This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The Spiral Optimization Algorithm: Convergence Conditions and Settings

Kenichi Tamura, *Associate Member, IEEE*, and Keiichiro Yasuda, *Member, IEEE*

***Abstract*—The spiral optimization (SPO) algorithm proposed by Tamura and Yasuda is a relatively novel and simple search concept inspired by natural spiral phenomena. This algorithm searches continuous space using no gradient and only spiral trajectories composed of spiral vectors generated by deterministic spiral models. The primary purpose of this paper is to propose conditions and settings that mathematically ensure the convergence of the SPO algorithm to a stationary point. The conditions relating to the sizes and directions of the spiral vectors and the initial search points are based on direct search theory and recent SPO algorithm theories. The presented convergence was numerically verified using test functions with different properties.**

***Index Terms*—Convergence, direct search, nature-inspired computation, spiral optimization (SPO) algorithm, stationary point.**

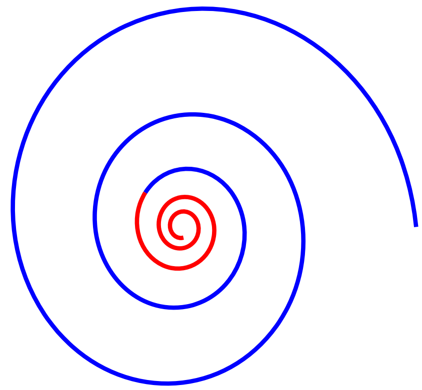


Fig. 1. Logarithmic spiral.

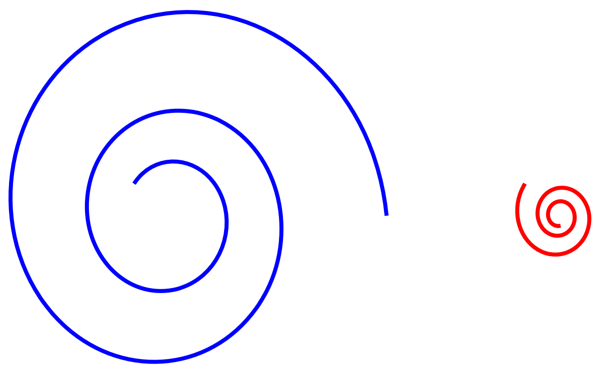


Fig. 2. Interpretation of logarithmic spiral.

I. Introduction

THE spiral optimization (SPO) algorithm was originally proposed by Tamura and Yasuda [1], [2] as an uncomplicated metaheuristic concept inspired by spiral phenomena in nature. The motivation for focusing on spiral phenomena was due to the insight that the dynamics that generate logarithmic spirals (Fig. 1) share the diversification and intensification (Fig. 2), which are important metaheuristic characteristics [3]. Structurally, the SPO algorithm is a multipoint search algorithm that has no objective function gradient, which uses multiple spiral models that can be described as deterministic dynamical systems. As search points follow logarithmic spiral trajectories toward the common center, defined as the current best point, better solutions can be found and the common center can be updated.

Given its origins, the SPO algorithm belongs to the nature-inspired metaheuristics category. Examples include the particle swarm optimization (PSO) [4], [5], which was inspired by bird flocking and fish schooling; the artificial bee colony algorithm [6], [7], inspired by the gulping behavior of bees; and the cuckoo search [8], inspired by the brood parasitism of cuckoos. SPO is comparatively similar to PSO, as both are dynamical system models that do not use evolutionary operations; however, both dynamical systems have different structures and use different information. These algorithms have been applied to various practical system designs (e.g., [9]–[11]) and extended to different problem classes (e.g., [12]–[14]), as they have flexible structures and do not impose severe conditions on the objective function. However, their approaches appear to be theoretically immature, and their convergence to an optimal solution, in particular, has not been strictly clarified.

Considering the structural perspective, the SPO algorithm can belong to the direct search field of nonlinear programming, which comprises methods that exclusively use the relative rank of objective values, rather than the gradient or the approximate gradient [15]. Examples range from the

Manuscript received January 24, 2017; accepted April 4, 2017. This paper was recommended by Associate Editor L. Fang. *(Corresponding author: Kenichi Tamura.)*

The authors are with the Department of Electrical and Electronic Engineering, Tokyo Metropolitan University, Tokyo 1920397, Japan (e-mail: tamu@tmu.ac.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2017.2695577

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

well-known classical pattern search [16] and the simplex algorithm of [17], to modern examples like the multidirectional search [18] and the generalized pattern search [19]. In particular, the generalized pattern search has been generally and strictly proven to converge to a stationary point that is a candidate for optimal solutions, under certain assumptions. These include the continuous differentiability of the objective function.

Following are the main fundamental studies concerning the SPO algorithm. The first SPO algorithm was proposed for 2-D unconstrained optimization [1] based on 2-D spiral models. This was extended to n-dimensional problems by generalizing the 2-D spiral model to an n-dimensional spiral model [2]. The spiral model is comprised of a composite rotation matrix and a step rate, the settings of which become crucial research topics along with initial search points placement, as they define the spiral trajectories used to search. Based on a statistical approach toward many numerical experiments, a step rate setting was proposed [20]. Moreover, an approach to set the step rate was proposed based on a stability analysis [21], [22]. In another proposal, the composite rotation matrix was set and the initial points were placed to ensure having descent directions periodically generated by spiral dynamics [23]. Many extended studies have also been conducted on the SPO due to its simple structure and concept; these studies have helped improve its global search performance [24]–[26] and proposed novel applications [25]–[30].

However, SPO algorithm convergence to an optimal solution has not yet been analyzed as thoroughly as other continuous nature-inspired algorithms. Given that the SPO algorithm deals with optimization problems, its potential to find optimal solutions is an obvious topic for investigation. This paper aims to establish the conditions and settings for the rotation matrix and the step rate with the initial points placement, which would enable the SPO algorithm to converge at a stationary point that is a candidate for optimal solutions. Starting from the insight that the SPO algorithm can be located in the direct search field, we focus on the convergence theory for the generalized pattern search [19] and extract the following two points for convergence relating search vectors added to the current best point.

- Point 1: The direction of at least one search vector is a descent direction that is bounded away from the orthogonal direction to the best point's gradient at any iteration.
- Point 2: The sizes of the search vectors, including at least one descent direction, are decreased only when the searches fail.

The convergence conditions and settings for the SPO algorithm are addressed by applying our previously published results [22], [23] for both points.

In Section II, we define the general form for the SPO algorithm and introduce the conventional setting methods [20]–[23]. Section III establishes the conditions relevant to the rotation matrix and the initial points from point 1 and about the step rate from point 2 according to the recent results [22], [23]. By analyzing Section IV, the SPO algorithm that satisfies the established conditions, denoted as the convergent SPO algorithm, is proven to converge to a stationary point in Section V. The analysis of Section IV focuses on the dynamics of the search points following the convergent SPO algorithm. The proof of Section V is finally conducted by a contradiction method relating to the step rate based on Section IV's analysis. Section VI proposes concrete settings that realize the established convergence conditions and numerically verifies the convergence of the convergent SPO algorithm for test functions.

The following notations are used in this paper. $\mathbb{R}$: the set of real numbers, $\mathbb{Z}$: the set of integer numbers, $\mathbb{N}$: the set of natural numbers, I_n: the $n \times n$ unit identity matrix, $\|\boldsymbol{x}\|$: Euclidean norm of $\boldsymbol{x} \in \mathbb{R}^n$, $\mathbb{L}(\boldsymbol{y}) := \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) \leqq f(\boldsymbol{y})\}$, $\mathbb{C}(\boldsymbol{y}) := \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) = f(\boldsymbol{y})\}$, $\mathbb{B}(\boldsymbol{y}, c) := \{\boldsymbol{x} \in \mathbb{R}^n | \|\boldsymbol{x} - \boldsymbol{y}\| < c\}$, $c > 0$, $\mathbb{X}_\sharp := \{\boldsymbol{x} \in \mathbb{R}^n | \nabla f(\boldsymbol{x}) = \boldsymbol{0}\}$, and dist $(\mathbb{X}, \mathbb{Y}) := \inf\{\|\boldsymbol{x} - \boldsymbol{y}\| | \boldsymbol{x} \in \mathbb{X}, \boldsymbol{y} \in \mathbb{Y}\}$.

II. Definition of the SPO Algorithm

This section introduces the SPO algorithm [2], [22], [23] in a general form following the formulation of the spiral model and its simulations. Then, the conventional setting methods for the algorithm as a metaheuristic are explained.

A. Spiral Model

The SPO algorithm utilizes logarithmic spirals generated by n-dimensional spiral models. The spiral model is a dynamical system whose state $\boldsymbol{x}(k) \in \mathbb{R}^n$ $(n \geqq 2)$ converges to a center $\boldsymbol{x}^\star \in \mathbb{R}^n$ from an initial point $\boldsymbol{x}(0)$ with a logarithmic spiral trajectory

$$\boldsymbol{x}(k+1) = \boldsymbol{x}^\star + rR(\theta)\big(\boldsymbol{x}(k) - \boldsymbol{x}^\star\big) \ (k = 0, 1, 2, \ldots) \tag{1}$$

where $rR(\theta)(\boldsymbol{x}(k) - \boldsymbol{x}^\star)$ is the spiral vector, $r > 0$ is the step rate of the distance between $\boldsymbol{x}(k)$ and $\boldsymbol{x}^\star$ per k, $\theta \in (-\pi, \pi]$ is the rotation rate of $\boldsymbol{x}(k)$ around the center $\boldsymbol{x}^\star$ per k, and $R(\theta) \in \mathbb{R}^{n\times n}$ is the composite rotation matrix generally defined by arbitrarily multiplying τ types of basic rotation matrices $R_{i_\ell, j_\ell}(\theta) \in \mathbb{R}^{n\times n}$ $(\ell = 1, \ldots, \tau)$ as follows:

$$R(\theta) = (-1)^\beta R_{i_1,j_1}(\theta) \times \cdots \times R_{i_\tau,j_\tau}(\theta), \quad \beta \in \{0, 1\}$$

$$R_{i_\ell,j_\ell}(\theta) = \begin{array}{cc} & \begin{array}{ccccccccccc} & & & i_\ell & & & & j_\ell & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i_\ell \\ \\ \\ \\ j_\ell \\ \\ \\ \\ \end{array} & \begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & \cos\theta & & & & -\sin\theta & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & \sin\theta & & & & \cos\theta & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix} \end{array} \tag{2}$$

where $i_\ell, j_\ell \in \{1, \ldots, n\}$, $i_\ell < j_\ell$, $(\ell = 1, \ldots, \tau)$ and the blank elements indicate 0.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

