Broyden's method in nuclear structure calculations

Andrzej Baran,[1,2,3] Aurel Bulgac,[4] Michael McNeil Forbes,[4] Gaute Hagen,[2] Witold Nazarewicz,[1,2,5,6] Nicolas Schunck,[1,2] and Mario V. Stoitsov[1,2,7]

[1]*Department of Physics & Astronomy, University of Tennessee, Knoxville, Tennessee 37996, USA*
[2]*Physics Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831, USA*
[3]*Institute of Physics, University of M. Curie-Sklodowska, ul. Radziszewskiego 10, 20-031 Lublin, Poland*
[4]*Department of Physics, University of Washington, Seattle, Washington 98195-1560, USA*
[5]*Institute of Theoretical Physics, Warsaw University, ul. Hoża 69, 00-681 Warsaw, Poland*
[6]*School of Engineering and Science, University of the West of Scotland, Paisley PA1 2BE, United Kingdom*
[7]*Institute of Nuclear Research and Nuclear Energy, Bulgarian Academy of Sciences, Sofia, Bulgaria*

(Received 28 May 2008; published 24 July 2008)

Broyden's method, widely used in quantum chemistry electronic-structure calculations for the numerical solution of nonlinear equations in many variables, is applied in the context of the nuclear many-body problem. Examples include the unitary gas problem, the nuclear density functional theory with Skyrme functionals, and the nuclear coupled-cluster theory. The stability of the method, its ease of use, and its rapid convergence rates make Broyden's method a tool of choice for large-scale nuclear structure calculations.

DOI: 10.1103/PhysRevC.78.014318 PACS number(s): 21.10.Dr, 21.60.Jz, 21.60.De, 71.15.Mb

I. INTRODUCTION

The nuclear many-body problem is undergoing a renaissance. Hand in hand with the experimental developments in the science of rare isotopes, a qualitative change in theoretical modeling of the nucleus is taking place. Developments of powerful conceptual, analytic, algorithmic, and computational tools enable scientists to probe the inner workings of nuclei with far greater precision than previously possible. These new tools, including terascale computer platforms, make researchers optimistic that the goal of developing a comprehensive, quantitative, and predictive theory of the nucleus and nucleonic matter is achievable [1,2].

The research challenges faced by the nuclear structure community are often interdisciplinary. Indeed, the quantum many-body problem represents one of the great intellectual and numerical challenges for nuclear and hadronic physics, quantum chemistry, atomic and condensed matter physics, and materials sciences. Often, to solve problems in many-body science, close interactions of physicists with computer scientists and mathematicians are required.

A theoretical framework aiming at the microscopic description of many-body systems and capable of extrapolating into unknown regions must fulfill several requirements. Namely, (i) it must be general enough to be confidently applied to unknown species whose properties are largely unknown; (ii) it should be capable of handling symmetry-breaking effects inherent to finite systems; (iii) it should be able to describe both finite systems and bulk matter; and (iv) in addition to observables, the method should provide associated error bars. These requirements are met by the density functional theory (DFT) in the formulation of Kohn and Sham [3].

The generalization of the DFT to the case of fermionic pairing was formulated for electronic superconductors [4]. The resulting Hartree-Fock-Bogoliubov (HFB) or Bogoliubov de-Gennes (BDG) equations can be viewed as the generalized Kohn-Sham equations of the standard DFT.

The nuclear DFT provides a reliable method for calculating properties of nuclei across the whole nuclear mass chart. The main ingredient of the nuclear DFT [5] is the energy density functional that depends on proton and neutron densities and currents representing distributions of nucleonic matter, spins, momentum, and kinetic energy, as well as their derivatives (gradient terms). If pairing correlations are present, the particle densities are augmented by pairing densities representing correlated nucleonic pairs [6]. Initially, attempts to build effective energy functionals were rooted in the empirical zero-range Skyrme interaction treated within the Hartree-Fock (HF) or HFB approximation. Following the success of DFT in atomic and molecular physics and, more recently, in the physics of ultracold atoms, it was realized that the interaction could be secondary to the functional, and recent progress has been directed toward extensions of the energy functionals to less traditional forms.

In parallel, much effort is put into trying to connect the nuclear DFT to more fundamental many-body theory based on internucleon interactions. To extend the *ab initio* program to the medium-mass region of the nuclear chart, a method that scales softly with system size and has a controllable error estimate is called for. The coupled-cluster (CC) theory is ideal in this respect. It scales polynomial with system size and it is *size extensive*, meaning that the energy scales correctly with increasing system size.

This work represents a collaborative effort performed under the Universal Nuclear Energy Density Functional (UNEDF) project [2]. The goal of UNEDF is to develop a new-generation theoretical framework that will describe all nuclei and their reactions by making use of the massive computer power. One of the objectives of UNEDF is to develop computational infrastructure for leadership-class computers; hence the algorithmic and computational focus of this paper.

The common feature of DFT and CC is to solve the quantum many-body problem for interacting fermions through self-consistent convergence schemes. The computational cost

of these iterative procedures can become very expensive, especially when the size of the model space or the number of nuclei processed simultaneously is huge. The advent of teraflop computers make such large-scale calculations feasible, but to take full advantage of unique resources, better convergence algorithms are called for. The aim of the present article is to report the implementation of Broyden's method [7], a quasi-Newton algorithm to solve large sets of nonlinear equations, in nuclear structure calculations. The HFB (or BDG) equations—the Kohn-Sham equations of the superfluid DFT—represent a coupled system of nonlinear equations for densities or mean fields. The nonlinearity enters through the self-consistency, i.e., the dependence of mean fields on densities. The CC equations have in fact a very similar structure. Here, the nonlinearity is hidden in the CC intermediates that depend on the particle-hole amplitudes.

Self-consistent formulations of these approaches usually diverge when using straight iterations. The simplest method commonly used to avoid divergences is the so-called linear mixing, in which input and output at a given iteration of the self-consistent process are mixed to provide input for the next step. We show that Broyden's method leads to much improved convergence rates. In many cases, the number of iterations required to reach convergence is lower by a factor of 3 to 10 with respect to linear-mixing iterative schemes.

The article is organized as follows. Broyden's method is outlined in Sec. II, both in its standard version (Sec. II A) that applies to a small number of unknowns, and in a modified version (Sec. II B) that is used to crunch very large sets of nonlinear equations. The examples of implementations come next. Section III discusses the performance of the standard Broyden's method for the case of the unitary gas of two fermionic superfluids in spherical geometry. In the following examples the modified Broyden's method has been used, as the size of the problem is intractable with the standard approach. The general case of the symmetry-unrestricted Skyrme-DFT theory is discussed in Sec. IV. Section V illustrates the particular case of an axially symmetric Skyrme-DFT problem and discusses the ability of the method to converge to the local energy extrema (i.e, minima, maxima, and saddle points). The performance of the method in *ab initio* coupled-cluster calculations, where the number of unknowns can be as large as 10^8, is presented in Sec. VI. Finally, the conclusions of our work are contained in Sec. VII.

II. BROYDEN'S MIXING METHOD

When solving a system of self-consistent equations, one usually begins with a set of initial conditions (in nuclear structure: single-particle wave functions, densities, fields, etc.) that linearize the problem. These initial conditions can be represented formally by an N-dimensional vector $\boldsymbol{V}_{\text{in}}^{(0)}$. Solving the self-consistent equations with these initial conditions defines a new vector

$$\boldsymbol{V}_{\text{out}}^{(m)} = \boldsymbol{I}\left[\boldsymbol{V}_{\text{in}}^{(m)}\right], \tag{1}$$

where $\boldsymbol{I}(\boldsymbol{V})$ is a function of the initial conditions. The self-consistency condition requires that the solution $\boldsymbol{V}^*$ be a fixed-point of this iteration: $\boldsymbol{I}(\boldsymbol{V}^*) = \boldsymbol{V}^*$.

