

## ARTICLE OPEN



# An automated predictor for identifying transition states in solids

Ketao Yin<sup>1,2,3,5</sup>, Pengyue Gao<sup>1,5</sup>, Xuecheng Shao<sup>1</sup>, Bo Gao<sup>1</sup>, Hanyu Liu<sup>1</sup>, Jian Lv<sup>1</sup>, John S. Tse<sup>1,3,4</sup>, Yanchao Wang<sup>1</sup> and Yanming Ma<sup>1,4</sup>

The minimum energy path (MEP) and transition state are two key parameters in the investigation of the mechanisms of chemical reactions and structural phase transformations. However, determination of transition paths in solids is challenging. Here, we present an evolutionary method to search for the lowest energy path and the transition state for pressure-induced structural transformations in solids without any user input or prior knowledge of possible paths. Instead, the initial paths are chosen stochastically by connecting randomly selected atoms from the initial to final structure. The MEP of these trials paths were computed and ranked in order of their energies. The matrix particle swarm optimization algorithm is then used to generate improved transition paths. The procedure is repeated until the lowest energy MEP is found. This method is validated by reproducing results of several known systems. The new method also successfully located the MEP for the direct low-temperature pressure induced transformation of face centered-cubic (FCC) silicon to the simple hexagonal(sh) phase and FCC lithium to a complex body centered-cubic c16 high-pressure phase. The proposed method provides a convenient, robust, and reliable approach to identify the MEP of phase transformations. The method is general and applicable to a variety of problems requiring the location of the transition state.

*npj Computational Materials* (2020)6:16; <https://doi.org/10.1038/s41524-020-0286-9>

## INTRODUCTION

Phase transitions of solids are a critical part of nature. Identification of the atomistic mechanism of a solid–solid phase transition is critical to understanding the physical process, and is thus essential to the development of synthesis strategies for new materials<sup>1,2</sup>. The atomistic mechanism of the phase transition is characterized by the trajectory with the lowest potential energy profile, or the minimum energy path (MEP), and the highest energy (saddle) point along this path is the transition state (TS)<sup>3–13</sup>. Several experimental techniques have been developed, for example, the recent cooled anion precursors and high-resolution electron spectroscopy were used to study molecular chemical reactions and led to the direct observation of reactive resonances of  $F+H_2$  (ref. <sup>3</sup>). However, these techniques are limited to simple molecules and are not applicable to the solid state. The experimental determination of the MEP and TS for solid–solid phase transitions is still challenging, and theoretical prediction of the TS of structural transitions has become a topic of great importance.

Considerable progress has been made in the development of algorithms for identifying the TS in phase transitions. The search strategy generally consists of two parts: saddle point searching and transition path searching. For the first part, the string method<sup>4</sup>, nudged elastic band (NEB) method<sup>5</sup>, doubly NEB method<sup>6–8</sup>, climbing image-NEB method<sup>9</sup>, and solid-state NEB method<sup>10</sup> have been used to locate the saddle point on the potential energy surface (PES) of a given transition path<sup>14–23</sup>. The second part consists of a search for the lowest-energy path connecting the initial and final structures. Various methods have been proposed for this search. The most notable approaches are the dimer method<sup>11</sup>, transition path sampling<sup>12</sup>, and stochastic

surface walking (SSW)<sup>13</sup>. The latter involves exploration of the entire PES using molecule dynamic (MD) simulations. Many successes have been achieved using these methods<sup>24</sup>. However, they are often feasible for simple structures which is limited by either (i) the need to guess the initial transition paths<sup>25</sup> or (ii) the substantial computational cost of long MD simulations. For example, direct simulation of the classical dynamics would require  $10^{12}$  force evaluations within a 0.5 eV energy-barrier window<sup>26</sup>. Thus, it is impractical to map the entire PES using traditional MD simulations.

The fundamental difficulty in determining the MEP is that the number of possible transition paths grows factorially ( $N!$ ) for a system consisting of  $N$  atoms given a chosen set of lattice vectors. Thus, a straightforward implementation of an exhaustive search strategy is impractical because of the astronomically large number of possible solutions. To address this problem, a new method for identifying the MEP and the TS for solid–solid phase transitions is proposed here that do not require any prior knowledge of probable reaction paths. The only required information is the initial states (IS) and final states (FS) from which a set of initial random paths are selected based on permutation of the relevant atoms connecting the IS and FS. The energy profile of each path is then examined using NEB calculations. Paths with low-energy barriers are retained and improved using the global particle swarm optimization (PSO) scheme. The procedure is repeated until the lowest MEP is determined. In addition, we find that the TS of a given transition path can be determined more rapidly by removing the fictitious “spring constraint” but allowing the direction of force acting on the image to be determined on-the-fly based on the energies of two neighboring images. The

<sup>1</sup>International Center of Computational Method and Software & State Key Lab of Superhard Materials, College of Physics, Jilin University, 130012 Changchun, China. <sup>2</sup>School of Physics and Electronic Engineering, Linyi University, 276005 Linyi, China. <sup>3</sup>Department of Physics and Engineering Physics, University of Saskatchewan, Saskatoon, SK S7N 5E2, Canada. <sup>4</sup>International Center of Future Science, Jilin University, 130012 Changchun, China. <sup>5</sup>These authors contributed equally: Ketao Yin, Pengyue Gao email: john.tse@usask.ca; wyc@calypso.cn; mym@jlu.edu.cn

reliability and performance of the method is validated by successful prediction of selected known phase transformations. The paper is organized as follows: first, the methodology and improvements made to the NEB method are described. Second, the method is applied to the known cases of the pressure-induced transformations of graphitic carbon to cubic and hexagonal diamond and the B4 (wurtzite) to B1 (rocksalt) phase transition in GaN. Third, the method is applied to provide insight into the unknown transition mechanism of the recently observed, low-temperature, kinetically controlled transition of FCC silicon to an sh structure and the more challenging long-sought-after pathway and associated change in the electronic structure during the high-pressure transition from a simple FCC lithium to a complex c16 structure.

## RESULTS

### Evolution strategy for transition state

The search for the MEP and corresponding TS for a solid–solid phase transition is a complex multi-dimensional problem with many local extrema. It is not possible to adopt simple exhaustive search strategies to explore the entire PES because of the extremely high computational cost. A global optimization algorithm with high efficiency is required to remove the randomness of the searched paths. The PSO algorithm has been successfully applied in many systems to resolve multi-dimensional problems. Its implementation has successfully predicted the structures of 3D crystals, 2D surface reconstructions and layer materials, and isolated clusters and molecules. It has also aided the search for functional materials (e.g. superhard and electride materials)<sup>27–38</sup>. However, unlike structure prediction involving continuous variables, the search for the MEP is a discrete global optimization problem for solid-to-solid phase transitions. In this manuscript, a new matrix PSO (MPSO) algorithm for discrete global optimization is proposed for the selection of an MEP from among a collection of candidate transition paths. An improved NEB method yielding significantly faster convergence to the TS was also developed.

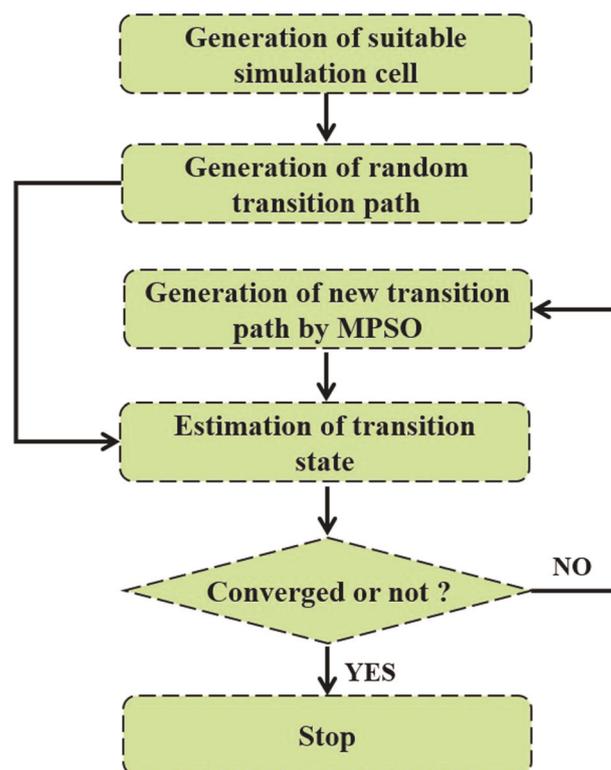
The proposed approach comprises four key steps depicted in the flow chart in Fig. 1: (i) generation of suitable simulation cells commensurate with the initial and final structures, (ii) generation of random transition paths, (iii) estimation of the TS and activation energy of the transition paths, and (iv) generation of improved transition paths using the MPSO method.