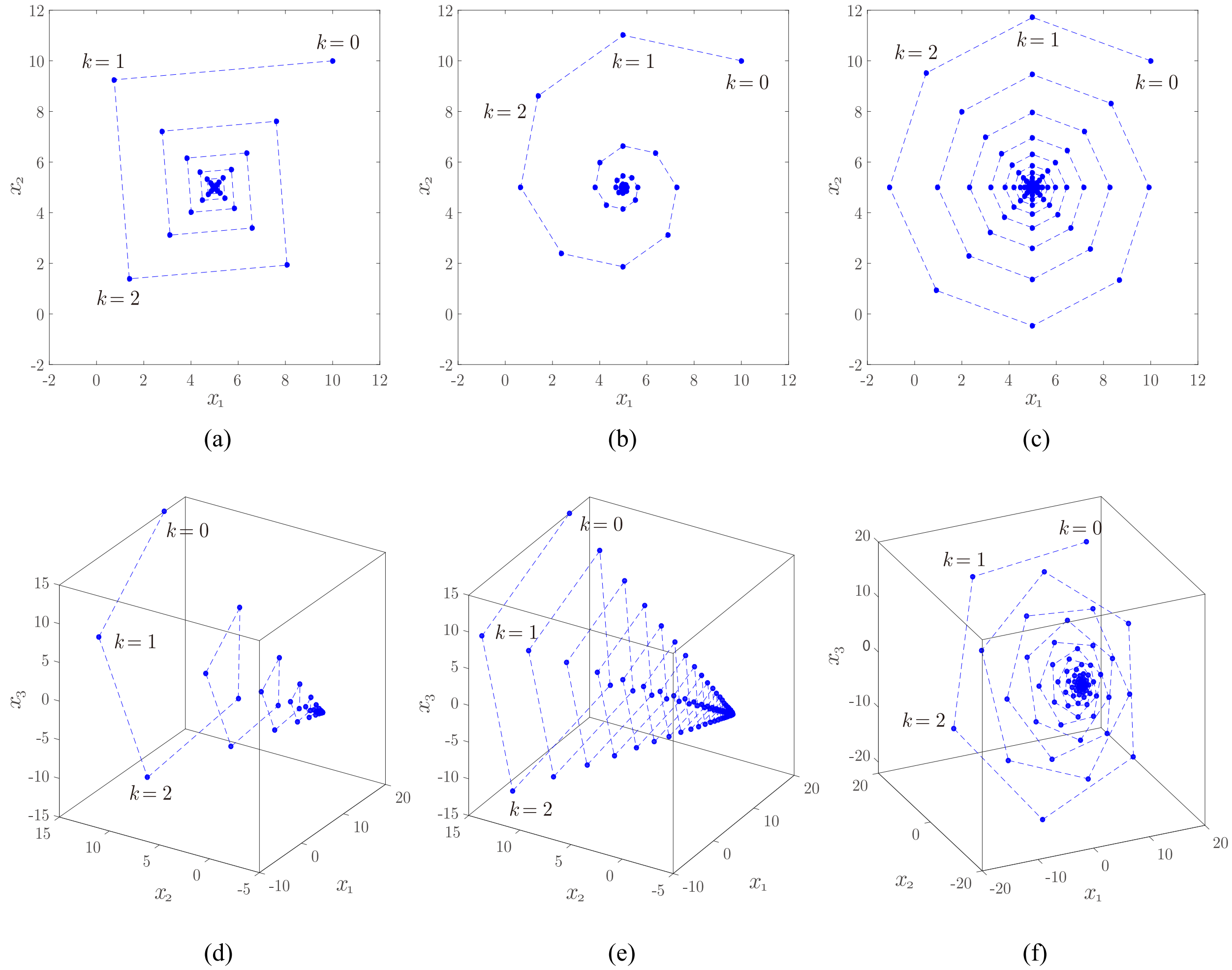


Fig. 3. Examples of trajectories of (1). (a) $n = 2$, $\theta = \pi/2$, $r = 0.85$. (b) $n = 2$, $\theta = \pi/4$, $r = 0.85$. (c) $n = 2$, $\theta = \pi/4$, $r = 0.95$. (d) $n = 3$, $\theta = \pi/2$, $r = 0.85$. (e) $n = 3$, $\theta = \pi/2$, $r = 0.95$. (f) $n = 3$, $\theta = \pi/4$, $r = 0.95$.

Fig. 3 shows examples of the trajectories of (1) with $R(\theta) = R_{1,2}(\theta), \boldsymbol{x}^\star = [5\ 5]^\top, \boldsymbol{x}(0) = [10\ 10]^\top$ and the three types of parameter settings Fig. 3(a)–(c) in a 2-D space and with $R(\theta) = R_{2,3}(\theta)R_{1,3}(\theta)R_{1,2}(\theta), \boldsymbol{x}^\star = [5\ 0\ 5]^\top, \boldsymbol{x}(0) = [15\ 15\ 15]^\top$ and the three types of parameter settings Fig. 3(d)–(f) in a 3-D space. In both cases, we can observe the spiral trajectories generated around $\boldsymbol{x}^\star$ and the parameters' effects.

Note that the conventional approach [2], [22], [23] has been to use an expression of the dynamical system form $\boldsymbol{x}(k+1) = rR(\theta)\boldsymbol{x}(k) + (I_n - rR(\theta))\boldsymbol{x}^\star$ that is essentially the same as (1).

B. General SPO Algorithm

Spiral trajectories are interesting natural phenomena to search because they are characterized by diversification and intensification properties that are important for metaheuristics [3]. Intensification intends to search locally and intensively, while diversification intends to search globally and roughly.

The SPO algorithm focuses on utilizing spiral trajectories that have both search diversification and intensification aspects. For the SPO algorithm to minimize an objective function $f : \mathbb{R}^n \to \mathbb{R}$ $(n \geqq 2)$

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}}\ f(\boldsymbol{x})$$

is generally described as Algorithm 1. This algorithm uses m $(\geqq 2)$ of the spiral model (1) by replacing the fixed constant step rate r and the given constant center $\boldsymbol{x}^\star$ with the tunable step rate $r(k)$ and the common center $\boldsymbol{x}^\star(k)$ defined as the best solution until the kth iteration, respectively.

Fig. 4 illustrates this search process such that the search points starting from the initial state can proceed and update the center defined as the current best solution toward the optimal solution using the spiral trajectories from diversification to intensification.

Remark 1: This algorithm can be further generalized by selecting a different setting for each search point, as in $R_i(\theta_i)$ and $r_i(k)$ $(i = 1, \ldots, m)$. This paper considers the simple case of Algorithm 1.

C. Conventional Settings for Metaheuristic

The performance of Algorithm 1 is strongly influenced by the settings of the composite rotation matrix $R(\theta)$ and the step rate $r(k)$, with each initial point placement $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ as each setting specifies a spiral trajectory; hence,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

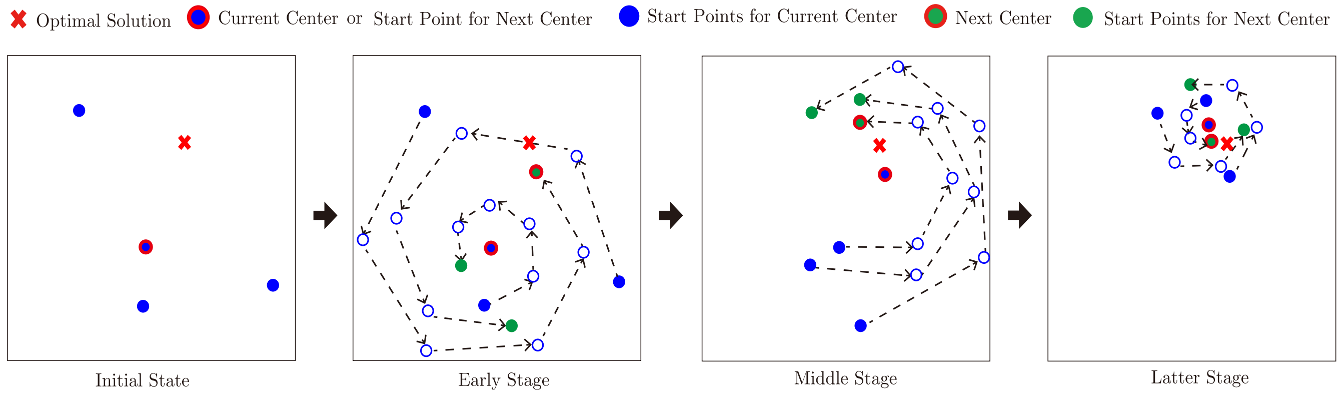


Fig. 4. Illustration of SPO algorithm's search process from diversification to intensification.

Algorithm 1 General SPO Algorithm

Step 0. Set the number of search points $m \geqq 2$, the rotation matrix $R(\theta)$, the setting rule for the step rate $r(k)$, and the termination criterion.

Step 1. Specify the initial search points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ with different objective function values and determine the center $\boldsymbol{x}^\star(0) = \boldsymbol{x}_{i_\mathrm{b}}(0)$, $i_\mathrm{b} = \min\{\operatorname{argmin}_{i=1,\ldots,m}\{f(\boldsymbol{x}_i(0))\}\}$, and then set $k = 0$.

Step 2. Determine the step rate $r(k)$ according to the setting rule.

Step 3. Update the search points as follows:

$$\boldsymbol{x}_i(k+1) = \boldsymbol{x}^\star(k) + r(k)R(\theta)(\boldsymbol{x}_i(k) - \boldsymbol{x}^\star(k)) \quad (i = 1, \ldots, m). \tag{3}$$

Step 4. Update the center as follows:

$$\boldsymbol{x}^\star(k+1) = \begin{cases} \boldsymbol{x}_{i_\mathrm{b}}(k+1) \ \left(\text{if } f(\boldsymbol{x}_{i_\mathrm{b}}(k+1)) < f(\boldsymbol{x}^\star(k))\right) \\ \boldsymbol{x}^\star(k) \ (\text{otherwise}) \end{cases}$$

where $i_\mathrm{b} = \min\{\operatorname{argmin}_{i=1,\ldots,m}\{f(\boldsymbol{x}_i(k+1))\}\}$.

Step 5. Set $k := k+1$. If the termination criterion is satisfied then terminate and output $\boldsymbol{x}^\star(k)$. Otherwise, return to Step 2.

their design is a crucial factor for this algorithm. Conventional methods for Algorithm 1 have been studied as practical metaheuristics for cases when the termination criterion is the maximum iteration number $k_{\max}$ supplied by the user.

1) Tamura and Yasuda [20] proposed a constant setting for the step rate $r(k)$ based on its quantitative analysis through repeated numerical experimentation. The design method, which is a function of $k_{\max}$, was derived to have that the average distance curve between the center and each point matched to an ideal curve based on a large statistical data set.
2) Tamura and Yasuda [21], [22] demonstrated a constant setting for the step rate $r(k)$ based on stability analysis of the center $\boldsymbol{x}^\star(k)$. This method allows the algorithm to progress from diversification in the early stage to intensification in the latter phase, making it suitable for the $k_{\max}$.
3) Tamura and Yasuda [23] proposed a setting for the rotation matrix $R(\theta)$ with the initial points placement $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ to periodically generate at least one descent direction with a $2n$ period, and a constant setting for the step rate $r(k)$ to utilize the descent direction effectively in the later stages to approach $k_{\max}$.

