Structure search method for atomic clusters based on the dividing rectangles algorithm

Kansei Kanayama [ORCID],[*] Atsuto Seko [ORCID], and Kazuaki Toyoura [ORCID][†]

Department of Materials Science and Engineering, Kyoto University, Kyoto 606-8501, Japan

The Dividing Rectangles (DIRECT) algorithm is a deterministic optimization method to explore optimal solutions by repeatedly dividing a given hyperrectangle search space into subhyperrectangles. Herein, we propose a structure search method for atomic clusters based on the DIRECT algorithm in combination with a gradient-based local optimizer to enable an efficient structure search in high-dimensional search spaces. We use the Z-matrix representation for defining the hyperrectangle search space, in which the bond lengths, bond angles, and dihedral angles specify a cluster structure. To evaluate its performance, we applied the proposed method to the Lennard-Jones clusters and two kinds of real atomic clusters with many metastable structures, i.e., phosphorus and sulfur clusters, and compared the results with those of conventional methods. The proposed method exhibits a higher efficiency than random search and a comparable efficiency to basin hopping.

I. INTRODUCTION

Material structures in stable and low-energy metastable states govern physical and chemical properties. Consequently, several conventional techniques have been developed for material structure determination such as x-ray or neutron crystallography, nuclear magnetic resonance spectroscopy, and electron microscopy. However, experimental information alone is typically insufficient to determine unknown material structures. A theoretical structure search using empirical potentials or first-principles calculations is a useful tool. It not only provides candidate structures prior to experiments but also can predict materials with unknown structures and properties. Particularly, first-principles calculations without any empirical parameters are powerful tools due to the development of high-performance computing systems and highly efficient codes.

Theoretical structure searches for atomic clusters and crystalline systems explore global and local minima on the potential energy surface (PES), which is defined in the configuration space. The configuration space has $3N - 6$ dimensions in an N-atom system, corresponding to the degree of freedom (DOF) of the atomic coordinates $3N$ subtracted by the DOF of translations and rotations of the whole system -6. Therefore, structure searches for systems consisting of many atoms are difficult due to the high-dimensional PES with numerous local minima.

Simple methods for enumerating low-energy local minima include grid search (GS) [1] and random search (RS) [2–4]. These uniformly explore a given search space without any prior and posterior information about the potential energies and their gradients. In contrast, sophisticated methods (e.g., basin hopping (BH) [5–7], BH parallel tempering (BHPT)

[8], simulated annealing (SA) [9–11], minima hopping (MH) [12,13], genetic algorithm (GA) [14–16], and particle swarm optimization (PSO) [17,18]) exploit prior and posterior information to enhance the search efficiency. These methods are all heuristic and widely used for structure search.

Among the sophisticated methods, BH, BHPT, MH, and SA are based on the Monte Carlo (MC) or the molecular dynamics (MD) method. They typically use a single walker to explore a given PES. GA and PSO are nature-inspired optimization methods, which imitate the evolution and behavior of biological systems, respectively. They perform simulated biological operations for a given population of structures at every step. As these sophisticated methods use multiple parameters, their search performances are sensitive to the given parameters. For example, BH, which combines MC with a local optimizer, uses at least two parameters (i.e., the maximum displacement of the walker on the PES and the temperature in the Metropolis criterion). In the case of a small displacement or low temperature, the walker cannot escape from a single basin or funnel, whereas the use of a high temperature prevents the walker from effectively exploiting information about the potential energies. Thus, the parameters must be carefully set for the target system in these methods.

To overcome this limitation, a parameter-free or parameter-insensitive structure search method is desirable. In this paper, we investigate a modification of the DIRECT algorithm [19], which is an abbreviation for Dividing Rectangles. This algorithm takes a deterministic approach that does not require multiple runs. DIRECT uses a parameter to determine the solution accuracy, but the algorithm works efficiently without fine parameter tuning. The original DIRECT algorithm does not exploit information about the gradients of the objective function. However, the potential energy gradients are available without additional computational costs in common first-principles calculation packages [20–24], implying that DIRECT can be improved. Herein, we investigate a modification that combines a gradient-based local optimizer to enhance

---
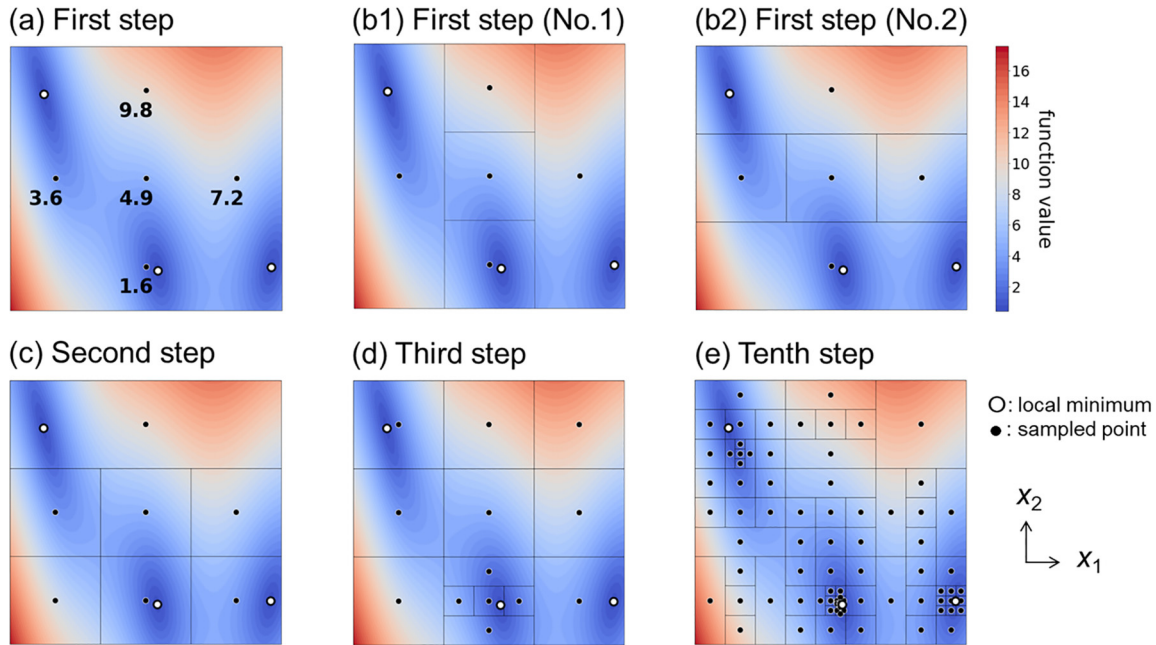[*]kanayama.kansei.67z@st.kyoto-u.ac.jp
[†]toyoura.kazuaki.5r@kyoto-u.ac.jp

FIG. 1. Example of exploring optimal solutions by DIRECT for a synthetic two-dimensional objective function in a square region ($0 \leqslant x_1 \leqslant 1, 0 \leqslant x_2 \leqslant 1$). Black and white dots represent the sampled points and the local minima, respectively. (a) Computed function values at four points, ($\frac{1}{2} \pm \frac{1}{3}, \frac{1}{2}$) and ($\frac{1}{2}, \frac{1}{2} \pm \frac{1}{3}$), in the first step. [(b1), (b2)] Of the two possible choices to divide the square into five subrectangles at the first step, we choose the second one (b2) because the subrectangle with the lowest function value becomes the largest one. Sampling points and divisions at the (c) second, (e) third, and (f) 10th steps.

the performance of the DIRECT algorithm. The combination manner is inspired by the strategy of BH, in which a local optimizer makes the search efficiency higher than the simple MC [6]. Although several modifications of DIRECT using local optimizers have been proposed in the past few decades, our proposed method differs in terms of the combination manner. To demonstrate its potential, the performance of the proposed method is compared with those of two conventional methods (RS and BH) when applied to the Lennard-Jones clusters and two kinds of real atomic clusters, i.e., phosphorus and sulfur clusters.

II. PROPOSED METHOD

The structure search for atomic clusters has three steps. (i) For a chemical formula with a given total number of atoms, all possible structural formulas corresponding to adjacency relations of atoms are enumerated. Since a cyclic structure can be represented as an acyclic structure by neglecting part of adjacency relations, it is sufficient to enumerate only acyclic formulas. See Ref. [25] for the conventional enumeration method of structural formulas. (ii) For each structural formula, the search space (the hyperrectangle domain) is defined using the Z-matrix representation. A single structural formula has multiple Z-matrix representations. Consequently, a representative must be selected from the Z matrices. In addition, structural symmetry indicates that the search space can be restricted to the irreducible one. The general definition of the Z matrix and the specifications of search space in this paper are described in Secs. II C and II D, respectively. (iii) The proposed method is applied to a structure search, in which the DIRECT algorithm is combined with a local optimizer. The

original DIRECT algorithm and our modification are detailed in Secs. II A and II B, respectively.

A. Original DIRECT algorithm

DIRECT [19] is a modification of the Lipschitzian approach [26–28]. This approach explores optimal solutions of an objective function $f(\mathbf{x})$ with the Lipschitz continuity, which is mathematically expressed in d dimensions as

$$|f(\mathbf{x}) - f(\mathbf{x}')| \leqslant K\|\mathbf{x} - \mathbf{x}'\| \quad \text{for } {}^\forall \mathbf{x} \in \mathbb{R}^d, {}^\forall \mathbf{x}' \in \mathbb{R}^d (\mathbf{x} \neq \mathbf{x}'),$$
(1)

where the absolute value of the slope of every two points is less than a certain positive real number K, called the *Lipschitz constant*. The Lipschitzian approach explores optimal solutions by repeating the selective division of a given domain using the Lipschitz continuity. One limitation is that K is generally unknown or does not exist in practical problems. In contrast, the Lipschitz constant does not have to be specified in DIRECT, allowing global optimization for black-box objective functions.

