Chemical Physics Letters 647 (2016) 31–35

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Periodic Pulay method for robust and efficient convergence acceleration of self-consistent field iterations

Amartya S. Banerjee[a], Phanish Suryanarayana[b,*], John E. Pask[c]

[a] Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[b] College of Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c] Physics Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

ARTICLE INFO

Article history:
Received 15 November 2015
In final form 13 January 2016
Available online 21 January 2016

ABSTRACT

Pulay's Direct Inversion in the Iterative Subspace (DIIS) method is one of the most widely used mixing schemes for accelerating the self-consistent solution of electronic structure problems. In this work, we propose a simple generalization of DIIS in which Pulay extrapolation is performed at periodic intervals rather than on every self-consistent field iteration, and linear mixing is performed on all other iterations. We demonstrate through numerical tests on a wide variety of materials systems in the framework of density functional theory that the proposed generalization of Pulay's method significantly improves its robustness and efficiency.

1. Introduction

Nonlinear equations are often posed as fixed point problems that lend themselves to solution via self-consistent iterations [1,2]. This is the commonly adopted practice in electronic structure calculations such as those based on density functional theory (DFT) [3,4], where additionally so-called *mixing schemes* are routinely employed to accelerate convergence [5,6]. The simplest of all mixing schemes is *linear mixing*, which is an under-relaxed fixed-point iteration. Although linear mixing can be guaranteed to converge for many systems with the choice of small enough mixing parameter [7], it tends to perform rather poorly in practice. Since the computational cost of electronic structure calculations is directly proportional to the number of self-consistent field (SCF) iterations required, considerable effort has been devoted to the formulation of more effective mixing schemes over the years, see, e.g., [5,6] and references therein.

Perhaps the most widely used mixing scheme is Pulay's Direct Inversion in the Iterative Subspace (DIIS) [8,9], based on Anderson's extrapolation [10]. Pulay's technique represents a specific variant of Broyden's quasi-Newton approach [11–13] and falls into the broad category of multisecant methods [14]. Although the relative simplicity and overall performance of DIIS [15] make it an attractive choice, it has been observed that Pulay mixing can stagnate and/or otherwise perform poorly in calculations involving certain metallic and/or inhomogeneous systems [7,16]. This has motivated the development of a number of alternative approaches, including variants of Broyden's method [13,17–19], the Relaxed Constrained Algorithm (RCA) [20,21], and a variety of preconditioning techniques [7,22–25]. However, while the improvements demonstrated have been in some cases substantial, increased complexity, additional parameters, and/or lack of transferability have hindered adoption in practice.

In this work, we introduce a simple generalization of the DIIS method for accelerating self-consistent field iterations, which we refer to as the Periodic Pulay method. The approach can be understood as the application of the recently developed Alternating Anderson-Jacobi (AAJ) technique [26]—an efficient solver for large-scale linear systems in the framework of the classical Jacobi fixed-point iteration—to SCF iterations in electronic structure calculations. Contrary to the conventional wisdom that DIIS generally far outperforms linear mixing, the central idea of the Periodic Pulay method is to employ Pulay extrapolation only once every few SCF iterations, and use linear mixing on all other iterations. We find that this simple generalization not only improves the efficiency of DIIS, but also makes it more robust. In addition, since the majority of electronic structure codes in current use (e.g., [27–31]) already employ Pulay mixing, the proposed technique can be easily incorporated.

* Corresponding author.
E-mail addresses: baner041@umn.edu (A.S. Banerjee), phanish.suryanarayana@ce.gatech.edu (P. Suryanarayana), pask1@llnl.gov (J.E. Pask).

http://dx.doi.org/10.1016/j.cplett.2016.01.033
0009-2614/

The remainder of this Letter is organized as follows. In Section 2, we present the Periodic Pulay method. In Section 3, we examine its performance for a wide range of materials systems in the context of DFT. Finally, we provide concluding remarks in Section 4.

