.* and proposed to describe the many-body wave function by a linear combination of the antisymmetrized product of the geminals. This is probably the earliest formulation of the interacting APG, which I will study in this thesis.

Because of the complicated calculus associated with APG, Arai [33] introduced a drastic simplification called "strong orthogonality" as follows. Although geminals are comprised of two fermions, they do not behave as boson because, for geminals defined by $\hat{F}[i] = \sum F_{ab}[i]\hat{a}_a^\dagger\hat{a}_b^\dagger$, the commutation relation is not equal to δ_{ik} but is given by

$$\left[\hat{F}^\dagger[i], \hat{F}[k]\right]_- = -2\sum_{a,b} F_{ab}[i]F_{ab}[k] + 4\sum_{a,b,c} F_{ab}[i]F_{cb}[k]\hat{a}_c^\dagger\hat{a}_a \equiv \hat{Q}_{ik}. \tag{2.6}$$

However, they showed that the bosonic property appears when one applies the following restriction for different geminals indexed by i and k

$$\sum_c F_{ca}[i]F_{cb}[k] = 0 \tag{2.7}$$

for all a and b. This condition is called "strong orthogonality". The APG wave function constructed from the strongly orthogonal geminals is called "antisymmetrized product of strongly orthogonal geminals (APSG)" wave function and is denoted as

$$|\psi^{\mathrm{APSG}}\rangle = \prod_m \hat{F}[m]|0\rangle \equiv |F[1]F[2]\cdots\rangle. \tag{2.8}$$

When the geminals are strongly orthogonal, one can significantly simplify the matrix element of an operator, say $\hat{a}_p^\dagger\hat{a}_q$. That is, those geminals that do not have amplitude at the sites indexed by p and q factor out and the matrix element is simplified as

$$\left\langle F[1]F[2]\cdots\left|\hat{a}_p^\dagger\hat{a}_q\right|F[1]F[2]\cdots\right\rangle = \left\langle F\left[s_1\right]F\left[s_2\right]\cdots\left|\hat{a}_p^\dagger\hat{a}_q\right|F\left[s_1\right]F\left[s_2\right]\cdots\right\rangle, \tag{2.9}$$

where the indices $s_1, s_2, \ldots$, which are subset of $1, 2, \cdots$, specify all the geminals having nonzero amplitude at p or q. This fact greatly simplifies the expression of the matrix element of operators when geminals are localized spatially since only a few of them appear in the expression.

For a certain system consisting of four electrons, the strong orthogonality restriction was found moderate although being severe in general [34]. Kutzelnigg [17] reformulated the APSG wave function using the natural spin orbitals (NSOs) and then derived integrodifferential equations to determine NSOs. The author also noticed similarity with the BCS trial wave function and, in addition, proposed another version of APG where each geminal is restricted to form a singlet pair by sharing the same spatial orbital for the up and down spins.

This approximate APG is called the antisymmetrized product of interacting geminals (APIG). APIG is more complex than APSG but numerically more tractable [22]. Nicely

and Harrison [35] applied APIG to diatomic molecules and, for BH, APIG recovers 98% of the correlation energy while APSG does only 80%. In this calculation, the APIG wave function was expanded into Slater determinants to calculate the matrix elements of the total energy, which is feasible only when the number of the Slater determinants is moderate. With this in mind, Carrington and Doggest [36] proposed to restrict the geminals to be spatially localized to within a few sites from the central site. By this, the geminals in APIG are strongly orthogonal when they do not overlap, with an effect of reducing the computational complexity. The resulting geminals, in addition, have a property that can be regarded as a valence bond, as illustrated using LiH. More recent activities can be found in the review paper by Johnson *et al.* [37].

Very recently, a growing number of literatures demonstrate that APSG can capture the static correlation reasonably well while cannot the dynamic one. Here, the static correlation indicates such correlation that essentially requires mixing of qualitatively different Slater determinants, while the dynamical one is such that can be treated by perturbation theory, such as the many-body perturbation theory of a finite order and the random phase approximation. Recent progress in the APSG-based perturbation theory was reviewed by Jeszenszki *et al.* [38].

2.3 JAPG theory in condensed matter physics community

One can introduce the correlation of the electron pairs by applying the Jastrow factor to the AGP wave function. Note that this is an extension of the conventional variational Monte Carlo simulation where the single Slater determinant is used instead of AGP. Casula and Sorella [39] used the Jastrow function consisting of a pair-wise function of the distance

$$J\left(r_1, \cdots, r_N\right) = \prod_{i.j} \exp\left[f\left(|r_i - r_j|\right)\right] \tag{2.10}$$

to form the Jastrow AGP (JAGP) trial wave function as

$$J\left(r_1, \cdots, r_N\right) \psi^{\mathrm{AGP}}\left(r_1, \cdots, r_N\right). \tag{2.11}$$

The geminals constituting AGP and the parameters for f were then variationally determined. They expanded the JAGP into a linear combination of Slater determinants to evaluate the total energy. Because of the very large number of the determinants thereby generated, they sampled them using a Monte Carlo simulation technique. They demonstrated the accuracy using isolated atoms.

Tahara and Imada [40] applied the particle number projection method and the symmetry projection method to the JAGP method. The method was subsequently applied to strongly correlated-condensed matter problem such as superconductivity [See, for example [41]].

2.4 Fermion pair theory in nuclear physics community

Onishi and Yoshida [23] provided detailed formulation for the generator coordinate method, a model for the deformed nucleus proposed in 1950s, using the BCS trial wave function. They provided a formula for the matrix element of the overlap and the Hamiltonian, which is called the Onishi formula or the Onishi-Yoshida formula. It is noteworthy that they also used the particle number and symmetry projection methods. This method also affected many of the subsequent formulations done in condensed-matter physics, quantum chemistry, and nuclear physics.

After determining the BCS wave function, which is called the Hartree-Fock-Bogoliubov (HFB) wave function in nuclear physics, it is used as the HFB vacuum to obtain various excited configurations. These configurations are used to do a CI calculation to obtain the absorption spectrum. The associated Hamiltonian matrix element needs apparently complicated formula but a compact expression was very recently given, for example, by Mizusaki and Oi [42].

2.5 AGP-CI

Here I review AGP-CI by Uemura *et al.* [24] to clarify the difference in property between APG and AGP-CI. When the Slater determinant is used, the convergence of the CI series (Eq. (1.7)) is very slow even if one optimizes the MOs. However, when using AGPs (Eq. (1.10)), one can greatly speed up the convergence. They applied AGP-CI to the water molecule and the Hubbard model. In the case of water molecule, they got very accurate results (almost the same as full-CI) by using only 10 terms. Namely, by using the anti-symmetric tensor decomposition like Eq. (1.8), the wave function can be represented compactly.

The AGP-CI was shown unstable numerically, and in addition, the variational calculation requires many iterations. Considering that the AGP-CI series grows exponentially with the system size and hence that the number of variational parameters also grows exponentially, the instability and the slow convergence problem will affect the calculation more severely. In this context, it is important to reduce the variational parameter. I will show below that the polynomial APG is advantageous in this respect.

2.6 HFB and GCM

In the mid-1950s, the BCS theory was proposed and this BCS wave function or its generalization, the Hartree-Fock-Bogoliubov (HFB) wave function have been used not only in the studies of superconductivity but also in the studies of other many-body quantum

systems. In those studies, one needs to determine antisymmetric matrices constituting the BCS wave functions [43, 44].

As shown in section 2.1, the BCS wave function can be written as

$$|\psi^{\mathrm{BCS}}\rangle = \exp\left(\frac{1}{2}\sum_{ab} F_{ab} a_a^\dagger a_b^\dagger\right)|0\rangle. \tag{2.12}$$

This BCS wave function has a form very similar to that of the AGP wave function. When one transforms the creation operator $a^\dagger$ as

$$\overline{a}_m^\dagger = \sum_{m'} U_{mm'}^* a_{m'}^\dagger, \tag{2.13}$$

one can rewrite the BCS wave function as

$$|\psi^{\mathrm{BCS}}\rangle = \prod_{m>0}\left(1 + s_m c_m \overline{a}_{\tilde{m}}^\dagger \overline{a}_m^\dagger\right)|0\rangle. \tag{2.14}$$

where U is a unitary matrix, while s and c have the following relationship:

$$\left(U^T F U\right)_{mn} = s_n^* c_n \delta_{m\tilde{n}}. \tag{2.15}$$

Here $\tilde{m}$ is the partner canonical to m. Dobaczewski showed how to convert the different antisymmetric matrix, $F[\mu]$ and $F[\nu]$, to the canonical form simultaneously [45].

In the field of nuclear physics, to calculate many-body quantum systems, the generator coordinate method (GCM) was developed. In nuclear many-body problems, it is important which type of the trial wave function to choose in GCM. After the advent of BCS theory, the HFB wave function has been used for the trial wave function of GCM in the calculation of the closed shell model.

In this thesis, I follow the version of the GCM theories, where the Onishi-Yoshida formula [23] and the Mizusaki-Oi formula [42], are used to calculate the overlap between the HFB wave function, or AGP wave function as well as the Hamiltonian matrix.

In the Onishi-Yoshida formula, the overlap of the AGP wave function becomes

$$\langle F[\lambda]|F[\mu]\rangle = \exp\left(\frac{1}{2}\mathrm{tr}\left[\ln(1 + F[\mu]F[\lambda]^T t)\right]\right)\Big|_{t^{\frac{n}{2}}}. \tag{2.16}$$

Here, T means the transpose of a matrix and $|_{t^{n/2}}$ means to extract the $(n/2)$ th-order coefficient of the polynomial (with respect to the auxiliary variable t). The matrix elements of the Hamiltonian was also given by the Onishi-Yoshida formula (see 3.3 for details). These formulae made it possible to obtain the AGP wave function via variational calculation. It is also possible to obtain formula for the derivatives although it is more convenient to use the Mizusaki-Oi formula to calculate the high-order derivatives (see 3.3 for details).

A similar formula for the APG wave function is not known, so that I propose a method to transform the APG wave function to a linear combination of the AGP wave functions as detailed in the subsequent chapter.

2.7 Embedding theory

Here I explain the embedding theory developed so far to treat strongly correlated impurity sites surrounded by a weakly correlated environment. The basic idea is as follows.

When one defines the impurity state as α and the bath state as β, one can decompose the entire wave function as

$$|\Psi\rangle = \sum_{i=1}^{M} \lambda_i |\alpha_i\rangle |\beta_i\rangle . \tag{2.17}$$

This is the Schmidt decomposition. The proof of the Schmidt decomposition is as follows.

Let $|a\rangle$ and $|b\rangle$ be orthonormal bases of A and B respectively. The basis of the quantum system AB then becomes

$$|\Psi^{AB}\rangle = \sum_{j=1}\sum_{k=1} \psi_{jk}|a_j\rangle|b_k\rangle, \tag{2.18}$$

where ψ_{jk} is a matrix representing the probability amplitude. One can rewrite ψ_{jk} using the singular value decomposition as

$$\psi_{jk} = \sum_i u_{ji}\lambda_i v_{ik}, \tag{2.19}$$

where u and v are unitary matrices and λ is the singular value. Substituting Eq. (2.19) into Eq. (2.18), it becomes

$$\begin{aligned} |\Psi^{AB}\rangle &= \sum_{j=1}\sum_{k=1}\left(\sum_i u_{ji}\lambda_i v_{ik}\right)|a_j\rangle|b_k\rangle \\ &= \sum_i \lambda_i \left(\sum_{j=1} u_{ji}|a_j\rangle\right)\left(\sum_{k=1} v_{ik}|b_k\rangle\right). \end{aligned} \tag{2.20}$$

By redefining the state vectors as

$$|\alpha_i\rangle = \sum_{j=1} u_{ji}|a_j\rangle \tag{2.21}$$

$$|\beta_i\rangle = \sum_{k=1} v_{ik}|b_k\rangle \tag{2.22}$$

then one arrives at the Schmidt decomposition (Eq. (2.17)).

The number of the bath state is larger than that of the impurity state in this Schmidt decomposition, but one only needs to take the sum of the number of impurity state M. Because of this fact, one can reduce the calculation cost very much.

The density matrix embedding method (DMET) [7, 8] uses the Schmidt decomposition and obtains the impurity state by requiring its density matrix to match that of the bath.

DMET is not a variational method. DMET is similar to the dynamical mean-field theory (DMFT), where Green's function is embedded instead of the wave function, but is simpler than DMFT in that the time-dependence is lacking in DMET. The computational cost of DMET is relatively low, while the accuracy is reasonably good for a number of systems. DMET has a known problem when HF is chosen for the bath and full-CI is used for the impurity. This is because the HF density matrix can describe only a subset of all possible forms for the density matrices.

To avoid the problem originated from the inherent structure of HF, Tsuchimochi *et al.* developed a new method of embedding the full-CI in AGP [46], which does not have such a structural problem. They applied this full-CI in AGP method to the Hubbard model or a Hydrogen ring and got accurate results. They noted, however, the application to multi-impurity systems is not easy. They used the projection method to fix the number of electrons, where a virtual phase is introduced into the HFB wave function and integration is done with respect to the phase. However, when this operation is performed, information is numerically lost leading to an inaccurate result for the multi-impurity systems. Therefore, I consider that one needs to avoid the projection method. In this thesis, instead of the projection method, I explicitly collect terms having n creation operators by introducing an auxiliary variable. The details are shown in section 3.7.1.

2.8 Hubbard model

I demonstrate the performance of geminal wave functions using the Hubbard model in this study. The details are shown in section 3.9. Here I review the Hubbard model. The Hubbard model was developed by John Hubbard in 1963 [47] to describe correlated electrons in solids. The Hamiltonian of the Hubbard model has the nearest-neighbor hopping term and the Coulomb interaction term. It assumes that the Coulomb interaction works only between electrons on the same site. Despite such a simple form, the Hubbard model has explained many behaviors of strongly correlated electrons. Also in the computational physics, the Hubbard model has been used as a benchmark to demonstrate or compare the performance of various methods. In particular, many DMRG studies used the Hubbard model as a benchmark. Therefore, I also used the Hubbard model to compare several geminal methods. Note that, there is a more general form Hamiltonian of the Hubbard model, which includes long-range Coulomb interaction, called the Pariser-Parr-Pople (PPP) Hamiltonian [48, 49]. The PPP Hamiltonian has been applied to describe electrons in molecular orbitals in chemistry.

Chapter 3

Formulation

3.1 Expression of geminal theories

Here I summarize the expression of the geminal wave function used in this thesis.

The geminal $\hat{F}$ can be defined as

$$\hat{F}[k] \equiv \sum_{a,b} F[k]_{ab} c_a^\dagger c_b^\dagger \tag{3.1}$$

where F is an antisymmetric matrix, $c^\dagger$ is the creation operator and k shows the types of geminal. The AGP wave function is then written as

$$|\Psi_{\mathrm{AGP}}\rangle = \hat{F}^{\frac{n}{2}} |0\rangle \equiv |F\rangle \,, \tag{3.2}$$

where n is the number of electrons. One can rewrite the AGP wave function as

$$|F\rangle \propto \exp[\hat{F}t]\Big|_{t^{\frac{n}{2}}} \,, \tag{3.3}$$

where $|_{t^N}$ means to extract the coefficient of t^N. In this thesis, I also represent the AGP wave function as $|tF\rangle$ to emphasize the auxiliary variable t. Note that the AGP wave function is the Hartree-Fock-Bogoliubov (HFB) wave function with a fixed number of electrons.

As the geminal describes the electron correlation within the pair, the AGP wave function is a mean-field theory of an electron pair. The AGP wave function is comprised of only one type of geminal. One can alternatively construct the wave function using different geminals for different pairs as

$$|\Psi_{\mathrm{APG}}\rangle = \hat{F}[1]\hat{F}[2]\cdots\hat{F}[n/2]\,|0\rangle \,. \tag{3.4}$$

The resulting wave function is called the APG wave function.

One can classify the geminal theories from the viewpoint of tensor decomposition. The full-CI wave function is written as

$$|\Psi\rangle = \sum_{i_1,\cdots,i_n} A_{i_1\cdots i_n} c_{i_1}^\dagger \cdots c_{i_n}^\dagger |0\rangle \,, \tag{3.5}$$

where $A_{i_1\cdots i_n}$ is an antisymmetric tensor of degree n. Using the antisymmetric tensor decomposition, one can represent the tensor using the canonical format as

$$A_{i_1\cdots i_n} = \sum_{r=1}^{R} \left(\sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) F[1]^{(r)}_{\sigma(i_1)\sigma(i_2)} F[2]^{(r)}_{\sigma(i_3)\sigma(i_4)} \cdots F[n/2]^{(r)}_{\sigma(i_{n-1})\sigma(i_n)} \right), \tag{3.6}$$

where $S_{n/2}$ is the permutation group of degree $n/2$. The minimum number of R is called the rank of the tensor, and the minimum number to approximate the tensor is called the approximate rank. When truncating the series at $R = 1$, it becomes the APG wave function and when one further assumes all the geminals to be identical to each other as $F[1] = F[2] = \cdots = F[n/2]$, it becomes the AGP wave function.

One can alternatively decompose the tensor using the Tucker format [50] as

$$A = \sum_{r_1,\cdots,r_{n/2}} C_{r_1\cdots r_{n/2}} F[1]^{(r_1)} F[2]^{(r_2)} \cdots F[n/2]^{(r_{n/2})}, \tag{3.7}$$

where C is called the core tensor. In the canonical format, the core tensor is diagonal, so one can recognize the canonical format is a special case of the Tucker format. I will use this geminal wave function in 3.2.3 and 3.2.4.

When all geminals in Eq. (3.6) are the same ($F[1] = F[2] = \cdots = F[n/2]$) and $R \geq 1$, it becomes the AGP-CI wave function [24, 51], which is a linear combination of AGPs as

$$|\Psi_{\mathrm{AGP-CI}}\rangle = \sum_{k=1}^{R} \hat{F}[k]^{\frac{n}{2}} |0\rangle \,. \tag{3.8}$$

I will represent the AGP-CI of R terms as AGP-CI(R) hereafter.

Instead of taking a linear combination, one can let electron pairs correlate by multiplying AGP with the correlation, as detailed in 3.7.

3.2 Polynomial decomposition

The APG wave function has so far been thought too complex to be determined variationally. I overcome this problem by using the polynomial decomposition [52], which is a tensor decomposition method developed in mathematics.

In mathematics, much effort has been made to find the minimum number of term required for the decomposition. The polynomial decomposition means to describe an M

th-order polynomial $\sum_{n_1\cdots n_N}^{\sum n_i=M} a_{n_1\cdots n_N} x_1^{n_1}\cdots x_N^{n_N}$ as a sum of multiples of a linear form as $\sum_{n_1\cdots n_N}^{\sum n_i=M} a_{n_1\cdots n_N} x_1^{n_1}\cdots x_N^{n_N} = \sum_{j=1}^{R} \lambda_j (\sum_k C_{jk} x_k)^M$. The decomposition is called the Waring decomposition [53, 54, 55]. The Waring decomposition is also called sum-power decomposition, the canonical decomposition and the rank-1 decomposition.

The simplest Waring decomposition is for monomial. Fischer's formula [56] can be applied to a monomial of the geminal, $\hat{F}[1]\cdots\hat{F}[N]$, to yield

$$\begin{aligned}&\hat{F}[1]\hat{F}[2]\cdots\hat{F}[N]\\ &= \frac{1}{2^{N-1}N!}\sum_{i_2=0}^{1}\cdots\sum_{i_N=0}^{1}(-1)^{i_2+\cdots+i_N}\left(\hat{F}[1]+(-1)^{i_2}\hat{F}[2]+\cdots+(-1)^{i_N}\hat{F}[N]\right)^N\end{aligned} \tag{3.9}$$

This indicates that APG can be transformed to a liner combination of AGPs, so that one can apply the Onishi-Yoshida formula to do a variational calculation. Likewise, one can do a variational calculation if the trial wave function has a polynomial form of the geminals.

Below, I briefly show the polynomials used in this thesis, although more detailed expression is shown in Appendix A.

3.2.1 Elementary symmetric polynomial

The APG wave function is a monomial of degree N. The most natural extension of APG is probably to use M ($M \geq N$) different geminals to form a symmetric polynomial as

$$\sum_{1\leq i_1<i_2<\cdots<i_N\leq M}\hat{F}[i_1]\hat{F}[i_2]\cdots\hat{F}[i_N] \quad\equiv\quad e_N(\hat{F}[1],\hat{F}[2],\cdots,\hat{F}[M]). \tag{3.10}$$

This polynomial is known as the elementary symmetric polynomial. Note that all terms are superlinear; multiple product of a geminal does not appear. The elementary symmetric polynomial is a special example of the tensor decomposition (Eq. (3.6)).

The Waring decomposition of an elementary symmetric polynomial is as follows [57]: for even N,

$$\begin{aligned}&e_N(\hat{F}[1],\hat{F}[2],\cdots,\hat{F}[M])\\ =\;& \frac{1}{2^N(M-N)N!}\sum_{i_1=0}^{1}\cdots\sum_{i_M=0}^{1}(-1)^{i_1+\cdots+i_M}\begin{pmatrix} M-N/2-(i_1+\cdots+i_M)-1\\ N/2-(i_1+\cdots+i_M)\end{pmatrix}\\ &\times(M-2(i_1+\cdots+i_M))\left((-1)^{i_1}\hat{F}[1]+(-1)^{i_2}\hat{F}[2]+\cdots+(-1)^{i_M}\hat{F}[M]\right)^N,\end{aligned} \tag{3.11}$$

where $i_1 + \cdots + i_M \leq N/2$ is assumed. For odd N,

$$\begin{aligned} & e_N(\hat{F}[1], \hat{F}[2], \cdots, \hat{F}[M]) \\ = & \frac{1}{2^{N-1}N!} \sum_{i_1=0}^{1} \cdots \sum_{i_M=0}^{1} (-1)^{i_1+\cdots+i_M} \begin{pmatrix} M - (N-1)/2 - (i_1 + \cdots + i_M) - 1 \\ (N-1)/2 - (i_1 + \cdots + i_M) \end{pmatrix} \\ & \times \left((-1)^{i_1} \hat{F}[1] + (-1)^{i_2} \hat{F}[2] + \cdots + (-1)^{i_M} \hat{F}[M] \right)^N, \end{aligned} \tag{3.12}$$

where $i_1 + \cdots + i_M \leq (N-1)/2$ is assumed.

3.2.2 Complete homogeneous symmetric polynomial

By allowing multiproduct of a geminal in an elementary symmetric polynomial, one can make the complete homogeneous symmetric polynomial,

$$\sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_N \leq M} \hat{F}[i_1]\hat{F}[i_2] \cdots \hat{F}[i_N] \equiv h_N(\hat{F}[1], \hat{F}[2], \cdots, \hat{F}[M]). \tag{3.13}$$

In a complete homogeneous symmetric polynomial, there is no known formula for the optimal Waring decomposition, so that I apply the Fischer formula to decompose each term.

3.2.3 Permanent polynomial

Here I make polynomials from Eq. (3.7). When all the coefficients, C, are equal to 1, Eq. (3.7) becomes a permanent polynomial,

$$\begin{aligned} \sum_{\sigma \in S_N} \hat{F}[1]^{(\sigma(1))} \hat{F}[2]^{(\sigma(2))} \cdots \hat{F}[N]^{(\sigma(N))} & \equiv \sum_{\sigma \in S_N} \hat{F}[1, \sigma(1)]\hat{F}[2, \sigma(2)] \cdots \hat{F}[N, \sigma(N)] \\ & \equiv \mathrm{perm}_N(\hat{F}[1,1], \cdots, \hat{F}[N,N]). \end{aligned} \tag{3.14}$$

To decompose a permanent polynomial [58], first I rewrite it as

$$\mathrm{perm}_N(\hat{F}[1,1], \cdots, \hat{F}[N,N]) = \frac{1}{2^{N-1}} \sum_{\epsilon = \{-1,1\}, \epsilon_1 = 1} \prod_{1 \leq i \leq N} \sum_{1 \leq j \leq N} \epsilon_i \epsilon_j \hat{F}[i,j]. \tag{3.15}$$

Afterward, I apply the Fischer formula term by term.