The goal of a fixed-point iteration algorithm is to construct a new guess $\boldsymbol{V}_{\text{in}}^{(m+1)}$ in such a way that the iteration converges,

$$\boldsymbol{F}^{(m)} \equiv \boldsymbol{V}_{\text{out}}^{(m)} - \boldsymbol{V}_{\text{in}}^{(m)} \to 0, \tag{2}$$

within the required tolerance. Using the previous output as the next input, $\boldsymbol{V}_{\text{in}}^{(m+1)} = \boldsymbol{V}_{\text{out}}^{(m)}$, does not guarantee convergence. If the resulting divergence exhibits oscillatory behavior, the usual ansatz is to mix the input at iteration m to define the input at iteration $m+1$:

$$\boldsymbol{V}_{\text{in}}^{(m+1)} = \alpha \boldsymbol{V}_{\text{out}}^{(m)} + (1-\alpha)\boldsymbol{V}_{\text{in}}^{(m)} \tag{3a}$$

$$= \boldsymbol{V}_{\text{in}}^{(m)} + \alpha \boldsymbol{F}^{(m)}. \tag{3b}$$

This *simple mixing* slows down the iteration, and with a suitable choice of the mixing parameter $\alpha \in [0, 1]$, convergence may be achieved. However, even if one can find α that leads to a convergent solution, there are many instances where the convergence is very slow. In large-scale calculations involving huge numbers of cases, the running time to find poorly converging solutions may become prohibitive.

If the vector field $\boldsymbol{I}(\boldsymbol{V})$ is differentiable, one can use derivative information to improve convergence. The idea is to regard the self-consistent condition $\boldsymbol{F}^{(m)} = 0$ as a set of nonlinear equations and to approach the problem from the point of view of finding the root: $\boldsymbol{F}(\boldsymbol{V}^*) = \boldsymbol{I}(\boldsymbol{V}^*) - \boldsymbol{V}^* = 0$. This set of nonlinear equations can be solved using the multidimensional Newton-Raphson method, in which the next guess can be approximated by

$$\boldsymbol{V}_{\text{in}}^{(m+1)} = \boldsymbol{V}_{\text{in}}^{(m)} - \mathbf{B}^{(m)} \boldsymbol{F}^{(m)}, \tag{4}$$

where $\mathbf{B}^{(m)} = (\mathbf{J}^{(m)})^{-1}$ and $J_{jk}^{(m)} = \partial F_j^{(m)}(\boldsymbol{V})/\partial V_k$ is the Jacobian matrix of the nonlinear equations at $\boldsymbol{V} = \boldsymbol{V}_{\text{in}}^{(m)}$. For sufficiently smooth functions, this method has a quadratic convergence. However, this approach is very expensive as it needs the explicit derivative evaluation.

A. Standard Broyden's method

In many situations, the quick evaluation of the inverse Jacobian is not possible. In such cases one can often achieve superlinear convergence by using a multidimensional generalization of the secant method, whereby an approximate matrix $\mathbf{B}^{(m)}$ is computed using a secant approximation based on the information obtained in previous iterations. This method is usually referred to as the quasi-Newton method. The difficulty in higher dimensions is that the secant update is not unique, and naïve implementations fail to converge. Broyden [7] suggested using the Sherman-Morrison formula to form an update to the inverse of the Jacobian that has good convergence properties:

$$\mathbf{B}^{(m+1)} = \mathbf{B}^{(m)} + \frac{|\delta \boldsymbol{V}_{\text{in}}\rangle - \mathbf{B}^{(m)}|\delta \boldsymbol{F}\rangle}{\langle \delta \boldsymbol{V}_{\text{in}}|\mathbf{B}^{(m)}|\delta \boldsymbol{F}\rangle} \langle \delta \boldsymbol{V}_{\text{in}}|\mathbf{B}^{(m)}, \tag{5}$$

where

$$|\delta \boldsymbol{V}_{\rm in}\rangle = \boldsymbol{V}_{\rm in}^{(m+1)} - \boldsymbol{V}_{\rm in}^{(m)},$$
$$|\delta \boldsymbol{F}\rangle = \boldsymbol{F}^{(m+1)} - \boldsymbol{F}^{(m)}.$$

One can start the iterative process with the initial guess $\mathbf{B}^{(0)} = \alpha \mathbf{1}$, as this is equivalent to the simple mixing (3). At any given step, one may also check the full quasi-Newton step (4): if it produces larger residuals $|\boldsymbol{F}^{(m+1)}|$ or increases the energy of the solution, then one might revert to simple mixing (3). This can be especially important if the function $\boldsymbol{F}(\boldsymbol{V})$ varies significantly with the step size because the secant approximation may then be extremely poor. More sophisticated, but globally convergent algorithms, are discussed in Ref. [8].

B. Modified Broyden's method

If the self-consistent equations involve many variables, i.e., N is large, storing the $N \times N$ matrix elements of the inverse Jacobian and performing the $N \times N$ matrix multiplications might be prohibitive; hence, further improvements are needed. To deal with the size issues and to improve performance, Broyden's method has been modified [9–11]. The modification by Johnson [10] is now widely used in quantum chemistry. This *modified Broyden's method* only utilizes information obtained in M previous iterations, and this information is used in the update of the approximate inverse Jacobian. We recall here the final expressions of this modified Broyden mixing procedure (a detailed derivation is given in Ref. [10]):

$$\boldsymbol{V}_{\rm in}^{(m+1)} = \boldsymbol{V}_{\rm in}^{(m)} + \alpha \boldsymbol{F}^{(m)} - \sum_{n=\tilde{m}}^{m-1} w_n \gamma_{mn} \mathbf{u}^{(n)}, \tag{6}$$

with $\tilde{m} = \max(1, m - M)$ and

$$\gamma_{mn} = \sum_{k=\tilde{m}}^{m-1} c_k^m \beta_{kn}, \quad \beta_{kn} = \left(w_0^2 \boldsymbol{I} + \mathbf{a}\right)^{-1}_{kn},$$
$$c_k^m = w_k [\Delta \boldsymbol{F}^{(k)}]^\dagger \boldsymbol{F}^{(m)}, \tag{7}$$
$$\mathbf{a}_{kn} = w_k w_n (\Delta F^{(n)})^\dagger \Delta F^{(k)},$$

where

$$\mathbf{u}^{(n)} = \alpha \Delta \boldsymbol{F}^{(n)} + \Delta \boldsymbol{V}^{(n)}$$
$$\Delta \boldsymbol{V}^{(n)} = \frac{\boldsymbol{V}_{\rm in}^{(n+1)} - \boldsymbol{V}_{\rm in}^{(n)}}{|\boldsymbol{F}^{(n+1)} - \boldsymbol{F}^{(n)}|}, \tag{8}$$
$$\Delta \boldsymbol{F}^{(n)} = \frac{\boldsymbol{F}^{(n+1)} - \boldsymbol{F}^{(n)}}{|\boldsymbol{F}^{(n+1)} - \boldsymbol{F}^{(n)}|}.$$

The weights $w_n (n = 1, \ldots, M)$ are associated with each previous iteration and the values $w_n = 1$ are usually chosen. The weight w_0 is assigned to the error in the inverse Jacobian and the value $w_0 = 0.01$ proposed in Ref. [10] gives stable results. The first two terms in Eq. (6) are simply the linear mixing of Eq. (3) with a mixing parameter α, whereas the last term is an additional correction. As mentioned in Ref. [10], the parameter α can be chosen to be rather large (in the examples of this article we have used $\alpha = 0.7$) compared to the usual values used in the simple mixing approach.

Equations (6)–(8) are all that is required to update $\boldsymbol{V}$. In the modified Broyden's method, large $N \times N$ matrices do not appear. It is only one $M \times M$ matrix $\mathbf{a}_{kl}$ and M vectors of length N that need to be stored. The iterative procedure starts with an initial input guess $\boldsymbol{V}_{\rm in}^{(0)}$ that generates the first output solution $\boldsymbol{V}_{\rm out}^{(0)}$. In the next iteration, $m = 1$, $\boldsymbol{V}_{\rm in}^{(1)}$ is a linear combination of $\boldsymbol{V}_{\rm in}^{(0)}$ and $\boldsymbol{V}_{\rm out}^{(0)}$ and the correction term is zero. For $m = 2$, the new $\boldsymbol{V}_{\rm in}^{(2)}$ contains the Broyden correction, including the information obtained in the previous ($m = 1$) iterations. The process is repeated until convergence is reached.

