Application of the variational principle to a coherent-pair condensate: The BCS case

L. Y. Jia^{*}

Department of Physics, University of Shanghai for Science and Technology, Shanghai 200093, People's Republic of China

(Received 9 August 2018; published 3 January 2019)

I propose a new algorithm that applies the variational principle to a coherent-pair condensate. The result is equivalent to that of the so-called variation after particle-number projection in the BCS case, but now the particle number is always conserved and the time-consuming projection is avoided. I derive analytical expressions for the average energy and its gradient in terms of the coherent-pair structure. In addition, I give analytically the pair structure at the energy minimum, and discuss its asymptotic behavior away from the Fermi surface, which looks quite simple and allows easy physical interpretations. The new algorithm iterates these pair-structure expressions to minimize energy. I demonstrate the new algorithm in a semirealistic example using the realistic V_{low-k} interaction and large model spaces (up to 15 harmonic-oscillator major shells). The energy can be minimized to practically arbitrary precision. The result shows that the realistic V_{low-k} interaction does not cause divergences in the pairing channel, although tiny occupation numbers (for example smaller than 10^{-5}) contribute to the energy (by a few tens of keV).

DOI: 10.1103/PhysRevC.99.014302

I. INTRODUCTION

Many nuclear structure theories start from a mean-field picture [1]. The choices for the mean field can be either phenomenological or microscopic. The phenomenological type includes the widely used spherical harmonic oscillator, Woods-Saxon, and Nilsson mean field. The microscopic theory is the Hartree-Fock (HF) method.

The pairing correlation has long been recognized [2] and influences practically all nuclei across the nuclear chart [3,4]. To incorporate pairing into the mean field, one introduces quasiparticles and uses, for example, Nilsson+BCS, HF+BCS, or Hartree-Fock-Bogoliubov (HFB) theory. These theories are examples of the variational principle, where the trial wave function is the quasiparticle vacuum. In the BCS case only the pair structure (occupation probability) is varied, whereas in the HFB case the pair structure is varied together with the canonical basis.

However, the BCS or HFB theory has the drawback of breaking the exact particle number [4]. Only the average particle number is guaranteed by the chemical potential. Effectively, one replaces the original microcanonical ensemble by the grand-canonical ensemble (at zero temperature). The two ensembles are equivalent in the thermodynamic limit, but differ in a nucleus having a finite number of nucleons. Especially in phase transition regions of sharp property changes, the differences may be large.

To cue the problems, projection onto good particle number is necessary. It is usually done by numerical integration over the gauge angle [4,5], and the result is a coherent-pair condensate [of generalized seniority zero, see Eq. (1)]. The projection can be done after or before the variation [1,4]. In this work I apply the variational principle directly to the coherent-pair condensate (VDPC). The particle number is always conserved and the time-consuming projection is avoided. (This feature is emphasized by the word "directly" in the name "VDPC".) This work considers VDPC+BCS that varies the coherent-pair structure v_{α} (3), and the result is equivalent to that of VAP+BCS. My next work will extend to VDPC+HFB that varies v_{α} together with the canonical basis, and is equivalent to VAP+HFB. (I name the new method VDPC because VAP may be misleading: there is no projection at all.)

I derive analytical expressions for the average energy and its gradient in terms of v_{α} . Requiring the gradient vanishes, I get the analytical expression of v_{α} at the energy minimum, and discuss its asymptotic behavior away from (above or below) the Fermi surface. The new VDPC algorithm iterates these v_{α} expressions to minimize the energy until practically arbitrary precision. Without integration over gauge angle (necessary in the VAP formalism), the analytical expressions of this work look quite simple and allow easy physical interpretations. I

Projection after variation (PAV) restores the correct particle number and improves binding energy [4,6]. But it violates the variational principle so the energy is not at minimum. Moreover, many realistic nuclei have weak pairing and the BCS or HFB solution collapses [7], then further projection gets nothing. The variation after projection (VAP) [5] is preferred over PAV when feasible. VAP is simply the variational principle using the coherent-pair condensate as the trial wave function, and produces the best energy. For VAP+BCS see, for example, Refs. [5,8–10]; for VAP+HFB see, for example, Refs. [6,11–14]. The practical difficulty of VAP is that numerical projection by integration is time consuming [15] and needed many times in the current VAP procedure. In the literature there are far fewer realistic applications of VAP than those of the HF+BCS or HFB theory without projection.

^{*}liyuan.jia@usst.edu.cn

demonstrate the new algorithm in a semirealistic example using the realistic $V_{\text{low-}k}$ interaction [16] and large model spaces (up to 15 harmonic-oscillator major shells). The energy-convergence pattern and actual computer time cost are given in detail. It is well known that zero-range pairing, frequently used together with the Skyrme density functional, diverges in the pairing channel [17,18]; so the pairing-active space needs a phenomenological cutoff [13,15]. My result shows that the realistic $V_{\text{low-}k}$ interaction converges naturally in the pairing channel, although tiny occupation numbers (for example smaller than 10^{-5}) contribute to the energy (by a few tens of keV).

This work relates to Refs. [8,10,14,19,20]. The average energy of the pair condensate (1) has been derived in terms of v_{α} , for the pairing Hamiltonian [8] and a general Hamiltonian [19]. However, the gradient of energy has not been derived. Using this energy expression (without gradient) to numerically minimize energy is quick in small model spaces [10,14]; but the numerical sign problem arises in large model spaces with many particles. For realistic applications, currently the pairing-channel model space has been limited to a 10 MeV window around the Fermi surface [10,14]. (5 MeV above and 5 MeV below, dimension is approximately that of one major shell in atomic nuclei.) Only for the state-independent pairing Hamiltonian with uniformly spaced single-particle energies (this Hamiltonian has only one parameter of the pairing strength in unit of the single-particle energy spacing), large model spaces have been used [8,10]. In this simple schematic model, v_{α} (and the occupation number) decreases monotonically as the single-particle energy increases, so probably the solution can be regulated to avoid the sign problem. Modern mean-field theories use large model spaces, and VAP has been done only by the gauge-angle integration [6,12,13]. One aim of this work is to propose the new VDPC algorithm.

I also mention the Lipkin-Nogami prescription to restore the particle number approximately [21-23]. It has been widely used because the exact VAP is computationally expensive [1,15]. There are ongoing efforts to improve the Lipkin method [15].

The paper is organized as follows. Section II briefly reviews the formalism for the condensate of coherent pairs. This is the kinematics of the theory. Section III derives the analytical expression for the average energy. Section IV derives the gradient of energy, v_{α} at the energy minimum, and discusses v_{α} 's asymptotic behavior away from the Fermi surface. The VDPC+BCS algorithm is described in Sec. V, and applied to a semirealistic example in Sec. VI. Section VII summarizes the work.

II. COHERENT-PAIR CONDENSATE

In this section I briefly review the formalism for the condensate of coherent pairs (state of zero generalized seniority [24]). For clarity I consider one kind of nucleon, the extension to active protons and neutrons is quite simple: the existence of protons simply provides a correction to the neutron singleparticle energy, through the two-body proton-neutron interaction. I assume time-reversal self-consistent symmetry [4,25], and the single-particle basis state $|\alpha\rangle$ is Kramers degenerate with its time-reversed partner $|\tilde{\alpha}\rangle$ $(|\tilde{\tilde{\alpha}}\rangle = -|\alpha\rangle)$. No other symmetry is assumed.

The ground state of the 2N-particle system is assumed to be an N-pair condensate,

$$|\phi_N\rangle = \frac{1}{\sqrt{\chi_N}} (P^{\dagger})^N |0\rangle, \qquad (1)$$

where

$$\chi_N = \langle 0 | P^N (P^{\dagger})^N | 0 \rangle \tag{2}$$

is the normalization factor. The coherent pair-creation operator is

)

$$P^{\dagger} = \frac{1}{2} \sum_{\alpha} v_{\alpha} a^{\dagger}_{\alpha} a^{\dagger}_{\tilde{\alpha}} = \sum_{\alpha \in \Theta} v_{\alpha} P^{\dagger}_{\alpha}, \qquad (3)$$

where

$$P^{\dagger}_{\alpha} = a^{\dagger}_{\alpha} a^{\dagger}_{\tilde{\alpha}} = P^{\dagger}_{\tilde{\alpha}} \tag{4}$$

creates a pair on $|\alpha\rangle$ and $|\tilde{\alpha}\rangle$. In Eq. (3), Θ is the set of pair indices that picks only one from each degenerate pair $|\alpha\rangle$ and $|\tilde{\alpha}\rangle$. With axial symmetry, orbits of a positive magnetic quantum number are a choice for Θ . \sum_{α} and $\sum_{\alpha \in \Theta}$ mean summing over single-particle indices and pair indices. Requiring $|\phi_N\rangle$ to be time even implies that the pair structure v_{α} (3) is real.