**Generation of a suitable simulation cell.** A crystal structure is represented by a lattice (unit cell) and the atom positions. Thus, the MEP for a solid–solid transition should be a function of the atom positions and lattice vectors. The key component in the search for an MEP is identification of the lowest-energy pathway connecting the initial and final structures. If the lattice mismatch between the initial and final states is large and non-negligible, the calculations will inevitably lead to an incorrect transition path with a higher energy barrier and inaccurate TS. To quantify the lattice mismatch, we use the following expression<sup>39</sup>:

$$\text{Min} = \sum_x |(X_{IS} - X_{FS})/X_{FS}|, \quad (1)$$

where  $X_{IS}$  and  $X_{FS}$  are the lattice parameters ( $a, b, c, \alpha, \beta, \gamma$ ) and volume of the initial and final structures, respectively. Supercells were constructed to find a suitable superlattice commensurate and with the smallest mismatch between the initial and final structures. The initial and final structures must have the same number of atoms. Once the simulation unit cell is determined, the goal of the global search approach is to determine the pathway with the lowest energy barrier to “move” atoms in the initial to the final structure.

**Generation of random transition path.** For a system containing  $N$  atoms, once a suitable simulation (supercell) cell is selected, the atoms in the model (i.e. initial and final structures) are numbered



**Fig. 1** Flow chart of the MPSO method. An overview of the proposed approach composed by four key steps.

as  $\{1, 2, 3, \dots, N\}$ . A permutation group is used to establish the mapping of atoms in the initial structure to the corresponding atoms in the final structure. The group element permuting the set of atoms can be represented using Cauchy’s two-line or matrix notation. For example, a system containing two atoms, designated as 1, 2 in the initial structure, that can map to 2, 1 in the final structure can be represented by the column vector  $(2 \ 1)$  or the transformation  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Any permutation of  $N$  atoms  $\{1, 2, \dots, N\}$  can be described by a similar permutation matrix (denoted by  $M$  in the following). Note that the matrix is orthogonal and that the order of the group is related to the number of atoms in the unit cell. A trivial permutation matrix is the  $N \times N$  identity matrix, which represents the mapping of 1, 2, 3, ...,  $N$  atomic positions in the initial structure to the same positions in the final structure.

To generate the stochastic transition paths, permutation matrices with one matrix element of every row or column is chosen randomly and assigned to 1 with the other matrix elements are set to 0. Because of the orthogonal condition of the matrices, each matrix represents a unique one-to-one correspondence between the atoms in the initial and final structures. After the random permutation matrices were constructed, the corresponding random candidate transition paths were generated.

**Estimation of the transition state and energy of the MEP path.** The TS is the highest energy point of a path on the PES. The energy barrier is the energy difference between the TS and the product. The transition paths, identified by their barrier energy bands, are used in the next step. Many methods can be used to estimate the TS, including the string method<sup>4</sup> and NEB method<sup>5</sup>. In this work, we employed the NEB method<sup>5</sup>. As will be described later, we observed that convergence to the TS can be improved by removing the artificial force constant and redefining the direction of the force toward the MEP.

**Evolution of transition path using MPSO.** The search for the lowest-energy transition path is similar to the traveling salesman problem (TSP)<sup>40</sup>. Instead of finding the shortest path between cities, the objective here is to locate the lowest energy path connecting the initial and final structures. Transition path sampling is a powerful tool in the search for the lowest energy path between defined endpoints. However, this method requires many trial trajectories shooting from configurations along the existing transition paths and involves the computation of the relevant dynamics which are prohibitively expensive. In comparison, to explore the PES the MPSO procedure from the TSP problem can be adopted to evolve the transition paths. In this case, a candidate transition path is designated as a “particle”, with the “pathway” represented by a (permutation) matrix  $M$  defining the linkages between atoms in the initial and corresponding new structure. In the initialization step, these matrices are constructed from the random permutations of the atoms in the supercell. A set of individual particles is called a population or a generation, where each particle is represented by an  $N \times N$  matrix ( $N$  is the number of atoms in the simulation cell). The energy barrier for a given transition path is used as the fitness descriptor.

After the initialization described above, the energy barrier of each candidate pathway is computed using the NEB method (see below). Subsequently, the pathways (permutation matrices) are ranked in order of increasing energy. The next step is to construct new pathways (permutation matrices) by rearranging the atom sequence to diversify the search for low-energy pathways. These matrices are constructed using knowledge from the most favorable pathway at the current generation “pbest” (represented by  $M_{i,pbest,t}$ ) and the global position with the best fitness value for the entire population “gbest” ( $M_{i,gbest,t}$ ) using Eq. (2):

$$v_{i,t+1} = v_{i,t} \otimes M_{i,pbest,t} \otimes M_{i,gbest,t}. \quad (2)$$

The position,  $x_{i,t+1}$ , of the next step is updated (evolved) from the previous position ( $x_{i,t}$ ) using a “velocity matrix” ( $v_{i,t+1}$ ) as defined in Eq. (3):

$$x_{i,t+1} = x_{i,t} \otimes v_{i,t+1}. \quad (3)$$

For this purpose, non-trivial new pathways are generated from a series of  $v_{i,t+1}$  matrices constructed by exchanging two arbitrarily

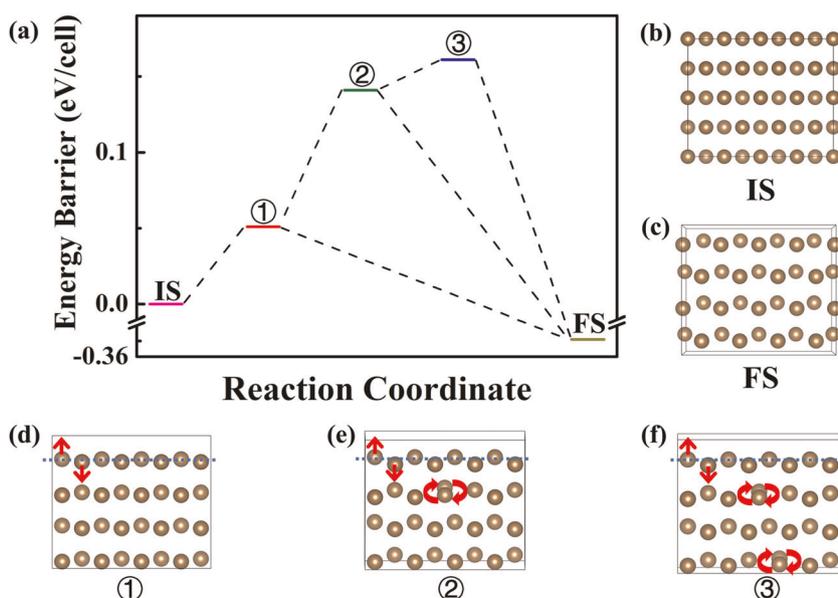
selected rows of atoms from  $M_{i,pbest,t}$  and  $M_{i,gbest,t}$  while keeping the other matrix elements subject to the condition that the product of these  $v$  matrices is the same as  $M_{i,pbest,t}$  or  $M_{i,gbest,t}$ . These operations maintain the “good” portion of the previously found low-energy path(s) and diversify the exploration of the PES by permutation of some atoms. Transition barrier calculations are performed on new paths (selected as those having the least forces acting on the atoms). Paths with high initial forces are deemed to be unrealistic and discarded. This procedure is repeated until no new lower-energy pathway is found for at least 30 generations.