Our actual implementation of the modified Broyden's method is based on a modified subroutine from the package PWSCF (plane-wave self-consistent field) for electronic-structure DFT calculations using pseudopotentials and a plane-wave basis set [12]. The code is published under the GNU General Public License. The Broyden mixing routine in PWSCF is a simple module of about 20 statements that uses the BLAS and LAPACK libraries.

III. SPHERICAL DFT THEORY WITH MEAN-FIELD MIXING: TWO-COMPONENT UNITARY FERMI GAS

As a first application of the standard Broyden's method in nuclear structure calculations, we describe the numerical solution of a two-component superfluid in a spherically symmetric harmonic trap using the superfluid local density approximation (SLDA) of DFT. The system is described by a three-parameter local energy density functional [13] adjusted to results of Monte Carlo calculations for homogeneous matter. The functional depends on the three local densities: $\rho(\boldsymbol{r})$ (particle density), $\tau(\boldsymbol{r})$ (kinetic energy density), and $\kappa(\boldsymbol{r})$ (pairing tensor), which are constructed from the usual Bogoliubov two-component quasiparticle wave functions $[U_k(\boldsymbol{r}), V_k(\boldsymbol{r})]$:

$$\rho_\uparrow(\boldsymbol{r}) = \sum_k |U_k|^2 f_T(E_k), \quad \rho_\downarrow(\boldsymbol{r}) = \sum_k |V_k|^2 f_T(-E_k),$$
$$\tau_\uparrow(\boldsymbol{r}) = \sum_k |\nabla U_k|^2 f_T(E_k), \quad \tau_\downarrow(\boldsymbol{r}) = \sum_k |\nabla V_k|^2 f_T(-E_k),$$
$$\kappa(\boldsymbol{r}) = \sum_k U_k V_k^* \frac{f_T(-E_k) - f_T(E_k)}{2}, \tag{9}$$

where k runs over the quasiparticle states, $\rho = \rho_\uparrow + \rho_\downarrow$, and the function $f_T(E) = [1 + \exp(E/T)]^{-1}$ is the thermal distribution function for fermions. The resulting Kohn-Sham equations can be written in the usual HFB form,

$$\mathcal{H}[h(\boldsymbol{r}), \Delta(\boldsymbol{r})] \begin{bmatrix} V_k \boldsymbol{r}) \\ U_k(\boldsymbol{r}) \end{bmatrix} = E_k \begin{bmatrix} V_k(\boldsymbol{r}) \\ U_k(\boldsymbol{r}) \end{bmatrix}, \tag{10}$$

with the HFB matrix given by

$$\mathcal{H} = \begin{bmatrix} h(\boldsymbol{r}) - \mu_\uparrow & \Delta(\boldsymbol{r}) \\ \Delta^\dagger(\boldsymbol{r}) & -h(\boldsymbol{r}) + \mu_\downarrow \end{bmatrix}, \tag{11}$$

where the particle-hole potential $\Gamma(\boldsymbol{r})$, which enters the particle-hole Hamiltonian

$$h(\boldsymbol{r}) = -\frac{\hbar^2}{2m^*}\nabla^2 + \Gamma(\boldsymbol{r}), \tag{12}$$

and the pairing potential $\Delta(\boldsymbol{r})$ are both functions of the local densities. (The value of the effective mass m^* in Eq. (12) has been determined in Ref. [13].) The introduction of two chemical potentials $\mu_\uparrow$ and $\mu_\downarrow$ eschews the need to use a different formalism (Pauli blocking) for dealing with odd systems [13,14].

The spherical HFB equations (10) are solved by discretizing $\mathcal{H}$ in the radial coordinate r using the so-called discrete variable representation DVR basis [15–17]. We use two sets of abscissa – $\{r_{0,n}\}(n = 1, \ldots, N_r)$ for even-ℓ and $\{r_{1,n}\}$ for odd-ℓ partial waves. Thus, to represent the functions $\Delta(r)$, $\Gamma(r)$, etc., we need only store $2N_r$ sets of data.

The chemical potentials for the two species ($\mu_\uparrow$ and $\mu_\downarrow$) can be viewed as the Lagrange multipliers that enter the grand thermodynamic potential $\mathcal{E} - \mu_\uparrow\rho_\uparrow - \mu_\downarrow\rho_\downarrow$. In the asymmetric situation, $\delta\mu = \mu_\uparrow - \mu_\downarrow \neq 0$, the spectra for the species are shifted so that the thermal distribution functions $f_T(E)$ appearing in Eq. (9) results in different numbers of each species. At zero temperature this produces a quantized step as $\delta\mu$ approaches the pairing gap. By introducing a small temperature, however, this discontinuity can be smoothed without noticeably affecting the physics. This smoothing is important for Broyden's method to achieve convergence.

Instead of treating the chemical potentials as external parameters and adjusting them separately during the iteration process, we have found that they may be adjusted *during the iteration* in such a way that the process converges to a solution with the desired particle number constraints. After some experimenting, we have found that to speed up the convergence the chemical potentials should be updated after every iteration according to a simple ansatz:

$$\mu_\sigma \mapsto \mu_\sigma + \eta\frac{N_\sigma^0 - N_\sigma}{N_+^0}\mu_{\mathrm{TF}}\big(N_\sigma^0\big), \tag{13}$$

where $\sigma = \downarrow$ or $\uparrow$, N_σ^0 is the desired particle number, $\mu_{\mathrm{TF}}(N)$ is some typical positive scale for the chemical potentials, and N_+^0 is some typical positive scale for the particle numbers (such as the total desired particle number). The parameter η can be adjusted to ensure convergence. Note that the computed particle number N_σ enters in exactly one place so that the update is zero if and only if $N_\sigma = N_\sigma^0$. Thermodynamics ensures that the chemical potential is adjusted in the appropriate direction, so as long as the scales in Eq. (13) are appropriately chosen the process converges nicely. The only problems we have encountered are (i) if the updated steps are too large, the system might oscillate wildly. This is easily compensated for by choosing a smaller value of η; (ii) if the temperature is too low, the resulting function may be discontinuous, in which case the update (13) may cause an oscillation. Introducing a small, but finite, temperature (smaller than any relevant scale) cures this problem.

The Broyden vector $\boldsymbol{V}^{(m)}$ in our SLDA calculations contains the values of $\Gamma(r)$ and $\Delta(r)$ on two DVR grids, and two chemical potentials:

$$\boldsymbol{V} \equiv \{\Gamma(r_{0,n}), \Gamma(r_{1,n}), \Delta(r_{0,n}), \Delta(r_{1,n}), \mu_\downarrow, \mu_\uparrow\}. \tag{14}$$

Thus, for a basis of size N_r, the size of the Broyden vector is $N = 4N_r + 2$.

We found that to ensure consistency in the above scheme, *all* parameters used in the iteration should appear in the Broyden process. Thus, even though our description in terms of the DVR basis is somewhat redundant with the function evaluated at two sets of abscissa $\{r_{0,n}\}$ and $\{r_{1,n}\}$, one must keep track of all the information describing the Jacobian. Also, one must be extremely careful to include any external parameters that are adjusted during the iteration in the Jacobian, as well as any internal parameters that are allowed to relax but that might add hysteresis to the system. The output of a single iteration must be a smooth and uniquely defined function of only the input vector $\boldsymbol{V}_{\mathrm{in}}^{(m)}$.