In practice, the canonical single-particle basis could be the self-consistent HF mean field [9,10] or the phenomenological Nilsson level [11,26]. In this work I vary v_{α} on this fixed canonical basis $|\alpha\rangle$ to minimize the average energy $\bar{E} = \langle \phi_N | H | \phi_N \rangle$. Varying the canonical basis $|\alpha\rangle$ will be in my next work of this series.

I review necessary techniques. References [27,28] introduced the many-pair density matrix

$$\begin{aligned} t^{[\gamma_{1}\gamma_{2}...\gamma_{r}],M}_{\alpha_{1}\alpha_{2}...\alpha_{p};\beta_{1}\beta_{2}...\beta_{q}} &\equiv \langle 0|P^{M-p}P_{\gamma_{1}}P_{\gamma_{2}}\dots P_{\gamma_{r}} \\ &\times P_{\alpha_{1}}P_{\alpha_{2}}\dots P_{\alpha_{p}}P_{\beta_{1}}^{\dagger}P_{\beta_{2}}^{\dagger}\dots P_{\beta_{q}}^{\dagger} \\ &\times P_{\gamma_{1}}^{\dagger}P_{\gamma_{2}}^{\dagger}\dots P_{\gamma_{r}}^{\dagger}(P^{\dagger})^{M-q}|0\rangle. \end{aligned}$$
(5)

All the pair indices $\alpha_1, \alpha_2, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q, \gamma_1, \gamma_2, \ldots, \gamma_r$ are distinct: Eq. (5) vanishes if there are duplicated P_{μ}^{\dagger} operators, or duplicated P_{μ}^{\dagger} operators, owing to the Pauli principle; and I require by definition that $\alpha_1, \alpha_2, \ldots, \alpha_p$ and $\beta_1, \beta_2, \ldots, \beta_q$ have no common index (the common ones have been moved to $\gamma_1, \gamma_2, \ldots, \gamma_r$). p, q, r are the number of α, β, γ indices. M equals to the total number of pair-creation operators minus r. Physically, the γ levels are Pauli blocked and inactive. For convenience I introduce $\{[\gamma_1\gamma_2\ldots\gamma_r]\}$ to represent a subspace of the original single-particle space, by removing Kramers pairs of single-particle levels $\gamma_1, \tilde{\gamma}_1, \gamma_2, \tilde{\gamma}_2, \ldots, \gamma_r, \tilde{\gamma}_r$ from the latter. Equation (5) is the pair-hopping amplitude of $P_{\alpha_1}P_{\alpha_2}\ldots P_{\alpha_p}P_{\beta_1}^{\dagger}P_{\beta_2}^{\dagger}\ldots P_{\beta_q}^{\dagger}$ in the blocked subspace $\{[\gamma_1\gamma_2\ldots\gamma_r]\}$.

Reference [28] introduced Pauli-blocked normalizations as a special case of Eq. (5) when p = q = 0,

$$\chi_{M}^{[\gamma_{1}\gamma_{2}...\gamma_{r}]} \equiv t_{;}^{[\gamma_{1}\gamma_{2}...\gamma_{r}],M}$$
$$= \langle 0|P^{M}P_{\gamma_{1}}P_{\gamma_{2}}\dots P_{\gamma_{r}}P_{\gamma_{1}}^{\dagger}P_{\gamma_{2}}^{\dagger}\dots P_{\gamma_{r}}^{\dagger}(P^{\dagger})^{M}|0\rangle.$$
(6)

It is the normalization in the blocked subspace $\{[\gamma_1 \gamma_2 \dots \gamma_r]\}$. Then Ref. [28] expressed many-pair density matrix (5) by normalizations (6),

$$t_{\alpha_{1}\alpha_{2}...\alpha_{p};\beta_{1}\beta_{2}...\beta_{q}}^{[\gamma_{1}\gamma_{2}...\gamma_{p}],M} = \frac{(M-p)!(M-q)!}{[(M-p-q)!]^{2}} \times v_{\alpha_{1}}v_{\alpha_{2}}\ldots v_{\alpha_{p}}v_{\beta_{1}}v_{\beta_{2}}\ldots v_{\beta_{q}} \times \chi_{M-p-q}^{[\alpha_{1}\alpha_{2}...\alpha_{p}\beta_{1}\beta_{2}...\beta_{q}\gamma_{1}\gamma_{2}...\gamma_{r}]}.$$
(7)

Normalizations (6) can be computed by recursive relations,

$$\chi_N = N \sum_{\alpha \in \Theta} (v_\alpha)^2 \chi_{N-1}^{[\alpha]}, \tag{8}$$

$$\chi_N - \chi_N^{[\alpha]} = (N v_\alpha)^2 \chi_{N-1}^{[\alpha]} = \chi_N \langle \phi_N | \hat{n}_\alpha | \phi_N \rangle, \qquad (9)$$

with initial value $\chi_{N=0}^{[\alpha]} = 1$. Knowing $\chi_{N-1}^{[\alpha]}$'s, I compute χ_N by Eq. (8), and then $\chi_N^{[\alpha]}$'s by Eq. (9). $\langle \phi_N | \hat{n}_\alpha | \phi_N \rangle = \langle 0 | P^N a_\alpha^{\dagger} a_\alpha (P^{\dagger})^N | 0 \rangle / \chi_N$ is the average occupation number. Equations (8) and (9) are just Eqs. (22)–(24) of Ref. [29], using $t_{\alpha;}^N = N v_\alpha \chi_{N-1}^{[\alpha]}$ implied from Eq. (7). Equations (8) and (9) are also valid in the blocked subspaces {[$\gamma_1 \gamma_2 \dots \gamma_r$]}. For example, in {[β]} they read

$$\chi_N^{[\beta]} = N \sum_{\alpha \in \Theta}^{P_\alpha \neq P_\beta} (v_\alpha)^2 \chi_{N-1}^{[\alpha\beta]}, \qquad (10)$$

$$\chi_{N}^{[\beta]} - \chi_{N}^{[\alpha\beta]} = (Nv_{\alpha})^{2} \chi_{N-1}^{[\alpha\beta]}$$
$$= \chi_{N}^{[\beta]} \langle \phi_{N}^{[\beta]} | \hat{n}_{\alpha} | \phi_{N}^{[\beta]} \rangle \ (P_{\alpha} \neq P_{\beta}), \quad (11)$$

where

$$\left|\phi_{N}^{\left[\beta\right]}\right\rangle \equiv \frac{1}{\sqrt{\chi_{N}^{\left[\beta\right]}}} \left(P^{\dagger} - v_{\beta} P_{\beta}^{\dagger}\right)^{N} \left|0\right\rangle \tag{12}$$

is the pair condensate with β and $\tilde{\beta}$ blocked. In {[$\beta\gamma$]} ($P_{\beta} \neq P_{\gamma}$) they read

$$\chi_N^{[\beta\gamma]} = N \sum_{\alpha \in \Theta}^{P_\alpha \neq P_\beta, P_\gamma} (v_\alpha)^2 \chi_{N-1}^{[\alpha\beta\gamma]}, \tag{13}$$

$$\chi_N^{[\beta\gamma]} - \chi_N^{[\alpha\beta\gamma]} = (Nv_\alpha)^2 \chi_{N-1}^{[\alpha\beta\gamma]} \ (P_\alpha \neq P_\beta, P_\gamma).$$
(14)

Later the new VDPC algorithm needs $\chi_N^{[\alpha]}$, $\chi_N^{[\alpha\beta]}$, and $\chi_N^{[\alpha\beta\gamma]}$. $\chi_N^{[\alpha]}$ must be computed by the recursive relation (9). $\chi_N^{[\alpha\beta]}$ could be computed by the recursive relation (11), but I find it simpler and quicker to use an alternative formula $(P_{\alpha} \neq P_{\beta})$:

$$(v_{\alpha})^{2} \chi_{N}^{[\alpha]} - (v_{\beta})^{2} \chi_{N}^{[\beta]} = [(v_{\alpha})^{2} - (v_{\beta})^{2}] \chi_{N}^{[\alpha\beta]}.$$
 (15)

Deriving Eq. (15) has only one step using Eq. (11). Similarly, $\chi_N^{[\alpha\beta\gamma]}$ could be computed by the recursive relation (14), but it is simpler and quicker to use $(P_\alpha, P_\beta, P_\gamma)$ are all different)

$$(v_{\alpha})^{2}\chi_{N}^{[\alpha\gamma]} - (v_{\beta})^{2}\chi_{N}^{[\beta\gamma]} = [(v_{\alpha})^{2} - (v_{\beta})^{2}]\chi_{N}^{[\alpha\beta\gamma]}.$$
 (16)

This section discusses the kinematics of the formalism, next I discuss the dynamics.