The algorithm can be succinctly summarized as follows:

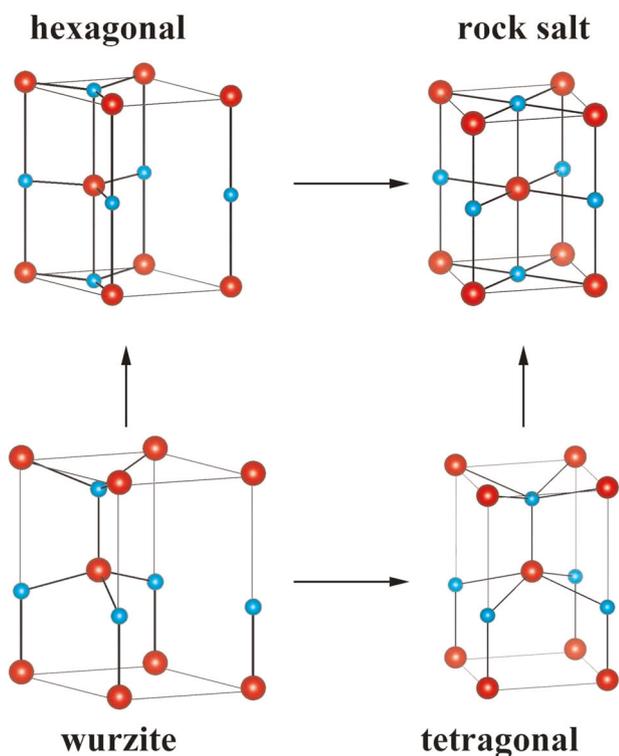
1. Initialization:
  - (a) set the initial “velocity” ( $v$ ) to the identity matrix and
  - (b) construct position vectors  $\{x_1, x_2, \dots, x_n\}$  by randomizing the  $n$  atom positions in the model.
2. Construct random paths (permutation matrices) linking the initial and final atom positions  $\{M_1, M_2, \dots, M_m\}$ , where  $m$  is population size. Compute and rank the energy at the TS of each path computed using the NEB method  $\{E_1, E_2, \dots, E_m\}$ .
3. Determine the local best  $M_{pbest}$  and global  $M_{gbest}$  matrices (paths).
4. Construct new paths ( $M$  matrices) by maintaining some part of  $M_{pbest}$  (i.e., fix selected rows but interchange the remaining rows randomly).
5. Compute new “velocity” using  $v_{i+1} = v_i \otimes M_{pbest} \otimes M_{gbest}$  and update the position vector,  $x_{i+1} = x_i \otimes v_{i+1}$ .
6. Repeat steps 2–6 using the updated  $M$  and  $x$  until convergence is achieved.

Validation by well-known transition paths

**Au.** The first example is meant to illustrate the process of MPSO in the search for the TS of the transformation from a cubic structure to a low-energy hexagonal structure predicted by structural search of a gold (Au) solid. The supercell consisting of 32 atoms is commensurate with the IS (Fig. 2b) and FS (Fig. 2c). The glue potential model<sup>41</sup> was used to describe the Au–Au interactions. Geometry optimizations were performed with the GULP<sup>42</sup> code. Several transition paths with low-energy barrier were found by MPSO and the relative energetics are compared in



**Fig. 2** Transition paths and structures for the transformation of Au. **a** Predicted energy barriers of transition states from IS to FS. The crystal structures of the IS **b** and FS **c**. Several transition paths **d–f** found by the MPSO are shown in the bottom panels.



**Fig. 3** Two transition paths for the transformation from the initial wurtzite structure of ZnO to the final rocksalt structure. The respective transition states structures, tetragonal (bottom) and hexagonal (top) are illustrated. The red and cyan balls represent Zn and O atoms, respectively.

Fig. 2a. The corresponding transition state structures are shown in Fig. 2d–g. Among these paths, path I has the lowest energy barrier of 0.091 V/cell and involves puckering of Au atoms in the same layer (Fig. 2a). Path II has a slightly higher barrier (0.141 eV/cell). In addition to puckering of Au atoms in one layer, two adjacent atoms in the same but different layer were exchanged. Path III with even higher energy barrier of 0.162 V/cell involves three processes: puckering of two Au atoms and exchange of two pair of adjacent atoms in different layers. This simple example shows how MPSO can be used efficiently to explore the PES to reveal possible low activation energy phase transition paths.

**ZnO.** To validate the proposed method, the well-known transformation from wurtzite (B4) ZnO to the rocksalt (B1) structure was studied. For this system, a supercell commensurate with the initial and final structure is a model with eight formula units (f.u.) (or 16 atoms) of ZnO. Previous studies have suggested that there are two candidate paths through the tetragonal or hexagonal intermediates<sup>43</sup> (Fig. 3). First, random reconstruction paths were generated using the MPSO, and these paths were then optimized using the NEB and NEB-solid state (NEB-SSNEB) methods with the spring constraint (to maintain the system on the MEP) to locate the TS. The TS with a tetragonal structure was found. The alternate path with the orthorhombic structure is 0.041 eV/atom higher in energy. Both the MEP and energetics are in good agreement with the findings from a previous study<sup>43</sup>. In this example, about 100 paths were sampled before convergence was reached.

After some experimentation, we observed that the convergence of locating the TS on the MEP can be accelerated by removing the spring constraint and using linear interpolation<sup>9,44–46</sup> to determine the direction of the image on the PES. In this method, a string with  $n + 1$  images denoted by  $(R_0, R_1, R_2, \dots, R_n)$  was used to estimate the transition path.  $R_0$  and  $R_N$  are fixed as they are the endpoints

of the initial and final structures. The  $n - 1$  intermediate images are adjusted using the steepest descent minimization algorithm shown in Fig. 4a. The key to the rapid convergence of the linear interpolation relies on the estimated tangent (direction) to optimize the TS (intermediate structure). In most proposed methods<sup>30–32</sup>, a local tangent to the path is estimated as the line segment connecting the previous and subsequent images in the chain according to Eq. (4),

$$\hat{\tau} = \frac{R_{i+1} - R_{i-1}}{|R_{i+1} - R_{i-1}|}. \quad (4)$$

Then the image structure is optimized. As shown pictorially in Fig. 4b, owing to the parallel component of the true force, relaxation along the tangent may lead to a structure falling back to the endpoints  $R_0$  or  $R_N$ . Thus, a correction term (fake force), such as the spring force in the NEB method, Lagrange multiplier of the string method, or Gaussian function in SSW, is employed to avoid this problem. However, the introduction of fake forces may impede the optimization of the transition path, even leading to imprecise results in some cases<sup>25</sup>.

An alternate tangent can be defined as in Eq. (5) with the direction of the forces used to optimize the images calculated according to Eq. (6). The optimized force in Fig. 4b does not have a parallel component. Thus, the optimization along the vertical direction of the tangent will not fall back to the initial or final state, thereby eliminating the need for a fake constraint force.

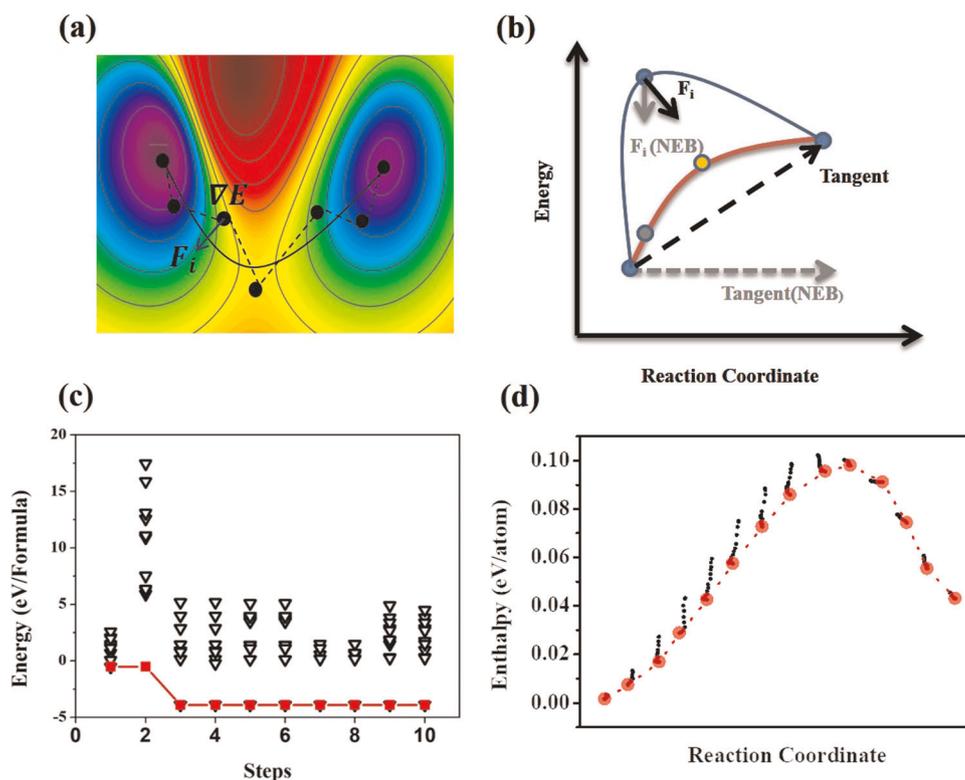
$$\hat{\tau} = \frac{F(R_{i+1}) - F(R_i)}{|F(R_{i+1}) - F(R_i)|} + \frac{F(R_i) - F(R_{i-1})}{|F(R_i) - F(R_{i-1})|}, \quad (5)$$

$$F_i^\perp = \nabla E(R_i) - \nabla E(R_i) \cdot \hat{\tau}. \quad (6)$$

To demonstrate the effectiveness of the modified approach used in conjunction with MPSO, the transition path for the transformation from wurtzite-to-rocksalt structure in ZnO using the same MPSO procedure was repeated. The evolution history for the search of the low-energy reaction pathway is shown in Fig. 4c. The energies of the optimized images obtained using the present method are shown in Fig. 4d. The images are optimized along the perpendicular direction to the energy surface; therefore, the distance to the initial state remained almost the same, even without the spring force. The modified NEB scheme converged to the same MEP as the conventional NEB method but with fewer iteration steps. The improvement is quite efficient as the string geometries (images) created from linear interpolation of the initial and final state structures already provide a well-enough approximation for the transition path. Therefore, the computation requirement of one transition path is no more than one structure optimization.