Figure 1 displays some typical results. Here we solve the problem of 60, 59, and 45 particles in a harmonic trap. We consider the cases with equal numbers $(N_\uparrow, N_\downarrow) = (30, 30)$, one particle extra $(N_\uparrow, N_\downarrow) = (30, 29)$, and

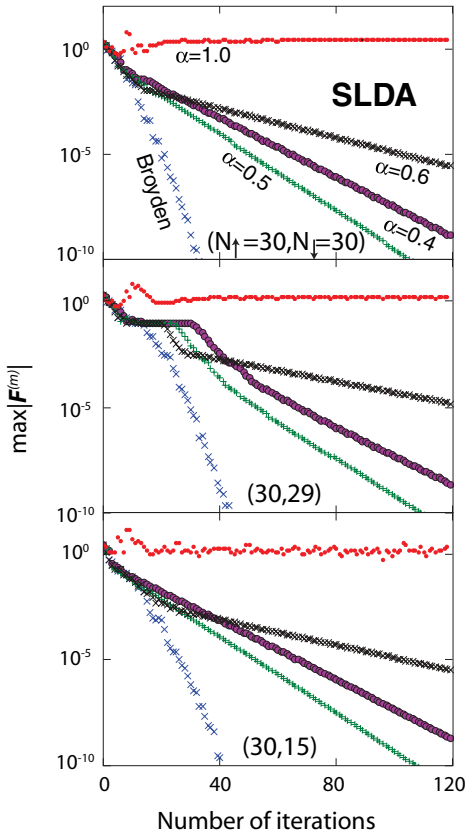

FIG. 1. (Color online) Comparison between the linear mixing algorithm and the standard Broyden's method in SLDA for systems of $(N_\uparrow, N_\downarrow) = (30, 30)$ (top), $(N_\uparrow, N_\downarrow) = (30, 29)$ (middle), and $(N_\uparrow, N_\downarrow) = (30, 15)$ (bottom). The basis size is $N_r = 40$ and the convergence parameter of Eq. (13) is $\eta = 1$. For the asymmetric systems, the equations exhibit a discontinuity responsible for small plateaus in the plots. Introducing a small temperature $T = 0.02\epsilon_F$ restores the continuity. Once the approximate solution is found, the temperature may be set back to zero.

half-paired $(N_\uparrow, N_\downarrow) = (30, 15)$. The size of the Broyden vector is $N = 162$. We see that in all cases, Broyden's method performs substantially better than the linear mixing method.

Note that there are many cases in which the full-step algorithm ($\alpha = 1$) does not converge. This is typically because the system exhibits some form of oscillation and the full step overshoots each time. In our problem this is due to the chemical potential update and indicates that we should probably choose a smaller factor η in Eq. (13). Broyden's method, however, automatically detects this and compensates accordingly, thus achieving good convergence even if η is not properly chosen.

In the odd-particle case (30, 29), a plateau appears where the convergence improves very slowly. This marks the point at which the chemical potentials adjust so as to split one state off from the Fermi sea. Prior to this, the large-scale structure of the initial state is corrected, but once this occurs, the chemical potentials must be finely tuned so that exactly one extra state is occupied. Broyden's method has the same problem because the approximate gradient changes rapidly. To deal with this, a small temperature is assumed. This makes the functions smooth and facilitates convergence. Once the steps are small enough that the energies do not cross zero, the oscillations terminate and convergence is resumed. If problems persist, the strategy is to start with a fairly large temperature to isolate the approximate values of the chemical potentials (only achieving rough convergence) and then iterate with these values at much lower temperatures to reach the desired tolerance.

IV. SYMMETRY-UNCONSTRAINED HFB THEORY WITH MEAN-FIELD MIXING

In this section we discuss the adaptation and performance of the modified Broyden's method in self-consistent nuclear structure calculations with the HFODD code of Refs. [18,19]. The code solves the Skyrme-HFB problem by expanding the solutions on the anisotropic Cartesian harmonic oscillator (HO) basis. HFODD does not have any self-consistent symmetry built in; hence, all deformation degrees of freedom can be taken into account. As the time-reversal symmetry can be broken in HFODD, the code can be used to describe nuclear rotation using the cranking approximation.

The building blocks of the Skyrme-HFB approach are the particle density $\rho(\boldsymbol{r}, \boldsymbol{r}')$ and the pairing density $\tilde{\rho}(\boldsymbol{r}, \boldsymbol{r}')$. The latter is related to the pairing tensor κ of Sec. III by an antiunitary transformation (see, e.g., Ref. [20]).

From ρ and $\tilde{\rho}$, other local densities and currents are constructed by considering spin and isospin degrees of freedom and by taking derivatives up to the second order [6]. By coupling the local densities and currents, one constructs the energy density $\mathcal{E}(\boldsymbol{r})$ that is scalar, time-even, and isoscalar. In its current version, HFODD does not consider explicit proton-neutron mixing; hence, the HFB equations are solved for protons and neutrons separately. A variation of the total energy with respect to the single-particle wave functions yields the mean field h and pairing field $\tilde{h}$ that enter the HFB Hamiltonian matrix:

$$\mathcal{H} = \begin{bmatrix} h(\boldsymbol{r}) - \lambda & \tilde{h}(\boldsymbol{r}) \\ \tilde{h}(\boldsymbol{r}) & -h(\boldsymbol{r}) + \lambda \end{bmatrix}. \tag{15}$$

The particle-hole mean-field Hamiltonian can be expressed as:

$$h(\boldsymbol{r}) = -\frac{\hbar^2}{2m}\Delta + \Gamma_0^{\text{even}} + \Gamma_0^{\text{odd}} + \Gamma_1^{\text{even}} + \Gamma_1^{\text{odd}} + U^{\text{Coul}}, \tag{16}$$

where the subscript $t = 0$ (1) denotes isoscalar (isovector) fields, whereas the superscript even (odd) labels time-even (time-odd) fields.

The pairing field $\tilde{h}$ in HFODD is related to the usual pairing field Δ in Eq. (11) by the same antiunitary transformation that relates $\tilde{\rho}$ and κ. In our HFB calculations, we employed the density-dependent δ interaction in the particle-particle channel. The corresponding particle-particle mean-field Hamiltonian reads:

$$\tilde{h}(\boldsymbol{r}) = \frac{1}{2}V_0\left[1 - V_1\frac{\rho(\boldsymbol{r})}{\rho_0}\right]\tilde{\rho}(\boldsymbol{r}), \tag{17}$$

where $\rho_0 = 0.16$ fm^{-3}. In this article we adopted the value of $V_1 = 0.5$ corresponding to the so-called mixed pairing (see Ref. [21] and references therein). Due to the zero-range of the pairing interaction, a cut-off energy of 60 MeV is used when summing up the contributions of the HFB quasiparticle states to the density matrices. For a given cut-off, the pairing strength V_0 is determined such as to reproduce the experimental neutron pairing gap in ^{120}Sn, $\Delta_n = 1.245$ MeV. As discussed in Ref. [22], this renormalization procedure gives very similar results as the pairing regularization method [23].

The mean-field Hamiltonian (16) depends on the following self-consistent potentials built of the local densities and currents:

$$\begin{aligned} \Gamma_t^{\text{even}}(\boldsymbol{r}) &= -\boldsymbol{\nabla}\cdot[M_t(\boldsymbol{r})\boldsymbol{\nabla}] + U_t(\boldsymbol{r}) \\ &\quad + \frac{1}{2i}[\overleftrightarrow{\nabla}\sigma\cdot\overleftrightarrow{B}_t(\boldsymbol{r}) + \overleftrightarrow{B}_t(\boldsymbol{r})\cdot\overleftrightarrow{\nabla}\sigma] \\ \Gamma_t^{\text{odd}}(\boldsymbol{r}) &= -\boldsymbol{\nabla}\cdot[(\boldsymbol{\sigma}\cdot\boldsymbol{C}(\boldsymbol{r})\boldsymbol{\nabla})] + \boldsymbol{\sigma}\cdot\boldsymbol{\Sigma}_t(\boldsymbol{r}) \\ &\quad + \frac{1}{2i}[\boldsymbol{\nabla}\sigma\cdot\boldsymbol{I}_t(\boldsymbol{r}) + \boldsymbol{I}_t(\boldsymbol{r})\cdot\boldsymbol{\nabla}]. \end{aligned} \tag{18}$$

The potentials $U_t(\boldsymbol{r})$ and $M_t(\boldsymbol{r})$ are real scalar fields (1 component), $\boldsymbol{\Sigma}_t(\boldsymbol{r})$, $\boldsymbol{C}_t(\boldsymbol{r})$, and $\boldsymbol{I}_t(\boldsymbol{r})$ are real vector fields (three components), and $\overleftrightarrow{B}_t(\boldsymbol{r})$ is a real second-rank tensor field (nine components). The pairing potential $\tilde{h}_t(\boldsymbol{r})$ is a two-component complex scalar field.

The HO eigenfunctions are Hermite polynomials weighted by a Gaussian factor. Consequently, all integrations in HFODD are carried out by using the Gauss-Hermite quadrature [18]. The mesh of the Gauss-Hermite nodes defines the discretized grid, and it is sufficient to define mean-field potentials on this grid to compute the HFB matrix.

