Projected gradient descent algorithm for *ab initio* crystal structure relaxation under a fixed unit cell volume

Yukuan Hu,[1,2] Junlei Yin,[3,4] Xingyu Gao,[5,∗] Xin Liu,[1,2,†] and Haifeng Song[5,‡]

[1]*State Key Laboratory of Scientific and Engineering Computing,*
Institute of Computational Mathematics and Scientific/Engineering Computing,
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[2]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*State Key Laboratory of Solidification Processing, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China*
[4]*Innovation Center, NPU Chongqing, Chongqing 401135, China*
[5]*Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, Beijing 100088, China*

This paper is concerned with *ab initio* crystal structure relaxation under a fixed unit cell volume, which is a step in calculating the static equations of state and forms the basis of thermodynamic property calculations for materials. The task can be formulated as an energy minimization with a determinant constraint. Widely used line minimization-based methods (e.g., conjugate gradient method) lack both efficiency and convergence guarantees due to the nonconvex nature of the feasible region as well as the significant differences in the curvatures of the potential energy surface with respect to atomic and lattice components. To this end, we propose a projected gradient descent algorithm named PANBB. It is equipped with (i) search direction projections onto the tangent spaces of the nonconvex feasible region for lattice vectors, (ii) distinct curvature-aware initial trial step sizes for atomic and lattice updates, and (iii) a nonrestrictive line minimization criterion as the stopping rule for the inner loop. It can be proved that PANBB favors theoretical convergence to equilibrium states. Across a benchmark set containing 223 structures from various categories, PANBB achieves average speedup factors of approximately 1.41 and 1.45 over the conjugate gradient method and direct inversion in the iterative subspace implemented in off-the-shelf simulation software, respectively. Moreover, it normally converges on all the systems, manifesting its unparalleled robustness. As an application, we calculate the static equations of state for the high-entropy alloy AlCoCrFeNi, which remains elusive owing to 160 atoms representing both chemical and magnetic disorder and the strong local lattice distortion. The results are consistent with the previous calculations and are further validated by experimental thermodynamic data.

I. INTRODUCTION

Crystal structure relaxation seeks equilibrium states on the potential energy surface (PES). It underpins various applications such as crystal structure predictions [1–4] and high-throughput calculations in material design [5–9]. Relaxing crystal structures under a fixed unit cell volume is a step in calculating the static equations of state (EOS) for materials [10, 11] and also presents the volume-dependent configurations for thermal models such as the quasiharmonic approximation [12, 13], classical mean-field potential [14–18], and electronic excitation based on Mermin statistics [19, 20], etc.

Mathematically, relaxing the crystal structure under a fixed unit cell volume can be formulated as minimizing the potential energy with a determinant constraint, where a nonconvex feasible region arises. Developing convergent methods for general nonconvex constrained optimization problems remains challenging, since classical tools (e.g., projection) and theories (e.g., duality) become inapplicable. As far as we know, little attention has been paid to the determinant-constrained optimization.

In practice, line minimization (LM)-based methods are usual choices for this purpose. A flowchart of the widely used LM-based methods is depicted in Fig. 1. Each iteration of these methods primarily consists of updating the search directions followed by an LM process to determine the step sizes.

Among the popular representatives in this class are the conjugate gradient method (CG) [21] and direct inversion in the iterative subspace (DIIS) [22]. Specifically, CG searches along the conjugate gradient directions made up by previous directions and current forces, while DIIS employs historical information to construct the quasi-Newton directions. In most cases, CG exhibits steady convergence but can be unsatisfactory in efficiency. DIIS converges rapidly in the small neighborhoods of local minimizers but can easily diverge when starting far from equilibrium states.

As shown in Fig. 1, the performance of the LM-based methods crucially depends on three key factors: search directions, initial trial step sizes, and LM criterion. However, the fixed-volume relaxation presents substantial difficulties on their designs. Firstly, due to the nonconvex nature of the feasible region, employing search directions for lattice vectors without tailored modifications can result in considerable variations in the configurations, giving rise to drastic fluctuations in the potential energies; for an illustration, please refer to Fig. 2 later. Secondly, on account of the significant differences in the curvatures of PES with respect to the atomic and lattice components, adopting shared initial trial step sizes, as in the widely used LM-based methods (see S2 in Fig. 1), can be detrimental to algorithm efficiency. This point is recognized through extensive numerical tests, where the relative displacements in the lattice vectors during relaxation are typically much smaller than those in the atomic positions. We refer readers to Fig. 3 later for an illustration. Thirdly, the monotone criterion utilized in the widely used LM-based methods mandates a strict reduction in the potential energy after each iteration, often making the entire procedure trapped in the LM process.
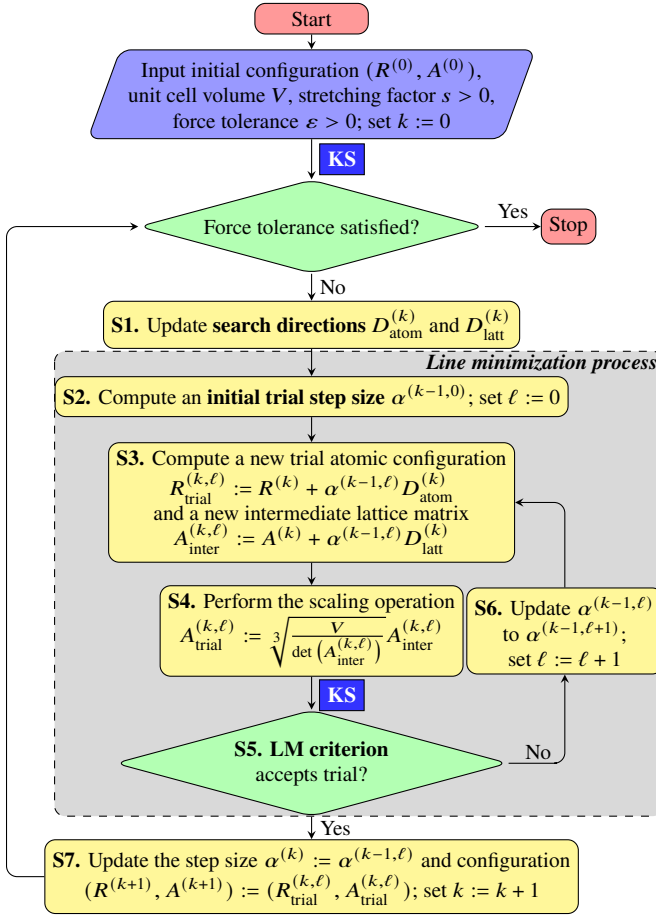
FIG. 1. A flowchart of the widely used LM-based methods. The bold text "**KS**" refers to solving Kohn-Sham (KS) equations to obtain new energy, atomic forces, and lattice stress whenever the configuration gets updated. The LM process is marked out by the shaded box. The subscript "inter" indicates the intermediate lattice matrix.

According to incomplete statistics, the CG implemented in off-the-shelf simulation software spends approximately 60% of its overhead on the trials rejected by the LM criterion, as supported by the Supplemental Material. It is worth noting that the authors of [23] devise a nonmonotone LM criterion but only consider atomic degrees of freedom.

In this work, we propose a projected gradient descent algorithm, called PANBB, that addresses all the aforementioned difficulties. Specifically, PANBB (i) employs search direction projections onto the tangent spaces of the nonconvex feasible region for lattice vectors, (ii) adopts distinct curvature-aware step sizes for the atomic and lattice components, and (iii) incorporates a nonmonotone LM criterion that encompasses both atomic and lattice degress of freedom. The convergence of PANBB to equilibrium states has been theoretically established.

PANBB has been tested across a benchmark set containing 223 systems from various categories, with approximately 68.6% of them are metallic systems, for which the EOS calculations are more demanding. PANBB exhibits average speedup factors of approximately 1.41 and 1.45 over the CG and DIIS implemented in off-the-shelf simulation software, respectively,

in terms of running time. In contrast to the failure rates of approximately 4.9% for CG and 25.1% for DIIS, PANBB consistently converges to the equilibrium states across the broad, manifesting its unparalleled robustness. As an application of PANBB, we calculate the static EOS of the high-entropy alloy (HEA) AlCoCrFeNi, whose fixed-volume relaxation remains daunting owing to the strong local lattice distortion (LLD) [24]. The numerical results are consistent with the previous calculations [25] obtained by relaxing only atomic positions. We further apply the modified mean field potential (MMFP) approach [18] to the thermal EOS, which is validated by the X-ray diffraction and diamond anvil cell experiments [26].

The notations are gathered as follows. We use $N \in \mathbb{N}$ for the number of atoms, and use $R := [\mathbf{r}_1, \ldots, \mathbf{r}_N] \in \mathbb{R}^{3 \times N}$ and $A := [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] \in \mathbb{R}^{3 \times 3}$ to denote the Cartesian atomic positions and lattice matrix, respectively. The reciprocal lattice matrix $A^{-\top}$ is abbreviated as $B \in \mathbb{R}^{3 \times 3}$. We denote by $E(R, A) \in \mathbb{R}$, $F_{\text{atom}}(R, A) \in \mathbb{R}^{3 \times N}$, $F_{\text{latt}}(R, A) \in \mathbb{R}^{3 \times 3}$, $\Sigma(R, A) \in \mathbb{R}^{3 \times 3}$, and $\Sigma_{\text{dev}}(R, A) \in \mathbb{R}^{3 \times 3}$, respectively, the potential energy, atomic forces, lattice force, full stress tensor, and deviatoric stress tensor evaluated at the configuration (R, A); if no confusion arises, we omit specifying (R, A). From [27, 28], one obtains the relation

$$F_{\text{latt}} = V \Sigma B - F_{\text{atom}} R^\top B. \tag{1}$$

When describing algorithms, we use the first and second superscripts in brackets to indicate respectively the outer and inner iteration numbers (e.g., $E^{(k)}$ and $\alpha^{(k-1,\ell)}$). We denote by "**I**" the identity mapping from $\mathbb{R}^{3 \times N}$ to $\mathbb{R}^{3 \times N}$. The notation "$\langle \cdot, \cdot \rangle$" refers to the standard inner product between two matrices, defined as $\langle M, N \rangle := \text{Tr}(M^\top N)$, whereas "$\| \cdot \|_{\text{F}}$" yields the matrix Frobenius norm as $\|M\|_{\text{F}} := \sqrt{\langle M, M \rangle}$.

II. ALGORITHMIC DEVELOPMENTS

In the sequel, we expound on our algorithmic developments from three aspects: search directions, initial trial step sizes, and LM algorithm and criterion.

II.1. Search Directions

For the atomic part, we simply take the steepest descent directions, i.e.,

$$D_{\text{atom}}^{(k)} := F_{\text{atom}}^{(k)} \text{ for all } k \geq 0.$$

As for the lattice part, we first note that the steepest descent direction $F_{\text{latt}}^{(k)}$ (see Eq. (1)) cannot be used directly due to the volume constraint; otherwise, there can be unexpected large variations in the lattice vectors, resulting in drastic fluctuations in the potential energies. For an illustration in two-dimensional context, see the left panel of Fig. 2 where one can observe a large distance between the last iterate $A^{(k)}$ (denoted by the black square) and the next iterate using $F_{\text{latt}}^{(k)}$ (denoted by the black circle).

