

Responses to the Comments

November 2, 2024

Response to Referee #1

The manuscript “Conjugate gradient direct minimization on the complex Stiefel manifold in Kohn-Sham density functional theory for finite and extended systems” presents an energy minimization approach to Kohn-Sham (KS) density functional theory (DFT) calculations, as an alternative to current state-of-the-art self-consistent-field (SCF) methods. The method is applied to small molecular and extended (crystalline) systems. To carry out the minimization subject to orthogonality constraints on the KS orbitals, a Riemannian conjugate gradient (RCG) method is employed. To account for partial occupation, Rayleigh quotients are computed to approximate eigenvalues from eigenvector approximations. The paper is interesting, though in its present form raises a number of questions/concerns. If these can be fully addressed, the work can be considered for publication in Computer Physics Communications.

Response: We deeply thank you for your efforts in carefully reading our manuscript, characterizing our work positively, and providing critical feedback in a professional manner. In response to the comments, significant improvements have been made, and all issues have been addressed point by point. We believe that the resubmitted version meets the strict standards for publication in the **Computer Physics Communications**. We cordially invite you to share any further questions or concerns you may have.

1. *Such matrix manifold minimization approaches have been around since the at least the late 1990’s, e.g., Refs. 25-32 in the manuscript. And they have yet to replace the standard SCF method in any widely used codes (ABINIT, Quantum Espresso, VASP, Elk, Wien2k, Gaussian, NWChem, GAMESS, etc.). What is new and significant in the present work relative to previous works [25-32]? Mathematical formulation? Algorithm? Implementation? It was not clear to me, as currently written, what is actually new and why it is significant.*

Response: We thank the referee for pointing out the absence of a clear significance statement in our current manuscript. Indeed, it is true that the manifold minimization method has not yet been incorporated into the standard codes. While many researchers have explored this approach extensively in recent years, to the best of our knowledge, we are the first to offer a **unified, basis-adaptable, algebraic** formulation for both finite and extended systems within the manifold minimization method applied to the Kohn-Sham model.

Moreover, we have demonstrated that our implementation of this formulation performs well in a non-orthogonal basis for both finite and extended systems, for both gapped and non-gapped systems with some caveats. It’s also important to note that we are only working with Stiefel

manifold such that non-idempotent density matrix can be treated, in contrast to previous work on the Grassmann manifold.

Given our ultimate goal of implementing the reduced density matrix functional theory (RDMFT), we find this formulation particularly advantageous, as it naturally handles orthogonality constraints required for natural orbitals. It is also convenient that this type of direct minimization suits for our goal where the traditional iterative diagonalization meets difficulty. We are the first to propose to use manifold minimization for RDMFT calculations in this work. We believe that the success in the Kohn-Sham problem represents a major stepping stone and is worth disseminating to the computational physics community.

Action: We have added a clear significance statement in the introduction and modified the motivation accordingly.

2. *With inconsistent notation (e.g., M basis functions and N orbitals early in the manuscript switching to n basis functions and p orbitals later; H is a differential operator in Eq. 7 but has “size of the basis sets M ”) and multiple layers of said notation to make expressions compact, I found the manuscript notably hard to follow. To clarify, the inconsistencies need to be resolved and compact notation expanded. The latter issue can perhaps be most straightforwardly addressed by elaborating Algorithm 1, or adding another more detailed version, which: (a) applies to the Kohn-Sham solution in extended systems with partial occupations; (b) uses notation consistent with the preceding mathematical formulation; (c) translates all notation to the actual linear algebra (matrix/vector/scalar) operations performed. The “ $\text{grad } f()$ ”, “ $R_{X_k}()$ ”, and “ $T_{\alpha_k d_k}()$ ” notations need to be translated to corresponding linear algebra operations in particular.*

This should clarify what the proposed method is actually doing. As is, it is not entirely clear to me that the method is doing anything more than explicit orthogonalization after standard conjugate gradient operations as needed, as done long before formulation in terms of Grassmann/Stiefel manifolds.