The Broyden vector $\boldsymbol{V}^{(m)}$ in HFODD contains the values of the 44 self-consistent field components on the Gauss-Hermite grid $\boldsymbol{r}_i$:

$$\begin{aligned} \boldsymbol{V} \equiv \{&U_t(\boldsymbol{r}_i), M_t(\boldsymbol{r}_i), \boldsymbol{\Sigma}_t(\boldsymbol{r}_i), \boldsymbol{C}_t(\boldsymbol{r}_i), \boldsymbol{I}_t(\boldsymbol{r}_i), \\ &\overleftrightarrow{B}_t(\boldsymbol{r}_i), \Re[\tilde{h}_t(\boldsymbol{r}_i)], \Im[\tilde{h}_t(\boldsymbol{r}_i)]\}. \end{aligned} \tag{19}$$

The advantage of dealing with the fields is that one does not have to worry about the unitarity condition for the generalized density matrix.

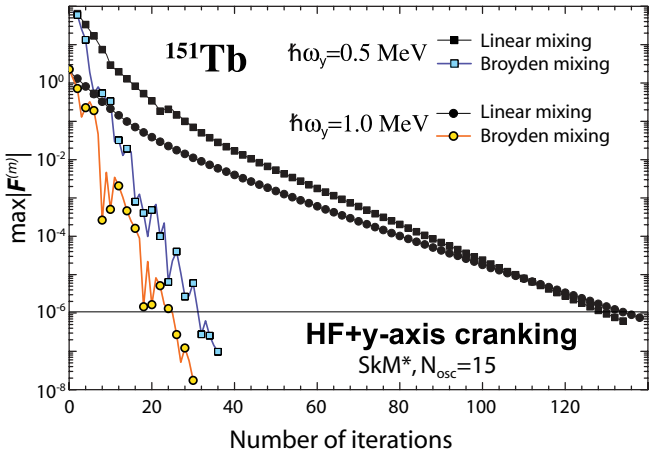

FIG. 2. (Color online) Comparison between linear mixing ($\alpha = 0.5$, black symbols) and the modified Broyden's method ($\alpha = 0.7$, $M = 7$) in HFODD for a high-spin superdeformed band in ^{151}Tb at two values of rotational frequencies $\hbar\omega_y = 0.5$ and 1 MeV.

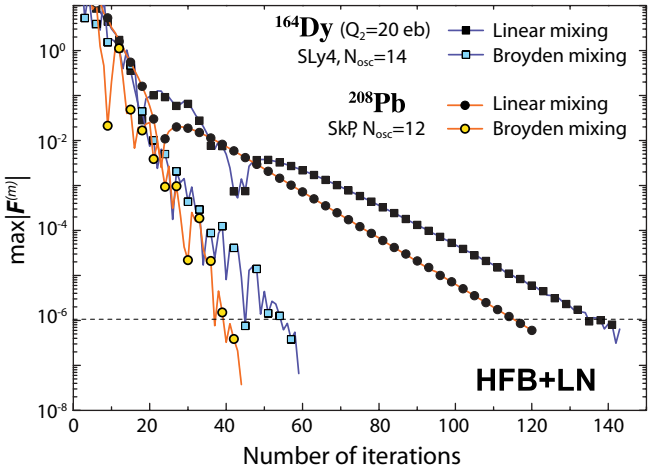

FIG. 3. (Color online) Comparison between linear mixing ($\alpha = 0.5$, black symbols) and Broyden's method ($\alpha = 0.7$, $M = 7$) in HFODD for the superdeformed minimum in ^{152}Dy at the quadrupole moment of $Q_2 = 20$ eb (SLy4 Skyrme functional, $N_{\rm osc} = 14$) and the spherical ground state of ^{208}Pb (SkP Skyrme functional, $N_{\rm osc} = 12$).

If N_x, N_y, and N_z are the numbers of Gauss-Hermite nodes in the x, y, and z directions, respectively, the total size of the Broyden vector in HFODD is $N = (N_x N_y N_z) \times 44$. For a spherical HO basis with $N_{\rm osc} = 12$, the number of Gauss-Hermite points required to perform exact integrations is $N_x = N_y = N_z = 26$, and the size of the Broyden vector becomes $N = 773,344$. In double precision floating point arithmetics, this represents approximately 6.2 MB of memory. For $M = 8$ vectors retained in the Broyden history, this represents a 100 MB memory overhead. Calculations involving larger bases are often required for studies of, e.g., fission pathways, and it is not unusual to deal with $N_{\rm osc} = 20$ for which the size of the Broyden vector becomes as large as $N = 3,259,872$, corresponding to ∼416 MB of additional memory.

Figure 2 illustrates the use of Broyden's method for a cranked Hartree-Fock calculation (without pairing) for the lowest high-spin superdeformed band of ^{151}Tb. The SKM* Skyrme functional [24] has been used. The rotation is generated by adding to the mean-field Hamiltonian (16) the cranking term $-\omega_y \hat{J}_y$, with ω_y being the rotational frequency. As the time-reversal symmetry is broken in HF, both time-even and time-odd fields appear. The calculation at $\hbar\omega_y = 0.5$ MeV was initialized with the unconstrained HF results for the superdeformed vacuum configuration of ^{151}Tb (with all the lowest single-particle HF Routhians occupied) at $\omega_y = 0$. The calculations at $\hbar\omega_y = 1.0$ MeV were initialized with the converged results at $\hbar\omega_y = 0.5$ MeV. The same stretched HO basis containing $N_{\rm osc} = 15$ shells and a deformation of $\beta_2 = 0.61$ and $\beta_4 = 0.1$ was used in both cases.

The advantage of the modified Broyden's method over the simple mixing is evident. Although in the standard calculations the required tolerance of 10^{-6} is achieved after $m \sim 30$ iterations, using the linear mixing, the same outcome is achieved after 130 steps.

Figure 3 illustrates the constrained HFODD calculations with pairing for the superdeformed minimum in ^{152}Dy at the quadrupole moment of $Q_2 = 20$ eb (using the SLy4 Skyrme functional [25]) and also for the spherical ground state of ^{208}Pb (using the SkP Skyrme functional [26]). Both calculations were carried out with mixed pairing and the Lipkin-Nogami (LN) particle-number corrections (see, e.g., Refs. [27,28]). To obtain stability of the results, the proton and neutron λ_2 parameters of the LN model had to be incorporated into the Broyden vector. Again, as in Fig. 2, the modified Broyden's method provides an impressive improvement over the linear mixing scheme.

Before concluding this section, we would like to make an important technical remark. In Cartesian coordinates, the HO wave functions are products of the Hermite polynomials and the Gaussians. In the code HFODD the Gaussians do not have to be calculated explicitly, as they are already incorporated into the Gauss-Hermite quadrature weights [18]. However, this implies that the mean fields used in HFODD are Gaussian scaled. Consequently, they can take nonzero (and possibly large) values at some points of the Gauss-Hermite grid where the physical potentials practically vanish. Using the scaled potentials as inputs to Broyden's method can be a source of very significant numerical instability. Indeed if the scaled potentials (or densities) are used, numerical and physically insignificant fluctuations of the real potentials are artificially enhanced, thereby introducing considerable (and artificial) numerical noise. This instability is not present when the physical mean-field potentials are used.

V. SYMMETRY-CONSTRAINED HFB THEORY WITH HAMILTONIAN MATRIX MIXING

In this section we discuss the performance of the modified Broyden's method as applied to the axial HFB problem. To this end, we employ the HFBTHO code [28,29] that solves the Skyrme-HFB equations by expanding the eigenstates in the cylindrical transformed harmonic oscillator (THO) basis. In the examples shown in this section, we use a SLy4 Skyrme energy density functional augmented by mixed pairing (17).

The HFB Hamiltonian matrix (15) is represented in a THO basis.