Figure 1 illustrates how DIRECT works for a synthetic objective function in a normalized square domain. In the initial state, only the function value at the center of the square ($\frac{1}{2}$, $\frac{1}{2}$) is computed. The first step computes the function values at four points [i.e., ($\frac{1}{2} \pm \frac{1}{3}, \frac{1}{2}$) and ($\frac{1}{2}, \frac{1}{2} \pm \frac{1}{3}$)], which are $\pm\frac{1}{3}$ displaced points from the center of the domain along either the horizontal or vertical side [Fig. 1(a)]. Next, the square domain is divided into five subrectangles since each one has a single point already computed at the center. Specifically, the square is divided into thirds along the two sides. In principle,
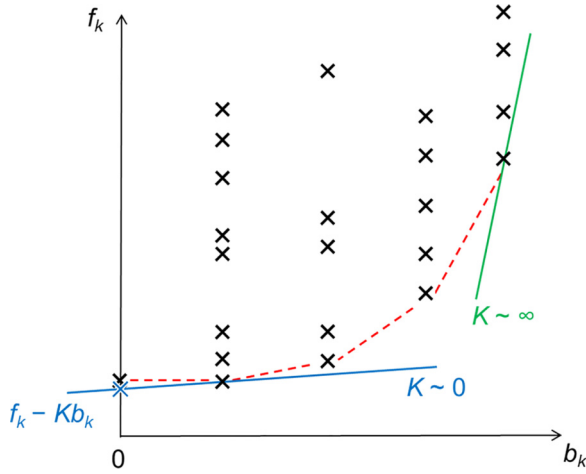
FIG. 2. Example of the b_k–f_k diagram at a step. If K is small and close to zero (blue line), the most likely optimal hyperrectangle is that with a low f_k. If K is large and close to infinity (green line), the most likely optimal hyperrectangle is that with a large b_k. Potentially optimal hyperrectangles correspond to all points on the lower-right convex hull (red broken line).

the square division has two options [Figs. 1(b1) and 1(b2)]. For a given synthetic function, we choose the second one [Fig. 1(b2)], where the square is tridivided in the order of vertical and horizontal dimensions so that the subrectangle with the lowest function value becomes the largest. DIRECT alternates between computing the objective function values and dividing the domain [Figs. 1(a)–1(e)]. Except for the first step, the potentially optimal rectangles should be selected for further division. Qualitatively, multiple rectangles with a large size or a low function value at the center are selected as the potentially optimal ones at each step, which is detailed later. Thus, DIRECT explores the optimal solutions for the global minima in function $f(\mathbf{x})$ by iteratively performing the three processes: (i) selection of potentially optimal rectangles (*selection process*), (ii) computation of function values in the selected rectangles (*computation process*), and (iii) division of the selected rectangles (*division process*).

Here, the three processes in DIRECT are described for the general case of a d-dimensional hyperrectangle domain. In the selection process, a potentially optimal hyperrectangle, which is a hyperrectangle with the lowest lower bound of the function value at a Lipschitz constant K, is selected. Considering the Lipschitz continuity, the lower bound in the kth hyperrectangle $\mathbf{D}_k$ is estimated by the function value at the center $f_k[=f(\mathbf{c}_k)]$ and the center-vertex distance b_k as follows:

$$f(\mathbf{x}) \geqslant f(\mathbf{c}_k) - K\|\mathbf{x} - \mathbf{c}_k\| \geqslant f_k - Kb_k \quad \text{for } ^\forall\mathbf{x} \in \mathbf{D}_k. \tag{2}$$

However, a single hyperrectangle with the smallest lower bounds cannot be uniquely identified because the estimated lower bound depends on the unknown constant K. Hence, a common strategy is employed in which all hyperrectangles to be the potentially optimal one for a positive Lipschitz constant are selected. The b_k-f_k diagram can easily identify such hyperrectangles (Fig. 2). Each point in the diagram corresponds to a hyperrectangle at a step, and the f_k intercept of the straight

line through each point with slope K is the estimated lower bound of the corresponding hyperrectangle for a Lipschitz constant K. If K is small and close to zero, a hyperrectangle with a low f_k is likely to be optimal, while a hyperrectangle with a large b_k is likely to be optimal if K is large and close to infinity. The potentially optimal hyperrectangles correspond to all points on the lower-right convex hull in the diagram (red broken line). This strategy makes it possible to balance local and global searches. The original DIRECT requires a lower limit on the Lipschitz constant K to determine the solution accuracy. In contrast, the lower limit on the hyperrectangle size is specified and hyperrectangles with a subthreshold size are not divided in this paper.

In the computation process, the function values at the two points $\mathbf{x} = \mathbf{c}_{\pm i}$ are computed along the longest side for each potentially optimal hyperrectangle. The points $\mathbf{c}_{\pm i}$ are given by

$$\mathbf{c}_{\pm i} = \mathbf{c}_0 \pm \left(\tfrac{1}{3}\right) a_i \mathbf{e}_i \quad i \in \{i | a_i = \max(a_1, a_2, \ldots, a_d)\}, \tag{3}$$

where a_i and $\mathbf{e}_i$ are the length and basis vector along the longest sides in the hyperrectangle, respectively, and $\mathbf{c}_0$ is the center of the hyperrectangle.

In the division process, each potentially optimal hyperrectangle is divided into $2d' + 1$ subhyperrectangles based on the newly computed $2d'$ function values, where d' is the number of the longest sides. Specifically, each original hyperrectangle is divided into thirds along the longest sides in the ascending order of $\min[f(\mathbf{c}_{-i}), f(\mathbf{c}_{+i})]$. Then the subhyperrectangle with the lowest function value tends to be the largest.

B. Modification of the DIRECT algorithm

To explore the optimal solution (the local minimum point) in a basin, conventional local optimizers based on gradient methods are more efficient than DIRECT. A simple modification is to alternate between DIRECT for several steps and a local optimization from the current solution using a local optimizer [29,30]. Several DIRECT-based methods with a local optimizer have been proposed [29–34]. These should be more efficient because they avoid the overrefining problem that occurs in the original DIRECT. For example, DIRMIN performs local optimizations from the centers in all potentially optimal hyperrectangles at every step. DIRMIN outperformed the original DIRECT for several test functions in various dimensions [31].

The proposed method, which is like DIRMIN, is a modification to overcome the curse of dimensionality. Unlike DIRMIN, which only considers potentially optimal hyperrectangles, this modification performs local optimizations at the centers in all hyperrectangles at every step. Second, the local minimum values after local optimizations are referenced for subdivision in DIRECT unlike DIRMIN, which uses the unoptimized values. This method is equivalent to the original DIRECT on the transformed step function $F(\mathbf{x})$ (Fig. 3). In the transformed step function, the function values are constant at the local minimum value within each basin in the original function $f(\mathbf{x})$. This combination manner is inspired by the strategy of BH. BH also explores the local minima on the transformed step function $F(\mathbf{x})$ rather than the original func-
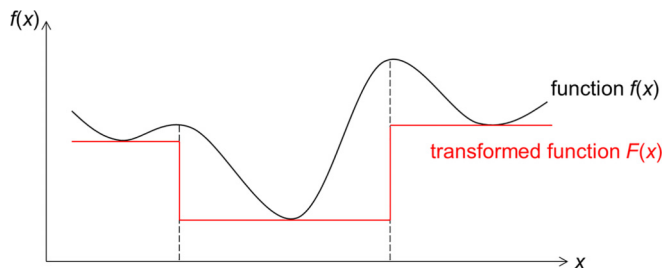
FIG. 3. Conceptual diagram of the function transformation from the original function $f(\mathbf{x})$ to the step function $F(\mathbf{x})$. In the transformed step function $F(\mathbf{x})$, the function values are constant at the local minimum value within each basin in the original function $f(\mathbf{x})$.

tion $f(\mathbf{x})$, making the search efficiency higher than the simple MC. We call the proposed method tDIRECT.

C. General definition of the Z-matrix representation

The hyperrectangle search space must be defined to apply the DIRECT algorithm to structure searches. Although an atomic cluster structure can be described in various manners, in this paper, we employ the Z-matrix representation using internal coordinates (i.e., bond lengths, bond angles, and dihedral angles) [35,36]. The internal coordinates specify the relative positions of the constituent atoms. Consequently, the search space satisfying the adjacency relations of atoms corresponding to a given structural formula can easily be defined by the internal coordinates, whereas using absolute coordinates such as Cartesian coordinates to specify adjacency relations is difficult. The distance matrix also explicitly indicates the relative positions of atoms. A distance matrix is often used for structure searches of molecules (e.g., distance geometry method) [37,38]. Because the number of interatomic distances in an N-atom system is $N(N-1)/2$, the dimension of a search space using the distance matrix is beyond that of configuration space, $3N - 6$, at $N > 4$. This disadvantage becomes more serious as the number of atoms in the system increases. In contrast, the number of internal coordinates in a Z matrix is equal to the dimension of the configuration space, $3N - 6$, satisfying the translational and rotational invariances without any constraints. Herein, the Z-matrix representation is employed for atomic clusters.