Instead, we search along the tangent lines of the feasible region at the current configuration. In mathematics, this amounts

Comparison between iterations w/ and w/o gradient projection
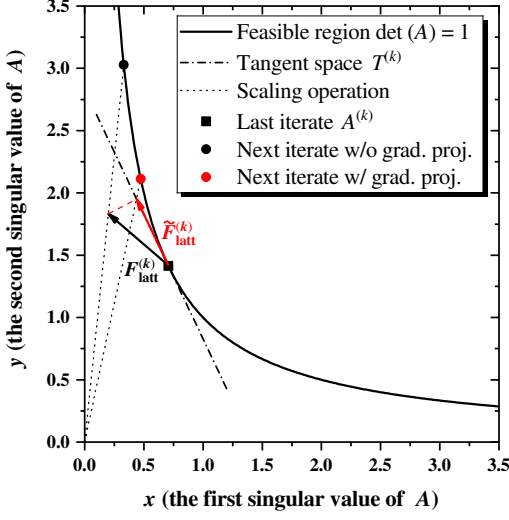
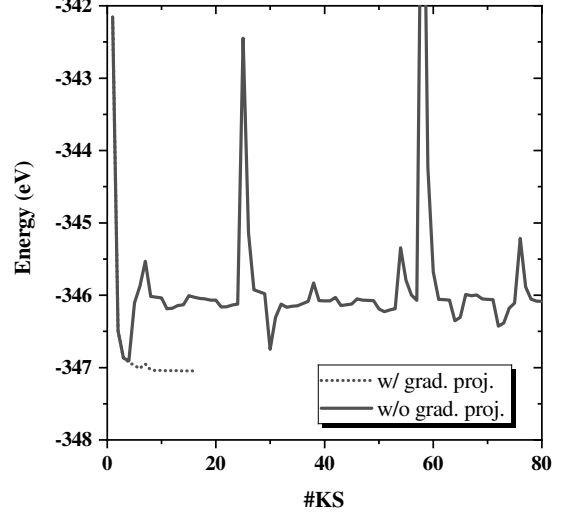**Relaxation history of PANBB on an 8*1*1 silicon supercell**

FIG. 2. Left: a comparison between iterations with and without the gradient projection (3) in two-dimensional context. The x- and y-axes refer to the first and second singular values of lattice matrix A, respectively. The black solid line stands for the feasible region $\det(A) = 1$, the black square denotes the last iterate $A^{(k)}$, and the black dashdotted line refers to the tangent space $\mathcal{T}^{(k)}$ of feasible region at $A^{(k)}$. The black and red arrows represent respectively the lattice force $F_{\text{latt}}^{(k)}$ and its projection onto $\mathcal{T}^{(k)}$, namely, $\tilde{F}_{\text{latt}}^{(k)}$. The black and red circles show the next iterates without and with the gradient projection (3), respectively, after performing the scaling operation indicated by the black dotted lines. Right: the relaxation history of PANBB on an $8 \times 1 \times 1$ silicon supercell (64 atoms), where "#KS" denotes the number of solving the KS equations, "w/ grad. proj." and "w/o grad. proj." stand for the results using $\tilde{F}_{\text{latt}}^{(k)}$ and $F_{\text{latt}}^{(k)}$, respectively.

to projecting the lattice force $F_{\text{latt}}^{(k)}$ onto the tangent space $\mathcal{T}^{(k)} \subseteq \mathbb{R}^{3\times3}$ [29] of $\mathcal{F} := \{A \in \mathbb{R}^{3\times3} : \det(A) = V\}$ at $A^{(k)}$, where

$$\mathcal{T}^{(k)} := \left\{ D \in \mathbb{R}^{3\times3} : \left\langle B^{(k)}, D \right\rangle = 0 \right\}. \tag{2}$$

The projection favors a closed-form expression:

$$\tilde{F}_{\text{latt}}^{(k)} := F_{\text{latt}}^{(k)} - \frac{\left\langle B^{(k)}, F_{\text{latt}}^{(k)} \right\rangle}{\|B^{(k)}\|_{\text{F}}^2} B^{(k)}. \tag{3}$$

We then take $D_{\text{latt}}^{(k)} := \tilde{F}_{\text{latt}}^{(k)}$ for $k \geq 0$. A comparison between the iterations with and without the gradient projection (3) can be found in the left panel of Fig. 2. After projecting $F_{\text{latt}}^{(k)}$ (denoted by the black arrow) onto $\mathcal{T}^{(k)}$, we obtain $\tilde{F}_{\text{latt}}^{(k)}$ (denoted by the red arrow) and then an iterate closer to $A^{(k)}$ (denoted by the red circle).

It can be shown theoretically that the gradient projection (3) eliminates the potential first-order energy increment induced by the scaling operation (S4 in Fig. 1). We shall point out that the strategy does not necessarily apply to general nonconvex constrained optimization problems. The desired theoretical results are achieved by fully exploiting the algebraic characterization of the tangent space (2). For technical details, please refer to the proof of Lemma 2 in Appendix A. We also demonstrate the merit of the gradient projection (3) through numerical comparisons; see the right panel of Fig. 2.

II.2. Initial Trial Step Sizes

Efficient initial trial step sizes are essential for improving the overall performance. Moreover, our extensive numerical tests reveal that the relative displacements in the lattice vectors are typically much smaller than those in the atomic positions during relaxation, particularly for large systems. This observation underscores the substantial differences in the local curvatures of PES with respect to the atomic positions and lattice vectors, or in other words, ill conditions underlying the relaxation tasks. In light of this, we shall treat the atomic and lattice components separately and leverage their respective curvature information.

Motivated by the above discussions, we introduce curvature-aware step sizes for the atomic and lattice components, respectively. For the atomic part, we adopt the alternating Barzilai-Borwein (ABB) step sizes [23, 30, 31], defined as

$$\alpha_{\text{atom}}^{(k-1,0)} = \alpha_{\text{atom,ABB}}^{(k)} := \begin{cases} \alpha_{\text{atom,BB1}}^{(k)}, & \text{if } \mathrm{mod}(k,2) = 0; \\ \alpha_{\text{atom,BB2}}^{(k)}, & \text{if } \mathrm{mod}(k,2) = 1 \end{cases} \tag{4}$$

for $k \geq 1$, where

$$\alpha_{\text{atom,BB1}}^{(k)} := \frac{\|S_{\text{atom}}^{(k-1)}\|_{\text{F}}^2}{\left\langle S_{\text{atom}}^{(k-1)}, Y_{\text{atom}}^{(k-1)} \right\rangle}, \quad \alpha_{\text{atom,BB2}}^{(k)} := \frac{\left\langle S_{\text{atom}}^{(k-1)}, Y_{\text{atom}}^{(k-1)} \right\rangle}{\|Y_{\text{atom}}^{(k-1)}\|_{\text{F}}^2},$$

$S_{\text{atom}}^{(k-1)} := R^{(k)} - R^{(k-1)}$, and $Y_{\text{atom}}^{(k-1)} := F_{\text{atom}}^{(k-1)} - F_{\text{atom}}^{(k)}$. It is

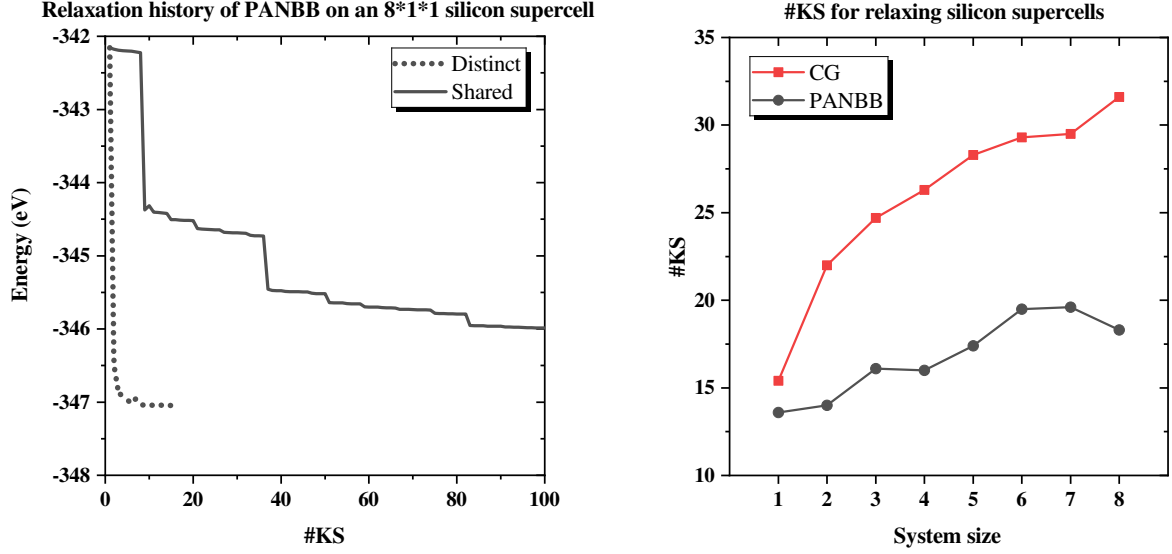

FIG. 3. The relaxation on silicon supercells ($8 \sim 64$ atoms). Left: the relaxation history of PANBB on an $8 \times 1 \times 1$ silicon supercell. "Distinct" and "Shared" stand for the results of PANBB using distinct and shared step sizes for atomic and lattice components, respectively. Right: the average #KS by CG and PANBB for relaxing silicon supercells of different sizes. Each system size is associated with 10 initial configurations, which are generated by random perturbations from equilibrium states; the deviations in the atomic positions are less than 0.1 Å and the deviations in the lattice vectors are less than 0.02 Å.

not difficult to verify that

$$\alpha_{\text{atom,BB1}}^{(k)} = \arg\min_{\alpha} \left\| \alpha^{-1} S_{\text{atom}}^{(k-1)} - Y_{\text{atom}}^{(k-1)} \right\|_{\text{F}},$$

$$\alpha_{\text{atom,BB2}}^{(k)} = \arg\min_{\alpha} \left\| \alpha Y_{\text{atom}}^{(k-1)} - S_{\text{atom}}^{(k-1)} \right\|_{\text{F}}.$$

In other words, $\alpha_{\text{atom,BB1}}^{(k)}\mathbf{I}$ and $\alpha_{\text{atom,BB2}}^{(k)}\mathbf{I}$ approximate the inverse atomic Hessian of potential energy around $R^{(k)}$. The gradient descent with the above BB step sizes thus resembles the Newton iteration for solving the nonlinear equations $F_{\text{atom}} = 0$.