**Carbon.** Because of its ability to form  $sp$ ,  $sp^2$ , and  $sp^3$  hybridized bonds, elemental carbon can occur in a wide variety of allotropes, including hexagonal graphite (HG), rhombohedral graphite (RG), cubic diamond (CD), and hexagonal diamond (HD). These available allotropes have been used variously as conductors or wide-gap insulators. Understanding the atomistic mechanism of the phase transitions of these structures is critical for synthesis design of these carbon allotropes. The atomic mechanisms of these phase transitions, especially for the HG-to-CD transition, have not been thoroughly understood until recently.

We applied the MPSO method to determine the MEP and TS of the carbon phase transformation from HG to CD at 7.2 GPa. The supercell with minimum lattice mismatch contains eight carbon atoms and is generated from superlattice transformation and of the primitive cell for HG and CD, respectively. The configurations and corresponding enthalpies of the images on the MEP are presented in Fig. 5a, b. The transformation of HG to CD was not



**Fig. 4** Exploration of PES with improved NEB method and examination of validity. **a** Schematic diagram of linear interpolation on a PES. The black line, black dots, and  $F_i$  represent the MEP, intermediate images, and optimized direction, respectively. **b** The difference between the improved NEB and original NEB method about the tangent direction leads to the effect on the optimization direction  $F_i$ . The blue and yellow balls represent images of the transition path. **c** Evolution history of searching for the transition path from the wurtzite-to-rocksalt structure of ZnO. The solid black and red lines represent the energy of the initial and optimized candidate transition state, respectively. **d** Optimization of images along the perpendicular direction of the tangents for the wurtzite-to-rocksalt transition. The x-axis is defined by the displacement from the initial state.

direct but occurred through a two-step process involving a stable rhombohedral (RG) intermediate<sup>47</sup>. This finding is in complete agreement with experimental findings. In the first step, the HG with AB stacking transforms to the RG structure with ABC stacking by overcoming a very small energy barrier ( $\sim 2$  meV/atom), where only a layer dislocation occurs without cleavage or the formation of chemical bonds. Concomitantly, the cell angle between lattice vectors  $a$  and  $c$  decreases from  $90^\circ$  to  $79^\circ$ , resulting in shortening of the interlayer distance. It is highly significant that this low-energy intermediate is correctly located by the stochastically chosen starting point of the transition paths by the present MPPO. In the second step, RG transforms to CD. In this process, a slight distortion along the  $c$ -axis occurs, and the carbon atoms in each layer of the RG structure and the distorted directions for the nearest two carbon atoms are completely opposite. The distortions accompanied the formation of new C–C bonds between the layers and led to the final CD structure. As new chemical bonds are formed during this process, the TS has a large energy barrier of  $\sim 0.283$  eV/atom, which agrees well with the value of 0.273 eV/atom estimated from the extrapolation of the high-pressure (to 7.2 GPa) values reported in a previous study<sup>48</sup>.

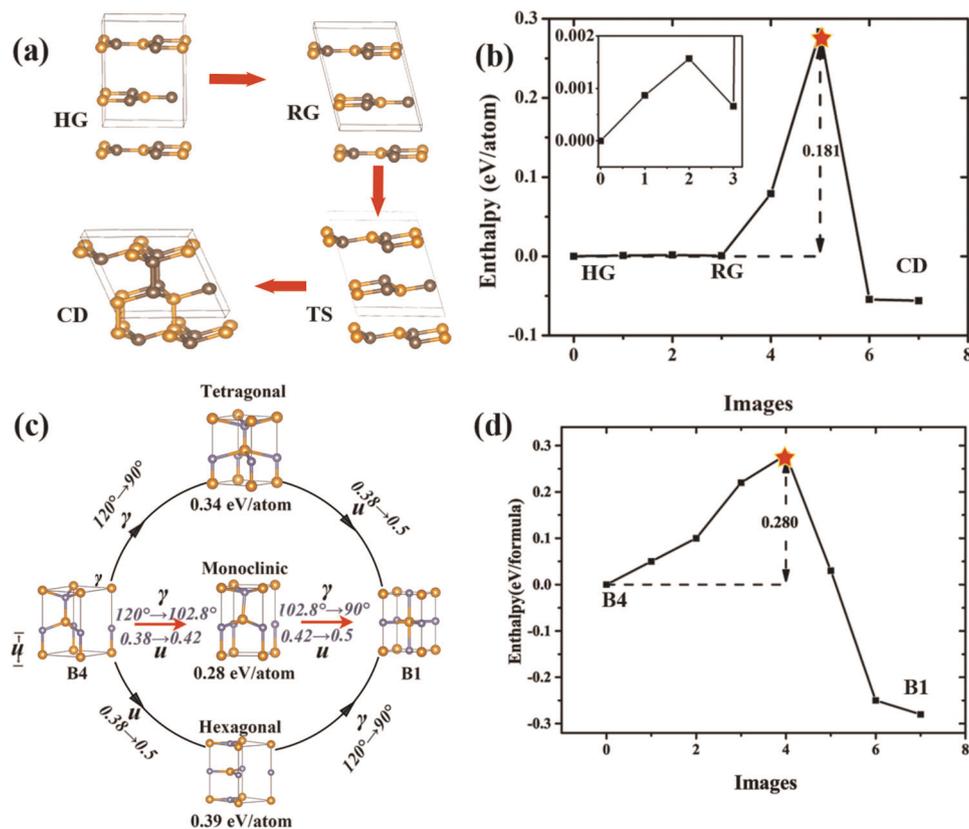
**GaN.** Compounds of post-transition metals have rich physical properties that make them potentially useful in electronic devices. Therefore, the pressure-induced phase transition of gallium nitride (GaN), a wide-band-gap semiconductor, from its ambient hexagonal wurtzite structure (B4, space group:  $P6_3mc$ ) to a cubic rocksalt structure (B1, space group:  $Fm-3m$ ) at  $\sim 50$  GPa has attracted substantial attention over the last few years<sup>49,50</sup>. Two possible mechanisms have been proposed. They are the hexagonal and tetragonal paths, which are named after their

intermediate structures (Fig. 5c). To facilitate the description of the transition path, two internal parameters ( $u$ ,  $\gamma$ ) are often defined to represent the relative positions of the Ga and N sublattices ( $u$ ) and the angle ( $\gamma$ ) between lattice vectors  $b$  and  $c$ .  $u$  and  $\gamma$  are (0.377,  $120^\circ$ ) and (0.5,  $90^\circ$ ) for B4 and B1 structures, respectively. In this case, a supercell of 8 f.u. of GaN was used. As shown in Fig. 5c, the two proposed paths consist of two separate steps: i.e., increase in  $u$  and closure of  $\gamma$ . For the hexagonal path<sup>51</sup>, the hexagonal h-MgO type structure is the TS, and can be transformed from the B4 structure via  $u$  increasing from 0.377 to 0.5 in the first step and the closure of  $\gamma$  from  $120^\circ$  to  $90^\circ$  in the second step. The order of the transformation steps is the reverse of the tetragonal path<sup>52</sup>. Previous calculations have revealed that the energy barriers at 45.7 GPa are 0.34 and 0.39 eV/f.u. for the tetragonal and hexagonal path, respectively<sup>53</sup>.

We revisited the B4–B1 transitions using the present MPPO approach, and found a new lower-energy MEP, which follows the monoclinic path shown in Fig. 5c. Unlike previous reports<sup>52</sup>, the TS has a monoclinic structure ( $Cm$ ) with  $u = 0.42$  and  $\gamma = 102.8^\circ$ <sup>54</sup>. The computed energy barrier of this path is 0.280 eV/f.u. (Fig. 5d), which is lower than those of the previously predicted tetragonal path (0.34 eV/f.u.<sup>9</sup>) and hexagonal path (0.39 eV/f.u.). The discovery of a hidden lower-energy pathway illustrates the robustness of the MPPO method, which does not assume any crystal symmetry for the structural phase transformation. The presumption of even intuitively logical high-symmetry transformation paths may not lead to the correct result.