The Broyden vector $\boldsymbol{V}^{(m)}$ in HFBTHO contains the matrix elements of the self-consistent fields h and $\tilde{h}$ for neutrons and protons:

$$\boldsymbol{V} \equiv \left\{h^n_{ij}, h^p_{ij}, \tilde{h}^n_{ij}, \tilde{h}^p_{ij}\right\}. \tag{20}$$

Because h and $\tilde{h}$ are Hermitian, $i \leqslant j$. The chemical potentials for neutrons and protons are adjusted at each step m to the correct number of particles. The matrix formulation (20) makes it easy to efficiently incorporate nonlocal terms that appear, e.g., in the particle-number restoration schemes. Some results of our Broyden HFBTHO calculations have been reported in Ref. [30].

Because HFBTHO strictly preserves axial symmetry, parity, and time reversal, the number of independent matrix elements is significantly reduced as compared to the full symmetry-unconstrained case of Sec. IV. For instance, for the basis space of $N_{\rm osc} = 20$ principal oscillator shells, the length of the Broyden vector is $N = 261,228$. For the modified Broyden's method with $M = 8$, the required memory increase in HFBTHO as compared to the simple linear mixing algorithm is usually less than 33 MB, i.e., is practically negligible.

In the implementation of the linear mixing procedure in HFBTHO, the mixing parameter α varies. Initially its value is set at $\alpha = 0.1$. In the next iteration, if the largest element of $\boldsymbol{F}^{(m)}$, $\max|\boldsymbol{F}^{(m)}|$, is less than $\max|\boldsymbol{F}^{(m-1)}|$, α is multiplied by a factor of 1.13 until it reaches the maximum value of $\alpha = 1.0$. When the condition $\max|\boldsymbol{F}^{(m)}| > \max|\boldsymbol{F}^{(m-1)}|$ is met, α is returned to its initial value of 0.1, and the process starts all over again. Such a technique has yielded stable results for all cases considered.

Figure 4 compares linear mixing and Broyden's method for the spherical nucleus ^{120}Sn, which is often used to determine the pairing strength of the functional. Both calculations are initiated from the same Woods-Saxon fields [29]. Usually convergence of all physical quantities of interest is achieved when $\max|\boldsymbol{F}^{(m)}|$ is less than 10^{-5}. For the nucleus ^{120}Sn, which represents a typical case, Broyden's method increases the convergence rate by a factor of $\sim$3, as compared to the linear mixing. Although the absolute convergence rate does depend on an individual nuclear configuration, Broyden's method has proved faster than the linear mixing in all the cases investigated. In particular, for some pathological configurations the improvement is staggering. Such an example is shown in Fig. 4 (bottom) for the prolate configuration in ^{194}Rn.

Here, the efficiency of Broyden's algorithm is almost a hundred times better than the linear mixing. An additional 30% gain can be made by increasing the number of vectors retained in the Broyden history from $M = 3$ to $M = 7$.

Broyden's method proves especially helpful for constrained calculations in which the total energy is calculated as a function of some collective variables (usually expectation values of selected one-body operators). A typical example is shown in Fig. 5 that displays the quadrupole deformation energy curve for ^{212}Ra. The quadrupole deformation β is related to the total

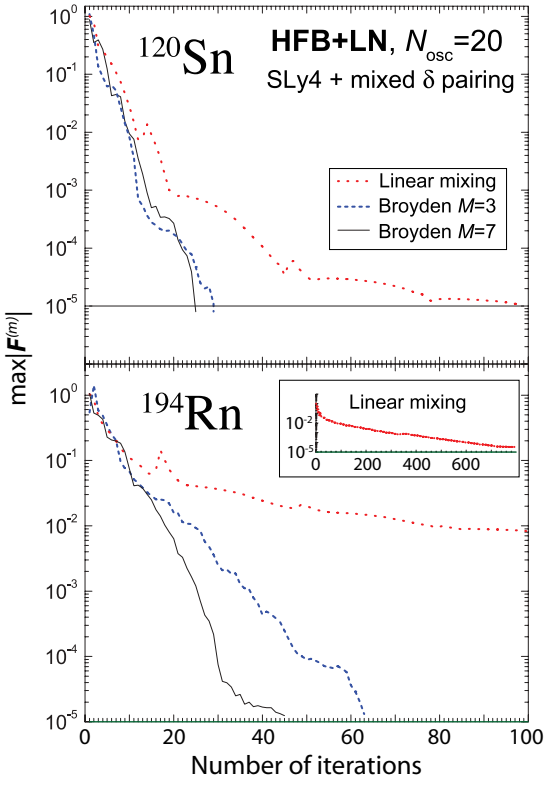

FIG. 4. (Color online) Comparison between linear mixing (dotted line) and Broyden's method in HFBTHO for ^{120}Sn (top) and ^{194}Rn (bottom). The largest element of $\boldsymbol{F}^{(m)}$ is shown as a function of m. The modified Broyden mixing was carried out with $\alpha = 0.7$, $w_0 = 0.01$, and $w_n = 1$ for $M = 3$ (dashed line) and $M = 7$ (solid line). The inset shows the slow convergence of the linear mixing for the pathological case of ^{194}Rn. Here, the tolerance 10^{-5} is reached after 4,345 iterations.

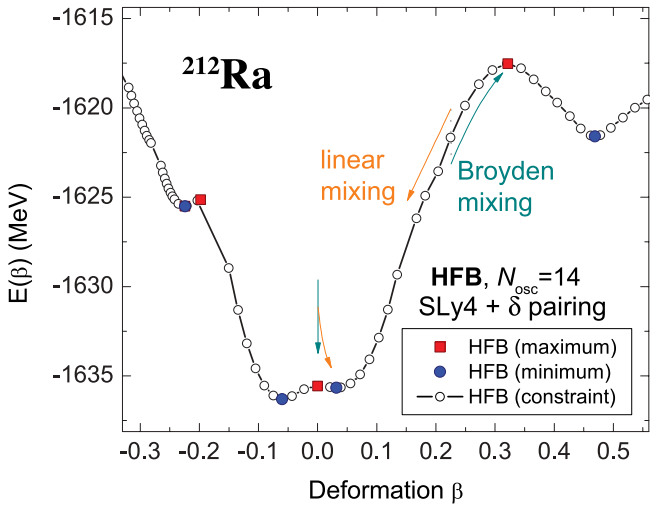

FIG. 5. (Color online) Energy curve of ^{212}Ra versus quadrupole deformation obtained in the constrained HFBTHO calculations. The local minima (maxima) are marked by dots (squares). While in the unconstrained calculations the linear mixing converges at the local minima only, Broyden's method yields *both* local minima and maxima, as indicated by arrows.

quadrupole moment Q_{20} via the usual relation:

$$Q_{20} = \sqrt{\frac{5}{\pi}}\langle r^2\rangle\beta. \tag{21}$$

Starting from the converged solution at the neighboring point of the deformation curve, the calculation converges in no more than 20–30 iterations. The number of iterations is reduced by a factor of 2–3 with respect to the linear mixing used under the same conditions. Since constrained calculations involve a large number of points along the collective path, and in many applications several collective coordinates are involved, such a gain greatly improves the scalability of computations.

A feature of Broyden's method, which is not present when using the linear mixing, is the ability to converge at both minima and maxima of the energy curve, which represent solutions of unconstrained HFB calculations. The extrema in the energy curve of ^{212}Ra are marked in Fig. 5 by black symbols. Starting the unconstrained calculation from a constrained solution at $\beta = 0.25$ and using the linear mixing, one obtains the local minimum at $\beta \approx 0.05$. Under the same conditions, the Broyden mixing yields the maximum at $\beta \approx 0.3$. A similar situation happens when starting with an initial guess at zero deformation: the linear mixing converges to the slightly deformed prolate minimum, whereas the Broyden mixing yields the spherical maximum of the energy curve.

This interesting feature of Broyden's method may be particularly useful for many-body tunneling calculations (e.g., fission), because it allows us to compute the potential barrier height with a much better accuracy than when using the linear mixing in a constrained mode. However, it also implies that a more detailed knowledge of the potential energy surface is required to compute unconstrained minima, because some of the converged solutions may turn out to be local maxima. A possible solution to this problem is to make use of the curvature condition

$$C^{(m)} = \left[\boldsymbol{V}_{\text{out}}^{(m-1)} - \boldsymbol{V}^{(m)}\right]^{\dagger}\mathbf{B}^{(m)}\boldsymbol{F}^{(m)} > 0, \tag{22}$$

which guarantees that the Broyden step is downhill, i.e., toward the minimum. In practical calculations, we accept the Broyden step if Eq. (22) is satisfied. Otherwise, we apply the linear mixing while keeping the current iteration in the Broyden history. Such a procedure bypasses maxima or inflection points and properly converges to the nearest local minimum.