3.2.4 Determinant polynomial

By introducing the sign to the coefficient of the permanent polynomial Eq. (3.14), one can make a determinant polynomial,

$$\begin{aligned} \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma) \hat{F}[1]^{(\sigma(1))} \hat{F}[2]^{(\sigma(2))} \cdots \hat{F}[N]^{(\sigma(N))} & \equiv \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma) \hat{F}[1, \sigma(1)]\hat{F}[2, \sigma(2)] \cdots \hat{F}[N, \sigma(N)] \\ & \equiv \det_N(\hat{F}[1,1], \cdots, \hat{F}[N,N]) \, |0\rangle \, . \end{aligned} \tag{3.16}$$

The determinant polynomial is the most studied polynomial in mathematics. However, a general form for the Waring decomposition is not known. In the case of $N = 3$, one can conveniently apply the formula by Derksen [59]:

$$\begin{aligned}\det{}_3(\hat{F}[1,1],\cdots,\hat{F}[3,3]) = \frac{1}{2}\Big[\quad & (F[1,3]+F[1,2])(F[2,1]-F[2,2])(F[3,1]+F[3,2]) \\ & +(F[1,1]+F[1,2])(F[2,2]-F[2,3])(F[3,2]+F[3,3]) \\ & +2F[1,2](F[2,3]-F[2,1])(F[3,3]+F[3,1]) \\ & +(F[1,3]-F[1,2])(F[2,2]+F[2,1])(F[3,2]-F[3,1]) \\ & +(F[1,1]-F[1,2])(F[2,3]+F[2,2])(F[3,3]-F[3,2])\Big],\end{aligned} \tag{3.17}$$

before using the Fischer formula term by term. By this, one can reduce the number of terms from 6 to 5; compare the left-hand side and the right-hand side of Eq. (3.17). Then I use the Fischer formula term by term.

3.3 Total energy formula

Here I show the details of deriving matrix elements of Hamiltonian and their derivatives. When using the Onishi-Yoshida commutation relation formula [23],

$$\left[c_\alpha, \exp\left(\hat{F}\right)\right] = \sum_\delta F_{\alpha\delta} c_\delta^\dagger \exp\left(\hat{F}\right), \tag{3.18}$$

one can derive

$$c_\alpha \left|tF\right\rangle = \sum_\gamma tF_{\alpha\gamma} c_\gamma^\dagger \left|tF\right\rangle. \tag{3.19}$$

This makes it possible to transform the annihilation operator into a linear combination of the creation operators. Using this relationship, one can obtain an analytic formula for the matrix element of the one-body term of the Hamiltonian as

$$\begin{aligned}\left\langle tF[\lambda]\right| c_\alpha^\dagger c_\beta \left|tF[\mu]\right\rangle &= \left\langle tF[\lambda]\right| c_\alpha^\dagger \sum_\gamma tF[\mu]_{\beta\gamma} c_\gamma^\dagger \left|tF[\mu]\right\rangle \\ &= \sum_\gamma F[\mu]_{\beta\gamma} \frac{\partial}{\partial F[\mu]_{\alpha\gamma}} \left\langle tF[\lambda]|tF[\mu]\right\rangle,\end{aligned} \tag{3.20}$$

which yields

$$\left\langle tF[\lambda]\right| c_a^\dagger c_b \left|tF[\mu]\right\rangle\Big|_{t^N} = \left(\frac{F[\mu]F[\lambda]^\dagger t^2}{1+F[\mu]F[\lambda]^\dagger t^2}\right)_{ba} \exp\left(\frac{1}{2}\mathrm{tr}\left[\ln(1+F[\mu]F[\lambda]^\dagger t^2)\right]\right)\Big|_{t^N}. \tag{3.21}$$

Similarly one can obtain the formula for the two-body term as

$$\begin{aligned}\langle tF[\lambda]|\, c_p^\dagger c_q^\dagger c_s c_r\, |tF[\mu]\rangle\Big|_{t^N} \;=\; & \left(\left[\frac{F[\mu]F[\lambda]^\dagger t^2}{1+F[\mu]F[\lambda]^\dagger t^2}\right]_{rp} \left[\frac{F[\mu]F[\lambda]^\dagger t^2}{1+F[\mu]F[\lambda]^\dagger t^2}\right]_{sq} \right. \\ & - \left[\frac{F[\mu]F[\lambda]^\dagger t^2}{1+F[\mu]F[\lambda]^\dagger t^2}\right]_{rq} \left[\frac{F[\mu]F[\lambda]^\dagger t^2}{1+F[\mu]F[\lambda]^\dagger t^2}\right]_{sp} \\ & \left. + \left[\frac{t}{1+F[\mu]F[\lambda]^\dagger t^2}F[\mu]\right]_{rs} \left[F[\lambda]^\dagger \frac{t}{1+F[\mu]F[\lambda]^\dagger t^2}\right]_{qp} \right) \\ & \times \exp\left(\frac{1}{2}\mathrm{tr}\left[\ln(1+F[\mu]F[\lambda]^\dagger t^2)\right]\right)\Big|_{t^N}. \end{aligned} \tag{3.22}$$

It is very hard to directly differentiate the overlap matrix and the matrix elements of Hamiltonian with respect to F. It is simpler to replace the derivatives as

$$\frac{1}{t}\frac{\partial}{\partial F[\mu]_{cd}}\langle tF[\lambda]|tF[\mu]\rangle = \langle tF[\lambda]|\, c_c^\dagger c_d^\dagger\, |tF[\mu]\rangle\,. \tag{3.23}$$

So far, I have shown what is needed to do the variational calculation. Although they are convenient in getting the total energy, it is not the case when there are many creation and annihilation operators, because one needs to differentiate the Onishi-Yoshida formula many times using Eq. (3.20).

Hence, I use another formula which was developed in the field of nuclear physics. Mizusaki and Oi [42] showed a formula

$$\begin{aligned} & \langle F[\lambda]|\, c_{a_1}c_{a_2}\cdots c_{a_{2n-1}}c_{a_{2n}}c_{b_1}^\dagger c_{b_2}^\dagger \cdots c_{b_{2m-1}}^\dagger c_{b_{2m}}^\dagger\, |F[\mu]\rangle \\ = & \sum_{\sigma\in S_{2m+2n}} \mathrm{sgn}(\sigma) X^{-1}_{\sigma(b_{2m})\sigma(b_{2m-1})}\cdots X^{-1}_{\sigma(b_2)\sigma(b_1)} X^{-1}_{\sigma(a_{2n}+M)\sigma(a_{2n-1}+M)}\cdots X^{-1}_{\sigma(a_2+M)\sigma(a_1+M)} \\ & \times \langle F[\lambda]|F[\mu]\rangle\,, \end{aligned} \tag{3.24}$$

where X is the $2M\times 2M$ matrix of the form

$$X = \begin{pmatrix} tF[\mu] & 1 \\ -1 & tF[\lambda]^\dagger \end{pmatrix} \tag{3.25}$$

$$X^{-1} = \begin{pmatrix} F[\lambda]^\dagger t\frac{1}{1+F[\mu]F[\lambda]^\dagger t^2} & \frac{1}{1+F[\lambda]^\dagger F[\mu]t^2} \\ -\frac{1}{1+F[\mu]F[\lambda]^\dagger t^2} & \frac{1}{1+F[\mu]F[\lambda]^\dagger t^2}F[\mu]t \end{pmatrix} \equiv \begin{pmatrix} Z^1 & Z^3 \\ Z^4 & Z^2 \end{pmatrix}. \tag{3.26}$$

The creation operators in Eq. (3.24) give rise to the former half elements (1 to M) of X^{-1} while the annihilation operators give rise to the latter half ($M+1$ to $2M$).

With this Mizusaki Oi formula, one can get the derivatives easier than with the Onishi Yoshida formula.

Here I detail the formula further. With the notations, $Z^1 \sim Z^4$, given in Eq. (3.26), one can derive the following equations,

$$\langle F[\lambda]|\, c_i^\dagger c_j^\dagger \,|F[\mu]\rangle = Z^1_{ji}\, \langle F[\lambda]|F[\mu]\rangle \tag{3.27}$$

$$\langle F[\lambda]|\, c_i c_j \,|F[\mu]\rangle = Z^2_{ji}\, \langle F[\lambda]|F[\mu]\rangle \tag{3.28}$$

$$\langle F[\lambda]|\, c_i c_j^\dagger \,|F[\mu]\rangle = Z^3_{ji}\, \langle F[\lambda]|F[\mu]\rangle \tag{3.29}$$

$$\langle F[\lambda]|\, c_a c_b c_c^\dagger c_d^\dagger \,|F[\mu]\rangle = (Z^1_{dc}Z^2_{ba} - Z^3_{db}Z^3_{ca} + Z^3_{da}Z^3_{cb})\, \langle F[\lambda]|F[\mu]\rangle \tag{3.30}$$

$$\begin{aligned} \langle F[\lambda]|\, c_a c_b c_c^\dagger c_d^\dagger c_e^\dagger c_f^\dagger \,|F[\mu]\rangle \;=\; & \Big[Z^1_{fe}(Z^1_{dc}Z^2_{ba} - Z^3_{db}Z^3_{ca} + Z^3_{da}Z^3_{cb}) \\ & - Z^1_{fd}(Z^1_{ec}Z^2_{ba} - Z^3_{eb}Z^3_{ca} + Z^3_{ea}Z^3_{cb}) \\ & + Z^1_{fc}(Z^1_{ed}Z^2_{ba} - Z^3_{eb}Z^3_{da} + Z^3_{ea}Z^3_{db}) \\ & - Z^3_{fb}(Z^1_{ed}Z^3_{ca} - Z^1_{ec}Z^3_{da} + Z^3_{ea}Z^1_{dc}) \\ & + Z^3_{fa}(Z^1_{ed}Z^3_{cb} - Z^1_{ec}Z^3_{db} + Z^3_{eb}Z^1_{dc}) \Big]\, \langle F[\lambda]|F[\mu]\rangle \,. \end{aligned} \tag{3.31}$$

Then, one can rewrite $Z^1 \sim Z^4$ as a Taylor series as

$$Z^1 = \sum_{n=0} F[\lambda]^\dagger (-1)^n \left(F[\mu]F[\lambda]^\dagger\right)^n t^{2n+1} \tag{3.32}$$

$$Z^2 = \sum_{n=0} (-1)^n \left(F[\mu]F[\lambda]^\dagger\right)^n F[\mu] t^{2n+1} \tag{3.33}$$

$$Z^3 = \sum_{n=0} (-1)^n \left(F[\lambda]^\dagger F[\mu]\right)^n t^{2n} \tag{3.34}$$

$$Z^4 = -\sum_{n=0} (-1)^n \left(F[\mu]F[\lambda]^\dagger\right)^n t^{2n}. \tag{3.35}$$

Therefore the formulae are simplified, for example, as

$$\begin{aligned} \langle F[\lambda]|\, c_i c_j^\dagger \,|F[\mu]\rangle\Big|_{t^2} &= Z^3_{ji}\, \langle F[\lambda]|F[\mu]\rangle\big|_{t^2} \\ &= Z^3_{ji}\big|_{t^2} + \langle F[\lambda]|F[\mu]\rangle|_{t^2} \\ &= \left(F[\lambda]^\dagger F[\mu]\right)_{ji} + \frac{1}{2}\mathrm{tr}[F[\lambda]^\dagger F[\mu]]. \end{aligned} \tag{3.36}$$

This is the working equation for the total energy used in my work.

3.4 Fredholm Pfaffian

The Fredholm Pfaffian is defined as

$$\mathrm{pf}(1+At) = \exp\left(\frac{1}{2}\mathrm{tr}[\ln(1+At)]\right). \tag{3.37}$$

Following the fact that $\det(1+At) = \exp(\mathrm{tr}[\ln(1+At)])$ is called the Fredholm determinant, I call Eq. (3.37) the Fredholm Pfaffian. The relationship between determinant and Pfaffian is shown in Appendix B.

Note that the Fredholm Pfaffian is identical to the overlap of the AGP wave function in the Onishi-Yoshida formula. In this section, I will see the details on the calculation of the Fredholm Pfaffian.

To calculate the Fredholm Pfaffian, one use the Taylor series,

$$e^x = \sum_{n=0} \frac{x^n}{n!} \tag{3.38}$$

and

$$\ln(1+x) = \sum_{n=1} \frac{(-1)^{n+1}}{n} x^n. \tag{3.39}$$

Using Eqs. (3.38) and (3.39), Eq. (3.37) is rewritten as

$$\exp\left(\frac{1}{2}\mathrm{tr}[\ln(1+At)]\right) = \sum_{n=0} \frac{1}{n!}\frac{1}{2^n}\left(\mathrm{tr}\left[\sum_{m=1} \frac{(-1)^{m+1}}{m} A^m t^m\right]\right)^n. \tag{3.40}$$

When I calculate the overlap of the AGP wave function, I use the n th-order coefficient of the right-hand side of Eq. (3.40). For example, the coefficient of t^4 is

$$\frac{1}{384}\mathrm{tr}[A]^4 - \frac{1}{32}\mathrm{tr}[A]^2\mathrm{tr}[A^2] + \frac{1}{32}\mathrm{tr}[A^2]^2 + \frac{1}{12}\mathrm{tr}[A]\mathrm{tr}[A^3] - \frac{1}{8}\mathrm{tr}[A^4]. \tag{3.41}$$

There is a more sophisticated method derived using the Cauchy integral although I could not achieve sufficient numerical precision by this method. Nevertheless it is instructive to briefly introduce the method as follows.

Suppose that there is a polynomial $f(z)$,

$$f(z) = \sum_{k=0}^{\infty} a_k z^k; \tag{3.42}$$

the coefficient a_n can be written as

$$\begin{aligned} a_n &= \frac{f^{(n)}(0)}{n!} \\ &= \frac{1}{2\pi i}\int_{|z|=r} \frac{f(z)}{z^{n+1}} dz \\ &= \frac{1}{2\pi r^n}\int_0^{2\pi} e^{-in\theta} f(re^{i\theta}) d\theta. \end{aligned} \tag{3.43}$$

The last equation of Eq. (3.43) is the Cauchy integral. One can apply this method to get the coefficient of Eq. (3.40). For that purpose, one needs to provide accurate derivatives. Usually, it is done using numerical differentiation, such as

$$F'(x_0) = \frac{F(x_0+h) - F(x_0-h)}{2h} + O(h^2). \tag{3.44}$$

One is faced with the information loss when taking very small h.

One can overcome this problem partially by using a technique [60] as

$$F'(x_0) = \frac{\mathrm{Im}(F(x_0+ih))}{h}. \tag{3.45}$$

This equation is obtained using the Taylor series of $F(x_0+ih)$,

$$F(x_0+ih) = F(x_0) + ihF'(x_0) - \frac{1}{2!}h^2F^{(2)}(x_0) - \frac{1}{3!}ih^3F^{(3)}(x_0) + \cdots, \tag{3.46}$$

which yields

$$\frac{\mathrm{Im}(F(x_0+ih))}{h} = F'(x_0) - \frac{1}{3!}h^2F^{(3)}(x_0) + \cdots. \tag{3.47}$$

This formula allows one to calculate the derivatives with the error $\mathrm{O}(h^2)$. Since this method does not include the difference, one can expect to retain the accuracy even when taking very small h.

Here I test the accuracy of this complex numerical difference method by applying to $\frac{\partial}{\partial F}\langle F|F\rangle$. The difference from the derivatives obtained using the analytical derivative is -2×10^{-16} when using $h = 1\times10^{-7}$ for a case where the difference is 1.8×10^{-8} when using the simple numerical differentiation. I can thus get much more accurate derivatives from the complex numerical difference method. Note that, I cannot reduce the error in usual numerical difference by using smaller h.

One can apply this derivatives method to the Fredholm Pfaffian as

$$\begin{aligned}\exp\left[\frac{1}{2}\mathrm{tr}[\ln(1+F^\dagger Ft)]\right]\Big|_{t^N} &\equiv f(t)|_{t^N} \\ &= \frac{f^{(N)}(0)}{N!} \\ &= \frac{1}{2\pi i}\int_{|z|=r}\frac{f(z)}{z^{N+1}}dz.\end{aligned} \tag{3.48}$$

The integration in Eq. (3.48) can be done using the trapezoidal sums as

$$\begin{aligned}a_0 &= \frac{1}{2\pi r^n}\int_0^{2\pi} e^{-in\theta}f(re^{i\theta})d\theta \\ &= \frac{1}{2\pi r^n}\sum_{k=0}^{N-1} g_k,\end{aligned} \tag{3.49}$$

where

$$g(\theta) \equiv \frac{1}{2\pi r^n} \int_0^{2\pi} e^{-in\theta} f(re^{i\theta}) d\theta g_k \equiv \int_{\theta_k}^{\theta_{k+1}} g(\theta) d\theta. \tag{3.50}$$

The trapezoidal sums can be done using the approximation

$$g(\theta) \approx \frac{(g(\theta_k) + g(\theta_{k+1}))\Delta\theta}{2}, \tag{3.51}$$

where $\theta_k \equiv \frac{2\pi k}{N}$ and $\Delta\theta \equiv \theta_{k+1} - \theta_k = \frac{2\pi}{N}$. Then one gets a formula, for example, for a_0 as

$$\begin{aligned} a_0 &\approx \frac{1}{2\pi r^n} \sum_{k=0}^{N-1} \frac{(g(\theta_k) + g(\theta_{k+1}))}{2} \frac{2\pi}{N} \\ &= \frac{1}{Nr^n} \sum_{k=0}^{N-1} e^{\frac{-in2\pi k}{N}} f(re^{\frac{i2\pi k}{N}}). \end{aligned} \tag{3.52}$$

This works well for a_0, but unexpectedly I cannot achieve enough accuracy when calculating higher-order derivatives.

3.5 Low-rank geminal matrix

In the variational calculation of the geminal wave function, the matrix elements of the Hamiltonian and their derivatives are much heavier than those of the overlap of wave function from the viewpoint of the computational cost. The reason is that, as one can see in Eqs. (3.21) and (3.22), the matrix elements of the Hamiltonian requires to calculate not only the Fredholm Pfaffian but also products of the geminal matrices. The calculation cost of the Fredholm Pfaffian is much lower than that of the product of Fredholm Pfaffian and the matrix. Therefore, if one can write the matrix elements of the Hamiltonian only as the Fredholm Pfaffian, the calculation cost can be reduced. In this context, I use low-rank geminal matrices to reformulate the matrix elements of the Hamiltonian and their derivatives only using the form of Fredholm Pfaffian.

In this section, I reduce the freedom of the geminal matrix f using sum of rank-1 matrices as

$$\begin{aligned} f_{pq} &= \sum_{1 \le k \le n} \left(u_{pk} u_{q\bar{k}} - u_{p\bar{k}} u_{qk} \right) \\ &\equiv \sum_{1 \le k \le n} u_{pk} u_{q\bar{k}} \epsilon_{k\bar{k}} \equiv \sum_{1 \le k \le n} f_{pq}^{(k)}, \end{aligned} \tag{3.53}$$

where u_{ak} is a vector of length M ($1 \le a \le M$) and ϵ is the 2×2 Levi-Civita tensor. Here M is the number of bases and n is the number of geminal pairs. As u is rank-1 matrix, the rank of f is n, while the rank of the original one is $M/2$. One can also write f as

$$f = \sum_{1 \le k \le n} |u_k\rangle \, \epsilon_{k\bar{k}} \, \langle u_{\bar{k}}| \, . \tag{3.54}$$

Note that when using a rank-n matrix, rank-n AGP is equivalent to rank-1 APG:

$$|\Psi\rangle = \frac{1}{n!}\left(\sum_{1\le k\le n}\hat{f}^{(k)}\right)^n|0\rangle = \prod_{1\le k\le n}\hat{f}^{(k)}|0\rangle \tag{3.55}$$

To calculate the Fredholm Pfaffian ($\mathrm{pf}(1+f_1f_2^\dagger)$), one needs to derive $x \equiv f_1 f_2^\dagger$, where the superscript $\dagger$ means to take the Hermitian conjugate. It can be written as

$$\begin{aligned}
x &= \sum_{1\le k,k'\le n}|u_k\rangle\,\epsilon_{k\bar{k}}\,\langle u_{\bar{k}}|v_{k'}\rangle\,\epsilon_{k'\bar{k}'}\,\langle v_{\bar{k}'}| \\
&= \sum_{1\le k,k'\le n}\left(|u_k^1\rangle\,\langle u_{\bar{k}}^2|v_{k'}^1\rangle\,\langle v_{k'}^2| - |u_k^1\rangle\,\langle u_{\bar{k}}^2|v_{k'}^2\rangle\,\langle v_{\bar{k}'}^1| - |u_k^2\rangle\,\langle u_{\bar{k}}^1|v_{k'}^1\rangle\,\langle v_{\bar{k}'}^2| + |u_k^2\rangle\,\langle u_{\bar{k}}^1|v_{k'}^2\rangle\,\langle v_{\bar{k}'}^1|\right) \\
&= \sum_{k,k',i,j}|u_k^i\rangle\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}_{ip}\langle u_k^p|u_{k'}^q\rangle\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}_{qj}\langle v_{k'}^j|\,.
\end{aligned} \tag{3.56}$$

It is found that x is an $n\times n$ matrix, and therefore one can calculate the Fredholm Pfaffian by evaluating an $n\times n$ matrix. It can reduce the calculation cost, because originally x is an $M\times M$ matrix.

Below, I rewrite all the elements needed in the variational calculation (matrix elements of the Hamiltonian and their derivatives) in terms of the Fredholm Pfaffian.

First, the derivative of f with respect to u becomes

$$\frac{\partial f_{pq}}{\partial u_{ak}} = \delta_{ap}u_{q\bar{k}} - \delta_{aq}u_{p\bar{k}} \equiv (\delta_{ak}\star f)_{pq} \tag{3.57}$$

$$\frac{\partial f_{pq}}{\partial u_{a\bar{k}}} = -\delta_{ap}u_{qk} + \delta_{aq}u_{pk} \equiv (\delta_{a\bar{k}}\star f)_{pq}. \tag{3.58}$$

Then,

$$\sum_{pq}\frac{\partial f_{pq}}{\partial u_{ak}}c_p^\dagger c_q^\dagger = 2c_a^\dagger\sum_q c_q^\dagger u_{q\bar{k}} \tag{3.59}$$

$$\sum_{pq}\frac{\partial f_{pq}}{\partial u_{a\bar{k}}}c_p^\dagger c_q^\dagger = -2c_a^\dagger\sum_q c_q^\dagger u_{qk}. \tag{3.60}$$

Therefore, the derivative of the AGP wave function ($|\Psi\rangle = \prod_{1\le l\le n}\hat{f}^{(l)}|0\rangle$) is

$$\frac{\partial}{\partial u_{ak}}|\Psi\rangle = \delta_{ak}\hat{f}^{(k)}\prod_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}|0\rangle = \frac{1}{n!}\left(\delta_{ak}\star\hat{f}^{(k)} + \sum_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}\right)^n|0\rangle \tag{3.61}$$

$$\frac{\partial}{\partial u_{a\bar{k}}}|\Psi\rangle = \delta_{a\bar{k}}\hat{f}^{(k)}\prod_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}|0\rangle = \frac{1}{n!}\left(\delta_{a\bar{k}}\star\hat{f}^{(k)} + \sum_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}\right)^n|0\rangle\,, \tag{3.62}$$

and the derivative of the overlap becomes

$$\frac{\partial}{\partial u_{ak}}\langle\Psi|\Psi\rangle = 2\sum_{f_2\in\{F\}}\mathrm{pf}\left(1+\left(\delta_{ak}\star\hat{f}^{(k)}+\sum_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}\right)f_2^\dagger t\right)|0\rangle \tag{3.63}$$

$$\frac{\partial}{\partial u_{a\bar{k}}}\langle\Psi|\Psi\rangle = 2\sum_{f_2\in\{F\}}\mathrm{pf}\left(1+\left(\delta_{a\bar{k}}\star\hat{f}^{(k)}+\sum_{1\le l\le n}^{l\neq k}\hat{f}^{(l)}\right)f_2^\dagger t\right)|0\rangle\,. \tag{3.64}$$

In this way, one can write derivatives only using sum of the Fredholm Pfaffian.