III. AVERAGE ENERGY

In this section I derive analytically the average energy of the pair condensate. The antisymmetrized two-body Hamiltonian is

$$H = \sum_{\alpha\beta} \epsilon_{\alpha\beta} a^{\dagger}_{\alpha} a_{\beta} + \frac{1}{4} \sum_{\alpha\beta\gamma\mu} V_{\alpha\beta\gamma\mu} a^{\dagger}_{\alpha} a^{\dagger}_{\beta} a_{\gamma} a_{\mu}.$$
(17)

Note the ordering of $\alpha\beta\gamma\mu$, thus $V_{\alpha\beta\gamma\mu} = -\langle\alpha\beta|V|\gamma\mu\rangle$. I assume time-even $H(\epsilon_{\alpha\beta} = \epsilon_{\tilde{\beta}\tilde{\alpha}}, V_{\alpha\beta\gamma\mu} = V_{\tilde{\mu}\tilde{\gamma}\tilde{\beta}\tilde{\alpha}})$ and real $\epsilon_{\alpha\beta}$ and $V_{\alpha\beta\gamma\mu}$. No other symmetry of H is assumed.

I compute the average energy $\bar{E} = \langle \phi_N | H | \phi_N \rangle$ in the canonical basis (3). For the one-body $\epsilon_{\alpha\beta}$ part, only the diagonal type $a^{\dagger}_{\alpha}a_{\alpha}$ contributes,

$$\langle 0|P^N a^{\dagger}_{\alpha} a_{\alpha} (P^{\dagger})^N |0\rangle = \chi_N \langle \phi_N |\hat{n}_{\alpha} |\phi_N \rangle = (N v_{\alpha})^2 \chi^{[\alpha]}_{N-1},$$
(18)

which is Eq. (9). For the two-body $V_{\alpha\beta\gamma\mu}$ part, only three mutually exclusive types contribute $(P_{\alpha} \neq P_{\beta})$: $a_{\alpha}^{\dagger} a_{\bar{\alpha}}^{\dagger} a_{\bar{\alpha}} a_{\alpha}$, $a_{\alpha}^{\dagger} a_{\bar{\alpha}}^{\dagger} a_{\bar{\alpha}} a_{\beta}$, and $a_{\alpha}^{\dagger} a_{\beta}^{\dagger} a_{\beta} a_{\alpha}$. The first type is

$$type1 = \langle 0|P^{N}a_{\alpha}^{\dagger}a_{\tilde{\alpha}}^{\dagger}a_{\tilde{\alpha}}a_{\alpha}(P^{\dagger})^{N}|0\rangle = \langle 0|P^{N}a_{\alpha}^{\dagger}a_{\alpha}(P^{\dagger})^{N}|0\rangle$$
$$= \chi_{N}\langle\phi_{N}|\hat{n}_{\alpha}|\phi_{N}\rangle, \tag{19}$$

because $|\alpha\rangle$ and $|\tilde{\alpha}\rangle$ are either both occupied or both empty in $(P^{\dagger})^{N}|0\rangle$. The second type $a_{\alpha}^{\dagger}a_{\tilde{\alpha}}^{\dagger}a_{\beta}a_{\beta} = P_{\alpha}^{\dagger}P_{\beta} = P_{\beta}P_{\alpha}^{\dagger}$, so Eqs. (5) and (7) imply

$$type2 = \langle 0|P^{N}a_{\alpha}^{\dagger}a_{\tilde{\alpha}}^{\dagger}a_{\tilde{\beta}}a_{\beta}(P^{\dagger})^{N}|0\rangle = t_{\beta;\alpha}^{N+1}$$
$$= N^{2}v_{\alpha}v_{\beta}\chi_{N-1}^{[\alpha\beta]}.$$
(20)

The third type $a^{\dagger}_{\alpha}a^{\dagger}_{\beta}a_{\beta}a_{\alpha} = 1 - a_{\alpha}a^{\dagger}_{\alpha} - a_{\beta}a^{\dagger}_{\beta} + a_{\alpha}a_{\beta}a^{\dagger}_{\beta}a^{\dagger}_{\alpha}$ by basic anticommutation relation, so definition (6) implies

$$type3 = \langle 0|P^{N}a_{\alpha}^{\dagger}a_{\beta}^{\dagger}a_{\beta}a_{\alpha}(P^{\dagger})^{N}|0\rangle$$
$$= \chi_{N} - \chi_{N}^{[\alpha]} - \chi_{N}^{[\beta]} + \chi_{N}^{[\alpha\beta]}.$$

Using $\chi_N - \chi_N^{[\alpha]} = N^2 (v_\alpha)^2 \chi_{N-1}^{[\alpha]}$ [Eq. (9)] and $\chi_N^{[\beta]} - \chi_N^{[\alpha\beta]} = N^2 (v_\alpha)^2 \chi_{N-1}^{[\alpha\beta]}$ [Eq. (11)], then factorizing out $N^2 (v_\alpha)^2$, I have

$$type3 = N^2 (v_{\alpha})^2 (\chi_{N-1}^{[\alpha]} - \chi_{N-1}^{[\alpha\beta]}).$$

In Eq. (11) I replace N by N - 1 ($N \rightarrow N - 1$) and exchange α and β ($\alpha \leftrightarrow \beta$),

$$\chi_{N-1}^{[\alpha]} - \chi_{N-1}^{[\alpha\beta]} = (N-1)^2 (v_\beta)^2 \chi_{N-2}^{[\alpha\beta]}$$
$$= \chi_{N-1}^{[\alpha]} \langle \phi_{N-1}^{[\alpha]} | \hat{n}_\beta | \phi_{N-1}^{[\alpha]} \rangle.$$
(21)

Thus I have two equivalent expressions,

$$type3 = N^{2}(N-1)^{2}(v_{\alpha}v_{\beta})^{2}\chi_{N-2}^{[\alpha\beta]}, \qquad (22)$$

or

$$type3 = N^{2}(v_{\alpha})^{2} \chi_{N-1}^{[\alpha]} \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle$$
$$= \chi_{N} \langle \phi_{N} | \hat{n}_{\alpha} | \phi_{N} \rangle \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle.$$
(23)

The last step is Eq. (9).

The expectation value of H(17) is

$$\begin{split} \langle \phi_N | H | \phi_N \rangle &= \sum_{\alpha \in \Theta} 2\epsilon_{\alpha\alpha} \langle \phi_N | a^{\dagger}_{\alpha} a_{\alpha} | \phi_N \rangle \\ &+ \sum_{\alpha \in \Theta} V_{a\tilde{\alpha}\tilde{\alpha}\alpha} \langle \phi_N | a^{\dagger}_{\alpha} a^{\dagger}_{\bar{\alpha}} a_{\bar{\alpha}} a_{\alpha} | \phi_N \rangle \\ &+ \sum_{\alpha,\beta \in \Theta}^{\alpha \neq \beta} V_{\alpha\tilde{\alpha}\tilde{\beta}\beta} \langle \phi_N | a^{\dagger}_{\alpha} a^{\dagger}_{\bar{\alpha}} a_{\beta} a_{\beta} | \phi_N \rangle \\ &+ \sum_{\alpha,\beta \in \Theta}^{\alpha \neq \beta} (V_{\alpha\beta\beta\alpha} + V_{\alpha\tilde{\beta}\tilde{\beta}\alpha}) \langle \phi_N | a^{\dagger}_{\alpha} a^{\dagger}_{\beta} a_{\beta} a_{\alpha} | \phi_N \rangle. \end{split}$$

$$(24)$$

The summations run over pair index α and β . The first term collects two equal contributions by $\epsilon_{\alpha\alpha} = \epsilon_{\tilde{\alpha}\tilde{\alpha}}$, which gives the factor 2. The second term collects four equal contributions by $V_{\alpha\tilde{\alpha}\tilde{\alpha}\alpha} = -V_{\alpha\tilde{\alpha}\alpha\tilde{\alpha}} = -V_{\tilde{\alpha}\alpha\tilde{\alpha}\alpha}$, which cancels the factor 1/4 in the Hamiltonian (17). The third term collects $V_{\alpha\tilde{\alpha}\tilde{\beta}\beta} = -V_{\alpha\tilde{\alpha}\beta\tilde{\beta}} = -V_{\tilde{\alpha}\alpha\tilde{\beta}\beta} = V_{\tilde{\alpha}\alpha\beta\tilde{\beta}}$. The fourth term collects $V_{\alpha\beta\alpha\beta} = -V_{\alpha\beta\alpha\beta} = -V_{\tilde{\alpha}\beta\tilde{\beta}\tilde{\alpha}} = -V_{\tilde{\alpha}\tilde{\beta}\tilde{\alpha}\tilde{\beta}}$ and $V_{\alpha\tilde{\beta}\tilde{\beta}\alpha} = -V_{\alpha\tilde{\beta}\alpha\tilde{\beta}} = -V_{\tilde{\alpha}\beta\tilde{\alpha}\beta}$.