Response: We apologize for the confusion caused by the mixed notations in the manuscript. To clarify, we have adhered to the general convention for the manifold, where the matrix dimensions are $n \times p$. We only specify the matrix size as $M \times N$ when discussing the electronic structure problem in the context of a particular basis set, where M represents the number of basis functions and N the number of orbitals.

In Algorithm 1, the relevant linear algebra operations are denoted. For instance, in line 3: the operation is performed “using Eq. 34,” and in line 5, “using Eq. 36.” We believe the missing detail pertains to the Riemannian gradient and we have added the relevant operation.

Our proposed method differs from Gillan’s approach, where repeated explicit orthogonalization is required. Instead, our method explicitly leverages the algebraic structure of the underlying manifold on which the orbitals reside. Due to the curvature of the manifold, the line search direction ought to be modified and the movement need to stay on the manifold. What is common is that both approaches aim to preserve the orthogonality condition throughout the minimization process. In applied mathematics, this is often referred to as the “feasibility” method in contrast to other non-feasible methods.

Action: We have inserted an explicit statement in the notation section to clarify where we use n -by- p and where we use M -by- N to make it consistent throughout the manuscript. We have also changed “The dimension of the Hamiltonian matrix $H_{ij} \equiv \langle \psi_i | h | \psi_j \rangle$ is the size of

the basis functions M .” We modified all the equations to have the k dependence all the way and deleted the subsection of “Extended system with multiple k points”. We have also made change to “grad $f()$ ” with operations of Eq. 37 in Algorithm 1.

3. *How does the proposed method account for possible degeneracies in the highest computed eigenvalues/vectors? For example, in a non-spin-polarized calculation with 1s, 2s, and 2p states, with 5 electrons and p_x, p_y , and p_z states degenerate, if only the lowest 3 KS orbitals are computed (to contain 5 electrons), then some arbitrary mixture of p_x, p_y , and p_z states may be computed rather than computing (and using) all of them. And this may in turn result in slow convergence, if convergence occurs at all, and wrong answers. Not accounting for such degeneracies (and near degeneracies) may, for example, be a source of slow- and failed convergence noted Section 4.2.*

Response: Thanks for the question. The case of possible degeneracies (and near degeneracies) is taken care of with fractional occupation, where the Stiefel manifold is required. The use of the Grassmann manifold could not deal with such cases. Choosing a set of smearing method and parameters can definitely handle the scenario described here, which both the SCF and the RCG methods work well. The only requirement is to have the number of bands larger than dimension of the occupied space. The failed convergence noted in Section 4.2 is most probably linked to initialization. From our tests, it has no problem in treating degeneracies. For the problem at hand, we performed a spin-unpolarized calculation for boron atom, sitting within a box size of 15.0 angstrom. Using the default 12 bands and Γ point, the eigenvalues and occupation numbers are listed in Table 1. Both the SCF and RCG methods give identical

Table 1: Eigenvalues and occupations for boron (B) atom for SCF and RCG.

| State index | Eigenvalue (eV) | Occupation |
|-------------|-----------------|------------|
| 1 | -9.26491 | 2.00000 |
| 2 | -3.38837 | 0.333333 |
| 3 | -3.38837 | 0.333333 |
| 4 | -3.38837 | 0.333333 |
| 5 | 11.0252 | 0.00000 |
| 6 | 11.0252 | 0.00000 |
| 7 | 11.0252 | 0.00000 |
| 8 | 19.6559 | 0.00000 |
| 9 | 19.6559 | 0.00000 |
| 10 | 19.6559 | 0.00000 |
| 11 | 19.6559 | 0.00000 |
| 12 | 19.6559 | 0.00000 |

eigenvalues and occupation numbers and the same total energy -70.679567 eV.

Action: We have added a sentence regarding the treatment of possible degeneracies (and near degeneracies).