Next, I show how to describe the first-order density matrix E_1,

$$E_1=\sum_{ab}\frac{\langle\Psi|\,h_{ab}c_a^\dagger c_b\,|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{3.65}$$

In the E_1, one can rewrite $h_{ab}c_a^\dagger c_b\hat{f}^{(1)}$ as (here I omit the sum symbol $\sum$ for simplicity),

$$\begin{aligned}\left(h_{ab}c_a^\dagger c_b\right)\left(f^{(1)}_{p_1q_1}c^\dagger_{p_1}c^\dagger_{q_1}\right) &= h_{ab}f^{(1)}_{p_1q_1}c_a^\dagger\left(\delta_{bp_1}-c^\dagger_{p_1}c_b\right)c^\dagger_{q_1}\\ &= h_{ab}f^{(1)}_{bq_1}c_a^\dagger c^\dagger_{q_1}-h_{ab}f^{(1)}_{p_1q_1}c_a^\dagger c^\dagger_{p_1}c_bc^\dagger_{q_1}\\ &= h_{ac}f^{(1)}_{cb}c_a^\dagger c_b^\dagger-h_{ab}f^{(1)}_{p_1b}c_a^\dagger c^\dagger_{p_1}+h_{ab}f^{(1)}_{p_1q_1}c_a^\dagger c^\dagger_{p_1}c^\dagger_{q_1}c_b\\ &= 2h_{ac}f^{(1)}_{cb}c_a^\dagger c_b^\dagger+\left(f^{(1)}_{p_1}q_1c^\dagger_{p_1}c^\dagger_{q_1}\right)h_{ab}c_a^\dagger c_b\\ &= \left(hf^{(1)}+f^{(1)}h\right)_{ab}c_a^\dagger c_b^\dagger+\hat{f}^{(1)}\hat{h}.\end{aligned} \tag{3.66}$$

Then the state $h_{ab}c_a^\dagger c_b\,|\Psi\rangle$ becomes

$$\begin{aligned}\hat{h}|\Psi\rangle=\hat{h}\hat{f}^{(1)}\cdots\hat{f}^{(n)}|0\rangle &= \left(\left(hf^{(1)}+f^{(1)}h\right)_{ab}c_a^\dagger c_b^\dagger+\hat{f}^{(1)}\hat{h}\right)\hat{f}^{(2)}\cdots\hat{f}^{(n)}|0\rangle\\ &= \left(hf^{(1)}+f^{(1)}h\right)_{ab}c_a^\dagger\hat{f}^{(2)}\cdots\hat{f}^{(n)}|0\rangle\\ &\quad+\hat{f}^{(1)}\left(hf^{(2)}+f^{(2)}h\right)_{ab}c_b^\dagger c_b^\dagger\hat{f}^{(3)}\cdots\hat{f}^{(n)}|0\rangle\\ &\quad+\cdots+\hat{f}^{(1)}\cdots\hat{f}^{(n-1)}\left(hf^{(n)}+f^{(n)}h\right)_{ab}c_a^\dagger c_b^\dagger|0\rangle.\end{aligned} \tag{3.67}$$

Since $(x^{(1)})_{ab}\equiv(hf^{(1)}+f^{(1)}h)_{ab}$ is a rank-2 matrix, one can use decomposition as

$$\hat{x}^{(1)}\hat{f}^{(2)}\cdots\hat{f}^{(n)}=\frac{1}{2\times n!}\left[\left(\hat{x}^{(1)}+\hat{f}^{(2)}+\cdots+\hat{f}^{(n)}\right)^n-\left(-\hat{x}^{(1)}+\hat{f}^{(2)}+\cdots+\hat{f}^{(n)}\right)^n\right]. \tag{3.68}$$

Here I rewrite $x^{(1)}$ as

$$\begin{aligned}x^{(1)} &= |hu_1\rangle\langle u_{\overline{1}}|-|u_1\rangle\langle u_1h|+|u_1\rangle\langle u_1h|-|hu_{\overline{1}}\rangle\langle u_1|\\ &\equiv |v_{01}\rangle\langle v_{02}|-|v_{02}\rangle\langle v_{01}|+|v_{11}\rangle\langle v_{12}|-|v_{12}\rangle\langle v_{11}|.\end{aligned} \tag{3.69}$$

In this way, if one uses v, one can use the simplification,

$$x^{(1)} + f^{(2)} + \cdots + f^{(n)} = \sum_{k\geq 0}(|v_{k1}\rangle \langle v_{k2}| - |v_{k2}\rangle \langle v_{k1}|) \equiv f_h. \tag{3.70}$$

Here f_h is a rank-$(n+1)$ matrix. When using this f_h in place of f, one can describe E_1 only by the Fredholm Pfaffian.

Next, by looking at the derivative of E_1,

$$\langle\Psi|\hat{h}\left(\frac{\partial}{\partial u_{a,2k-1}}|\Psi\rangle\right) + \left(\frac{\partial}{\partial u_{a,2k-1}}\langle\Psi|\right)\hat{h}|\Psi\rangle, \tag{3.71}$$

one can transform Eq. (3.71) to the form of the Fredholm Pfaffian by using Eqs. (3.59) - (3.62).

Finally, let me investigate the second-order density matrix and their derivatives. Since the coefficient of the second-order density matrix can be decomposed as $V_{pqrs} = \sum_\nu v^\nu_{pq} v^\nu_{rs}$, the Hamiltonian becomes

$$\begin{aligned} H &= \sum_{pq} t_{pq} c_p^\dagger c_q + \sum_{pqrs} V_{pqrs} c_p^\dagger c_q^\dagger c_s c_r \\ &= -2\sum_{pq} V_{pqpq} + \sum_{pq}\left(t_{pq} + 4\sum_r V_{prqr}\right) c_p^\dagger c_q + \sum_\nu (v^\nu_{rs} c_s c_r)\left(v^\nu_{pq} c_p^\dagger c_q^\dagger\right) \\ &\equiv \sum_{pq} h_{pq} c_p^\dagger c_q + \sum_\nu \hat{v}^{(\nu)\dagger}\hat{v}^{(\nu)} + \text{ const }. \end{aligned} \tag{3.72}$$

Therefore, to calculate the second-order density matrix, one needs to calculate

$$\sum_\nu \left\langle 0 \left| \hat{f}_2^{(n)\dagger} \cdots \hat{f}_2^{(1)\dagger} \hat{v}^{(\nu)\dagger}\hat{v}^{(\nu)} \hat{f}_1^{(1)} \cdots \hat{f}_1^{(n)} \right| 0 \right\rangle, \tag{3.73}$$

which can be written using the Fredholm Pfaffian. Especially for the Hubbard model, the second-order density matrix E_2 becomes

$$E_2 = U \frac{1}{(n+1)!^2} \frac{1}{\langle\Psi|\Psi\rangle} \sum_{f_1,f_2\in\{F\}} \sum_a \left\langle 0 \left| \left(\hat{f}_2^\dagger + \left(c_{a\sigma}^\dagger c_{a\bar{\sigma}}^\dagger\right)^\dagger\right)^{n+1} \left(\hat{f}_1 + c_{a\sigma}^\dagger c_{a\bar{\sigma}}^\dagger\right)^{n+1} \right| 0 \right\rangle, \tag{3.74}$$

where U is the on-site Coulomb potential. Then all the terms that consist of rank-$(n+1)$ matrices can be described using the Fredholm Pfaffian.

With this low-rank approximation, one can do the APG calculations easier. This is also the case for the polynomial APG calculation. Here I focus only on the determinant-

polynomial wave function. I make the calculation simpler by using

$$|\Psi\rangle = \det \begin{pmatrix} \hat{f}_0[1] & \hat{f}_{-1}[2] & 0 & \cdots & \cdots & 0 & \hat{f}_1[n] \\ \hat{f}_1[1] & \hat{f}_0[2] & \hat{f}_{-1}[3] & 0 & 0 & 0 & 0 \\ 0 & \hat{f}_1[2] & \hat{f}_0[3] & \ddots & 0 & \vdots & \vdots \\ \vdots & 0 & \hat{f}_1[3] & \ddots & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & \ddots & \hat{f}_{-1}[n-1] & 0 \\ 0 & 0 & 0 & 0 & \ddots & \hat{f}_0[n-1] & \hat{f}_{-1}[n] \\ \hat{f}_{-1}[1] & 0 & 0 & 0 & 0 & \hat{f}_1[n-1] & \hat{f}_0[n] \end{pmatrix} |0\rangle \,. \tag{3.75}$$

This determinant-polynomial wave function is expected to describe fluctuation missing in APG,

$$\hat{f}_0[1]\hat{f}_0[2]\cdots\hat{f}_0[n], \tag{3.76}$$

to some extent. Since the determinant can be written as a trace of products of matrices as

$$\begin{aligned} \det \;=\; & \mathrm{tr}\left[\begin{pmatrix} \hat{f}_0[1] & -\hat{f}_1[1]\hat{f}_{-1}[2] \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{f}_0[2] & -\hat{f}_1[2]\hat{f}_{-1}[3] \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} \hat{f}_0[n] & -\hat{f}_1[n]\hat{f}_{-1}[1] \\ 1 & 0 \end{pmatrix} \right] \\ & - \prod_{i=1}^{n} \hat{f}_1[i] - \prod_{i=1}^{n} \hat{f}_{-1}[i]. \end{aligned} \tag{3.77}$$

One can apply the formulation made above for the low-rank AGP.

3.6 Schur decomposition

One can decompose a $2n \times 2n$ real antisymmetric matrix x into the Schur form, $x = USU^{-1}$, where U is a unitary matrix and S is a band-diagonal matrix,

$$S = \begin{pmatrix} 0 & \lambda_1 & 0 & 0 & 0 \\ -\lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & 0 & \lambda_n \\ 0 & 0 & 0 & -\lambda_n & 0 \end{pmatrix}, \tag{3.78}$$

where λ_m is the eigenvalues of x. Using this Schur decomposition, one can get the products as $x^N = US^NU^{-1}$. Therefore, using

$$\mathrm{tr}\left[x^N\right] = \mathrm{tr}\left[S^N\right] = \sum_{m=1}^{n} \lambda_m^N, \tag{3.79}$$

one can calculate the trace easily and thus simplify the calculation of the Fredholm Pfaffian: The most time-consuming operation then becomes single matrix diagonalization although many matrix products are required originally.

Now I extend the formula to APG. To do it, I transform the AGP wave function by applying the Schur decomposition to every geminal matrix after rewriting the AGP as

$$|\Psi\rangle = (1+\lambda_1 a_1^\dagger a_{\bar{1}}^\dagger t)(1+\lambda_2 a_2^\dagger a_{\bar{2}}^\dagger t)\cdots(1+\lambda_N a_N^\dagger a_{\bar{N}}^\dagger t)\,|0\rangle\Big|_{t^N}. \tag{3.80}$$

Then the corresponding APG wave function becomes

$$\begin{aligned}|\Psi\rangle \;=\; & (1+\lambda_1 a[1]_1^\dagger a[1]_{\bar{1}}^\dagger t)(1+\lambda_2 a[1]_2^\dagger a[1]_{\bar{2}}^\dagger t)\cdots \\ & \times(1+\lambda_1 a[2]_1^\dagger a[2]_{\bar{1}}^\dagger t)\cdots \\ & \cdots \\ & \times(1+\lambda_1 a[N]_1^\dagger a[N]_{\bar{1}}^\dagger t)\cdots(1+\lambda_N a[N]_N^\dagger a[N]_{\bar{N}}^\dagger t)\,|0\rangle\Big|_{t^N}.\end{aligned} \tag{3.81}$$

Note that in Eq. (3.80), I use a single type of the creation operator while in Eq. (3.81), I use different creation operators for different pairs. When one introduces an abbreviation, $\lambda_1 a[N]_1^\dagger a[N]_{\bar{1}}^\dagger = \hat{\lambda}[N]_1$, one can expand Eq. (3.81) as

$$\hat{\lambda}[1]_1\hat{\lambda}[1]_2\hat{\lambda}[2]_1\hat{\lambda}[3]_1\cdots|0\rangle + \hat{\lambda}[1]_1\hat{\lambda}[2]_1\hat{\lambda}[4]_1\cdots|0\rangle + \cdots. \tag{3.82}$$

If one assumes that there is only one nonzero eigenvalue ($\lambda[m]_1$ have a finite value and others are zero), for the sake of reducing the computational cost, which I call the rank-1 approximation, the wave function is greatly simplified and is written, for example for $N=3$, as

$$|\Psi\rangle = \lambda[1]_1 a^\dagger[1]_1 a^\dagger[1]_{\bar{1}}\lambda[2]_1 a^\dagger[2]_1 a^\dagger[2]_{\bar{1}}\lambda[3]_1 a^\dagger[3]_1 a^\dagger[3]_{\bar{1}}\,|0\rangle. \tag{3.83}$$

Likewise one can represent the elementary symmetric polynomial wave functions as

$$\begin{aligned}|\Psi\rangle \;=\; & (1+\lambda_1 a[1]_1^\dagger a[1]_{\bar{1}}^\dagger t)(1+\lambda_2 a[1]_2^\dagger a[1]_{\bar{2}}^\dagger t)\cdots \\ & \times(1+\lambda_1 a[2]_1^\dagger a[2]_{\bar{1}}^\dagger t)\cdots \\ & \cdots \\ & \times(1+\lambda_1 a[M]_1^\dagger a[M]_{\bar{1}}^\dagger t)\cdots(1+\lambda_M a[M]_M^\dagger a[M]_{\bar{M}}^\dagger t)\,|0\rangle\Big|_{t^N},\end{aligned} \tag{3.84}$$

where M is the number of types of geminals. Then for $N=3$ and $M=6$, for example, one gets

$$|\Psi\rangle = \lambda[1]_1 a^\dagger[1]_1 a^\dagger[1]_{\bar{1}}\lambda[2]_1 a^\dagger[2]_1 a^\dagger[2]_{\bar{1}}\lambda[3]_1 a^\dagger[3]_1 a^\dagger[3]_{\bar{1}}\,|0\rangle + \cdots. \tag{3.85}$$

In this example, it has ${}_6C_3 = 20$ terms.

3.7 Embedding theory

I introduce an embedding method called the four-body embedding method. The physical meaning of the four-body correlation is the correlation between two-body and two pairs, or two geminals. The method is used to describe the correlation between geminals within an impurity site that is embedded in a medium described by the AGP. The correlation is described by multiplying a correlation factor consisting of "quadruplet" $g_4(r_1, r_2, r_3, r_4)$, or an antisymmetric tensor of degree four. That is, the trial wave function is prepared as the n th-order coefficient of

$$\exp\left[\sum_{ijkl} G_{ijkl} c_i^\dagger c_j^\dagger c_k^\dagger c_l^\dagger t^2\right] \exp\left[\sum_{pq} F_{pq} c_p^\dagger c_q^\dagger t\right]. \tag{3.86}$$

It is noted that the impurity site does not necessarily consist of a single site. One could extend the method by using antisymmetric tensors of higher degree, but I leave the extension as a target of future study. Figure 3.1 shows a schematics of the method. As shown in this figure, one can include the correlation between geminals in the impurity site by introducing the four-body correlation factor.

The present method has similarity with an existing method. In the Jastrow AGP (JAGP) method [39, 40], the Jastrow factor is multiplied to the AGP wave function of the type

$$\exp\left(\sum_{ij} f(r_{ij}) \hat{n}_i \hat{n}_j\right) \exp\left[\sum_{pq} F_{pq} c_p^\dagger c_q^\dagger\right]. \tag{3.87}$$

Note that the Jastrow factor $f(r_{ij})$ has been used for strongly correlated systems, such as superconductivity. To optimize the JAGP wave function, stochastic methods are used. Neuscamman adapted JAGP to molecules [61]. Also, he developed a method by combining the coupled cluster (CC) and JAGP (CJAGP) [62] as

$$|\Psi\rangle = \exp(\hat{T})|\Phi\rangle, \tag{3.88}$$

where $|\Phi\rangle$ is the JAGP wave function and $\hat{T}$ is a CC operator as

$$\hat{T} = \sum_{\sigma,\tau\in\{\uparrow,\downarrow\}} \sum_{i,j,k,l} T_{i_\sigma j_\tau}^{k_\sigma l_\tau} c_{k_\sigma}^\dagger c_{i_\sigma} c_{l_\tau}^\dagger c_{j_\tau}. \tag{3.89}$$

JAGP and CJAGP methods were shown effective in obtaining the total energy of molecules, but are not so in obtaining the atomic force because of the statistical fluctuation inherent to stochastic methods. This is problematic in studying chemistry because the atomic force and the hessians are crucially important to determine the relaxed structure and to perform a molecular dynamics simulation. Only recently, there was an effort to reduce the effect

of fluctuation [63, 64]. The four-body correlation in the AGP method, on the contrary, determines the wave function variationally, so that the atomic force can be determined accurately. It is also worth stressing that it is much easier to analyze the wave function when not using the stochastic method.

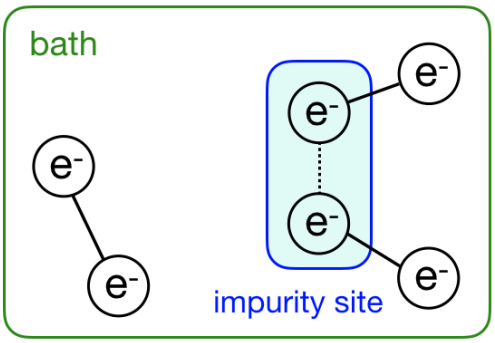

Figure 3.1: A schematics of four-body correlation in AGP embedding. There is an impurity site in the bath. In the bath, I consider the correlation within the electron pair and in the impurity site, I consider the correlation between two electron pairs too.

3.7.1 Four-body correlation

The four-body correlation factor to be applied to the AGP wave function is

$$\exp\left[\sum_{pqrs} G_{pqrs} c_p^\dagger c_q^\dagger c_r^\dagger c_s^\dagger t^2\right], \tag{3.90}$$

where G is the antisymmetric tensor of degree four. To describe the correlation in the region A where the electron repulsion is significant, I restrict the summation and rewrite the wave function as

$$|\Psi\rangle = \exp\left[\sum_{ij}^{\mathrm{A}} \sum_{pq}^{\mathrm{all}} G_{ijpq} c_i^\dagger c_j^\dagger c_p^\dagger c_q^\dagger t^2\right] |tF\rangle \Big|_{t^{\frac{n}{2}}}, \tag{3.91}$$

where the subscripts i, j are limited in the area A. I call this wave function AGP4. If one takes only site $1\uparrow$ and $1\downarrow$ for A, the wave function becomes

$$\begin{aligned} |\Psi\rangle &= \left(1 + \sum_{pq}^{\mathrm{all}} G_{1\uparrow 1\downarrow pq} c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_p^\dagger c_q^\dagger t^2\right) |tF\rangle \Big|_{t^{\frac{n}{2}}} \\ &\equiv \left(1 + \sum_{pq}^{\mathrm{all}} g_{1\uparrow 1\downarrow} G_{pq} c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_p^\dagger c_q^\dagger t^2\right) |tF\rangle \Big|_{t^{\frac{n}{2}}}, \end{aligned} \tag{3.92}$$

where G is an antisymmetric matrix and g is

$$g_{kl} = \begin{cases} 0.5 & (k,l) = (1\uparrow, 1\downarrow) \\ -0.5 & (k,l) = (1\downarrow, 1\uparrow) \\ 0 & (\text{otherwise}) \end{cases} . \tag{3.93}$$

Then the overlap becomes

$$\langle tF| \left(1 + \sum_{pq} g_{1\uparrow 1\downarrow} G_{pq} c^\dagger_{1\uparrow} c^\dagger_{1\downarrow} c^\dagger_p c^\dagger_q t^2 + \sum_{rs} g^\dagger_{1\uparrow 1\downarrow} G^\dagger_{rs} c_s c_r c_{1\downarrow} c_{1\uparrow} t^2 \right.$$
$$\left. + \sum_{pqrs} g^\dagger_{1\uparrow 1\downarrow} G^\dagger_{rs} g_{1\uparrow 1\downarrow} G_{pq} c_s c_r c_{1\downarrow} c_{1\uparrow} c^\dagger_{1\uparrow} c^\dagger_{1\downarrow} c^\dagger_p c^\dagger_q t^4 \right) |tF\rangle . \tag{3.94}$$

As I did for the AGP formulation, I define

$$\hat{g} \equiv \sum_{kl} g_{kl} c^\dagger_k c^\dagger_l , \tag{3.95}$$

and

$$\hat{G} \equiv \sum_{pq} G_{pq} c^\dagger_p c^\dagger_q . \tag{3.96}$$

Then one obtains the four-body wave function as

$$|\Psi\rangle = \hat{g}\hat{G}\hat{F}^N |0\rangle . \tag{3.97}$$

Since it has a structure similar to the APG wave function, one can apply the Fischer formula for the Waring decomposition. Contrary to APG, all the geminals are not different, so that one can simplify the Fischer formula as

$$\hat{g}\hat{G}\hat{F}^N = \sum_{k=0}^{N} \sum_{l=0}^{1} \frac{1}{2^{N+1}(N+1)(N+2)} \frac{(-1)^{k+l}}{(n-k)!k!} \left(\hat{g} + (-1)^l \hat{G} + (N-2k)\hat{F} \right)^{N+2} . \tag{3.98}$$

Since

$$\begin{aligned} \hat{g}^2\hat{G}^2\hat{F}^N &= \sum_{k=0}^{N} \sum_{l=0}^{1} \sum_{m=0}^{2} \frac{1}{2^{N+2}(N+1)(N+2)(N+3)(N+4)} \frac{(-1)^{k+l+m}}{(n-k)!k!(2-m)!m!} \\ &\quad \times \left(\hat{g} + (-1)^l \hat{g} + (2-2m)\hat{G} + (N-2k)\hat{F} \right)^{N+4} , \end{aligned} \tag{3.99}$$

one can generally obtain

$$\begin{aligned} \hat{g}^p\hat{G}^p\hat{F}^N &= \sum_{k=0}^{p-1} \sum_{l=0}^{p} \sum_{m=0}^{N} \frac{(p-1)!p!N!}{2^{2p+N-1}(2p+N)!} \frac{(-1)^{k+l+m}}{(p-1-k)!k!(p-l)!l!(N-m)!m!} \\ &\quad \times \left((p-2k)\hat{g} + (p-2l)\hat{G} + (N-2k)\hat{F} \right)^{2p+N} . \end{aligned} \tag{3.100}$$

In AGP4, if one takes $N = 3$, the wave function becomes $|\Psi\rangle = \hat{g}\hat{G}\hat{F}\,|0\rangle$, which has the same form as that of APG although the matrix element for g is restricted: g_{ab} is nonzero only when one of the indices belongs to the region A. They are different in that g is a limited matrix. Note that if the restriction is removed, or if g is a full antisymmetric matrix, the AGP4 constitutes a complete homogeneous symmetric polynomial wave function. Therefore, the four-body correlation factor is a kind of polynomial APG.

In the above, I have assumed only one site in the strongly correlated area, but one can extend the four-body theory to include more than one orbitals (multi-orbital AGP4). I will explain it using a two-orbital AGP4 as an example. Suppose that two orbitals exist in the correlated site 1 as shown in FIG. 3.2.

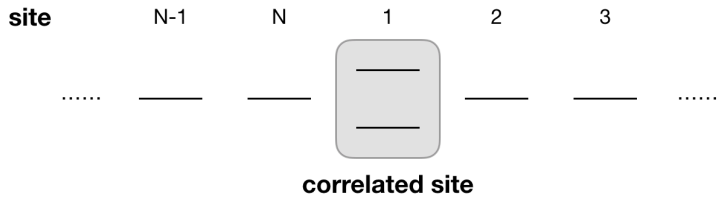

Figure 3.2: 2 orbitals are existed in correlated site 1.