As a simple example, we illustrate a Z matrix for a methane molecule with a regular tetrahedral coordination of H atoms around the central C atom. The Z matrix is expressed as

Row 1:	C						
Row 2:	H	1	r_2				
Row 3:	H	1	r_3	2	φ_3		
Row 4:	H	1	r_4	2	φ_4	3	θ_4
Row 5:	H	1	r_5	2	φ_5	3	θ_5

where $r_n = 1.09\,\text{Å}$ $(n = 2, \ldots, 5)$, $\varphi_n = 109.5°$ $(n = 3, \ldots, 5)$, and $\theta_n = 120.0°$ $(n = 4, \ldots, 5)$, according to the literature [39]. The species and position of the nth atom in the nth row of the Z matrix are defined by describing the internal coordinates. The first row specifies only the atomic species C in the first column and not other internal

coordinates. In the second row, the bond length between the first and second atoms r_2 is necessary to determine the relative position. In the third row, the bond angle between bonds 1–3 and 1–2, φ_3, is required as well as the 1–3 bond length r_3. In the fourth row, the dihedral angle between the planes 4–1–2 and 1–2–3, θ_4, the bond length r_4, and the bond angle φ_4 are essential, and similarly after the fifth row. In general, an N-atom cluster has $3N - 6$ internal coordinates in total [i.e., bond lengths r_n $(r_n > 0, n = 2, \ldots, N)$, bond angles φ_n $(0° \leqslant \varphi_n \leqslant 180°, n = 3, \ldots, N)$, and dihedral angles θ_n $(0° \leqslant \theta_n < 360°, n = 4, \ldots, N)$], corresponding to the hyperrectangle domain for the atomic cluster structure. Thus, the Z matrix specifies each atomic position in a cluster as the relative positions to already specified atoms.

D. Specification of the search space for atomic cluster structures

The Z matrix has multiple equivalent representations for a given structural formula. The representation depends on which atom is assigned to each row of the Z matrix and which adjacent atoms define the bond length, the bond angle, and the dihedral angle. As an example, consider a methane molecule. A different Z-matrix representation can be expressed as

Row 1:	H						
Row 2:	C	1	r_2				
Row 3:	H	2	r_3	1	φ_3		
Row 4:	H	2	r_4	1	φ_4	3	θ_4
Row 5:	H	2	r_5	4	φ_5	3	θ_5

In this representation, the first and second rows specify the hydrogen and carbon atoms, respectively. In addition, the bond angle in the fifth row is defined as 5–2–4 instead of 5–2–1. Below, we explain the scheme to select a representative Z matrix in this paper.

First, the atomic indices corresponding to the row number of the Z matrix are determined. For atomic indexing, the structural formula is considered as a graph, which is represented by two sets of nodes and edges. The graph is acyclic, and every two adjacent nodes in the structural formula are connected by a single edge, even if the chemical-bonding state contains a multiple bond. That is, the graph is an undirected tree. The priority of node i, π_i, $(i = 1, \ldots, N)$ is determined using criteria (a) and (b). These are the two criteria in canonical numbering for molecular graphs proposed by Jochum and Gasteiger [40].

Criterion (a): Eccentricity of node i, ε_i. In graph theory, the eccentricity of a node is defined as the maximum distance to the other nodes. The distance is defined as the number of edges in the shortest path connecting two nodes. If $\varepsilon_i < \varepsilon_j$, then $\pi_i > \pi_j$, meaning that the node closer to the center of an atomic cluster has a higher priority.

Criterion (b): Degree of node i, k_i. In graph theory, the degree of a node is defined as the number of connected edges. If $k_i > k_j$, then $\pi_i > \pi_j$, meaning that the node with more adjacent atoms has a higher priority.

The priority is initially determined by criterion (a). Criterion (b) is used only when multiple nodes have equal priorities in criterion (a). The atoms assigned to the individual nodes are arranged in descending order of π_i in the Z matrix. Table I

TABLE I. Atomic indexing for the structural formula shown in Fig. 4. The structural formula consists of nine X atoms. Eccentricity, degree, priority, and index of each node are also shown.

	X^a	X^b	X^c	X^d	X^e	X^f	X^g	X^h	X^i
(a) Eccentricity ε_i	6	5	4	3	4	5	6	5	6
(b) Degree k_i	1	2	2	2	3	2	1	2	1
Priority π_i	1	2	3	5	4	2	1	2	1
Index	9	6	3	1	2	4	7	5	8

demonstrates determining the atomic indices based on the criteria for an acyclic structural formula composed of nine atoms, as shown in Fig. 4. When atoms have the same priority, the atom adjacent to a smaller-indexed atom is indexed. If equivalent atoms are adjacent to the same atom, they are indexed arbitrarily. For example, among X^b, X^f, and X^h with the same priority in Table I, X^b is assigned a larger index than X^f and X^h due to the adjacency to X^c with a larger index, while X^f and X^h are indexed arbitrarily.

Second, the induced subtree defining the bond angle φ_n ($n = 3, \ldots, N$) and dihedral angle θ_n ($n = 4, \ldots, N$) is selected for every atom. Note those defining the bond lengths are uniquely determined in the case of acyclic structural formulas. Induced subtrees for bond and dihedral angles are selected as the smallest sum of atomic indices in the possible trees. Using the acyclic structural formula in Fig. 4 as an example, there are two possible induced subtrees defining the dihedral angle specifying the position of X^i (atomic index: 8) [i.e., X^i-X^h-X^e-X^d (atomic indices: 8-5-2-1) or X^i-X^h-X^e-X^f (atomic indices: 8-5-2-4)]. In this case, we select the first tree according to the above rule.

Once a procedure is constructed to select a representative of the Z matrix in the above scheme, the hyperrectangle search space can be specified, i.e., $r_{n,LB} \leqslant r_n \leqslant r_{n,UB}$ ($n = 2, \ldots, N$), $0° \leqslant \varphi_n \leqslant 180°$ ($n = 3, \ldots, N$), and $0° \leqslant \theta_n < 360°$ ($n = 4, \ldots, N$). Here, $r_{n,LB}$ and $r_{n,UB}$ are the lower and upper bounds for the bond length r_n, respectively. These can be restricted to a short range by referring to the bond radii reported in the literature. Because in this paper we set $r_{n,LB}$ and $r_{n,UB}$ to the same length, every bond length is fixed to a constant value. For example, the constant value is set to the average bond length of 100 cluster structures obtained by the preliminary RS. As a result, the search space dimension is reduced by $N − 1$ from $3N − 6$ to $2N − 5$. The reduced search space by the constant bond lengths is reasonable because the proposed method incorporates a local optimizer to optimize the bond lengths.

In the $(2N − 5)$-dimensional hyperrectangle search space for bond and dihedral angles, any point corresponds to an atomic cluster structure. Because each point in this domain may not necessarily correspond to different structures, there
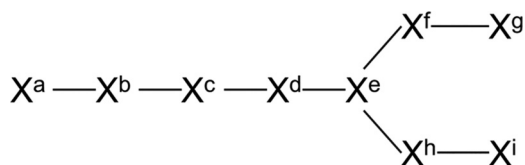

Z-matrix

(a)
Row 1: X
Row 2: X r_2 1
Row 3: X r_3 1 φ_3 2
Row 4: X r_4 2 φ_4 1 θ_4 3
Row 5: X r_5 2 φ_5 1 θ_5 3
Row 6: X r_6 3 φ_6 1 θ_6 2
Row 7: X r_7 4 φ_7 2 θ_7 1
Row 8: X r_8 5 φ_8 2 θ_8 1
Row 9: X r_9 6 φ_9 3 θ_9 1

(b)
Row 1: X
Row 2: X r_2 1
Row 3: X r_3 1 φ_3 2
Row 4: X r_5 2 φ_5 1 θ_5 3
Row 5: X r_4 2 φ_4 1 θ_4 3
Row 6: X r_6 3 φ_6 1 θ_6 2
Row 7: X r_8 4 φ_8 2 θ_8 1
Row 8: X r_7 5 φ_7 2 θ_7 1
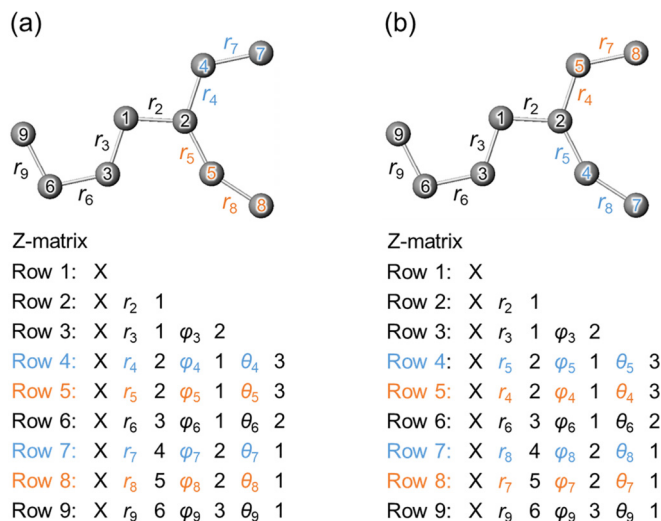Row 9: X r_9 6 φ_9 3 θ_9 1

FIG. 5. Two Z-matrix representations for identical cluster structures, in which the internal coordinates of the two bijective subtrees, 4–7 and 5–8, are exchanged. Number for each atom is the atomic index corresponding to the row number in the Z matrix.

may be multiple equivalent points. Using the structural formula in Fig. 4 as an example, any structure has two different representations in the Z matrix, as shown in Fig. 5. The two representations coincide with each other by the permutation of the individual internal coordinates of two bijective subtrees 4–7 and 5–8. To exclude duplicate points, the search space is restricted to the irreducible parts of the domain (irreducible search space). Specifically, constraints are imposed on the bond or dihedral angles in every pair of representative atoms in such bijective subtrees [e.g., $\theta_7 < \theta_8$ in Fig. 5(a)]. In tDIRECT, local optimizations for the points outside the irreducible search space are not performed. Additionally, the hyperrectangles that do not overlap with the irreducible search space are not further divided.

In addition, mirror symmetry reduces the search space. The search space is defined by the dihedral angle in the fourth atom as a range of $0° \leqslant \theta \leqslant 180°$, which restricts the position of the fourth atom to one side of the plane formed by the first, second, and third atoms. The two structures in Fig. 6 are equivalent when all the other coordinates are equal except for the dihedral angle $(360° − \theta)$ on the fourth atom. Note that the difference between the right- and left-handed systems in this paper are ignored.
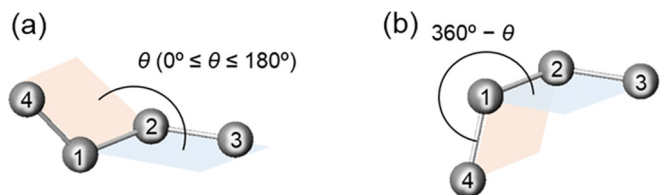

FIG. 4. Acyclic structural formula with nine X atoms. These atoms are tentatively labeled alphabetically.

