

Broyden Mixing for Nuclear Density Functional Calculations

M.V. Stoitsov

¹ Department of Physics and Astronomy, University of Tennessee,
Knoxville, TN 37996, USA

² Physics Division, Oak Ridge National Laboratory,
P.O.B. 2008, Oak Ridge, TN 37831, USA

³ Joint Institute for Heavy Ion Research,
Oak Ridge, TN 37831, USA

⁴ Institute of Nuclear Research and Nuclear Energy,
Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria

Abstract. The modified Broyden mixing method, which is widely used in quantum chemistry to perform first-principles electronic-structure calculations, is applied to nuclear density functional calculations. A much better convergence rate is achieved which, in some of the cases, is ranged between times to hundreds of times in comparison with the so-called linear mixing procedure used so far. A new feature which is not present when using the linear mixing is that the Broyden mixing gives solutions in all the points where the deformation energy curve has extremum. The stability of the Broyden mixing procedure and the fastest convergence achieved makes it a candidate for a standard tool nowadays in nuclear density functional calculations.

1 Introduction

The development of radioactive beam facilities significantly improves the opportunity to measure nuclei close to nuclear drip lines. A recent example is the observation of nuclei as ^{40}Mg and ^{42}Al , which have relatively large mass number and are almost at the neutron drip line [1]. This increases the demand for theoretical predictions for a broad range of nuclei across the nuclear mass chart.

Nuclear density functional theory (DFT) is one of the most reliable methods for calculating properties of nuclei within the whole nuclear mass chart. DFT is based on the idea that there exists a universal nuclear energy density functional (UNEDF) which represents the total energy of the system as a functional $E[\rho, \tilde{\rho}]$ of the normal ρ and pairing $\tilde{\rho}$ one-body density matrices, defining various local densities and currents. After minimizing the functional $E[\rho, \tilde{\rho}]$, one ends up with a highly nonlinear system of integro-differential equations which are solved in a self-consistent manner.

Initially, attempts to build a UNEDF were rooted in the zero-range Skyrme interaction treated within the Hartree-Fock (HF) or Hartree-Fock-Bugoliubov (HFB) approximation. However, it was realized afterwards that an effective interaction could be secondary to the functional, i.e., it is the UNEDF that defines the force. This is

the strategy followed by modern nuclear DFT applications. Extensive discussions on contemporary achievements and problems under way to build a UNEDF are in [2].

The variety of nuclear deformations and different kinds of states represents a real challenge when doing large-scale mass table calculations. For example, considering even-even nuclei only, one usually has to calculate a couple of thousands of them. Each nucleus can have more solutions. Depending on the deformation, some of the nuclei can have three or more minima - spherical states with zero deformation, prolate states with positive deformation, or oblate states with negative deformation. One should calculate all of them in order to find the solution with the lowest total energy which defines the ground state of the nucleus.

One can calculate thousands of nuclei only using the recently developed parallel supercomputers as, for example, the Cray T4 Jaguar supercomputer at CCS ORNL with more than 40000 processors. In this context, it is important to realize the complexity of the problem when applying DFT to large-scale nuclear mass table calculations.

For some nuclei the deformation energy surface is rather soft. This makes the convergence of the self-consistent procedure rather slow. Approaching the drip lines the interplay between nuclear deformation and the pairing effects leads to some fine cancelations which generate some staggering during the iterative process and also leads to quite lengthy and time-consuming iterations. The same happens when some level crossing exists close to the top of the pairing window used.

A possible solution is to increase the limit of iterations to a very large number. A typical run for the whole even-even mass chart contains about 2737 different bound nuclear states which identify the ground states for all 1527 even-even nuclei entering the mass chart between nuclear drip lines. At the end of the run, one can see that 2032 of the nuclei converge for up to 500 iterations, but there are still 404 nuclei that converge up to 1000 iterations, another 123 nuclei that converge up to 2000 iterations, and even 152 nuclei that converge up to 6000 iterations. At the end, there are 26 nuclei which still do not converge even for the limit of 7000 iterations.

These figures obviously demonstrate one needs better reliable algorithms that are almost always converging in the broad varieties of different nuclei and nuclear states. If not, one is forced to recalculate many of the remaining unsuccessful solutions which are difficult and often are a source of additional errors.

The modified Broyden mixing method [4] is an algorithm with faster convergence which is already a standard tool in atomic and molecular DFT calculations, but it was never attempted in nuclear DFT calculations. The attention to Broyden mixing methods in the context of nuclear density functional calculations has been attracted by a recent investigation of A. Bulgac [5]. The aim of the present paper is to report the encouraging results from a particular implementation of the Broyden mixing method to nuclear DFT calculations.

