Computer Physics Communications 312 (2025) 109596

Contents lists available at ScienceDirect

Computer Physics Communications

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

Computational Physics

Direct minimization on the complex Stiefel manifold in Kohn-Sham density functional theory for finite and extended systems

Kai Luo [a,*], Tingguang Wang [a], Xinguo Ren [b]

[a] Department of Applied Physics, School of Physics, Nanjing University of Science and Technology, Nanjing 210094, China
[b] Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

The review of this paper was arranged by Prof. W. Jong

Keywords:
Density functional theory
Stiefel manifold
Direct minimization
Conjugate gradient
Extended systems

ABSTRACT

Direct minimization method on the complex Stiefel manifold in Kohn-Sham density functional theory is formulated to treat both finite and extended systems in a unified manner. This formulation is well-suited for scenarios where straightforward iterative diagonalization becomes challenging, especially when the Aufbau principle is not applicable. We present the theoretical foundation and numerical implementation of the Riemannian conjugate gradient (RCG) within a localized non-orthogonal basis set. Riemannian Broyden-Fletcher-Goldfarb-Shanno (RBFGS) method is tentatively implemented. Extensive testing compares the performance of the proposed methods and highlights that the quasi-Newton method is more efficient. However, for extended systems, the computational time required grows rapidly with respect to the number of **k**-points.

1. Introduction

Density functional theory (DFT) stands as a highly utilized method for simulating a wide range of physical systems, including atoms, molecules, clusters, solids, and other complex forms of matter. This popularity stems from its exceptional balance between accuracy and computational efficiency. The ingenious implementation of Kohn-Sham (KS) theory within DFT precisely addresses the non-interacting kinetic energy through the solution of a one-body Schrödinger equation, incorporating an effective potential inclusive of exchange-correlation (xc) effects [1,2].

Presently, the self-consistent field (SCF) algorithm derived from the first-order necessary optimality condition is the prevailing method for tackling the Kohn-Sham problem. It revolves around identifying eigenvalues and corresponding eigenvectors within the occupied space of the Hamiltonian matrix. However, while widely adopted, this approach is susceptible to convergence issues. A well-designed density update scheme is thus indispensable for achieving convergence within a reasonable number of iterations [3,4]. Meanwhile, the converged solution may occasionally be a saddle point of the energy surface rather than a minimum [5].

As an alternative, the Kohn-Sham problem can be treated as an optimization problem. One approach is the direct minimization of the Kohn-Sham energy functional with respect to the electronic degrees of freedom. This method requires ensuring that the orbitals remain orthonormal, thereby framing the task as a constrained optimization problem. This can be achieved by applying explicit orthonormalization, such as Gram-Schmidt or QR orthonormalization to the updated orbitals after each iteration [6,7]. This constrained problem can also be reformulated into an unconstrained optimization problem using penalty function methods [8–11] or augmented Lagrangian (ALM) methods [8,12,13,9,14,15]. However, the smoothness requirement for penalty functions necessitates that the original objective function has high-order smoothness. In practice, calculating the exact gradients of these penalty functions for non-convex problem is often computationally expensive, and selecting appropriate penalty parameters can be challenging. Recent parallelizable frameworks within the ALM method demonstrated effectiveness and high scalability, showing promise in electronic structure calculations [14,15].

The constraint can also be fulfilled by a unitary transformation matrix, which is applied to a set of orthonormal reference orbitals and then optimized. Using the exponential transformation with a skew-Hermitian matrix exponent in a compact basis set such as linear combination of atomic orbitals (LCAO), it has been shown that good performance can be achieved compared to the SCF method for both finite and extended systems [16–24]. However, for non-compact basis functions of size M (e.g. plane waves), computing the exponential of these matrices typically scales as $\mathcal{O}(M^3)$ and thus making it computationally expensive.

* Corresponding author.
E-mail address: kluo@njust.edu.cn (K. Luo).

https://doi.org/10.1016/j.cpc.2025.109596
Received 4 November 2024; Received in revised form 24 March 2025; Accepted 26 March 2025

Available online 28 March 2025

In recent years, considerable progress has been made in the area of Riemannian optimization for the electronic structure theory. The non-convex problem can be converted to a geodesically convex problem in a curved space. Since the Kohn-Sham energy is defined on a curved space (Riemannian manifold), the optimization has to take the curvature into account [25,26]. The extension of unconstrained optimization from Euclidean spaces leads to Riemannian optimization. In addition to its effective utilization in various classical optimization problems with geometric restrictions, Riemannian optimization has proven highly beneficial in electronic structure computations [27–33]. In these works by Wen et al. [29], Zhang et al. [30], and Dai et al. [33], finite systems were treated. Ref. [33] did not address periodic systems. As a result, the underlying manifold is thus real in principle and only when the basis (e.g. planewaves) is itself complex does the manifold become complex. However, in the periodic systems, complexity is required due to the intrinsic complexity of the Bloch states. We note that a recent 2023 paper [34], which is motivated by metallic systems and employs a plane-wave basis, uses Bloch-periodic boundary conditions for periodic calculations, as we do here; whereas our formulation is in terms of a general, non-orthogonal basis rather than a global, orthonormal one as in Ref. [34].

Despite of enormous success of the Kohn-Sham density functional theory, it finds difficulty in handling systems with strong correlations, such as Mott insulators. One promising theory for this matter going beyond KSDFT is the reduced density matrix functional theory (RDMFT) [35,36], where the traditional iterative diagonalization meets difficulty and the Aufbau principle is not applicable. The orthogonality constraint for the natural orbitals (eigenfunctions of the one-particle reduced density matrix (1RDM)), can be easily integrated in the Stiefel manifold.

In this study, we introduce a unified formulation adaptable to any basis for both finite and extended systems within the manifold minimization method. An implementation based on inexact line search of the conjugate gradient (CG) (and tentatively Broyden-Fletcher-Goldfarb-Shanno (BFGS)), for the Kohn-Sham problem is provided within a non-orthogonal local basis set. By incorporating fractional occupation, both metallic and degenerate (or nearly degenerate) systems can be treated, which is only possible in the Stiefel manifold but not in the Grassmann manifold. Their performances on finite systems and extended systems are compared against the standard SCF method and discussed. Its success on the Kohn-Sham problem lays a solid foundation for the ongoing development of RDMFT and other theories that require non-idempotent density matrix.

2. Theory formulation

2.1. Notations

For a complex matrix $A \in \mathbb{C}^{m\times n}$, matrices $A^\dagger$ and A^{-1} denote the complex conjugate transpose and inverse of A, respectively. For a vector $d \in \mathbb{C}^n$, the operator $\mathrm{Diag}(d)$ returns a square diagonal matrix in $\mathbb{C}^{n\times n}$ with the elements of d on the main diagonal, while conversely $\mathrm{diag}(A)$ returns the vector in $\mathbb{C}^n$ containing the main diagonal elements of the square matrix $A \in \mathbb{C}^{n\times n}$. I_p is an identity matrix of size $p \times p$. The symmetrized matrix of a square matrix A is denoted as $\mathrm{sym}(A) = (A + A^\dagger)/2$. The trace of A, i.e., the sum of the elements on the main diagonal of a square matrix $A \in \mathbb{C}^{n\times n}$, is denoted by $\mathrm{tr}(A)$. The Frobenius inner product in Euclidean space for matrices $A, B \in \mathbb{C}^{m\times n}$ is defined as $\langle A, B\rangle_e = \mathrm{tr}\left(A^\dagger B\right)$, and the corresponding Frobenius norm $\|\cdot\|_F$ is given by $\|X\|_F = \langle A, A\rangle^{1/2} = \left(\sum_{i,j}\left|A_{ij}\right|^2\right)^{1/2}$. For a k indexed matrix A_k, the boldface notation $\mathbf{A}$ is to denote $(A_1, A_2, \ldots, A_K)$ and is of size K. When we are specifically dealing with electronic structure problems, we may use indices such as M, N. Otherwise, m, n, p will be used for a general matrix.

2.2. Continuous Kohn-Sham DFT model

In Kohn-Sham density functional theory, the central quantity in the variational principle is the energy functional

$$E_{\mathrm{KS}}[\{\psi_i(\mathbf{r})\}] = -\frac{1}{2}\sum_i f_i \int \mathrm{d}\mathbf{r}\, \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r}) + E_{\mathrm{H}}[n] + E_{\mathrm{xc}}[n] + \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})v_{\mathrm{ext}}(\mathbf{r}), \tag{1}$$

where the Hartree energy is given by

$$E_{\mathrm{H}}[n] = \frac{1}{2}\int\int \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \tag{2}$$

and $E_{\mathrm{xc}}[n]$ is the exchange-correlation energy functional, which has to be approximated in practice.

Here and throughout this article we work in atomic units and therefore have set $\hbar = m_e = e = 1$, where m_e is the electron mass and e is the charge of the proton. v_{ext} is the external potential for electron-nuclei interaction.