FIG. 6. Structures (a) and (b) are equivalent when the dihedral angles specifying the fourth atom are θ and $360° − \theta$, respectively, and all other coordinates are equal. Note that the difference between the right- and left-handed systems is ignored in this paper.
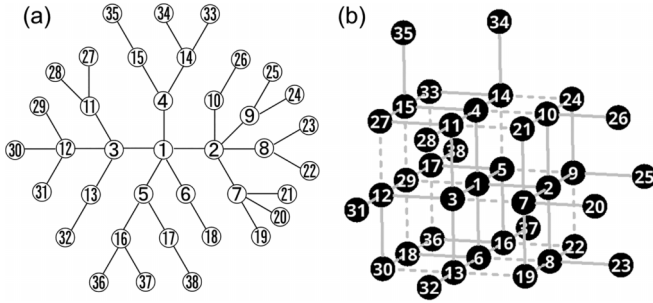
FIG. 7. (a) Structural formula labeled with the atomic indices of LJ$_{38}$ cluster for constructing the Z matrix. Structural formula of each LJ$_N$ cluster ($N \leqslant 36$) corresponds to the induced subtree consisting of all nodes with a label in the range of 1 and N. (b) The LJ$_{38}$ structure corresponding to the center of the search space, defined as $r_n = 2^{-1/6}$ ($n = 2, \ldots, 38$), $0° \leqslant \varphi_n \leqslant 180°$ ($n = 3, \ldots, 38$), $0° \leqslant \theta_4 < 180°$, and $\theta_{n,0} \leqslant \theta_n < \theta_{n,0} + 360°$ ($\theta_{n,0} = 0°, -90°, -180°$, or $-270°$ for $n = 5, \ldots, 38$) in Z-matrix representation, in which all atoms are located on the simple-cubic lattice.

III. RESULTS AND DISCUSSION

A. Application to Lennard-Jones clusters

First, we applied tDIRECT to the global optimization problems for Lennard-Jones clusters LJ$_N$ ($n = 8, 10, \ldots, 38$). Both the potential well depth and the equilibrium distance were set to unity in the LJ potential. In these applications, cluster structures were explored only in a single search space constructed from the structural formula shown in Fig. 7(a). The structural formula of the LJ$_N$ cluster corresponds to the induced subtree consisting of all nodes with a label in the range of 1 and N. Although the structural formula appears to be a complicated two-dimensional graph, it is constructed based on a simple-cubic structure, as shown in Fig. 7(b). The center of the search space, corresponding to the initial structure in tDIRECT, becomes the simple-cubic cluster by appropriately setting the ranges of dihedral angles with periodicity. Specifically, four types of ranges of dihedral angles $\theta_{n,0} \leqslant \theta_n < \theta_{n,0} + 360°$ ($\theta_{n,0} = 0°, -90°, -180°$, or $-270°$) were appropriately chosen for $n = 5, \ldots, N$. The bond lengths r_n ($n = 2, \ldots, N$) were fixed at theoretically optimal values l_{fix} as $2^{-1/6}$, and the bond angles φ_n ($n = 3, \ldots, N$) and the dihedral angle θ_4 were defined by $0° \leqslant \varphi_n, \theta_4 \leqslant 180°$. The maximum coordination number of the structures corresponding to the center of the search space is six, which is different from those of the stable LJ$_N$ clusters forming an icosahedral- or face-centered-cubic-based structure [5]. The employed search spaces, therefore, never work in favor of tDIRECT in the structure search for LJ clusters.

Local optimizations were performed only if the shortest atomic distance was $>0.5\,l_{\mathrm{fix}}$ (called the *initial condition*) to avoid local optimizations for unrealistic initial structures. The conjugate gradient algorithm [41] was used for local optimizations implemented in LAMMPS [42]. The convergence criterion was that all the force components on any atom are $<1 \times 10^{-4}$ in force units. A local optimization not converging within the maximum number of total energy evaluations ($=100\,N$) was terminated as an error.

The threshold for the hyperrectangle size was set to 30° in the search space defined by the bond and dihedral angles. Thus, hyperrectangles with a subthreshold size were not divided into subhyperrectangles. We set a high energy, 1000 in energy units, at the center value in a hyperrectangle when the initial structure did not satisfy the initial condition.

For comparison, we also applied RS and BH to the LJ clusters. In RS, local optimizations were performed for the randomly generated structures satisfying the initial condition. In BH, we performed 10 runs, in which each run started at a local minimum obtained by a local optimization from a randomly generated initial structure satisfying the initial condition. BH repeated the following three processes for the initial structure: (i) randomly displace each atom in an atomic cluster, (ii) perform a local optimization, and (iii) accept or reject the transition to the new structure with the atomic displacements. In process (i), each atom was randomly displaced in Cartesian coordinates. The displacement width was determined as $l_{\mathrm{fix}}\alpha\eta$ ($0 \leqslant \eta \leqslant 1$: uniform random number) in a random direction, where the scaling factor α was 0.7 or 1.0. In process (iii), we accepted or rejected the structural transition according to the Metropolis criterion, in which the temperature parameter was set as 0.05 per atom in energy units.

Figure 8 shows (a) the number of geometry optimizations and (b) the number of total energy evaluations until finding the stable structure for each LJ$_N$ cluster ($N = 8, 10, \ldots, 38$). The RS profiles are shown as the expected values estimated after finding the stable structure multiple times in a single long run. The BH profiles are shown by the boxplots of 100 runs, in which labels BH_1 and BH_07 correspond to the different scaling factors α for atomic displacements 1.0 and 0.7, respectively. For $N \leqslant 16$, the differences in the number of geometry optimizations or total energy evaluations are small between the three methods of tDIRECT, RS, and BH, probably because the PESs of the relatively small LJ clusters have not so many local minima [43]. For $N \geqslant 18$, the three methods have a clear difference in performance. tDIRECT is more efficient than RS and slightly less efficient than BH on average. However, the performance of BH depends on the scaling factor α and largely fluctuates even under the same scaling factor due to the probabilistic nature. By contrast, tDIRECT exhibits consistent performance due to the deterministic nature, which is the superiority of tDIRECT over BH. Note that tDIRECT shows excellent performance for the LJ$_{38}$ cluster. It is well known that LJ$_{38}$ has a PES with an archetypal double funnel [44], which makes it difficult to find the global minimum by conventional global optimization methods. Table II shows the comparison of the performances for the LJ$_{38}$ cluster between tDIRECT with the benchmarks from Ref. [45]. tDIRECT is more efficient than BH and GA, whose computational cost for finding the stable structure is half and quarter in comparison with the BH and GA, respectively. The performances of BH and GA were reported to be improved with the symmetrization schemes, which reduce the computational costs by more than an order of magnitude. Although tDIRECT is less efficient than the symmetrized BH and GA, the performance of tDIRECT should be notable in terms of deterministic approach with a single insensitive parameter, which may be due to the well-balanced strategy for global and local searches.
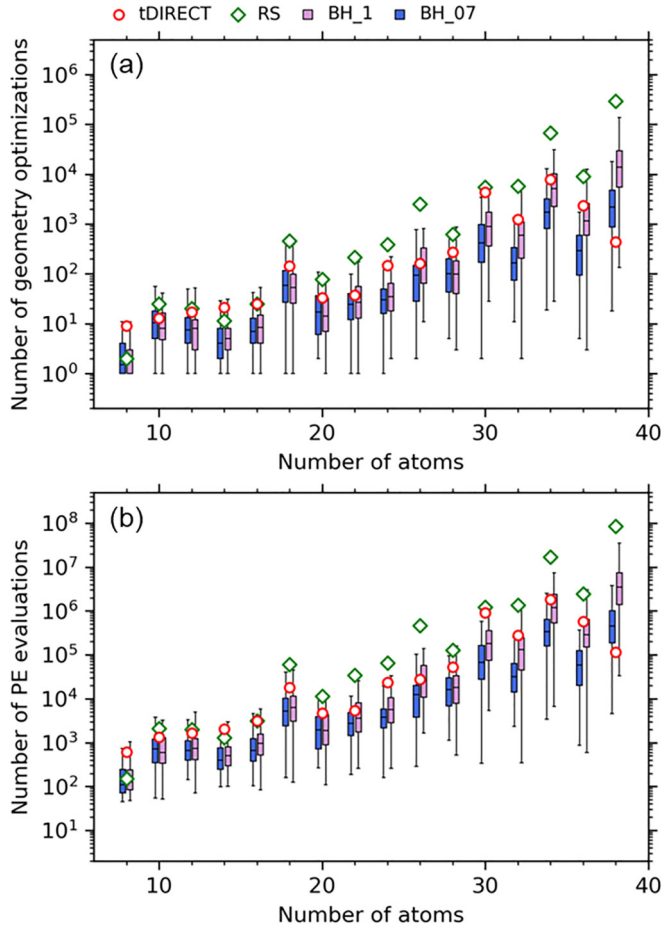
FIG. 8. (a) The number of geometry optimizations and (b) the number of total energy evaluations in tDIRECT (red circle) until finding the stable structure for each LJ$_N$ cluster ($N = 8, 10, \ldots, 38$). The performances of random search (RS) and basin hopping (BH) are shown for comparison. The RS profiles (green diamond) are shown as the expected values estimated after finding the stable structure multiple times in a single long run. The BH profiles are shown as boxplots of a hundred runs. The pink (right) and blue (left) boxplots at each number of atoms correspond to the results of BH with scaling factors $\alpha = 1.0$ and 0.7 (BH_1 and BH_07), respectively.