The paper is organized in the following way. The Broyden mixing procedure is outlined in Sec. 2. Its implementation and numerical results from nuclear density functional calculations are discussed in Sec. 3. Conclusions are given in Sec. 4.

2 Broyden mixing method

When solving HFB equations, one usually starts with given initial potentials $V^{(0)}$ which linearize the HFB system of equations. Its solution gives us an output of nuclear densities which define new potentials $V^{(1)}$ used to substitute $V^{(0)}$ as a new input for the next iteration. This iterative process continues until self-consistency is reached, i.e., until the differences $F^{(m)} = V^{(m)} - V^{(m-1)}$ on two consecutive iterations become negligible. Note, it is equally valid to consider densities or some other quantities in this self-consistent context, as will be discussed later.

Typically, the self-consistent procedure diverges when using straight iterations. The most simple-minded method to avoid this is to use the linear mixing, i.e., as an input on the next iteration to use the linear combination

$$V^{(m+1)} = (1 - \alpha)V^{(m-1)} + \alpha V^{(m)} = V^{(m-1)} + \alpha F^{(m)}, \quad (1)$$

where the superscript m denotes the particular iteration. By a suitable choice of the constant $\alpha \in [0, 1]$, convergence is usually achieved. However, as already mentioned, there are many instances where the convergence proceeds very slowly with an excessive number of iterations.

The linear mixing is usually used in nuclear DFT calculations and the best one can do in order to speed up the convergence is to change the values of α during the iterations, i.e., increasing α when $|F^{(m)}| \leq |F^{(m-1)}|$ or decreasing it otherwise. Obviously a better iterative method is necessary.

The main idea of Broyden mixing methods is that the self-consistent condition $F^{(m)} = 0$ can be viewed as a nonlinear equation which should be satisfied at the end of the self-consistent procedure. The Broyden method, usually referred to as the Newton-Raphson Jacobian update method [3], uses mixing of the type

$$V^{(m+1)} = V^{(m-1)} - (J^{(m)})^{-1} F^{(m)}, \quad (2)$$

where $J^{(m)}$ is the Jacobian matrix between successive iterations starting from a good initial guess. However a shortcoming of the method is the prohibitive storage required in keeping the updated $N \times N$ Jacobian matrix as well as $N \times N$ matrix multiplications required after each iteration. N is the length of the vector V and it can be a very large number (in our case the number of coordinate points of the fields times the number of the fields to be updated).

The Broyden method has been further modified by D.D. Johnson [4], and this modification is now widely used in quantum chemistry to perform first-principles electronic-structure calculations. It incorporates information from previous M iterations, and this information is used in the update of the Jacobian matrix. The final expressions of this modified Broyden mixing procedure we are going to use (one can see a detailed derivation in [4]) reads

$$V^{(m+1)} = V^{(m-1)} + \alpha F^{(m)} - \sum_{n=\tilde{m}}^{m-1} w_n \gamma_{mn} u^{(n)}, \quad (3)$$

with

$$\gamma_{mn} = \sum_{k=\tilde{m}}^{m-1} c_k^m \beta_{kn}, \quad \beta_{kn} = (w_0^2 I + a)_{kn}^{-1}, \quad c_k^m = w_k \langle \Delta F^{(k)} | F^{(m)} \rangle, \quad (4)$$

$$a_{kn} = w_k w_n |\Delta F^{(n)} \rangle \langle \Delta F^{(n)}|, \quad u^{(n)} = \alpha (\Delta F^{(n)} + \Delta V^{(n)}), \quad (5)$$

where

$$\Delta V^{(n)} = \frac{V^{(n+1)} - V^{(n)}}{|F^{(n+1)} - F^{(n)}|}, \quad \Delta F^{(n)} = \frac{F^{(n+1)} - F^{(n)}}{|F^{(n+1)} - F^{(n)}|}. \quad (6)$$

The weights w_n ($n = 1, \dots, M$) are associated with each previous iteration and usually values $w_n = 1$ work well. The weight w_0 is assigned to the error in the inverse Jacobian and the value $w_0 = 0.01$, proposed in [4], is giving stable results. The first two terms in eq. (3) are simply the linear mixing of eq. (1), with a mixing parameter α , while the last term is the correction to this. As mentioned in [4], the parameter α can be chosen to be rather large ($\alpha = 0.7$ in our calculations) compared to that in the linear mixing.