#### Applications on unsolved systems

**Direct FCC  $\rightarrow$  SC transformation at low temperature.** The successive crystal-crystal transformations of elemental silicon under

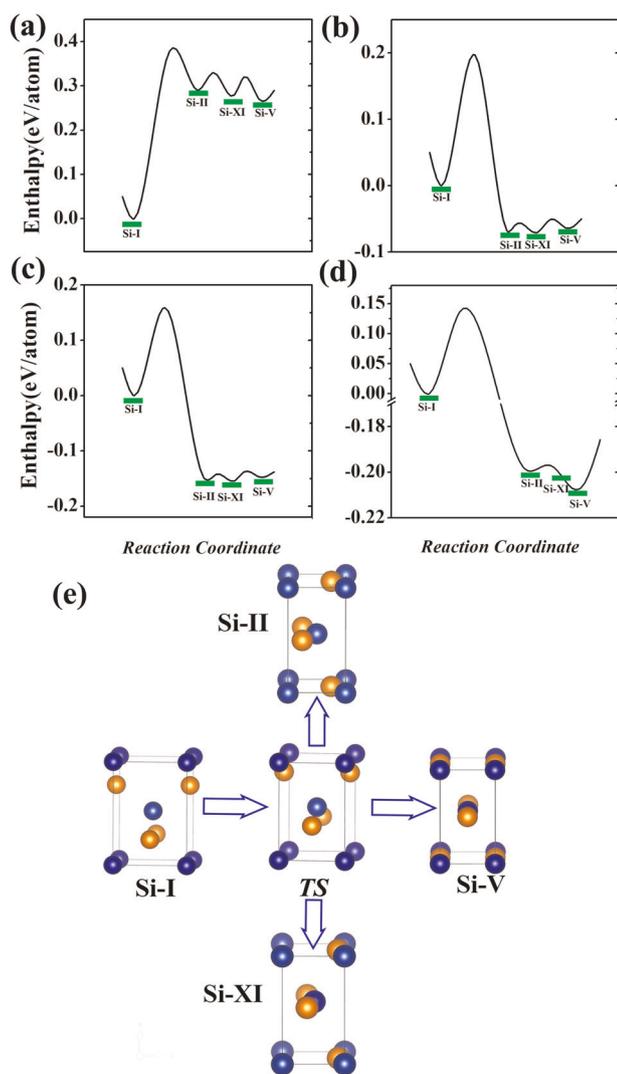


**Fig. 5** Phase transformations of carbon and GaN. **a** Structure change from HG to CD through RG and TS of carbon. The orange and brown balls both represent carbon atoms. **b** MEP from HG to CD of carbon. Part of the MEP is enlarged in the inset. **c** Phase transformations from B4 to B1 in GaN. The orange and light blue balls represent Ga and N atoms, respectively. **d** MEP from B4 to B1 transformation of GaN. Note that the TS is denoted by a red star.

compression are the archetypes of structural diversity at high pressure. Under ambient conditions, Si adopts an FCC structure (Si-I). When compressed at room temperature, Si-I undergoes a well-characterized sequence of phase transitions first to the metallic  $\beta$ -tin structure (Si-II) at 11 GPa<sup>55</sup>, followed by a transition at 13 GPa to the *Imma* structure (Si-XI)<sup>56</sup>, which transforms to a simple hexagonal phase (Si-V) at 16 GPa<sup>57</sup>. Therefore, it is surprising that when Si was compressed at low temperature (80 K) in a hydrostatic medium<sup>58</sup>, a direct transformation from the Si-I to Si-V structure was observed at 17 GPa, bypassing the two thermodynamically stable intermediate phases Si-II and Si-XI. The phase transition is clearly kinetically controlled. The puzzling aspect is understanding why this transition occurs at low temperature. To resolve this issue, it is essential to obtain information on the MEPs and energy barriers of the room-temperature phase transitions. We used MPSO to compute the MEP and TS of the successive structural transformations. First, we calculated the MEP of the transformation from Si-I to Si-II, Si-XI, and Si-V at 0, 12, 15, and 17 GPa using the PBE functional. The results are presented in Fig. 6a–d, respectively. At 0 GPa (Fig. 6a), the total energy of Si-II is higher than that of Si-XI. This finding is consistent with previous LDA calculations showing that the *Imma* (Si-XI) structure is slightly more stable than the  $\beta$ -tin polymorph (Si-II), which is contrary to the observed transformation sequence. As the main concern here is the energy barriers associated with the transitions, the small total energy differences between the polymorphs are not detrimental to the discussion presented below. Exploration of the MEP used a simulation cell with minimum lattice mismatch containing four atoms for the initial (Si-I) and final states (Si-II, Si-XI, and Si-V). The predicted MEPs of the Si-I  $\rightarrow$  Si-II, Si-I  $\rightarrow$  Si-XI, and Si-I  $\rightarrow$  Si-V phase transitions

at 0, 12, 15, and 17 GPa are presented in Fig. 6a–d, respectively. A 8-atom supercell cell commensurate with all the Si polymorphs of interest here was used in the calculations. The purpose of computing the activation barriers at different pressures was to determine the energies required for the structural transformations when the original FCC is compressed to that pressure. As expected, the activation barrier decreased with increasing pressure, because the *PV* work helps to reduce the energy required to overcome the structural transformation barrier. At 0, 12, 15, and 17 GPa, the barrier heights are 0.378, 0.197, 0.199, and 0.159 eV, respectively. At 0 GPa, an activation energy of 0.378 eV/atom is needed to overcome the barrier(s) of transformation to Si-II, Si-XI, and Si-V in the order of stability. At 12 GPa (the expected transformation pressure from Si-I to Si-II), the calculated energy barrier is reduced to 0.197 eV/atom but is not zero. This result indicates that the *PV* work alone is not sufficient to overcome the barriers and that thermal activation is needed to overcome the residual energy barrier. Another piece of information derived from the MEP analysis is that the stability of the final structure resulted in the compression of Si-I at different pressures.

At 12 GPa, as anticipated, Si-II is the most stable form; however, the enthalpy of Si-V is already lower than that of Si-II. Further compression to 15 and 17 GPa led to transformation to Si-XI and Si-V, respectively. At 17 GPa, the barrier is 0.136 eV/atom, which is 0.242 eV/atom lower than that at 0 GPa; however, the most stable final structure is now Si-V. This result indicates that if Si-I (FCC) is compressed at low temperature (80 K), there may be insufficient thermal energy to overcome the transition barrier to Si-II or Si-XI. Therefore, the FCC structure is maintained until 17 GPa, when the activation barrier is substantially reduced and then transforms directly into Si-V (sh). Notably, all the structural transformations



**Fig. 6** Phase transformations of silicon. **a** 0, **b** 12, **c** 15, and **d** 17 GPa. **e** Structural transformations from Si-I to Si-II, Si-XI, and Si-V. **f** Two-dimensional PES of silicon at 17 GPa.

starting from Si-I share a common TS with an intermediate orthogonal structure (Fig. 6e). This observation is consistent with the finding from electronic structure calculations that the instability of the Si-I (FCC) phase results from the increased participation of  $s$ - $d$  hybridization resulting in the tetragonal distortion of the local Si tetrahedral ( $sp^3$ ) bonding<sup>59,60</sup>.