VI. COUPLED-CLUSTER APPLICATIONS

In this section we illustrate the efficiency of the modified Broyden's method in the context of the nuclear CC theory. Coupled-cluster theory originated in nuclear structure; it was pioneered by Coester and Kümmel [31,32] in the early 1960s. In the past two decades, it has been applied mostly in quantum chemistry and today it defines the state-of-the-art many-body theory in this field (see Ref. [33] for a recent overview). Recently, the CC approach has seen a revival in nuclear structure [34]. In Ref. [35] it was applied to the description of loosely bound and unbound helium isotopes; in Ref. [36] it was shown that coupled cluster meets few-body benchmarks and converged results for the ground-state energy of ^{40}Ca were given; in Ref. [37] coupled-cluster theory was extended to three-body Hamiltonians, and the first CC calculations with three-nucleon forces were presented.

Within the CC theory, the nuclear ground-state wave function is written as

$$|\psi\rangle = e^{\hat{T}}|\phi\rangle\,, \tag{23}$$

where $|\phi\rangle = \prod_{i=1}^{A}\hat{a}_i^{\dagger}|0\rangle$ is a single-particle product state expressed in some convenient representation and

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_A \tag{24}$$

is a correlation operator that can be expressed in terms of k-body operators

$$\hat{T}_k = \frac{1}{(k!)^2}\sum_{i_1,\ldots,i_k;a_1,\ldots,a_k} t_{i_1\ldots i_k}^{a_1\ldots a_k}\hat{a}_{a_1}^{\dagger}\ldots\hat{a}_{a_k}^{\dagger}\hat{a}_{i_k}\ldots\hat{a}_{i_1} \tag{25}$$

representing particle-hole (p-h) excitations with respect to the reference state $|\phi\rangle$. In Eq. (25) and in the following, $i, j, k, \ldots$ label hole orbitals, whereas $a, b, c, \ldots$ refer to particle states.

The many-body correlations atop the reference state are introduced through the exponentiated correlation operator. If no truncations are introduced in $\hat{T}$, the theory is exact and the choice of starting reference state is arbitrary. However, in practice, one typically restricts expansion (24) to the two leading terms, $\hat{T} \approx \hat{T}_1 + \hat{T}_2$. Within the resulting coupled-cluster singles and doubles (CCSD) model, the CC equations can be written as

$$E = \langle\phi|\overline{H}|\phi\rangle, \tag{26}$$

$$0 = \left\langle\phi_i^a\right|\overline{H}|\phi\rangle, \tag{27}$$

$$0 = \left\langle\phi_{ij}^{ab}\right|\overline{H}|\phi\rangle. \tag{28}$$

Here $|\phi_{i_1\ldots i_n}^{a_1\ldots a_n}\rangle = \hat{a}_{a_n}^{\dagger}\ldots\hat{a}_{a_1}^{\dagger}\hat{a}_{i_1}\ldots\hat{a}_{i_n}|\phi\rangle$ is a $np-nh$ excitation of the reference state $|\phi\rangle$, and

$$\overline{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}} = (\hat{H}e^{\hat{T}})_c \tag{29}$$

is the similarity-transformed Hamiltonian (note that $\overline{H}$ is non-Hermitian). The last expression on the right-hand side of Eq. (29) indicates that only fully connected diagrams contribute to the construction. This ensures that no unlinked diagrams enter the CC wave function, regardless of the truncation level in $\hat{T}$, and this property makes CC theory size extensive [33].

To determine the ground-state energy E in Eq. (26), the p-h excitation amplitudes t_i^a and t_{ij}^{ab} have to be evaluated through a nonlinear set of equations (27) and (28). By making use of diagram rules and defining intermediates by writing the diagrams in factorized form, the equations for the amplitudes can be written in a quasilinearized form. As an example, Eq. (27) can be written as:

$$\begin{aligned} 0 = {} & f_a^i + I_a'^e t_e^i - \bar{h}_m^i t_a^m - v_{ma}^{ie} t_e^m + \bar{h}_m^e t_{ea}^{mi} \\ & - \tfrac{1}{2}\bar{h}_{mn}^{ie} t_{ae}^{mn} + \tfrac{1}{2} v_{am}^{ef} t_{ef}^{im}. \end{aligned} \tag{30}$$

In Eq. (30), f_a^i is the Fock matrix and $I_a'^e$, $\bar{h}_m^i$, and $\bar{h}_{mn}^{ie}$ are the intermediates [38]. The nonlinearity is hidden in the intermediates that depend on the amplitudes t_i^a and t_{ij}^{ab}.

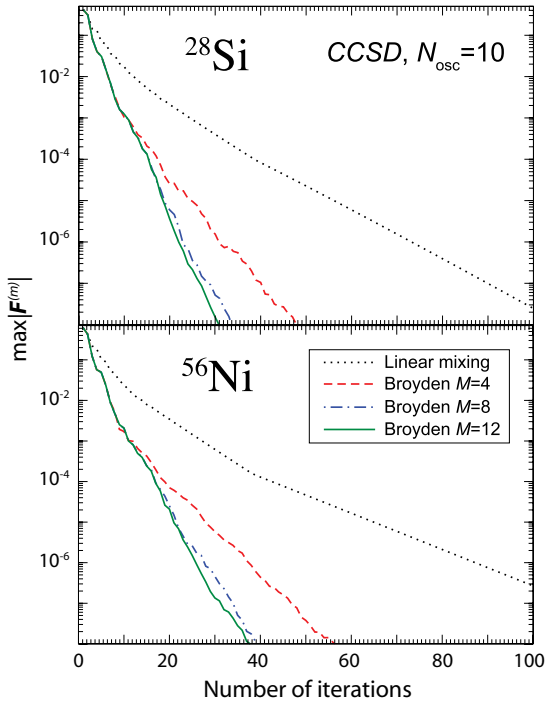

FIG. 6. (Color online) Convergence of CCSD ground-state energy of ^{28}Si (top) and ^{56}Ni (bottom) using the modified Broyden's method with $M = 4$, 8, and 12 compared to the simple mixing with $\alpha = 0.7$. The calculations were carried out in a model space of $N_{\rm osc} = 10$ major oscillator shells.

Typically the number of unknowns is too large to allow for a direct solution of the CC equations; hence, one has to resort to iterative approaches. In the largest calculation we have performed to date [37], we solved Eqs. (27) and (28) for $\sim 10^8$ number of unknowns, on 1000 2-GB parallel processors. Each iterative step costs, at the CCSD level, $\sim n_h^2 n_p^4$ computational cycles (n_h/n_p is the number of hole/particle orbitals).

Considering the number of unknowns and the highly nonlinear nature of CC equations, the use of iterative convergence accelerators, such as the modified Broyden's method, is a must. In our nuclear CC calculations, the Broyden vector is given by the amplitudes:

$$\boldsymbol{V} \equiv \left\{ t_i^a, t_{ij}^{ab} \right\}. \tag{31}$$

Figure 6 compares the performance of the modified Broyden's method with $M = 4$, 8, and 12 to the simple mixing ansatz (3). To solve the CCSD equations for the CCSD ground state energies of ^{28}Si (top) and ^{56}Ni (bottom), we have used an oscillator representation of the Hamiltonian in a model space of $N_{\rm osc} = 10$ major oscillator shells with $\hbar\omega_{\rm osc} = 32$ MeV. The Hamiltonian consists of the kinetic-energy operator (minus the center-of-mass correction) and the N^{3}LO nucleon-nucleon interaction of Ref. [39].

It is seen that utilizing Broyden's method increases the convergence dramatically as compared to the simple mixing. It is found that for $M = 12$, iterations converge more rapidly, and the tolerance of 10^{-8} is reached typically after 30–40 iterations. For the linear mixing to achieve similar accuracy, 130–140 iterations are required.

VII. CONCLUSIONS

Broyden's method, a quasi-Newton method for the numerical solution of nonlinear equations in many variables, is widely used in quantum chemistry to perform electronic-structure calculations. In this work, we applied Broyden's method (both in its standard and modified variants) to several nuclear structure problems. The examples range from several DFT applications in various geometries to *ab initio* CC calculations.