Substituting Eqs. (18), (19), (20), (22) into Eq. (24),

$$\begin{aligned} \langle \phi_N | H | \phi_N \rangle &= \frac{N^2}{\chi_N} \Biggl(\sum_{\alpha \in \Theta} (2\epsilon_{\alpha\alpha} + G_{\alpha\alpha}) (v_\alpha)^2 \chi_{N-1}^{[\alpha]} \\ &+ \sum_{\alpha,\beta \in \Theta}^{\alpha \neq \beta} G_{\alpha\beta} v_\alpha v_\beta \chi_{N-1}^{[\alpha\beta]} \\ &+ (N-1)^2 \sum_{\alpha,\beta \in \Theta}^{\alpha \neq \beta} \Lambda_{\alpha\beta} (v_\alpha v_\beta)^2 \chi_{N-2}^{[\alpha\beta]} \Biggr), \end{aligned}$$
(25)

where I introduce the paring matrix elements $G_{\alpha\beta}$ and the monopole matrix elements $\Lambda_{\alpha\beta}$ as

$$G_{\alpha\beta} \equiv V_{\alpha\tilde{\alpha}\tilde{\beta}\beta},\tag{26}$$

$$\Lambda_{\alpha\beta} = V_{\alpha\beta\beta\alpha} + V_{\alpha\tilde{\beta}\tilde{\beta}\alpha}.$$
 (27)

Note $G_{\alpha\beta} = G_{\beta\alpha} = G_{\alpha\bar{\beta}}$, $\Lambda_{\alpha\beta} = \Lambda_{\beta\alpha} = \Lambda_{\alpha\bar{\beta}}$, and $G_{\alpha\alpha} = \Lambda_{\alpha\alpha}$. Equation (25) expresses the average energy by normalizations and is adopted in coding. Another equivalent expression by occupation numbers reveals more physics. Using Eqs. (18) and (23),

$$\begin{split} \langle \phi_N | H | \phi_N \rangle &= \sum_{\alpha \in \Theta} (2\epsilon_{\alpha\alpha} + G_{\alpha\alpha}) \langle \phi_N | \hat{n}_{\alpha} | \phi_N \rangle \\ &+ N^2 \sum_{\alpha, \beta \in \Theta}^{\alpha \neq \beta} G_{\alpha\beta} v_{\alpha} v_{\beta} \frac{\chi_{N-1}^{[\alpha\beta]}}{\chi_N} \\ &+ \sum_{\alpha, \beta \in \Theta}^{\alpha \neq \beta} \Lambda_{\alpha\beta} \langle \phi_N | \hat{n}_{\alpha} | \phi_N \rangle \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle. \tag{28}$$

The analytical expression for the average energy has been given in a slightly different form in Ref. [19]. The gradient

of energy and others, crucial to the new VDPC algorithm, are new results of this work as given in the next section.

IV. GRADIENT OF ENERGY

In this section I derive the gradient of average energy with respect to the pair structure v_{α} (3). Moreover, I give the analytical formula of v_{α} that minimizes the energy. Finally I discuss the asymptotic behavior of v_{α} away from (above or below) the Fermi surface.

Equation (25) expresses the average energy \overline{E} in terms of (Pauli-blocked) normalizations χ_N . To compute gradient of \overline{E} , I first compute gradient of χ_N . Under infinitesimal change δv_{α} of a single v_{α} , the variation of P^{\dagger} (3) and $(P^{\dagger})^N$ are

$$\delta P^{\dagger} = (\delta v_{\alpha}) P_{\alpha}^{\dagger},$$

$$\delta[(P^{\dagger})^{N}] = N(P^{\dagger})^{N-1} \delta P^{\dagger} = N(\delta v_{\alpha}) P_{\alpha}^{\dagger} (P^{\dagger})^{N-1}.$$

Thus the variation of χ_N (2) is

$$\begin{split} \delta\chi_N &= \langle 0|\delta[P^N](P^{\dagger})^N|0\rangle + \langle 0|P^N\delta[(P^{\dagger})^N]|0\rangle \\ &= \langle 0|P^N\delta[(P^{\dagger})^N]|0\rangle + \text{H.c.} \\ &= N\langle 0|P^NP_{\alpha}^{\dagger}(P^{\dagger})^{N-1}|0\rangle\delta v_{\alpha} + \text{H.c.} \end{split}$$

H.c. means Hermitian conjugate, which in fact equals to the first term because v_{α} is real. Using Eqs. (5) and (7), I have

$$\delta\chi_{N} = 2Nt_{\alpha}^{N}\delta v_{\alpha} = 2N^{2}v_{\alpha}\chi_{N-1}^{[\alpha]}\delta v_{\alpha}$$
$$= \frac{2\chi_{N}}{v_{\alpha}}\langle\phi_{N}|\hat{n}_{\alpha}|\phi_{N}\rangle\delta v_{\alpha} = \frac{2}{v_{\alpha}}(\chi_{N}-\chi_{N}^{[\alpha]})\delta v_{\alpha}.$$
 (29)

The last two steps use Eq. (9).

If I Pauli block the β index $(P_{\beta} \neq P_{\alpha})$ from the very beginning, the derivation remains valid, so Eq. (29) implies

$$\delta \chi_N^{[\beta]} = 2N^2 v_\alpha \chi_{N-1}^{[\alpha\beta]} \delta v_\alpha = \frac{2\chi_N^{[\beta]}}{v_\alpha} \langle \phi_N^{[\beta]} | \hat{n}_\alpha | \phi_N^{[\beta]} \rangle \delta v_\alpha$$
$$= \frac{2}{v_\alpha} (\chi_N^{[\beta]} - \chi_N^{[\alpha\beta]}) \delta v_\alpha, \quad P_\beta \neq P_\alpha. \tag{30}$$

And of course $\delta \chi_N^{[\alpha]} = 0$. Similarly, I easily obtain $\delta \chi_N^{[\beta\gamma]}$ by Pauli blocking the two indices β and γ from the very beginning. Substituting $\delta \chi_N$ (29), $\delta \chi_N^{[\beta]}$ (30), and $\delta \chi_N^{[\beta\gamma]}$ into Eq. (25), using basic calculus then collecting similar terms, a two-page long derivation gives the energy gradient,

$$\frac{\partial \bar{E}}{\partial v_{\alpha}} = \frac{\partial [\langle \phi_{N} | H | \phi_{N} \rangle]}{\partial v_{\alpha}}$$

$$\stackrel{\exp 1}{=} \frac{-2N^{2}}{\chi_{N}} \left[\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]} + \frac{\chi_{N}^{[\alpha]}}{N^{2} v_{\alpha}} (\langle \phi_{N}^{[\alpha]} | H | \phi_{N}^{[\alpha]} \rangle - \bar{E}) \right]$$

$$\stackrel{\exp 2}{=} \frac{2N^{2}}{\chi_{N}} \left[\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]} + v_{\alpha} \chi_{N-1}^{[\alpha]} (d_{\alpha} + \langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle - \bar{E}) \right], \quad (32)$$

where

$$d_{\alpha} = 2\epsilon_{\alpha\alpha} + G_{\alpha\alpha} + 2\sum_{\beta\in\Theta}^{P_{\beta}\neq P_{\alpha}} \Lambda_{\alpha\beta} \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle \qquad (33)$$

$$= 2\epsilon_{\alpha\alpha} + G_{\alpha\alpha} + 2(N-1)^2 \sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} \Lambda_{\alpha\beta}(v_{\beta})^2 \frac{\chi_{N-2}^{[\alpha\beta]}}{\chi_{N-1}^{[\alpha]}}.$$
(34)

The two gradient expressions (31) and (32) are equivalent. d_{α} is the single-pair energy similar to the common single-particle HF energy. In Eq. (33), $2\epsilon_{\alpha\alpha} + G_{\alpha\alpha}$ is the unperturbed energy for a pair on orbits $|\alpha\rangle$ and $|\tilde{\alpha}\rangle$, this pair interacts with other pairs by energy $2\sum_{\beta\in\Theta}^{P_{\beta}\neq P_{\alpha}} \Lambda_{\alpha\beta} \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle$. Equation (34) is equivalent to Eq. (33), based on Eq. (11).