These practical metaheuristic methods could not be mathematically guaranteed to converge to a stationary point.

III. Convergent Spiral Optimization Algorithm

This section proposes the SPO algorithm's setup conditions to converge at a stationary point of the objective function. The conditions are derived based on the direct search theory and the previous theories on the SPO algorithm. The convergence proof will be supplied in Section V according to the analysis in Section IV.

A. Clues to Convergence

The search performance of Algorithm 1 depends on setting the composite rotation matrix $R(\theta)$, the step rate $r(k)$, and the initial points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$. The purpose of this section is to find clues to these setting conditions that enable Algorithm 1 to converge to a stationary point.

As described in Section I, structurally the SPO algorithm can belong to the direct search field in nonlinear programming. To date, many direct search algorithms have been studied [16]–[19]. In particular, the generalized pattern search [19], which is relatively modern, has been generally and strictly proven to converge to a stationary point in the case of a continuous differentiable objective function.

The generalized pattern search iteratively proceeds by adding some finite search vectors to the current best point and adjusting the vector sizes based on the objective values. By analyzing this mechanism's theory, we can extract the following two points for convergence in terms of the directions and sizes of the search vectors added to the current best point.

Point 1: The direction of at least one search vector is a descent direction that is bounded away from orthogonal direction of the gradient at the current best point for any iteration.

Point 2: The sizes of the search vectors, including at least one descent direction, are decreased only when the searches fail.

Point 1 is typical in convergence proofs of minimization algorithms with descent directions in nonlinear programming [31], and point 2 is natural for algorithms with descent directions because a descent direction vector reduces the objective function when its size is sufficiently small [31], [32].

Now, based on the two points, we consider setup conditions about the composite rotation matrix $R(\theta)$, the step rate $r(k)$ and the initial points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$. From the updating laws in step 3, because $\boldsymbol{x}^\star(k)$ is defined as the current best point, the spiral vectors $r(k)R(\theta)(\boldsymbol{x}_i(k) - \boldsymbol{x}^\star(k))$ $(i = 1, \ldots, m)$ dynamically work as the search vectors added to the current best point and then should be analyzed in terms of their directions and sizes.

To analyze them, defining the different vectors between any two search points $\boldsymbol{d}_{i,j}(k) := \boldsymbol{x}_i(k) - \boldsymbol{x}_j(k)$ $(i, j = 1, \ldots, m)$ and the index $i_k^\star$ such that $\boldsymbol{x}^\star(k) = \boldsymbol{x}_{i_k^\star}(k)$, we represent the updating laws (3) as

$$\boldsymbol{x}_i(k+1) = \boldsymbol{x}^\star(k) + r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k) \ \ (i = 1, \ldots, m). \quad (4)$$

From this, we can derive

$$\boldsymbol{d}_{i,j}(k) = \gamma(k-1)R(\theta)^k \boldsymbol{d}_{i,j}(0) \ \ (i, j = 1, \ldots, m) \quad (5)$$

where

$$\gamma(k-1) := \prod_{s=0}^{k-1} r(s). \quad (6)$$

This means the dynamical relation between $\boldsymbol{x}_i(k)$ and $\boldsymbol{x}_j(k)$ is invariant for any center.

Then, based on the relation (5), the spiral vectors $r(k)R(\theta)(\boldsymbol{x}_i(k) - \boldsymbol{x}^\star(k)) = r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k)$ $(i = 1, \ldots, m)$ are

$$r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k) = \gamma(k)R(\theta)^{k+1}\boldsymbol{d}_{i,i_k^\star}(0) \quad (7)$$

and their norms are

$$\left\| \gamma(k)R(\theta)^{k+1}\boldsymbol{d}_{i,i_k^\star}(0) \| = \gamma(k) \|\boldsymbol{d}_{i,i_k^\star}(0) \right\|. \quad (8)$$

From (6) and (8), $r(k)$ determines the sizes of the spiral vectors in the search process, but the rotation matrix $R(\theta)$ and the initial points $\boldsymbol{x}_i(0)$ cannot affect the sizes. Meanwhile, from (7), both rotation matrix $R(\theta)$ and initial points $\boldsymbol{x}_i(0)$ determine the directions of the spiral vectors, but $r(k)$ (> 0) cannot change the directions. Thus, we can consider conditions on the initial points $\boldsymbol{x}_i(0)$ and $R(\theta)$ from point 1 and on $r(k)$ from point 2 independently.

B. Conditions From Point 1

Here, we consider conditions on the composite rotation matrix $R(\theta)$ and the initial points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ so that the dynamics of spiral vectors satisfy point 1. To approach this task, we focus on the following two conditions established in [23], under which at least one of the spiral vectors generates a descent direction every $2n$ iterations at the center $\boldsymbol{x}^\star(k_0)$ for any iteration k_0.

Condition 1: Set the composite rotation matrix $R(\theta)$, i.e., $(i_1, j_1), \ldots, (i_\tau, j_\tau)$, β and θ in (2), to satisfy the following conditions.

1) It is a periodic matrix with period $2n$

$$R(\theta)^{2n} = I_n.$$

2) For any $i \in \mathbb{I} = \{0, \ldots, 2n-1\}$, there is a unique $j \in \mathbb{I}$ such that

$$R(\theta)^i = -R(\theta)^j.$$

Condition 2: Set the initial points $\boldsymbol{x}_i(0) \in \mathbb{R}^n$ to satisfy the following: for each $i \in \{1, \ldots, m\}$, there exists $p \in \{1, \ldots, m\}$, $\ell_1, \ldots, \ell_n \in \mathbb{I} = \{0, \ldots, 2n-1\}$ such that

$$R(\theta)^{\ell_1}\boldsymbol{d}_{p,i}(0), \ldots, R(\theta)^{\ell_n}\boldsymbol{d}_{p,i}(0) \quad (9)$$

are linearly independent, where $\boldsymbol{d}_{p,i}(0) = \boldsymbol{x}_p(0) - \boldsymbol{x}_i(0)$.

However, in [23], this proof did not address the gist of point 1, "bounded away from the orthogonal direction of the gradient." In the following theorem, we solve this gist by proving that the two conditions enable at least one of the spiral vectors (7) to have a descent direction with boundedness of the inner product away from zero against $\nabla f(\boldsymbol{x}^\star(k_0))$ from any iteration k_0 within $2n$ iterations.

Theorem 1: Let $f : \mathbb{R}^n \to \mathbb{R}$ $(n \geqq 2)$ be continuously differentiable. Suppose that Algorithm 1 satisfies conditions 1 and 2. There exists a constant $\sigma > 0$ such that, for the gradient $\nabla f(\boldsymbol{x}^\star(k_0)) \neq \boldsymbol{0}$ at any iteration k_0, there exists $p \in \{0, \ldots, m\}$ and $\overline{q} \in \{0, \ldots, 2n-1\}$ such that the spiral vector $r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i_{k_0}^\star}(k_0 + \overline{q})$ satisfies

$$\frac{\nabla f(\boldsymbol{x}^\star(k_0))^\top}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|} \frac{r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i_{k_0}^\star}(k_0 + \overline{q})}{\left\| r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i_{k_0}^\star}(k_0 + \overline{q}) \right\|} \leqq -\sigma. \quad (10)$$

Proof: First let us more strictly represent p and $\ell_1, \ldots, \ell_n$ for each i in condition 2 as $p(i)$ and $\ell_1(i), \ldots, \ell_n(i)$, respectively.

From the assumption of condition 2, for an index $i_0^\star$ of the center $\boldsymbol{x}^\star(k_0) = \boldsymbol{x}_{i_{k_0}^\star}(k_0)$, there exists an index $p(i_{k_0}^\star)$ whose initial point satisfies the linear independence of (9). From (5), the spiral vector between $\boldsymbol{x}_{p(i_{k_0}^\star)}(k)$ and $\boldsymbol{x}_{i_{k_0}^\star}(k)$ is given by

$$r(k)R(\theta)\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(k) = \gamma(k)R(\theta)^{k+1}\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(0).$$

Thus, the unit vectors of $2n$ spiral vectors $r(k_0 + s)R(\theta)\boldsymbol{d}_{p(i_{k_0}^\star),i_{k_0}^\star}(k_0 + s)$ $(s = 0, \ldots, 2n-1)$ are described as

$$\begin{aligned} &\frac{r(k_0 + s)R(\theta)\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(k_0 + s)}{\left\| r(k_0 + s)R(\theta)\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(k_0 + s) \right\|} \\ &= \frac{R(\theta)^{k_0+s+1}\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(0)}{\left\| R(\theta)^{k_0+s+1}\boldsymbol{d}_{p\left(i_{k_0}^\star\right),i_{k_0}^\star}(0) \right\|} \ \ (s = 0, \ldots, 2n-1). \quad (11) \end{aligned}$$

Further, from conditions 1 and 2, these vectors satisfy

$$\frac{R(\theta)^{k_0+s+1}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)}{\left\|R(\theta)^{k_0+s+1}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|} \in \left\{\frac{R(\theta)^{0}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)}{\left\|R(\theta)^{0}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|},\ldots,\frac{R(\theta)^{2n-1}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)}{\left\|R(\theta)^{2n-1}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|}\right\}$$
$$= \left\{\frac{\pm R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)}{\left\|R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|}\ (j=1,\ldots,n)\right\}$$
$$(s=0,\ldots,2n-1). \tag{12}$$

Next, we consider the absolute value of the inner products of the unit vector of the gradient $\nabla f(\boldsymbol{x}^\star(k_0)) \neq \boldsymbol{0}$ and the unit vectors of the $2n$ spiral vectors $r(k_0+s)R(\theta)\boldsymbol{d}_{p(i^\star_{k_0}),i^\star_{k_0}}(k_0+s)$ $(s=0,\ldots,2n-1)$:

$$\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right\|}$$
$$(s=0,\ldots,2n-1).$$

Using (11) and (12)

$$\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right\|}$$
$$\in \left\{\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|}\ (j=1,\ldots,n)\right\}(s=0,\ldots,2n-1). \tag{13}$$

Define the square matrix

$$D\left(i^\star_{k_0}\right) = \left[R(\theta)^{\ell_1(i^\star_{k_0})}\boldsymbol{d}_{p(i^\star_{k_0}),i^\star_{k_0}}(0)\ \ \ldots\ \ R(\theta)^{\ell_n(i^\star_{k_0})}\boldsymbol{d}_{p(i^\star_{k_0}),i^\star_{k_0}}(0)\right]$$

which is nonsingular according to condition 2. Using this matrix and a unit vector $\boldsymbol{e}_j$ whose jth element is 1

$$R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0) = D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\ (j=1,\ldots,n).$$

Thus, since each element of the set in (13) is given by

$$\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))^\top\|\left\|R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|} = \frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))^\top\|\left\|D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\right\|}$$

we define $\boldsymbol{\eta} = D(i^\star_{k_0})^\top\nabla f(\boldsymbol{x}^\star(k_0))$ and use the basic matrix norm property $\|A\boldsymbol{y}\| \leqq \|A\|\|\boldsymbol{y}\|(\forall A \in \mathbb{R}^{n\times n}, \boldsymbol{y} \in \mathbb{R}^n)$ to obtain

$$\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\right\|} = \frac{|\boldsymbol{\eta}^\top\boldsymbol{e}_j|}{\left\|D\left(i^\star_{k_0}\right)^{-\top}\boldsymbol{\eta}\right\|\left\|D\left(i^\star_{k_0}\right)\boldsymbol{e}_j\right\|} \geqq \frac{1}{\zeta\left(D\left(i^\star_{k_0}\right)\right)}\left(\frac{|\boldsymbol{\eta}^\top\boldsymbol{e}_j|}{\|\boldsymbol{\eta}\|\|\boldsymbol{e}_j\|}\right)$$

where $\zeta(D(i^\star_{k_0})) := \|D(i^\star_{k_0})^{-1}\|\|D(i^\star_{k_0})\|$ is a condition number and only dependent on the initial search points.