4. *Issue of significant practical importance: How can the proposed method be applied to spin-polarized calculations of magnetic systems, where multiple solutions are possible corresponding to different spin configurations for a given atomic/crystal structure, e.g., ferromagnetic, antiferromagnetic, high-spin, low-spin, sometimes close in energy? The usual way to reach such*

different solutions using the standard SCF method is to provide an initial guess closer to one solution than the others. Can the proposed approach be expected to reach different solutions in the same way? Why or why not?

Response: We appreciate your question on the magnetic systems. The current approach could be easily extended to the spin-polarized calculations where essentially each wavefunction needs to become a two component quantity. However, this approach might find the same trouble in discerning states with very close energies as in the SCF method. For the energy landscape where multiple local minima are not well separated from a given initial state, this approach may reach different solutions for depending on the initial direction (though usually the initial direction takes the steepest descent direction). This would be true as well for the SCF method with different density mixing schemes. Both methods can only give a local minimum solution for a given initial state.

5. *What pseudopotentials were used in each case? Established nonlocal potentials, e.g., PseudoDojo, SG15, SSSP, SPMS? If only local potentials were used (i.e., lacking nonlocal projectors), results could show significant differences from established nonlocal potentials and/or reference all-electron results.*

Response: We have used the standard SG15-V1.0 pseudopotentials for all the calculations.

Action: We have inserted "SG15-V1.0" in the pseudopotential description to make it explicit.

6. *In the molecular calculations, what were the domain sizes chosen? How were they determined to be sufficiently large for wavefunctions and charge density to vanish sufficiently for the energy accuracy desired? What energy accuracy was targeted?*

Response: In the reported calculations, the domain size is quite large, which is a cubic box of size 15 angstrom. Conventionally, this is much bigger than necessary for the charge density to decay. We believe the critical issue is not about how accurate the calculation we try to achieve. Even in the traditional SCF (diagonalization) method, it depends on many factors including quality of pseudopotentials, basis-sets, simulation domain size, effective energy cutoff, etc. Instead, whether we can reproduce the results of the SCF method is the ultimate test of the proposed approach. In any rate, the energy accuracy (given a certain energy tolerance) has been matched to the SCF method within 0.003 meV for tested molecules.

Action: We have added clarifications about domain size.

7. *In the molecular calculations, what were the boundary conditions on the computed wavefunctions and potential? Both Dirichlet? Both vanishing at the domain boundary or only wavefunctions vanishing? If the latter, what was the boundary condition on the potential and how was it determined?*

Response: The boundary conditions for the computed wavefunctions and potential are periodic boundary conditions.

Action: We have added clarifications about boundary conditions (also for next question).

8. *In the extended (solid) system calculations, what were the unit cells chosen in each case? What were the boundary conditions on the computed wavefunctions and potential?*

Response: To answer the question, we listed the lattice vectors and lattice constants for these systems.

Periodic boundary conditions are used.

Table 2: Unit cells used for the tested solids.

| System name | Lattice vectors | Lattice constant |
|-------------|--|------------------|
| Cu | 0.50 0.50 0.00 0.50 0.00 0.50 0.00 0.50 0.50 | 6.9164 bohr |
| LiF | 1.00 0.00 0.00 0.00 1.00 0.00 0.00 0.00 1.00 | 4.028 Å |
| Mg | 1.00 0.00 0.00 0.00 1.00 0.00 0.00 0.00 1.00 | 3.55 Å |
| MgO | 1.00 0.00 0.00 0.00 1.00 0.00 0.00 0.00 1.00 | 4.25648 Å |
| NaCl | 1.00 0.00 0.00 0.00 1.00 0.00 0.00 0.00 1.00 | 5.620 Å |
| SiC | 1.00 0.00 0.00 0.00 1.00 0.00 0.00 0.00 1.00 | 4.348 Å |