Because \overline{E} is independent of an overall scaling of v_{α} , the gradient of \overline{E} is perpendicular to \vec{v} ,

$$\nabla \bar{E} \cdot \vec{v} = \sum_{\alpha \in \Theta} v_{\alpha} \frac{\partial \bar{E}}{\partial v_{\alpha}} = 0$$

This identity is used to numerically check the computer code. Later I also need the HF single-particle energy

$$e_{\alpha} = \epsilon_{\alpha\alpha} + \sum_{\beta \in \text{SD}} V_{\alpha\beta\beta\alpha} = \epsilon_{\alpha\alpha} + \sum_{\beta \in \Theta}^{\beta \in \text{SD}} \Lambda_{\alpha\beta}, \quad (35)$$

where $\beta \in SD$ means the β orbit is occupied in the HF Slater determinant. The Fermi energy $e_F \equiv (e_{h.o.} + e_{l.e.})/2$, where $e_{h.o.}$ and $e_{l.e.}$ are e_{α} of the highest (energy e_{α}) occupied and the lowest empty HF single-particle level. In Eq. (3), if I set v_{α} to 1 for occupied orbits and to 0 for empty orbits, the pair condensate (1) reduces to the HF Slater determinant. In this case $d_{\alpha} \approx 2e_{\alpha}$.

At energy minimum, the gradient (31) and (32) vanish, which implies

$$v_{\alpha} \stackrel{\exp 1}{=} \frac{\langle \phi_{N}^{[\alpha]} | H | \phi_{N}^{[\alpha]} \rangle - \bar{E}}{-N^{2} \left(\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]} \right) / \chi_{N}^{[\alpha]}}$$

$$= \frac{\langle \phi_{N}^{[\alpha]} | H | \phi_{N}^{[\alpha]} \rangle - \bar{E}}{-\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} \frac{1}{v_{\beta}} \langle \phi_{N}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N}^{[\alpha]} \rangle}$$

$$\stackrel{\exp 2}{=} \frac{-\left(\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]} \right) / \chi_{N-1}^{[\alpha]}}{d_{\alpha} + \langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle - \bar{E}}$$

$$= \frac{-\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} (1 - \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]}) \right)}{d_{\alpha} + \langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle - \bar{E}}$$
(37)

The last step of Eqs. (36) and (37) use Eq. (11). Analytically, Eq. (31) is equivalent to Eq. (32), and Eq. (36) is equivalent to Eq. (37). Numerically, I prefer Eq. (31) or Eq. (36) when $e_{\alpha} \ll e_F$ (e_F is the Fermi energy, here \ll means the α orbit is well below the Fermi surface), and prefer Eq. (32) or Eq. (37) when $e_{\alpha} \gg e_F$. When $e_{\alpha} \ll e_F$, physical arguments imply $\langle \phi_N^{[\alpha]} | H | \phi_N^{[\alpha]} \rangle - \bar{E} \approx 2(e_F - e_{\alpha})$ and $d_{\alpha} + \langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle - \bar{E} \approx 0$. I want to avoid the \approx 0 case to avoid the numerical sign problem (subtract two very closed numbers, $d_{\alpha} + d_{\alpha} = 0$.

 $\langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle$ and \bar{E}), so I prefer Eq. (31) or Eq. (36). When $e_{\alpha} \gg e_{F}$, physical arguments imply $\langle \phi_{N}^{[\alpha]} | H | \phi_{N}^{[\alpha]} \rangle - \bar{E} \approx 0$ and $d_{\alpha} + \langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle - \bar{E} \approx 2(e_{\alpha} - e_{F})$. I want to avoid the $\langle \phi_{N}^{[\alpha]} | H | \phi_{N}^{[\alpha]} \rangle - \bar{E} \approx 0$ case, so prefer Eq. (32) or Eq. (37).

The above analysis also implies the asymptotic behavior of v_{α} away from (above or below) the Fermi surface,

$$v_{\alpha} \approx \frac{2(e_{F} - e_{\alpha})}{-N^{2} \left(\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]} \right) / \chi_{N}^{[\alpha]}}$$
$$= \frac{2(e_{F} - e_{\alpha})}{-\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} \frac{1}{v_{\beta}} \langle \phi_{N}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N}^{[\alpha]} \rangle}, \ e_{\alpha} \ll e_{F}, \quad (38)$$

and

$$v_{\alpha} \approx \frac{-\left(\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]}\right) / \chi_{N-1}^{[\alpha]}}{2(e_{\alpha} - e_{F})} = \frac{-\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \left(1 - \left\langle \phi_{N-1}^{[\alpha]} \middle| \hat{n}_{\beta} \middle| \phi_{N-1}^{[\alpha]} \right\rangle \right)}{2(e_{\alpha} - e_{F})}, \ e_{\alpha} \gg e_{F}.$$

$$(39)$$

In Eqs. (38) and (39), the summation

$$\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} \chi_{N-1}^{[\alpha\beta]}$$

$$= \frac{\chi_{N}^{[\alpha]}}{N^{2}} \sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} \frac{1}{v_{\beta}} \langle \phi_{N}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N}^{[\alpha]} \rangle$$

$$P_{\alpha} \neq P_{\alpha}$$
(40)

$$=\chi_{N-1}^{[\alpha]}\sum_{\beta\in\Theta}^{P_{\alpha}}G_{\alpha\beta}v_{\beta}\left(1-\left\langle\phi_{N-1}^{[\alpha]}\middle|\hat{n}_{\beta}\middle|\phi_{N-1}^{[\alpha]}\right\rangle\right) \quad (41)$$

depends on the details of interaction, and should have important contributions when the β orbit is near the Fermi surface. If $e_{\beta} \ll e_F$, v_{β} is very large and $\langle \phi_N^{[\alpha]} | \hat{n}_{\beta} | \phi_N^{[\alpha]} \rangle \approx 1$, Eq. (40) shows that this $G_{\alpha\beta}$ term is suppressed by the factor $1/v_{\beta}$, compared with those terms near the Fermi surface. If $e_{\beta} \gg$ e_F , v_{β} is very small and $\langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle \approx 0$, Eq. (41) shows that this $G_{\alpha\beta}$ term is suppressed by the factor v_{β} , compared with those terms near the Fermi surface.

When $e_{\alpha} \ll e_F$, generally v_{α} should increase when e_{α} decreases, and the linear factor $e_F - e_{\alpha}$ in the numerator of Eq. (38) contributes to this trend. The other factor $\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} \frac{1}{v_{\beta}} \langle \phi_N^{[\alpha]} | \hat{n}_{\beta} | \phi_N^{[\alpha]} \rangle$ in the denominator should also contribute to this trend by the decaying of $G_{\alpha\beta}$, because β was near the Fermi surface as explained above. When $e_{\alpha} \gg e_F$, generally v_{α} should decrease when e_{α} increases, and the inverse-linear factor $1/(e_{\alpha} - e_F)$ in the denominator of Eq. (39) contributes to this trend. The other factor $\sum_{\beta \in \Theta}^{P_{\beta} \neq P_{\alpha}} G_{\alpha\beta} v_{\beta} (1 - \langle \phi_{N-1}^{[\alpha]} | \hat{n}_{\beta} | \phi_{N-1}^{[\alpha]} \rangle)$ in the numerator should also contribute to this trend by the decaying of $G_{\alpha\beta}$.

The exact [Eqs. (36) and (37)] and asymptotic [Eqs. (38) and (39)] expressions for v_{α} are key to the new VDPC algorithm as given in the next section.

V. COMPUTER ALGORITHM

In this section I design the computer algorithm to minimize the average energy \bar{E} . The variational parameters are the pair structure v_{α} (3). I have expressed \bar{E} (25) and its gradient $\frac{\partial \bar{E}}{\partial v_{\alpha}}$ [Eqs. (31) and (32)] in terms of v_{α} . In principle, coding these expressions and choosing an available minimizer (for example FMINUNC in MATLAB) solve the problem.

Practically, in large model spaces the sign problem may arise. If I normalize v_{α} such that $v_{\alpha} \approx 1$ (the order of magnitude) near the Fermi surface, v_{α} could be very large (small) for $e_{\alpha} \ll e_F (e_{\alpha} \gg e_F)$ orbits: typically $v_{\alpha} \approx 10 (v_{\alpha} \approx 0.01)$ near $e_{\alpha} = e_F - 20$ MeV ($e_{\alpha} = e_F + 20$ MeV). Then the sign problem may arise when recursively computing $\chi_N^{[\alpha]}$ by Eq. (9).

Most numerical softwares run very quickly using numbers of double-precision floating-point format. (Because usually the machine precision is double precision.) So in practice, if I use double-precision numbers, how large could the model space be? My experience is that there is no sign problem up to the case of 2N = 24 particles in a single-particle space of dimension $D = 2\Omega = 48$. (The model space is half-filled; the particle number is the largest considering the particlehole symmetry [30,31]. Ω is the number of vacancies for Kramers pairs.) In this case MATLAB FMINUNC costs typically 0.3 second to minimize \overline{E} , on a laptop by serial computing (not in parallel).