Now let us derive the ABB step sizes for the lattice part. Due to the volume constraint, the equilibrium condition on the lattice becomes $\Sigma_{\text{dev}} = 0$, namely, the shear stress vanishes and the normal stress cancels out the external isotropic pressure. This is equivalent to $\tilde{F}_{\text{latt}} = 0$ in view of Eqs. (1) and (3), provided that the atomic forces vanish. Recalling the Newton iteration for solving the nonlinear equations $\tilde{F}_{\text{latt}} = 0$, we obtain the following ABB step sizes for the lattice vectors

$$\alpha_{\text{latt}}^{(k-1,0)} = \alpha_{\text{latt,ABB}}^{(k)} := \begin{cases} \alpha_{\text{latt,BB1}}^{(k)}, & \text{if mod}(k,2) = 0; \\ \alpha_{\text{latt,BB2}}^{(k)}, & \text{if mod}(k,2) = 1 \end{cases} \quad (5)$$

for $k \geq 1$, where

$$\alpha_{\text{latt,BB1}}^{(k)} := \frac{\|S_{\text{latt}}^{(k-1)}\|_{\text{F}}^2}{\left\langle S_{\text{latt}}^{(k-1)}, Y_{\text{latt}}^{(k-1)} \right\rangle}, \quad \alpha_{\text{latt,BB2}}^{(k)} := \frac{\left\langle S_{\text{latt}}^{(k-1)}, Y_{\text{latt}}^{(k-1)} \right\rangle}{\|Y_{\text{latt}}^{(k-1)}\|_{\text{F}}^2},$$

$S_{\text{latt}}^{(k-1)} := A^{(k)} - A^{(k-1)}$, and $Y_{\text{latt}}^{(k-1)} := \tilde{F}_{\text{latt}}^{(k-1)} - \tilde{F}_{\text{latt}}^{(k)}$. With

these distinct step sizes, we update the trials as

$$\begin{aligned} R_{\text{trial}}^{(k,\ell)} &:= R^{(k)} + \alpha_{\text{atom}}^{(k-1,\ell)} D_{\text{atom}}^{(k)}, \\ A_{\text{inter}}^{(k,\ell)} &:= A^{(k)} + \alpha_{\text{latt}}^{(k-1,\ell)} D_{\text{latt}}^{(k)}, \qquad \ell = 0, 1, \dots. \quad (6) \\ A_{\text{trial}}^{(k,\ell)} &:= \sqrt[3]{\frac{V}{\det\left(A_{\text{inter}}^{(k,\ell)}\right)}} A_{\text{inter}}^{(k,\ell)}, \end{aligned}$$

As shown in Fig. 3, we demonstrate the benefits brought by the distinct ABB step sizes (4) and (5) through the relaxation of silicon supercells. From the left panel, one can observe the considerable acceleration of the distinct version over the shared one. In the latter case, the ABB step sizes are shared and computed by the quantities from both the atomic and lattice components. In the right panel, we test our new algorithm equipped with the distinct step sizes on the silicon supercells of different sizes ($1 \times 1 \times 1 \sim 8 \times 1 \times 1$). Compared with that of the CG built in off-the-shelf simulation software, the number of solving the KS equations of the new algorithm increases slowly as the system size grows, which indicates a preconditioning-like effect of the our strategy.

II.3. LM Algorithm and Criterion

With the search directions in place, the LM algorithm starts from the initial step sizes and attempts to locate an appropriate trial configuration that meets certain LM criterion. Capturing the local PES information, however, the ABB step sizes can occasionally incur small increments on energies [23]. Extensive numerical simulations suggest that keeping the ABB step sizes intact is more beneficial [23, 32]. The monotone LM criterion (S5 in Fig. 1) adopted in the widely used LM-based methods will lead to unnecessary computational expenditure

and undermine the merits of the ABB step sizes. Notably, the authors of [23] propose a nonmonotone criterion which encompasses only atomic degress of freedom.

In the following, a nonrestrictive nonmonotone LM criterion is presented for both atomic and lattice degrees of freedom and the simple backtracking is taken as the LM algorithm. Specifically, the nonmonotone criterion accepts the trial configuration if

$$E_{\text{trial}}^{(k,\ell)} \le \bar{E}^{(k)} - \eta \big(\alpha_{\text{atom}}^{(k-1,\ell)} \|F_{\text{atom}}^{(k)}\|_F^2 + \alpha_{\text{latt}}^{(k-1,\ell)} \|\tilde{F}_{\text{latt}}^{(k)}\|_F^2 \big), \quad (7)$$

where $E_{\text{trial}}^{(k,\ell)} := E(R_{\text{trial}}^{(k,\ell)}, A_{\text{trial}}^{(k,\ell)})$, $\eta \in (0,1)$ is a constant, and $\{\bar{E}^{(k)}\}$ is a surrogate sequence, defined recursively as

$$\bar{E}^{(k+1)} := \frac{\bar{E}^{(k)} + \mu^{(k)} q^{(k)} E^{(k+1)}}{1 + \mu^{(k)} q^{(k)}}, \quad (8)$$
$$q^{(k+1)} := \mu^{(k)} q^{(k)} + 1,$$

with $\bar{E}^{(0)} := E^{(0)}$, $q^{(0)} := 1$, and $\mu^{(k)} \in [0,1]$. The existence of $\alpha_{\text{atom}}^{(k-1,\ell)}$ and $\alpha_{\text{latt}}^{(k-1,\ell)}$ satisfying Eq. (7) is proved in Lemma 2 of Appendix A. It is not hard to verify that $\bar{E}^{(k)}$ is a weighted average of the past energies $\{E^{(j)}\}_{j=0}^{k}$ and is always larger than $E^{(k)}$. Hence, small energy increments are tolerable. That being said, the criterion (7) is sufficient to ensure the convergence of the entire procedure; see Theorem 1 in Appendix A.

By virtue of the nonmonotone nature of the criterion (7), we take the simple backtracking as the LM algorithm. Whenever the last trial configuration is rejected, we set

$$\alpha_{\text{atom}}^{(k-1,\ell)} := \alpha_{\text{atom}}^{(k-1,\ell-1)} \delta_{\text{atom}}, \quad \alpha_{\text{latt}}^{(k-1,\ell)} := \alpha_{\text{latt}}^{(k-1,\ell-1)} \delta_{\text{latt}} \quad (9)$$

where $\delta_{\text{atom}}, \delta_{\text{latt}} \in (0,1)$ are constants. Without carefully selected δ_{atom} and δ_{latt}, we observe in most simulations that the ABB step sizes can directly meet the nonrestrictive criterion (7) and the rejected trials account for approximately 1.8% of computational costs, in stark contrast to approximately 58.8% in the widely used CG.

II.4. Implementation Details

With the aforementioned developments, we have devised a projected gradient descent algorithm called PANBB for crystal structure relaxation under a fixed unit cell volume. More specifically, for the initial trial step sizes, we set $\alpha_{\text{atom}}^{(-1,0)} = 4.8 \times 10^{-2}$ Å²/eV, $\alpha_{\text{latt}}^{(-1,0)} = 10^{-6}$ Å²/eV during the first iteration [33] and compute in the subsequent iterations the truncated absolute values of the ABB step sizes:

$$\alpha_{\text{atom}}^{(k-1,0)} = \max \left\{ \min \left\{ \left| \alpha_{\text{atom,ABB}}^{(k)} \right|, \tau_{\text{atom}}^{(k)}, 10 \right\}, 10^{-5} \right\}, \quad (10a)$$

$$\alpha_{\text{latt}}^{(k-1,0)} = \max \left\{ \min \left\{ \left| \alpha_{\text{latt,ABB}}^{(k)} \right|, \tau_{\text{latt}}^{(k)}, 0.1 \right\}, 10^{-7} \right\}, \quad (10b)$$

where

$$\tau_{\text{atom}}^{(k)} := \gamma_{\text{atom}}^{(k)} \max \left\{ -\log_{10} \left(\|F_{\text{atom}}^{(k)}\|_F / N \right), 1 \right\},$$
$$\tau_{\text{latt}}^{(k)} := \gamma_{\text{latt}}^{(k)} \max \left\{ -\log_{10} \left(\|\tilde{F}_{\text{latt}}^{(k)}\|_F / N \right), 1 \right\}.$$

The factors $\gamma_{\text{atom}}^{(k)}$, $\gamma_{\text{latt}}^{(k)}$ take respectively the values of 1, 10^{-3} for $k = 0$ and later are modified according to the previous iterations: if they appear to be stringent from the recent history for $\alpha_{\text{atom}}^{(k-1,0)}$ and $\alpha_{\text{latt}}^{(k-1,0)}$ to pass the criterion (7), we increase them; otherwise, we decrease them. For instance,

$$\gamma_{\text{atom}}^{(k)} := \begin{cases} 2\gamma_{\text{atom}}^{(k-1)}, & \text{Case I;} \\ 0.5\gamma_{\text{atom}}^{(k-1)}, & \text{Case II;} \\ \gamma_{\text{atom}}^{(k-1)}, & \text{otherwise,} \end{cases}$$

where

- **Case I:** the truncation has been active but $\alpha_{\text{atom}}^{(k-1,0)}$ has directly passed the criterion (7) for twice in the past $\min\{k - \hat{k}, 20\}$ iterations;

- **Case II:** $\alpha_{\text{atom}}^{(k-1,0)}$ has failed to pass the criterion (7) for twice in the past $\min\{k - \hat{k}, 20\}$ iterations.

Initially, we set $\hat{k} := 0$; whenever the factor $\gamma_{\text{atom}}^{(k)}$ gets changed, we set $\hat{k} := k$. The update on $\gamma_{\text{latt}}^{(k)}$ follows similar rules. Note that the box boundaries in Eq. (10), such as 10^{-5} and 10 for $\alpha_{\text{atom}}^{(k-1,0)}$, are present for the ease of theoretical analyses. In practice, these boundaries are rarely touched. The criterion (7) is equipped with $\eta = 10^{-4}$ and $\mu^{(k)} \equiv 0.05$ following [23]. We take $\delta_{\text{atom}} = 0.1$ and $\delta_{\text{latt}} = 0.5$ for the backtracking (9) [34]. Under mild conditions, we establish the convergence of PANBB to equilibrium states; see Theorem 1 in Appendix A.

III. NUMERICAL SIMULATIONS

III.1. Experimental Settings

We have implemented PANBB in the in-house plane-wave code CESSP [35–38]. The exchange-correlation energy is described by the generalized gradient approximation [39]. Electron-ion interactions are treated with the projector augmented-wave (PAW) potentials based on the open-source ABINIT Jollet-Torrent-Holzwarth data set library in the PAW-XML format [40]. The total energies are calculated using the Monkhorst-Pack mesh [41] with a k-mesh spacing of $0.12 \sim 0.20$ Å^{-1}. The typical plane-wave energy cutoffs are around $500 \sim 600$ eV. The KS equations are solved by the preconditioned self-consistent field (SCF) iteration [37].