From this inequality, the relation (13), and Lemma A in Appendix A, for any $\nabla f(\boldsymbol{x}^\star(k_0)) \neq \boldsymbol{0}$

$$\max\left\{\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))^\top r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|r(k_0+s)R(\theta)\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(k_0+s)\right\|}\ (s=0,\ldots,2n-1)\right\}$$
$$= \max\left\{\frac{\left|\nabla f(\boldsymbol{x}^\star(k_0))R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right|}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|\left\|R(\theta)^{\ell_j\left(i^\star_{k_0}\right)}\boldsymbol{d}_{p\left(i^\star_{k_0}\right),i^\star_{k_0}}(0)\right\|}\ (j=1,\ldots,n)\right\} \geqq \frac{1}{\zeta\left(D\left(i^\star_{k_0}\right)\right)}\frac{1}{\sqrt{n}} > 0$$

holds. According to condition 1-2), there exists $\overline{q} \in \{0,\ldots,2n-1\}$ such that

$$\frac{\nabla f(\boldsymbol{x}^\star(k_0))^\top}{\|\nabla f(\boldsymbol{x}^\star(k_0))\|}\frac{r(k_0+\overline{q})R(\theta)\boldsymbol{d}_{p(i^\star_{k_0}),i^\star_{k_0}}(k_0+\overline{q})}{\left\|r(k_0+\overline{q})R(\theta)\boldsymbol{d}_{p(i^\star_{k_0}),i^\star_{k_0}}(k_0+\overline{q})\right\|} \leqq -\frac{1}{\zeta\left(D\left(i^\star_{k_0}\right)\right)}\frac{1}{\sqrt{n}} < 0.$$

Finally, defining

$$\sigma = \frac{1}{\zeta_{\max}}\frac{1}{\sqrt{n}},\ \ \zeta_{\max} = \max\{\zeta(D(j))\ (j=1,\ldots,m)\} \tag{14}$$

that is dependent on all initial search points, we have proven this theorem. ■

C. Conditions From Point 2

Here, we consider setting conditions on the step rate $r(k)$ from point 2.

First, we consider a criterion to judge the success or failure of the search in the iteration process since the spiral vectors dynamically work as search vectors added to the current best point. To achieve this, we use $2n$ iterations span, which guarantees a descent direction on point 1 by Theorem 1, as the criterion to judge the success or failure. That is, we judge the search process to be successful if the current best point is updated within the $2n$ iterations after it has been newly set; otherwise, the search process is judged to be unsuccessful.

Then, we consider switched values of $r(k)$ for the decrease in case of failure or for the nondecrease in case of success, respectively. From (8), we can easily state a policy that switches $r(k)$ as $r(k) = h \in (0, 1)$ when the search fails or as $r(k) = h \geq 1$ when it succeeds. Meanwhile, by extending the stability analysis [22], relations between $r = h$ and the behavior of spiral model (1) are investigated.

1) *The Center Is Asymptotically Stable:* $\|\boldsymbol{x}(k) - \boldsymbol{x}^\star\| \to 0 \ \ (k \to \infty)$ if $r = h \in (0, 1)$ (converging spirals).
2) *The Center Is Neutrally Stable:* $\|\boldsymbol{x}(k) - \boldsymbol{x}^\star\| \equiv \|\boldsymbol{x}(0) - \boldsymbol{x}^\star\| (k \to \infty)$ if $r = h = 1$ (oscillating spirals).
3) *The Center Is Unstable:* $\|\boldsymbol{x}(k) - \boldsymbol{x}^\star\| \to \infty \ (k \to \infty)$ if $r = h > 1$ (diverging spirals).

Thus, in other words, this policy means that not only convergent but also nonconvergent spiral trajectories are used for the switch depending on the search situation.

From the above consideration, we propose the following condition on the step rate $r(k)$ from point 2.

Condition 3: Define an iteration when the center is newly updated as $k^\star$. Set the parameter $r(k)$ by the following rule:

$$r(k) = \begin{cases} 1 & \left(k^\star \leqq k \leqq k^\star + 2n - 1\right) \\ h & \left(k \geqq k^\star + 2n\right) \end{cases} \tag{15}$$

where $h \in (0, 1) \subset \mathbb{R}$.

This means that h is a constant parameter set by the user prior to running the algorithm. Here, we do not use divergent spirals with $h > 1$, because the analysis of the following sections is more complex.

D. Convergent SPO Algorithm

Algorithm 1 with conditions 1–3, called convergent SPO algorithm, is represented in Algorithm 2.

We finalize this section by showing the main theorem for Algorithm 2's convergence that will be proved in Section V based on the analysis in Section IV.

Theorem 2: Define the worst initial search point by $\boldsymbol{x}_\star(0) := \boldsymbol{x}_{i_{\rm w}}(0)$, $i_{\rm w} := \max\{\underset{i=1,\ldots,m}{\operatorname{argmax}}\{f(\boldsymbol{x}_i(0))\}\}$. Suppose that $\mathbb{L}(\boldsymbol{x}_\star(0))$ is compact and $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on a neighborhood of $\mathbb{L}(\boldsymbol{x}_\star(0))$. Then, the sequence of the center $\{\boldsymbol{x}^\star(k)\}$ produced by Algorithm 2 satisfies

$$\liminf_{k\to+\infty} \left\|\nabla f\left(\boldsymbol{x}^\star(k)\right)\right\| = 0. \tag{16}$$

Proof: This proof will be at the end of Section V. ■

Algorithm 2 Convergent SPO Algorithm

Step 0. Set the number of search points $m \geqq 2$ and the termination criterion. Set the composite rotation matrix $R(\theta)$ with Condition 1 and $h \in (0, 1)$ of Condition 3.

Step 1. Specify the initial search points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ with Condition 2 and different objective function values and determine the center $\boldsymbol{x}^\star(0) = \boldsymbol{x}_{i_{\rm b}}(0), i_{\rm b} = \min\{\underset{i=1,\ldots,m}{\operatorname{argmin}}\{f(\boldsymbol{x}_i(0))\}\}$, and then set $k = 0$ and $k^\star = 0$.

Step 2. Set the parameter $r(k)$ as follows:

$$r(k) = \begin{cases} 1 & (k^\star \leqq k \leqq k^\star + 2n - 1) \\ h & (k \geqq k^\star + 2n). \end{cases}$$

Step 3. Update the search points as follows:

$$\boldsymbol{x}_i(k+1) = \boldsymbol{x}^\star(k) + r(k)R(\theta)(\boldsymbol{x}_i(k) - \boldsymbol{x}^\star(k))$$
$$(i = 1, \ldots, m).$$

Step 4. Update the center as follows:

$$\boldsymbol{x}^\star(k+1) = \begin{cases} \boldsymbol{x}_{i_{\rm b}}(k+1) & \left(\text{if } f(\boldsymbol{x}_{i_{\rm b}}(k+1)) < f(\boldsymbol{x}^\star(k))\right) \\ \boldsymbol{x}^\star(k) & (\text{otherwise}) \end{cases}$$

where $i_{\rm b} = \min\{\underset{i=1,\ldots,m}{\operatorname{argmin}}\{f(\boldsymbol{x}_i(k+1))\}\}$. Furthermore, if $\boldsymbol{x}^\star(k+1) \neq \boldsymbol{x}^\star(k)$, then $k^\star = k + 1$.

Step 5. Set $k := k+1$. If the termination criterion is satisfied, then terminate and output $\boldsymbol{x}^\star(k)$. Otherwise, return to Step 2.

IV. Analysis of Dynamics

In this section, we analyze the dynamics of the search points following Algorithm 2 as a background for the proof presented in Section V. The following results are unique to Algorithm 2 and independent of the objective function.

Here, we use the description of the updating laws of step 3 as in Section III-A

$$\boldsymbol{x}_i(k+1) = \boldsymbol{x}^\star(k) + r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k) \ \ (i = 1, \ldots, m) \tag{17}$$

where $\boldsymbol{d}_{i,j}(k) = \boldsymbol{x}_i(k) - \boldsymbol{x}_j(k)$ and $\boldsymbol{x}^\star(k) = \boldsymbol{x}_{i_k^\star}(k)$. Furthermore, from Section III-A, we recall that their spiral vectors $r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k)$ $(i = 1, \ldots, m)$ can be expressed as follows:

$$r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k) = \gamma(k)R(\theta)^{k+1}\boldsymbol{d}_{i,i_k^\star}(0) \tag{18}$$

where

$$\gamma(k) = \prod_{s=0}^{k} r(s). \tag{19}$$

The next lemma represents the search points $\boldsymbol{x}_i(k)$ $(i = 1, \ldots, m;\ k \geqq 1)$ that follow Algorithm 2.

Lemma 1: The search points $\boldsymbol{x}_i(k)$ $(i = 1, \ldots, m;\ k \geqq 1)$ of Algorithm 2 are of the form

$$\boldsymbol{x}_i(k) = \boldsymbol{x}^\star(0) + \sum_{s=0}^{k-2} \gamma(s)R(\theta)^{s+1}\boldsymbol{d}_{i_{s+1}^\star, i_s^\star}(0) + \gamma(k-1)R(\theta)^k \boldsymbol{d}_{i,i_{k-1}^\star}(0) \ \ (i = 1, \ldots, m). \tag{20}$$

Note that the second term in the right-hand side is ignored when $k=1$.

Proof: This proof procedure is conducted by mathematical induction. First, from (17), we have

$$\boldsymbol{x}_i(1)=\boldsymbol{x}^\star(0)+r(0)R(\theta)\boldsymbol{d}_{i,i_0^\star}(0)\ \ (i=1,\ldots,m)$$

and (20) holds when $k=1$.