The four-body wave function becomes

$$\begin{aligned} |\Psi\rangle &= \Big(\hat{F}^N + \hat{G}[11]\hat{F}^{N-2} + \hat{G}[22]\hat{F}^{N-2} \\ &\quad + \hat{G}[12]\hat{F}^{N-2} + \hat{G}[11]\hat{G}[22]\hat{F}^{N-4} + \hat{G}[12]\hat{G}[22]\hat{F}^{N-4}\Big)\,|0\rangle\,. \end{aligned} \tag{3.101}$$

where $\hat{G}[11]$ is the four-body correlation factor for the interaction within the orbital-1, $\hat{G}[22]$ is that within the orbital-2 and $\hat{G}[12]$ is that between the orbital-1 and orbital-2 of the site 1. When this wave function is rewritten in the same way as before, it becomes

$$\begin{aligned} |\Psi\rangle &= \hat{F}^N\,|0\rangle + \sum_{ab}\sum_{\{p,q\}} G_{ab} g_{pq} c_a^\dagger c_b^\dagger c_p^\dagger c_q^\dagger \hat{F}^{N-2}\,|0\rangle \\ &\quad + \sum_{abcd}\sum_{\{p,q,r,s\}} G_{ab} G_{cd} g_{pq} g_{rs} c_a^\dagger c_b^\dagger c_p^\dagger c_q^\dagger c_c^\dagger c_d^\dagger c_r^\dagger c_s^\dagger \hat{F}^{N-4}\,|0\rangle\,, \end{aligned} \tag{3.102}$$

where the sum sets $\{p,q\} = \{1,2\},\{3,4\},\{1,3\},\{1,4\},\{2,3\},\{2,4\}$ and $\{p,q,r,s\} = \{1,2,3,4\},\{1,3,2,4\},\{1,4,2,3\}$. Here the subscript 1 denotes a composite of site 1, orbital 1, and up spin. Likewise, the subscript 2 means a composite of site 1, orbital 1, and down spin; the subscript 3 is a composite of site 1, orbital 2, and up spin. Then one can use Eq. (3.99) for decomposition.

3.8 Optimization

In this study, I use variational calculation to optimize the trial wave function and get the ground state energy. The conjugate gradient (CG) method and the Newton method are available as methods of finding the minimum value. However, optimization of a function of high-dimensional variables is often problematic because there are many local minima and saddle-points.

If there are local minima, it is difficult to achieve global minimum using the CG method because it is easily trapped at a local minimum. However, if there exist saddle points only, it is in principle possible to search global minimum even by the modified Newton method but the Newton method is very time-consuming.

In the study of Dauphin *et al.* [65], if there are saddle points, they claimed that the minimizing direction is correct in the CG optimization, but it needs too many steps to achieve global minimum.

In this context, I choose the CG method. The CG method requires only the first-derivatives, which is advantageous in simplifying the program. In my experience, I need about 1000 to 10000 iterations to reduce the magnitude of the gradient. This indicates that there are many saddle points. Also, the corresponding total energies are often different depending on the initial conditions, indicating there are local minima.

To avoid the local minima at least partially within a reasonable computational time, I prepare 100 different random initial conditions and do the CG optimization.

In the rest of this subsection, I will show that the total energy is not a convex function of the geminal matrix. This can be explained by the fact that the simple Newton step does not reduce the total energy. Here I detail this problem.

As the AGP wave function can be written as

$$|Ft\rangle = e^{\sum_{ab} F_{ab} c_a^\dagger c_b^\dagger t} |0\rangle = e^{\hat{F}t} |0\rangle , \tag{3.103}$$

the derivative with respect to F_{ab} is

$$\frac{\partial}{\partial F_{ab}} |Ft\rangle = c_a^\dagger c_b^\dagger t e^{\hat{F}t} |0\rangle . \tag{3.104}$$

When taking a sum after multiplying F_{ab}, I get

$$\sum_{ab} F_{ab} \frac{\partial}{\partial F_{ab}} |Ft\rangle = \hat{F}t |Ft\rangle . \tag{3.105}$$

Also, the second derivatives become

$$\frac{\partial}{\partial F_{cd}} \frac{\partial}{\partial F_{ab}} |Ft\rangle = c_a^\dagger c_b^\dagger c_c^\dagger c_d^\dagger t^2 e^{\hat{F}t} |0\rangle . \tag{3.106}$$

When taking a sum after multiplying F_{ab}, I get

$$\sum_{ab} F_{ab} \frac{\partial^2}{\partial F_{ab} \partial F_{cd}} |Ft\rangle = \hat{F} c_c^\dagger c_d^\dagger t^2 e^{\hat{F}t} |0\rangle . \tag{3.107}$$

When I use the Taylor series of $e^{\hat{F}t}$ as

$$e^{\hat{F}t} |0\rangle \Big|_{t^N} = \left(1 + \hat{F}t + \frac{1}{2}\hat{F}^2 t^2 + \cdots + \frac{1}{n!}\hat{F}^n t^n + \cdots \right) |0\rangle \Big|_{t^N} , \tag{3.108}$$

I can derive the following relationship using Eq. (3.107),

$$\begin{aligned}
\sum_{ab} F_{ab} \frac{\partial^2}{\partial F_{ab} \partial F_{cd}} |Ft\rangle &= \hat{F} c_c^\dagger c_d^\dagger t^2 e^{\hat{F}t} |0\rangle \Big|_{t^N} \\
&= \hat{F} c_c^\dagger c_d^\dagger \frac{1}{(N-1)!} \hat{F}^{N-2} |0\rangle \\
&= (N-1) \frac{1}{(N-1)!} \hat{F}^{N-1} c_c^\dagger c_d^\dagger |0\rangle \\
&= (N-1) c_c^\dagger c_d^\dagger e^{\hat{F}t} \Big|_{t^{N-1}} \\
&= (N-1) c_c^\dagger c_d^\dagger t e^{\hat{F}t} \Big|_{t^N} \\
&= (N-1) \frac{\partial}{\partial F_{cd}} |Ft\rangle \Big|_{t^N} .
\end{aligned} \tag{3.109}$$

Therefore, I get the equation,

$$\sum_{ab} F_{ab} \frac{\partial^2}{\partial F_{ab} \partial F_{cd}} |Ft\rangle \Big|_{t^N} = (N-1) \frac{\partial}{\partial F_{cd}} |Ft\rangle \Big|_{t^N} . \tag{3.110}$$

It means that the second derivatives are written by using the first derivatives. So in geminal wave function theory, the naive Newton method does not work because the second derivatives have no more information than the first derivatives. This problem may be overcome as follows.

Usually, the Newton method updates the next step by using the Hessian as

$$\delta F = -H^\dagger G, \tag{3.111}$$

where H is the Hessian and G is the gradient. One can modify the Hessian by using the absolute value of the Hessian matrix as

$$\delta F \simeq |H|^\dagger G. \tag{3.112}$$

In my preliminary calculation of a six-electron system, I achieved the minimum without being trapped at the local minimum. This suggests that the existence of saddle points is the major reason for the problem.

3.9 Calculation model

In this section, I introduce a few models that are used in this thesis. Also, I describe the total-energy expression.

3.9.1 Hubbard model

The Hamiltonian of the one-dimensional Hubbard model becomes

$$H = t \sum_{<i,j>} c_i^\dagger c_j + U \sum_{(i,j)} c_i^\dagger c_j^\dagger c_j c_i, \tag{3.113}$$

where the element $< i, j >$ is the nearest-neighbor pair of the same spin and (i, j) is the on-site pair.

3.9.2 Anderson model

When I put the on-site Coulomb interaction U to only one site, it becomes the one-dimensional Anderson model and the Hamiltonian is

$$H = t \sum_{<i,j>} c_i^\dagger c_j + U c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_{1\downarrow} c_{1\uparrow}, \tag{3.114}$$

where the subscript $1\uparrow$ denotes the site 1 with up spin and $1\downarrow$ the site 1 with down spin. In section 4.7, I introduce only one impurity site (site 1) and use the periodic boundary condition (FIG. 3.3).

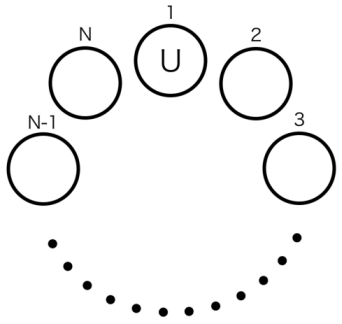

Figure 3.3: One-dimensional Anderson model with the periodic boundary condition. Only on the first site a nonzero value U is assigned as the on-site Coulomb interaction.

When I consider the two-orbital Anderson model like FIG. 3.2, the Hamiltonian becomes

$$H = t \sum_{\{i,j\}} c_i^\dagger c_j + U_1 c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_{1\downarrow} c_{1\uparrow} + U_2 c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger c_{2\downarrow} c_{2\uparrow}, \tag{3.115}$$

where the set $\{i,j\} = \{1,3\},\{1,5\},\{2,4\},\{2,6\},\cdots,\{2N+2,2\},\{2N+2,4\}$, which each number includes site and spin information $(1 = 1\uparrow, 2 = 1\downarrow, \cdots)$.

3.9.3 Total energy

The total energy is

$$E = \frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{3.116}$$

and their derivatives are

$$\begin{aligned} \frac{\partial E}{\partial F} &= \frac{\partial}{\partial F}\left(\frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi|\Psi\rangle}\right) \\ &= \frac{\left(\frac{\partial}{\partial F}\langle\Psi| H |\Psi\rangle\right)\langle\Psi|\Psi\rangle - \left(\frac{\partial}{\partial F}\langle\Psi|\Psi\rangle\right)\langle\Psi| H |\Psi\rangle}{\left(\langle\Psi|\Psi\rangle\right)^2}. \end{aligned} \tag{3.117}$$

The unit of all results I show is $(/t)$ and the calculation is done with the half-filling condition.

Chapter 4

Result

4.1 Property of APG

I will show the property of APG using the one-demensional Hubbard model (Eq. (3.113), $U = 10$), when the number of electrons is taken to be 6 - 12. First, I compare the APG with the Hartree-Fock (HF) and the AGP. I use HΦ [66] to do the exact diagonalization and mVMC [67] to do the HF calculation in this thesis. Figure 4.1 shows the residual error in the total energy referred to the exact value per electron. It shows that the APG provides a significantly lower energy and also that the residual error increases moderately with the number of electrons. Next, I compare the APG with the AGP-CI (FIG. 4.2). For fair comparison, I use the same number of geminals both for the AGP-CI and the APG. The error is smaller for the APG than for the AGP-CI. In spite of the superiority of the APG, the error of the APG is 0.0235 even for six-electron system. Note that, the calculation cost of the AGP is $O(N^5)$ and that of the AGP-CI(K) is $O(N^5K^2)$, while those of the APG and polynomial APG methods are exponentially. Also, TABLE. 4.1 shows the total energy of the exact diagonalization.

It is worth mentioning if the residual error is small enough or not. Typical energy scale often used in this context is the chemical accuracy, which corresponds to the thermal energy at room temperature ($k_BT = 25$ meV). This is important in discussing the chemical equilibrium. The energy scale of 25 meV corresponds to 0.025 in unit of t and thus to 0.0125 when t is taken to be 2 eV. The residual error of the APG, which amounts to 0.14, is about ten times larger. This does not immediately mean that the APG is inaccurate by one order of magnitude however. It was reported in the literature that the error in the dissociation energy is one order of magnitude smaller than the residual error when the APGS was applied to a water molecule and an ethylene molecule [68]. This demonstrates that it is too strict to compare the residual error with the thermal energy. Instead, I consider it reasonable to relax the standard by several times larger and thus require the

residual error be less than about $0.03 - 0.05$ in the present case. With this criterion in mind, I conclude that the error of the APG, 0.14, is not satisfactory.

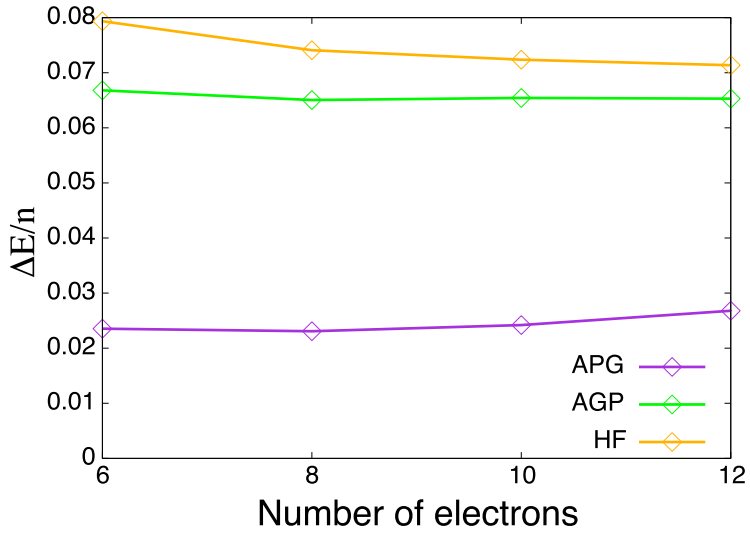

Figure 4.1: (Color online) Residual error of APG, AGP and Hartree-Fock in the total energy per electron $\Delta E/n$ plotted against the number of electrons n.

Number of electrons	6	8	10	12
Total energy	-1.664	-2.177	-2.704	-3.232

Table 4.1: Total energy of the exact diagonalization for the 6 to 12-electron systems.

To see the details of APG, here I show the diagonal elements of the first-order density matrix,

$$\rho^{(1)}_{i\sigma i\sigma} = \frac{\langle\Psi|\, c^\dagger_{i\sigma} c_{i\sigma}\, |\Psi\rangle}{\langle\Psi|\Psi\rangle}, \tag{4.1}$$

of APG for $n = 12$ (FIG. 4.3). The exact value is 0.5 at each site while the values of APG strongly fluctuate around that value. The values of APG are located at around 0.5 when averaged over the spin at the same site, while the amplitudes of the fluctuation are different at different site breaking the translational symmetry of the model. This is the most important sign that APG is not appropriate for strongly correlated systems, which is likely to be overcome using polynomial APGs as will be discussed in section 4.2.

Also, FIG. 4.4 shows the pair correlation function at the same site, $f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$,

$$f_i \equiv \frac{\langle\Psi|\, c^\dagger_{i\uparrow} c_{i\uparrow} c^\dagger_{i\downarrow} c_{i\downarrow}\, |\Psi\rangle}{\langle\Psi|\, c^\dagger_{i\uparrow} c_{i\uparrow}\, |\Psi\rangle \langle\Psi|\, c^\dagger_{i\downarrow} c_{i\downarrow}\, |\Psi\rangle}. \tag{4.2}$$

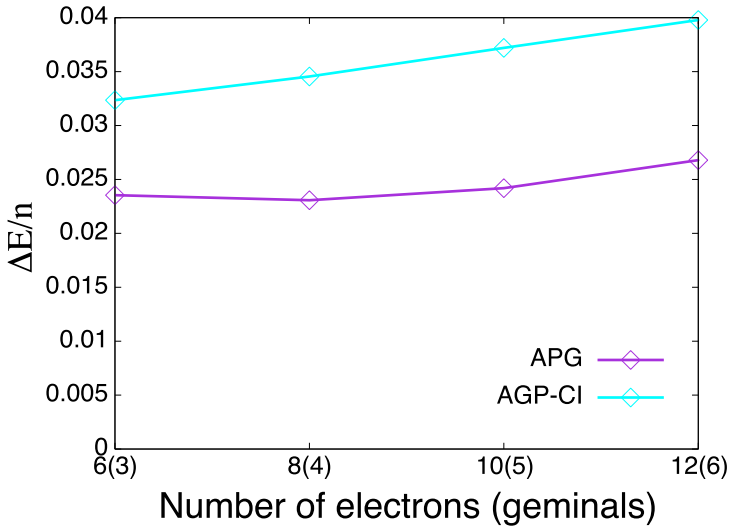

Figure 4.2: (Color online) Residual error of APG and AGP-CI in the total energy per electron $\Delta E/n$ plotted against the number of electrons n.

This pair correlation function (f_i) represents the probability of double occupancy in site i. I found that the values of APG is higher than those of the exact diagonalization, especially at the sites of 3 and 4. This result shows that the electron repulsion is incorrectly large at the sites 3 and 4 although the repulsion should be equally small at all the site as indicated by the exact diagonalization. This is considered as the main reason why APG provides a higher value for the total energy.

Next, I check how accurate the pair correlation function of APG is as the number of electron increases. Figures 4.5, 4.6 and 4.7 are the pair correlation function of the exact diagonalization and APG for the systems with the number of electrons 6, 8 and 10, respectively. As one increases the number of electrons, the error of the pair correlation function of APG increases. APG tends to overestimate the pair correlation function while it sometimes underestimates. Therefore, the APG description deteriorates rather rapidly with increasing n.

4.2 Property of polynomial APG

In the previous section, it is found that the accuracy of APG is still insufficient from the viewpoint of the total energy, the first-order density matrix and pair correlation function. Now, I compare the accuracy of the polynomial APG wave functions. Here, I use the one-dimensional Hubbard model (Eq. (3.113), $U = 10$) and $n = 6$. First I compare the result of e_3 and h_3 using the number of geminal types $M = 3$. Note that, for $n = 6$ and

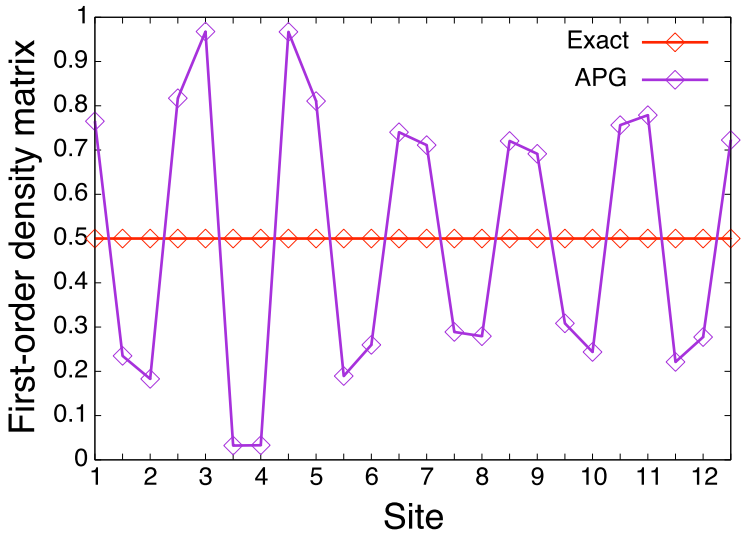

Figure 4.3: (Color online) Diagonal elements of the first-order density matrix of exact diagonalization and APG for the 12-electron system. The horizontal axis indicates the site and spin index; the up spin at the second site, for example, is shown at 2 on the axis while the corresponding down spin is shown at 2.5 on the axis.

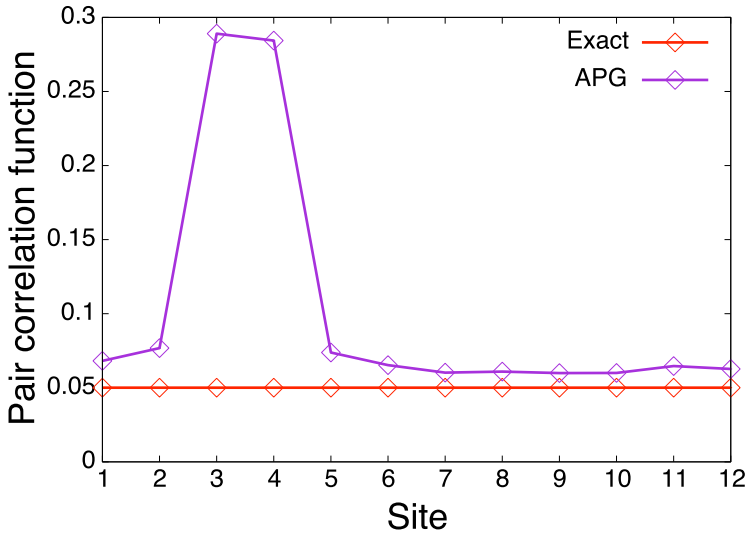

Figure 4.4: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization and APG for the 12-electron system.

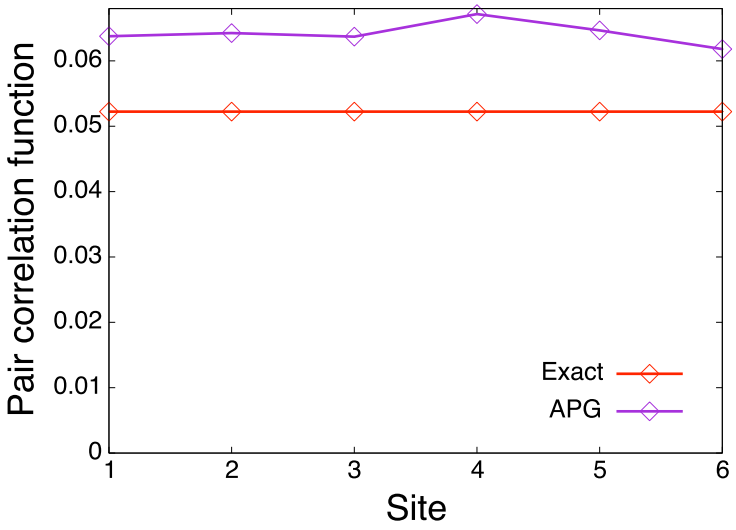

Figure 4.5: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization and APG for the 6-electron system.

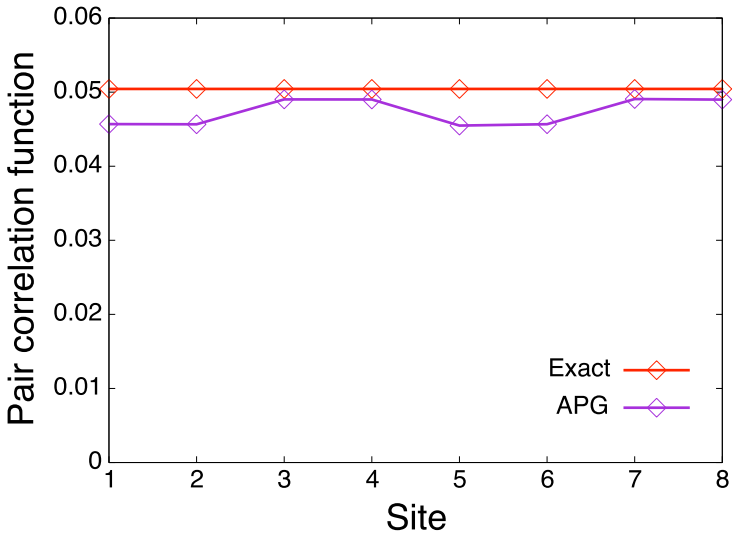

Figure 4.6: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization and APG for the 8-electron system.

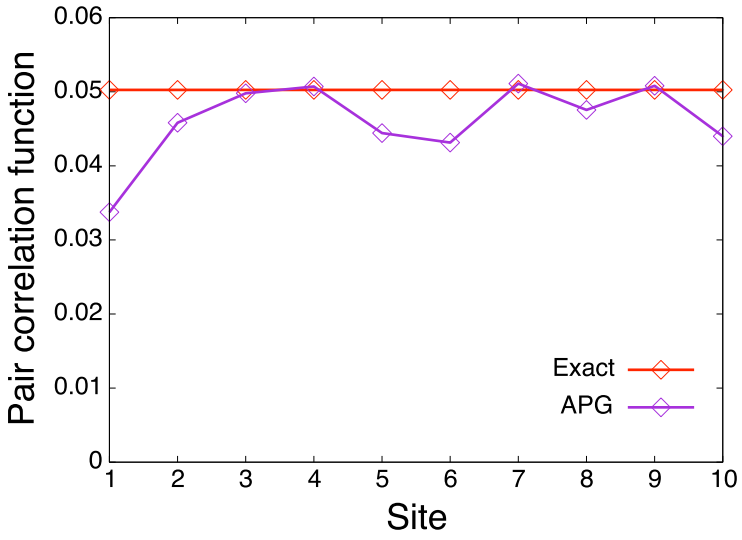

Figure 4.7: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization and APG for the 10-electron system.