2. Periodic Pulay method

The self-consistent field (SCF) method casts the equations for the electronic ground-state as the fixed-point problem

$$\boldsymbol{\rho} = \mathbf{g}(\boldsymbol{\rho}), \tag{1}$$

where $\boldsymbol{\rho} \in \mathbb{R}^{N\times 1}$ is the electron density, and the nonlinear mapping $\mathbf{g} : \mathbb{R}^{N\times 1} \to \mathbb{R}^{N\times 1}$ is composed of the effective potential evaluation for a given electron density and electron density evaluation for the associated Hamiltonian. The convergence properties of the SCF iteration in the vicinity of the solution are determined by the properties of the Jacobian of the residual function $\mathbf{f}(\boldsymbol{\rho}) = \mathbf{g}(\boldsymbol{\rho}) - \boldsymbol{\rho}$ [7]. Therefore, a strategy that leads to improved conditioning/solvability of the linear system associated with the Jacobian may also lead to improved convergence of the SCF iteration [32]. In this context, the effectiveness of the GMRES approach [33] in solving linear systems is closely related to the success of the DIIS method in accelerating SCF iterations [14,34,35]. Similarly, well established ideas for accelerating the solution of linear systems through preconditioning have found their counterparts in electronic structure calculations [7,22,23,25].

In recent work [26], a new solver for large-scale linear systems of equations has been developed in which Anderson extrapolation is performed at periodic intervals within the classical Jacobi fixed-point iteration. On one hand, under-relaxed Jacobi iterations are well known to rapidly damp higher-frequency components of the residual [36]. On the other hand, periodic application of Anderson extrapolation has the effect of damping lower-frequency components. Therefore, the simultaneous application of these two methods stands to efficiently reduce the overall norm of the residual, thus leading to the success of the so called Alternating Anderson-Jacobi (AAJ) method [26]. In particular, AAJ has been found to significantly outperform both GMRES and Anderson-accelerated Jacobi methods. This provides the motivation for the Periodic Pulay mixing scheme proposed here, which can be viewed as the extension of the AAJ method to SCF fixed-point iterations in electronic structure calculations.

In the Periodic Pulay method, Eq. (1) is solved using the following fixed-point iteration:

$$\boldsymbol{\rho}_{i+1} = \boldsymbol{\rho}_i + \mathbf{C}_i \mathbf{f}_i, \tag{2}$$

where the subscript i denotes the iteration number, $\mathbf{f}_i = \mathbf{f}(\boldsymbol{\rho}_i)$, and the matrix

$$\mathbf{C}_i = \begin{cases} \alpha \mathbf{I} & \text{if } (i+1)/k \notin \mathbb{N}, \quad \text{(Linear mixing)} \\ \alpha \mathbf{I} - (\mathbf{R}_i + \alpha \mathbf{F}_i)(\mathbf{F}_i^T \mathbf{F}_i)^{-1} \mathbf{F}_i^T & \text{if } (i+1)/k \in \mathbb{N}. \quad \text{(Pulay mixing)} \end{cases} \tag{3}$$

In the above expression, $\mathbf{R}_i$ and $\mathbf{F}_i$ denote the iterate and residual histories:

$$\mathbf{R}_i = [\Delta\boldsymbol{\rho}_{i-n+1}, \Delta\boldsymbol{\rho}_{i-n+2}, \ldots, \Delta\boldsymbol{\rho}_i] \in \mathbb{R}^{N\times n}, \tag{4}$$

$$\mathbf{F}_i = [\Delta\mathbf{f}_{i-n+1}, \Delta\mathbf{f}_{i-n+2}, \ldots, \Delta\mathbf{f}_i] \in \mathbb{R}^{N\times n}, \tag{5}$$

where $\Delta\boldsymbol{\rho}_i = \boldsymbol{\rho}_i - \boldsymbol{\rho}_{i-1}$ and $\Delta\mathbf{f}_i = \mathbf{f}_i - \mathbf{f}_{i-1}$. In addition, α is the mixing parameter, n is the size of the mixing history, and k is the frequency of Pulay extrapolation. We outline the aforedescribed approach in Algorithm 1, wherein *tol* specifies the residual convergence criterion.