Then, assuming that (20) holds when $k=t(>1)$, we consider the case when $k=t+1$. Because

$$\boldsymbol{x}_i(t+1)=\boldsymbol{x}^\star(t)+r(t)R(\theta)\boldsymbol{d}_{i,i_t^\star}(t)\ \ (i=1,\ldots,m)$$

from (17), and based on our assumption we obtain the following:

$$\begin{aligned}\boldsymbol{x}_i(t+1)&=\boldsymbol{x}^\star(0)+\sum_{s=0}^{t-2}\gamma(s)R(\theta)^{s+1}\boldsymbol{d}_{i_{s+1}^\star,i_s^\star}(0)\\&\quad+\gamma(t-1)R(\theta)^t\boldsymbol{d}_{i_t^\star,i_{t-1}^\star}(0)+r(t)R(\theta)\boldsymbol{d}_{i,i_t^\star}(t)\\&=\boldsymbol{x}^\star(0)+\sum_{s=0}^{t-1}\gamma(s)R(\theta)^{s+1}\boldsymbol{d}_{i_{s+1}^\star,i_s^\star}(0)\\&\quad+\gamma(t)R(\theta)^{t+1}\boldsymbol{d}_{i,i_t^\star}(0).\end{aligned}$$

This result equals (20) at $k=t+1$, as a result, the lemma has be proven. ■

The next lemma describes the center $\boldsymbol{x}^\star(k)=\boldsymbol{x}_{i_k^\star}(k)$ based on Lemma 1.

Lemma 2: The center $\boldsymbol{x}^\star(k)$ of Algorithm 2 is represented as follows:

$$\boldsymbol{x}^\star(k)=\boldsymbol{x}^\star(0)+\sum_{s=0}^{k-1}\gamma(s)\boldsymbol{\eta}(s),\quad \boldsymbol{\eta}(s)\in\mathbb{M} \tag{21}$$

where

$$\mathbb{M}=\left\{R(\theta)^\ell\boldsymbol{d}_{p,q}(0)(\ell=0,\ldots,2n-1;\ p,q=1,\ldots,m)\right\}. \tag{22}$$

Note that the second term in the right-hand side of (21) is ignored when $k=0$.

Proof: By this lemma, it is obvious that (21) holds when $k=0$. From Lemma 1, when $k\geqq 1$ we have

$$\boldsymbol{x}^\star(k)=\boldsymbol{x}^\star(0)+\sum_{s=0}^{k-1}\gamma(s)R(\theta)^{s+1}\boldsymbol{d}_{i_{s+1}^\star,i_s^\star}(0).$$

Moreover, since from condition 1-1) for any $i,j\in\{1,\ldots,m\}$, and for any $u\in\mathbb{Z}_{\geqq 0}$

$$R(\theta)^u\boldsymbol{d}_{i,j}(0)\in\mathbb{M}$$

holds; hence, (21) holds. ■

The next lemma shows the boundedness of the spiral vectors $r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k)$ of (18).

Lemma 3: In Algorithm 2, there is a positive constant $\overline{d}$ such that

$$\left\|r(k)R(\theta)\boldsymbol{d}_{i,i_k^\star}(k)\right\|\leqq\gamma(k)\overline{d}\ \ (i=1,\ldots,m). \tag{23}$$

Proof: For any difference vectors $\boldsymbol{d}_{i,j}(k)=\boldsymbol{x}_i(k)-\boldsymbol{x}_j(k)\ (i,j=1,\ldots,m)$

$$r(k)R(\theta)\boldsymbol{d}_{i,j}(k)=\gamma(k)R(\theta)^{k+1}\boldsymbol{d}_{i,j}(0)\ \ (i,j=1,\ldots,m)$$

hold from (5) and their norm become

$$\left\|r(k)R(\theta)\boldsymbol{d}_{i,j}(k)\right\|=\gamma(k)\left\|\boldsymbol{d}_{i,j}(0)\right\|\ \ (i,j=1,\ldots,m).$$

Hence, defining

$$\overline{d}=\max\left\{\left\|\boldsymbol{d}_{i,j}(0)\right\|\ \ (i,j=1,\ldots,m)\right\}$$

we obtain the inequality (23). ■

The following lemma shows the behavior of $\gamma(k)$ of (19) that determines the sizes of the spiral vectors.

Lemma 4: Suppose $\mathbb{L}(\boldsymbol{x}^\star(0))$ is compact. For Algorithm 2, $\gamma(k)$ of (19) satisfies

$$\liminf_{k\to+\infty}\gamma(k)=0. \tag{24}$$

Proof: The proof is conducted by contradiction. First, suppose that there exists $\underline{\gamma}>0$ for which $\gamma(k)\geqq\underline{\gamma}$. Therefore, from its definition, the minimum of $\gamma(k)$ is $\gamma_{\min}:=\min_{0\leq k<\infty}\{\gamma(k)\}\geqq\underline{\gamma}$. Therefore, we can define $k_{\min}=\min\{k\ |\gamma(k)=\gamma_{\min}\}$ and hold

$$\gamma_{\min}=\gamma(k)=\gamma(k+1)=\cdots\left(k\geqq k_{\min}\right). \tag{25}$$

According to step 2, this means that the center is updated infinitely at least once every $2n$ iterations.

The following shows a contradiction of the above assumption. From Lemma 2, the center $\boldsymbol{x}^\star(k)=\boldsymbol{x}_{i_k^\star}(k)$ is expressed as follows:

$$\boldsymbol{x}^\star(k)=\boldsymbol{x}^\star(0)+\sum_{s=0}^{k-1}\gamma(s)\boldsymbol{\eta}(s),\ \ \boldsymbol{\eta}(s)\in\mathbb{M}$$

where $\mathbb{M}$ is a finite vector set from (22). Thus, from (25), for all $k>k_{\min}$, we have

$$\boldsymbol{x}^\star(k)=\underline{\boldsymbol{x}}+\gamma_{\min}\sum_{s=k_{\min}}^{k-1}\boldsymbol{\eta}(s),\quad\boldsymbol{\eta}(s)\in\mathbb{M} \tag{26}$$

where $\underline{\boldsymbol{x}}$ is a constant vector defined as

$$\underline{\boldsymbol{x}}=\boldsymbol{x}^\star(0)+\sum_{s=0}^{k_{\min}-1}\gamma(s)\boldsymbol{\eta}(s).$$

In (26), since $\mathbb{M}$ is a finite vector set, we can represent $\mathbb{M}$ as $\mathbb{M}=\{z_1,\ldots,z_P\}$ where $z_i\in\mathbb{R}^n$, $P\in\mathbb{N}$. Furthermore, due to $\boldsymbol{\eta}(s)\in\mathbb{M}$ in (26) and by defining the infinite discrete vector set $\hat{\mathbb{T}}:=\{\sum_{i=1}^{P}a_iz_i|a_i\in\mathbb{Z}\ \ (i=1,\ldots,P)\}$, we obtain the relation $\sum_{s=k_{\min}}^{k-1}\boldsymbol{\eta}(s)\in\hat{\mathbb{T}}$. Thus, defining the infinite discrete set $\mathbb{T}:=\underline{\boldsymbol{x}}+\gamma_{\min}\hat{\mathbb{T}}$, which is translated from $\hat{\mathbb{T}}$ by the constant vector $\underline{\boldsymbol{x}}$ and the constant scalar $\gamma_{\min}$ in (26), we understand that $\boldsymbol{x}^\star(k)$ of (26) stays in the translated infinite discrete set $\mathbb{T}$ for $k>k_{\min}$, i.e., $\boldsymbol{x}^\star(k)\in\mathbb{T}\ (k>k_{\min})$.

From the definition of Algorithm 2, $\boldsymbol{x}^\star(k)\in\mathbb{L}(\boldsymbol{x}^\star(0))$ is satisfied. Therefore, we have $\boldsymbol{x}^\star(k)\in\mathbb{L}(\boldsymbol{x}^\star(0))\cap\mathbb{T}(k>k_{\min})$. Since $\mathbb{L}(\boldsymbol{x}^\star(0))$ is compact, $\mathbb{L}(\boldsymbol{x}^\star(0))\cap\mathbb{T}$ becomes a finite discrete set, and it is impossible for Algorithm 2 to continue updating the center $\boldsymbol{x}^\star(k)\in\mathbb{L}(\boldsymbol{x}^\star(0))\cap\mathbb{T}$ indefinitely. Thus, from step 2, $\gamma(k)\to 0$ is confirmed. However, this contradicts the assumption $\gamma(k)\geqq\underline{\gamma}>0$. ■

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

V. Proof of Convergence

In this section, using Theorem 1 and based on the analysis provided in Section IV, we prove Theorem 2 that Algorithm 2 guarantees $\liminf_{k\to+\infty} \|\nabla f(\boldsymbol{x}^\star(k))\| = 0$. This condition is finally proved by a contradiction regarding the step rate. The contradiction technique is seen in [19].

We first define the worst initial search point by

$$\boldsymbol{x}_\star(0) = \boldsymbol{x}_{i_{\rm w}}(0),\;\; i_{\rm w} = \max\left\{\operatorname*{argmax}_{i=1,\dots,m}\{f(\boldsymbol{x}_i(0))\}\right\}$$

that exists from definition of Algorithm 2.

The following proposition indicates the relation between $\gamma(k)$ and the updating center $\boldsymbol{x}^\star(k)$ within a finite number of iterations in Algorithm 2.

Proposition 1: Suppose that $\mathbb{L}(\boldsymbol{x}_\star(0))$ is compact, and $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on a neighborhood of $\mathbb{L}(\boldsymbol{x}_\star(0))$. For $\varepsilon > 0$, $\mathbb{E}(\varepsilon)$ is defined as follows:

$$\mathbb{E}(\varepsilon) = \left\{\boldsymbol{x} \in \mathbb{L}(\boldsymbol{x}_\star(0)) \;\middle|\operatorname{dist}\left(\boldsymbol{x}, \mathbb{X}_\sharp\right) \geqq \varepsilon\right\}.$$

There exists a constant $\delta > 0$ such that if $\boldsymbol{x}^\star(k_0) \in \mathbb{E}(\varepsilon)$ and $\gamma(k_0) < \delta$ for any iteration k_0 then Algorithm 2 will attain $f(\boldsymbol{x}^\star(k_0 + q + 1)) < f(\boldsymbol{x}^\star(k_0))$ $\;(\exists q \in \{0, 1, \dots, 2n - 1\})$.

Proof: To prove this, we take the following two steps.