B. Application to phosphorus and sulfur clusters

1. Model systems and computational conditions

We applied tDIRECT to practical structure search problems for phosphorus clusters P_N and sulfur clusters S_N ($N = 8, 10,$ and 12). For comparison, we also applied RS and BH to these clusters. For the application of BH, the scaling factor α was set to 0.4, 0.5, 0.7, or 1.0, and the temperature parameter was set as 0.1 eV/atom. In these applications, cluster structures were searched only in the single search space constructed from a linear structural formula. Figure 9 shows the atomic indices for the linear structural formula X_N. Because any linear structural formula can be described with two different Z-matrix representations due to the two bijective subtrees (Fig. 9, subtrees 1 and 2), we constrained the dihedral angles specifying the atoms at the ends of the individual subtrees, i.e., $\theta_{N-1} \leqslant \theta_N$. As a result, the irreducible search space is defined by $0° \leqslant \varphi_n \leqslant 180°$ ($n = 3, \ldots, N$), $0° \leqslant \theta_4 < 180°$,

TABLE II. Comparison of the performances for the LJ$_{38}$ cluster between tDIRECT and the benchmarks from [45]. The benchmarks are BH, GA, BH with the core orbits symmetrization schemes (BH-CO), GA with CO schemes (GA-CO), and BH with the continuous symmetry measure schemes (BH-CSM), corresponding to the statistics of 100 random starting points.

Method	Geometry optimization	Total energy evaluations
tDIRECT	441	114738
BH [45]	1271	185493
GA [45]	2885	404825
BH-CO [45]	142	20655
GA-CO [45]	105	25365
BH-CSM [45]	34	4369

$0° \leqslant \theta_n < 360°$ ($n = 5, \ldots, N$), and $\theta_{N-1} \leqslant \theta_N$. The bond lengths r_n ($n = 2, \ldots, N$) were fixed at constant values l_{fix} of 2.35, 2.37, and 2.39 Å in the phosphorus clusters P$_8$, P$_{10}$, and P$_{12}$, respectively, while l_{fix} was set to 2.36 Å for all sulfur clusters. These values were obtained by averaging the interatomic distances under 1.5 l_{min} of 100 optimized structures in the preliminary RS runs, where l_{min} is the shortest interatomic distance of each cluster structure.

The total energies and atomic forces of structures for each atomic cluster were calculated based on density functional theory (DFT) implemented in Gaussian 16 [20]. We employed the PBE0 exchange-correlation term [46] and the LanL2DZ basis functions [47–49], assuming a singlet electronic state. The rational function optimization algorithm [50–53] was used for local optimizations. Local optimization has two convergence criteria: (i) the root mean square of the first derivative of the total energy is $<3 \times 10^{-4}$ E_{h}/a_0 (1.543×10^{-2} eV/Å), and its maximum is $<4.5 \times 10^{-4}$ E_{h}/a_0 (2.314×10^{-2} eV/Å); and (ii) the root-mean-square deviation (RMSD) of the atomic displacement is $<1.2 \times 10^{-3}$ a_0 (6.350×10^{-4} Å), and its maximum is $<1.8 \times 10^{-3}$ a_0 (9.526×10^{-4} Å). Note E_{h} (27.2116 eV) is Hartree energy and a_0 (0.5292 Å) is the Bohr radius.

In the DFT calculations, local optimizations that satisfied either of the following error criteria were aborted: (i) The
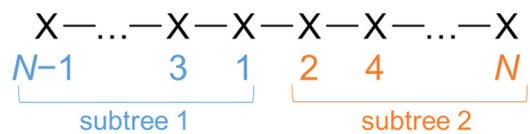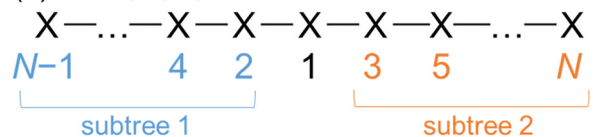

FIG. 9. Atomic indices for the linear structural formula of X_N in the cases of (a) even N and (b) odd N. Any structural formula in both cases can be described with two different Z-matrix representations due to the two bijective subtrees 1 and 2. We imposed a constraint on the dihedral angles, $\theta_{N-1} \leqslant \theta_N$.
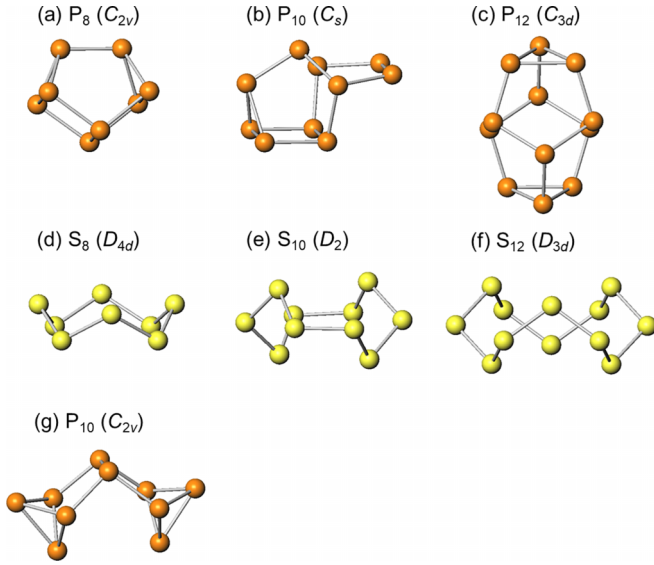
FIG. 10. Stable structures of (a) P_8, (b) P_{10}, (c) P_{12}, (d) S_8, (e) S_{10}, and (f) S_{12} obtained by the proposed methods using density functional theory (DFT) calculations with the PBE0 exchange-correlation term and the LanL2DZ basis functions. (g) Reported stable structure of P_{10} obtained at the SCF/SVP (split valence plus polarization) level in Ref. [57]. Parentheses indicate the point group of each cluster structure.

output structure is separated into multiple clusters, or (ii) the number of total energy evaluations in a local optimization is $>50\,N$ times. The criterion of cluster separation was that the interdistances of subclusters were $>1.5\,l_{\min}$. If a local optimization yielded an error in the self-consistent field (SCF) convergence or a structure with imaginary vibration modes, the local optimization was repeated after randomly displacing every atom in the final structure, as the RMSD of the atomic displacement was $<0.1\,\text{Å}$. If the same SCF error occurred more than once, the local optimization was terminated as an error.

In structure search for these practical problems, low-energy local minima are important as well as the global minimum from the thermodynamic point of view. The low-energy metastable structures may exist at a rate proportional to the Boltzmann constant at finite temperatures, and one of them may become more stable at high temperatures than the global

TABLE III. Methods of the total energy calculations for the structure search of the phosphorus clusters P_N and sulfur clusters S_N ($N = 8, 10, 12$) in the literature.

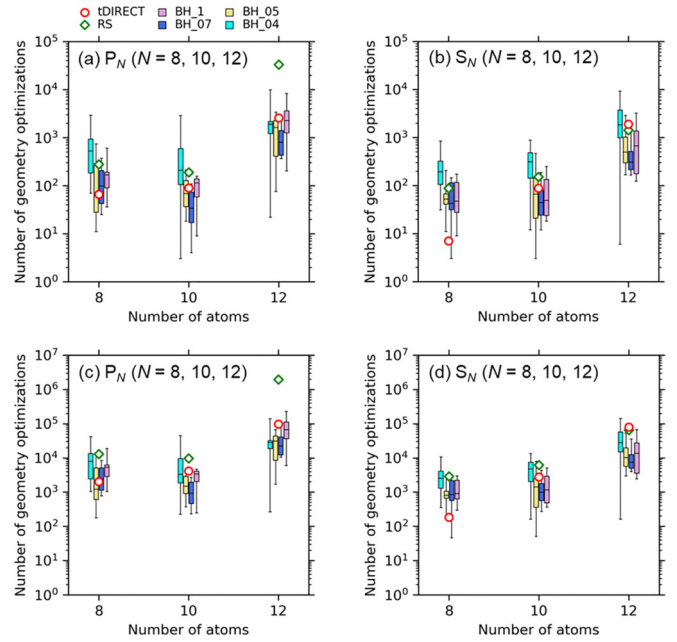| Literature	Calculation method	Clusters
Ref. [55]	MD-DF (LSDA)	P_8
Ref. [56]	MD-DF (LSDA)	P_{10}
Ref. [57]	SCF/SVP	P_8, P_{10}, P_{12}
Ref. [58]	B3LYP/6–311G*	P_8, P_{10}, P_{12}
Ref. [59]	MD-DF (LSDA)	S_8, S_{10}, S_{12}
Ref. [60]	MP4/6-31G*//HF/3-21G*	S_8
	MP3/6-31G*//HF/3-21G*	S_{10}, S_{12}
Ref. [61]	B3LYP/6-31G*	S_8, S_{10}, S_{12}

FIG. 11. [(a), (b)] The number of geometry optimizations and [(c), (d)] the number of total energy evaluations in tDIRECT (red circle) until finding the stable structure of the P_N clusters and the S_N clusters ($N = 8, 10, 12$). The performances of random search (RS) and basin hopping (BH) are shown for comparison. The RS profiles (green diamond) are shown as the expected values estimated after finding the stable structure multiple times in a single long run. The BH profiles are shown by the boxplots of 10 runs. The four boxplots from right to left at each number of atoms are BH_1 (pink), BH_07 (blue), BH_05 (yellow), and BH_04 (light blue), corresponding to the results of BH with scaling factors $\alpha = 1.0, 0.7, 0.5$, and 0.4, respectively.

minimum due to the larger contribution of rotational and/or vibrational free energies. The Gehrke's criterion [54] was employed after each structure search to determine the structural coincidence between structures A and B. The two structures were considered identical when the following condition was satisfied:

$$\frac{\sum (l_{A,i} - l_{B,i})^2}{\sum \left(l_{A,i}^2 + l_{B,i}^2\right)} < 10^{-4}, \tag{4}$$

where $l_{A,i}$ and $l_{B,i}$ are the interatomic distances of structures A and B sorted in ascending order, respectively. Note that the optimized structure must maintain a cluster form. That is, the structure must correspond to a connected graph. We assumed that two atoms were adjacent if the sorted interatomic distance $l_{X,i}$ satisfied the following condition:

$$l_{X,j} \leqslant l_{X,j-1} + 0.5\,\text{Å} \quad \text{for all } j = 1, 2, \ldots, i, \tag{5}$$

where $l_{X,0}$ is the shortest interatomic distance $l_{\min}$ of structure X.