9. *For meaningful comparisons of the proposed RCG method to the standard SCF method, the major operations of each method need to be specified, so that it is clear which operations are the same (hopefully most) and which are different. The manuscript states that the implementation was in the ABACUS software package. Please summarize the major operations of the standard SCF implementation in ABACUS, and the differences from that corresponding to the RCG implementation.*

Response: Thanks for the suggestions. Here we list the major operations of the standard SCF implementation in ABACUS ver.3.5.0. Since both the atomic basis and the planewave basis are supported, in the initialization step, the quantities for both schemes are prepared. In the SCF process, the major steps include initialization of SCF (including overlap matrix, Hamiltonian matrix, superposition of atomic density etc.) in `beforescf()`. Diagonalization of matrix and density matrix formation are performed in `hamilt2density()`. A mixing scheme for the electron density is handled by `mix_rho()`, Kohn-Sham potential is updated in `update_pot()` and energy components are calculated iteratively in `eachiterfinish()`.

In the RCG implementation, the main omission is the diagonalization of the matrix, which is replaced with the line search on the manifold. The associated cost in line search is the evaluation of the objective function and possibly its gradient.

Action: We have inserted a flowchart for the RCG method.

10. *For meaningful comparison to numbers of iterations in the standard SCF method, one must use a reasonable choice of SCF mixing in the standard SCF method: e.g., Anderson/Pulay mixing with history length > 5 and model dielectric function preconditioner is common (a simple case being Kerker preconditioner for metallic systems; see, e.g., <https://docs.abinit.org/variables/gstate/#dielng>). What mixing method and parameters were used in the SCF calculations being compared to? How are the number of iterations affected by the choice?*

Response: In the SCF method, the simplified modified Broyden mixing (D. D. Johnson Phys. Rev. B 38, 12807, (1988)) with default history length `mixing_ndim=8` and mixing parameter `mixing_beta=0.8` (for `nspin=1`) were used. Normally, the Broyden mixing is slightly faster than the Pulay mixing. The default mixing setup has been tested to function well if not optimally. For each case, one could in principle select the optimal parameters that uses less computational time but that is not what we are aiming for. We note that the number of iterations with Kerker preconditioner on by default for molecular system is slightly larger (1-2 more iterations) than when we turned off the preconditioner. However, turning off the preconditioner, three cases in SCF fail to converge (CCH radical, CH3CH2O radical and Si2). Previous reported numbers of iterations essentially does not change.

Action: We have modified the description to include more details on the SCF setup.

11. *For meaningful comparisons of timings, it is essential to know the key aspects of the algorithms, implementations, and computational resources being compared. For the present ABACUS implementations of SCF and RCG methods, what is the timing breakdown of key computational steps? E.g., Initialization (before SCF iterations begin), construction of initial guess, computation of potential, construction of Hamiltonian, diagonalization of Hamiltonian (or RCG minimization step), construction of density, computation of total energy. For such small problems, diagonalization should take negligible time and so it is not clear to me why RCG should be faster – assuming reasonable SCF mixing, comparably tight tolerances for SCF and RCG, etc. It is particularly important to break out initialization time, since that can vary widely depending on implementation.*

Response: Thank you so much for the suggestions. To begin with, the computational resources used for both methods are identical, as indicated in the manuscript where multi-threading was disabled and a single core used. As pointed out in previous major operations in ABACUS implementation, we have followed the your suggestions and break down the timing of each operation. We realize that in previous calculations for SCF methods, by mistakenly setting to output files onto the filesystem, the computational time reported is greatly overestimated. Same setting in the RCG input was not effective though. Therefore, we have reanalyzed these calculations and retrieved the observations that RCG is more efficient than SCF.

Based on the current implementation, for most of the molecular calculations RCG takes more time than that of SCF and only about 18 cases RCG takes less time. In our preliminary implementation of BFGS, we found that 81 cases where BFGS is faster than that of SCF.