$M = 3$, e_3 is identical to APG. And then, I compare results obtained by using the h_3 augmented with coefficient (denoted as h_3 with C) as

$$\sum_{1\le i_1\le i_2\le\cdots\le i_N\le M} C_{\{i_1,i_2,\cdots,i_N\}}\hat{F}[i_1]\hat{F}[i_2]\cdots\hat{F}[i_N]. \tag{4.3}$$

Here I optimize not only F but also the coefficient C. TABLE 4.2 shows the result of the residual error in the correlation energy, ΔEc. Here the number of AGPs means the number of terms appearing after the Waring decomposition. It is found that the error is smaller for e_3 than for h_3. Also, the result of h_3 with C is only slightly better than e_3. It is worth emphasizing that although one might expect that h_3 can describe correlation better because h_3 contains more terms such as the square of a geminal, which are lacking in e_3, the result of h_3 is worse on the contrary. Even the result of optimizing h_3 with coefficients is not much different from e_3. This suggests unfavorable contribution of the square terms and the coefficient C just reduces the amplitude of the multiple product terms. This also suggests that reasonably important terms are included in e_3 for the six-electron system. I will analyze the reason of this in section 4.4. For now, it is noted that the accuracy is not very good when taking $M = 3$.

Then I increase the number of geminal types to $M = 9$ (TABLE 4.3). One can see that the accuracy is improved. In particular, the determinant polynomial shows the result much better than others. It is important that the result depends significantly on the type of polynomial, indicating it is important to use appropriate polynomial.

I also show the average time per iteration in TABLEs 4.2 and 4.3. When comparing

the same number of electrons, the calculation cost is related to the number of AGPs. However, in this calculation, the computing time is not strictly correspond to the number of AGPs. This is because the coding has not been optimized yet.

Polynomial type ($M = 3$)	ΔE_c (%)	Number of AGPs	Time / iteration (s)
e_3	29.67	4	0.132
h_3	36.14	7	0.665
h_3 with C (Eq. (4.3))	29.47	19	3.58

Table 4.2: Residual error in ΔE_c, the number of AGPs generated after the Waring decomposition and the average time one iteration.

Polynomial type ($M = 9$)	ΔE_c (%)	Number of AGPs	Time / iteration (s)
e_3	7.135E-05	10	2.23
h_3	9.145E-05	19	4.28
perm_3	9.895E-05	16	2.27
$\det_3$	2.769E-07	20	2.83

Table 4.3: Residual error in ΔE_c, the number of AGPs generated after the Waring decomposition and the average time one iteration.

I next investigate the effect of the polynomial type of the wave function on the pair correlation function. Figure 4.8 shows the pair correlation function of the exact diagonalization, APG, perm_3 and $\det_3$. The values of perm_3 and $\det_3$ are almost the same as those of the exact diagonalization while APG provides a much larger value. With the polynomial-type wave function, one can thus treat the electron repulsion much better than with APG. To compare the result in more detail, I focus on perm_3 and $\det_3$ in FIG. 4.9. Although there is almost no difference, it is found that the results of $\det_3$ is closer to the exact values at more sites. The $\det_3$ wave function is thus a better method not only in terms of the total energy but also in terms of the pair correlation function.

I also show how the results of e_3 and h_3 change as the number of geminals grows (FIG. 4.10). The error is reduced as one increases the number of geminals for both polynomials but e_3 keeps superiority to h_3. From these results I can conclude that, for the six-electron system, one can achieve enough accuracy when using six types of geminals in both e_3 and h_3.

4.3 U dependency

In this section I investigate the performance of APG and AGP-CI wave functions by changing the on-site Coulomb potential U from 0 to 10 in the Hubbard model (Eq. (3.113)).

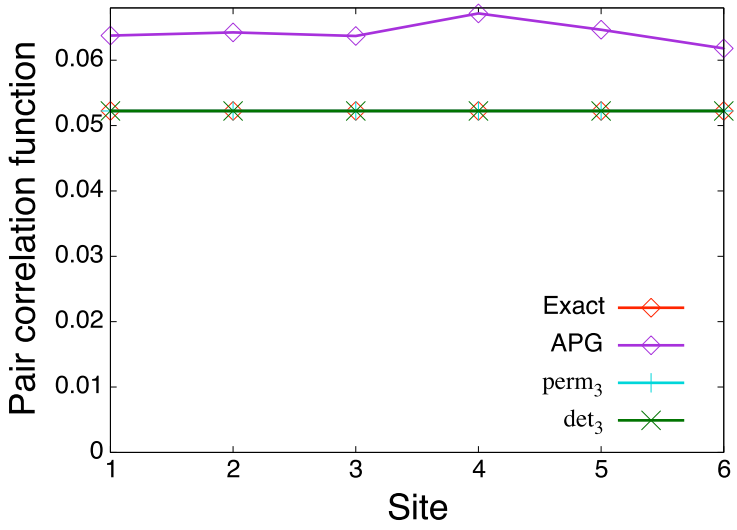

Figure 4.8: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization, APG, perm_3 and det_3 for the 6-electron system.

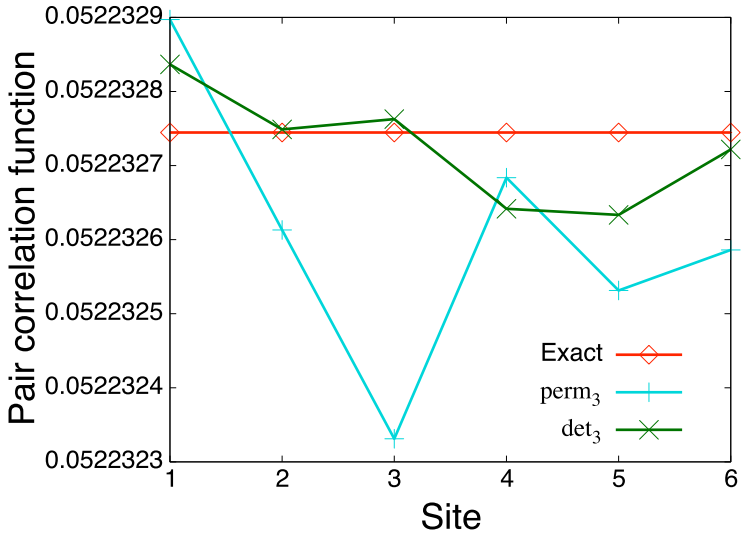

Figure 4.9: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the exact diagonalization, perm_3 and det_3 for the 6-electron system.

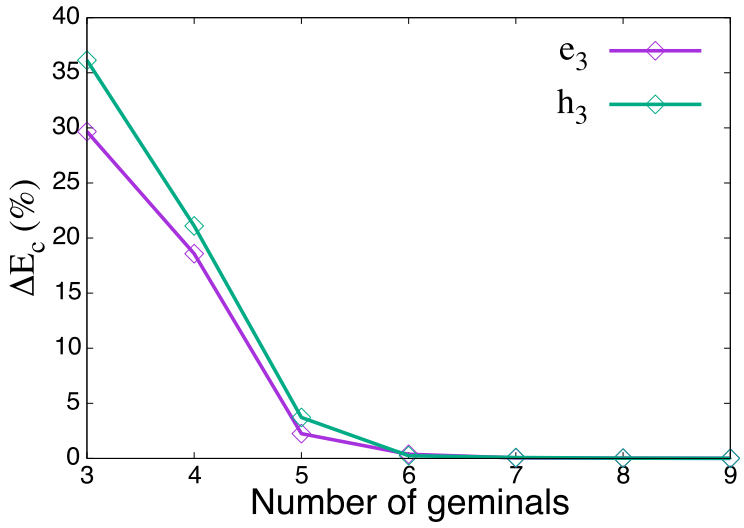

Figure 4.10: (Color online) Residual error in the correlation energy ΔE_c plotted against the number of geminals M.

Figure 4.11 shows that the U dependency of APG and AGP-CI(6) for the 12-electron system. The vertical axis shows the error of E_c (%) and the horizontal axis shows U. As the value of U decreases from 10, the error tends to decrease both for APG and AGP-CI keeping the error of APG smaller than that of AGP-CI. The error is about 18% in APG and 28% in AGP-CI even for $U = 1$ indicating an inferior aspect of the both methods in representing electron correlation. Since the errors show a nonlinear dependence on U, I expect a reasonable description as U is reduced below 1.

As I commented in section 4.1, the required energy error is less than about $0.03 - 0.05$. The total energy error for the 12-electron system and $U = 1$ of the APG is 0.03. Therefore, it is estimated that the required precision is reached when the value of U is less than 1.

4.4 Schur decomposition

I also analyze the properties of the APG wave function and the geminal polynomial wave function. I show the eigenvalues and eigenvectors of the geminal matrices obtained by applying the Schur decomposition. The matrices used in this section are those obtained in section 4.1 $\sim$ 4.3.

Figures 4.12 and 4.13 show the absolute values of the eigenvalues obtained from the APG calculation using 6 and 12 electrons, respectively. It is found that only a few eigenvalues have nonzero values and others are almost zero. Also FIG. 4.14 shows the absolute values of the eigenvalues obtained from h_3 using 6 electrons and 3 geminal types ($M = 3$).

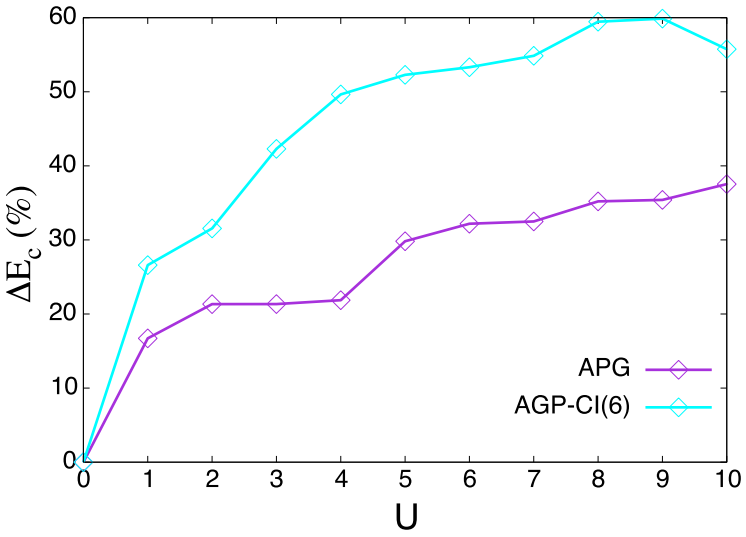

Figure 4.11: (Color online) Residual error in the $\Delta E_c(\%)$ of APG and AGP-CI (12-electron system).

Note that, for the 6-electron system, APG is identical to e_3. By comparing FIGs. 4.12 and 4.14, it is found that, for both e_3 and h_3, there are only a few nonzero eigenvalues. The similarity in the structure of the eigenvalues would probably mean that the multiple products such as $\hat{F}^3[1]$ and $\hat{F}^2[1]\hat{F}[2]$, which are contained only in h_3, do not contribute to changing the structure of the eigenvalues; the structure may be rather insensitive to the polynomial type.

From here, I analyze the eigenvector of the largest eigenvalue considering that other states have appreciably smaller eigenvalues and thus are expected to give minor contribution. Figure 4.15 shows the absolute values of eigenvectors for APG; the number of electron is 6. It is found that the eigenvector is localized at the third and fourth sites although all sites are equivalent. Next, I show the eigenvectors of perm_3 (FIG. 4.16) and $\det_3$ (FIG. 4.17). Compared with the eigenvectors of APG (FIG. 4.15), those of $\det_3$ and perm_3 are not localized to specific sites but all sites are covered by more than one eigenstates. I conjecture that the APG wave function is not flexible enough to describe the system well so that the unnaturally localized geminals are obtained by variation. This is not the case for $\det_3$ and perm_3 possibly because the "multi-APG" is appropriate for this purpose.

I further analyze the results of APG by changing the value of U. Figure 4.18 is the eigenvectors of the geminal of APG; the number of electron is 12 and $U = 10$. It is seen that the wave functions are very localized as having been seen for $n = 6$. Then I change U to 1 (FIG. 4.19). Contrary to the case of $U = 10$, the eigenvectors are localized within 3 to

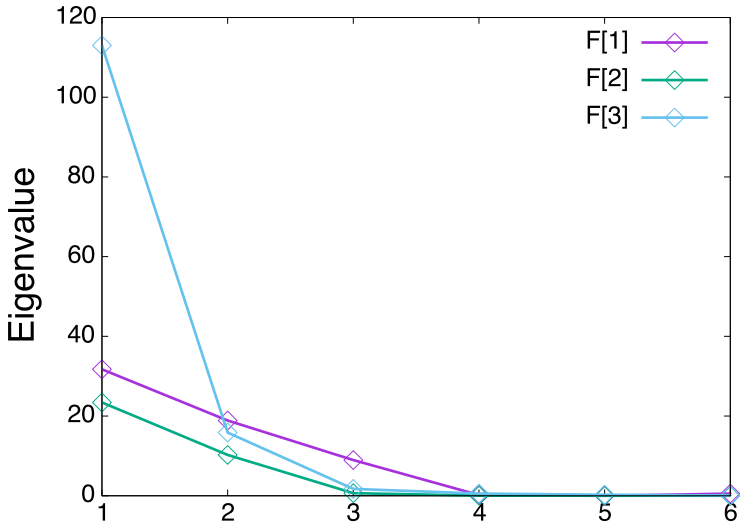

Figure 4.12: (Color online) The absolute eigenvalues of geminal matrices in APG (6-electron system).

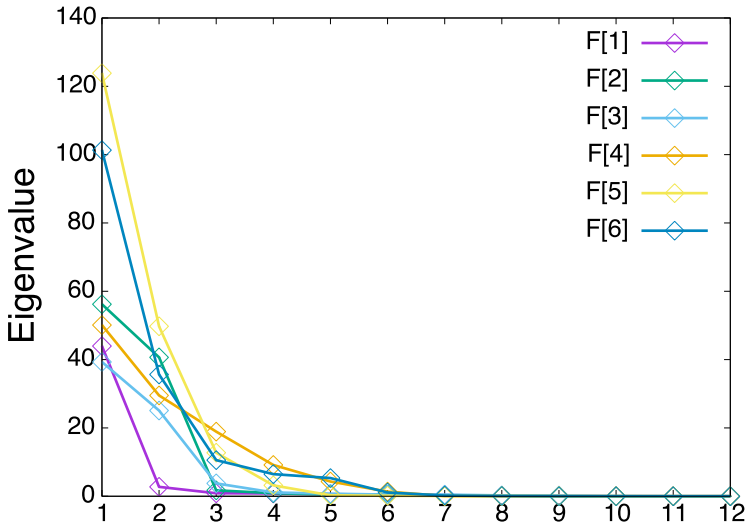

Figure 4.13: (Color online) The absolute eigenvalues of geminal matrices in APG (12-electron system).

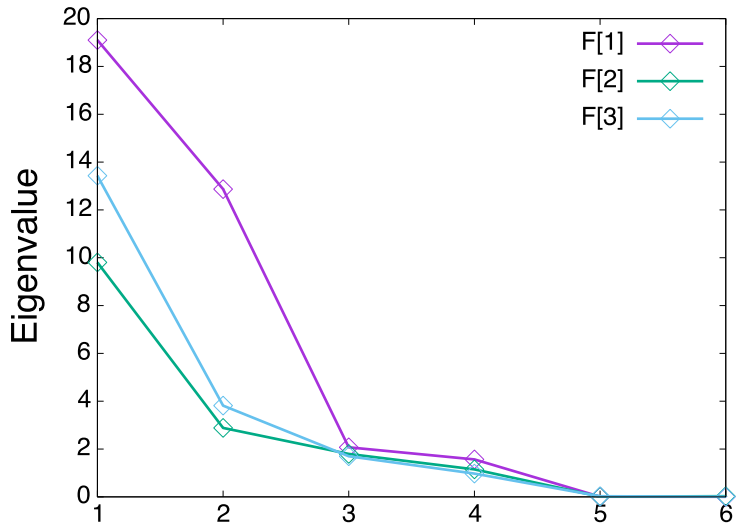

Figure 4.14: (Color online) The absolute eigenvalues of geminal matrices in h_3 (6-electron system, $M = 3$).

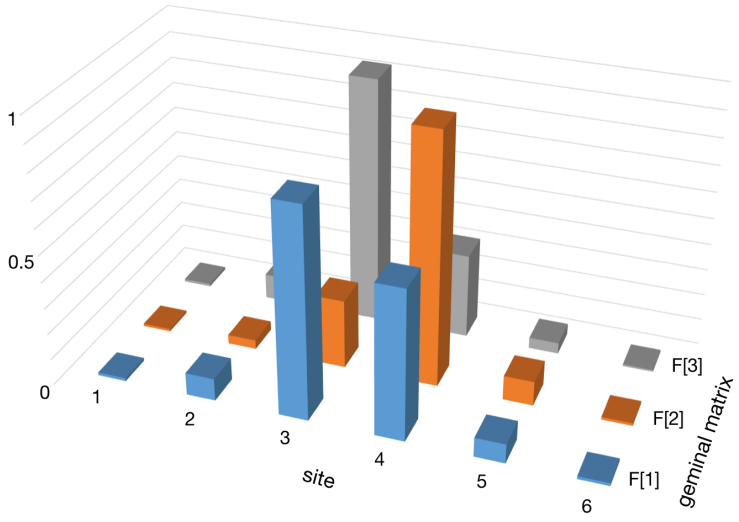

Figure 4.15: (Color online) The absolute eigenvectors of geminal matrices in APG ($n = 6$).

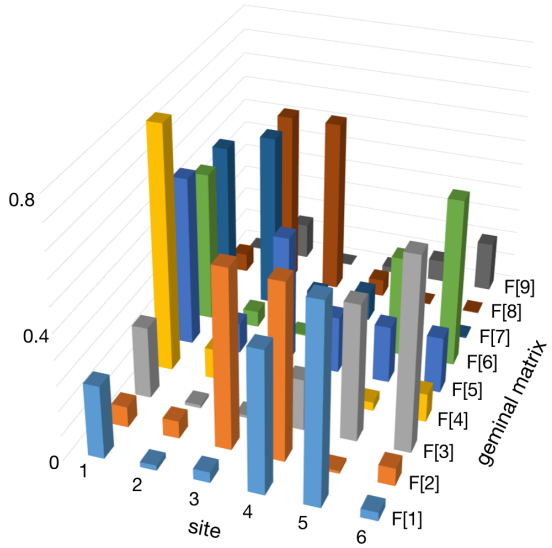

Figure 4.16: (Color online) The absolute eigenvectors of geminal matrices in perm_3 ($n = 6$).

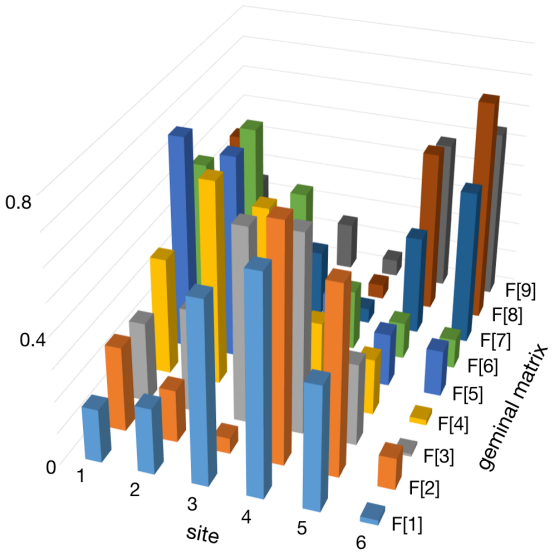

Figure 4.17: (Color online) The absolute eigenvectors of geminal matrices in $\det_3$ ($n = 6$).

4 sites and all sites are covered by at least one eigenstate. This suggests that the unnatural localization of the eigenvectors is typical of strong correlation and is not seen for weakly correlated systems. As a reference, I show the results of HF: Figure 4.20 is for the case of $U = 10$ and FIG. 4.21 is for the case of $U = 1$. Here I use the geminal matrix made from the Slater determinant. The relationship between the geminal and Slater determinant is described in Appendix D. From the viewpoint of the distribution of the wave function, HF wave function thus behaves similarly to the APG wave function.

I consider FIG. 4.17 provides an important implication for describing the strongly correlated system in terms of the valence bond. Each geminal, $F[1]$ to $F[9]$, has a center at a different site and is rather localized within a few sites. Looking at the geminals in more detail, the geminals $F[1]$ to $F[3]$ have the center at the sites 3 to 6, the geminals $F[4]$ to $F[6]$ are centered at around the sites 1 to 3, and the geminals $F[7]$ to $F[9]$ are centered around 5, 6, and 1. The wave function is constructed by terms containing one geminal from $\hat{F}[1]$ to $\hat{F}[3]$, one from $\hat{F}[4]$ to $\hat{F}[6]$, and one from $\hat{F}[7]$ to $\hat{F}[9]$. This result may be regarded as a description of the wave function in terms of the resonating valence bond. This will be an extension of the valence bond theory to strong correlated systems.

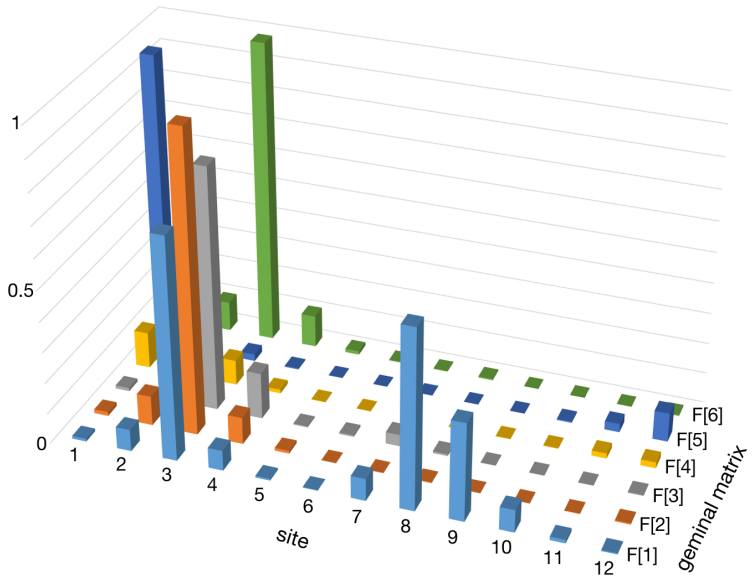

Figure 4.18: (Color online) The absolute eigenvectors of geminal matrices in APG ($n = 12, U = 10$).

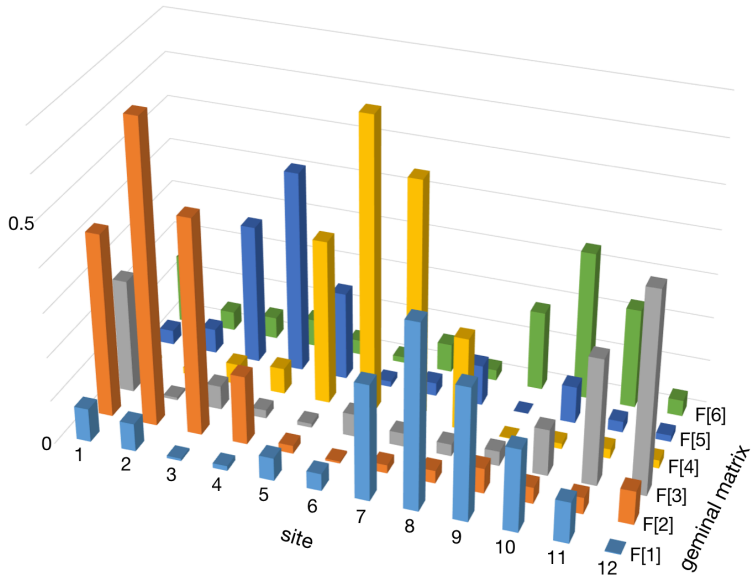

Figure 4.19: (Color online) The absolute eigenvectors of geminal matrices in APG ($n = 12, U = 1$).