Step 1: From the definition of Algorithm 2 and the assumptions, $\mathbb{C}(\boldsymbol{x}_\star(0)) = \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) = f(\boldsymbol{x}_\star(0))\}$ and $\mathbb{L}(\boldsymbol{x}^\star(0)) = \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) \leqq f(\boldsymbol{x}^\star(0))\}$ are compact and disjoint, and there exists $\eta := \operatorname{dist}(\mathbb{L}(\boldsymbol{x}^\star(0)), \mathbb{C}(\boldsymbol{x}_\star(0))) > 0$. Thus, if $\gamma(k_0) < (\eta/2\overline{d})$ where $\overline{d}$ is in Lemma 3, we have

$$\left\|\boldsymbol{x}_i(k_0 + 1) - \boldsymbol{x}^\star(k_0)\right\| < \eta/2$$

and $\boldsymbol{x}_i(k_0 + 1) \in \mathbb{L}(\boldsymbol{x}_\star(0))$.

We can also define $\alpha = \min_{\boldsymbol{x}\in\mathbb{E}(\varepsilon)} \|\nabla f(\boldsymbol{x})\| > 0$ from the assumptions. Here, since f is continuously differentiable on a neighborhood of $\mathbb{L}(\boldsymbol{x}_\star(0))$, ∇f is continuous on $\mathbb{L}(\boldsymbol{x}_\star(0))$. Furthermore, since $\mathbb{L}(\boldsymbol{x}_\star(0))$ is compact, ∇f is uniformly continuous on $\mathbb{L}(\boldsymbol{x}_\star(0))$. Thus, for $(\sigma\alpha/2) > 0$ where σ is from Theorem 1, there exists $\tau > 0$ such that if $\|\boldsymbol{y} - \boldsymbol{x}^\star(k_0)\| < \tau$ $\;(\boldsymbol{y} \in \mathbb{L}(\boldsymbol{x}_\star(0)))$, then

$$\left\|\nabla f(\boldsymbol{y}) - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\| < \frac{\sigma\alpha}{2}.$$

Based on the above analysis, defining $\delta = (1/\overline{d}) \min\{(\eta/2),\; \tau\}$, if

$$\gamma(k_0) < \delta \tag{27}$$

holds for any iteration k_0, we have

$$\boldsymbol{x}_i(k_0 + 1) \in \mathbb{L}(\boldsymbol{x}_\star(0)) \tag{28}$$
$$\left\|\boldsymbol{x}_i(k_0 + 1) - \boldsymbol{x}^\star(k_0)\right\| < \eta/2 \tag{29}$$

and

$$\left\|\nabla f(\boldsymbol{x}_i(k_0 + 1)) - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\| < \frac{\sigma\alpha}{2}\;\; (i = 1, \dots, m). \tag{30}$$

Step 2: This step focuses on the two points $\boldsymbol{x}_{i^\star_{k_0}}(k_0) = \boldsymbol{x}^\star(k_0)$ and $\boldsymbol{x}_p(k_0)$ satisfying the inequality (10) at any iteration k_0 in Theorem 1, and assumes that the center $\boldsymbol{x}^\star(k_0)$ is not updated until $k_0 + \overline{q}$ where $\overline{q}$ is of Theorem 1, that is

$$\boldsymbol{x}^\star(k_0) = \cdots = \boldsymbol{x}^\star(k_0 + \overline{q}). \tag{31}$$

Under this assumption, we have

$$\boldsymbol{x}_p(k_0 + \overline{q} + 1) = \boldsymbol{x}^\star(k_0) + r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q}) \tag{32}$$

from (17).

From (28), (29), and (31), $\boldsymbol{x}_p(k_0+\overline{q}+1) \in \mathbb{B}(\boldsymbol{x}^\star(k_0), \eta/2) \subset \mathbb{L}(\boldsymbol{x}_\star(0))$ is derived, and by applying the mean value theorem in Appendix A to $\boldsymbol{x}_p(k_0 + \overline{q} + 1)$ and $\boldsymbol{x}^\star(k_0)$, there exists a $\lambda \in (0, 1)$ for which $f(\boldsymbol{x}_p(k_0 + \overline{q} + 1)) - f(\boldsymbol{x}^\star(k_0)) = \nabla f(\boldsymbol{w})^\top(\boldsymbol{x}_p(k_0 + \overline{q} + 1) - \boldsymbol{x}^\star(k_0))$ is satisfied considering $\boldsymbol{w} = \lambda\boldsymbol{x}_p(k_0 + \overline{q} + 1) + (1 - \lambda)\boldsymbol{x}^\star(k_0) \in \mathbb{B}(\boldsymbol{x}^\star(k_0), \eta/2)$. Given that, we have the following:

$$\begin{aligned}
&f\left(\boldsymbol{x}_p(k_0 + \overline{q} + 1)\right) - f\left(\boldsymbol{x}^\star(k_0)\right)\\
&\quad= \nabla f\left(\boldsymbol{x}^\star(k_0)\right)^\top\left(\boldsymbol{x}_p(k + \overline{q} + 1) - \boldsymbol{x}^\star(k_0)\right)\\
&\qquad+ \left(\nabla f(\boldsymbol{w})^\top - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)^\top\right)\left(\boldsymbol{x}_p(k_0 + \overline{q} + 1) - \boldsymbol{x}^\star(k_0)\right)\\
&\quad= \nabla f\left(\boldsymbol{x}^\star(k_0)\right)^\top r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\\
&\qquad+ \left(\nabla f(\boldsymbol{w})^\top - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)^\top\right)r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q}).
\end{aligned} \tag{33}$$

Here, from (10), we have the following relation for the first term in the right-hand side of (33):

$$\begin{aligned}
&\nabla f\left(\boldsymbol{x}^\star(k_0)\right)^\top r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\\
&\quad\leqq -\sigma r(k_0 + \overline{q})\left\|\nabla f(\boldsymbol{x}^\star(k_0))\right\|\left\|R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\right\|.
\end{aligned} \tag{34}$$

Furthermore, based on the Cauchy–Schwarz inequality (Appendix A) for the second term on the right-hand side of (33), we have

$$\begin{aligned}
&\left|\left(\nabla f(\boldsymbol{w})^\top - \nabla f(\boldsymbol{x}^\star(k_0))^\top\right)r(k_0 + \overline{q})R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\right|\\
&\quad\leqq r(k_0 + \overline{q})\left\|\nabla f(\boldsymbol{w}) - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\|\left\|R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\right\|.
\end{aligned} \tag{35}$$

Thus, by combining (34) and (35), (33) is expressed as

$$\begin{aligned}
&f\left(\boldsymbol{x}_p(k_0 + \overline{q} + 1)\right) - f\left(\boldsymbol{x}^\star(k_0)\right)\\
&\quad\leqq -\sigma r(k_0 + \overline{q})\left\|\nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\|\left\|R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\right\|\\
&\qquad+ (k_0 + \overline{q})\left\|\nabla f(\boldsymbol{w}) - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\|\left\|R(\theta)\boldsymbol{d}_{p,i^\star_{k_0}}(k_0 + \overline{q})\right\|\\
&\quad= \left(-\sigma\left\|\nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\| + \left\|\nabla f(\boldsymbol{w}) - \nabla f\left(\boldsymbol{x}^\star(k_0)\right)\right\|\right)\\
&\qquad\times\left\|\boldsymbol{x}_p(k_0 + \overline{q} + 1) - \boldsymbol{x}^\star(k_0)\right\|.
\end{aligned} \tag{36}$$

The relation $\|\nabla f(\boldsymbol{w}) - \nabla f(\boldsymbol{x}^\star(k_0))\| < (\sigma\alpha/2)$ is satisfied from $\|\boldsymbol{w} - \boldsymbol{x}^\star(k_0)\| < \|\boldsymbol{x}_p(k_0 + \overline{q} + 1) - \boldsymbol{x}^\star(k_0)\|$, (30), and (31), and the inequality (36) becomes

$$\begin{aligned} & f\big(\boldsymbol{x}_p(k_0+\overline{q}+1)\big)-f\big(\boldsymbol{x}^\star(k_0)\big)\\ &\quad < \Big(-\sigma\big\|\nabla f\big(\boldsymbol{x}^\star(k_0)\big)\big\|+\frac{\sigma\alpha}{2}\Big)\big\|\boldsymbol{x}_p(k_0+\overline{q}+1)-\boldsymbol{x}^\star(k_0)\big\|\\ &\quad < \Big(-\sigma\big\|\nabla f\big(\boldsymbol{x}^\star(k_0)\big)\big\|+\frac{\sigma}{2}\big\|\nabla f\big(\boldsymbol{x}^\star(k_0)\big)\big\|\Big)\\ &\qquad \times \big\|\boldsymbol{x}_p(k_0+\overline{q}+1)-\boldsymbol{x}^\star(k_0)\big\|\\ &\quad = -\frac{\sigma}{2}\big\|\nabla f\big(\boldsymbol{x}^\star(k_0)\big)\big\|\big\|\boldsymbol{x}_p(k_0+\overline{q}+1)-\boldsymbol{x}^\star(k_0)\big\|<0 \end{aligned} \tag{37}$$

that is $f(\boldsymbol{x}_p(k_0+\overline{q}+1)) < f(\boldsymbol{x}^\star(k_0))$. Considering that this inequality is satisfied under the assumption (31), we have proven that Algorithm 2 gives

$$f(\boldsymbol{x}^\star(k_0+q+1)) < f\big(\boldsymbol{x}^\star(k_0)\big)$$

where $q \in \{0, 1, \dots, \overline{q}\} \subseteq \{0, 1, \dots, 2n\}$. ■

In brief, Proposition 1 suggests that if $\gamma(k)$ is sufficiently small, Algorithm 2 continues to update the center within $2n$ iterations. Using this proposition, we understand the following lemma used for proof of Theorem 2.

Lemma 5: Assume that $\mathbb{L}(\boldsymbol{x}_\star(0))$ is compact, and $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable on a neighborhood of $\mathbb{L}(\boldsymbol{x}_\star(0))$. In addition, suppose that $\liminf_{k\to+\infty} \|\nabla f(\boldsymbol{x}^\star(k))\| \neq 0$. There exists a constant $\underline{\gamma}$ for which over all k, $\gamma(k) > \underline{\gamma}$ in Algorithm 2 is satisfied.

Proof: Since $\liminf_{k\to+\infty} \|\nabla f(\boldsymbol{x}^\star(k))\| \neq 0$, there exists $\varepsilon > 0$ such that for all k

$$\boldsymbol{x}^\star(k) \in \mathbb{E}(\varepsilon) = \big\{\boldsymbol{x} \in \mathbb{L}(\boldsymbol{x}_\star(0)) \,\big|\, \mathrm{dist}\big(\boldsymbol{x}, \mathbb{X}_\sharp\big) \geqq \varepsilon\big\}. \tag{38}$$

From Proposition 1, there exists a constant δ for which if $\gamma(k) = \prod_{i=0}^{k} r(i) < \delta$ then Algorithm 2 will continue to update $\boldsymbol{x}^\star(k)$ within $2n$ iterations. Thus, based on the adjusting rule for $r(k)$ in step 2, defining $\underline{\gamma} = h^{\bar{\rho}+2n+1}$ with $\bar{\rho} := \min\{\rho \mid h^\rho \leqq \delta\}$, we have $\gamma(k) > \underline{\gamma}$ for all k. ■

Finally, using contradiction, we prove Theorem 2 from Section III-D.