The electron density $n(\mathbf{r})$ is the sum of the squared norm of the Kohn–Sham wave functions $\psi_i(\mathbf{r})$ weighted by the smearing function $f(\epsilon, \mu)$, (e.g. the Fermi–Dirac distribution or the Gaussian smearing function)

$$n(\mathbf{r}) = \sum_i^\infty f(\epsilon_i, \mu)\left|\psi_i(\mathbf{r})\right|^2. \tag{3}$$

The chemical potential μ is chosen such that the total number of electrons is N_e. In many density-matrix based formulations, it is useful to have the single-particle density matrix

$$\gamma(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^\infty f(\epsilon_i, \mu)\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}'), \tag{4}$$

whose diagonal is the electron density

$$n(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r}). \tag{5}$$

Minimization of the energy functional subject to the orthogonality condition

$$\int \mathrm{d}\mathbf{r}\,\psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) = \delta_{ij}, \tag{6}$$

leads to the Kohn-Sham equation,

$$h\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{7}$$

$$\nabla^2 v_{\mathrm{H}}(\mathbf{r}) = 4\pi n(\mathbf{r}) \tag{8}$$

$$v_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \tag{9}$$

where the single-particle Hamiltonian is $h = -\frac{1}{2}\nabla^2 + v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r})$. The Hartree potential v_{H} may be obtained by solving the Poisson equation (see Eq. (8)). The dimension of the Hamiltonian matrix $H_{ij} \equiv \langle\psi_i|h|\psi_j\rangle$ is the size of the basis functions M. However, due to the fast decaying property of the smearing function, only N lowest eigenstates are needed. Typically N is smaller than M by a few orders of magnitude, especially for the case of non-compact basis, such as the plane-wave basis. Diagonalization of the Hamiltonian gives eigenvalues ϵ_i's which are arranged from the smallest to the largest as $\epsilon_1 \le \epsilon_2 \le \cdots \le \epsilon_N$. This equation has to be solved iteratively due to the orbital dependence of the single-particle effective potential.

2.3. Matrix formulation

The continuous Kohn-Sham model can be conveniently expressed in matrix notations and solved on a computer. To unify the treatment of both finite and extended systems, we explicitly include the $\mathbf{k}$ dependence in the formulation. For each wave-vector $\mathbf{k}$ in the 1st Brillouin zone, due to symmetry, there is a weight $\omega_{\mathbf{k}}$ associated with it according

to the space-group of the underlying structure. Normally, Kohn-Sham eigenstates are Bloch orbitals which can be expanded in terms of a possibly non-orthogonal basis functions $\{\chi_{\mu\mathbf{k}}\}$ of size M,

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{\mu=1}^{M} C_{\mu i\mathbf{k}}\chi_{\mu\mathbf{k}}(\mathbf{r}). \tag{10}$$

Here i and $\mathbf{k}$ are the band index and the Bloch wave vector. The expansion coefficients $C_{\mu i\mathbf{k}}$ can be regarded as a matrix whose ith column contains the expansion coefficients of the ith wave function indexed by $\mathbf{k}$. For a Γ-point calculation in real basis, real coefficients can be used for memory saving and speedup. To keep it general, we have $\mathbf{C} \in \mathbb{C}^{M\times N\times K}$ (bold symbol for this size, see notations in Section 2.1). Popular choices of basis functions are plane waves [7], Gaussian orbitals [37,38], (numerical) LCAO [39], multiresolution analyses [40], or finite-difference/finite-element real-space grids [41,42].

Within a local basis set, we might express the basis functions as

$$\chi_{\mu\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} \exp(i\mathbf{k}\cdot\mathbf{R})\phi_{\mu}(\mathbf{r}-\tau_{\mu}-\mathbf{R}) \tag{11}$$

where $\phi_{\mu}(\mathbf{r}-\tau_{\mu}-\mathbf{R})$ are the atomic orbitals centering on an atom in the unit cell $\mathbf{R}$. The index μ enumerates the atomic orbitals.

The total density matrix is the sum

$$P = \sum_{\mathbf{k}} P_{\mathbf{k}}, \tag{12}$$

where the density matrix $P_{\mathbf{k}}$ in the matrix representation is

$$P_{\mathbf{k}} = \omega_{\mathbf{k}} C_{\mathbf{k}} F_{\mathbf{k}} C_{\mathbf{k}}^{\dagger} \tag{13}$$

with occupation matrix elements $F_{ij\mathbf{k}} = f(\epsilon_{i\mathbf{k}},\mu)\delta_{ij}$. The charge density can be expressed as the diagonal of the density matrix,

$$n(\mathbf{r}) = \mathrm{diag}(\langle\mathbf{r}|P|\mathbf{r}'\rangle), \tag{14}$$

in which P is evaluated on grid points. The chemical potential μ can be determined by satisfying $N_e = \int \mathrm{d}\mathbf{r}\, n(\mathbf{r})$, with N_e electrons in the unit cell.

The Kohn-Sham equation (see Eq. (7)) can be cast into a generalized matrix eigenvalue problem,

$$H_{\mathbf{k}} C_{\mathbf{k}} = E_{\mathbf{k}} S_{\mathbf{k}} C_{\mathbf{k}}, \tag{15}$$

where $H_{\mathbf{k}}$, $S_{\mathbf{k}}$, and $C_{\mathbf{k}}$ are the Hamiltonian matrix, overlap matrix and eigenvectors at a given $\mathbf{k}$-point, respectively. The energy matrix $E_{\mathbf{k}}$ is a diagonal matrix with elements $E_{ij\mathbf{k}} = \epsilon_{i\mathbf{k}}\delta_{ij}$. In the case of norm-conserving pseudopotentials, the external potential can be split into local part v_{α}^{L} and nonlocal part v_{α}^{NL}, $v_{\mathrm{ext},\alpha} = v_{\alpha}^{L} + v_{\alpha}^{NL}$. The nonlocality of pseudopotential v^{NL} is included via the standard nonlocal projectors [43],

$$v_{\alpha}^{NL} = \sum_{l=0}^{l_{\max}}\sum_{m=-l}^{l}\sum_{n=1}^{n_{\max}} |\chi_{\alpha lmn}\rangle\langle\chi_{\alpha lmn}| \tag{16}$$

where $|\chi_{\alpha lmn}\rangle$ are non-local projectors. l and m are the azimuthal and magnetic quantum numbers, respectively, and n is the multiplicity of projectors. $l_{\max}$ and $n_{\max}$ are the maximal angular momentum and the maximal multiplicity of projectors for each angular momentum channel.

The Hamiltonian matrix $H_{\mathbf{k}}$ and the overlap matrix $S_{\mathbf{k}}$ are both of size $M\times M$,

$$H_{\alpha\beta\mathbf{k}} = \int \mathrm{d}\mathbf{r}\, \chi_{\alpha\mathbf{k}}^{*}(\mathbf{r}) H_{\mathbf{k}}\, \chi_{\beta\mathbf{k}}(\mathbf{r}) \tag{17}$$

and

$$S_{\alpha\beta\mathbf{k}} = \int \mathrm{d}\mathbf{r}\, \chi_{\alpha\mathbf{k}}^{*}(\mathbf{r})\, \chi_{\beta\mathbf{k}}(\mathbf{r}). \tag{18}$$

For orthonormal basis functions, such as the plane wave basis, $S_{\mathbf{k}} = I_M$. In general, the overlap matrix $S_{\mathbf{k}}$ is a symmetric positive definite matrix, which assures a Cholesky decomposition

$$S_{\mathbf{k}} = U_{\mathbf{k}}^{\dagger} U_{\mathbf{k}}. \tag{19}$$

The orthogonal constraint imposed on the coefficient matrix reads

$$C_{\mathbf{k}}^{\dagger} S_{\mathbf{k}} C_{\mathbf{k}\dagger} = I_N. \tag{20}$$

The total energy is thus a function of the coefficient matrix $\mathbf{C}$, $E_{\mathrm{KS}}(\mathbf{C})$.

2.4. Matrix optimization with orthogonality constraints

As introduced above, instead of diagonalizing the Hamiltonian matrix, an alternative method is called the direct minimization, which was discussed in details in Ref. [7]. The standard optimization problem with orthogonality constraints is

$$\min_{X\in\mathbb{C}^{n\times p}} \mathcal{F}(X), \text{ such that } X^{\dagger}X = I_p, \tag{21}$$

where $\mathcal{F}(X) : \mathbb{C}^{n\times p} \to \mathbb{R}$ is a differentiable function. Meanwhile, the KS minimization problem becomes

$$\min_{C_{\mathbf{k}}\in\mathbb{C}^{M\times N}} E_{\mathrm{KS}}(\mathbf{C}), \text{ such that } C_{\mathbf{k}}^{\dagger} S_{\mathbf{k}} C_{\mathbf{k}} = I_N. \tag{22}$$

It can be easily verified that the KS problem can be adapted to the standard form, with the auxiliary transformation matrix $U_{\mathbf{k}}$ (see Eq. (19)),

$$X_{\mathbf{k}} = U_{\mathbf{k}} C_{\mathbf{k}} \quad \text{or} \quad C_{\mathbf{k}} = U_{\mathbf{k}}^{-1} X_{\mathbf{k}}. \tag{23}$$

The total energy can be written as a sum

$$E_{\mathrm{KS}} = E_b[P] + \Phi[n], \tag{24}$$

where

$$E_b[P] = \mathrm{tr}(\sum_{\mathbf{k}} P_{\mathbf{k}} H_{\mathbf{k}}), \tag{25}$$

is the band energy and $\Phi[n]$ is a density-dependent quantity. The derivative of the energy functional is essential in the optimization. The energy variation can be computed with the variation of the density matrix

$$\mathrm{d}E_{\mathrm{KS}} = \mathrm{tr}(\sum_{\mathbf{k}} H_{\mathbf{k}} \mathrm{d}P_{\mathbf{k}}) \tag{26}$$

where, substituting Eq. (23) into Eq. (13),

$$\begin{aligned}\mathrm{d}P_{\mathbf{k}} &= \omega_{\mathbf{k}}\left[\mathrm{d}C_{\mathbf{k}}\, F_{\mathbf{k}}\, C_{\mathbf{k}}^{\dagger} + C_{\mathbf{k}}\, F_{\mathbf{k}}\, \mathrm{d}C_{\mathbf{k}}^{\dagger}\right] \\ &= \omega_{\mathbf{k}}\left[U_{\mathbf{k}}^{-1}\left(\mathrm{d}X_{\mathbf{k}}\, F_{\mathbf{k}}\, X_{\mathbf{k}}^{\dagger} + X_{\mathbf{k}}\, F_{\mathbf{k}}\, \mathrm{d}X_{\mathbf{k}}^{\dagger}\right)(U_{\mathbf{k}}^{-1})^{\dagger}\right].\end{aligned} \tag{27}$$