Table 3: The class name for SCF is `ESolver_KS_LCAO::`, for the case of $\text{CH}_2=\text{CH}-\text{CH}=\text{CH}_2$ molecule. The total time for the main program using `Driver::driver_line()` is 42.57 s. The number of iteration is 23, and on average the time per iteration is about 1.81 s.

| function name | number of calls | total time (s) | average time per call (s) |
|-------------------------------|-----------------|----------------|---------------------------|
| <code>Init()</code> | 1 | 0.78 | 0.78 |
| <code>beforescf()</code> | 1 | 1.44 | 1.44 |
| <code>hamilt2density()</code> | 23 | 17.93 | 0.78 |
| <code>mix_rho()</code> | 22 | 3.69 | 0.17 |
| <code>updatepot()</code> | 23 | 18.44 | 0.80 |
| <code>eachiterfinish()</code> | 0 | 0.00 | 0.00 |

In contrast, the total time of RCG is 72.15 s within 31 iterations. The average time per iteration is 2.32 s. The breakdown is not easily compared toe-to-toe with that of SCF approach (see Figure 1). The number of function evaluations, gradient evaluations, retractions evaluations, and vector transport evaluations is 47, 33, 46, and 96, respectively. From our output, however, we see that the time spent on the manifold related operations is nothing. Out of these, the main consumption is inside the function evaluation (or energy calculation) and potential update. Our implementation has used mostly available routines already in the SCF procedure, and there are probably many redundant calls that could be avoided to increase its efficiency.

Action: We have appended details and the flowchart for RCG method.

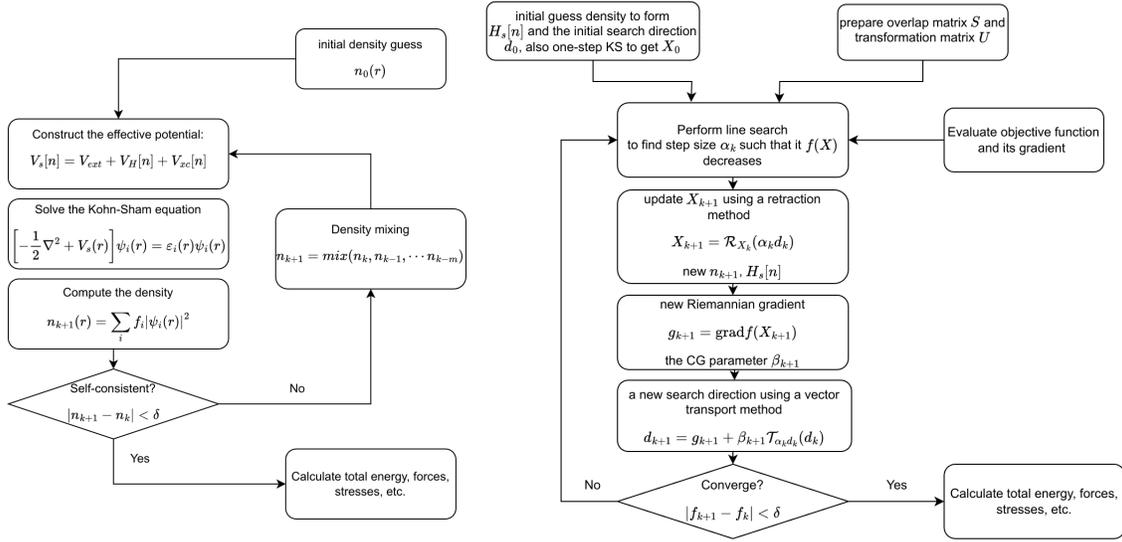


Figure 1: The flowchar for the SCF and the RCG methods.

12. Also, for meaningful comparisons of timings, it is important that the same stopping criteria are used for the methods being compared, and that those criteria are sound. For example, energy convergence may stall, giving small differences from one iteration to the next, before charge density is well converged. What were the stopping criteria for the SCF and RCG calculations whose timings were compared? How did the convergence of density compare in each case?