Proof of Theorem 2: Suppose that $\liminf_{k\to+\infty} \|\nabla f(\boldsymbol{x}^\star(k))\| \neq 0$. From Lemma 5, there exists $\underline{\gamma}$ for which $\gamma(k) > \underline{\gamma}$. However, this contradicts Lemma 4. ■

VI. Settings and Verifications

In this section, we show specific setting methods to realize conditions 1–3 and verify the convergence of Algorithm 2 with the proposed setting methods using numerical experiments.

A. Setting Methods

To execute Algorithm 2, we must show specific setting methods to satisfy conditions 1–3. Conditions 1 and 2 are abstract, and their setting methods are as follows.

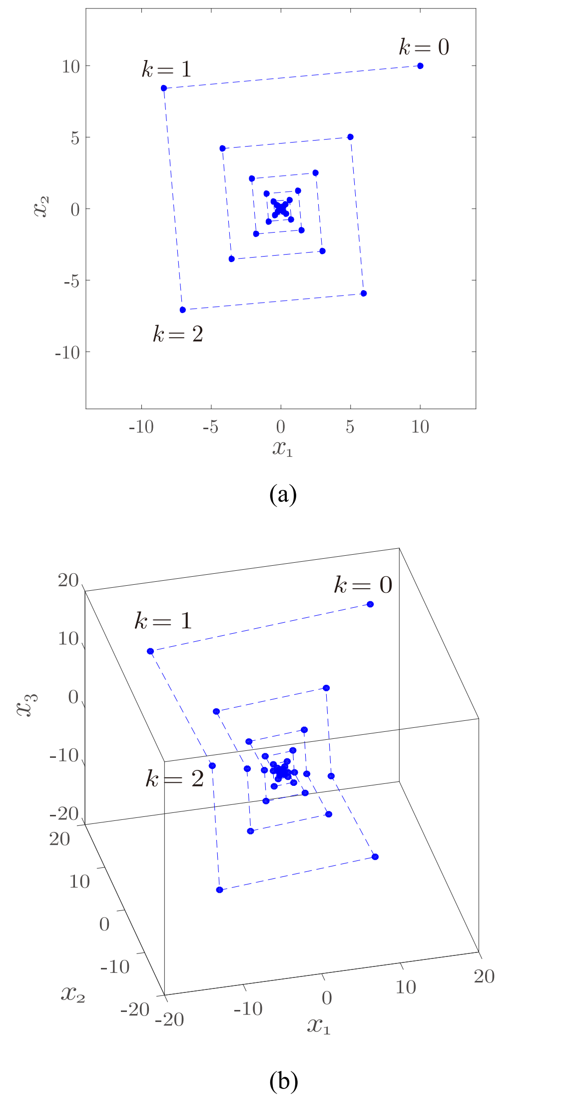


Fig. 5. Examples of trajectories of (1) with the proposed settings (39)–(42). (a) $n = 2$, $\boldsymbol{x}(0) = [10\ 10]^\top$, $\boldsymbol{x}^\star = [0\ 0]^\top$, $\omega = 1/2$. (b) $n = 3$, $\boldsymbol{x}(0) = [15\ 15\ 15]^\top$, $\boldsymbol{x}^\star = [0\ 0\ 0]^\top$, $\omega = 1/2$.

1) Setting for Condition 1: Set $R(\theta)$ as follows [23]:

$$\begin{cases} \theta = (-1)^n \pi/2 \\ R(\theta) = (-1)^n \prod_{i=1}^{n-1} R_{i,i+1}(\theta). \end{cases} \tag{39}$$

For any n ($\geqq 2$), by calculating $R(\theta)$ in (39), we can easily prove that

$$R(\theta) = \begin{bmatrix} \boldsymbol{0}_{n-1}^\top & -1 \\ I_{n-1} & \boldsymbol{0}_{n-1} \end{bmatrix} \tag{40}$$

where $\boldsymbol{0}_{n-1}$ is the $(n-1)\times 1$ zero vector. Then, calculating $R^0(\theta), \dots, R^{2n+1}(\theta)$, we can confirm that this $R(\theta)$ satisfies condition 1.

2) Setting for Condition 2: Set initial points $\boldsymbol{x}_i(0) \in \mathbb{R}^n$ ($i = 1, \dots, m$) as follows:

$$\begin{aligned} \min_{i=1,\dots,m} \Big\{ \max_{j=1,\dots,m} \big\{ \mathrm{rank}\big[& \boldsymbol{d}_{j,i}(0)\ R(\theta)\boldsymbol{d}_{j,i}(0) \\ & \dots\ R(\theta)^{2n-1}\boldsymbol{d}_{j,i}(0) \big]\big\}\Big\} = n \end{aligned} \tag{41}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
INITIAL SEARCH POINTS $x_{i,j}(0)$ FOR THE gTH TRIAL

Trial g	Search Point i	Vector Element j: 1	2	3	4	5	6	7	8	9	10
1	1	-4.22004	-4.43306	1.46808	1.00802	-4.79016	3.82103	-4.24861	4.87139	-3.15169	-1.45231
	2	2.23931	-4.37168	0.02599	2.89043	-0.12748	2.08955	1.69252	4.12032	-4.07013	-3.25452
	3	4.48789	-0.8699	-0.06337	-0.88402	3.73448	4.27544	-1.60755	-1.00921	-4.70991	1.04058
	4	0.94101	-4.64321	2.3607	-3.77581	-1.63237	-4.99407	-1.03908	2.42633	3.45057	0.24583
	5	0.88299	1.33911	-2.72838	-1.79162	1.84316	4.44124	1.16128	0.37587	0.12825	2.57706
2	1	1.84506	2.6751	4.73598	-3.68541	-1.14586	0.84816	2.2787	-4.38481	-3.49128	-1.62975
	2	-0.45352	-1.30991	2.66749	-0.32414	-2.56523	-1.28674	0.28078	-2.67346	3.12286	2.52816
	3	2.95469	3.75256	-4.26768	-4.39062	4.77511	2.93761	0.14838	-4.52012	-0.21258	2.21409
	4	-3.42863	2.89575	-3.87142	-2.31188	4.20285	-1.61126	-3.35253	2.0309	4.0657	3.72703
	5	2.5761	-1.5734	-4.74408	4.91001	0.12443	0.89526	-4.908	4.09556	0.64274	-2.8211
3	1	1.63017	1.14126	-1.09286	1.05019	-3.74824	4.12974	1.73001	-1.56196	-3.21603	-4.43466
	2	0.68537	4.76908	-2.03722	-3.96213	-2.50636	-2.04892	3.97158	4.7126	3.42971	-2.87723
	3	3.01578	-1.37534	3.94389	-2.05455	0.59397	4.01866	2.76394	0.59928	-3.41615	-2.66455
	4	-4.43384	-1.06257	-1.93506	4.63912	1.50658	1.7338	4.25401	0.95553	4.97053	-0.16168
	5	-4.40388	1.72287	0.63935	1.81715	4.45607	-2.31744	-4.49973	0.43554	2.80225	-0.79473
4	1	1.6479	3.44814	2.5669	-0.87844	4.49979	0.72489	4.55151	3.56562	3.23939	-2.18392
	2	-1.27779	4.30796	3.87753	-2.76341	4.26695	0.5487	-1.51707	0.05733	-4.40827	0.93141
	3	1.81967	3.14818	1.64717	-1.30581	-1.0006	-1.12983	0.29084	-2.58613	3.19782	0.39089
	4	-3.09689	-3.43731	2.5373	4.07278	-4.31418	1.56112	-0.92396	3.05473	3.72127	-4.61191
	5	-3.31152	1.74387	4.89018	-2.28265	2.38338	-2.878	2.75173	4.55063	-0.85288	-0.13902
5	1	-3.13258	0.20195	2.79648	-4.6382	3.29929	4.30458	4.37003	2.50365	0.79894	0.02452
	2	-4.23178	-3.77876	2.27884	1.95158	4.66224	-3.81355	0.78115	-3.89078	-1.65081	3.3614
	3	-0.60946	1.11138	2.25576	-2.52609	-1.23698	-1.72363	-1.47535	-3.45593	4.69454	4.18173
	4	-0.40249	0.12108	4.57523	1.54462	0.87331	0.85748	3.2477	-0.23381	-3.42576	-4.59523
	5	3.41856	3.14179	-1.6973	4.85113	-3.60978	-2.56759	-0.94027	-2.72541	-4.25147	-2.37901
6	1	-1.31907	-3.86803	-0.05923	2.93208	2.41991	3.96197	1.97953	-2.61234	-1.42451	-1.93937
	2	-0.28293	-0.66762	-0.3186	-4.34979	-3.21456	4.65295	-2.20819	0.24051	-2.30065	-0.49476
	3	-1.69803	-3.41015	-3.3815	3.87136	2.51475	-3.69704	-0.98382	3.22203	2.38737	0.53401
	4	-2.14347	-3.6898	-0.71316	-2.16964	4.35646	2.45871	-3.18264	-3.21111	1.12555	-4.92236
	5	2.57772	-1.81321	-3.0299	1.46959	-0.88743	-1.33276	-4.98625	-2.15884	2.46282	-2.51048
7	1	-2.0649	0.9456	3.89398	2.08561	3.92843	-0.4814	-3.80219	-4.35499	0.03965	-1.30082
	2	1.06169	-3.96673	0.14817	-3.65338	-0.35407	2.94546	-0.44125	-1.05539	3.27228	-0.10009
	3	4.59393	4.11128	-4.97441	-1.08289	-0.81867	2.48444	3.59686	-2.2672	0.03017	-0.19288
	4	-3.11895	-4.79801	1.667	-0.95462	1.92893	-4.12266	1.20622	-4.52212	0.27421	0.13021
	5	1.93454	0.40774	-4.65408	4.96244	-3.38697	-3.28976	-3.47391	-0.23829	0.16209	-4.51463
8	1	1.49343	-2.48937	-4.65487	-2.55912	-3.31956	-0.94212	-0.13459	0.24321	-2.92539	-1.39678
	2	-0.84633	-3.9663	-2.82072	-3.46988	3.82154	0.52098	-3.21315	-0.73236	-3.47935	0.73818
	3	4.40665	-2.03552	-1.62127	-0.63478	1.59496	3.1514	-2.27733	-4.4645	2.64687	-3.8193
	4	-2.88758	-0.64821	4.1807	-2.74474	-3.70378	4.38143	-4.57704	2.09387	-1.95796	-0.98915
	5	-4.85583	-2.66909	-0.35843	-3.62417	3.53123	2.32265	4.62461	-1.81492	0.90107	-2.93533
9	1	-0.07496	3.87327	-1.413	0.92949	1.67343	3.21905	4.68846	-4.55533	2.1361	2.11885
	2	-4.161	4.34057	-1.84929	-3.36859	0.51618	4.64091	0.85275	2.69782	-4.42891	3.33339
	3	0.83135	0.36578	-1.68472	2.42372	0.7287	-0.83987	-2.33554	0.51815	1.31746	3.87312
	4	1.7595	-0.30987	4.01416	-3.95603	-2.48469	3.46232	2.23627	-3.91002	-1.8535	3.15499
	5	2.30496	1.28238	3.315	-2.89268	4.2502	3.77796	-4.43021	1.45471	-4.32015	-1.34752
10	1	-0.90153	-3.98499	3.49113	1.84347	-1.92838	2.17965	-0.74066	-1.40371	-4.90353	3.30056
	2	0.71006	0.31543	-1.63666	1.94878	-4.49027	3.97307	3.2856	-4.1772	2.27552	3.37359
	3	1.68102	3.04838	0.09269	-0.32657	2.07032	-1.97444	4.03643	3.12816	3.40157	-1.48234
	4	-2.98205	-4.6214	-3.64667	0.33337	-4.90806	2.47003	-2.43483	0.32496	4.64854	-3.47824
	5	-3.52899	-1.12417	-4.70721	0.14194	-4.50502	3.28207	4.23723	2.92891	3.89192	0.50147