Equations (3) - (6) constitute all information that is required for updating of the vector V . The algorithm needs a storage of one $M \times M$ matrix a_{kl} and M vectors with length N . No storage or multiplications of large $N \times N$ matrices is necessary. One simply starts with an initial guess $V^{(0)}$ which generates the first iteration solution $V^{(1)}$. For the next iteration, $m = 1$ defines $V^{(2)}$ as a linear mixing between $V^{(0)}$ and $V^{(1)}$. The correction term is zero since the lower boundary of the summation starts from $\tilde{m} = \max(1, m - M)$. Then, at $m = 2$, the new $V^{(3)}$ already contains the Broyden mixing, including the information from the previous ($m = 1$) iterations and so on until convergence is achieved.

3 Numerical Results

Present numerical results are obtained as in [6] with a SLY4 Skyrme energy density functional and mixed δ -pairing using a HFBTHO code published in [7] with linear mixing, recently upgraded with the Broyden mixing procedure as described in the previous section.

In the HFBTHO code [7] the mixing is done with respect to all matrix elements calculated in the (transformed) harmonic oscillator (HO) basis. This choice is almost equivalent to mixing the HFB fields since the matrix elements are linear with respect to the HFB fields. The advantage is that one can efficiently mix also non-local terms as they appear when using, for example, the approximate Lipkin-Nogami particle-number projection.

In other words, the vector $V^{(m)}$ includes the matrix elements of the HFB field for neutrons, followed by the same matrix elements except for protons, then by the matrix elements of the pairing field for neutrons and protons. For $N_{sh} = 20$ major oscillator basis states, the length N of vector for updating is $N = 261228$. If one keeps the Broyden history for the previous $M = 8$ iterations, it constitutes an

increase of the memory used by the HFBTHO code of up to about 33 Mb, which is an acceptable number. The numerical overhead is not noticeable compared with the linear mixing.

The actual implementation of the Broyden mixing procedure is based on a subroutine modified from the GNU General Public License computer code PWscf for electronic-structure calculations within Density-Functional Theory and Density-Functional Perturbation Theory using pseudo-potentials and a plane-wave basis set [8]. It is using routines from the standard BLAS and LAPACK libraries only.

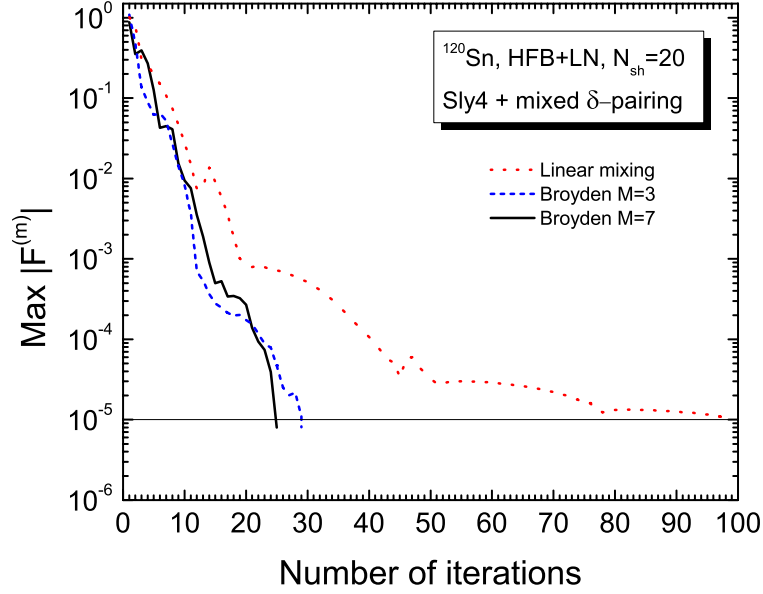


Figure 1. The maximal value of $|F^{(m)}|$ (in logarithmic scale) as a function of the number of iterations m for the nucleus ^{120}Sn . Linear mixing results (dotted line) are compared with Broyden mixing ($\alpha = 0.7$, $w_0 = 0.01$, $w_n = 1$) with $M = 3$ (dashed line) and $M = 7$ (solid line).