The modern mean-field theory uses large model spaces (for example, 15 harmonic-oscillator major shells). In this case double precision is not enough, and I resort to softwares that run quickly using arbitrary-precision numbers, for example, Mathematica. In principle, any algorithm running into the sign problem could use this strategy of increasing precision. However, in practice, computing with arbitrary-precision numbers is usually much slower than with double-precision numbers; so the formulas for coding must be simple so that the total computer time cost is feasible. In this work I use 120 effective digits (by *Mathematica* function SetPrecision[120]) to compute the average energy \overline{E} (25), which overcomes the sign problem. Then I switch to double precision and compute the gradient $\frac{\partial \bar{E}}{\partial v_{\alpha}}$ [Eqs. (31) and (32)] and the new v_{α} [Eqs. (36) and (37)], which has no sign problem at all. Computing \overline{E} (by 120 effective digits) costs negligible time compared to computing $\frac{\partial \bar{E}}{\partial v_{\alpha}}$ and the new v_{α} (by double precision), so using arbitrary-precision numbers hardly affects the total time cost. And for this reason I do not fine tune precision in this work but always use 120 effective digits, which is more than enough to overcome the sign problem when computing \bar{E} .

The VDPC algorithm is designed to increase the valence space gradually: first minimizes \overline{E} in a small valence space (of dimension $2\Omega \approx 50$ to use double precision) around e_F to quickly get the big picture, next refines the solution in larger valence spaces until the desired convergence. The Nilsson levels below the valence space are completely filled and form an inert core, the core simply corrects the valence-space single-particle energy by its HF mean field. Specifically, the algorithm has five steps:

(i) I sort the single-particle basis states $|\alpha\rangle$ by their HF energy e_{α} (35), and occupy the lowest 2N basis states. In other words, I solve the HF equation but without mixing the basis states. (This work is VDPC+BCS, not VDPC+HFB.) This step is not needed if the input single-particle basis is already the HF basis. (ii) I select around e_F the first valence space (VS1) of dimension $2\Omega \approx 50$ (to use double precision). Then I input both the energy [Eq. (25)] and the gradient [Eqs. (31) and (32)] into MATLAB FMINUNC, to quickly minimize \overline{E} . The resultant v_{α} of VS1 is called $v_{\alpha}^{(1)}$. (iii) I select around e_F the second valence space (VS2) of dimension $2\Omega \approx 200$. I initialize v_{α} of VS2 to be $v_{\alpha}^{(1)}$ if $|\alpha\rangle$ belongs to VS1, and to be a very large (small) number if $|\alpha\rangle$ is not in VS1 and $e_{\alpha} < e_F$ $(e_{\alpha} > e_F)$ so that $n_{\alpha} \approx 1$ $(n_{\alpha} \approx 0)$. Then I use the analytical formulas (36) and (37) to iterate v_{α} until convergence (usually 10 iterations are enough). The resultant v_{α} of VS2 is called $v_{\alpha}^{(2)}$. VS2 is large enough so that $v_{\alpha}^{(2)}$ is very close to the final solution. (iv) For all the basis states $|\beta\rangle$ not in VS2, estimate v_{β} . I substitute $v_{\alpha}^{(2)}$ into the asymptotic expressions (38) and (39) to compute v_{β} . (This is the first-order perturbation: determine v_{β} from $v_{\alpha}^{(2)}$ of VS2.) Next I substitute $v_{\alpha}^{(2)}$ and v_{β} into Eqs. (38) and (39) again, to compute the final v_{β} , labeled as v_{β}^{est} . (This is the second-order perturbation: consider corrections from mutual interactions among v_{β} .) The corresponding occupation number is n_{β}^{est} . (v) Choose two cutoffs n_{\min} and n_{max} (for example $n_{\text{min}} = 10^{-6}$ and $n_{\text{max}} = 1 - 10^{-7}$), and select the third valence space (VS3): VS3 consists of VS2 and those basis states $|\beta\rangle$ satisfying $n_{\min} \leq n_{\beta}^{\text{est}} \leq n_{\max}$. I initialize v_{α} of VS3 to be $v_{\alpha}^{(2)}$ if $|\alpha\rangle$ belongs to VS2, and to be v_{β}^{est} if $|\alpha\rangle$ is not in VS2. Then I use the analytical formulas (36) and (37) to iterate v_{α} in VS3 until the desired convergence. The resultant v_{α} of VS3 is called $v_{\alpha}^{(3)}$. This finishes the VDPC algorithm.

The next section demonstrates the algorithm in a semirealistic example, giving the actual time cost and energyconvergence pattern.

VI. REALISTIC EXAMPLE

In this section I apply the VDPC+BCS algorithm to the semirealistic example of the rare-earth nucleus ${}^{158}_{64}$ Gd₉₄. (This example has been used in my recent paper [28] on deformed generalized seniority.) The purpose is to demonstrate the effectiveness of the algorithm under realistic interactions. For simplicity, I consider only the neutron degree of freedom, governed by the antisymmetrized two-body Hamiltonian

$$H = \sum_{\alpha} \epsilon_{\alpha} a_{\alpha}^{\dagger} a_{\alpha} + \frac{1}{4} \sum_{\alpha\beta\gamma\mu} V_{\alpha\beta\gamma\mu} a_{\alpha}^{\dagger} a_{\beta}^{\dagger} a_{\gamma} a_{\mu}.$$
(42)

The single-particle levels ϵ_{α} are eigenstates of the Nilsson model [32]. The Nilsson parameters are the same as in Ref. [28], here I only repeat $\beta = 0.349$ (the experimental quadrupole deformation [33]). The neutron residual interaction $V_{\alpha\beta\gamma\mu}$ is the low-momentum *NN* interaction $V_{\text{low-}k}$ [16] derived from the free-space N³LO potential [34].

Specifically, I use the code distributed by Hjorth-Jensen [35] to compute (without Coulomb, charge-symmetry breaking, or charge-independence breaking) the two-body matrix elements of $V_{\text{low-}k}$ in the spherical harmonic oscillator basis up to (including) the $\mathcal{N} = 14$ major shell, with the standard momentum cutoff 2.1 fm⁻¹. ($\mathcal{N} = 2n_r + l$ is the major-shell quantum number.) Then the Nilsson model is diagonalized in this spherical $\mathcal{N} \leq 14$ basis, the eigenenergies are ϵ_{α} and the eigen wave functions transform the spherical two-body matrix elements into those on the Nilsson basis as used in the Hamiltonian (42).

This procedure has several assumptions. Mainly the proton-neutron interaction generates the static deformation and self-consistently the Nilsson mean field. The residual proton-neutron interaction is neglected, and in the Hamiltonian (42) the part of the neutron-neutron interaction already included in the Nilsson mean field ϵ_{α} is not removed from $V_{\alpha\beta\gamma\mu}$. These assumptions make the example semirealistic. My goal is to demonstrate the effectiveness of the VDPC algorithm, not to accurately reproduce the experimental data.

All the numerical calculations of this work were done on a laptop. The laptop has one quad-core CPU (Intel Core i7-4710MQ @ 2.5 GHz), but I used only serial computing on a single core (no parallel computing). All time costs plotted in the figures or given in the text are the actual time costs spent on this laptop. This work uses MATLAB R2015a and *Mathematica* 10.2, to give the actual software version.

I follow the steps listed in Sec. V, and an example of a complete run is shown in Fig. 1. In less than 30 seconds, the energy error reduces to about 3 keV. (The exact minimum E(exact), converged energy in the full space, is given by a separate calculation. The full space consists of all Nilsson



FIG. 1. Energy and time in a complete run. The horizontal axis shows the five steps listed in Sec. V, where step (v) is divided into five iterations. The cross symbols correspond to the left axis and show the energy error after each step or iteration, relative to the exact minimum E(exact) (converged energy in the full space). The circle symbols correspond to the right axis and show the computer time spent by each step or iteration. All time costs in this work refer to that by serial computing on a laptop (CPU is Intel Core i7-4710MQ @ 2.5 GHz, no parallel computing).



FIG. 2. Convergence of energy in VS1 by MATLAB. The horizontal axis shows the number of iterations. The vertical axis shows the energy at each iteration E(iter), relative to the final converged energy E(final).

levels after diagonalizing in the spherical $\mathcal{N} \leq 14$ basis.) Below I explain and discuss each step.

Step (i) sorts the Nilsson basis by their HF energy e_{α} (35). This step costs negligible time. Figure 1 shows that the pairing correlation energy is 1.83 MeV (defined as the energy difference between the HF Slater determinant and the final coherent-pair condensate).