2. Obtained stable and metastable structures of the phosphorus and sulfur clusters

Figure 10 shows the stable structures of the phosphorus and sulfur clusters obtained in this paper, while Table III summarizes the calculation methods in the previous reports
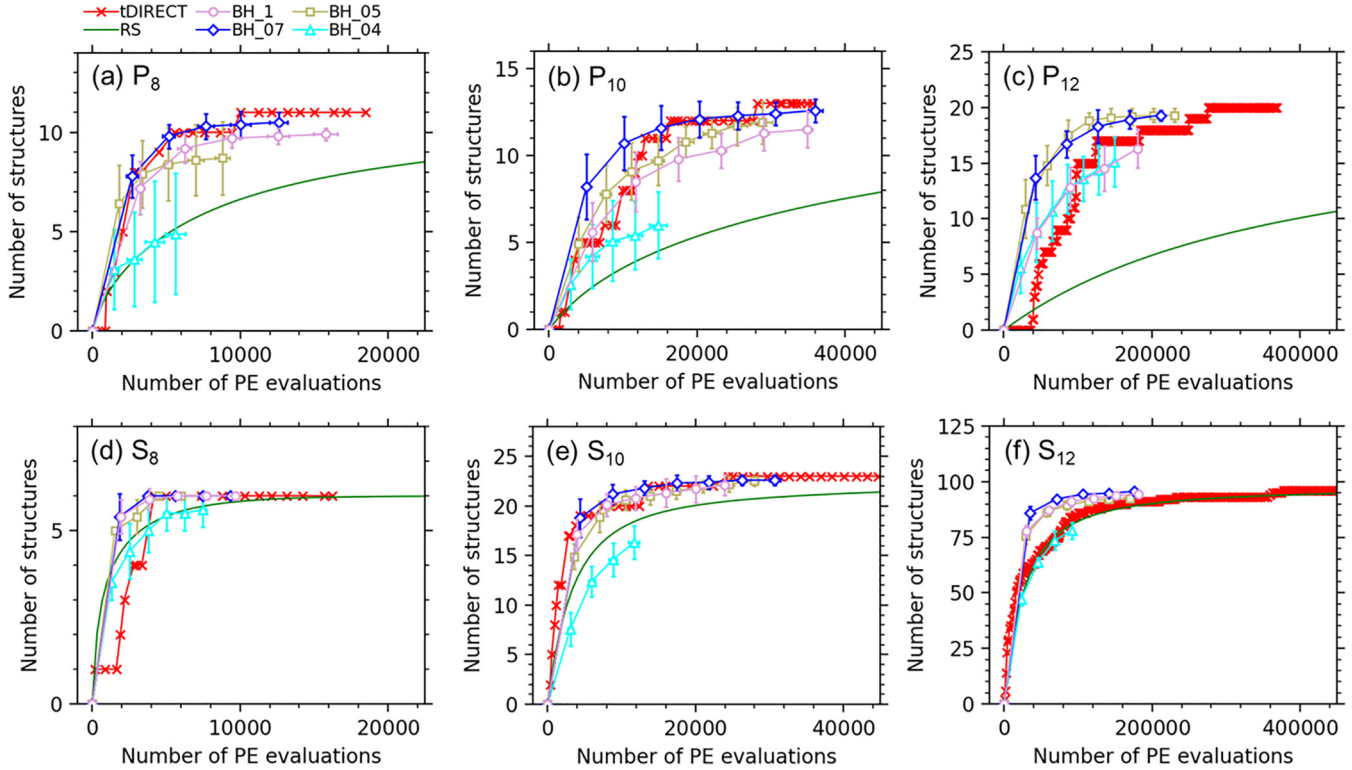
FIG. 12. Performances of tDIRECT (red cross) in the structure search for (a) P_8, (b) P_{10}, (c) P_{12}, (d) S_8, (e) S_{10}, and (f) S_{12}, in which the number of obtained stable and low-energy metastable structures are shown as a function of the number of total energy evaluations. Performances of random search (RS) and basin hopping (BH) are shown for comparison. The RS profiles (green lines) are shown as the expected values estimated from the results of N_{smp} local optimizations. BH profiles are the average values of 10 runs with error bars ($\pm\sigma$). BH_1 (pink circle), BH_07 (blue diamond), BH_05 (yellow square), and BH_04 (light blue triangle) correspond to the BH profiles with scaling factors $\alpha = 1.0, 0.7, 0.5$, and 0.4, respectively.

on structure searches of these clusters [55–61]. The obtained stable structures were consistent with those found in the literature, except for that of P_{10} in Ref. [57]. The inconsistent structure in the literature [Fig. 10(g)] was obtained at the SCF/SVP (split valence plus polarization) level, which was a metastable structure at the PBE0/LanL2DZ level in this paper with a relative energy of 0.064 eV/atom vs the stable structures. The stable structures of S_8, S_{10}, and S_{12} were consistent also with the reported structures characterized by x-ray structural analyses [62].

The literature contained 14 low-energy metastable structures for the phosphorus and sulfur clusters (<0.1 eV/atom vs the stable structure). Of these, 12 were obtained in this paper. One of the two unobtained structures had a relative energy >0.1 eV/atom at the PBE0/LanL2DZ level, while the other was not identified due to the unclear description in the literature. On the contrary, the proposed method obtained many unreported low-energy structures (see Fig. 12).

3. Performance of tDIRECT

Figure 11 shows the numbers of geometry optimizations and total energy evaluations until finding the stable structure of the phosphorus clusters P_N and sulfur clusters S_N ($N = 8$, 10, 12). The RS profiles are shown as the expected values estimated after finding the stable structure multiple times in a single long run. The BH profiles are shown by the box-

plots of 10 runs, in which labels BH_1, BH_07, BH_05, and BH_04 correspond to the different scaling factors α for atomic displacements, 1.0, 0.7, 0.5, and 0.4, respectively. Overall, tDIRECT is more efficient than RS and slightly less efficient than BH. Since the numbers of local minima of these clusters (see Table IV) are larger than that of the LJ cluster with the same size [43], these profiles exhibit a similar tendency to those of relatively large LJ clusters ($N \geqslant 18$).

Figure 12 shows the performances of tDIRECT for enumerating the low-energy structures of the phosphorus clusters P_N and sulfur clusters S_N. The number of obtained stable and

TABLE IV. Initial setting and results of RS for the phosphorus and sulfur clusters. N_{smp}, N_{min}, and $N_{min(low)}$ are the number of initial structures, the number of obtained stable and metastable structures, and the number of those with a lower energy (<0.1 eV/atom vs the stable structure), respectively.

Clusters	Dimension (2N-5)	N_{smp}	N_{min}	$N_{min(low)}$
P_8	11	5000	136	11
P_{10}	15	25 000	1370	13
P_{12}	19	100 000	12 917	20
S_8	11	5000	103	6
S_{10}	15	25 000	1069	23
S_{12}	19	100 000	8666	97

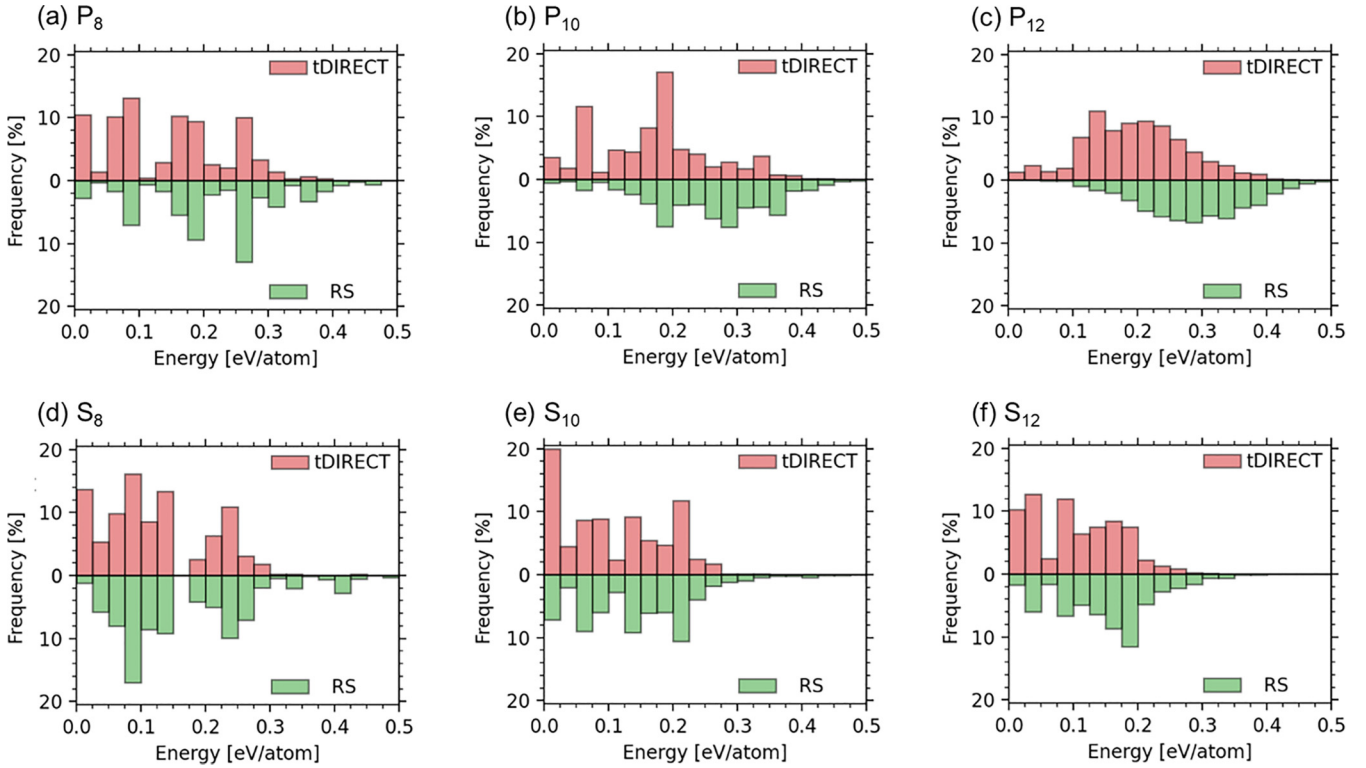FIG. 13. Frequency histograms of convergence to the stable and metastable structures with a total energy in each range (width of 0.025 eV/atom). (a) P_8, (b) P_{10}, (c) P_{12}, (d) S_8, (e) S_{10}, and (f) S_{12} clusters, and the upper and lower histograms in each figure correspond to the ones at the ends of the structure searches in tDIRECT and random search (RS), respectively.