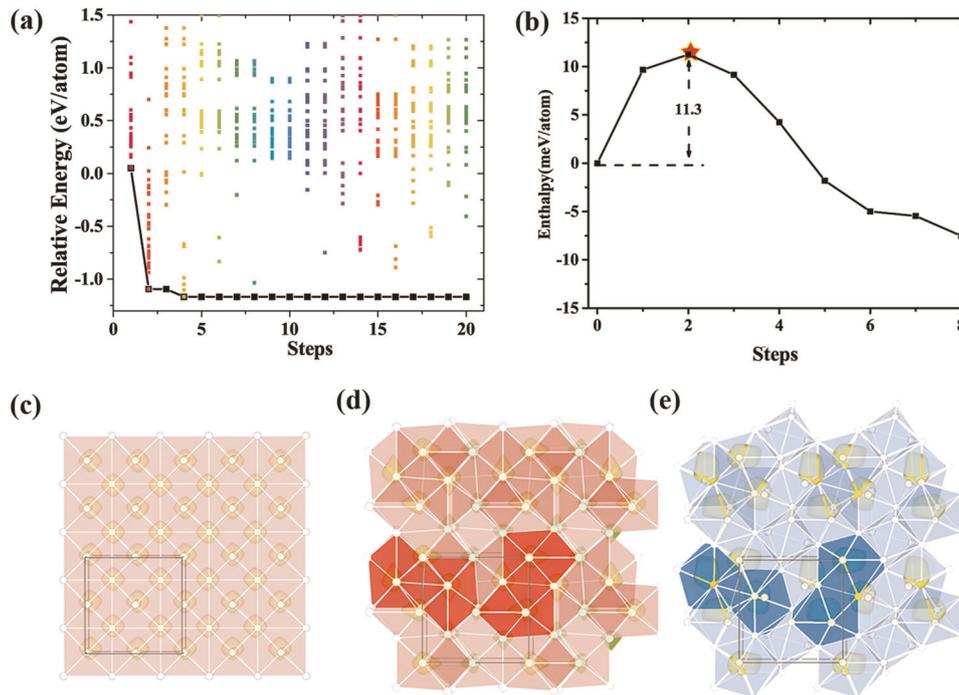
**FCC  $\rightarrow$  BCC transition in high-pressure lithium.** The examples given above, in principle, can be resolved by using relatively small supercells (with up to eight atoms). Therefore, such systems may be adequately handled using existing MEP search techniques with guesses of the appropriate transition paths. In such circumstances the MPSO method is not designed to compete with these established approaches. However, the strength of the present method lies in its use for the study of much larger systems, for which it is difficult or even impossible to make judicial guesses of possible initial transition paths. One example is the high-pressure structure of elemental lithium (Li). Li is considered a simple metal with a simple FCC structure<sup>61</sup>. However, it undergoes a structural transition to the unexpectedly complex  $c/16$  structure at 39 GPa. In fact, most group I and II elements exhibit similar structural transformations from a simple structure to novel and

complex structures. The mechanisms and underlying electronic factors for such transformations are still unknown. As the simplest group I element, the determination of the atomistic mechanism of the FCC ( $Fd3m$ )  $\rightarrow$   $c/16$  ( $I$ ) structural transformation in Li may provide a basis for a deeper understanding of the complex transformation behaviors of other elements at high pressure. Here, we employ the MPSO approach to investigate the MEP and TS of the phase transformation at 44 GPa (Fig. 7a–e). The minimum mismatch supercell commensurate with the initial and final structures consists of 16 Li atoms and was constructed by transforming the FCC structure in BCC settings (i.e.,  $1 \times 1 \times 2$  replication of  $\mathbf{a}_x' = \frac{1}{2}(\mathbf{a}_x + \mathbf{a}_y)$  and  $\mathbf{a}_y' = \frac{1}{2}(\mathbf{a}_x - \mathbf{a}_y)$ , see inset of Fig. 8). The  $c/16$  structure is very complex, and the selection of a suitable path for the transformation a priori is not obvious. It is noteworthy that, in principle, the number of possible transition paths is  $\sim 3 \times 10^{16}$ .

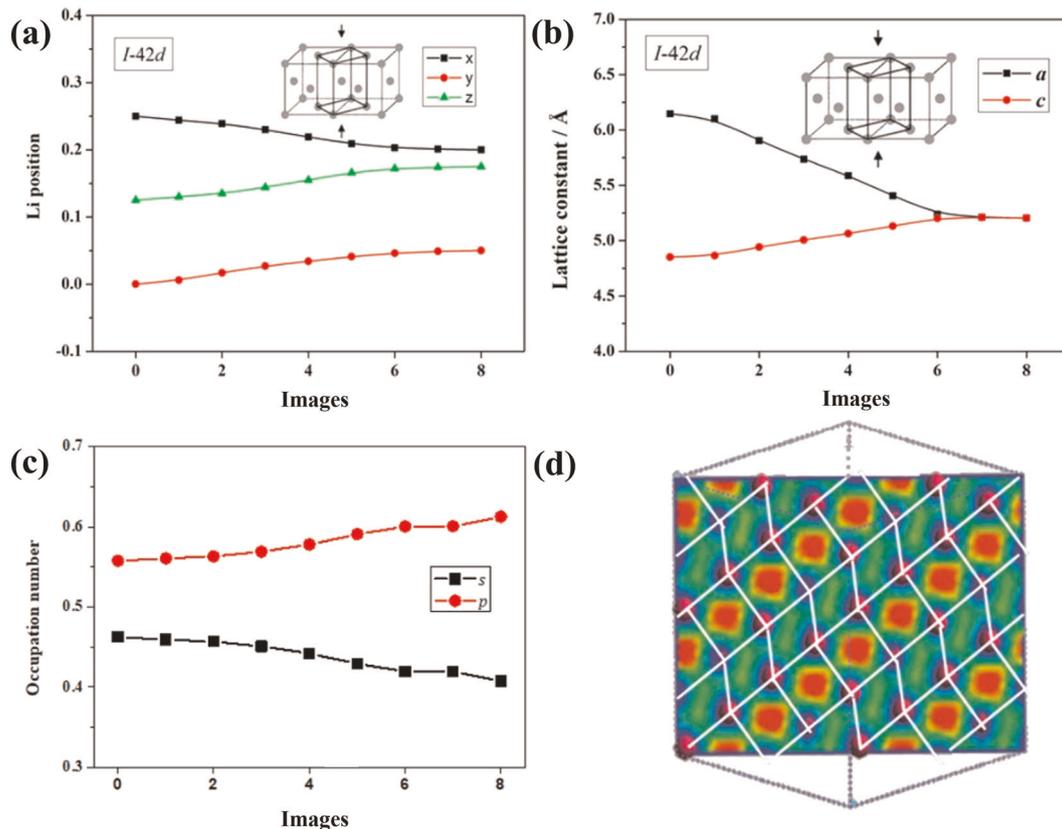
The evolution history of the MPSO calculations is shown in Fig. 7a. The MPSO algorithm converged in only the fourth generation. It yielded several energetically competitive MEP transition paths, including that with the lowest energy barrier (11.3 meV/atom at 44 GPa) shown in Fig. 7b. The energy differences between the intermediate structures (images) and the initial FCC structure are very small ( $<15$  meV/atom), suggesting that the PES for the FCC  $\rightarrow$  BCC Li transition is relatively flat at 44 GPa. FCC, intermediate, and  $c/16$  structures are pretended in Fig. 7c–e. The intermediate structures along the transformation path and at the TS have the  $I-42d$  symmetry. In this space group, there is only one unique Li position (Wyckoff 16e). The evolutions of the Li position and cell axes along the MEP are shown in Fig. 8a, b, respectively. Despite the complexity of the  $c/16$  structure, the transformation mechanism is relatively simple. Under pressure, in the FCC structure, the “ $c$ ”-axis is compressed, and, concomitantly, the Li atoms move closer to each other. As shown in Fig. 7c–e, FCC formed by eightfold-coordinated octahedral unit transforms to  $c/16$  structure formed by fivefold-coordinated dodecahedral unit and electron charge mainly concentrates at the center, however, the energy barrier is only 11.3 meV/atom. The small energy difference arises from changes in the nature of the Li–Li chemical bonds. It is known that for group I and II elements under pressure, the atoms can utilize empty higher azimuthal angular momentum orbitals for chemical interaction; hence, the atoms typically undergo  $s$ - $p$  or  $s$ - $d$  hybridization. In this case, the Li  $p$  orbital participating in the chemical bond becomes more important as the pressure increases. This phenomenon is confirmed by the calculation of the  $s$  and  $p$  occupation number of the nature bond orbital (NBO) of the Li atom<sup>62,63</sup>, where the  $p/s$  ratio increases from 1.16 (0.65/0.56) in the FCC structure to 1.55 (0.62/0.41) in the  $c/16$  structure, indicating significant  $s$ - $p$  hybridization. The overlap between the spatially diffuse Li  $p$  orbitals leads to the migration of the electron density to the interstitial regions, forming “electrides” in the intermediate images (Fig. 8c) and the interstitial sites of dodecahedra of the images. Eventually, a very simple bonding pattern emerges for  $c/16$  (Fig. 8d). The Li atoms forming strips of connected squares with electrons localized in the center result from multi-center bonding through predominantly Li  $p$  orbitals. As there are not enough electrons to completely fill all the  $sp$ -hybridized orbitals, the square strips “slide” against each other, resulting in weaker interaction between them.

## DISCUSSION

A new method of searching for the MEP and TS of solid–solid transformations based on the global MPSO scheme was presented. This method is robust and does not require a priori knowledge on the mechanism or guesses of possible transition paths. The method was validated by successfully reproducing the mechanism of the well-known MEP and TS of the B4  $\rightarrow$  B1 transformation in ZnO and the HG  $\rightarrow$  CD transformation in carbon.