Therefore, the derivative of the energy functional can be derived as follows [20]

$$(\nabla E_{\mathrm{KS}})_{\mathbf{k}} \equiv \frac{\partial E_{\mathrm{KS}}}{\partial X_{\mathbf{k}}^{\dagger}} = \omega_{\mathbf{k}}(U_{\mathbf{k}}^{-1})^{\dagger} H_{\mathbf{k}} U_{\mathbf{k}}^{-1} X_{\mathbf{k}} F_{\mathbf{k}}, \tag{28}$$

following the definition

$$\mathrm{d}E_{\mathrm{KS}} = \mathrm{tr}\left[\mathrm{d}X_{\mathbf{k}}^{\dagger}\left(\frac{\partial E_{\mathrm{KS}}}{\partial X_{\mathbf{k}}^{\dagger}}\right) + \left(\frac{\partial E_{\mathrm{KS}}}{\partial X_{\mathbf{k}}^{\dagger}}\right)^{\dagger}\mathrm{d}X_{\mathbf{k}}\right] \tag{29}$$

and the application of cyclicity of the trace $\mathrm{tr}(XY) = \mathrm{tr}(YX)$.

2.5. Complex Stiefel manifold

Classical methods for unconstrained optimization in Euclidean space, such as steepest descent, conjugate gradient, quasi-Newton methods, and trust-region methods, can be generalized to optimization on

Riemannian manifolds. For the KS problem (22), the underlying manifold is a complex Stiefel manifold, which is the space of matrices defined as

$$\mathrm{St}_n^p := \{X \in \mathbb{C}^{n\times p} : X^\dagger X = I_p\}. \tag{30}$$

We denote the manifold as St for brevity. The Stiefel manifold may be embedded in the np-dimensional Euclidean space of n-by-p matrices. When $p = 1$, the Stiefel manifold reduces to a sphere, and when $p = n$, it corresponds to the group of orthogonal matrices, known as O_n. At $X \in \mathrm{St}$, we have the tangent space

$$T_X\mathrm{St} = \left\{Y = XB + Z \mid B^\dagger = -B, Z^\dagger X = 0\right\}, \tag{31}$$

where $Y, Z \in \mathbb{C}^{n\times p}$, $B \in \mathbb{C}^{p\times p}$. Here, B is a skew-Hermitian matrix and Z is a matrix orthogonal to X.

The orthogonal projection of any vector $V \in \mathbb{C}^{n\times p}$ onto $T_X\mathrm{St}$ is

$$\pi_X(V) = V - X\,\mathrm{sym}(X^\dagger V). \tag{32}$$

There are two commonly used metrics for the tangent space: the Euclidean metric $\langle U, V\rangle_X^e = \mathrm{tr}(U^\dagger V)$ and the canonical metric

$$\langle U, V\rangle_X^c = \mathrm{tr}\left[U^\dagger\left(I_n - \frac{1}{2}XX^\dagger\right)V\right] \tag{33}$$

where $U, V \in T_X\mathrm{St}$ [25].

In Riemannian optimization, two fundamental components are needed. The first component is the "retraction", which smoothly maps a point from the tangent space to the manifold. In notations, a mapping $\mathcal{R}$ from the tangent space $T\mathrm{St}$ into St is a retraction, which satisfies

$$\mathcal{R}_X(0) = X, \forall X \in \mathrm{St}, \tag{34a}$$

$$\frac{d}{dt}\mathcal{R}_X(tZ)\,|_{t=0} = Z, \forall X \in \mathrm{St}, \forall Z \in T_X\mathrm{St}. \tag{34b}$$

Common retractions for Stiefel manifold include the exponential mapping

$$\mathcal{R}_X^{exp}(U) = (X \quad U)\left(\exp\begin{pmatrix} A & -S \\ I_p & A\end{pmatrix}\right)\begin{pmatrix} I_p \\ 0\end{pmatrix}\exp(-A) \tag{35}$$

where $X \in \mathrm{St}, U \in T_X\mathrm{St}, A = X^\dagger U$, and $S = U^\dagger U$ [25]. This retraction requires geodesics along the manifold, where matrix exponential is required and thus computationally expensive. In contrast, projection-like retractions such as the QR decomposition can be viewed as the first-order approximations to the exponential one, which is preferred in many practical applications. The QR decomposition retraction is

$$\mathcal{R}_X(U) = \mathrm{qf}(X + U) \tag{36}$$

where $\mathrm{qf}(\cdot)$ denotes the Q factor of the QR decomposition with non-negative elements on the diagonal of R. The polar decomposition

$$\mathcal{R}_X(U) = (X + U)\left(I_p + U^\dagger U\right)^{-1/2} \tag{37}$$

is another common retraction choice, which is second-order.

The second component is the "vector transport", which transfers a vector from the tangent space of an adjacent point to the same tangent space at a given point. It is a computationally affordable approximation to the "parallel transport". This is essential in the optimization approaches such as the conjugate gradient method, because otherwise vectors from different tangent spaces are not directly computable. The vector transport by projection is denoted by $\mathcal{T}^P$, as in Eq. (32),

$$\mathcal{T}_U^P(V) = V - Y\,\mathrm{sym}(Y^\dagger V) \tag{38}$$

where $U, V \in T_X\mathrm{St}, Y = \mathcal{R}_X(U)$, and $\mathcal{R}$ is the associated retraction. Alternatively, the vector transport by differentiated retraction $\mathcal{T}_U^R(V)$ could be used accordingly [44].

2.6. Riemannian conjugate gradient methods

Conjugate gradient methods offer significant advantages by efficiently handling the curvature and geometric structure of the manifold, requiring low memory, and achieving faster convergence. They avoid the need for matrix inversions and are adaptable with retraction methods, making them versatile and powerful tools for manifold-based optimization problems in various scientific and engineering applications.

Similar to the Euclidean case, the Riemannian conjugate gradient (RCG) methods require the essential ingredient, the Riemannian gradient $g = \mathrm{grad}\, f$. In the Stiefel manifold, it can be computed with the Euclidean gradient ∇f

$$\mathrm{grad}\, f = \nabla f - X\left(\nabla f\right)^\dagger X. \tag{39}$$

Keeping the conjugacy, the new search direction d_{k+1} is computed as

$$d_{k+1} = -g_{k+1} + \beta_{k+1}\mathcal{T}_{\alpha_k d_k}(d_k) \tag{40}$$

where g_k denotes the gradient at iteration k and $\mathcal{T}_{\alpha_k d_k}(d_k)$ in this work is the projection base formula in Eq. (38). In RCG methods, the parameter β_{k+1} for the conjugate gradient direction in each iteration can take on various forms. To facilitate this, it is often convenient to define the quantity:

$$y_{k+1} = g_{k+1} - \mathcal{T}_{\alpha_k d_k}(g_k) \tag{41}$$

With this, schemes by Fletcher-Reeves [45], Polak-Ribière [46] and Polyak [47], Dai-Yuan [48], Hestenes-Stiefel [49], Liu-Storey [50], Hager-Zhang [51] are commonly adopted algorithms. Some examples of these adapted forms are listed below.

$$\beta_{k+1}^{\mathrm{FR}} = \frac{\langle g_{k+1}, g_{k+1}\rangle_{X_{k+1}}}{\langle g_k, g_k\rangle_{X_k}}, \tag{42a}$$

$$\beta_{k+1}^{\mathrm{PR-P}} = \frac{\langle g_{k+1}, y_{k+1}\rangle_{X_{k+1}}}{\langle g_k, g_k\rangle_{X_k}}, \tag{42b}$$

$$\beta_{k+1}^{\mathrm{DY}} = \frac{\langle g_{k+1}, g_{k+1}\rangle_{X_{k+1}}}{\langle y_{k+1}, \mathcal{T}_{\alpha_k d_k}(d_k)\rangle_{X_{k+1}}}, \tag{42c}$$

$$\beta_{k+1}^{\mathrm{HS}} = \frac{\langle g_{k+1}, y_{k+1}\rangle_{X_{k+1}}}{\langle y_{k+1}, \mathcal{T}_{\alpha_k d_k}(d_k)\rangle_{X_{k+1}}}. \tag{42d}$$

The canonical metric in Eq. (33) is adopted in computing the vector product, e.g. $\langle y_{k+1}, \mathcal{T}_{\alpha_k d_k}(d_k)\rangle_{X_{k+1}}$ and hence all superscripts c is omitted. With these, the RCG method is summarized in Algorithm 1. In this algorithm, we give an example of QR decomposition retraction in the language of linear algebra operations. For $\mathcal{R}_{X_k}(\alpha_k d_k) = \mathrm{qf}(X_k + \alpha_k d_k)$, $\mathrm{qf}(X_k + \alpha_k d_k)$ is the Q factor of the QR decomposition of $X_k + \alpha_k d_k$. For other retractions and vector transports, one needs to apply the corresponding linear algebra operations in Table 1.