In addition to PANBB, we test the performance of the CG and DIIS built in CESSP. Notably, compared with the one documented in textbook, the CG in test uses deviatoric stress tensors in search direction updates for the fixed-volume relaxation. We terminate all the three methods whenever the number of solving the Kohn-Sham equations (abbreviated as #KS) reaches 1000 or the maximum atomic force and the maximum stress components divided by N fall below 0.01 eV/Å. The convergence criterion for the SCF iteration is 10^{-5} eV. To evaluate the performance of the three methods, we record the #KS and running time in seconds (abbreviated as T) required for fulfilling the force tolerance.
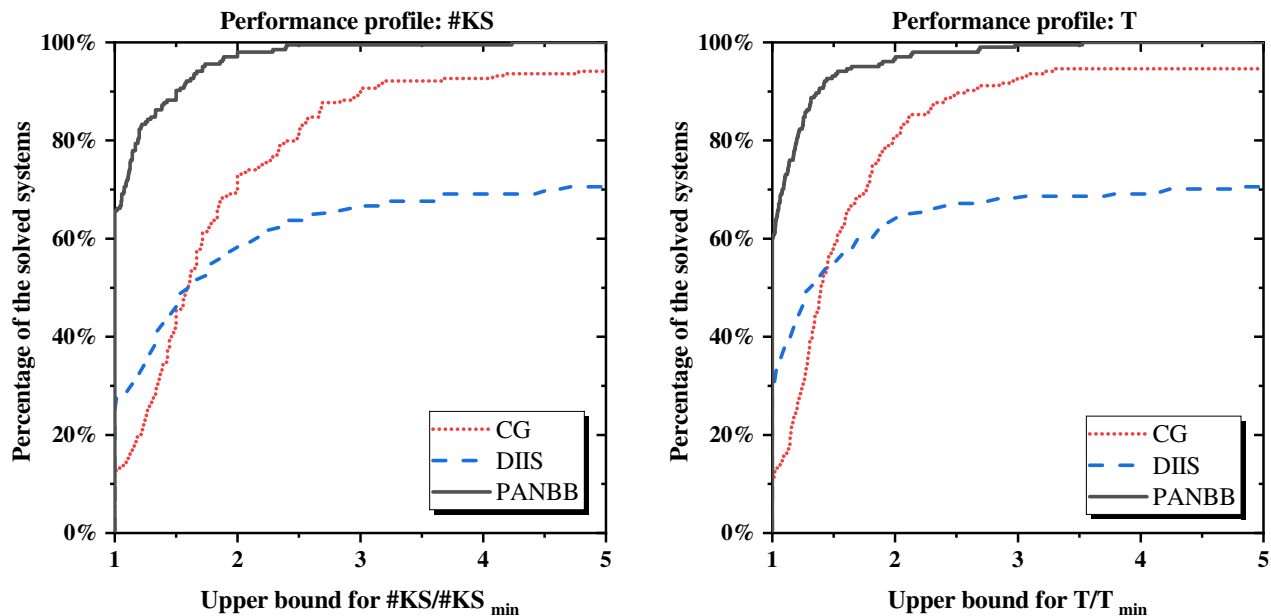
FIG. 4. The performance profiles of CG, DIIS, and PANBB, where $\#KS_{min}$ and T_{min} denote the minimum #KS and T required by the three methods on a given system, respectively. To reach a fair comparison and illustrate the algorithm robustness at the same time, we only retain (i) the systems where the difference between the converged energies obtained by each pair of methods does not exceed 3 meV per atom; (ii) the systems on which at least one method diverges or fails to converge before #KS exceeds 1000, and set the associated #KS and T to ∞. The final number of the systems used for this figure is 204. Left: #KS. Right: T.

III.2. Benchmark Test

We have performed numerical comparisons among CG, DIIS, and PANBB on a benchmark set consisting of 223 systems from various categories, such as metallic systems, organic molecules, semiconductors, ABX_3 perovskites, heterostructures, and 2D materials. About 68.6% of them are metallic systems, for which the EOS calculations are more demanding. The number of atoms in one system ranges from 2 to 215. Some of them are available from the Materials Project [42] and the Organic Materials Database [43]. In the beginning, these systems are at their ideal configurations for defects, heterostructures, and substitutional alloys, or simply place a molecule on top of a surface. More information about the benchmark set can be found in the Supplemental Material.

To provide a comprehensive comparison among CG, DIIS, and PANBB over the benchmark set, we adopt the performance profile [44] as shown in Fig. 4, where $\#KS_{min}$ and T_{min} represent the minimum #KS and T required by the three methods on a given system, respectively. For example, in the right panel of Fig. 4, the intercepts of the three curves indicate that CG, DIIS, and PANBB are the fastest on approximately 11.3%, 28.9%, and 59.8% of all the systems, respectively. The y-coordinate touched by the curve of PANBB when the x-coordinate equals two shows that the running time of PANBB is not twice larger than those of the other two methods on approximately 96.5% of all the systems. Basically, a larger area below the curve implies better overall performance of the associated method. We can therefore conclude from Fig. 4 that PANBB favors the best overall performance. In addition, the Supplemental Material reveals that PANBB is respectively faster than CG and DIIS on approximately 85.2% and 68.8% of all the systems. The

average speedup factors of PANBB over CG and DIIS, in terms of running time, are around 1.41 and 1.45, respectively. We also calculate the average speedup factors of PANBB over CG by the system category; see Fig. 5 for an illustration. Among others, the acceleration brought by PANBB is the most evident when relaxing organic molecules, ABX_3 perovskites, and metallic systems, with speedup factors of 1.67, 1.50, and 1.40, respectively, in terms of running time. These statistics provide strong evidence for the universal efficiency superiority of PANBB. The robustness of PANBB also deserves a whistle. Thanks to its theoretical convergence guarantees (see Theorem 1 in Appendix A), PANBB converges consistently across the benchmark set, whereas CG fails on 11 systems due to the breakdown of the LM algorithm, and DIIS diverges on 56 systems.

III.3. Calculations of EOS on the AlCoCrFeNi HEA

As an application of crystal structure relaxation under a fixed unit cell volume, we employ PANBB to calculate the static EOS up to approximately 30 GPa for the body-centered cubic AlCoCrFeNi HEA. The local lattice distortion (LLD), induced mainly by the large atomic size mismatch of the alloy components, is one of the core effects responsible for the unprecedented mechanical behaviors of HEAs. The LLD in some equimolar HEAs under ambient pressure is studied by *ab initio* approach [24]. The volume-dependent LLD is investigated and is helpful for understanding the structural properties of HEAs under pressure. The thermodynamic properties have also been investigated in comparison with experimental data for validation.
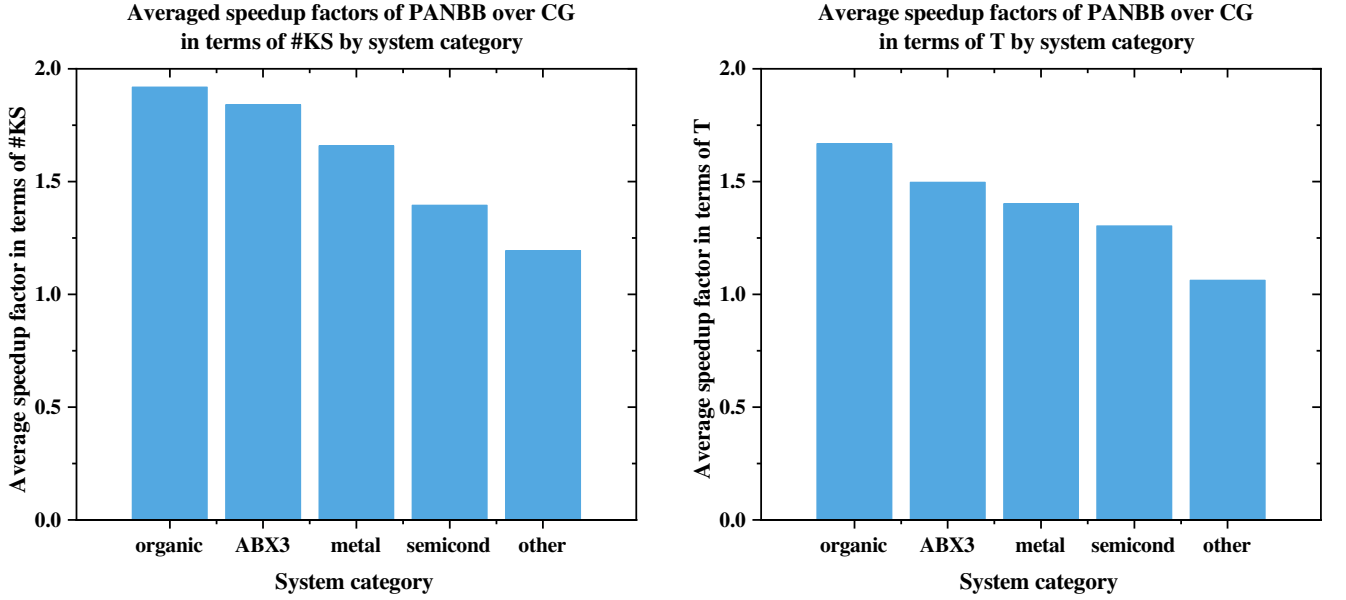
FIG. 5. The average speedup factors of PANBB over CG by the system category. To achieve a fair comparison, we only include the systems where both PANBB and CG converge normally and the converged energy differences per atom are not larger than 3 meV. Categories whose numbers of the included systems are less than 5 are merged together into "other". Left: #KS. Right: T.
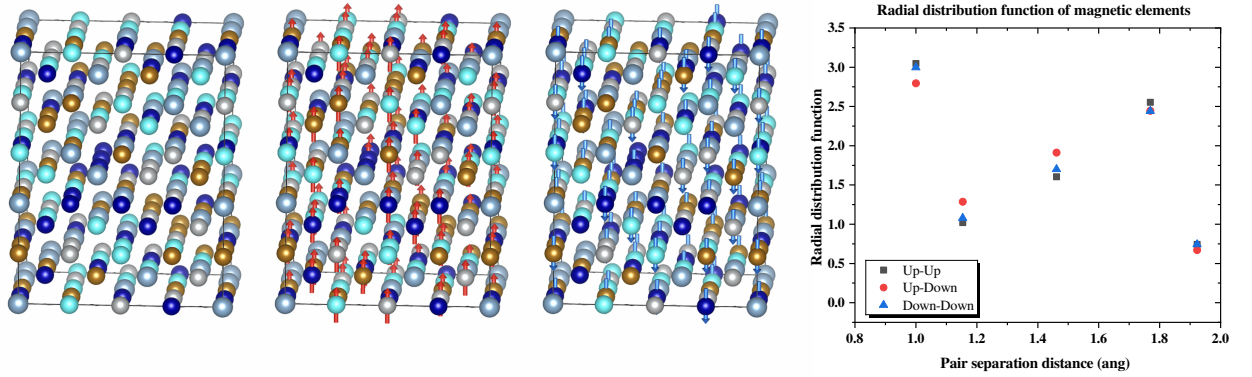

FIG. 6. The generated atomic and magnetic configurations. From left to right: the atomic configuration generated by the SAE method, the 64 spin-up and 64 spin-down sites chosen by the integer programming approach for magnetic configuration, and the radial distribution function of the magnetic elements.