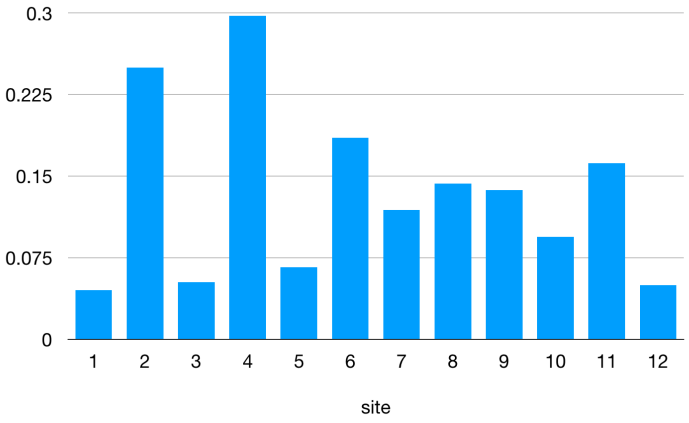

Figure 4.20: (Color online) The absolute eigenvectors of geminal matrix in HF ($n = 12, U = 10$).

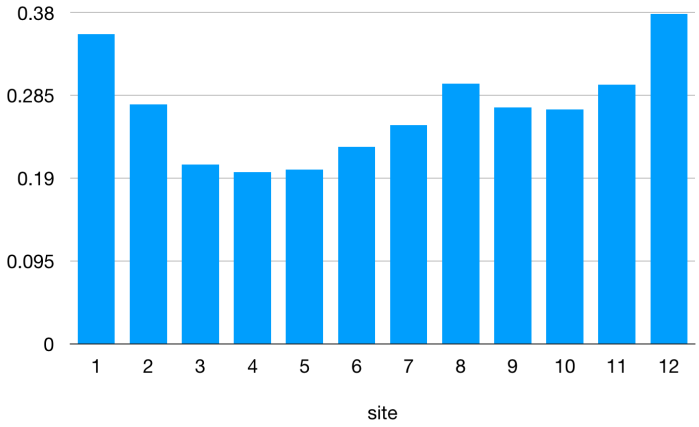

Figure 4.21: (Color online) The absolute eigenvectors of geminal matrix in HF ($n = 12, U = 1$).

4.5 AGP-CI

In the AGP4, APG and other geminal polynomial methods, the wave function is expanded into a linear combination of AGPs, or AGP-CI, by applying the Waring decomposition. In this section, I show how the accuracy improves as the number of types of geminals increases in AGP-CI by using the one-dimensional Hubbard model (Eq. (3.113), $U = 10$).

Figure 4.22 shows the error in the energy per electron of AGP-CI, calculated by comparing the energy of the exact diagonalization; here I use a six-electron system and the number of types of geminals is increased from 1 to 10. AGP-CI(1) equals to AGP. It is found that the error is reduced as the number of types of the geminals is increased. When using 8 geminals, the error is 0.000209, which is significantly smaller than the error of the 6-geminal calculation.

Figure 4.23 shows the results of the energy error, obtained by using a larger number of electrons, 8 to 12. Here the number of types of the geminals is increased from 10 to 30. As the number of the geminal types increases, the error decreases, but the error does not necessarily decrease very much. In the 8-electron system using 30 types of geminals, the error is 6.68×10^{-5}, which is sufficiently small. However, in 12-electron system, even with 30 types of geminals, the error is 0.0233 per electron, which is not so small compared with the required accuracy, less than about $0.03 - 0.05$. In addition, the error does not decrease appreciably in going from 20 types of geminals to 30 types of geminals, suggesting that it may require much more types of geminals to get accurate results in 12-electron system. It is thus conjectured that the number of types of geminals require to achieve enough accuracy increases rapidly as the number of electron increases and that AGP-CI is not so

efficient in those systems.

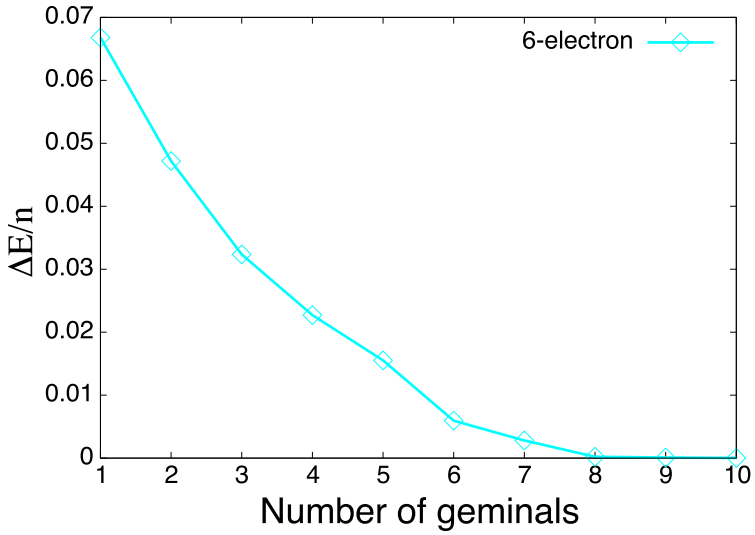

Figure 4.22: (Color online) The total energy error of AGP-CI(K) from the result of the exact diagonalization ($K = 1 \sim 10$) for the 6-electron system.

4.6 Low-rank geminal

To simplify the calculation, I try to use the geminals described by the rank-n matrices as shown in section 3.5. Figure 4.24 shows that the total energy error (per electron) of the original AGP and rank-n AGP, or rank-1 APG, by applying the one-dimensional Hubbard model (Eq. (3.113), $U = 10$). The error is smaller for the original by about 10% although the degrees of freedom is smaller.

Note that the number of variational variables in the original AGP is $M(M-1)/2 = \frac{1}{2}M^2 - \frac{1}{2}M$, where M is the number of state, $M = 2n$ (half-filling), and in the rank-n AGP the number is $\frac{1}{2}M^2$. This result indicates that the advantage of using the rank-n matrices is only in the simplification of the algorithm.

I further examine the low-rank determinant wave function as

$$|\Psi\rangle = \det \begin{pmatrix} \hat{f}_0[1] & \hat{f}_{-1}[2] & 0 & \hat{f}_1[4] \\ \hat{f}_1[1] & \hat{f}_0[2] & \hat{f}_{-1}[3] & 0 \\ 0 & \hat{f}_1[2] & \hat{f}_0[3] & \hat{f}_{-1}[4] \\ \hat{f}_{-1}[1] & 0 & \hat{f}_1[3] & \hat{f}_0[4] \end{pmatrix} |0\rangle \, . \tag{4.4}$$

In TABLE 4.4 I compare the energy errors for the AGP, rank-n AGP (rank-1 APG), AGP-CI(3), rank-1 determinant and APG. I use the Hubbard model (Eq. (3.113), $U = 10$) with eight electrons. It is found that by using the determinant, the error is reduced and

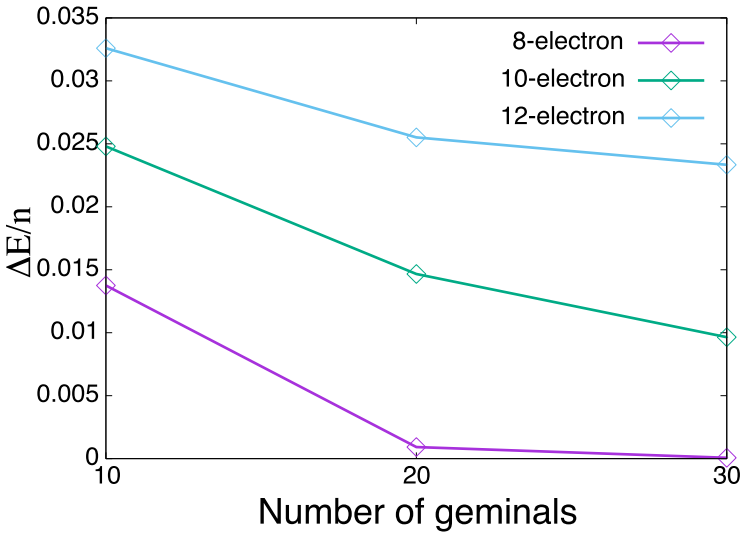

Figure 4.23: (Color online) The total energy error of AGP-CI(K) from the result of the exact diagonalization ($K = 10 \sim 30$) for the 8 to 12-electron system.

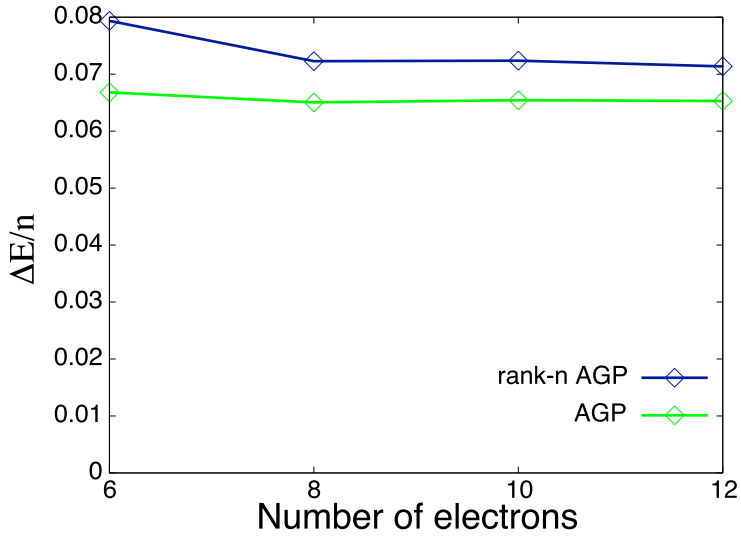

Figure 4.24: (Color online) The total-energy error vs exact diagonalization per electron of the rank-n AGP and the original AGP.

becomes similar to that of the full-rank APG. Note that both AGP-CI(3) and the rank-n determinant consist of three types of geminals. The total energy is lower for the rank-1 determinant in spite of the fact that the rank-1 matrices are used for the geminals. Considering that the algorithm is simplified by using the rank-1 matrices and that small number of geminals is required to get lower total energy when using Eq. (4.4), the low-rank approximation may be a promising way to go, although admittedly one needs further study to obtain a conclusive remark. Note that, in the rank-1 determinant, the number of variational variables is $3 \times 2 \times M \times n = 3M^2$.

8-electron	$\Delta E/n$	Number of geminals
AGP	0.0650	1
rank-n AGP (rank-1 APG)	0.0723	1
AGP-CI(3)	0.0425	3
rank-1 $\det_4$	0.0348	3
APG	0.0231	4

Table 4.4: Energy error per electron of AGP, rank-n AGP (rank-1 APG), AGP-CI(3), rank-1 $\det_4$ and APG for the 8-electron system.

4.7 Four-body correlation

I examine the property of AGP4 by comparing with AGP-CI. In AGP-CI, I use 1 to 3 types of geminals (AGP-CI(1) $\sim$ AGP-CI(3)). The variational variables of AGP4 are G and F in Eq. 3.97, so that the number of variational variables is the same as that of AGP-CI(2). Figure 4.25 shows the error of the total energy using the one-dimensional Anderson model (Eq. (3.114), $U = 10$); the number of electron is $8 \sim 20$. As I increase the number of electrons, the errors of AGP-CI increase but the error of AGP4 does not increase like AGP-CI. The error of AGP-CI(1) is already large for the 8-electron system, also the errors of AGP-CI(2) and AGP-CI(3) rapidly increase from the 16-electron system, but AGP4 remains accurate even for the 20-electron system (the value of error is 0.00048).

Figure 4.26 shows the first-order density matrix of the diagonal elements ($\rho^{(1)}_{i\sigma i\sigma}$) of the 18-electron system. The results of the exact diagonalization and AGP4 are almost the same, but the error of AGP-CI(1) (=AGP) is large especially at the site 1; at this site, which is assigned with strong repulsion U, the spin symmetry of AGP-CI(1) is broken. On the contrary, AGP4 does not break the spin symmetry though I do not impose any spin restriction. In FIG. 4.27 and FIG. 4.28, I also show the pair correlation function of the 18-electron system, with respect to the site 1 accommodating the up spin ($f_{i\sigma} \equiv \rho^{(2)}_{1\uparrow i\sigma}/2\rho^{(1)}_{1\uparrow 1\uparrow}\rho^{(1)}_{i\sigma i\sigma}$), where $\rho^{(1)}$ and $\rho^{(2)}$ are the first- and second-order density ma-

trix, respectively. I find again that the results of the exact diagonalization and AGP4 are almost the same and that $f_{1\downarrow}$ of AGP-CI(1) deviates. It means that AGP-CI(1) cannot describe the electron repulsion on site of $U = 10$ very well allowing a larger probability for the double occupancy. Moreover, the values of $f_{2\downarrow}$ and $f_{18\downarrow}$ in AGP-CI(1) are different, which indicates that the symmetry is broken. Therefore, by using AGP4, I was able to successfully express spin symmetry in this periodic system.

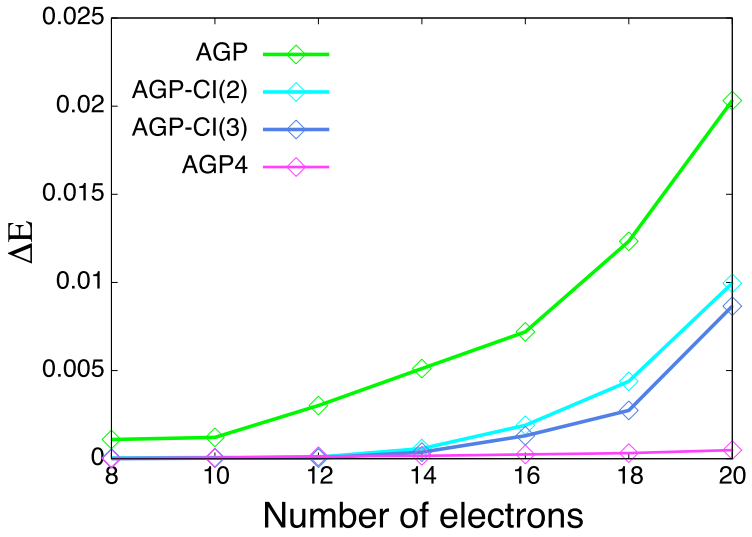

Figure 4.25: (Color online) The residual error in the total energy plotted against the number of electrons. ($U = 10$)

I also show the result for another system having two correlated spots (Eq. (3.115), $U_1 = 10, U_2 = 10$). Figure 4.29 shows the error in the total energy of AGP4 and AGP ($n = 6 \sim 16$). It is found that the errors in AGP4 remain small even when increasing the number of electrons, but the error in AGP is not very small for the 6-electron system and the error rapidly increases when exceeding $n = 14$. Here again I have found superior properties of AGP4.

Note that TABLE. 4.5 shows the total energy of the exact diagonalization.

Number of electrons	6	8	10	12	14	16	18	20
Single impurity		-9.332	-12.37	-14.54	-17.42	-19.69	-22.49	-24.81
Two impurities	-7.130	-8.866	-12.21	-14.23	-17.29	-19.46		

Table 4.5: Total energy of the exact diagonalization for the single and two-orbital impurity systems.

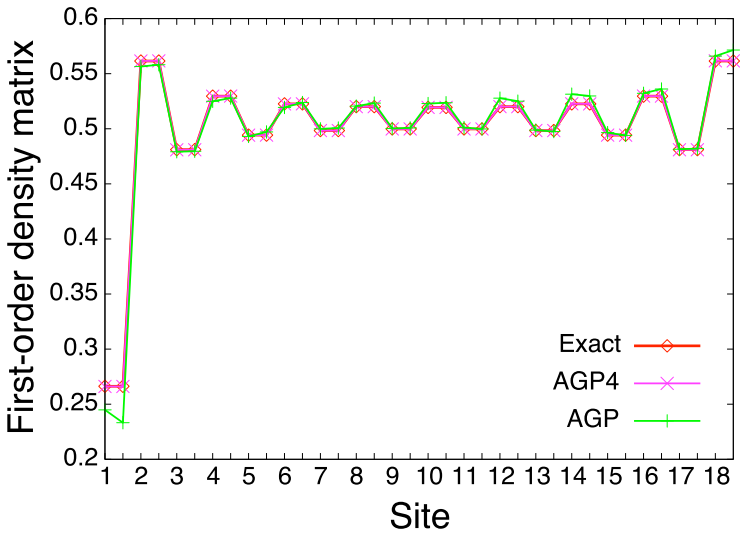

Figure 4.26: (Color online) The first-order density matrix ($\rho^{(1)}_{i\sigma i\sigma}$) of AGP, AGP4 and the exact diagonalization. ($n = 18$)

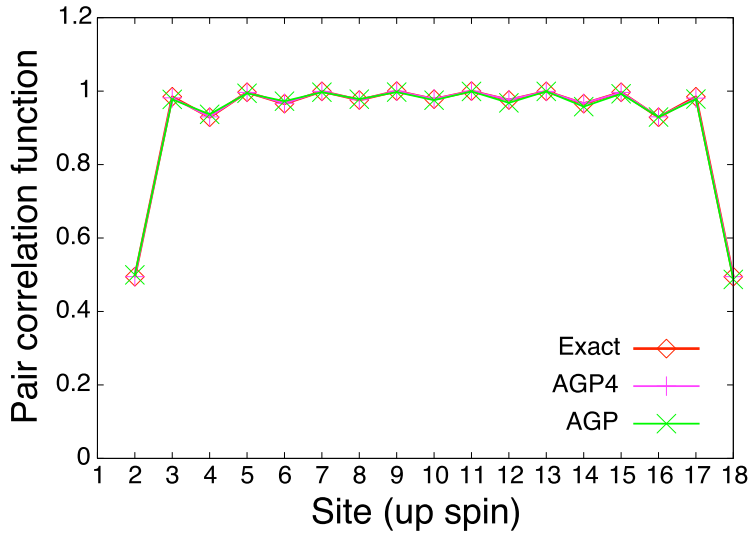

Figure 4.27: (Color online) The pair correlation function with up spin ($f_{i\sigma} \equiv \rho^{(2)}_{1\uparrow i\uparrow}/2\rho^{(1)}_{1\uparrow 1\uparrow}\rho^{(1)}_{i\uparrow i\uparrow}$) of AGP, AGP4 and the exact diagonalization. ($n = 18$)

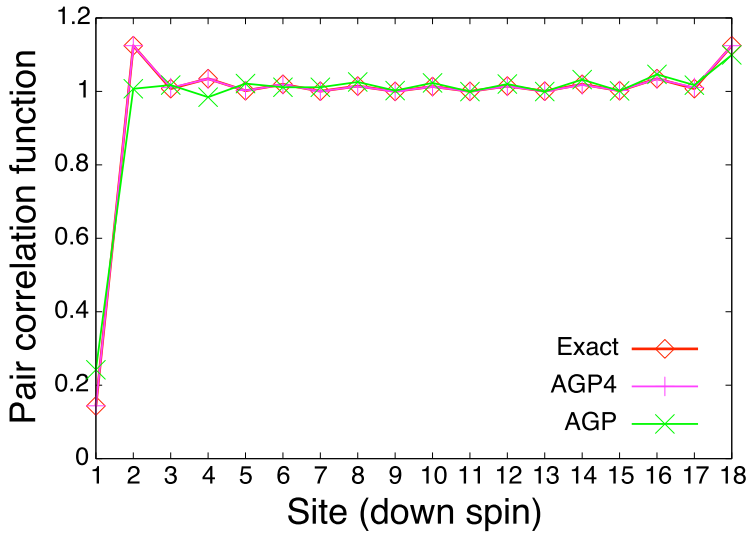

Figure 4.28: (Color online) The pair correlation function with down spin ($f_{i\sigma} \equiv \rho^{(2)}_{1\uparrow i\downarrow}/2\rho^{(1)}_{1\uparrow 1\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of AGP, AGP4 and the exact diagonalization. ($n = 18$)

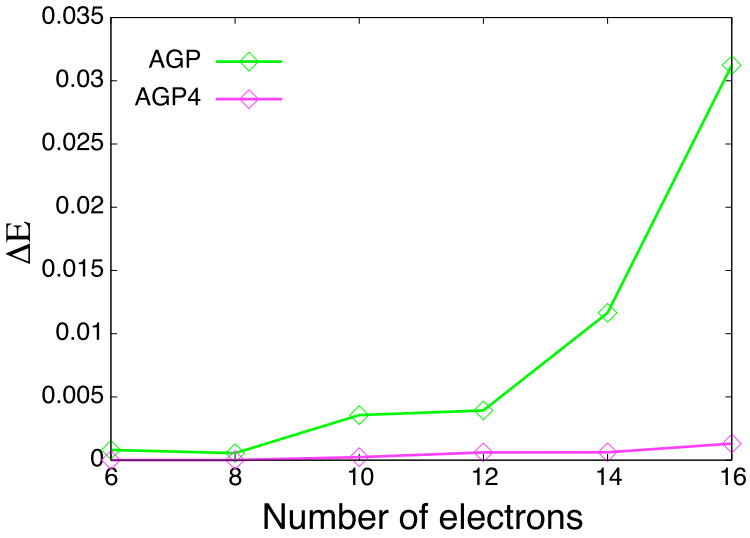

Figure 4.29: (Color online) The residual error in the total energy plotted against the number of electrons. ($U_1 = 10, U_2 = 10$)

4.8 Discussion

For extending the resonating-valence-bond picture to the strongly correlated regime, I found that the determinant polynomial APG is the most suitable method as far as I investigated for the six-electron system that may represent a benzene molecule. Note that the valence bond model commonly assign two resonating Kekuré structures to characterize the electronic structure of benzene. I found that the eigenstates of the geminal localize within a few sites around the center and the central positions are distributed without showing a bias towards specific sites. One can thus recognize the geminals as representing a valence bond. I used nine geminals for the calculation and categorized the resulting geminals into three groups each containing three geminals. Each APG term constituting the polynomial APG consists of three geminals with each geminal belonging to a different group. I can recognize the APG as a component of the resonance and, since the determinant polynomial consists of six terms, I find that the electronic structure as resonance of six components. The system under study requires six components although the conventional valence bond theory requires only two components. The difference may have come from different strength of correlation. This result suggests that the resonating-valence-bond picture may also be valid in larger systems although I could not perform corresponding simulation because of the limited computational time available; the computation is currently quite time consuming.

I have also developed a method of reducing the computational time as well as the complexity of the algorithm. Because of the structure of the eigenvalues of the geminal, I found that the low-rank method as a reasonable approximation. Although the accuracy of the APG calculation was inferior to the original APG, I could improve the accuracy by combining the low-rank method with the determinant-type polynomial showing a promise of further reduction of the computational time. In the APG calculations, I have not explicitly used the fact that the eigenvectors are localized spatially and thus I may reduce the time further by restricting the degrees of freedom. Reduction of the time is one of the most important problems to be solved for further development of the theory.

Towards the goal of embedding the interacting geminals in a medium of non-interacting geminals, I have tested the four-body correlation embedding method. The method works well for the problem I have investigated. There is much room for extension of the method and possible application fields would be the defect physics and molecular science.

Throughout the research, I found that the low convergence of the variational calculation is the most urgent problem. As far as I tested for a small number of systems, the calculation was not trapped at local minima when using the modified Newton method, suggesting that one can expect speed-up of the calculation by adopting an algorithm to avoid the saddle points. Note that considerable attention is paid recently to the development of new algorithms to avoid the saddle points.

What I achieved in this thesis is to formulate a variational method for APG and its extension, which has been thought too demanding because of the complexity of the algorithm. I overcame the difficulty by introducing the Waring decomposition to transform APG to a linear combination of AGPs although the resulting algorithm requires the computational time that scales combinatorially with increasing particle number. I made clear the disadvantage of using APG for the development of valence-bond picture and showed the need for using the polynomial APG although I was able to demonstrate the performance only for six-electron system. I believe that the present work will be an important step to understand the strongly correlated few-body systems.