For multiple **k**-points, all quantities are indexed by $k_i = 1, 2, \cdots, K$, where K is its total size. Therefore, the concept of the product of manifolds naturally fits into the description. A product of Stiefel manifolds is denoted by $\mathrm{St} = \mathrm{St}_1 \times \mathrm{St}_2 \times \cdots \times \mathrm{St}_K$, where St_i is a sub-manifold. An element $\mathbf{X}$ in St is denoted by

$$\mathbf{X} = \left(X_1^T, X_2^T, \cdots, X_K^T\right)^T, \tag{43}$$

where $X_i \in \mathrm{St}_i$. The tangent space of St is

$$T_X\mathrm{St} = T_{X_1}\mathrm{St}_1 \times T_{X_2}\mathrm{St}_2 \times \cdots \times T_{X_K}\mathrm{St}_K, \tag{44}$$

whose norm is thus the sum of all the norms of each sub-manifold.

To apply to the multiple **k**-points cases, one simply expands the dimension of the pertinent X, by stacking K copies of X, each of the same size. In the algorithm, one needs to adapt X and g into bold symbols $\mathbf{X}$ and $\mathbf{g}$ (see above). For each **k**-point, taking care of the orthogonal constraint $X_{k_i}^\dagger X_{k_i} = I_p$, the related operations in the algorithm have to be performed within the corresponding submanifold. The modification to

Table 1
Example of key linear algebra operations for Riemannian optimization on the Stiefel manifold. Note, $Y = \mathcal{R}_X(U)$ in the vector transport formula.

Operation	Linear Algebra Formula	Key LAPACK Routine
Euclidean metric	$\langle U,V\rangle_X^e = \mathrm{tr}[U^\dagger V]$	ZGEMM
Canonical metric	$\langle U,V\rangle_X^c = \mathrm{tr}[U^\dagger(I-\frac{1}{2}XX^\dagger)V]$	ZGEMM
Riemannian gradient	$\mathrm{grad}\, f(X) = \nabla f - X(\nabla f)^\dagger X$	ZGEMM
QR retraction	$\mathcal{R}_X(U) = \mathrm{qf}(X+U)$	ZGEQRF
Polar retraction	$\mathcal{R}_X(U) = (X+U)(I_p + U^\dagger U)^{-1/2}$	ZHEEV
Vector transport by projection	$\mathcal{T}_U(V) = V - Y\,\mathrm{sym}(Y^\dagger V)$	ZGEMM
Tangent space projection	$\pi_X(V) = V - X\,\mathrm{sym}(X^\dagger V)$	ZGEMM

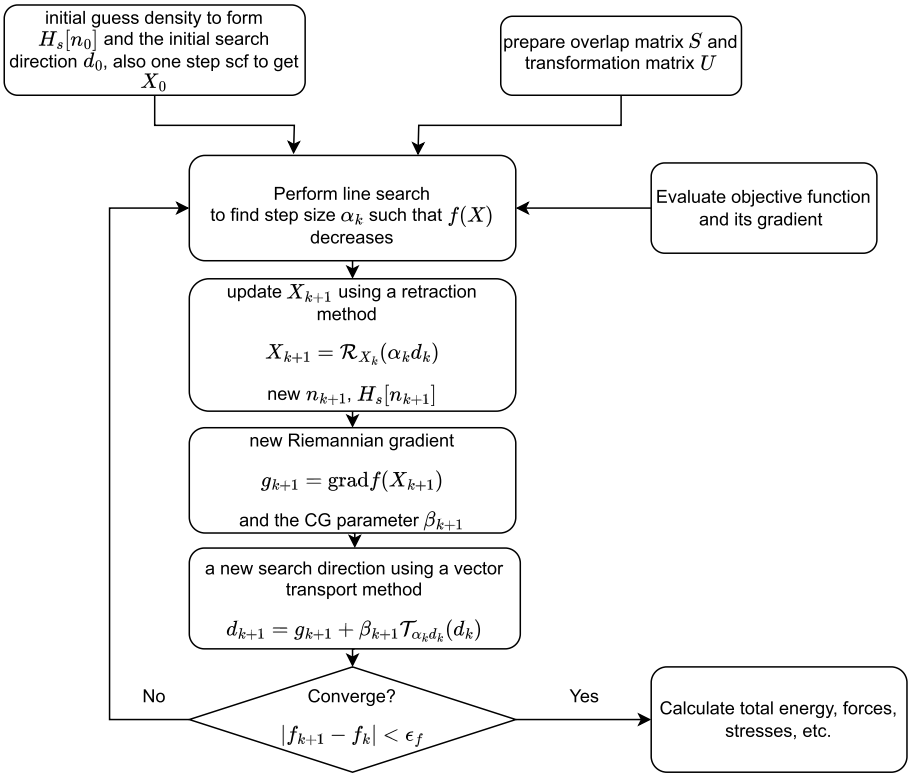

Fig. 1. The flowchart for the RCG method.

the norm is to use the proper norm for the product of manifolds (e.g. the maximum norm in Ref. [34]) when computing the CG parameters.

Algorithm 1 Conjugate gradient method for minimizing $f(X)$ on the Stiefel manifold.

1: Initialization: choose $X_0 \in \mathrm{St}$, $\epsilon_g, \epsilon_f > 0$, $k_{\max}$, $g_0 = \mathrm{grad}\, f(X_0)$, initial search direction $d_0 = g_0$
2: **while** $\|g_k\| > \epsilon_g$ (or $|f_{k+1} - f_k| > \epsilon_f$) **and** $k < k_{\max}$ **do**
3: Line search to find step size $\alpha_k > 0$, and update the point $X_{k+1} \leftarrow \mathcal{R}_{X_k}(\alpha_k d_k)$ using Eq. (36)
4: Compute new Riemannian gradient $g_{k+1} \leftarrow \mathrm{grad}\, f(X_{k+1})$ using Eq. (39) and the conjugate direction parameter β_{k+1} (e.g. using the FR scheme)

$$\beta_{k+1} \leftarrow \frac{\langle g_{k+1}, g_{k+1}\rangle_{X_{k+1}}}{\langle g_k, g_k\rangle_{X_k}}$$

5: Compute a search direction as $d_{k+1} \leftarrow -g_{k+1} + \beta_{k+1}\mathcal{T}_{\alpha_k d_k}(d_k)$ using Eq. (38)
6: $k \leftarrow k+1$
7: **end while**

More graphically, we represent this algorithm in the following flowchart (Fig. 1). The most time consuming part is in the line search part where function and its gradient evaluations are required to find the right step size. In contrast, the standard SCF method is rather different (see Fig. B.5 in Appendix B). It requires direct (or iterative) diagonalization of the Hamiltonian matrix and a proper density mixing scheme.

3. Implementation details

As a first step, we have implemented the conjugate gradient direct minimization on the complex Stiefel manifold algorithm within the open-source ABACUS software [52,39], which uses norm-conserving pseudo potentials to describe the interactions between nuclear ions and valence electrons. Currently, we have only implemented the numerical atomic basis set calculations. The same algorithm can be easily adapted to the plane-wave basis, which ABACUS also supports.

Taking advantage of the modular structure, the whole algorithm is integrated as a new inherited solver. In this solver, typical conjugate gradient schemes (42) can be chosen via variables within the input. Choices of retraction defaults to the projection type and the vector transport by projection $\mathcal{T}^P$ are used. The step length is chosen such that it satisfies the strong Wolfe conditions [44]

$$f\left(\mathcal{R}_{X_k}\left(\alpha_k d_k\right)\right) \leq f\left(X_k\right) + c_1\alpha_k\left\langle \nabla f\left(X_k\right), d_k\right\rangle_{X_k} \tag{45}$$

and

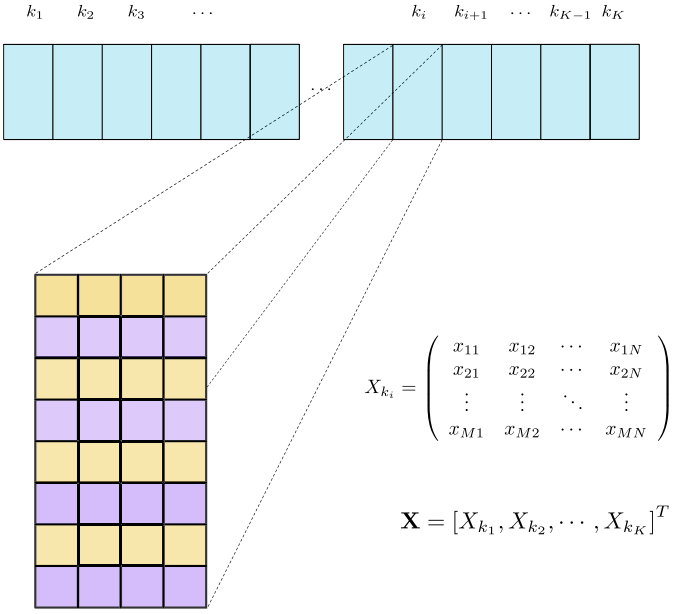

Fig. 2. Data structure for **k**-dependent variable **X** where each component of the vector indexed by k_i lives in a sub-manifold. Each matrix X_k is a $M \times N$ complex matrix. Same structure applies to **H** and **F**.

$$\left|\left\langle \operatorname{grad} f\left(\mathcal{R}_{X_k}\left(\alpha_k d_k\right)\right), \mathcal{T}_{\alpha_k d_k}(d_k)\right\rangle_{\mathcal{R}_{X_k}(\alpha_k d_k)}\right| \leq c_2\left|\left\langle \operatorname{grad} f\left(X_k\right), d_k\right\rangle_{X_k}\right| \tag{46}$$

where $0 < c_1 < c_2 < 1$. Default values $c_1 = 10^{-4}, c_2 = 0.9$ are used. The initial step length $\alpha_0 = 1.0$ is used as default and can be modified in the input. The line search begins with a trial estimate α^t, and keeps increasing the step length until it finds either an acceptable step length or an interval that brackets the desired step lengths. Once such an interval is established, the *zoom* algorithm, combined with cubic interpolation, is employed to refine the step length. This bracketing process continues until an acceptable step length is determined [8]. We have not done any preconditioning to speed up the convergence yet. An initial guess for the orbitals is taken to be the eigenvectors of the Hamiltonian obtained from a superposition of atomic densities. The implementation can be found in Ref. [53].