Step (ii) selects VS1, and uses MATLAB FMINUNC to minimize \bar{E} . Figure 1 chooses VS1 to be $(e_F - 5.38, e_F + 5.5)$, which consists of all Nilsson levels α satisfying $e_F - 5.38$ MeV $\langle e_{\alpha} \langle e_F + 5.5 \text{ MeV}$. VS1 has dimension $2\Omega = 48$ (24 Nilsson levels below e_F and 24 above e_F). Starting from a random initial v_{α} , MATLAB quickly minimizes \bar{E} in about 0.3 second. This process is plotted in Fig. 2: how the energy converges as the number of iterations increases. Figure 1 shows that VS1 has a cutoff (truncation) error of 1.33 MeV for energy. [The cutoff error is defined as the energy minimum (the converged answer) in VS1 minus E(exact).]

Step (iii) selects VS2, and uses *Mathematica* to minimize \overline{E} by iterating Eqs. (36) and (37). I choose VS2 to be $(-\infty, e_F + \Delta e)$, which consists of all Nilsson levels α satisfying $e_{\alpha} < e_F + \Delta e$. This is the conventional way of doing truncation: the valence space is limited to a finite energy window around the Fermi surface. Eight choices for the cutoff Δe generate eight different model spaces VS2, their dimensions and cutoff errors (relative to the full space when $\Delta e = +\infty$) are shown in Fig. 3. For two out of the eight choices, $\Delta e = 20$ MeV and 50 MeV, Fig. 4 plots the energy convergence patten. The accumulated computer time cost increases linearly with the number of iterations, so each iteration costs the same time approximately. The energy error decreases the fastest in the first few iterations (by several orders of magnitude). Overall, the curve is linear on the log-scale plot, so energy converges exponentially with the number of iterations. In plotting Fig. 1, I fix VS2 to be $(-\infty, e_F + 20)$ and do ten iterations: the time cost is 3.4 seconds and the energy cutoff error is 267 keV.



FIG. 3. The cutoff error and dimension of eight different model spaces (VS2). The horizontal axis shows the cutoff Δe , so the model space is $(-\infty, e_F + \Delta e)$. It consists of all Nilsson levels α satisfying $e_{\alpha} < e_F + \Delta e$, where e_F is the Fermi energy. The cross symbols correspond to the left axis and show the cutoff error: the energy of each model space $E(\Delta e)$ relative to the exact energy of the full space E(exact). The circle symbols correspond to the right axis and show the dimension of each model space.

Step (iv) estimates v_{α} in the full space by the asymptotic expressions (38) and (39). This step reduces the energy error to 18 keV in 2.4 seconds.

In step (v), I select VS3 by choosing two cutoffs n_{\min} and n_{\max} . There are 94 Nilsson levels below e_F , and in this work I include all of them by setting $n_{\max} = 1$. Ten choices of n_{\min} generate ten different VS3, their dimensions and cutoff errors (relative to the full space when $n_{\min} = 0$) are shown in Fig. 5. In plotting Fig. 1, I fix $n_{\min} = 3 \times 10^{-7}$ and do five iterations. The time cost is about 6 seconds per iteration, and the energy



FIG. 4. Energy and time in two different model spaces (VS2) by *Mathematica*. The solid and dotted lines correspond to the left vertical axis, and show the energy errors in the two model spaces. The circle and triangle symbols correspond to the right vertical axis, and show the accumulated computer time cost after each iteration.



FIG. 5. The cutoff error and dimension of ten different model spaces (VS3). The horizontal axis shows the cutoff n_{\min} , so the model space consists of all Nilsson levels α satisfying $n_{\alpha} \ge n_{\min}$. The cross and circle symbols have similar meanings to those of Fig. 3.

cutoff error is 2.4 keV. Figure 1 shows that step (v) costs the largest time, but the energy converges quickly so practically one needs very few iterations (around three iterations). In passing, extending to VDPC+HFB, one may not need the slowest step (v) when computing v_{α} on each intermediate canonical basis, because after step (iv) the energy error is already pretty small (18 keV). Only at the final iterations step (v) was needed to reach complete convergence.

This finishes my explanation of Fig. 1. Note that Fig. 1 is just an example; for a desired accuracy, fine tuning parameters in these five steps finds the shortest time cost.

The above discussions use two truncating schemes: by energy Δe and by occupation number n_{\min} . Figure 6 compares



FIG. 6. Compare the two truncating schemes shown in Figs. 3 and 5. The horizontal axis shows the cutoff error in each truncated subspace, and the vertical axis shows the subspace dimension.



FIG. 7. The error of the asymptotic v_{α} expression relative to the exact v_{α} expression. The relative error $RE \equiv (v_{\alpha}^{asymptotic}/v_{\alpha}^{exact}) - 1$. The horizontal axis shows $|v_{\alpha}^{exact}|$, the absolute value of the exact v_{α} . The dot symbols correspond to the left vertical axis and show |RE|. The cross symbols correspond to the right vertical axis and show the exact occupation number n_{α} .

them and shows that truncating by occupation number is more effective than by energy. For a desired accuracy (cutoff error), truncating by n_{\min} needs a smaller subspace dimension. The two schemes differ the most at small cutoff error or large dimension (the top-left corner). This means for Nilsson levels α well above the Fermi surface, n_{α} is not a smooth monotonically decreasing function of e_{α} , but fluctuating around an overall decreasing trend. Some Nilsson levels α are populated more (less) than their neighbours owing to large (small) magnitudes of $G_{\alpha\beta}$, as shown by Eq. (39). However at small dimension (the bottom-right corner), the two schemes differ little. Later I will show that time cost is proportional to the cube of dimension, so the advantage of truncating by n_{\min} over truncating by Δe is more apparent when evaluating the time cost instead of the dimension.

The asymptotic expressions (38) and (39) very well reproduce the exact v_{α} (36) and (37) away from the Fermi surface. Figure 7 compares them at the energy minimum in the full space. The horizontal axis shows $|v_{\alpha}^{exact}|$ instead of v_{α}^{exact} , because some v_{α}^{exact} are negative with very small magnitudes. (The range is $-2.40 \times 10^{-4} \le v_{\alpha}^{exact} \le 39.7$). Near the Fermi surface, the asymptotic expressions (38) and (39) are not justified so |RE| is big (|RE| is the absolute value of relative error). Going away from the Fermi surface, |RE| becomes smaller and smaller. There are 680 different v_{α} (the full-space dimension is 1360 = 680 × 2), 661 of them have |RE| <10%, 568 of them have |RE| < 1%. Figure 7 shows v_{α} at the energy minimum; near the minimum |RE| has a similar pattern, which makes step (iv) effective.

The new algorithm minimizes \overline{E} through iterating v_{α} , by the exact expressions (36) and (37) or the asymptotic



FIG. 8. Computer time cost of one iteration in different model spaces. The horizontal axis shows the dimension of each model space, and the vertical axis shows the time cost of one iteration (averaged over many iterations). The asterisk symbols iterate by the exact v_{α} expressions (36) and (37), and the circle symbols iterate by the asymptotic v_{α} expressions (38) and (39).

expressions (38) and (39). The computer time cost per each iteration mainly depends on the dimension of the (singleparticle) model space. Figure 8 shows that this time increases approximately linearly with dimension on the log-log plot. I perform a linear least-squares fit in the form $\log(T) = \alpha \log(D) + C$ (*T* is time in unit of second, *D* is dimension, α and *C* are fitting parameters). The result is $T = (D/339)^{2.96}$ for the exact v_{α} , and $T = (D/1245)^{1.76}$ for the asymptotic v_{α} . The latter is much quicker.

Finally I suggest some directions to further optimize the algorithm. First, Fig. 1 shows that step (v) costs the most time, in fact computing $\langle \phi_N^{[\alpha]} | H | \phi_N^{[\alpha]} \rangle$ and $\langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle$ in Eqs. (36) and (37) is very time consuming. Currently I use only serial computing on a single core; an easy speedup would be computing $\langle \phi_N^{[\alpha]} | H | \phi_N^{[\alpha]} \rangle$ and $\langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle$ in parallel, by distributing each of them (each α) to different cores. Second, in large model spaces (for example $\mathcal{N} \leq 14$ of this work) majority of α orbits are above the Fermi surface and computed by Eq. (37). For those $e_{\alpha} \gg e_F$ orbits, a very good approximation [better than Eq. (39)] to Eq. (37) would be replacing $\langle \phi_{N-1}^{[\alpha]} | H | \phi_{N-1}^{[\alpha]} \rangle$ by $\langle \phi_{N-1} | H | \phi_{N-1} \rangle$. If this approximation caused little error in the final average energy, it should be used to greatly reduce the time cost. Third, Fig. 1 shows that step (iv) [iterates asymptotic v_{α} expressions (38) and (39)] is very quick and very effective, it could be used many times at different places.