To incorporate the temperature-induced magnetic structural transition, we need to describe both the ferromagnetic (FM) and paramagnetic (PM) states of AlCoCrFeNi. To characterize the inherent chemical and magnetic disorder, we combine the SAE method [45] and an integer programming approach [46] to construct representative $4 \times 4 \times 5$ supercells with 160 atoms. Specifically, the SAE method is first invoked for chemical disorder modeling, where we seek to minimize

$$\sum_{d(C_2) < r_2^c} w_2 f(C_2, \sigma) + \sum_{d(C_3) < r_3^c} w_3 f(C_3, \sigma)$$

with respect to the atomic occupation σ. Here w_n ($n = 2, 3$) are appropriate weights and r_n^c ($n = 2, 3$) are cutoff radii. The notations $f(C_2, \sigma)$ and $f(C_3, \sigma)$ represent the similarity functions associated with the diatomic clusters C_2 and triatomic clusters C_3, respectively. The corresponding diameter of the cluster is denoted by $d(C_2)$ or $d(C_3)$. We perform 90,000 Monte carlo

samplings and minimize $f(C_2, \sigma)$ to approximately 0.06 and $f(C_3, \sigma)$ to approximately 0.18.

For magnetic disorder modeling, we refrain from using the SAE method again because it necessitates distinguishing spin-up and spin-down elements, leading to a higher-dimensional search space. Instead, we resort to an integer programming approach to distribute spin-up and spin-down sites uniformly in the configuration given by the SAE method, as a way to achieve the disordered local moment approximation [47]. We refer readers to Appendix B for more technical details. The generated atomic and magnetic configurations are illustrated in Fig. 6. The rightmost panel displays the radial distribution function of the magnetic elements, demonstrating a proper description of the magnetic disorder by the integer programming approach.

To calculate the static EOS for both the FM and PM states, we distribute 18 unit cell volumes in the range between 10.29

Å^3/atom and 13.12 Å^3/atom, corresponding to the static pressures from approximately 29.4 GPa to −12.9 GPa. For the FM state, we employ PANBB for the fixed-volume relaxation with colinear calculations. Following [25], the PM state is then described based on the relaxed FM structures, with the initial magnetic configuration determined by the integer programming approach. The energy-volume relations on both the FM and PM AlCoCrFeNi are depicted in Fig. 7, along with
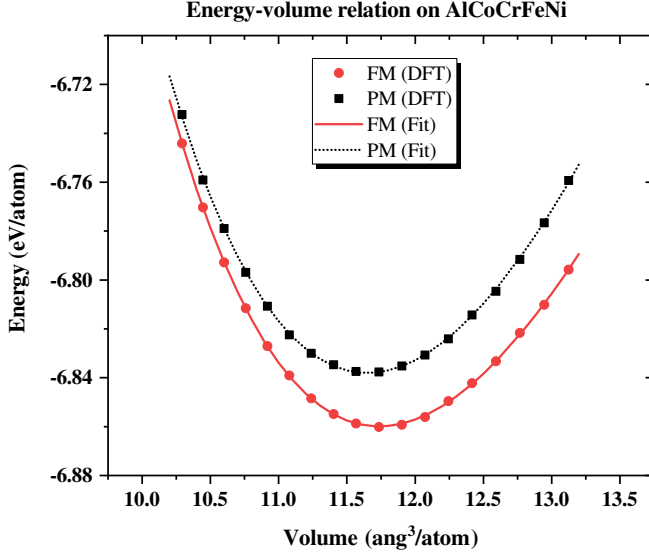

FIG. 7. The energy-volume relations on FM and PM AlCoCrFeNi. "DFT" indicates the *ab initio* results, while "Fit" is for the curve fitted to the static BM3 EOS. The red circles and solid line represent the results for the FM state, while the black squares and dotted line represent the results for the PM state.

the fitted third-order Birch-Murnaghan (BM3) EOS [48, 49]. We record the fitted parameters in Table I, where V_0 represents the equilibrium volume and E_0, B_0, and B_0' denotes the energy, bulk modulus, and its derivative with respect to pressure associated with V_0, respectively.

TABLE I. The fitted parameters of the static BM3 EOS.

References	V_0 (Å^3/atom)	E_0 (eV/atom)	B_0 (GPa)	B_0'
FM ([25])	11.77	-6.85	157.32	8.9
FM (PANBB)	11.74	-6.86	158.49	5.3
FM (CG)	11.73	-6.86	158.04	5.8
PM ([25])	11.70	-6.82	161.64	5.3
PM (PANBB)	11.65	-6.84	167.59	4.7
PM (CG)	11.65	-6.84	161.64	4.0

Compared with the previous calculations [25], our fitted V_0 and B_0 for both the FM and PM states, E_0 for the FM state, and B_0' for the PM state are comparable; see Table I. The gaps in B_0' for the FM state and E_0 for the PM state can result from two aspects: (i) we enhance magnetic disorder by the integer programming approach; (ii) [25] relaxes only the atomic positions, while PANBB optimizes both the atomic and lattice degrees of freedom. By the way, our value of B_0' appears to be more reasonable since it is closer to those

obtained for similar compositions [24]. Within the mean-field approximation [50], we can estimate the Curie temperature of AlCoCrFeNi as

$$\frac{2}{3(1-c)k_B}(E_0^{\mathrm{PM}} - E_0^{\mathrm{FM}}) \approx 210.8 \text{ K},$$

where c is the concentration of nonmagnetic atoms, k_B is the Boltzmann's constant, E_0^{FM} and E_0^{PM} are the energies of the FM and PM states associated with the equilibrium volumes from the BM3 EOS fitting based on the PANBB results, respectively; see Fig. 7. Our estimated Curie temperature lies in the interval of the critical temperature of Al$_x$CoCrFeNi [51], which is between 200 K ($x = 0.75$) and 400 K ($x = 1.25$), to some extent demonstrating rational descriptions of the FM and PM states.

To construct the EOS in the temperature (T) and pressure (P) space, the modified mean field potential (MMFP) approach [18] is used to describe the ion vibration energy as

$$F_{\mathrm{vib}}^{\mathrm{MMFP}} := -k_B T \left[\frac{3}{2} \ln \frac{M k_B T}{2\pi\hbar} + \ln v_f(V, T) \right],$$

$$v_f(V, T) := 4\pi \int \exp\left[\frac{p(r, V)}{k_B T} \right] r^2 \, dr,$$

$$p(r, V) := \frac{1}{2} [E_s(D + r) + E_s(D - r) - 2E_s(D)]$$
$$+ \frac{\lambda r}{2D} [E_s(D + r) - E_s(D - r)],$$

where $\hbar$ denotes the Planck constant, M is the relative atomic mass, D is the nearest atomic distance, r corresponds to the distance of the ion vibration away from its equilibrium position, E_s is the static EOS, and p is the constructed potential. Here, λ can take three values, −1, 0, 1, which originate from different forms of the Grüneisen coefficient. As a powerful tool to calculate the Helmholtz free energy, the MMFP effectively describes the ion vibration with the anharmonic effect, and has been widely used in thermodynamic physical properties simulations of HEAs [25, 45].

Since the effectiveness of the MMFP approach highly depends on the potential field constructed from the static EOS, we can validate the previous calculations by investigating the thermodynamic properties of AlCoCrFeNi. We first calculate the equilibrium volume and bulk modulus at ambient conditions with the MMFP approach using the PM PANBB results. Our values (approximately 11.83 Å^3/atom and 148.8 GPa) turn out to be close to those from the X-ray diffraction and diamond anvil cell experiments (11.79 Å^3/atom and 150±2.5 GPa) [26]. We then calculate the volume-pressure relation at 300 K with the PM PANBB results; see the solid line with label "300 K (MMFP)" in Fig. 8. At equivalent pressures, the maximum relative deviation of the volumes from the diamond anvil cell data [26] does not exceed 0.7%.

We would also like to compare PANBB with CG on the relaxation of the FM AlCoCrFeNi. PANBB converges normally at all volumes, while CG is found to fail at 4 volumes due to the breakdown of the LM process. Moreover, a sharp jump of c/a is found during the relaxation of CG when the pressure exceeds 21 GPa; see Fig. 9. The results of CG at the other 16 volumes are fitted by the BM3 EOS. The fitted parameters of
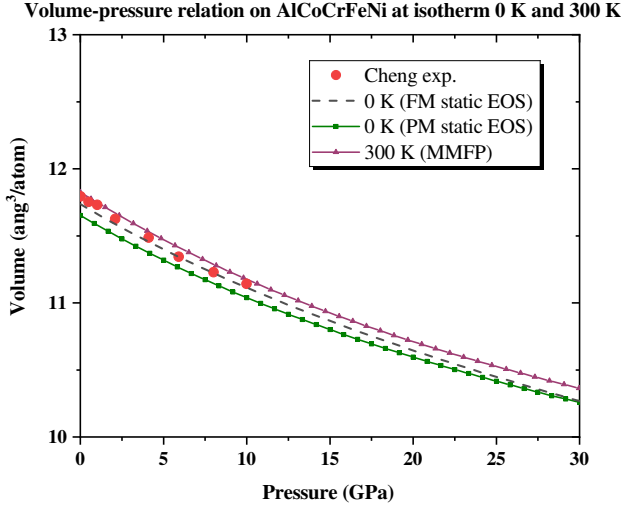
Volume-pressure relation on AlCoCrFeNi at isotherm 0 K and 300 K

FIG. 8. The volume-pressure relation on AlCoCrFeNi at 0 K and 300 K. "Cheng exp." represents the X-ray diffraction and diamond anvil cell experimental results [26], "0 K (FM static EOS)" refers to the relation by fitting the static EOS based on the FM PANBB results, "0 K (PM static EOS)" refers to the relation by fitting the static EOS based on the PM PANBB results, and "300 K (MMFP)" refers to the relation given by the MMFP approach based on the PM PANBB results.

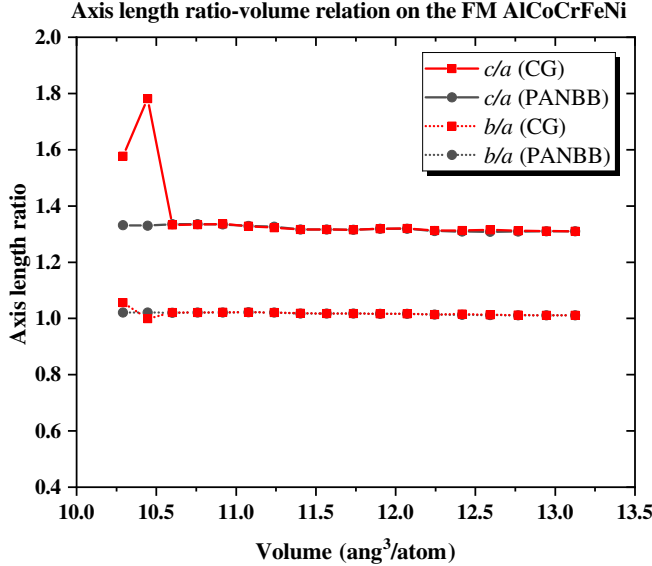**Axis length ratio-volume relation on the FM AlCoCrFeNi**

FIG. 9. The axis length ratio-volume relation on the FM AlCoCrFeNi. For the volumes where CG fails to converge normally, we use the axis length ratios of the final structures. The axis length ratios "c/a" and "b/a" are indicated by solid and dotted lines, respectively. The results of CG and PANBB are indicated by square and circle markers, respectively.

the static EOS of AlCoCrFeNi are summarized in Table I and consistent with those derived from the PANBB results. The average speedup factor of PANBB over CG in terms of running time is 1.36. Note that this speedup factor is calculated at the volumes where both methods converge normally and the energy differences per atom are not larger than 3 meV.