The number of unknown variables (i.e., the size of the Broyden vector) in our calculations ranges from $\sim$200 in SLDA calculations with the standard Broyden's method, to $\sim 300,000$ in HFBTHO, $\sim 3,000,000$ in HFODD, and $\sim 10^8$ in CC applications of the modified Broyden's algorithm. Much faster convergence has been achieved in comparison with the linear-mixing procedure that is often used in such types of calculations. We note that Broyden's method often achieves superlinear convergence, although mildly. If the iteration is not properly structured, then it might not correspond to a contractive mapping and may not converge. By constructing an approximation to the Jacobian, Broyden's method can determine the downhill direction even if the iteration itself does not provide this information.

The downhill direction chosen by Broyden's method is toward a fixed point, not necessarily toward a minimum of the functional. Thus, the iterative process might easily converge to a maximum or a saddle point even when the iteration would naturally head toward a minimum. If the goal is to find the minimum, a combination of the linear mixing and Broyden's method can be efficiently used, as demonstrated in the example of constrained HFB calculations of Fig. 5. However, the above feature of Broyden's method to converge to maxima or saddle points can be taken advantage of if the theoretical objective *is* to determine local maxima or saddle points (e.g., when studying fission pathways).

In summary, Broyden's method is easy to incorporate into fixed-point codes and provides impressive performance improvements in nuclear many-body calculations.

ACKNOWLEDGMENTS

Useful discussions with Jorge Moré and Jason Sarich are gratefully acknowledged. The UNEDF SciDAC Collaboration is supported by the U.S. Department of Energy under grant No. DE-FC02-07ER41457. This work was also supported by the U.S. Department of Energy under Contract Nos. DE-FG02-96ER40963 (University of Tennessee), DE-AC05-00OR22725 with UT-Battelle, LLC (Oak Ridge National Laboratory), DE-FG05-87ER40361 (Joint Institute for Heavy Ion Research), and DE-FG02-97ER41014 (University of Washington), by the National Nuclear Security Administration under the Stewardship Science Academic Alliances program through the U.S. Department of Energy Research Grant DE-FG03-03NA00083, and by the Polish Ministry of Science and Education under Contract N 202 179 31/3920. WN acknowledges support from the Carnegie Trust during his stay in Scotland.

[1] *RIA Theory Bluebook: A Road Map*, http://www.orau.org/ria/RIATG/Blue_Book_FINAL.pdf.

[2] http://unedf.org.

[3] W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).

[4] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Phys. Rev. Lett. **60**, 2430 (1988); N. N. Lathiotakis, M. A. L. Marques, M. Lüders, L. Fast, and E. K. U. Gross, Int. J. Quantum Chem. **99**, 790 (2004); M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross, Phys. Rev. B **72**, 024545 (2005).

[5] M. Bender, P.-H. Heenen, and P.-G. Reinhard, Rev. Mod. Phys. **75**, 121 (2003); *Extended Density Functionals in Nuclear Structure Physics*, edited by G. A. Lalazissis, P. Ring, and D. Vretenar (Springer Verlag, New York, 2004).

[6] E. Perlińska, S. G. Rohoziński, J. Dobaczewski, and W. Nazarewicz, Phys. Rev. C **69**, 014316 (2004).

[7] C. G. Broyden, Math. Comput. **19**, 577 (1965).

[8] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, New York, 1983); W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipies: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, New York, 2007).

[9] D. Vanderbilt and S. G. Louie, Phys. Rev. B **30**, 6118 (1984).

[10] D. D. Johnson, Phys. Rev. B **38**, 12807 (1988).

[11] V. Eyert, J. Comput. Phys. **124**, 271 (1996).

[12] S. Baroni, A. Dal Corso, S. de Gironcoli, P. Giannozzi, C. Cavazzoni, G. Ballabio, S. Scandolo, G. Chiarotti, P. Focher, A. Pasquarello, K. Laasonen, A. Trave, R. Car, N. Marzari, A. Kokalj, http://www.pwscf.org/.

[13] A. Bulgac, Phys. Rev. A **76**, 040502(R) (2007); A. Bulgac and M. M. Forbes, arXiv:0804.3364; A. Bulgac and M. M. Forbes, in preparation (2008).

[14] R. Sensarma, W. Schneider, R. B. Diener, and M. Randeria, arXiv:0706.1741 (2007).

[15] R. G. Littlejohn and M. Cargo, J. Chem. Phys. **117**, 27 (2002).

[16] R. G. Littlejohn, M. Cargo, J. Tucker Carrington, K. A. Mitchell, and B. Poirier, J. Chem. Phys. **116**, 8691 (2002).

[17] D. Baye, Phys. Status Solidi B **243**, 1095 (2006).

[18] J. Dobaczewski and J. Dudek, Comput. Phys. Commun. **102**, 166 (1997).

[19] J. Dobaczewski and J. Dudek, Comput. Phys. Commun. **102**, 183 (1997); **131**, 164 (2000); J. Dobaczewski and P. Olbratowski, Comput. Phys. Commun. **158**, 158 (2004); **167**, 214 (2005).

[20] J. Dobaczewski, W. Nazarewicz, T. R. Werner, J.-F. Berger, C. R. Chinn, and J. Dechargé, Phys. Rev. C **53**, 2809 (1996).

[21] J. Dobaczewski, W. Nazarewicz, and M. V. Stoitsov, Eur. Phys. J. A **15**, 21 (2002).

[22] P. J. Borycki, J. Dobaczewski, W. Nazarewicz, and M. V. Stoitsov Phys. Rev. C **73**, 044319 (2006).

[23] A. Bulgac, Phys. Rev. C **65**, 051305(R) (2002).

[24] J. Bartel, P. Quentin, M. Brack, C. Guet, and H. B. Håkansson, Nucl. Phys. **A386**, 79 (1982).

[25] E. Chabanat, P. Bonche, P. Haensel, J. Meyer, and R. Schaeffer, Nucl. Phys. **A627**, 710 (1997); **A635**, 231 (1998).

[26] J. Dobaczewski, H. Flocard, and J. Treiner, Nucl. Phys. **A422**, 103 (1984).

[27] M. V. Stoitsov, J. Dobaczewski, R. Kirchner, W. Nazarewicz, and J. Terasaki, Phys. Rev. C **76**, 014308 (2007).

[28] M. V. Stoitsov, J. Dobaczewski, W. Nazarewicz, S. Pittel, and D. J. Dean, Phys. Rev. C **68**, 054312 (2003).

[29] M. V. Stoitsov, J. Dobaczewski, W. Nazarewicz, and P. Ring, Comput. Phys. Commun. **167**, 43 (2005).

[30] M. V. Stoitsov, in *Nuclear Theory*, edited by S. Dimitrova, Proceedings of the 26th International Workshop on Nuclear Theory, Rila Mountains, Bulgaria, June 2007 (Institute for Nuclear Research and Nuclear Energy Bulgarian Academy of Sciences, Sofia, 2008), p. 13.

[31] F. Coester, Nucl. Phys. **7**, 421 (1958).

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

[33] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

[34] D. J. Dean, J. R. Gour, G. Hagen, M. Hjorth-Jensen, K. Kowalski, T. Papenbrock, P. Piecuch, and M. Włoch, Nucl. Phys. **A752**, 299 (2005).

[35] G. Hagen, D. J. Dean, M. Hjorth-Jensen, and T. Papenbrock, Phys. Lett. **B656**, 169 (2007).

[36] G. Hagen, T. Papenbrock, D. J. Dean, A. Schwenk, A. Nogga, M. Włoch, and P. Piecuch Phys. Rev. C **76**, 034302 (2007).

[37] G. Hagen, D. J. Dean, M. Hjorth-Jensen, T. Papenbrock, and A. Schwenk, Phys. Rev. C **76**, 044305 (2007).

[38] J. R. Gour, P. Piecuch, M. Hjorth-Jensen, M. Włoch, and D. J. Dean, Phys. Rev. C **74**, 024310 (2006).

[39] D. R. Entem and R. Machleidt, Phys. Lett. **B524**, 93 (2002).