Due to the **k** dependence in the orbitals $X_{\mathbf{k}}$, the dimension of the one-electron wave-function is of size $M \times N \times K$, where K denotes the number of irreducible **k**-points in the 1st Brillouin zone. For this, we use the concept of product of Stiefel manifolds as a single entity. Therefore, for each **k**, there is a sub-manifold associated with it. The data structure is sketched in Fig. 2. To assess the feasibility for varying K, the norm of **X** is defined as the sum of norms in each sub-manifold divided by the number of sub-manifolds, K. In other words, we take the average. In this way, its norm $\|\mathbf{X}\|$ should be close to 1 for varying K.

The structure is very similar to a typical representation of wave-functions. What is needed in Algorithm 1 is to empower the linear algebra operations to them. This could easily be realized with some open-source linear algebra libraries, such as *armadillo* [54] or *Eigen*[55] in C++ language. In our implementation, we used a vector of size K of complex matrices in *armadillo* to represent **X**. Each vector element is of size $M \times N$. The linear algebra operations are performed independently within each sub-manifold.

To form $F_{\mathbf{k}}$, the occupation number $f_{i\mathbf{k}}$ can be computed with only the diagonal elements of $C_{\mathbf{k}}^{\dagger} H_{\mathbf{k}} C_{\mathbf{k}}$. The chemical potential μ is determined by bisection method according to a smearing scheme such as the Fermi-Dirac smearing or the Gaussian smearing.

4. Results

All calculations were performed on a workstation with an Intel(R) Comet Lake Processor (at 3.80 GHz×8, 16 MB cache). The total number of cores is 8 and the total number of threads is 16. All codes were compiled with the Intel oneAPI compilers on Debian 12. To ensure fair comparisons, multi-threading was disabled, and only a single core was utilized for all computations.

4.1. Test problems

To test the RCG algorithm, we have applied the optimization procedure to two simple problems. The argument $X_k \in \mathbb{C}^{n\times p}$ in both problems is subject to the orthogonality constraint $X_k^{\dagger} X_k = I_p$. To mimic the **k** dependence in the electronic structure theory for periodic systems, we have extended these with a **k** dependence of minimum size 2, namely $K = 2$.

The first problem is the orthogonal Procrustes problem. In this problem, the objective function is $f(X) = \|AX - B\|_F$ and its gradient can be computed analytically $\frac{\partial f}{\partial X^{\dagger}} = AX - B$. The second problem is the eigenvalue problem, whose objective function is $f(X) = -\frac{1}{2}\operatorname{tr}\left(X^{\dagger} E X\right)$ and its gradient is $\frac{\partial f}{\partial X^{\dagger}} = -EX$. For each k, $A_k \in \mathbb{C}^{m\times n}$, $B_k \in \mathbb{C}^{m\times p}$, and $E_k \in \mathbb{C}^{n\times n}$. Setting random A and $B = AI$, an initial guess is chosen as the known solution I plus a perturbed random deviation $X_0 = I + \epsilon P$, where ϵ is a small number and P is a random matrix, the algorithm successfully finds the solution. Similarly, setting E as a Hermitian matrix, the algorithm also delivers the right solution against the standard eigensolver.

4.2. Molecules and solids

To identify the advantages and disadvantages of the RCG method, we performed single point ground state energy calculations for G2 data set of small molecules [56,57], and a few simple bulk solids. As a comparison, we have also provided a tentative implementation of the Riemannian BFGS (RBFGS) method (see detailed algorithm in Appendix A), which has to be considered preliminary (some more cases fail to converge). Structure files were converted to ABACUS format using the utility code *atomkit*. For molecules, a cubic supercell of length 15.0 Angstrom with Γ-point is always used and periodic boundary conditions are imposed. For solids, Monkhorst-Pack k-point meshes were generated with a spacing of 0.06 in the unit of $2\pi/$Å with *atomkit* as well. Norm-conserving pseudopotentials (optimized ONCV) [58] were used to describe the electron-ion interactions. The nonlocality of pseudopotential is included via the standard nonlocal projectors [43]. Double-ζ plus polarization function (DZP) basis set and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional were used [59]. The default SCF algorithm uses the Broyden density mixing of dimension 8 [60], with mixing parameter 0.8 for spin-unpolarized calculations. The convergence is reached when the relative density error, $\Delta\rho_R = \frac{1}{N_e}\int |\Delta\rho(r)| d^3 r$, is less than 10^{-6}. By default, the Kerker preconditioner is also turned on. The direct minimization stops when the change of the total energy between consecutive iterations is less than 5×10^{-9} a.u. Note here we use $|f_{k+1} - f_k| < \epsilon_f$ as the stopping criterion in Algorithm 1. After numerous tests, we found that the function tolerance ϵ_f is more effective in controlling the convergence than the gradient tolerance ϵ_g.

To begin with, we implemented four conjugate gradient schemes listed in Eq. (42). For a test molecule acetonitrile CH_3CN, the converged energy from the SCF calculation is taken as the minimum energy. The error in terms of iteration is shown on a log scale. It can be seen that all schemes have the super-linear convergence. For this case, Dai-Yuan and Polak-Rieère-Polyak are slightly faster in reaching the minimum. It is also observed that Dai-Yuan runs faster than others in many cases (see

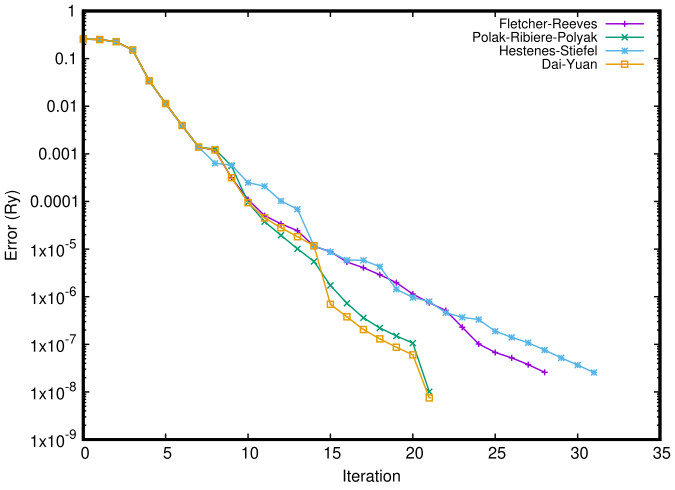

Fig. 3. Convergence rate comparison among four common conjugate gradient schemes for RCG calculation on acetonitrile molecule, CH_3CN. Error is taken with respect to the converged energy of SCF and is displayed on a log scale. Among these schemes, the Dai-Yuan method consistently consumes the least time in most scenarios.

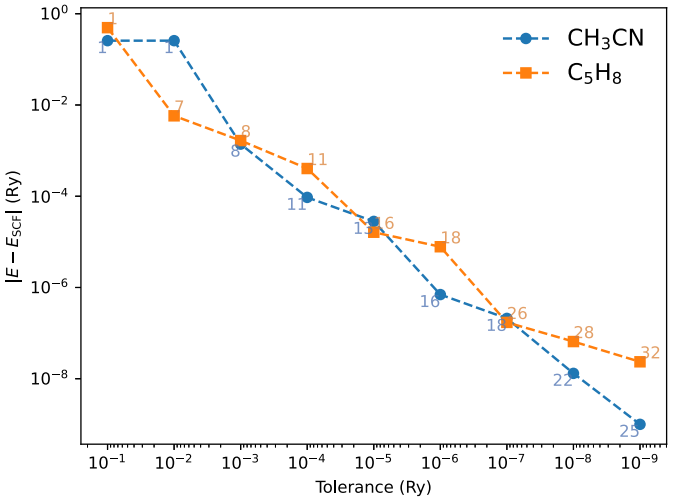

Fig. 4. The error of the ground state energy with respect to the different tolerances for CH_3CN and C_5H_8. E_{SCF} is the converged energy from SCF calculation. The number of iterations is shown beside the marker.

Fig. 3). Therefore, all our calculations used Dai-Yuan scheme in computing β_k.

Additionally, to examine how the error in the ground state energy changes based on different stopping criteria and tolerance levels in RCG, we present the results in Fig. 4 for the acetonitrile (CH_3CN) and spiropentane (C_5H_8) molecules. We observe that a smaller tolerance ϵ_f leads to a smaller error in the ground state energy, which aligns with our expectations. The number of iterations increases when the tolerance decreases, as anticipated.

Calculations for all 148 molecules in the G2 dataset are successfully converged using the SCF method. However, for some molecules, the minimization process of the RCG method halted after the first few iterations, resulting in higher energies compared to the SCF calculations. This issue may stem from the initial gradient being too small due to the atomic density initialization. For nearly all the remaining 134 molecules, the RCG method yielded total energies with an error of less than 0.003 meV (mostly 0.001 meV) compared to the SCF energies.

The number of iterations of RCG is compared to that of RBFGS. In Table 2, the statistics show that the RCG method has larger variations in the number of iterations. The average iterations of RBFGS are less than the RCG method. From these tests, the RBFGS method should be the method of choice.

Table 2
Statistics for the total number of iterations for RCG and RBFGS methods for molecules. Avg. and Std. stand for the mean and the standard deviation.