low-energy metastable structures is shown as a function of the number of total energy evaluations corresponding to the computational cost. For comparison, the performances of the two conventional methods RS and BH are also shown. The RS profiles are shown as the expected values estimated from the results of N_{smp} local optimizations. Specifically, the expected number of stable and low-energy metastable structures $\langle N_{min} \rangle$ in n_{smp} local optimizations is estimated as

$$\langle N_{min} \rangle = \sum_i \left[1 - \left(1 - \frac{N_{min,i}}{N_{smp}} \right)^{n_{smp}} \right], \quad (6)$$

where $N_{min,i}$ is the number of initial structures converging to the ith structure in N_{smp} local optimizations, which is proportional to the basin size around the ith structure. The expected number of total energy evaluations per local optimization was set as the average value in N_{smp} local optimizations. Table IV summarizes the initial setting and the results of RS for each cluster. The BH profiles are shown as the average of 10 runs with an error bar of $\pm\sigma$ (σ: standard deviation).

tDIRECT displayed a relatively higher efficiency than RS. The tDIRECT profiles for the phosphorus clusters showed an excellent performance compared with the RS profiles. However, the tDIRECT and RS profiles were similar for the sulfur clusters. The difference in the relative efficiency of tDIRECT to RS by cluster type may be due to the search space. The fraction of the low-energy region in the search space differed between the phosphorus and sulfur clusters. Figure 13 shows the frequency histograms of convergence to the stable and metastable structures with a total energy in each range

(width of 0.025 eV/atom). The upper and lower histograms correspond to the ones at the ends of the structure searches in tDIRECT and RS, respectively. In the RS histograms, the phosphorus clusters were distributed on the higher-energy side relative to those of the sulfur clusters, suggesting that the volume fractions of basins around local minima with a low energy were relatively small in the phosphorus clusters. In such a case, tDIRECT worked efficiently because it sampled the initial structures for local optimizations from potentially optimal hyperrectangles, i.e., possibly low-energy regions. The histograms of tDIRECT were located on the lower-energy side compared with those of RS for the phosphorus clusters, resulting in the higher efficiency of tDIRECT vs RS. In contrast, the RS histograms for the sulfur clusters were on the lower-energy side, and the differences in the histograms between tDIRECT and RS were relatively small. In fact, the sums of the absolute differences in the frequency histograms between the tDIRECT and RS were 61, 70, and 82% for P_8, P_{10}, and P_{12}, respectively. In contrast, these sums were 41, 39, and 48% for S_8, S_{10}, and S_{12}, respectively. Because the difference in sampling the initial structures for local optimizations was relatively small for the sulfur clusters, tDIRECT and RS had comparable efficiencies. Thus, preferential sampling of the initial structures for local optimizations from low-energy regions in tDIRECT worked effectively for search spaces with small low-energy regions.

Comparing the two conventional methods, BH outperformed RS on average, but there was a dependence on the parameter setting (Fig. 12). Among the four profiles with the different scaling factors α (profiles BH_1, BH_07, BH_05,
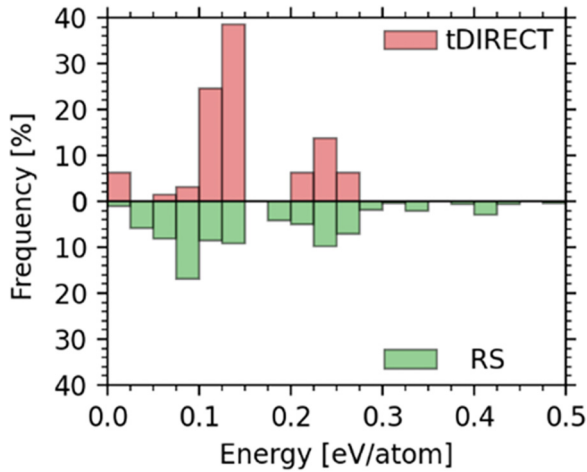
FIG. 14. Frequency histograms of convergence to the stable and metastable structures with a total energy in each range (width of 0.025 eV/atom) for the S_8 cluster. Upper and lower histograms correspond to the one at the early steps of $\sim$2000 energy evaluations in tDIRECT and the one at the end of the structure search in random search (RS), respectively.

and BH_04), the search efficiency was remarkably low at $\alpha = 0.4$ (BH_04) because the walker exploring the PES had difficulty escaping the current basin due to the small displacement. Among the other profiles with $\alpha = 1.0$, 0.7, and 0.5 (BH_1, BH_07, and BH_05), the differences in the search efficiency for the phosphorus clusters were relatively large compared with those for the sulfur clusters. The efficiencies for P_8, P_{10}, and P_{12} were in the order of BH_07 > BH_1 > BH_05, BH_07 > BH_05 > BH_1, and BH_05 > BH_07 > BH_1, respectively, where the appropriate scaling factor α depended on the number of atoms in the phosphorus cluster. Thus, an efficient structure search by BH required careful parameter tuning. Although it seems useful to adaptively change the scaling factor α instead of employing a fixed value, additional parameters such as the initial value or the changing rate of α are necessary.

tDIRECT and BH showed comparable efficiencies for each cluster, but there were a few exceptions. tDIRECT had a lower efficiency than that of BH for the S_8 and P_{12} clusters, particularly in the early steps. Since the DIRECT algorithm selected the hyperrectangles to divide according to the already-computed function values, tDIRECT tended to be less efficient when information about the potential energies was scarce, which occurred in the early steps. Scarce information also led to stagnation in the tDIRECT profiles due to excessively exploring basins around insignificant local minima. This behavior was clearly seen in the structure search for the S_8 cluster by tDIRECT. Figure 14 shows similar frequency histograms for the S_8 cluster in Fig. 13(d). The upper histogram corresponds to the one for the early steps with $\sim$2000 total energy evaluations in tDIRECT, while the lower shows the end of the structure search in RS. The fraction of the obtained stable and metastable structures in the range of 0.125–0.15 eV/atom reached $\sim$40% in the early steps,

which was beyond the definition of low-energy metastable structures. In contrast, BH only used the total energy difference between the current and the displaced points for PES exploration. When the parameters (T and α) were set carefully, BH had a higher search efficiency than tDIRECT and RS, especially in the early steps.

Nonetheless, tDIRECT has advantages to BH from two viewpoints. First, tDIRECT obtains stable and low-energy metastable structures deterministically, unlike the probabilistic BH. Second, tDIRECT is a simple method, as it uses a single insensitive parameter (angle threshold), whereas BH requires a minimum of two parameters (T and α). Considering these advantages, tDIRECT should be a useful method for structure searches.

IV. CONCLUSIONS

We proposed a DIRECT-based structure search method for atomic clusters combined with a local optimizer. tDIRECT performs local optimizations from the centers of all hyperrectangles during the DIRECT process, and the search space is divided according to the local minimum values after local optimizations. This is equivalent to the original DIRECT on a transformed step function, where the function values are constant at the local minimum value within each basin in the original function. The search space for atomic cluster structures is constructed using the Z-matrix representation, in which the atomic positions are specified by bond lengths, bond angles, and dihedral angles. To improve the search efficiency, the search space is restricted by using constant bond lengths and imposing constraints on the bond or dihedral angles.

We applied the proposed method to the Lennard-Jones clusters and two kinds of atomic clusters P_N and S_N, comparing the results with those of the conventional methods RS and BH. For the global optimization problems, tDIRECT is more efficient than RS and comparable to or slightly less efficient than BH on average. For the problems of enumerating the low-energy structures, tDIRECT has a comparable or higher efficiency than that of RS and comparable efficiency to BH with a few exceptions. Although it tends to exhibit a relatively lower efficiency to BH in the enumeration of low-energy structures at the early steps as well as in the global optimization, tDIRECT has advantages as a deterministic and parameter-insensitive structure search method.

[1] A. Törn and A. Žilinskas, *Global Optimization, Lecture Notes in Computer Science*, Vol. 350 (Springer, Berlin, 1989).

[2] R. S. Anderssen and P. Bloomfield, Properties of the random search in global optimization, J. Optim. Theory Appl. **16**, 383 (1975).

[3] W. L. Price, Global optimization by controlled random search, J. Optim. Theory Appl. **40**, 333 (1983).

[4] A. R. Oganov, *Modern Methods of Crystal Structure Prediction* (Wiley-VCH, Weinheim, 2011).

[5] D. J. Wales and J. P. K. Doye, Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms, J. Phys. Chem. A **101**, 5111 (1997).