In order to be as fair as possible in the comparison, we have to mention the way the linear mixing is implemented in the HFBTHO code. Initially it starts with a very small value of the mixing parameter $\alpha = 0.1$. On the next iteration, if the maximal value of $|F^{(m)}|$ is less than that of $|F^{(m-1)}|$, the parameter α is increased multiplying it by a factor of 1.13 until $\alpha = 1$. Then α stays at $\alpha = 1$ until the above requirement is satisfied. If not, the value is returned to its initial value $\alpha = 0.1$, and the process starts over again. Maybe this is not the most efficient strategy, but it gives stable results for all nuclei through the mass chart.

Fig. 1 shows a comparison between linear and Broyden mixing results for the benchmark spherical nucleus ^{120}Sn , which is usually used to fix the pairing strength of the functional. Both calculations are started from the same initial Woods-Saxon

fields [7]. The maximal value of $|F^{(m)}|$ is plotted (in logarithmic scale) as a function of the number of iterations m . Usually convergence of all physical characteristics is achieved when the maximal value of $|F^{(m)}|$, i.e., the maximal difference between all matrix elements on two consecutive iterations, is less than 10^{-5} .

From Fig. 1, it is seen that for the nucleus ^{120}Sn , the Broyden mixing leads to a faster convergence in about 3-4 times compared with the linear mixing. As one may expect, the convergence rate will depend on the nuclear state calculated. Nevertheless, in all the cases we have calculated, the Broyden mixing was giving faster convergence.

For nuclei where linear mixing works well, the Broyden mixing does not basically change the convergence rate a lot. But in the cases of “difficult” nuclei, the Broyden mixing is doing an amazing job. Such an example is shown in Fig. 2 for the prolate state of the nucleus ^{194}Rn .

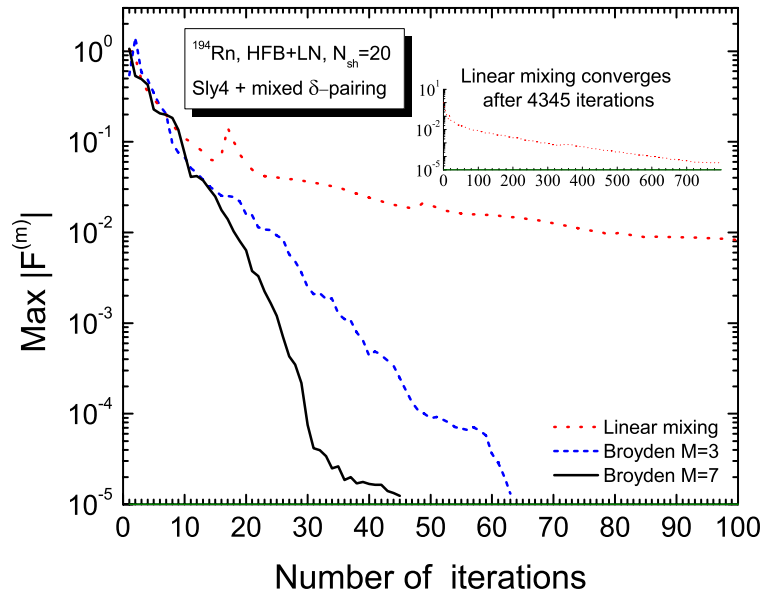


Figure 2. The same as in Fig. 1 except for the nucleus ^{194}Rn . The inset shows the slow convergence of the linear mixing which leads to the same results after 4345 iterations.

From Fig. 2, it is seen that for the nucleus ^{194}Rn the efficiency of the Broyden mixing is almost a hundred times better. One can see also that the use of 7 states in the Broyden history ($M = 7$) improves the convergence (compared with the case $M = 3$) with additional 30%.

We have found the Broyden mixing procedure especially helpful when doing constraint calculations. An example of constraint calculations is shown in Fig. 3 for the nucleus ^{122}Ra . Starting with the solution from the previous point of the deformation curve, the results converge on average for no more than 20-30 iterations.

Typically there is in average a factor of 2-3 comparing with the linear mixing used under the same conditions. This is important because in the constraint calculations one usually needs many points from the deformation energy curve.