Response: We agree with your general comment regarding stopping criterion. In the version where we implemented RCG, the SCF in ABACUS only uses the charge density criterion meaning that the charge density error between two sequential densities is less than 1.0×10^{-6} . In a later version, the energy criterion is implemented only recently. In contrast, we used the energy criterion in the RCG method, which means the energy error is less than 1.0×10^{-8} . We might also require the norm of gradient to be less than 1.0×10^{-5} . Though the criterion varies, the energy agree with each other to the accuray at least milli Hartree, which we believe is sufficiently tight and both tolerances are comparable.

Response to Referee #2

The paper deals with solving the Kohn-Sham DFT minimization problem by using a Riemannian conjugate gradient method on the complex matrix Stiefel manifold. Such a method has previously been considered and thoroughly analyzed in [X. Dai, Z. Liu, L. Zhang, and A. Zhou. A Conjugate Gradient Method for Electronic Structure Calculations. SIAM J. Sci. Comp. 39(6):A2702-A2740, 2017, DOI: 10.1137/16M1072929]. Since the paper does not contain any new results, I would recommend to reject it.

Response: Thank you for taking the time to review our manuscript. We appreciate your reference to the prior work by Dai et al. (2017), which has indeed provided a strong foundation for the use of Riemannian conjugate gradient (RCG) methods in electronic structure calculations.

However, we would like to clarify a few things. To our knowledge, the first paper that provides the framework that deals with the Kohn-Sham DFT minimization problem is by Edelman et al. [Alan Edelman, Tomás A. Arias, and Steven T. Smith, The Geometry of Algorithms with Orthogonality Constraints, *SIAM Journal on Matrix Analysis and Applications* 20, no. 2 (1998): 303-353] We would also point out the theory has been implemented faithfully using the geodesic approach using planewave basis in NWCHEM.

We also note that two papers by Zaiwen Wen have already considered the gradient method to electronic structure calculations. These two are [Zaiwen Wen, Wotao Yin, A feasible method for optimization with orthogonality constraints, *Math. Program., Ser. A* (2013) 142:397–434], and [Xin Zhang, Jinwei Zhu, Zaiwen Wen, and Aihui Zhou, Gradient Type Optimization Methods For Electronic Structure Calculations, *SIAM Journal on Scientific Computing* 36, no. 3 (2014): C265-C289]. Then we have the 2017 paper by X. Dai et al. which considered different orthogonality preserving strategies in conjugate gradient method.

These papers essentially all treated finite systems but periodic systems were not seriously considered. The manifold is thus real in principle. Those numerical experiments were done for finite systems only. The reason they considered complex manifold is only when the basis is itself complex. In contrast, the complexity of our manifold comes from the periodic treatment irrespective the complexity of the basis. These papers did not consider possible fractional occupations in forming the densities, and either the Grassmann or the Stiefel manifold could be used.

The work that most closely aligns with our approach is the recent 2023 paper by X. Dai [Xiaoying Dai, Stefano de Gironcoli, Bin Yang, and Aihui Zhou, Mathematical Analysis and Numerical Approximations of Density Functional Theory Models for Metallic Systems, *Multiscale Modeling and Simulation* 21, no. 3 (2023): 777-803.] In this work, the complex manifold for periodic systems were used and metallic systems with fractional occupation were the main focus. But even so, a distinction is that these work use orthogonal basis (real space as done in Octopus, or planewave as done in KSSOLV or modified Quantum Espresso). In our formulation, it is a basis adaptable algebraic formulation and our implementation used localized non-orthogonal numerical atomic basis. Lastly, our work represents a pioneering effort to forge a connection between physicists and theoretical chemists who are actively engaged in electronic structure methods and the applied mathematics community. By facilitating collaboration and communication across these disciplines, we aim to enhance the understanding and application of Riemannian optimization in the study of electronic structures.

While we acknowledge that our paper builds on the concepts introduced in the aforementioned work, we believe these several key aspects mentioned distinguish our study and offer novel contributions.

Action: We have added citations to these work and discussions on the major differences in the introduction.