By the way, we illustrate the LLD achieved by fixed-volume relaxation at pressures of approximately 0 GPa, 21 GPa, and 29 GPa using the smearing radial distribution function. This stands in sharp contrast to the peaks at coordination shells in the ideal lattice, as depicted in Fig. 10. Notably, when the pressure does not exceed 21 GPa, the radial distribution functions of the structures relaxed by PANBB and CG display a high degree of consistency. However, as the pressure rises to 29 GPa, the radial distribution function obtained by PANBB exhibits smeared peaks consistent with those of the ideal lattice, while that obtained by CG shows significant displacements. This can be attributed to the previously mentioned issue of strong lattice distortion during the relaxation process of CG. Whether such lattice distortion conforms to physics deserves future investigations.

IV. CONCLUSIONS

In this work, we introduce PANBB, a novel force-based algorithm for relaxing crystal structures under a fixed unit cell volume. By incorporating the gradient projections, distinct curvature-aware initial trial step sizes, and a nonmonotone LM criterion, PANBB exhibits universal superiority over the widely used methods in both efficiency and robustness, as demonstrated across a benchmark test. To the best of our knowledge, PANBB is the first algorithm that guarantees theoretical convergence for the determinant-constrained optimization. Moreover, the application in the AlCoCrFeNi HEAs also showcases its potential in facilitating calculations of the wide-range EOS and thermodynamical properties for multicomponent alloys.

Appendix A: Convergence Analyses

In what follows, we rigorously establish the convergence of PANBB to equilibrium states regardless of the initial configurations. The analyses depend on the following two reasonable
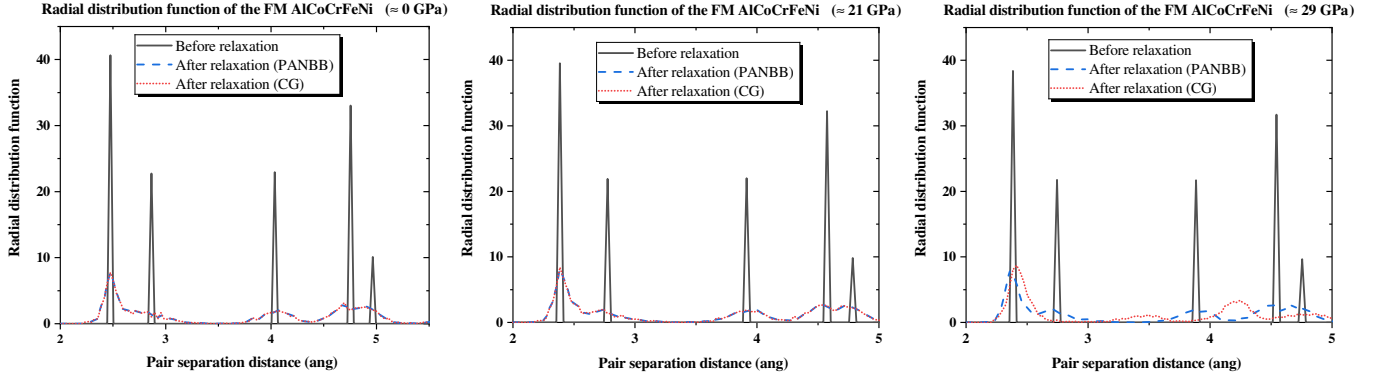
FIG. 10. Radial distribution functions of the FM AlCoCrFeNi at pressures of approximately 0 GPa (left), 21 GPa (middle), and 29 GPa (right). The black solid lines refer to the functions before relaxation, while the blue dashed and red dotted lines represent the functions after relaxation using PANBB and CG, respectively.

assumptions on the potential energy functional.

Assumption 1. *The potential energy functional E is locally Lipschitz continuously differentiable with respect to R and A. The stress tensor Σ is continuous with respect to R and A.*

Assumption 2. *The potential energy functional E is coercive over $\mathcal{F}$, namely, $E(R, A) \to \infty$ as $\|R\|_2 \to \infty$ or $\|A\|_2 \to \infty$ with $\det(A) = V$.*

For brevity, we make the following notation

$$\tilde{D}_{\text{latt}} := B^\top \tilde{F}_{\text{latt}}, \tag{A1}$$

where $\tilde{F}_{\text{latt}}$ is defined in Eq. (3), and define the level set

$$\mathcal{L} := \{(R, A) : E(R, A) \le E^{(0)}, \ \det(A) = V\}. \tag{A2}$$

In view of Eqs. (6) and (A1), $A_{\text{inter}}^{(k,\ell)}$ can be rewritten as

$$A_{\text{inter}}^{(k,\ell)} = A^{(k)}(I + \alpha_{\text{latt}}^{(k-1,\ell)} \tilde{D}_{\text{latt}}^{(k)}). \tag{A3}$$

Assumptions 1 and 2 directly imply the following lemma.

Lemma 1. *Suppose that Assumptions 1 and 2 hold. Then there exists an $M > 0$ such that if $(R, A) \in \mathcal{L}$, all of $\|R\|_{\text{F}}$, $\|A\|_{\text{F}}$, $\|B\|_{\text{F}}$, $\|F_{\text{atom}}\|_{\text{F}}$, and $\|\tilde{D}_{\text{latt}}\|_{\text{F}}$ are upper bounded by M.*

Proof. The existence of a uniform upper bound for $\|R\|_{\text{F}}$ and $\|A\|_{\text{F}}$ follows directly from Assumption 2 and the definition of the set $\mathcal{L}$ in Eq. (A2). That for $\|B\|_{\text{F}}$ is derived from $\det(A) = V$. The uniform boundedness of $\|F_{\text{atom}}\|_{\text{F}}$ and $\|\tilde{D}_{\text{latt}}\|_{\text{F}}$ are then deduced from Eqs. (1), (A1), and Assumption 1. ∎

Let $\bar{\alpha}_{\text{atom}}$, $\bar{\alpha}_{\text{latt}}$, $\underline{\alpha}_{\text{atom}}$, and $\underline{\alpha}_{\text{atom}}$ denote the upper and lower bounds for the ABB step sizes in Eq. (10). Based on Assumptions 1 and 2, some positive constants are defined using M in Lemma 1, δ_{atom}, δ_{latt} in Eq. (9) as well as $\bar{\alpha}_{\text{atom}}$ and $\bar{\alpha}_{\text{latt}}$ in Eq. (10).

$$L_{\nabla E} := \sup \left\{ \frac{\|\nabla E(R_1, A_1) - \nabla E(R_2, A_2)\|_{\text{F}}}{\|(R_1, A_1) - (R_2, A_2)\|_{\text{F}}} : \|R_i\|_{\text{F}} \le (1 + \bar{\alpha}_{\text{atom}})M, \ \|A_i\|_{\text{F}} \le \frac{3}{2}M, \ i = 1, 2 \right\},$$

$$L_{\text{latt}} := \sup \left\{ \frac{|E(R, A_1) - E(R, A_2)|}{\|A_1 - A_2\|_{\text{F}}} : \|R\|_{\text{F}} \le (1 + \bar{\alpha}_{\text{atom}})M, \ \|A_i\|_{\text{F}} \le 3M, \ i = 1, 2 \right\}, \tag{A4}$$

$$\bar{M} := \frac{48}{7}M^3 L_{\text{latt}}, \quad \bar{\ell} := \left\lceil \max \left\{ \frac{\log \frac{L_{\nabla E} \bar{\alpha}_{\text{atom}}}{2(1-\eta)}}{\log \frac{1}{\delta_{\text{atom}}}}, \frac{\log \frac{(L_{\nabla E} + 2\bar{M})\bar{\alpha}_{\text{latt}}}{2(1-\eta)}}{\log \frac{1}{\delta_{\text{latt}}}}, \frac{\log(2M\bar{\alpha}_{\text{latt}})}{\log \frac{1}{\delta_{\text{latt}}}} \right\} \right\rceil.$$

Lemma 2 below asserts that the LM processes in PANBB are always finite.

Lemma 2. *Suppose that Assumptions 1 and 2 hold. Assume that $(R^{(k)}, A^{(k)}) \in \mathcal{L}$. Then in the $(k+1)$th iteration, the LM criterion (7) must be satisfied after invoking the backtracking (9) for at most $\bar{\ell}$ times and $(R^{(k+1)}, A^{(k+1)}) \in \mathcal{L}$, where $\bar{\ell}$ is defined in Eq. (A4).*

Proof. We prove this lemma by showing that there is a sufficient reduction from $E^{(k)}$ to $E_{\text{trial}}^{(k,\ell)}$ provided that $\ell \ge \bar{\ell}$, which indicates the fulfillment of the criterion (7). We divide the proof into two steps.

Step 1. Estimate $E(R_{\text{trial}}^{(k,\ell)}, A_{\text{inter}}^{(k,\ell)}) - E^{(k)}$.

It follows from Eqs. (2) and (6) that

$$\left\langle F_{\text{atom}}^{(k)}, R_{\text{trial}}^{(k,\ell)} - R^{(k)} \right\rangle = -\frac{1}{\alpha_{\text{atom}}^{(k-1,\ell)}} \|R_{\text{trial}}^{(k,\ell)} - R^{(k)}\|_{\text{F}}^2,$$

$$\left\langle F_{\text{latt}}^{(k)}, A_{\text{inter}}^{(k,\ell)} - A^{(k)} \right\rangle = -\frac{1}{\alpha_{\text{latt}}^{(k-1,\ell)}} \|A_{\text{inter}}^{(k,\ell)} - A^{(k)}\|_{\text{F}}^2. \tag{A5}$$

Since $\ell \geq \bar{\ell}$, it is not hard to show $\alpha_{\text{latt}}^{(k-1,\ell)} \leq 1/(2M)$. From this, $\alpha_{\text{atom}}^{(k-1,\ell)} \leq \bar{\alpha}_{\text{atom}}$, and Lemma 1, we can infer $\alpha_{\text{latt}}^{(k-1,\ell)}\|\tilde{D}_{\text{latt}}^{(k)}\|_2 \leq 1/2$ and

$$\|R_{\text{trial}}^{(k,\ell)}\|_{\text{F}} \overset{(6)}{\leq} (1 + \bar{\alpha}_{\text{atom}})M, \tag{A6}$$

$$\|A_{\text{inter}}^{(k,\ell)}\|_{\text{F}} \overset{(A3)}{\leq} \|A^{(k)}\|_{\text{F}}(1 + \alpha_{\text{latt}}^{(k-1,\ell)}\|\tilde{D}_{\text{latt}}^{(k)}\|_2) \leq \frac{3}{2}M. \tag{A7}$$

By the definition of $L_{\nabla E}$ in Eq. (A4), one can derive from Eqs. (A5)-(A7) that $E(R_{\text{trial}}^{(k,\ell)}, A_{\text{inter}}^{(k,\ell)}) - E^{(k)}$ is upper bounded by

$$\left[\frac{L_{\nabla E}}{2} - \frac{1}{\alpha_{\text{atom}}^{(k-1,\ell)}}\right] \|R_{\text{trial}}^{(k,\ell)} - R^{(k)}\|_{\text{F}}^2$$

$$+ \left[\frac{L_{\nabla E}}{2} - \frac{1}{\alpha_{\text{latt}}^{(k-1,\ell)}}\right] \|A_{\text{inter}}^{(k,\ell)} - A^{(k)}\|_{\text{F}}^2. \tag{A8}$$

Step 2. Estimate $E_{\text{trial}}^{(k,\ell)} - E(R_{\text{trial}}^{(k,\ell)}, A_{\text{inter}}^{(k,\ell)})$.