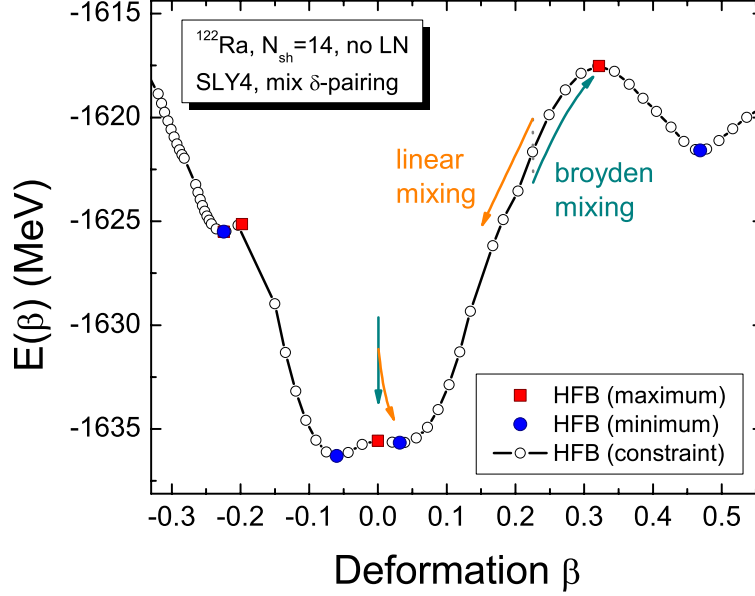


Figure 3. Deformation energy curve for the nucleus ^{122}Ra obtained in constrained DFT calculations. Linear mixing unconstrained solutions are found at the minimal points only (the solid circle symbols), while with Broyden mixing, solutions at the maximal points (the solid square symbols) also exist.

A new feature which is not present when using the linear mixing is illustrated in Fig. 3 as well. Basically, using the linear mixing, the unconstrained solution is usually found at the points where the total energy of the system has a minimum as a function of deformation (the solid circle symbols in Fig. 3). To find one or another minima depends on the initial starting point. In addition to that, however, the Broyden mixing gives solutions in all the points where the deformation energy curve has extremum including the maxima (the solid square symbols in Fig. 3).

For example, starting the unconstrained calculations with an initial solution obtained from constrained calculations with deformation $\beta = 0.25$ and using the linear mixing, one ends up with the solution at the minimum with $\beta \approx 0.05$. Using the Broyden mixing under the same conditions, one ends up with the solution at the maximum with $\beta \approx 0.3$. The same situation occurs when starting with an initial guess with zero deformation, as illustrated Fig. 3.

This new feature of DFT calculations, combined with the Broyden mixing, may be useful in the cases when one has to determine the position of fission barrier, for

example. However, it requires much detailed knowledge of the deformation energy curve in order to locate the energy minima required in the nuclear ground state calculations.

4 Conclusions

The modified Broyden mixing method [4], widely used in quantum chemistry to perform first-principles electronic-structure calculations, has been applied to nuclear density functional calculations. Much faster convergence is achieved in comparison with the linear mixing procedure which is presently in use in such types of calculations.

The convergence rate strongly depends on the nuclear state calculated. For nuclei where linear mixing works well, the Broyden mixing does not basically change the convergence rate. In the cases of nuclei where the linear mixing gives convergence in a thousand iterations, the efficiency of the Broyden mixing is almost a hundred times better. The convergence for about 26 nuclear states through the whole even-even mass chart still remains a problem.

A new feature which is not present when using the linear mixing is that the Broyden mixing gives solutions in all the points where the deformation energy curve has extremum. This requires much detailed knowledge of the deformation energy curve in order to locate the energy minima required in the nuclear ground state calculations.

In conclusion, the stability of the Broyden mixing procedure and the fastest convergence achieved makes it a candidate for a standard tool nowadays in nuclear DFT calculations.

Acknowledgments

This work was supported in part 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), and DE-FG05-87ER40361 (Joint Institute for Heavy Ion Research). The UNEDF SciDAC Collaboration is supported by the U.S. Department of Energy under grant No. DE-FC02-07ER41457.

References

1. T. Baumann, *et al.*, *Nature* **499**, 1022 (2007).
2. M.V. Stoitsov, J. Dobaczewski, W. Nazarewicz, P. Borycki, *Int. J. of Mass Spectrometry* **251**, 243 (2006).
3. C. G. Broyden, *Math. Comput.* **19**, 577 (1963).
4. D. D. Johnson, *Phys. Rev. C* **38**, 12808 (2003).
5. A. Bulgac, *Phys. Rev. A* **76**, 040502(R) (2007).

6. M. V. Stoitsov, J. Dobaczewski, W. Nazarewicz, S. Pittel, and D. J. Dean, *Phys. Rev. C* **68**, 054312 (2003).
7. M.V. Stoitsov, J. Dobaczewski, W. Nazarewicz, P. Ring, *Comp. Phys. Commun.* **167**, 43 (2005).
8. Plane-Wave Self-Consistent Field (PWscf) computer code, <http://www.pwscf.org>