**Fig. 7 Search of transition paths of lithium.** **a** Evolution history of search for the transition path from FCC to c16 structure of Li. The different colored dots indicate the energy barriers of the different paths examined in each step. **b** Transition path with the lowest energy barrier. The red star denotes the TS barrier for the MEP. **c–e** Calculated charge density distribution of **c** FCC, **d** the transition state (TS), and **e** c16 structures. The iso-surface value is set as 0.02. Note that the white balls represent Li atoms. Pink octahedral ( $\text{Li}_6$ ), blue dodecahedral ( $\text{Li}_8$ ).



**Fig. 8 Evolution of structural and electronic properties in lithium.** Evolution of **a** Li position, **b** length of the supercell axis, and **c** occupation of the *s* and *p* NBO orbitals (see text) of the images along the MEP. **d** Plot of the electron localization function along the [111] plane of c16, showing the 2D shear bonding pattern with electrons localized in the squares.

Moreover, in GaN, a hidden lower-energy MEP was discovered. The results demonstrate the reliability and predictive power of the proposed method. The surprising direct silicon FCC  $\rightarrow$  sh transformation at low temperature was explained by characterization and analysis of the MEP and TS of the known successive high-pressure phases at room temperature. A mechanism for the transformation from the low-pressure FCC structure to the high-pressure *c*/16 was revealed. The result is the first step for understanding the mechanisms of the transformations to the novel and very complex polymorph structures in group I and II elements at high pressure.

The performance of the proposed MPSO method for the search of transition pathways associated with structural transformations is measured by the reliability and efficient. In ZnO, GaN, and Si, the lowest energy path was found after sampling 100 paths. For more complex structural transformations, such as in diamond and high pressure Li, ca. 500 trials were needed before the respective correct transition states were determined. The initial guess paths were found to influence the rate of convergence. Thus, if the barrier energies of the initial paths were already low, it often leads to much faster convergence. In this aspect, it is noteworthy that the successful rate is dependent on the PSO ratio<sup>34</sup>, i.e. the ratio of the self-confidence factor and the swarm-confidence factor. In one test case, we have performed 10 calculations on a gold (Au) system containing 16 atoms using an embedded atom potential<sup>41</sup> with different randomly generated starting candidate paths. Depending on the choice of the PSO ratio, the MPSO method was successful in finding the lowest path eight times, i.e., the successful rate is  $\sim$ 80%. A set of inappropriate initial guess paths may lead to erroneous result. Our experience is that a high PSO ratio will lead to faster convergence but sometimes to an undesirable result.

The simplicity of MPSO method is that it eliminates the empirical choice of plausible transition paths. However, it still suffers from slow convergence for large systems (our experience is  $N > 20$ ) due to the excessive number of trial paths. This problem can be circumvented by imposing realistic geometric constraints in the construction of the trial transition paths. This approach does not alter the basic procedure described here. Preliminary calculations have shown that it can reduce the computation effort substantially. There are still rooms for further improvement and fine tuning of the MPSO method, Details of the implementation of this strategy and applications will be reported later.

The method proposed here, perhaps with suitable geometric constraints, is most suitable for this investigation as it is almost impossible to suggest possible pathways for these very complicated transformations from low-pressure simple cubic structures to very complex structures with unit cells consisting of hundreds of atoms at high pressures. For these systems, it may not currently be feasible to employ first-principles density functional theory calculations as were performed here. However, these large systems may be handled by employing accurate atomistic potentials derived from artificial neural networks<sup>64</sup>. It is noteworthy that an ab initio quality potential suitable for the study of sodium at high pressure is already available<sup>65</sup>.

## METHODS

All electronic calculations were performed with the Vienna Ab initio Simulation Package (VASP)<sup>66</sup> code. The effect of the atomic core of an element was replaced by the projector-augmented potential (PAW). A plane-wave cutoff energy 500 eV was employed to expand the wave functions. A *k*-point mesh of  $2\pi \times 0.06 \text{ \AA}^{-1}$  was used to sample the Brillouin zone during the transition path search. Once an approximate transition state was located, a denser *k*-mesh  $2\pi \times 0.02 \text{ \AA}^{-1}$  was used in subsequent climbing NEB (c-NEB) and energy-barrier calculations. In all the cases studied here, a population consisted of 30 trial paths was used. The search is assumed to converge if no new lower-energy pathway was found for 30 successive generations. Supercells with eight formula units was used in

the ZnO, C, GaN, and Si calculations. A 16 atoms model was used for lithium.

## DATA AVAILABILITY

All data generated or analyzed during this study are included in this published article.

## CODE AVAILABILITY

The procedure is now included in the CALYPSO software package and it is free to use by the community.

Received: 28 August 2019; Accepted: 19 February 2020;

Published online: 12 March 2020

## REFERENCES

- Eyring, H. The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107–115 (1935).
- Laidler, K. J. & King, M. C. Development of transition-state theory. *J. Phys. Chem.* **87**, 2657–2664 (1983).
- Kim, J. B. et al. Spectroscopic observation of resonances in the F + H<sub>2</sub> reaction. *Science* **349**, 510–513 (2015).
- E, W., Ren, W. & Vanden-Eijnden, E. String method for the study of rare events. *Phys. Rev. B* **66**, 4 (2002).
- Henkelman, G. & Jónsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **113**, 9978–9985 (2000).
- Trygubenko, S. A. & Wales, D. J. A doubly nudged elastic band method for finding transition states. *J. Chem. Phys.* **120**, 2082 (2004).
- Munro, L. J. & Wales, D. J. Defect migration in crystalline silicon. *Phys. Rev. B* **59**, 3969–3980 (1999).
- Parrish, R. M., Sherrill, C. D., Parrish, R. M. & Sherrill, C. D. Spatial assignment of symmetry adapted perturbation theory interaction energy components: the atomic SAPT partition. *J. Chem. Phys.* **141**, 044115 (2014).
- Henkelman, G. & Jo, H. Introduction. I. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9978–9985 (2000).
- Sheppard, D., Xiao, P., Chemelewski, W., Johnson, D. D. & Henkelman, G. A generalized solid-state nudged elastic band method. *J. Chem. Phys.* **136**, 074103 (2012).
- Henkelman, G. & Jónsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **111**, 7010 (1999).
- Bolhuis, P. G., Chandler, D., Dellago, C. & Geissler, P. L. Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Annu. Rev. Phys. Chem.* **53**, 291–318 (2002).
- Shang, C. & Liu, Z.-P. Stochastic surface walking method for structure prediction and pathway searching. *J. Chem. Theory Comput.* **9**, 1838–1845 (2013).
- Niblett, S. P., de Souza, V. K., Jack, R. L. & Wales, D. J. Effects of random pinning on the potential energy landscape of a supercooled liquid. *J. Chem. Phys.* **149**, 114503 (2018).
- Heuer, A. Exploring the potential energy landscape of glass-forming systems: from inherent structures via metabasins to macroscopic transport. *J. Phys. Condens. Matter* **20**, 373101 (2008).
- Swinburne, T. D. & Perez, D. Self-optimized construction of transition rate matrices from accelerated atomistic simulations with Bayesian uncertainty quantification. *Phys. Rev. Mater.* **2**, 053802 (2018).
- Niblett, S. P., de Souza, V. K., Stevenson, J. D. & Wales, D. J. Dynamics of a molecular glass former: energy landscapes for diffusion in ortho-terphenyl. *J. Chem. Phys.* **145**, 024505 (2016).
- De Souza, V. K. & Wales, D. J. The potential energy landscape for crystallisation of a Lennard-Jones fluid. *J. Stat. Mech. Theory Exp.* **7**, 074001 (2016).
- De Souza, V. K. & Wales, D. J. Connectivity in the potential energy landscape for binary Lennard-Jones systems. *J. Chem. Phys.* **130**, 194508 (2019).
- Middleton, T. F. & Wales, D. J. Energy landscapes of model glasses. II. Results for constant pressure. *J. Chem. Phys.* **118**, 4583–4593 (2003).
- Middleton, T. F. & Wales, D. J. Energy landscapes of some model glass formers. *Phys. Rev. B* **64**, 024205 (2001).
- Swinburne, T. D. & Marinica, M. C. Unsupervised calculation of free energy barriers in large crystalline systems. *Phys. Rev. Lett.* **120**, 135503 (2018).
- Saksaengwijit, A. & Heuer, A. Finite-size effects in silica: a landscape perspective. *J. Phys. Condens. Matter* **19**, 205143 (2007).