First notice from Eq. (A3) that $\det(A_{\text{inter}}^{(k,\ell)}) = V\det(I + \alpha_{\text{latt}}^{(k-1,\ell)}\tilde{D}_{\text{latt}}^{(k)})$. Due to $\alpha_{\text{latt}}^{(k-1,\ell)}\|\tilde{D}_{\text{latt}}^{(k)}\|_2 \leq 1/2$ and the fact that the modulus of any eigenvalue is not larger than the maximum singular value, one has

$$\det(A_{\text{inter}}^{(k,\ell)}) \in \left[\frac{1}{8}V, \frac{27}{8}V\right]. \tag{A9}$$

By the definition of the scaling operation (see S4 in Fig. 1), Eqs. (A7) and (A9), $\|A_{\text{trial}}^{(k,\ell)}\|_{\text{F}} \leq 3M$. From this,

$$E_{\text{trial}}^{(k,\ell)} - E(R_{\text{trial}}^{(k,\ell)}, A_{\text{inter}}^{(k,\ell)})$$

$$\overset{(A4),(A6)}{\leq} L_{\text{latt}}\|A_{\text{trial}}^{(k,\ell)} - A_{\text{inter}}^{(k,\ell)}\|_{\text{F}}$$

$$= L_{\text{latt}}\left|\sqrt[3]{\det(A_{\text{inter}}^{(k,\ell)})} - \sqrt[3]{V}\right| \frac{\|A_{\text{inter}}^{(k,\ell)}\|_{\text{F}}}{\sqrt[3]{\det(A_{\text{inter}}^{(k,\ell)})}}. \tag{A10}$$

For one thing, in view of Eqs. (A7) and (A9),

$$\frac{\|A_{\text{inter}}^{(k,\ell)}\|_{\text{F}}}{\sqrt[3]{\det(A_{\text{inter}}^{(k,\ell)})}} \leq \frac{3M}{\sqrt[3]{V}}. \tag{A11}$$

For another, again by Eq. (A9),

$$\left|\sqrt[3]{\det(A_{\text{inter}}^{(k,\ell)})} - \sqrt[3]{V}\right|$$

$$= \frac{\left|\det(A_{\text{inter}}^{(k,\ell)}) - V\right|}{\det(A_{\text{inter}}^{(k,\ell)})^{2/3} + \det(A_{\text{inter}}^{(k,\ell)})^{1/3}V^{1/3} + V^{2/3}}$$

$$\leq \frac{4\left|\det(A_{\text{inter}}^{(k,\ell)}) - V\right|}{7V^{2/3}}. \tag{A12}$$

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of $\alpha_{\text{latt}}^{(k-1,\ell)}\tilde{D}_{\text{latt}}^{(k)}$. Since $\tilde{F}_{\text{latt}}^{(k)}$ lies on $\mathcal{T}^{(k)}$ (see the definition in Eq. (2)), one has $\sum_{i=1}^3 \lambda_i = 0$ from Eq. (A1). Therefore,

$$\left|\det(A_{\text{inter}}^{(k,\ell)}) - V\right| \overset{(A3)}{=} V\left|\det(I + \alpha_{\text{latt}}^{(k-1,\ell)}\tilde{D}_{\text{latt}}^{(k)}) - 1\right|$$

$$= V\left|\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_3\right|.$$

Since $\max_i\{|\lambda_i|\} \leq \alpha_{\text{latt}}^{(k-1,\ell)}\|\tilde{D}_{\text{latt}}^{(k)}\|_2 \leq 1/2$,

$$\left|\det(A_{\text{inter}}^{(k,\ell)}) - V\right| \leq 4V\max_i\{|\lambda_i|\}^2 \leq 4V\left\|\alpha_{\text{latt}}^{(k-1,\ell)}\tilde{D}_{\text{latt}}^{(k)}\right\|_2^2$$

$$\overset{(A3)}{=} 4V\left\|B^{(k)\top}(A_{\text{inter}}^{(k,\ell)} - A^{(k)})\right\|_2^2$$

$$\leq 4M^2V\|A_{\text{inter}}^{(k,\ell)} - A^{(k)}\|_{\text{F}}^2, \tag{A13}$$

where the last inequality uses Lemma 1. Combining Eq. (A4) and Eqs. (A10)-(A13), one obtains

$$E_{\text{trial}}^{(k,\ell)} - E(R_{\text{trial}}^{(k,\ell)}, A_{\text{inter}}^{(k,\ell)}) \leq \bar{M}\|A_{\text{inter}}^{(k,\ell)} - A^{(k)}\|_{\text{F}}^2. \tag{A14}$$

Finally, putting Eqs. (A8) and (A14) together, we have

$$E_{\text{trial}}^{(k,\ell)} - E^{(k)} \leq \left[\frac{L_{\nabla E}}{2} - \frac{1}{\alpha_{\text{atom}}^{(k-1,\ell)}}\right]\|R_{\text{trial}}^{(k,\ell)} - R^{(k)}\|_{\text{F}}^2$$

$$+ \left[\frac{L_{\nabla E}}{2} + \bar{M} - \frac{1}{\alpha_{\text{latt}}^{(k-1,\ell)}}\right]\|A_{\text{inter}}^{(k,\ell)} - A^{(k)}\|_{\text{F}}^2.$$

By virtue of the criterion (7) and $E^{(k)} \leq \bar{E}^{(k)}$ [23, Lemma 2], the LM process must stop when

$$\begin{cases} \dfrac{L_{\nabla E}}{2} - \dfrac{1}{\alpha_{\text{atom}}^{(k-1,\ell)}} \leq -\dfrac{\eta}{\alpha_{\text{atom}}^{(k-1,\ell)}}, \\ \dfrac{L_{\nabla E}}{2} + \bar{M} - \dfrac{1}{\alpha_{\text{latt}}^{(k-1,\ell)}} \leq -\dfrac{\eta}{\alpha_{\text{latt}}^{(k-1,\ell)}}, \end{cases}$$

which is fulfilled once ℓ reaches $\bar{\ell}$. Since $E(R_{\text{trial}}^{(k,\ell)}, A_{\text{trial}}^{(k,\ell)}) \leq \bar{E}^{(k)} \leq E^{(0)}$, $(R^{(k+1)}, A^{(k+1)}) \in \mathcal{L}$. ∎

Leveraging Lemma 2, we are ready to establish the convergence of PANBB.

Theorem 1. *Suppose that Assumptions 1 and 2 hold. Let $\{(R^{(k)}, A^{(k)})\}$ be the configuration sequence generated by PANBB with $\mu_k \in [\mu_{\min}, \mu_{\max}] \subseteq (0, 1]$. Then we have $F_{\text{atom}}^{(k)} \to 0$ and $\tilde{F}_{\text{latt}}^{(k)} \to 0$ as $k \to \infty$. Moreover, there exists at least one limit point of $\{(R^{(k)}, A^{(k)})\}$ and any limit point $(R^\star, A^\star)$ satisfies $F_{\text{atom}}^\star = 0$, $\Sigma_{\text{dev}}^\star = 0$.*

Proof. By Lemma 2, the backtracking scheme (9), and $\alpha_{\text{atom}}^{(k-1,0)} \geq \underline{\alpha}_{\text{atom}}$, $\alpha_{\text{latt}}^{(k-1,0)} \geq \underline{\alpha}_{\text{latt}}$, it holds that

$$\alpha_{\text{atom}}^{(k)} \geq \beta_{\text{atom}} := \underline{\alpha}_{\text{atom}}\delta_{\text{atom}}^{\bar{\ell}}, \quad \alpha_{\text{latt}}^{(k)} \geq \beta_{\text{latt}} := \underline{\alpha}_{\text{latt}}\delta_{\text{latt}}^{\bar{\ell}}.$$

Plugging this into the criterion (7), we obtain

$$E^{(k+1)} \leq \bar{E}^{(k)} - \eta\left[\beta_{\text{atom}}\|F_{\text{atom}}^{(k)}\|_{\text{F}}^2 + \beta_{\text{latt}}\|\tilde{F}_{\text{latt}}^{(k)}\|_{\text{F}}^2\right]. \tag{A15}$$

Let $\Delta^{(k)} := \beta_{\text{atom}} \|F_{\text{atom}}^{(k)}\|_F^2 + \beta_{\text{latt}} \|\tilde{F}_{\text{latt}}^{(k)}\|_F^2$. From this, a recursion for the monitoring sequence follows.

$$
\begin{aligned}
\bar{E}^{(k+1)} &\overset{(8)}{=} \frac{\bar{E}^{(k)} + \mu^{(k)} q^{(k)} E^{(k+1)}}{q^{(k+1)}} \\
&\overset{(A15)}{\leq} \frac{\bar{E}^{(k)} + \mu^{(k)} q^{(k)} (\bar{E}^{(k)} - \eta \Delta^{(k)})}{q^{(k+1)}} \\
&= \bar{E}^{(k)} - \eta \frac{\mu^{(k)} q^{(k)}}{q^{(k+1)}} \Delta^{(k)}. \quad (A16)
\end{aligned}
$$

A byproduct of Eq. (A16) is the monotonicity of $\{\bar{E}^{(k)}\}$. Since E is bounded from below over $\mathcal{L}$ (by Assumption 1 and Lemma 1) and $E^{(k)} \leq \bar{E}^{(k)}$ for any $k \geq 0$ [23, Lemma 2], $\{\bar{E}^{(k)}\}$ is also bounded from below. Consequently,

$$
\sum_{k=0}^{\infty} \frac{\mu^{(k)} q^{(k)} \Delta^{(k)}}{q^{(k+1)}} \overset{(A16)}{\leq} \frac{1}{\eta} \sum_{k=0}^{\infty} (\bar{E}^{(k)} - \bar{E}^{(k+1)}) < +\infty. \quad (A17)
$$

Following from Eq. (8) and $\mu_{\max} \geq \mu^{(k)} \geq \mu_{\min} > 0$, $q^{(k)} \geq 1$, we can derive

$$
\frac{\mu^{(k)} q^{(k)}}{q^{(k+1)}} \geq \frac{\mu_{\min} q^{(k)}}{\mu^{(k)} q^{(k)} + 1} = \frac{\mu_{\min}}{\mu^{(k)} + 1/q^{(k)}} \geq \frac{\mu_{\min}}{\mu_{\max} + 1},
$$

which, together with Eq. (A17), yields $\Delta^{(k)} \to 0$ as $k \to \infty$. By $\beta_{\text{atom}}, \beta_{\text{latt}} > 0$, the definition of $\Delta^{(k)}$, and Lemma 1, one has $F_{\text{atom}}^{(k)} \to 0$ and $\tilde{F}_{\text{latt}}^{(k)} \to 0$ as $k \to \infty$.