In this thesis, I used the Hubbard model in which all bonds are uniform, or all t are identical. However, it may be better to use more than one type of t to study atoms and molecules quantitatively with including their details. Therefore, it is important to test such models for the better understanding of the molecular science, and this is one of the future tasks. As an example, I considered the alternating hopping Hubbard model in Appendix E.

Chapter 5

SUMMARY AND CONCLUSION

In this thesis, I developed a geminal wave function theory and investigated the possibility to characterize strongly correlated few-body systems in terms of the concept of the valence bond. This work is motivated by the need for treating a strongly correlated region surrounded by a weakly correlated medium, which is typical of point defects in semiconductors and reaction centers in catalysts. To proceed a step forward, I tried to advance a geminal theory to be able to simulate strongly correlated electrons either embedded in a weakly interacting medium or isolated from the medium. The contribution of this work is to provide a formulation to enable application of the geminal theory to strongly correlated systems. This work also demonstrated the possibility to represent the electronic structure in terms of the resonating valence bond.

In Chapter 1, I started by giving overview of my thesis work followed by more detailed explanations of the related theories for interacting Fermions. First, I explained existing geminal-based theories developed so far with an emphasis put on the valence bond theory. The most relevant theory is the antisymmetrized product of geminals (APG) theory, which is closely related to the antisymmetrized geminal powers (AGP) that has been developed as an extension of the BCS theory. The APG theory has been thought difficult to handle numerically and thus approximated versions have been developed so far, although the difficulty is overcome in this thesis by using the Waring decomposition to transform the APG wave function to the AGP-based wave function, for which the total-energy formula has been developed. I also explained need for extending APG to describe the resonance effect, which is realized in this thesis by using polynomial of geminals instead of monomial of geminals. I also proposed my idea of embedding a region where geminals are interacting in a medium where they are not interacting.

In Chapter 2, I reviewed the electron-pair theories in more detail to make the position of my thesis work clearer. Since the geminal theories have been developed rather independently to each other in different fields, I provided the explanation separately. First, I

explained the AGP theory developed in the chemical-physics community. There, AGP was recognized as an extension of the Hartree-Fock (HF) approximation and was formulated on the basis of the BCS theory. In the physical-chemistry community, the APG theory was developed for obtaining deeper understanding of the valence bond. Because of the difficulty to optimize the geminals, simpler versions of the APG theory were developed by restricting the degrees of freedom for the variational parameters. In condensed-matter community, the JAGP theory was developed as an extension of the HF-based variational Monte Carlo simulation. The theory was applied to superconductivity as well. In the nuclear-physics community, the GCM theory for deformed nuclei was developed using the HF-Bogoliubov (HFB) ground state and a few-particle excitations from HFB. The algorithms developed for GCM very importantly affected the development of geminal theories of different fields.

In Chapter 3, I explained detailed aspect of my formulation. I defined the geminal operator, on that basis, I represent the AGP and APG wave functions as well as their extensions called AGP-CI and polynomial APG. I introduced a method of transforming APG to AGP-CI. By using the Fischer form for the Waring decomposition, I transform the APG wave function to the AGP-CI, and similarly I transform the polynomial APG to AGP-CI. As the type of polynomial, I used the elementary symmetric polynomial, the complete homogeneous symmetric polynomial, the permanent polynomial, and the determinant polynomial.

Then, I presented formulae to determine the total energy variationally. In the formulation, I utilize the Onishi-Yoshida formula to obtain the overlap and the Hamiltonian matrix elements as well as their derivative with respect to the variational parameters. I also proposed a simplified version of the APG calculations where the geminal matrix is represented using low-rank antisymmetric matrix, instead of the general antisymmetric matrix used for the original APG. By this, one needs to evaluate the Fredholm Pfaffian only and thus one can reduce the complexity of the formulation very much. I also provided detailed formulation for the embedding theory, where the interacting region is described using interacting geminals and the environment is described by AGP. Then, I explained the conjugate gradient method used for optimizing the variational parameters as well as the models used for demonstrating the performance of my formulation, that is, one-dimensional (1D) Hubbard model and 1D Anderson model.

In Chapter 4, I presented the numerical results. I found that the total energy of APG is close to the exact one compared with AGP and HF, but the APG breaks the symmetry considerably. The geminals constituting the APG wave function are biased unnaturally to a few sites, suggesting that the APG theory is inappropriate as a theory for the valence bond. The polynomial APG, on the contrary, provides variationally superior result and, among the polynomials, the determinant polynomial was found to show the best result.

By analyzing the obtained geminals using the Schur decomposition, I found that the eigenstates are localized within a few sites from the center contrary to those of APG. The eigenvalues distribute covering all the sites without showing a bias to specific sites. For the determinant polynomial applied to the six-site model, the behavior of the geminal is such that I can recognize it as a representation of the resonating valence bond. Note that the resonating Kekuré structures are commonly assigned to understand the electronic structure of benzene. Similarly in the strongly interacting regime, the present result suggests that the electronic structure may be understood by resonance of nine geminals. Admittedly, I could not provide a clear and systematic image of the resonating valence bond using various models in the strongly correlated regime, but I suggested the possibility by enabling the calculation of polynomial APG.

I also tested the low-rank geminal method and demonstrated the possibility to reduce the computational cost. Finally, I found that the embedding scheme works well for describing the impurity problem. The total energy as well as the wave function are satisfactorily close to the result of exact diagonalization. This is found in the case not only for the model with a single-impurity site but also for the model with two impurity sites, showing a promise to further advance the method for the impurity problem.

In advancing the polynomial APG method, there are several problems to overcome. One is to speed up the variation by using a method robust against the saddle-point effect. Note that this nonconvex optimization problem has attracted considerable attention in many fields such as machine learning and various new algorithms have been proposed. The next problem is then to reduce the scaling of the calculation. Currently, the computational cost scales combinatorially because of the number of the AGP-CI terms generated by the Waring decomposition. If one can make explicit use of the fact that the eigenstates of the geminals are localized in space, one may reduce the cost by reusing the calculation of previous iteration as did in DMRG. One may also advance the mathematics of tensor decomposition to solve the problem. By improving the computational method as stated above, I believe one can achieve deeper understanding of the impurity problem, which has remained as an important problem in several fields of materials science.

Appendix A

Examples of decomposition

In this section, I show specific examples of polynomials and their decomposition. The general formula is shown in section 3.2. Here I use the dimension of polynomial $N = 3$, which means it is used for six-electron system.

A.1 APG

The APG uses monomial form as

$$\hat{F}[1]\hat{F}[2]\hat{F}[3] \tag{A.1}$$

and its decomposition by the Fischer formula is

$$\begin{aligned} & \hat{F}[1]\hat{F}[2]\hat{F}[3] \\ = \quad & \frac{1}{24}\Big[(\hat{F}[1]+\hat{F}[2]+\hat{F}[3])^3-(\hat{F}[1]+\hat{F}[2]-\hat{F}[3])^3 \\ & -(\hat{F}[1]-\hat{F}[2]+\hat{F}[3])^3+(\hat{F}[1]-\hat{F}[2]-\hat{F}[3])^3\Big]. \end{aligned} \tag{A.2}$$

A.2 Elementary symmetric polynomial

Here I consider the case of the number of types of geminals $M = 4$. Then the decomposition of elementary symmetric polynomial becomes

$$\begin{aligned} & \hat{F}[1]\hat{F}[2]\hat{F}[3]+\hat{F}[1]\hat{F}[2]\hat{F}[4]+\hat{F}[1]\hat{F}[3]\hat{F}[4]+\hat{F}[2]\hat{F}[3]\hat{F}[4] \\ = \frac{1}{24} \quad & \Big[2(\hat{F}[1]+\hat{F}[2]+\hat{F}[3]+\hat{F}[4])^3-(-\hat{F}[1]+\hat{F}[2]+\hat{F}[3]+\hat{F}[4])^3 \\ & -(\hat{F}[1]-\hat{F}[2]+\hat{F}[3]+\hat{F}[4])^3-(\hat{F}[1]+\hat{F}[2]-\hat{F}[3]+\hat{F}[4])^3 \\ & -(\hat{F}[1]+\hat{F}[2]+\hat{F}[3]-\hat{F}[4])^3\Big]. \end{aligned} \tag{A.3}$$

A.3 Complete homogeneous symmetric polynomial

As the same with previous section, I also consider the case $M = 4, N = 3$, and the decomposition of complete homogeneous symmetric polynomial becomes

$$\begin{aligned}
&\hat{F}[1]^3 + \hat{F}[1]^2\hat{F}[2] + \hat{F}[1]\hat{F}[2]^2 + \hat{F}[2]^3 + \hat{F}[1]^2\hat{F}[3] + \hat{F}[1]\hat{F}[2]\hat{F}[3] + \hat{F}[2]^2\hat{F}[3] + \hat{F}[1]\hat{F}[3]^2 \\
&+\hat{F}[2]\hat{F}[3]^2 + \hat{F}[3]^3 + \hat{F}[1]^2\hat{F}[4] + \hat{F}[1]\hat{F}[2]\hat{F}[4] + \hat{F}[2]^2\hat{F}[4] + \hat{F}[1]\hat{F}[3]\hat{F}[4] + \hat{F}[2]\hat{F}[3]\hat{F}[4] \\
&+\hat{F}[3]^2\hat{F}[4] + \hat{F}[1]\hat{F}[4]^2 + \hat{F}[2]\hat{F}[4]^2 + \hat{F}[3]\hat{F}[4]^2 + \hat{F}[4]^3 \\
= \;& \frac{1}{6}\Big[3\hat{F}[1]^3 + 3\hat{F}[2]^3 + 3\hat{F}[3]^3 + 3\hat{F}[4]^3 + (\hat{F}[2] + \hat{F}[3] + \hat{F}[4])^3 \\
&+(\hat{F}[1] + \hat{F}[3] + \hat{F}[4])^3 + (\hat{F}[1] + \hat{F}[2] + \hat{F}[4])^3 + (\hat{F}[1] + \hat{F}[2] + \hat{F}[3])^3\Big]. \qquad \text{(A.4)}
\end{aligned}$$

As I explained in section 3.2.2, there is no general formula for the complete homogeneous symmetric polynomial. I cannot find a general N formula, but I made only $N = 3$ and $N = 4$ formula for general M.

In $N = 3$ and $M \geq 4$, the decomposition becomes

$$\frac{1}{6}\left[\sum_i^M \hat{F}[i]^3 + (4 - M)\left(\sum_i^M \hat{F}[i]\right)^3 + \sum_j^M \left(\sum_i^M \hat{F}[i] - \hat{F}[j]\right)^3\right]. \qquad \text{(A.5)}$$

The number of terms after the decomposition is $2M + 1$ when $M \geq 5$ and $2M$ when $M = 4$.

In $N = 4$ and $M \geq 5$,

$$\begin{aligned}
\frac{1}{384} \quad & \Bigg[256\sum_i^M \hat{F}[i]^4 + \left(83 - \frac{1}{2}(M-4)(M-5) + 53(M-5)\right)\left(\sum_i^M \hat{F}[i]\right)^4 \\
&+64\sum_j^M \left(\sum_i^M \hat{F}[i] - \hat{F}[j]\right)^4 - (6 + M)\sum_j^M \left(\sum_i^M \hat{F}[i] - 2\hat{F}[j]\right)^4 \\
&+\sum_{jk}^M \left(\sum_i^M \hat{F}[i] - 2\hat{F}[j] - 2\hat{F}[k]\right)^4\Bigg]. \qquad \text{(A.6)}
\end{aligned}$$

The number of terms after the decomposition is $3M + 1 +_M C_2$.

Also, when I consider adding the coefficient to the complete homogeneous symmetric polynomial ($N = 3, M = 3$), I decompose by using the Fischer decomposition term by

term as

$$
\begin{aligned}
&\sum_{1\le i_1\le i_2\le\cdots\le i_N\le M} C_{\{i_1,i_2,\cdots,i_N\}}\hat{F}[i_1]\hat{F}[i_2]\cdots\hat{F}[i_N] \\
= \;& C_1\hat{F}[1]^3 + C_2\hat{F}[1]^2\hat{F}[2] + C_3\hat{F}[1]\hat{F}[2]^2 + C_4\hat{F}[2]^3 + C_5\hat{F}[1]^2\hat{F}[3] \\
& + C_6\hat{F}[1]\hat{F}[2]\hat{F}[3] + C_7\hat{F}[2]^2\hat{F}[3] + C_8\hat{F}[1]\hat{F}[3]^2 + C_9\hat{F}[2]\hat{F}[3]^2 + C_{10}\hat{F}[3]^3 \\
= \;& C_1\hat{F}[1]^3 + C_2\left[-\frac{1}{12}\hat{F}[2]^3 + \frac{1}{24}(-2\hat{F}[1]+\hat{F}[2])^3 + \frac{1}{24}(2\hat{F}[1]+\hat{F}[2])^3\right] \\
& + C_3\left[-\frac{1}{12}\hat{F}[1]^3 + \frac{1}{24}(-2\hat{F}[2]+\hat{F}[1])^3 + \frac{1}{24}(2\hat{F}[2]+\hat{F}[1])^3\right] + C_4\hat{F}[2]^3 \\
& + C_5\left[-\frac{1}{12}\hat{F}[3]^3 + \frac{1}{24}(-2\hat{F}[1]+\hat{F}[3])^3 + \frac{1}{24}(2\hat{F}[1]+\hat{F}[3])^3\right] \\
& + C_6\left[\frac{1}{24}(\hat{F}[1]+\hat{F}[2]+\hat{F}[3])^3 - \frac{1}{24}(-\hat{F}[1]+\hat{F}[2]+\hat{F}[3])^3\right. \\
& \left. - \frac{1}{24}(\hat{F}[1]-\hat{F}[2]+\hat{F}[3])^3 - 1/24(\hat{F}[1]+\hat{F}[2]-\hat{F}[3])^3\right] \\
& + C_7\left[-\frac{1}{12}\hat{F}[3]^3) + \frac{1}{24}(-2\hat{F}[2]+\hat{F}[3])^3 + \frac{1}{24}(2\hat{F}[2]+\hat{F}[3])^3\right] \\
& + C_8\left[-\frac{1}{12}\hat{F}[1]^3 + \frac{1}{24}(-2\hat{F}[3]+\hat{F}[1])^3 + \frac{1}{24}(2\hat{F}[3]+\hat{F}[1])^3\right] \\
& + C_9\left[-\frac{1}{12}\hat{F}[2]^3 + \frac{1}{24}(-2\hat{F}[3]+\hat{F}[2])^3 + \frac{1}{24}(2\hat{F}[3]+\hat{F}[2])^3\right] + C_{10}\hat{F}[3]^3. \quad \text{(A.7)}
\end{aligned}
$$

A.4 Permanent polynomial

In $N = 3$, the permanent polynomial uses nine types of geminals as

$$
\begin{aligned}
& \mathrm{perm}_3(\hat{F}[1,1],\cdots,\hat{F}[3,3]) \\
= \;& \hat{F}[1,1]\hat{F}[2,2]\hat{F}[3,3] + \hat{F}[1,2]\hat{F}[2,3]\hat{F}[3,1] + \hat{F}[1,3]\hat{F}[2,1]\hat{F}[3,2] \\
+ \;& \hat{F}[1,3]\hat{F}[2,2]\hat{F}[3,1] + \hat{F}[1,2]\hat{F}[2,1]\hat{F}[3,3] + \hat{F}[1,1]\hat{F}[2,3]\hat{F}[3,2]. \quad \text{(A.8)}
\end{aligned}
$$

Then one can decompose as

$$
\begin{aligned}
&\frac{1}{4}\Big[(\hat{F}[1,1]+\hat{F}[1,2]+\hat{F}[1,3])(\hat{F}[2,1]+\hat{F}[2,2]+\hat{F}[2,3])(\hat{F}[3,1]+\hat{F}[3,2]+\hat{F}[3,3]) \\
&-(\hat{F}[1,1]+\hat{F}[1,2]-\hat{F}[1,3])(\hat{F}[2,1]+\hat{F}[2,2]-\hat{F}[2,3])(\hat{F}[3,1]+\hat{F}[3,2]-\hat{F}[3,3]) \\
&-(\hat{F}[1,1]-\hat{F}[1,2]+\hat{F}[1,3])(\hat{F}[2,1]-\hat{F}[2,2]+\hat{F}[2,3])(\hat{F}[3,1]-\hat{F}[3,2]+\hat{F}[3,3]) \\
&+(\hat{F}[1,1]-\hat{F}[1,2]-\hat{F}[1,3])(\hat{F}[2,1]-\hat{F}[2,2]-\hat{F}[2,3])(\hat{F}[3,1]-\hat{F}[3,2]-\hat{F}[3,3])\Big]. \quad \text{(A.9)}
\end{aligned}
$$

After this decomposition, 2^{N-1} terms appeared. One can rename the geminals like APG, e_N and h_N as follows:

$$\begin{aligned}
& \mathrm{perm}_3(\hat{F}[1,1],\cdots,\hat{F}[3,3]) \\
= \; & \hat{F}[1,1]\hat{F}[2,2]\hat{F}[3,3] + \hat{F}[1,2]\hat{F}[2,3]\hat{F}[3,1] + \hat{F}[1,3]\hat{F}[2,1]\hat{F}[3,2] \\
& + \hat{F}[1,3]\hat{F}[2,2]\hat{F}[3,1] + \hat{F}[1,2]\hat{F}[2,1]\hat{F}[3,3] + \hat{F}[1,1]\hat{F}[2,3]\hat{F}[3,2] \\
= \; & \hat{F}[1]\hat{F}[5]\hat{F}[9] + \hat{F}[2]\hat{F}[6]\hat{F}[7] + \hat{F}[3]\hat{F}[4]\hat{F}[8] + \hat{F}[3]\hat{F}[5]\hat{F}[7] + \hat{F}[2]\hat{F}[4]\hat{F}[9] + \hat{F}[1]\hat{F}[6]\hat{F}[8]
\end{aligned} \tag{A.10}$$

A.5 Determinant polynomial

Apart from the Derksen formula, there is another decomposition formula for the determinant polynomial of $N = 3$.

$$\begin{aligned}
\det_3(\hat{F}[1,1],\cdots,\hat{F}[3,3]) \;=\; & (F[1,2]+F[1,3])F[2,1]F[3,2] \\
& -(F[1,1]+F[1,3])F[2,2]F[3,1] \\
& -F[1,2](F[2,1]+F[2,3])(F[3,2]+F[3,3]) \\
& +(F[1,2]-F[1,1])F[2,3](F[3,1]+F[3,2]+F[3,3]) \\
& +F[1,1](F[2,2]+F[2,3])(F[3,1]+F[3,3])
\end{aligned} \tag{A.11}$$

There is a permanent decomposition of the same form as

$$\begin{aligned}
\mathrm{perm}_3(\hat{F}[1,1],\cdots,\hat{F}[3,3]) \;=\; & (F[1,2]+F[1,3])F[2,1]F[3,2] \\
& +(F[1,1]+F[1,3])F[2,2]F[3,1] \\
& +F[1,2](F[2,1]+F[2,3])(F[3,3]-F[3,2]) \\
& +(F[1,1]+F[1,2])F[2,3](F[3,1]+F[3,2]-F[3,3]) \\
& +F[1,1](F[2,2]+F[2,3])(F[3,3]-F[3,1]).
\end{aligned} \tag{A.12}$$

Appendix B

Pfaffian, determinant and permanent

The Pfaffian is identical to the overlap of the AGP wave function as shown in section 3.4. In this section, I show the mathematical property of the Pfaffian.

The determinant can be defined as

$$\det A = \sum_{\sigma} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}, \tag{B.1}$$

where $A = a_{ij}, (1 \leq i, j \leq n)$ is a $n \times n$ square matrix. The determinant has the properties as follows:

$$\det(A) = \det(A^T) \tag{B.2}$$

$$\det(\alpha A) = (\alpha)^n \det(A^T) \tag{B.3}$$

, where α is a scalar.

Taking the sign from the determinant, it becomes permanent as

$$\mathrm{perm}\, A = \sum_{\sigma} \prod_{i=1}^{n} a_{i\sigma(i)}. \tag{B.4}$$

The calculation cost of the determinant is $O(n^3)$, but that of the permanent is exponentially.

In relation to the permanent, the hafnian can be written as

$$\mathrm{haf} A = \sum_{M} \prod_{(i,j)\in M} a_{ij}. \tag{B.5}$$

The relationship between the permanent and the hafnian becomes

$$\mathrm{perm}(A) = \mathrm{haf} \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}. \tag{B.6}$$

Therefore, the calculation cost of the hafnian is larger than that of permanent.

B.1 Pfaffian

The Pfaffian is defined using a $2N \times 2N$ antisymmetric matrix $A = a_{ij}$,

$$\mathrm{Pf}(A) = \sum_{\sigma \in S_{2N}} \mathrm{sgn}(\sigma) a_{\sigma(1)\sigma(2)} a_{\sigma(3)\sigma(4)} \cdots a_{\sigma(2N-1)\sigma(2N)}, \tag{B.7}$$

where $\mathrm{sgn}(\sigma)$ is the sign of N-pair $\{(i_1, j_1), \cdots, (i_N, j_N)\}$ replacement,

$$\begin{pmatrix} 1 & 2 & \cdots & 2N-1 & 2N \\ i_1 & j_1 & \cdots & i_N & j_N \end{pmatrix}. \tag{B.8}$$

For example, when A is a 4×4 matrix, the Pfaffian becomes

$$\mathrm{Pf}(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}. \tag{B.9}$$

The Pfaffian has the following properties:

$$\mathrm{Pf}(A)\mathrm{Pf}(B) = \exp\left(\frac{1}{2}\mathrm{tr}[\ln(A^T B)]\right), \tag{B.10}$$

$$\mathrm{Pf}(BAB^T) = \det(B)\mathrm{Pf}(A), \tag{B.11}$$

where B is $2N \times 2N$ matrix.

The relationship between the Pfaffian and the determinant is

$$\det A = (\mathrm{pf} A)^2. \tag{B.12}$$

B.2 Graph theory

The perfect matching is the concept in graph theory, which relates the Pfaffian very much as I will explain later, so I introduce the graph theory [69] in this section.

A graph is made of vertices and edges. Figures B.1 - B.4 show examples of graphs.

Figure B.1 shows a complete graph. Every vertex is connected to all other vertices by edges. In this example, there are 5 vertices, and here it is called K_5. Figure B.2 shows a bipartite graph. The vertices are divided into two groups and the edges exist only to connect a vertex from one group to a vertex from the other. Figure B.3 shows a planar graph. In the planar graph, it can be drown the edges not to cross any other edges. The plane graph is the graph which is actually drawn not to cross. The left graph of FIG. B.5 is a planar graph and the right one is a plane graph. Figure B.4 shows a tree graph. It does not have cycles.

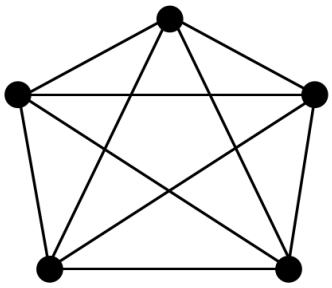

Figure B.1: Complete graph K_5

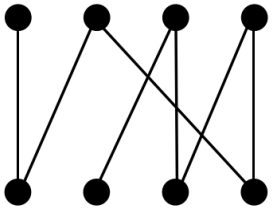

Figure B.2: Bipartite graph

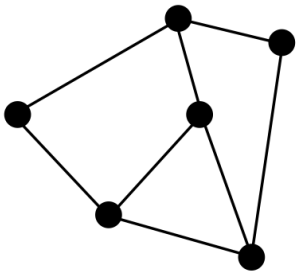

Figure B.3: Planar graph

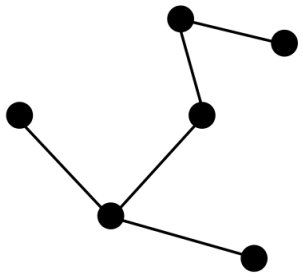

Figure B.4: Tree graph

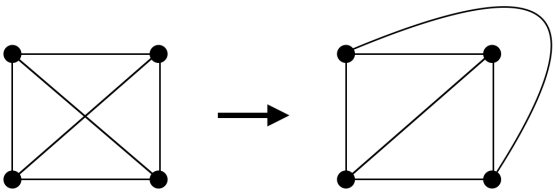

Figure B.5: Planar graph to plane graph.