My results suggest that the realistic $V_{\text{low-}k}$ interaction does not cause divergences in the pairing channel. In this work the full space ($\mathcal{N} \leq 14$) has dimension 1360. Figure 5 shows that in the truncated subspace of dimension 426 ($n_{\min} = 6 \times 10^{-6}$), the energy cutoff error is already less than 26 keV energy has converged, roughly speaking. In some cases, a few tens of keV may be important [36], then the tiny occupation numbers (for example smaller than 6×10^{-6}) can not be neglected and one should further enlarge the subspace.

What is the error of VDPC+BCS relative to the exact solution of the Hamiltonian (42)? References [9,10] studied this question using the state-independent pairing Hamiltonian (constant pairing strength $G_{\alpha\beta} = G$), where exact solution is available by the Richardson's method [37,38]. For this model Hamiltonian, VDPC+BCS is quite accurate in small spaces up to $\Omega \approx 20$ [10], but the error is evident in very large spaces [9,10]. In Sm isotopes, the error of pairing correlation energy is about 1 MeV using a space of 11 oscillator shells ($\Omega = 286$) [9]. However, this model Hamiltonian is not very realistic: for soft realistic interactions such as V_{low-k} , globally $G_{\alpha\beta}$ decays (magnitude decreases) as $|\epsilon_{\alpha} - \epsilon_{\beta}|$ increases, which limits the space size active in the pairing channel. The error of HF+VDPC+BCS in large spaces under realistic interactions remains an open question, because the exact solution is difficult. Section VII gives some suggestions to go beyond VDPC: break coherent pairs and perform configuration mixing; use two or more different coherent pairs (different v_{α}).

VII. CONCLUSIONS

This work proposes a new algorithm that applies the variational principle directly to the coherent-pair condensate (VDPC). It always conserves the particle number, and avoids the time-consuming particle-number projection by gauge-angle integration. Specifically, I derive analytical expressions for the average energy and its gradient in terms of the coherent-pair structure v_{α} . Requiring the gradient vanishes, I obtain the analytical expression of v_{α} at the energy minimum. The new VDPC algorithm iterates this v_{α} expression to minimize energy until practically arbitrary precision. I also find the asymptotic expression of v_{α} that is highly accurate (see Fig. 7) and numerically very fast (see Fig. 8). These analytical expressions look quite simple and allow easy physical interpretations.

I demonstrate the new VDPC algorithm in a semirealistic example using the realistic V_{low-k} interaction and large model spaces (up to 15 harmonic-oscillator major shells). The energy-convergence pattern and actual computer time cost are given in detail. Figure 1 shows an example run from beginning to end. How to organize the analytical results of this work into an optimal numerical algorithm remains an open question, and some suggestions are given at the end of Sec. VI.

It is a good property of a specific interaction to converge naturally in the pairing channel; otherwise a phenomenological cutoff is needed to truncate the pairing-active model space. The zero-range pairing, frequently used together with the Skyrme density functional, does not have this property. The Gogny force has this good property. My results show that the realistic $V_{\text{low-}k}$ interaction has this good property. However, tiny occupation numbers contribute to the energy (see Fig. 5), thus should be kept if the desired accuracy is high.

This work considers VDPC+BCS. Extending to VDPC+HFB needs the gradient of the average energy with respect to changes of the canonical single-particle basis. They have been derived analytically and will be the topic of my next work in the series.

I suggest two directions to go beyond VDPC to further improve the ground state (and low-lying states). First, one can break coherent pairs and perform configuration mixing [28], which is the generalized-seniority truncation [27,28,39–42]. Second, the variational ground state can have two or more different (different v_{α}) coherent pairs instead of only one. This is justified in for example neutron-halo nuclei, where the coherent pairs in and not in the halo may be different.

ACKNOWLEDGMENTS

Support is acknowledged from the National Natural Science Foundation of China No. 11405109.

- M. Bender, P.-H. Heenen, and P.-G. Reinhard, Rev. Mod. Phys. 75, 121 (2003).
- [2] A. Bohr, B. R. Mottelson, and D. Pines, Phys. Rev. 110, 936 (1958).
- [3] A. Bohr and B. Mottelson, *Nuclear Structure* (Benjamin, New York, 1975).
- [4] P. Ring and P. Schuck, *The Nuclear Many-Body Problem* (Springer-Verlag, Berlin, 1980).
- [5] K. Dietrich, H. J. Mang, and J. H. Pradal, Phys. Rev. 135, B22 (1964).
- [6] M. Anguiano, J. L. Egido, and L. M. Robledo, Phys. Lett. B 545, 62 (2002).
- [7] G. F. Bertsch, C. A. Bertulani, W. Nazarewicz, N. Schunck, and M. V. Stoitsov, Phys. Rev. C 79, 034306 (2009).
- [8] J. Dukelsky and G. Sierra, Phys. Rev. B 61, 12302 (2000).
- [9] G. G. Dussel, S. Pittel, J. Dukelsky, and P. Sarriguren, Phys. Rev. C 76, 011302(R) (2007).
- [10] N. Sandulescu and G. F. Bertsch, Phys. Rev. C 78, 064318 (2008).

- [11] J. Sheikh and P. Ring, Nucl. Phys. A 665, 71 (2000).
- [12] M. Anguiano, J. Egido, and L. M. Robledo, Nucl. Phys. A 696, 467 (2001).
- [13] M. V. Stoitsov, J. Dobaczewski, R. Kirchner, W. Nazarewicz, and J. Terasaki, Phys. Rev. C 76, 014308 (2007).
- [14] G. Hupin and D. Lacroix, Phys. Rev. C 86, 024309 (2012).
- [15] X. B. Wang, J. Dobaczewski, M. Kortelainen, L. F. Yu, and M. V. Stoitsov, Phys. Rev. C 90, 014312 (2014).
- [16] S. Bogner, T. T. S. Kuo, and A. Schwenk, Phys. Rep. 386, 1 (2003).
- [17] A. Bulgac and Y. Yu, Phys. Rev. Lett. 88, 042504 (2002).
- [18] A. Bulgac, Phys. Rev. C 65, 051305 (2002).
- [19] G. Hupin, D. Lacroix, and M. Bender, Phys. Rev. C 84, 014309 (2011).
- [20] G. Hupin and D. Lacroix, Phys. Rev. C 83, 024317 (2011).
- [21] H. J. Lipkin, Ann. Phys. (N.Y.) 9, 272 (1960).
- [22] Y. Nogami, Phys. Rev. 134, B313 (1964).
- [23] H. C. Pradhan, Y. Nogami, and J. Law, Nucl. Phys. A 201, 357 (1973).

- [24] Igal Talmi, Simple Models of Complex Nuclei: The Shell Model and Interacting Boson Model (Harwood Academic, Chur, 1993).
- [25] A. L. Goodman, Adv. Nucl. Phys. 11, 263 (1979).
- [26] J. A. Sheikh, P. Ring, E. Lopes, and R. Rossignoli, Phys. Rev. C 66, 044318 (2002).
- [27] L. Y. Jia, J. Phys. G: Nucl. Part. Phys. 42, 115105 (2015).
- [28] L. Y. Jia, Phys. Rev. C 96, 034313 (2017).
- [29] L. Y. Jia, Phys. Rev. C 88, 044303 (2013).
- [30] I. Talmi, Phys. Rev. C 25, 3189 (1982).
- [31] L. Y. Jia, Phys. Rev. C 93, 064307 (2016).
- [32] S. G. Nilsson, Mat. Fys. Medd. Dan. Vid. Selsk. 29, 16 (1955).
- [33] Data retrieved from the National Nuclear Data Center (Brookhaven National Laboratory), http://www.nndc.bnl.gov/.

- [34] D. R. Entem and R. Machleidt, Phys. Rev. C 68, 041001(R) (2003).
- [35] https://github.com/ManyBodyPhysics/ManybodyCodes/CENS
- [36] D. Lunney, J. M. Pearson, and C. Thibault, Rev. Mod. Phys. 75, 1021 (2003).
- [37] R. W. Richardson, Phys. Lett. 3, 277 (1963); Phys. Rev. 141, 949 (1966).
- [38] J. Dukelsky, S. Pittel, and G. Sierra, Rev. Mod. Phys. **76**, 643 (2004).
- [39] K. Allaart, E. Boeker, G. Bonsignori, M. Savoia, and Y. K. Gambhir, Phys. Rep. 169, 209 (1988).
- [40] Y. M. Zhao and A. Arima, Phys. Rep. 545, 1 (2014).
- [41] L. Y. Jia and C. Qi, Phys. Rev. C 94, 044312 (2016).
- [42] C. Qi, L. Y. Jia, and G. J. Fu, Phys. Rev. C 94, 014312 (2016).