Let $(R^\star, A^\star)$ be a limit point of $\{(R^{(k)}, A^{(k)})\}$, whose existence is guaranteed by Lemma 1. Then by the above arguments and Eq. (3), $F_{\text{atom}}^\star = 0$ and

$$
\begin{aligned}
0 = \tilde{F}_{\text{latt}}^\star &= F_{\text{latt}}^\star - \frac{\langle B^\star, F_{\text{latt}}^\star \rangle}{\|B^\star\|_F^2} B^\star \\
&= V \left[\Sigma^\star B^\star - \frac{\langle B^\star, \Sigma^\star B^\star \rangle}{\|B^\star\|_F^2} B^\star \right],
\end{aligned}
$$

which implies that $\Sigma^\star$ is a multiple of I and hence $\Sigma_{\text{dev}}^\star = 0$. ∎

Appendix B: Magnetic Disorder Modeling

We elaborate below on the technical details when modeling the magnetic disorder for the PM structures, in particular, how to solve the distribution problem for the spin-up or spin-down sites. We assume that an atomic configuration is available (e.g., given by the SAE method), and moreover introduce some notations:

- $S \in \mathbb{N}$: number of the atomic species (in the AlCoCrFeNi case, equals 5);

- $S_{\text{mag}} \in \mathbb{N}$: number of the magnetic atomic species (in the AlCoCrFeNi case, equals 4). For convenience, we assume that species $1, \ldots, S_{\text{mag}}$ are magnetic;

- $N_s \in \mathbb{N}$: number of the atoms in the unit cell belonging to species s (in the AlCoCrFeNi case, $N_1 = \cdots = N_5 = 32$) ($s \in \{1, \ldots, S\}$);

- $N_{\text{mag}} \in \mathbb{N}$: number of magnetic atoms in the unit cell (in the AlCoCrFeNi case, equals 64);

- $\ell_{\text{ax}} \in \mathbb{N}$: number of the atomic layers in the unit cell with axis $\text{ax} \in \{a, b, c\}$ as normal;

- $\mathcal{I}_s \subseteq \{1, \ldots, N\}$: indices of the sites occupied by species s ($s \in \{1, \ldots, S_{\text{mag}}\}$);

- $\mathcal{I}_{\text{ax},l,s} \subseteq \{1, \ldots, N\}$: indices of the sites occupied by species s located on the lth atomic layer with axis ax as normal ($\text{ax} \in \{a, b, c\}$, $l \in \{1, \ldots, \ell_{\text{ax}}\}$, $s \in \{1, \ldots, S_{\text{mag}}\}$);

- $\mathbf{y} \in \{0, 1\}^{N_{\text{mag}}}$: binary indicators for the spin-up configuration, where $y_n = 1$ ($n \in \{1, \ldots, N_{\text{mag}}\}$) means that the nth atom spins up. Conversely, one may interpret the zero entries as spin-down atoms;

- $\mathbf{y}^0 \in \{0, 1\}^{N_{\text{mag}}}$: binary indicators for the trial spin-up configuration. The meaning of each entry is analogous to that of $\mathbf{y}$.

With the above notations, we can formulate the spin-up distribution problem as the following binary integer programming:

$$
\begin{aligned}
\max_{\mathbf{y}} \quad & \langle \mathbf{y}^0, \mathbf{y} \rangle, \\
\text{s. t.} \quad & \lfloor N_s/(2\ell_{\text{ax}}) \rfloor \leq \sum_{n \in \mathcal{I}_{\text{ax},l,s}} y_n \leq \lceil N_s/(2\ell_{\text{ax}}) \rceil, \\
& l \in \{1, \ldots, \ell_{\text{ax}}\}, \ \text{ax} \in \{a, b, c\}, \ s \in \{1, \ldots, S_{\text{mag}}\}, \\
& \sum_{n \in \mathcal{I}_s} y_n = N_s/2, \ s \in \{1, \ldots, S_{\text{mag}}\}, \\
& \mathbf{y} \in \{0, 1\}^{N_{\text{mag}}}.
\end{aligned}
$$
$$\text{(B1)}$$

The integer programming (B1) maximizes the similarity between the trial and optimized spin-up structures, represented by the inner product of their corresponding binary indicators. In our experiments, we simply take $\mathbf{y}^0 = 0$. The first line of constraints enforces that the spin-up sites of each species are evenly distributed over each atomic layer. The second line of constraints imposes that the numbers of spin-up and spin-down sites for each species are identical. Upon solving the integer programming (B1) (e.g., by off-the-shelf software like MATLAB [52]), one can transform the solution $\mathbf{y}^\star$ to a concrete magnetic configuration.

[1] A. R. Oganov and C. W. Glass, J. Chem. Phys. **124**, 244704 (2006).

[2] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Phys. Rev. B **82**, 094116

(2010).

[3] Z. Chen, W. Jia, X. Jiang, S. Li, and L.-W. Wang, Comput. Phys. Commun. **219**, 35 (2017).

[4] Y. Wang, J. Lv, P. Gao, and Y. Ma, Acc. Chem. Res. **55**, 2068–2076 (2022).

[5] S. Goedecker, W. Hellmann, and T. Lenosky, Phys. Rev. Lett. **95**, 055501 (2005).

[6] L. B. Vilhelmsen and B. Hammer, Phys. Rev. Lett. **108**, 126101 (2012).

[7] C. H. Hu, A. R. Oganov, Q. Zhu, G. R. Qian, G. Frapper, A. O. Lyakhov, and H. Y. Zhou, Phys. Rev. Lett. **110**, 165504 (2013).

[8] Q. Li, D. Zhou, W. Zheng, Y. Ma, and C. Chen, Phys. Rev. Lett. **110**, 136403 (2013).

[9] L. Zhang, J. Luo, A. Saraiva, B. Koiller, and A. Zunger, Nat. Commun. **4**, 1 (2013).

[10] P. Vinet, J. H. Rose, J. Ferrante, and J. R. Smith, J. Phys.: Condens. Matter **1**, 1941 (1989).

[11] A. J. Jackson, J. M. Skelton, C. H. Hendon, K. T. Butler, and A. Walsh, J. Chem. Phys. **143**, 184101 (2015).

[12] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, NY, USA, 1996).

[13] P. Carrier, R. Wentzcovitch, and J. Tsuchiya, Phys. Rev. B **76**, 064116 (2007).

[14] Y. Wang, D. Chen, and X. Zhang, Phys. Rev. Lett. **84**, 3220 (2000).

[15] Y. Wang, Phys. Rev. B **61**, R11863 (2000).

[16] Y. Wang and L. Li, Phys. Rev. B **62**, 196 (2000).

[17] L. Li and Y. Wang, Phys. Rev. B **63**, 245108 (2001).

[18] H.-F. Song and H.-F. Liu, Phys. Rev. B **75**, 245126 (2007).

[19] E. Wasserman, L. Stixrude, and R. E. Cohen, Phys. Rev. B **53**, 8296 (1996).

[20] T. Jarlborg, E. G. Moroni, and G. Grimvall, Phys. Rev. B **55**, 1288 (1997).

[21] J. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Tech. Rep. (Carnegie Mellon University, Pittsburgh, PA, USA, 1994).

[22] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[23] Y. Hu, X. Gao, Y. Zhao, X. Liu, and H. Song, Phys. Rev. B **106**, 104101 (2022).

[24] H. Song, F. Tian, Q.-M. Hu, L. Vitos, Y. Wang, J. Shen, and N. Chen, Phys. Rev. Mater. **1**, 023404 (2017).

[25] J. Wu, Z. Yang, J. Xian, X. Gao, D. Lin, and H. Song, Front. Mater. **7**, 590143 (2020).

[26] B. Cheng, F. Zhang, H. Lou, X. Chen, P. K. Liaw, J. Yan, Z. Zeng, Y. Ding, and Q. Zeng, Scripta Mater. **161**, 88 (2019).

[27] V. Atalla, *Density-Functional Theory and Beyond for Organic Electronic Materials*, Ph.D. thesis, Technischen Universität Berlin, Berlin, Germany (2013).

[28] F. Knuth, *Strain and Stress: Derivation, Implementation, and Application to Organic Crystals*, Ph.D. thesis, Freie Universität Berlin, Berlin, Germany (2015).

[29] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed., Springer Series in Operations Research and Financial Engineer-ing (Springer, New York, NY, USA, 2006).

[30] J. Barzilai and J. Borwein, IMA J. Numer. Anal. **8**, 141 (1988).

[31] Y. Dai and R. Fletcher, Numer. Math. **100**, 21 (2005).

[32] R. Fletcher, in *Optimization and Control with Applications*, Applied Optimization, Vol. 96, edited by L. Qi, K. Teo, and X. Yang (Springer, Boston, MA, USA, 2005) pp. 235–256.

[33] The specific initial step size for the atomic part is also used by CG realized in vasp and cessp. The numerical results of CG and PANBB will become more comparable. A small initial step size is chosen for the lattice part to prevent unexpected breakdown. In simulations, the performance of PANBB are not sensitive to the two values.

[34] We use a small δ_{atom} to end the LM process more quickly and help the ensuing ABB step sizes learn local curvatures better. The value of δ_{latt} is not so critical, since usually $\alpha_{\text{latt}}^{(k-1,0)}$ is already sufficiently small.

[35] J. Fang, X. Gao, H. Song, and H. Wang, J. Chem. Phys. **144**, 244103 (2016).

[36] X. Gao, Z. Mo, J. Fang, H. Song, and H. Wang, Comput. Phys. Commun. **211**, 54 (2017).

[37] Y. Zhou, H. Wang, Y. Liu, X. Gao, and H. Song, Phys. Rev. E **97**, 033305 (2018).

[38] J. Fang, X. Gao, and H. Song, Commun. Comput. Phys. **26**, 1196 (2019).

[39] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[40] F. Jollet, M. Torrent, and N. Holzwarth, Comput. Phys. Commun. **185**, 1246 (2014).

[41] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[42] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, APL Mater. **1**, 011002 (2013).

[43] S. S. Borysov, R. M. Geilhufe, and A. V. Balatsky, PLoS ONE **12**, e0171501 (2017).

[44] E. Dolan and J. Moré, Math. Program. **91**, 201 (2002).

[45] F. Tian, D. Lin, X. Gao, Y. Zhao, and H. Song, J. Chem. Phys. **153**, 034101 (2020).

[46] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer Programming*, Graduate Texts in Mathematics, Vol. 271 (Springer, Cham, Switzerland, 2014).

[47] B. L. Gyorffy, A. J. Pindor, J. Staunton, G. M. Stocks, and H. Winter, J. Phys. F Met. Phys. **15**, 1337 (1985).

[48] F. Birch, Phys. Rev. **71**, 809 (1947).

[49] A. Otero-de-la-Roza and V. Luaña, Comput. Phys. Commun. **182**, 1708 (2011).

[50] H. Ge, H. Song, J. Shen, and F. Tian, Mater. Chem. Phys **210**, 320 (2018).

[51] Y.-F. Kao, S.-K. Chen, T.-J. Chen, P.-C. Chu, J.-W. Yeh, and S.-J. Lin, J. Alloys Compd. **509**, 1607 (2011).

[52] The MathWorks Inc., MATLAB version: 9.13.0 (R2022b) (2022).