Iterations	RCG	RBFGS
Avg.	26.5	16.1
Std.	13.6	4.2
Min	7	4
Max	84	29

The Hamiltonian is updated in both real space and reciprocal space, a process that involves computationally expensive numerical grid integrals and Fourier transforms. This step is common to both the direct minimization method and the iterative diagonalization method, resulting in similar computational costs per iteration. During the line search, the RCG and RBFGS method may require a couple of calls in evaluating the objective function and its gradient. In contrast, for small system sizes, such as those in the G2 dataset, the manifold-related operations (such as metric evaluation, retraction, and vector transport) in these methods account for only a small fraction of the overall computational time.

However, for the solids (Cu, LiF, Mg, MgO, NaCl, and SiC) tested, the computational time required grows rapidly with respect to the number of **k**-points. In contrast, the SCF method is less sensitive to the number of **k**-points. In this regard, the SCF method is more efficient than the RCG and RBFGS methods. We anticipate that the efficiency and robustness of direct minimization methods can be further enhanced through continuous effort on refinement of the implementation.

5. Discussions and conclusion

Direct minimization method on the complex Stiefel manifold in Kohn-Sham density functional theory is formulated to treat both finite and extended systems in a unified manner. Utilizing the product of Stiefel manifolds, we have demonstrated the feasibility of direct minimization calculations with a line search method for both finite and extended systems. In our pilot implementation of the conjugate gradient method and tentative version of the BFGS method on the complex Stiefel manifold within a compact basis set, we conducted comparison tests to reveal advantage and disadvantages of the Riemannian methods.

Without invoking any preconditioning in the manifold minimization method, we show it can deal with both finite systems compared to the standard SCF method. In fact, for finite systems with Γ-point calculation, it is not necessary to use complex Stiefel manifold. Reverting back to the real case, further speed-up is guaranteed by the dimension reduction. Unfortunately, it is rather slow for extended systems. The slow convergence problem for periodic systems should be alleviated with better preconditioning. Building upon the linear algebra formulation, the implementation of other second-order Riemannian optimization methods, such as trust-region methods, may potentially compete against the SCF method.

Currently, the default retraction and vector transport methods are projection-based. The main computational bottleneck for the KSDFT problem lies in the evaluation of the objective function and its gradient. For a compact basis set, the performance degradation is negligible even when using exact geodesic retraction and parallel transport. However, for future implementations involving a non-compact basis (such as a plane-wave basis), geodesic retraction may become increasingly demanding due to the necessity of evaluating the matrix exponential. Similarly, the recent Exponential Transformation Direct Minimization (ETDM) method [24] might also face challenges under these conditions. In such scenarios, the economical retraction and vector transport methods will demonstrate their superiority, offering a more efficient alternative without compromising performance. It is worth pointing out

that in the matrix exponential such as ETDM, the number of elements in the exponent to be optimized is $M(M+1)$ for each **k**-point, much larger than this work, which is $2MN$ when $M \gg N$. Normally, the associated operations scales as $O(MN^2)$, much faster than $O(M^3)$.

The framework laid out in this work offers potential conveniences for various other electronic structure problems with orthogonality constraints. For instance, the self-interaction corrected functional can be directly tested. Currently, we are actively investigating the implementation of reduced density matrix functional, where additional degrees of freedom, such as the natural occupation number, need to be optimized. For these computationally intensive objectives, simultaneous optimization should be employed to achieve more efficient calculations. Further research is needed to determine which method—iterative diagonalization or direct minimization—offers greater efficiency in RDMFT for periodic systems.

CRediT authorship contribution statement

Kai Luo: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Tingguang Wang:** Writing – review & editing, Validation, Investigation, Data curation. **Xinguo Ren:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

K. Luo and T.G. Wang were funded by the National Natural Science Foundation of China under Grant No. 12104230. X.G. Ren was funded by the National Natural Science Foundation of China under Grant No. 12134012, 12374067, and 12188101.

The authors express their gratitude to Tao Yan and Shimin Zhao for their invaluable explanations of nuanced details in Riemannian optimization. The valuable guidance provided by Daye Zheng on the data structure of ABACUS has been immensely beneficial. His insights have greatly enhanced our understanding and efficiency in working with the package, contributing significantly to the overall progress of our project. Additionally, they thank Michael Ulbrich and his colleagues for providing the modified KSSOLV code for ensemble DFT calculations. The authors are also grateful to Alan S. Edelman, Ross A. Lippert, Steven T. Smith, and Eric J. Bylaska for their generous insights. This work has greatly benefited from the insightful comments and suggestions of anonymous referees.

Appendix A. Riemannian BFGS method

We present the structure of the RBFGS algorithm [61,62] in Algorithm 2, where the search direction at kth iteration is obtained by

$$d_k = -B_k^{-1}\,\mathrm{grad}\, f(X_k), \tag{A.1}$$

with B_k^{-1} being a linear operator on $T_{X_k}\mathrm{St}$, which approximates the action of the inverse Hessian along the gradient $\mathrm{grad}\, f(X_k)$. The symmetry and positive definiteness of B_{k+1} [62] are ensured by

$$B_{k+1} = \begin{cases} \tilde{B}_k - \dfrac{\tilde{B}_k s_k (\tilde{B}_k^* s_k)^\dagger}{\left\langle \tilde{B}_k^* s_k, s_k \right\rangle} + \dfrac{y_k y_k^\dagger}{\langle y_k, s_k\rangle}, & \dfrac{\langle y_k, s_k\rangle}{\|s_k\|^2} \geq \vartheta(\|\,\mathrm{grad}\, f(X_k)\,\|), \\ \tilde{B}_k & \text{otherwise}, \end{cases} \tag{A.2}$$

where $\tilde{B}_k = \mathcal{T}_{\alpha_k d_k} \circ B_k \circ \left(\mathcal{T}_{\alpha_k d_k}\right)^{-1}$, $y_k = \beta_k^{-1}\,\mathrm{grad}\, f(X_{k+1}) - \mathcal{T}_{\alpha_k d_k}\left(\mathrm{grad}\, f(X_k)\right)$, $s_k = \mathcal{T}_{\alpha_k d_k}\left(\alpha_k d_k\right)$, $\beta_k = \frac{\|\alpha_k d_k\|}{\|\mathcal{T}^R_{\alpha_k d_k}(\alpha_k d_k)\|}$, ϑ is a function that strictly increasing at 0 and satisfying $\vartheta(0) = 0$, for instance, we can set $\vartheta(t) = 10^{-4}t$. Here, the symbol $\circ$ denotes operator composition (or operation composition), meaning that two operations are applied sequentially to an object [63,26]. $\mathcal{T}$ is vector transport by projection that its inverse is equal to its adjoint, $\mathcal{T}^R$ denotes the vector transport by differentiated retraction. Let A be a linear operator on $T_X\mathrm{St}$, A^* denotes the adjoint operator of A. When the retraction is based on the QR decomposition, for any $Z, U \in T_X\mathrm{St}$, $\mathcal{T}^R$ is [26]

$$\begin{aligned} \mathcal{T}^R_Z(U) = {} & \mathcal{R}_X(Z)\rho_{\mathrm{skew}}\left(\mathcal{R}_X(Z)^\dagger U \mathcal{R}_X(Z)^\dagger (X+Z)^{-1}\right) \\ & + \left(I - \mathcal{R}_X(Z)\mathcal{R}_X(Z)^\dagger\right) U \left(\mathcal{R}_X(Z)^\dagger (X+Z)^{-1}\right), \end{aligned} \tag{A.3}$$

$$(\rho_{\mathrm{skew}}(A))_{ij} = \begin{cases} A_{ij}, & \text{if } i > j, \\ 0, & \text{if } i = j, \\ -A_{ji}, & \text{if } i < j, \end{cases} \tag{A.4}$$

where $\mathcal{R}_X(\cdot)$ is the QR decomposition retraction in Eq. (36), $X \in \mathrm{St}$.

In the algorithm, the inverse Hessian approximation $H_k = B_k^{-1}$ is used instead of B_k, and the update formula for H_k [62] is

$$H_{k+1} = \tilde{H}_k - \frac{\tilde{H}_k y_k, \left(\tilde{H}_k^* y_k\right)^\dagger}{\langle \tilde{H}_k^* y_k, y_k\rangle} + \frac{s_k s_k^\dagger}{\langle s_k, y_k\rangle}, \quad \tilde{H}_k = \mathcal{T}_{\alpha_k d_k} \circ H_k \circ \left(\mathcal{T}_{\alpha_k d_k}\right)^{-1}, \tag{A.5}$$

thus, $d_k = -H_k\,\mathrm{grad}\, f(X_k)$, and computing H_k is easier than calculating the inverse of B_k. H_i and $\tilde{H}_i$ refer to the approximation to the inverse Hessian matrix, not the Hamiltonian matrix. Meanwhile, one has to choose the initial inverse Hessian approximation H_0. The widely used one is the scaled identity matrix

$$H_0 = \gamma I \tag{A.6}$$

where γ is a positive scalar and I is the identity matrix. One can either utilize gradient information or problem specific estimate of Hessian to initialize H_0. In this work, we used the most naïve choice of $\gamma = 1$. For multiple **k**-points, analogous extension to Algorithm 1 should be applied.

Algorithm 2 BFGS method for minimizing $f(X)$ on the Stiefel manifold.