[6] D. J. Wales and H. A. Scheraga, Global optimization of clusters, crystals, and biomolecules, Science **285**, 1368 (1999).

[7] M. Iwamatsu and Y. Okabe, Basin hopping with occasional jumping, Chem. Phys. Lett. **399**, 396 (2004).

[8] B. Strodel, J. W. L. Lee, C. S. Whittleston, and D. J. Wales, Transmembrane structures for Alzheimer's Aβ_{1-42} oligomers, J. Am. Chem. Soc. **132**, 13300 (2010).

[9] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, Science **220**, 671 (1983).

[10] V. Černý, Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm, J. Optim. Theory Appl. **45**, 41 (1985).

[11] J. Pannetier, J. Bassas-Alsina, J. Rodriguez-Carvajal, and V. Caignaert, Prediction of crystal structures from crystal chemistry rules by simulated annealing, Nature (London) **346**, 343 (1990).

[12] S. Goedecker, Minima hopping: an efficient search method for the global minimum of the potential energy surface of complex molecular systems, J. Chem. Phys. **120**, 9911 (2004).

[13] S. Goedecker, W. Hellmann, and T. Lenosky, Global Minimum Determination of the Born-Oppenheimer Surface within Density Functional Theory, Phys. Rev. Lett. **95**, 055501 (2005).

[14] D. M. Deaven and K. M. Ho, Molecular Geometry Optimization with a Genetic Algorithm, Phys. Rev. Lett. **75**, 288 (1995).

[15] R. L. Johnston, Evolving better nanoparticles: genetic algorithms for optimising cluster geometries, Dalton Trans. 4193 (2003).

[16] A. R. Oganov and C. W. Glass, Crystal structure prediction using *ab initio* evolutionary techniques: principles and applications, J. Chem. Phys. **124**, 244704 (2006).

[17] S. T. Call, D. Yu. Zubarev, and A. I. Boldyrev, Global minimum structure searches via particle swarm optimization, J. Comput. Chem. **28**, 1177 (2007).

[18] Y. Wang, J. Lv, L. Zhu, and Y. Ma, Crystal structure prediction via particle-swarm optimization, Phys. Rev. B **82**, 094116 (2010).

[19] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, Lipschitzian optimization without the Lipschitz constant, J. Optim. Theory Appl. **79**, 157 (1993).

[20] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. P. G. A. Barone, G. A. Petersson, H. J. R. A. Nakatsuji *et al.*, *Gaussian 16, Revision A.03, C02.* (Gaussian, Inc., Wallingford, 2016).

[21] G. Kresse and J. Hafner, *Ab initio* molecular dynamics for open-shell transition metals, Phys. Rev. B **48**, 13115 (1993).

[22] G. Kresse and J. Furthmüller, Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[23] G. M. J. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez *et al.*, Recent developments in the general atomic and molecular electronic structure system, J. Chem. Phys. **152**, 154102 (2020).

[24] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni *et al.*, Advanced capabilities for materials modelling with QUANTUM ESPRESSO, J. Phys. Condens. Matter **29**, 465901 (2017).

[25] J.-L. Faulon and A. Bender, *Handbook of Chemoinformatics Algorithms* (Chapman & Hall/CRC, London, 2010).

[26] B. O. Shubert, A sequential method seeking the global maximum of a function, SIAM J. Numer. Anal. **9**, 379 (1972).

[27] S. A. Piyavskii, An algorithm for finding the absolute extremum of a function, USSR Comput. Math. Math. Phys. **12**, 57 (1972).

[28] R. Horst and P. M. Pardalos, *Handbook of Global Optimization, Vol. 2 of Nonconvex Optimization and Its Applications* (Springer, New York, 1995).

[29] C. A. Floudas and P. M. Pardalos, *Encyclopedia of Optimization*, 2nd Ed. (Springer, New York, 2009).

[30] D. R. Jones and J. R. R. A. Martins, The DIRECT algorithm: 25 years later, J. Glob. Optim. **79**, 521 (2021).

[31] G. Liuzzi, S. Lucidi, and V. Piccialli, A DIRECT-based approach exploiting local minimizations for the solution of large-scale global optimization problems, Comput. Optim. Appl. **45**, 353 (2010).

[32] G. Liuzzi, S. Lucidi, and V. Piccialli, Exploiting derivative-free local searches in DIRECT-type algorithms for global optimization, Comput. Optim. Appl. **65**, 449 (2016).

[33] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, and J. Žilinskas, Globally-biased BIRECT algorithm with local accelerators for expensive global optimization, Expert Syst. Appl. **144**, 113052 (2020).

[34] L. M. Rios and N. V. Sahinidis, Derivative-free optimization: a review of algorithms and comparison of software implementations, J. Glob. Optim. **56**, 1247 (2013).

[35] M. Lipton and W. C. Still, The multiple minimum problem in molecular modeling. Tree searching internal coordinate conformational space, J. Comput. Chem. **9**, 343 (1988).

[36] M. Saunders, K. N. Houk, Y.-D. Wu, W. C. Still, M. Lipton, G. Chang, and W. C. Guida, Conformations of cycloheptadecane. A comparison of methods for conformational searching, J. Am. Chem. Soc. **112**, 1419 (1990).

[37] P. K. Weiner, S. Profeta, G. Wipff, T. Havel, I. D. Kuntz, R. Langridge, and P. A. Kollman, A distance geometry study of ring systems: application to cyclooctane, 18-crown-6, cyclododecane and androstanedione, Tetrahedron **39**, 1113 (1983).

[38] G. M. Crippen and T. F. Havel, *Distance Geometry and Molecular Conformation* (Wiley, New York, 1988), Vol. 15.

[39] M. W. Chase, *NIST-JANAF Thermochemical Tables*, 4th Ed. (American Chemical Society and American Institute of Physics, New York, 1998).

[40] C. Jochum and J. Gasteiger, Canonical numbering and constitutional symmetry, J. Chem. Inf. Comput. Sci. **17**, 113 (1977).

[41] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, Rev. Fr. Inform. Rech. Opér. **3**, 35 (1969).

[42] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen *et al.*, LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. **271**, 108171 (2022).

[43] J. Kostrowicki, L. Piela, B. J. Cherayil, and H. A. Scheraga, Performance of the diffusion equation method in searches for optimum structures of clusters of Lennard-Jones atoms, J. Phys. Chem. **95**, 4113 (1991).

[44] J. P. K. Doye, M. A. Miller, and D. J. Wales, The double-funnel energy landscape of the 38-atom Lennard-Jones cluster, J. Chem. Phys. **110**, 6896 (1999).

[45] M. T. Oakley, R. L. Johnston, and D. J. Wales, Symmetrisation schemes for global optimisation of atomic clusters, Phys. Chem. Chem. Phys. **15**, 3965 (2013).

[46] C. Adamo and V. Barone, Toward reliable density functional methods without adjustable parameters: the PBE0 model, J. Chem. Phys. **110**, 6158 (1999).

[47] P. J. Hay and W. R. Wadt, *Ab initio* effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg, J. Chem. Phys. **82**, 270 (1985).

[48] W. R. Wadt and P. J. Hay, *Ab initio* effective core potentials for molecular calculations. Potentials for main group elements Na to Bi, J. Chem. Phys. **82**, 284 (1985).

[49] P. J. Hay and W. R. Wadt, *Ab initio* effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals, J. Chem. Phys. **82**, 299 (1985).

[50] J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, Walking on potential energy surfaces, J. Phys. Chem. **87**, 2745 (1983).

[51] A. Banerjee, N. Adams, J. Simons, and R. Shepard, Search for stationary points on surfaces, J. Phys. Chem. **89**, 52 (1985).

[52] J. Baker, An algorithm for the location of transition states, J. Comput. Chem. **7**, 385 (1986).

[53] J. Baker, An algorithm for geometry optimization without analytical gradients, J. Comput. Chem. **8**, 563 (1987).

[54] R. Gehrke and K. Reuter, Assessing the efficiency of first-principles basin-hopping sampling, Phys. Rev. B **79**, 085412 (2009).

[55] R. O. Jones and D. Hohl, Structure of phosphorus clusters using simulated annealing—P2 to P8, J. Chem. Phys. **92**, 6710 (1990).

[56] R. O. Jones and G. Seifert, Structure of phosphorus clusters using simulated annealing. II. P_9, P_{10}, P_{11}, anions P_4^{2-}, P_{10}^{2-}, P_{11}^{3-}, and cations P_n^+ to $n = 11$, J. Chem. Phys. **96**, 7564 (1992).

[57] M. Häser, U. Schneider, and R. Ahlrichs, Clusters of phosphorus: a theoretical investigation, J. Am. Chem. Soc. **114**, 9551 (1992).

[58] L. Guo, H. Wu, and Z. Jin, First principles study of the evolution of the properties of neutral and charged phosphorus clusters, J. Mol. Struct. THEOCHEM **677**, 59 (2004).

[59] D. Hohl, R. O. Jones, R. Car, and M. Parrinello, Structure of sulfur clusters using simulated annealing: S_2 to S_{13}, J. Chem. Phys. **89**, 6823 (1988).

[60] K. Raghavachari, C. M. Rohlfing, and J. S. Binkley, Structures and stabilities of sulfur clusters, J. Chem. Phys. **93**, 5862 (1990).

[61] M. D. Chen, M. L. Liu, H. B. Luo, Q. E. Zhang, and C. T. Au, Geometric structures and structural stabilities of neutral sulfur clusters, J. Mol. Struct. THEOCHEM **548**, 133 (2001).

[62] R. Steudel, J. Steidel, and R. Reinhardt, X-ray structural analyses of cyclodecasulfur (S_{10}) and of a cyclohexasulfur-cyclodecasulfur molecular addition compound ($S_6 \cdots S_{10}$) [1], Z. Für Naturforschung B **38**, 1548 (1983).