B.3 Perfect matching

The matching is the set of edges that do not share vertices. The maximum matching means it has the largest number of matching in a graph. When one can get edges with all vertices, it is a perfect matching. Not all graph have the perfect matching and also to determine the number of perfect matching of a graph is a very difficult problem. However, one can get perfect matching for a planar graph by using the Pfaffian.

Figure B.6 shows an example of matching. The blue lines of the right graph show a matching of the left graph. This graph does not have a perfect matching.

Figure B.6: The blue lines in the right graph show the matching of the left graph.

First, I consider the way to express a graph in a matrix form. I prepare the vertex set

$V = \{v_1, v_n\}$ and make matrix A as

$$A_{ij} = \begin{cases} 1 & (\text{if } v_i \text{ and } v_j \text{ are connected by an edge}) \\ 0 & (\text{otherwise}) \end{cases} . \tag{B.13}$$

The matrix A becomes a symmetric matrix. For example, the matrix form of graph in FIG. B.7 is

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} . \tag{B.14}$$

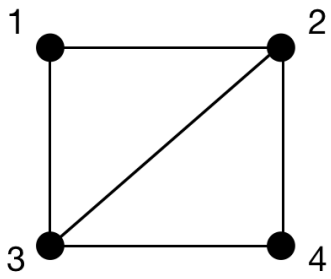

Figure B.7: Graph with no direction.

When one imposes direction of edges, the matrix A becomes an antisymmetric matrix as

$$A_{ij} = \begin{cases} 1 & (\text{if } v_i \text{ and } v_j \text{ are connected by forward direction edge}) \\ -1 & (\text{if } v_i \text{ and } v_j \text{ are connected by opposite direction edge}) \\ 0 & (\text{otherwise}) \end{cases} . \tag{B.15}$$

For example, the matrix form of graph in FIG. B.8 is

$$A = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \end{pmatrix} . \tag{B.16}$$

The important thing is that the Pfaffian corresponds to the perfect matching. This is called the Pfaffian orientation. Here I see the Pfaffian orientation by using a bipartite graph.

The degree of a vertex means the number of edges which the vertex has. The graph in which all degrees of vertices are the same is called the regular graph. There is an important

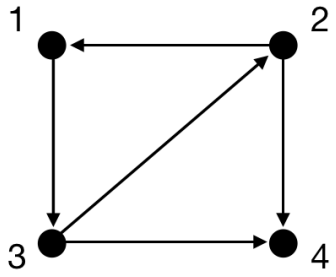

Figure B.8: Graph with direction.

theorem that every regular bipartite graph has a perfect matching. Before proving that theorem, I prove Hall's marriage theorem:

When there is a bipartite graph $G(V_1, V_2)$ in which the vertex set is divided into V_1 and V_2, the necessary and sufficient condition of (1) existing a matching that covers all vertices of V_1 is that (2) for any subset S of V_1, $|S| \leq |N(S)|$. Here $N(S)$ is the set of vertices adjacent to any vertex of S.

The proof of (1) $\rightarrow$ (2) is easy. I prove the contraposition of this. If $|S| > |N(S)|$, there is no such matching as covering S.

Then I prove (2) $\rightarrow$ (1) using the inductive method.

When $|V_1| = 1$, it is obvious. Suppose that Hall's marriage theorem holds when $|V_1| \leq k$. Consider the graph G which $|V_1| = k + 1$ and satisfies condition (2).

First, in the case of $|S| + 1 \leq |N(S)|$. When one makes a bipartite graph $G'(V_1', V_2')$ by removing one edge (v_1, v_2) from G, G' also satisfies the condition (2), and therefore there is a matching M covering V_1' according to the induction hypothesis. If one considers the union of M and (v_1, v_2) as the whole matching, one finds that the condition (1) holds.

Next, consider the case in that there is a subset A which satisfies $|A| = |N(A)|$. The bipartite graph composed of A and $N(A)$ satisfies the condition (2) and according to the induction hypothesis, there is a matching M_1 covering vertex A. Also, M_1 is the perfect matching. Therefore, if one shows that there is a matching M_2 covering V_1' in the remaining graph $G'(V_1', V_2')$, one can prove the theorem. Considering B, which is a subset of V_1', $|A \cup B| + 1 \leq |N(A \cup B)|$ holds. Therefore, there are at least $|B|$ vertices that do not belong to $N(A)$ in the destination of $A \cup B$. Since they are not connected to A, they are all connected to B. So, $|B| \leq |N(B)|$. Therefore. Since G' also satisfies the condition (2), there is a matching M_2 covering V_1' on the assumption of induction. Q.E.D.

Then I prove the theorem that *every regular bipartite graph* $G(V_1, V_2)$ *has a perfect matching.*

As G is a regular bipartite graph, $k|V_1| = k|V_2|$. Here k is the number of edges which one vertex has. So $|V_1| = |V_2|$. Therefore, if there is a matching that covers all vertices of V_1, it becomes the perfect matching. I define E_S to be the union of edges connected to S. Then $E_s \subseteq E_{N(S)}$, or $|E_s| \leq |E_{N(S)}|$. Also, $|E_s| = k|S|$ and $|E_{N(S)}| = k|N(S)|$. Therefore, $k|S| = |E_s| \leq |E_{N(S)}| = k|N(S)|$ and one gets $|S| \leq |N(S)|$. Therefore using Hall's marriage theorem, one can say that there is a matching that covers all vertices of V_1 and find that this matching is a perfect matching. Q.E.D.

The left graph of FIG. B.9 is a regular bipartite graph and it has the perfect matchings like the middle and the right graphs.

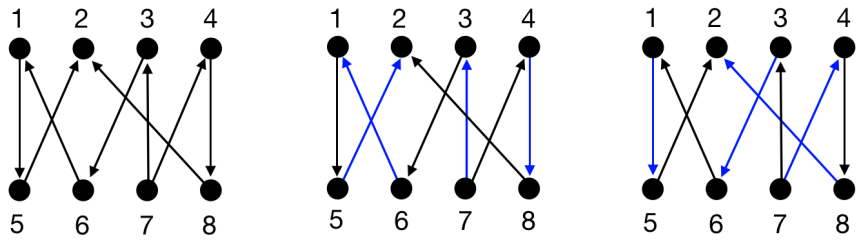

Figure B.9: The examples of perfect matching in the bipartite graph.

Here one can write the matrix A of this graph as

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{B.17}$$

The Pfaffian of A becomes

$$\mathrm{Pf}(A) = \sum_{\sigma \in \{5,6,7,8\}} \mathrm{sgm}(\sigma) a_{1\sigma(5)} a_{2\sigma(6)} a_{3\sigma(7)} a_{4\sigma(8)}. \tag{B.18}$$

The middle figure of FIG. B.9 corresponds to one of Eq. (B.18),

$$a_{16} a_{25} a_{37} a_{48} = -1. \tag{B.19}$$

Also, the right figure of FIG. B.9,

$$a_{15}a_{28}a_{36}a_{47} = 1. \tag{B.20}$$

If the number of intersecting points other than vertex is odd, $\mathrm{sgm}(\sigma)$ becomes -1 and if even, it becomes 1.

Appendix C

Diagonalization

Here I show the details of the diagonalization using the Schur decomposision in section 3.6.

Any square matrix A can be decomposed as

$$A = QUQ^{-1}, \tag{C.1}$$

where Q is a unitary matrix ($Q^{-1} = Q^{\dagger}$) and U is an upper triangular matrix called the Schur form. The eigenvalues of A is given by the diagonal elements of U. When A is a normal matrix ($A^{\dagger}A = AA^{\dagger}$), U becomes a diagonal matrix. Each column vector of Q is an eigenvector of A.

Also, the real square matrix B can be decomposed as

$$B = QUQ^{-1}, \tag{C.2}$$

where Q is a real orthogonal matrix and U is a block upper triangular matrix. The size of the diagonal block of U is less than 2.

Appendix D

Geminal and Slater determinant

In this section, I review one of the ways to express the Slater determinant by using the geminal matrix [67]. Figures 4.20 and 4.21 are derived by using this expression.

The AGP wave function and Slater determinant are defined as

$$|\Psi_{\mathrm{AGP}}\rangle = \left(\sum_{a,b} F_{ab} c_a^\dagger c_b^\dagger\right)^{\frac{n}{2}} |0\rangle\,, \tag{D.1}$$

$$\begin{aligned} |\Psi_{\mathrm{SL}}\rangle &= \prod_i^n \left(\sum_a \Phi_{ai} c_a^\dagger\right) |0\rangle \\ &= \prod_i^n \psi_i^\dagger |0\rangle\,. \end{aligned} \tag{D.2}$$

Here Φ is a normalized orthogonal basis as

$$\sum_a \Phi_{ai}^* \Phi_{aj} = \delta_{ij}. \tag{D.3}$$

Then, the Slater determinant becomes

$$\begin{aligned} |\Psi_{\mathrm{SL}}\rangle &= \prod_i^{n/2} \psi_{2i-1}^\dagger \psi_{2i}^\dagger |0\rangle \\ &= \left(\sum_i^{n/2} \psi_{2i-1}^\dagger \psi_{2i}^\dagger\right)^{\frac{n}{2}} |0\rangle \\ &= \left(\sum_i^{n/2} \left(\sum_a \Phi_{a,2i-1} c_a^\dagger\right)\left(\sum_b \Phi_{b,2i} c_b^\dagger\right)\right)^{\frac{n}{2}} |0\rangle \\ &= \left(\sum_{a,b} \left(\sum_i^{n/2} \Phi_{a,2i-1}\Phi_{b,2i}\right) c_a^\dagger c_b^\dagger\right)^{\frac{n}{2}} |0\rangle\,. \end{aligned} \tag{D.4}$$

Therefore, one can represent F as

$$F_{ab} = \sum_{i}^{n/2} \left(\Phi_{a,2i-1} \Phi_{b,2i} - \Phi_{b,2i-1} \Phi_{a,2i} \right). \tag{D.5}$$

Appendix E

Alternating hopping Hubbard model

In this study, I mainly used the homogenous Hubbard model defined in Eq. (3.113), which all t's are the same. However, to model actual atoms or molecules, it may be sometimes more suitable to use multiple t values. In this appendix, I study the alternating hopping Hubbard model defined by

$$H = t_1 \sum_{<i,j,\text{odd}>} c_i^\dagger c_j + t_2 \sum_{<i,j,\text{even}>} c_i^\dagger c_j + U \sum_{(i,j)} c_i^\dagger c_j^\dagger c_j c_i, \tag{E.1}$$

where the elements $< i, j, \text{odd} >$ and $< i, j, \text{even} >$ are the nearest-neighbor pairs related to the between sites of the odd-numbered and the even-numbered respectively as shown in FIG. E.1. This model has been known to describe the Bechgaard salts well [70].

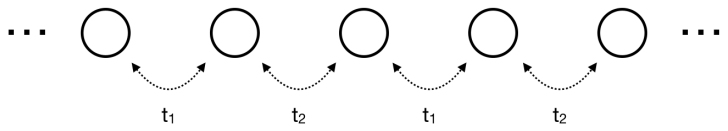

Figure E.1: One-dimensional alternating hopping Hubbard model with the periodic boundary condition.

Here I show results for the alternating hopping Hubbard model with $t_1 = 1$ and $t_2 = 0.5$, at half-filling. TABLE E.1 shows the residual error in the total energy of AGP, APG and $\det_3$ for the six-electron system. Each error is less than that for the homogeneous Hubbard model (Figure 4.1 and TABLE 4.3). It is found that the error of the APG is less than 0.01, which means it has enough accuracy. TABLE E.1 shows the residual error in the total energy of AGP and APG for the 12-electron system. Here also the error is less than that for the homogeneous Hubbard model (Figure 4.1).

6-electron	ΔE	Number of types of geminals
AGP	0.349	1
APG	0.00492	3
$\det_3$	2.04E-07	9

Table E.1: Residual error in the total energy for the six-electron system.

12-electron	ΔE	Number of types of geminals
AGP	0.877	1
APG	0.00869	6

Table E.2: Residual error in the total energy for the 12-electron system.

To see the details, I show the first-order density matrix of the APG for the 12-electron system (FIG. E.2). Comparing with FIG. 4.3, the result improves significantly. The result does not break the symmetry. Also, FIG. E.3 shows the pair correlation function of the APG for the 12-electron system. The result of the APG is also very close to the values of the exact diagonalization.

It is found that in the alternating hopping Hubbard model, the APG is sufficiently effective. It may be because there is no need for a resonance structure in this model. To confirm this property, I need more verifications. This is one of the future tasks.

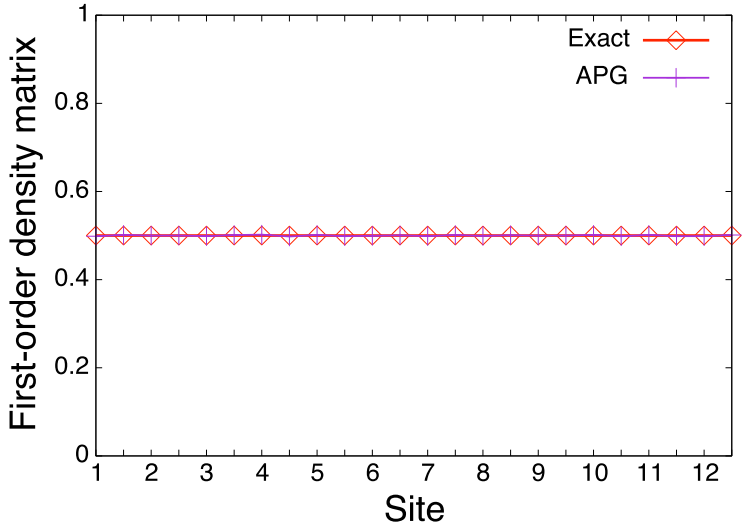

Figure E.2: (Color online) The first-order density matrix ($\rho^{(1)}_{i\sigma i\sigma}$) of the APG and the exact diagonalization for the 12-electron system.

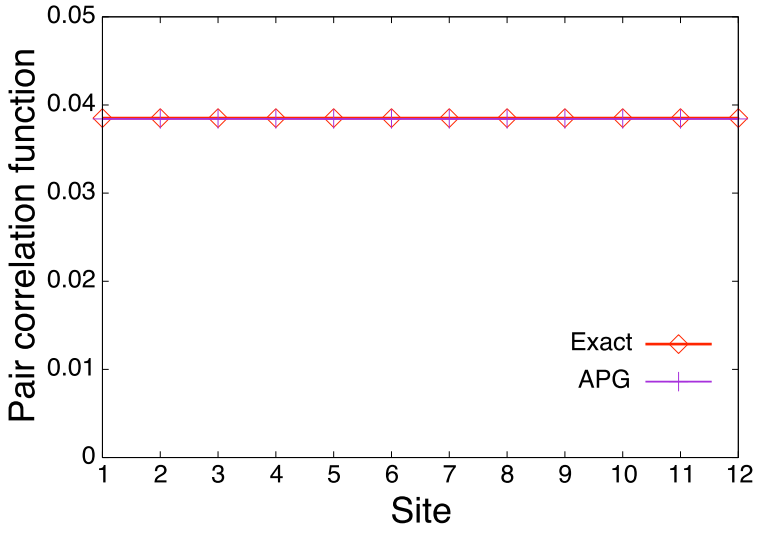

Figure E.3: (Color online) Pair correlation function ($f_i \equiv \rho^{(2)}_{i\uparrow i\downarrow}/2\rho^{(1)}_{i\uparrow i\uparrow}\rho^{(1)}_{i\downarrow i\downarrow}$) of the APG and the exact diagonalization for the 12-electron system.

Appendix F

Water molecule

I applied AGP, AGP-CI and APG to the water molecule (10-electron system). The geometry condition is O $= (0, 0, 0)$, H $= (-1.809, 0, 0), (0.453549, 1.751221, 0)$ and the basis set is STO-3G. The potentials are obtained from HORTON (Helpful Open-source Research TOol for N-fermion systems) program [71]. Figure F.1 shows the error of the total energy. The error is smaller for APG than for AGP and AGP-CI.

H_2O	ΔE	Number of geminals
AGP	0.0284	1
AGP-CI(2)	0.00604	2
AGP-CI(3)	0.00227	3
AGP-CI(4)	0.000563	4
AGP-CI(5)	0.000130	5
APG	5.09E-05	5

Table F.1: Residual error of AGP, AGP-CI and APG for the water molecule.

Bibliography

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] V. I. Anisimov, F. Aryasetiawan, and A. Lichtenstein, Journal of Physics: Condensed Matter **9**, 767 (1997).

[4] P. Eva, A. Lichtenstein, D. Vollhardt, and E. Koch, Theoretische Nanoelektronik (2011).

[5] M. Suzuki, *Quantum Monte Carlo methods in condensed matter physics* (World scientific, 1993).

[6] L. Hedin, Phys. Rev. **139**, 796 (1965).

[7] G. Knizia and G. K.-L. Chan, Phys. Rev. Lett. **109**, 186404 (2012).

[8] G. Knizia and G. K.-L. Chan, J. Chem. Theory Comput. **9**, 1428 (2013).

[9] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).

[10] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

[11] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[12] Ö. Legeza, J. Röder, and B. Hess, Mol. Phys. **101**, 2019 (2003).

[13] G. K.-L. Chan and S. Sharma, Annu. Rev. Phys. Chem. **62**, 465 (2011).

[14] R. J. Gillespie and I. Hargittai, *The VSEPR model of molecular geometry* (Courier Corporation, 2013).

[15] A. J. Coleman, J. Math. Phys. **6**, 1425 (1965).

[16] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[17] W. Kutzelnigg, J. Chem. Phys. **40**, 3640 (1964).

[18] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck, J. Chem. Theory Comput. **9**, 1394 (2013).

[19] A. Hurley, J. E. Lennard-Jones, and J. A. Pople, Proc. R. Soc. Lond. A **220**, 446 (1953).

[20] A. M. Tokmachev, Int. J. Quantum Chem. **116**, 265 (2016).

[21] M. Tarumi, M. Kobayashi, and H. Nakai, Int. J. Quantum Chem. **113**, 239 (2013).

[22] D. M. Silver, J. Chem. Phys. **50**, 5108 (1969).

[23] N. Onishi and S. Yoshida, Nucl. Phys. **80**, 367 (1966).

[24] W. Uemura, S. Kasamatsu, and O. Sugino, Phys. Rev. A **91**, 062504 (2015).

[25] T. G. Kolda and B. W. Bader, SIAM review **51**, 455 (2009).

[26] O. Goscinski, Int. J. Quantum Chem. **22**, 591 (1982).

[27] B. Weiner and O. Goscinski, Phys. Rev. A **27**, 57 (1983).

[28] D. A. Mazziotti, Chem. Phys. Lett. **338**, 323 (2001).

[29] D. Mazziotti, J. Chem. Phys. **112**, 10125 (2000).

[30] K. Hara, S. Iwasaki, and K. Tanabe, Nucl. Phys. A **332**, 69 (1979).

[31] H. Shull, J. Chem. Phys. **30**, 1405 (1959).

[32] R. G. Parr, F. O. Ellison, and P. G. Lykos, J. Chem. Phys. **24**, 1106 (1956).

[33] T. Arai, J. Chem. Phys. **33**, 95 (1960).

[34] R. McWeeny and B. Sutcliffe, Proc. R. Soc. Lond. A **273**, 103 (1963).

[35] V. A. Nicely and J. F. Harrison, J. Chem. Phys. **54**, 4363 (1971).

[36] P. Carrington and G. Doggett, Mol. Phys. **26**, 641 (1973).

[37] P. A. Johnson, P. W. Ayers, P. A. Limacher, S. De Baerdemacker, D. Van Neck, and P. Bultinck, Comput. Theor. Chem. **1003**, 101 (2013).

[38] P. Jeszenszki, P. R. Nagy, T. Zoboki, Á. Szabados, and P. R. Surján, Int. J. Quantum Chem. **114**, 1048 (2014).

[39] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).

[40] D. Tahara and M. Imada, J. Phys. Soc. Jpn. **77**, 114701 (2008).

[41] T. Yanagisawa, M. Miyazaki, and K. Yamaji, Int. J. Mod. Phys. B , 1840023 (2018).

[42] T. Mizusaki and M. Oi, Phys. Lett. B **715**, 219 (2012).

[43] B. Zumino, J. Math. Phys. **3**, 1055 (1962).

[44] C. Bloch, Nucl. Phys. **39**, 95 (1962).

[45] J. Dobaczewski, Phys. Rev. C **62**, 017301 (2000).

[46] T. Tsuchimochi, M. Welborn, and T. Van Voorhis, J. Chem. Phys. **143**, 024107 (2015).

[47] J. Hubbard, Proc. R. Soc. Lond. A **276**, 238 (1963).

[48] R. Pariser and R. G. Parr, J. Chem. Phys. **21**, 767 (1953).

[49] J. A. Pople, Trans. Faraday Soc. **49**, 1375 (1953).

[50] L. R. Tucker, Psychometrika **31**, 279 (1966).

[51] A. Kawasaki and O. Sugino, J. Chem. Phys. **145**, 244110 (2016).

[52] A. Kawasaki and O. Sugino, arXiv preprint arXiv:1805.06138 (2018).

[53] L. Oeding and G. Ottaviani, J. Symb. Comput. **54**, 9 (2013).

[54] P. Comon and B. Mourrain, Signal Processing **53**, 93 (1996).

[55] J. Alexander and A. Hirschowitz, J. Algebr. Geom. **4**, 201 (1995).

[56] I. Fischer, Mathematics Magazine **67**, 59 (1994).

[57] H. Lee, Linear Algebra Its Appl. **492**, 89 (2016).

[58] D. G. Glynn, Eur. J. Combin. **31**, 1887 (2010).

[59] H. Derksen, Foundations of Computational Mathematics **16**, 779 (2016).

[60] W. Squire and G. Trapp, SIAM review **40**, 110 (1998).

[61] E. Neuscamman, J. Chem. Phys. **139**, 194105 (2013).

[62] E. Neuscamman, Communication: A jastrow factor coupled cluster theory for weak and strong electron correlation, 2013.

[63] Y. Luo, A. Zen, and S. Sorella, J. Chem. Phys. **141**, 194112 (2014).

[64] A. Zen, Y. Luo, G. Mazzola, L. Guidoni, and S. Sorella, J. Chem. Phys. **142**, 144111 (2015).

[65] Y. N. Dauphin, R. Pascanu, C. Gulcehre, K. Cho, S. Ganguli, and Y. Bengio, Identifying and attacking the saddle point problem in high-dimensional non-convex optimization, in *Advances in neural information processing systems*, pp. 2933–2941, 2014.

[66] M. Kawamura, K. Yoshimi, T. Misawa, Y. Yamaji, S. Todo, and N. Kawashima, Computer Physics Communications **217**, 180 (2017).

[67] T. Misawa, S. Morita, K. Yoshimi, M. Kawamura, Y. Motoyama, I. Kota, T. Ohgoe, M. Imada, and T. Kato, Comput. Phys. Commun. (2018).

[68] E. Pastorczak and K. Pernal, Physical Chemistry Chemical Physics **17**, 8622 (2015).

[69] J. Nguyen, Bachelor Thesis (2008).

[70] K. Penc and F. Mila, Physical Review B **50**, 11429 (1994).

[71] T. Verstraelen, P. Tecmer, F. Heidar-Zadeh, C. E. González-Espinoza, M. Chan, T. D. Kim, K. Boguslawski, S. Fias, S. Vandenbrande, D. Berrocal, and P. W. Ayers, Horton 2.1.0, http://theochem.github.com/horton/, 2017.