1: Initialization: choose $X_0 \in \mathrm{St}$, $\epsilon_g, \epsilon_f > 0$, $k_{\max}$, $g_0 = \mathrm{grad}\, f(X_0)$, initial inverse Hessian approximation $H_0 = I$ that is symmetric positive definite with respect to the metric in Eq. (33)
2: **while** $\|g_k\| > \epsilon_g$ (or $|f_{k+1} - f_k| > \epsilon_f$) **and** $k < k_{\max}$ **do**
3: Compute a direction as $d_k \leftarrow -H_k g_k$
4: Line search to find step size $\alpha_k > 0$, and update the point $X_{k+1} \leftarrow \mathcal{R}_{X_k}(\alpha_k d_k)$ using Eq. (36)
5: Compute new Riemannian gradient $g_{k+1} \leftarrow \mathrm{grad}\, f(X_{k+1})$ and inverse Hessian approximation H_{k+1} using Eq. (A.5)
6: $k \leftarrow k+1$
7: **end while**

Since the retraction based on the QR decomposition and vector transport $\mathcal{T}$ do not satisfy the locking condition [61]

$$\mathcal{T}_\xi(\xi) = \beta \mathcal{T}^R_\xi(\xi), \ \beta = \frac{\|\,\xi\,\|}{\|\,\mathcal{T}^R_\xi(\xi)\,\|}, \tag{A.7}$$

the vector transport $\mathcal{T}$ needs to be modified as

$$\mathcal{T}_d(\xi) = \left(I - \frac{2v_2 v_2^\dagger}{\langle v_2, v_2\rangle}\right)\left(I - \frac{2v_1 v_1^\dagger}{\langle v_1, v_1\rangle}\right)\mathcal{T}_d(\xi), \tag{A.8}$$

where $v_1 = \xi_1 - \omega, v_2 = \omega - \xi_2, \xi_1 = \mathcal{T}_d(d), \xi_2 = \beta\mathcal{T}^R_d(d), Y = \mathcal{R}_X(d)$, d denotes search direction, $\xi \in T_X\mathrm{St}$, ω could be any vector in tangent space $T_Y\mathrm{St}$ that satisfies $\|\omega\| = \|\xi_1\| = \|\xi_2\|$, and we take $\omega = -\xi_1$. Of

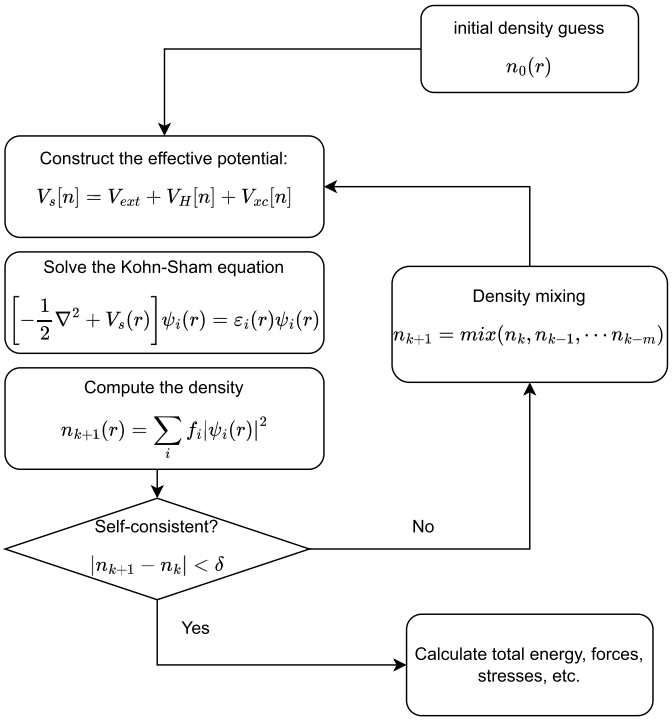

Fig. B.5. The flowchart for the SCF method. Different convergence criteria can be chosen.

course, there are other methods to ensure the locking condition [61]. In our computation, we use the modified vector transport to replace the original one. Although $\tilde{H}_k = \mathcal{T}_{\alpha_k d_k} \circ H_k \circ \left(\mathcal{T}_{\alpha_k d_k}\right)^{-1}$ is theoretically preferred, our calculations indicate that using $\tilde{H}_k = \mathcal{T}_{\alpha_k d_k} \circ H_k \circ \left(\mathcal{T}_{\alpha_k d_k}\right)^{-1}$ to compute the approximate inverse Hessian matrix results in a significant increase in the number of iterations and time consumption compared to directly using $\tilde{H}_k = H_k$. Additionally, for systems that cannot converge to the SCF calculation results with the latter approach, the former approach also does not lead to convergence. Thus, we adopted $\tilde{H}_k = H_k$.

Appendix B. Flowchart of the SCF method

To fully illustrate the difference between the SCF method and proposed RCG method, a flowchart of the SCF method is attached, which can be compared against Fig. 1.

Data availability

Data will be made available on request.

References

[1] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. 136 (1964) B864–B871, https://doi.org/10.1103/PhysRev.136.B864, https://link.aps.org/doi/10.1103/PhysRev.136.B864.

[2] W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. 140 (1965) A1133–A1138, https://doi.org/10.1103/PhysRev.140.A1133, https://link.aps.org/doi/10.1103/PhysRev.140.A1133.

[3] C.G. Broyden, A class of methods for solving nonlinear simultaneous equations, Math. Comput. 19 (92) (1965) 577–593.

[4] P. Pulay, Convergence acceleration of iterative sequences. The case of scf iteration, Chem. Phys. Lett. 73 (2) (1980) 393–398, https://doi.org/10.1016/0009-2614(80)80396-4, https://www.sciencedirect.com/science/article/pii/0009261480803964.

[5] A.C. Vaucher, M. Reiher, Steering orbital optimization out of local minima and saddle points toward lower energy, J. Chem. Theory Comput. 13 (3) (2017) 28207264, https://doi.org/10.1021/acs.jctc.7b00011.

[6] M.J. Gillan, Calculation of the vacancy formation energy in aluminium, J. Phys. Condens. Matter 1 (4) (1989) 689, https://doi.org/10.1088/0953-8984/1/4/005.

[7] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, J.D. Joannopoulos, Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients, Rev. Mod. Phys. 64 (1992) 1045–1097, https://doi.org/10.1103/RevModPhys.64.1045.

[8] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 1999.

[9] D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, 2014.

[10] Z. Wen, C. Yang, X. Liu, Y. Zhang, Trace-penalty minimization for large-scale eigenspace computation, J. Sci. Comput. 66 (3) (2016) 1175–1203, https://doi.org/10.1007/s10915-015-0061-0.

[11] N. Xiao, X. Liu, Y.-x. Yuan, Exact penalty function for $\ell_{2,1}$ norm minimization over the Stiefel manifold, SIAM J. Optim. 31 (4) (2021) 3097–3126, https://doi.org/10.1137/20M1354313.

[12] M.R. Hestenes, Multiplier and gradient methods, J. Optim. Theory Appl. 4 (5) (1969) 303–320, https://doi.org/10.1007/BF00927673.

[13] M.J. Powell, A method for nonlinear constraints in minimization problems, Optimization (1969) 283–298.

[14] B. Gao, X. Liu, Y.-x. Yuan, Parallelizable algorithms for optimization problems with orthogonality constraints, SIAM J. Sci. Comput. 41 (3) (2019) A1949–A1983, https://doi.org/10.1137/18M1221679.

[15] B. Gao, G. Hu, Y. Kuang, X. Liu, An orthogonalization-free parallelizable framework for all-electron calculations in density functional theory, SIAM J. Sci. Comput. 44 (3) (2022) B723–B745, https://doi.org/10.1137/20M1355884.

[16] J. Douady, Y. Ellinger, R. Subra, B. Levy, Exponential transformation of molecular orbitals: a quadratically convergent SCF procedure. I. General formulation and application to closed-shell ground states, J. Chem. Phys. 72 (3) (1980) 1452–1462, https://doi.org/10.1063/1.439369, https://pubs.aip.org/aip/jcp/article-pdf/72/3/1452/18921303/1452_1_online.pdf.

[17] J.F. Rico, J.M.G. De La Vega, J.I.F. Alonso, P. Fantucci, Restricted Hartree–Fock approximation. i. Techniques for the energy minimization, J. Comput. Chem. 4 (1) (1983) 33–40, https://doi.org/10.1002/jcc.540040106, https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.540040106, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540040106.

[18] J.F. Rico, M. Paniagua, J.I.F. Alonso, P. Fantucci, Restricted Hartree–Fock approximation. ii. Computational aspects of the direct minimization procedure, J. Comput. Chem. 4 (1) (1983) 41–47, https://doi.org/10.1002/jcc.540040107, https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.540040107, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540040107.

[19] M. Head-Gordon, J.A. Pople, Optimization of wave function and geometry in the finite basis Hartree-Fock method, J. Phys. Chem. 92 (11) (1988) 3063–3069, https://doi.org/10.1021/j100322a012.

[20] S. Ismail-Beigi, T. Arias, New algebraic formulation of density functional calculation, Comput. Phys. Commun. 128 (1) (2000) 1–45, https://doi.org/10.1016/S0010-4655(00)00072-2, https://www.sciencedirect.com/science/article/pii/S0010465500000722.

[21] T.V. Voorhis, M. Head-Gordon, A geometric approach to direct minimization, Mol. Phys. 100 (11) (2002) 1713–1721, https://doi.org/10.1080/00268970110103642.

[22] J. VandeVondele, J. Hutter, An efficient orbital transformation method for electronic structure calculations, J. Chem. Phys. 118 (10) (2003) 4365–4369, https://doi.org/10.1063/1.1543154, https://pubs.aip.org/aip/jcp/article-pdf/118/10/4365/19207648/4365_1_online.pdf.