Algorithm 1. Periodic Pulay method.

Input: $\boldsymbol{\rho}_0$, α, n, k, and tol
repeat $i = 0, 1, 2 \ldots$
  $\mathbf{f}_i = \mathbf{g}(\boldsymbol{\rho}_i) - \boldsymbol{\rho}_i$
  if $(i+1)/k \in \mathbb{N}$ **then**
    $\boldsymbol{\rho}_{i+1} = \boldsymbol{\rho}_i + \alpha\mathbf{f}_i - (\mathbf{R}_i + \alpha\mathbf{F}_i)(\mathbf{F}_i^T\mathbf{F}_i)^{-1}\mathbf{F}_i^T\mathbf{f}_i$
  else
    $\boldsymbol{\rho}_{i+1} = \boldsymbol{\rho}_i + \alpha\mathbf{f}_i$
until $\|\mathbf{f}_i\| < tol$;
Output: $\boldsymbol{\rho}_i$

The Periodic Pulay method can be considered as a generalization of both the classical Pulay and linear mixing schemes. Specifically, the classical Pulay scheme is recovered for frequency of extrapolation $k = 1$, while the linear mixing scheme is recovered as $k \to \infty$. In principle, the parameter k is arbitrary and independent of the mixing history size n. However, it is worthwhile to narrow the parameter space for k, particularly from the perspective of practical calculations. For this purpose, we note that larger values of k typically lead to more stable but slowly converging SCF iterations (a consequence of increased linear mixing), while smaller values of k tend to provide less damping of higher-frequency error components and correspondingly slower overall convergence as well. In practice, a balance between these limits is preferable and we have found that it is usually counterproductive to set $k > n/2$ when n is even, and $k > (n+1)/2$ when n is odd; and similarly counterproductive to set $k < 2$.

The effectiveness of alternating Pulay and linear-mixing iterations has been recognized before. In particular, the Guaranteed Reduction Pulay (GR-Pulay) scheme [30,37] alternates classical Pulay and linear mixing in successive SCF iterations with mixing parameter $\alpha = 1$. As such, GR-Pulay can be understood as a special case of Periodic Pulay with $\alpha = 1$ and $k = 2$. However, in more difficult cases, e.g., highly inhomogeneous and/or metallic systems at low temperature, fixing $\alpha = 1$ can degrade performance or lead to SCF divergence. The flexibility of reducing α in such cases is thus important to retain. The flexibility to vary k as well—within prescribed limits—can also be advantageous, as we show below.

While the mixing parameters in the Pulay and linear extrapolations are in general distinct, we use the same mixing parameter for both in the present work. Indeed, the approach can be further generalized by employing different values for the two parameters. However, we have found that such a strategy does not yield significant gains in practice, and therefore refrain from introducing this additional parameter here. Also, it is worth noting that the Periodic Pulay method does not rely on any heuristics based on the variation of the total energy or residual during the SCF iteration. Yet, as demonstrated in the next section, we find the method to be both robust and efficient for the full range of systems considered, even when classical Pulay fails to converge. Finally, though we have described the Periodic Pulay scheme in terms of density mixing, it is identically applicable to potential mixing, the expressions for which can be obtained by replacing the electron density appearing in Eqs. (1)–(5) with the relevant potential.

3. Results and discussion

In this section, we verify the robustness and efficiency of the Periodic Pulay method for accelerating the self-consistent field iteration in density functional theory calculations. For this purpose, we implement the Periodic Pulay scheme in the SIESTA code [31,38], wherein mixing is performed on the density matrix. To demonstrate the effectiveness of the method across diversephysical