24. Sheppard, D., Terrell, R. & Henkelman, G. Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **128**, 134106 (2008).
25. Quapp, W. & Bofill, J. M. A comment to the nudged elastic band method. *J. Comput. Chem.* **31**, 2526–2531 (2010).
26. Henkelman, G., Jóhannesson, G., Jónsson, H. in *Theoretical Methods in Condensed Phase Chemistry* (ed. Schwartz, S.D.) 269–302 (Kluwer Academic Publishers, Dordrecht, 2002).
27. Wang, Y. et al. Materials discovery via CALYPSO methodology. *J. Phys. Condens. Matter* **27**, 203203 (2015).
28. Li, Y., Hao, J., Liu, H., Li, Y. & Ma, Y. The metallization and superconductivity of dense hydrogen sulfide. *J. Chem. Phys.* **140**, 174712 (2014).
29. Wang, Y. et al. An effective structure prediction method for layered materials based on 2D particle swarm optimization algorithm. *J. Chem. Phys.* **137**, 224108 (2012).
30. Lu, S., Wang, Y., Liu, H., Miao, M.-S. & Ma, Y. Self-assembled ultrathin nanotubes on diamond (100) surface. *Nat. Commun.* **5**, 3666 (2015).
31. Zhu, L., Liu, H., Pickard, C. J., Zou, G. & Ma, Y. Reactions of xenon with iron and nickel are predicted in the earth's inner core. *Nat. Chem.* **6**, 644–648 (2014).
32. Zhang, X. et al. First-principles structural design of superhard materials. *J. Chem. Phys.* **138**, 114101 (2013).
33. Lv, J., Wang, Y., Zhu, L. & Ma, Y. Particle-swarm structure prediction on clusters. *J. Chem. Phys.* **137**, 084104 (2012).
34. Wang, Y., Lv, J., Zhu, L. & Ma, Y. CALYPSO: a method for crystal structure prediction. *Comput. Phys. Commun.* **183**, 2063–2070 (2012).
35. Zhu, L. et al. Substitutional alloy of Bi and Te at high pressure. *Phys. Rev. Lett.* **106**, 18–21 (2011).
36. Lv, J., Wang, Y., Zhu, L. & Ma, Y. Predicted novel high-pressure phases of lithium. *Phys. Rev. Lett.* **106**, 19–22 (2011).
37. Wang, H., Tse, J. S., Tanaka, K., Iitaka, T. & Ma, Y. Superconductive sodalite-like clathrate calcium hydride at high pressures. *Proc. Natl Acad. Sci. USA* **109**, 6463–6466 (2012).
38. Wang, Y., Lv, J., Zhu, L. & Ma, Y. Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B* **82**, 1–8 (2010).
39. Zur, A. & McGill, T. C. Lattice match: an application to heteroepitaxy. *J. Appl. Phys.* **55**, 378–386 (1984).
40. Goldberg, E. F. G., Goldberg, M. C. & de Souza, G. R. Particle swarm optimization algorithm for the traveling salesman problem in *Traveling Salesman Problem* (ed. Greco, F.) 75–96 (IntechOpen, 2008).
41. Ercolessi, F., Parrinello, M. & Tosatti, E. Simulation of gold in the glue model. *Philos. Mag. A Phys. Condens. Matter, Struct. Defects Mech. Prop.* **58**, 213–226 (1988).
42. Gale, J. D. & Rohl, A. L. The General Utility Lattice Program (GULP). *Mol. Simul.* **29**, 291–341 (2003).
43. Saitta, A. & Decremps, F. Unifying description of the wurtzite-to-rocksalt phase transition in wide-gap semiconductors: the effect of d electrons on the elastic constants. *Phys. Rev. B* **70**, 1–5 (2004).
44. Zarkevich, N. A. & Johnson, D. D. Titanium  $\alpha$ - $\omega$  phase transformation pathway and a predicted metastable structure. *Struct. Phys. Rev. B* **93**, 020104 (2016).
45. Yeo, S. C., Han, S. S. & Lee, H. M. Adsorption, dissociation, penetration, and diffusion of N<sub>2</sub> on and in Bcc Fe: first-principles calculations. *Phys. Chem. Chem. Phys.* **15**, 5186–5192 (2013).
46. Wang, J.-T., Chen, C. & Kawazoe, Y. Phase conversion from graphite toward a simple monoclinic Sp<sup>3</sup>-carbon allotrope. *J. Chem. Phys.* **137**, 024502 (2012).
47. Yang, G. W. & Wang, J. B. Pulsed-laser-induced transformation path of graphite to diamond via an intermediate rhombohedral graphite. *Appl. Phys. A Mater. Sci. Process.* **72**, 475–479 (2001).
48. Khaliullin, R. Z. et al. Nucleation mechanism for the direct graphite-to-diamond phase transition. *Nat. Mater.* **10**, 693–697 (2011).
49. Limpjumnong, S. & Lambrecht, W. Homogeneous strain deformation path for the wurtzite to rocksalt high-pressure phase transition in GaN. *Phys. Rev. Lett.* **86**, 91–94 (2001).
50. Cai, J. & Chen, N. Microscopic mechanism of the wurtzite-to-rocksalt phase transition of the group-III nitrides from first principles. *Phys. Rev. B - Condens. Phys.* **75**, 1–12 (2007).
51. Limpjumnong, S. & Lambrecht, W. Theoretical study of the relative stability of wurtzite and rocksalt phases in MgO and GaN. *Phys. Rev. B* **63**, 1–11 (2001).
52. Bouffelfel, S. E., Zahn, D., Grin, Y. & Leoni, S. Walking the path from B<sub>4</sub>- to B<sub>1</sub>-type structures in GaN. *Phys. Rev. Lett.* **99**, 2–5 (2007).
53. Qian, G. R. et al. Variable cell nudged elastic band method for studying solid-solid structural phase transitions. *Comput. Phys. Commun.* **184**, 2111–2118 (2013).
54. Sowa, H. On the transition from the wurtzite to the NaCl type. *Acta Crystallogr. Sect. A* **57**, 176–182 (2001).
55. Gaál-Nagy, K., Pavone, P. & Strauch, D. Ab initio study of the beta-Ti n  $\rightarrow$  Imma  $\rightarrow$  sh phase transitions in silicon and germanium. *Phys. Rev. B* **69**, 134112 (2004).
56. McMahon, M. & Nelmes, R. New high-pressure phase of Si. *Phys. Rev. B* **47**, 8337–8340 (1993).
57. Olijnyk, H., Sikka, S. & Holzapfel, W. Structural phase transitions in Si and Ge under pressures up to 50 GPa. *Phys. Lett. A* **51**, 1–4 (1984).
58. Tse, J. S. et al. Pressure-induced changes on the electronic structure and electron topology in the direct FCC  $\rightarrow$  SH transformation of silicon. *J. Phys. Chem. C* **18**, 1161–1166 (2014).
59. Tse, J. S., Boldyreva, E. Electron density topology of crystalline solids at high pressure. in *Modern Charge-Density Analysis* (eds. Gatti, C. & Macchi, P.) 573–623 (Springer, 2012).
60. Li, R., Liu, J., Bai, L., Tse, J. S. & Shen, G. Pressure-induced changes in the electron density distribution in  $\alpha$ -Ge near the  $\alpha$ - $\beta$  transition. *Appl. Phys. Lett.* **107**, 072109 (2015).
61. Ackland, G. J. et al. Quantum and isotope effects in lithium metal. *Science* **356**, 1254–1259 (2017).
62. Dunnington, B. D. & Schmidt, J. R. Generalization of natural bond orbital analysis to periodic systems: applications to solids and surfaces via plane-wave density functional theory. *J. Chem. Theory Comput.* **8**, 1902–1911 (2012).
63. Weinhold, F. Natural bond orbital analysis: a critical overview of relationships to alternative bonding perspectives. *J. Comput. Chem.* **33**, 2363–2379 (2012).
64. Behler, J. First principles neural network potentials for reactive simulations of large molecular and condensed systems. *Angew. Chem. - Int. Ed.* **56**, 12828–12840 (2017).
65. Khaliullin, R. Z., Eshet, H., Kühne, T. D., Behler, J. & Parrinello, M. Graphite-diamond phase coexistence study employing a neural-network mapping of the ab initio potential energy surface. *Phys. Rev. B Phys.* **81**, 18–21 (2010).
66. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).

## ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program of China under Grant Nos. 2016YFB0201201 and 2017YFB0701503, the National Natural Science Foundation of China (Grant Nos. 11534003, 11822404, 11904148, and 11774127), and Science Challenge Project No. TZ2016001. Parts of the calculations were performed at the High-Performance Computing Center of Jilin and computers at Compute Canada.

## AUTHOR CONTRIBUTIONS

J.S.T., Y.W., and Y.M. conceived the project. K.Y. proposed the method. K.Y. and P.G. carried out the first principles calculations. H.L. and J.L. helped to revise the manuscript. X.S. and B.G. advised the strategy of the research. All authors discussed the results and wrote the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to J.S.T., Y.W. or Y.M.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020