[23] C. Freysoldt, S. Boeck, J. Neugebauer, Direct minimization technique for metals in density functional theory, Phys. Rev. B 79 (2009) 241103, https://doi.org/10.1103/PhysRevB.79.241103, https://link.aps.org/doi/10.1103/PhysRevB.79.241103.

[24] A.V. Ivanov, E.Ö. Jónsson, T. Vegge, H. Jónsson, Direct energy minimization based on exponential transformation in density functional calculations of finite and extended systems, Comput. Phys. Commun. 267 (2021) 108047, https://doi.org/10.1016/j.cpc.2021.108047, https://www.sciencedirect.com/science/article/pii/S0010465521001594.

[25] A. Edelman, T.A. Arias, S.T. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (2) (1998) 303–353, https://doi.org/10.1137/S0895479895290954.

[26] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, 2008.

[27] D. Raczkowski, C.Y. Fong, P.A. Schultz, R.A. Lippert, E.B. Stechel, Unconstrained and constrained minimization, localization, and the Grassmann manifold: theory and application to electronic structure, Phys. Rev. B 64 (2001) 155203, https://doi.org/10.1103/PhysRevB.64.155203, https://link.aps.org/doi/10.1103/PhysRevB.64.155203.

[28] E.J. Bylaska, K. Tsemekhman, F. Gao, New development of self-interaction corrected dft for extended systems applied to the calculation of native defects in 3c–sic, Phys. Scr. 2006 (T124) (2006) 86, https://doi.org/10.1088/0031-8949/2006/T124/017.

[29] Z. Wen, W. Yin, A feasible method for optimization with orthogonality constraints, Math. Program. 142 (1) (2013) 397–434, https://doi.org/10.1007/s10107-012-0584-1.

[30] X. Zhang, J. Zhu, Z. Wen, A. Zhou, Gradient type optimization methods for electronic structure calculations, SIAM J. Sci. Comput. 36 (3) (2014) C265–C289, https://doi.org/10.1137/130932934.

[31] B. Jiang, Y.-H. Dai, A framework of constraint preserving update schemes for optimization on Stiefel manifold, Math. Program. 153 (2) (2015) 535–575, https://doi.org/10.1007/s10107-014-0816-7.

[32] K. Baarman, V. Havu, T. Eirola, Direct minimization for ensemble electronic structure calculations, J. Sci. Comput. 66 (3) (2016) 1218–1233, https://doi.org/10.1007/s10915-015-0058-8.

[33] X. Dai, Z. Liu, L. Zhang, A. Zhou, A conjugate gradient method for electronic structure calculations, SIAM Journal on Scientific Computing 39 (6) (2017) A2702–A2740, https://doi.org/10.1137/16M1072929.

[34] X. Dai, S. de Gironcoli, B. Yang, A. Zhou, Mathematical analysis and numerical approximations of density functional theory models for metallic systems, Multiscale Model. Simul. 21 (3) (2023) 777–803, https://doi.org/10.1137/22M1472103.

[35] S. Sharma, J.K. Dewhurst, N.N. Lathiotakis, E.K.U. Gross, Reduced density matrix functional for many-electron systems, Phys. Rev. B 78 (2008) 201103, https://doi.org/10.1103/PhysRevB.78.201103, https://link.aps.org/doi/10.1103/PhysRevB.78.201103.

[36] J. Wang, E.J. Baerends, Self-consistent-field method for correlated many-electron systems with an entropic cumulant energy, Phys. Rev. Lett. 128 (2022) 013001, https://doi.org/10.1103/PhysRevLett.128.013001, https://link.aps.org/doi/10.1103/PhysRevLett.128.013001.

[37] E.R. Davidson, D. Feller, Basis set selection for molecular calculations, Chem. Rev. 86 (4) (1986) 681–696, https://doi.org/10.1021/cr00074a002.

[38] S. Wilson, Basis Sets, John Wiley & Sons, Ltd, 1987, pp. 439–500, Ch. 8, https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470142936.ch8, https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470142936.ch8.

[39] P. Li, X. Liu, M. Chen, P. Lin, X. Ren, L. Lin, C. Yang, L. He, Large-scale ab initio simulations based on systematically improvable atomic basis, Comput. Mater. Sci. 112 (2016) 503–517, https://doi.org/10.1016/j.commatsci.2015.07.004, https://www.sciencedirect.com/science/article/pii/S0927025615004140.

[40] T.A. Arias, Multiresolution analysis of electronic structure: semicardinal and wavelet bases, Rev. Mod. Phys. 71 (1999) 267–311, https://doi.org/10.1103/RevModPhys.71.267, https://link.aps.org/doi/10.1103/RevModPhys.71.267.

[41] J.R. Chelikowsky, N. Troullier, Y. Saad, Finite-difference-pseudopotential method: electronic structure calculations without a basis, Phys. Rev. Lett. 72 (1994) 1240–1243, https://doi.org/10.1103/PhysRevLett.72.1240, https://link.aps.org/doi/10.1103/PhysRevLett.72.1240.

[42] P. Motamarri, S. Das, S. Rudraraju, K. Ghosh, D. Davydov, V. Gavini, Dftfe – a massively parallel adaptive finite-element code for large-scale density functional theory calculations, Comput. Phys. Commun. 246 (2020) 106853, https://doi.org/10.1016/j.cpc.2019.07.016, https://www.sciencedirect.com/science/article/pii/S0010465519302309.

[43] L. Kleinman, D.M. Bylander, Efficacious form for model pseudopotentials, Phys. Rev. Lett. 48 (1982) 1425–1428, https://doi.org/10.1103/PhysRevLett.48.1425, https://link.aps.org/doi/10.1103/PhysRevLett.48.1425.

[44] X. Zhu, A Riemannian conjugate gradient method for optimization on the Stiefel manifold, Comput. Optim. Appl. 67 (1) (2017) 73–110, https://doi.org/10.1007/s10589-016-9883-4.

[45] R. Fletcher, C.M. Reeves, Function minimization by conjugate gradients, Comput. J. 7 (2) (1964) 149–154, https://doi.org/10.1093/comjnl/7.2.149, https://academic.oup.com/comjnl/article-pdf/7/2/149/959725/070149.pdf.

[46] E. Polak, G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, Rev. Fr. Inform. Rech. Opér. Sér. Rouge 3 (16) (1969) 35–43.

[47] B. Polyak, The conjugate gradient method in extremal problems, USSR Comput. Math. Math. Phys. 9 (4) (1969) 94–112, https://doi.org/10.1016/0041-5553(69)90035-4, https://www.sciencedirect.com/science/article/pii/0041555369900354.

[48] Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM J. Optim. 10 (1) (1999) 177–182, https://doi.org/10.1137/S1052623497318992.

[49] M.R. Hestenes, E. Stiefel, et al., Methods of Conjugate Gradients for Solving Linear Systems, vol. 49, NBS, Washington, DC, 1952.

[50] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, part 1: theory, J. Optim. Theory Appl. 69 (1) (1991) 129–137, https://doi.org/10.1007/BF00940464.

[51] W.W. Hager, H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM J. Optim. 16 (1) (2005) 170–192.

[52] M. Chen, G.-C. Guo, L. He, Systematically improvable optimized atomic basis sets for ab initio calculations, J. Phys. Condens. Matter 22 (44) (2010) 445501, https://doi.org/10.1088/0953-8984/22/44/445501.

[53] https://github.com/kluophysics/directmin_abacus, 2025, this is only temporary and will eventually be merged into the main branch.

[54] C. Sanderson, R. Curtin, Armadillo: a template-based c++ library for linear algebra, J. Open Source Softw. 1 (2) (2016) 26.

[55] G. Guennebaud, B. Jacob, et al., Eigen v3, eigen, http://eigen.tuxfamily.org, 2010.

[56] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople, Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation, J. Chem. Phys. 106 (3) (1997) 1063–1079, https://doi.org/10.1063/1.473182, https://pubs.aip.org/aip/jcp/article-pdf/106/3/1063/19100768/1063_1_online.pdf.

[57] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J.A. Pople, Assessment of Gaussian-2 and density functional theories for the computation of ionization potentials and electron affinities, J. Chem. Phys. 109 (1) (1998) 42–55, https://doi.org/10.1063/1.476538, https://pubs.aip.org/aip/jcp/article-pdf/109/1/42/19137091/42_1_online.pdf.

[58] M. Schlipf, F. Gygi, Optimization algorithm for the generation of oncv pseudopotentials, Comput. Phys. Commun. 196 (2015) 36–44, https://doi.org/10.1016/j.cpc.2015.05.011, https://www.sciencedirect.com/science/article/pii/S0010465515001897.

[59] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, https://doi.org/10.1103/PhysRevLett.77.3865, https://link.aps.org/doi/10.1103/PhysRevLett.77.3865.

[60] D.D. Johnson, Modified Broyden's method for accelerating convergence in self-consistent calculations, Phys. Rev. B 38 (1988) 12807–12813, https://doi.org/10.1103/PhysRevB.38.12807, https://link.aps.org/doi/10.1103/PhysRevB.38.12807.

[61] W. Huang, K.A. Gallivan, P.-A. Absil, A Broyden class of quasi-Newton methods for Riemannian optimization, SIAM J. Optim. 25 (3) (2015) 1660–1685, https://doi.org/10.1137/140955483.

[62] W. Huang, P.-A. Absil, K.A. Gallivan, A Riemannian bfgs method without differentiated retraction for nonconvex optimization problems, SIAM J. Optim. 28 (1) (2018) 470–495, https://doi.org/10.1137/17M1127582.

[63] W. Ring, B. Wirth, Optimization methods on Riemannian manifolds and their application to shape space, SIAM J. Optim. 22 (2) (2012) 596–627, https://doi.org/10.1137/11082885X.