where $\boldsymbol{d}_{j,i}(0) = \boldsymbol{x}_j(0) - \boldsymbol{x}_i(0)$. This result is identical to condition 2 and its validity can be easily checked by software like MATLAB. Note that the existence of the initial points satisfying condition 2 has been proven [23]. In addition, we can confirm numerically that this condition is fully satisfied even when they are placed randomly.

Although condition 3 is easily satisfied, we propose a practical setting for h in condition 3.

3) Setting for Condition 3: Set the rate $h \in (0, 1)$ as follows for condition 3:

$$h = \sqrt[2n]{\omega}, \ \ \omega \in (0, 1). \tag{42}$$

Since ω and h are invertible, this condition is satisfied without loss of generality. The importance of this setting emanates from the determination of h based on $2n$ iterations, which is an important measure for the convergence from Theorem 1. This is practical because it helps the users to set h to prevent having a very small h^{2n} for the computers; hence, it can handle the algorithm without causing more effect of round error or taking very longer time for the convergence.

TABLE II
SETTINGS OF $R(\theta)$ AND ω USED FOR NUMERICAL VERIFICATION

ω	$1/2$, $1/5$
$R(\theta)$	Eq.(40) with $n = 5$, 10

TABLE III
CHECK RESULTS OF INITIAL SEARCH POINTS CONDITION (41)

Settings of m and n		Trial Number g: 1	2	3	4	5	6	7	8	9	10
$n = 5$	$m = 2$	5	5	5	5	5	5	5	5	5	5
	$m = 5$	5	5	5	5	5	5	5	5	5	5
$n = 10$	$m = 2$	10	10	10	10	10	10	10	10	10	10
	$m = 5$	10	10	10	10	10	10	10	10	10	10

Generally, analyzing a specific guide line of setting $\omega \in (0, 1)$ is difficult because it depends on complex interaction among the objective function structure, the dimension, the initial search points setting, and the number of search points. However, considering the purpose of this practical device (42), we can provide an abstract guideline that is not to set $\omega \in (0, 1)$ too small.

TABLE IV
RESULTS FOR THE SPHERE FUNCTION

Setup n	Setup m	Setup ω	Checked Items	Trial Number										10 Trials' Mean
				1	2	3	4	5	6	7	8	9	10	
$n=5$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.41.E-07	8.87.E-08	2.46.E-07	2.25.E-07	2.31.E-07	2.50.E-07	2.22.E-07	1.26.E-07	2.40.E-07	1.80.E-07	1.95.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	7.52.E-04	5.96.E-04	9.93.E-04	9.48.E-04	9.61.E-04	9.99.E-04	9.43.E-04	7.10.E-04	9.79.E-04	8.48.E-04	8.73.E-04
			k_{fin}	500	488	2773	1209	435	3081	835	496	576	613	1100.6
			Function call times	1000	976	5546	2418	870	6162	1670	992	1152	1226	2201.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.02.E-07	2.30.E-07	2.50.E-07	2.45.E-07	2.11.E-07	2.50.E-07	2.10.E-07	1.79.E-07	2.45.E-07	1.86.E-07	2.21.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.00.E-04	9.59.E-04	9.99.E-04	9.90.E-04	9.19.E-04	1.00.E-03	9.17.E-04	8.46.E-04	9.90.E-04	8.63.E-04	9.38.E-04
			k_{fin}	471	458	2405	1743	450	2537	497	370	507	356	979.4
			Function call times	942	916	4810	3486	900	5074	994	740	1014	712	1958.8
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.55.E-07	1.88.E-07	1.77.E-07	1.96.E-07	2.18.E-07	1.54.E-07	1.89.E-07	2.06.E-07	1.59.E-07	1.98.E-07	1.84.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	7.87.E-04	8.66.E-04	8.40.E-04	8.85.E-04	9.33.E-04	7.84.E-04	8.70.E-04	9.09.E-04	7.97.E-04	8.91.E-04	8.56.E-04
			k_{fin}	559	714	625	575	350	462	376	506	390	368	492.5
			Function call times	2795	3570	3125	2875	1750	2310	1880	2530	1950	1840	2462.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.61.E-07	1.47.E-07	2.38.E-07	1.38.E-07	9.42.E-08	8.48.E-08	2.04.E-07	1.83.E-07	2.19.E-07	2.23.E-07	1.69.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.02.E-04	7.66.E-04	9.75.E-04	7.43.E-04	6.14.E-04	5.83.E-04	9.03.E-04	8.57.E-04	9.36.E-04	9.43.E-04	8.12.E-04
			k_{fin}	446	382	295	323	283	312	385	310	359	324	341.9
			Function call times	2230	1910	1475	1615	1415	1560	1925	1550	1795	1620	1709.5
$n=10$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.32.E-07	2.37.E-07	2.06.E-07	2.01.E-07	2.18.E-07	2.46.E-07	2.45.E-07	1.83.E-07	1.98.E-07	2.35.E-07	2.20.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.64.E-04	9.73.E-04	9.08.E-04	8.97.E-04	9.33.E-04	9.92.E-04	9.89.E-04	8.55.E-04	8.89.E-04	9.70.E-04	9.37.E-04
			k_{fin}	1730	1767	1234	2242	1778	1981	2478	1473	1569	1749	1800.1
			Function call times	3460	3534	2468	4484	3556	3962	4956	2946	3138	3498	3600.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.43.E-07	2.50.E-07	1.86.E-07	2.16.E-07	1.61.E-07	1.82.E-07	2.49.E-07	1.85.E-07	2.37.E-07	1.65.E-07	2.07.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.87.E-04	9.99.E-04	8.64.E-04	9.29.E-04	8.03.E-04	8.54.E-04	9.99.E-04	8.60.E-04	9.73.E-04	8.11.E-04	9.08.E-04
			k_{fin}	1316	1927	1128	1371	1444	984	2209	1312	1850	1348	1488.9
			Function call times	2632	3854	2256	2742	2888	1968	4418	2624	3700	2696	2977.8
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.82.E-07	2.49.E-07	2.47.E-07	2.06.E-07	1.99.E-07	1.50.E-07	2.03.E-07	1.40.E-07	1.88.E-07	2.42.E-07	2.01.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.52.E-04	9.98.E-04	9.95.E-04	9.08.E-04	8.92.E-04	7.73.E-04	9.00.E-04	7.49.E-04	8.67.E-04	9.84.E-04	8.92.E-04
			k_{fin}	1723	1414	1520	1635	1656	1496	1344	1589	1339	1138	1485.4
			Function call times	8615	7070	7600	8175	8280	7480	6720	7945	6695	5690	7427
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.28.E-07	2.32.E-07	2.32.E-07	2.45.E-07	2.11.E-07	2.22.E-07	1.92.E-07	2.41.E-07	1.68.E-07	2.09.E-07	2.08.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	7.14.E-04	9.62.E-04	9.63.E-04	9.91.E-04	9.18.E-04	9.42.E-04	8.77.E-04	9.82.E-04	8.20.E-04	9.14.E-04	9.08.E-04
			k_{fin}	1392	1185	1254	1098	1021	983	1105	873	996	1127	1103.4
			Function call times	6960	5925	6270	5490	5105	4915	5525	4365	4980	5635	5517

TABLE V
RESULTS FOR THE SCHWEFEL FUNCTION

Setup n	Setup m	Setup ω	Checked Items	Trial Number										10 Trials' Mean
				1	2	3	4	5	6	7	8	9	10	
$n=5$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.48.E-07	8.34.E-07	8.52.E-07	7.13.E-07	8.01.E-07	8.89.E-07	4.56.E-07	2.23.E-07	7.94.E-07	2.39.E-07	5.95.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	7.72.E-04	9.94.E-04	9.95.E-04	9.64.E-04	9.98.E-04	9.99.E-04	9.96.E-04	8.39.E-04	9.63.E-04	7.59.E-04	9.28.E-04
			k_{fin}	619	471	9742	2984	3023	32756	3805	969	2456	590	5741.5
			Function call times	1238	942	19484	5968	6046	65512	7610	1938	4912	1180	11483.0
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	5.52.E-08	4.57.E-07	8.58.E-07	7.01.E-07	7.86.E-07	8.86.E-07	5.08.E-07	3.13.E-07	7.93.E-07	3.82.E-07	5.74.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	5.04.E-04	8.80.E-04	9.99.E-04	9.96.E-04	9.90.E-04	9.99.E-04	9.99.E-04	8.34.E-04	9.84.E-04	9.81.E-04	9.17.E-04
			k_{fin}	489	525	9687	2976	2503	35056	3790	1067	2361	476	5893
			Function call times	978	1050	19374	5952	5006	70112	7580	2134	4722	952	11786
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.73.E-07	6.14.E-08	5.83.E-08	3.50.E-07	1.64.E-07	3.93.E-07	5.12.E-08	4.80.E-07	3.09.E-07	2.82.E-07	2.32.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	6.04.E-04	5.21.E-04	5.82.E-04	9.29.E-04	6.64.E-04	8.35.E-04	5.78.E-04	8.90.E-04	9.84.E-04	8.65.E-04	7.45.E-04
			k_{fin}	494	518	422	469	433	378	475	662	348	341	454
			Function call times	2470	2590	2110	2345	2165	1890	2375	3310	1740	1705	2270
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	3.06.E-07	1.05.E-07	2.57.E-07	6.85.E-07	3.11.E-07	3.20.E-07	1.95.E-07	1.16.E-07	1.58.E-08	4.61.E-07	2.77.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.42.E-04	5.86.E-04	8.76.E-04	9.29.E-04	9.79.E-04	9.48.E-04	9.12.E-04	7.18.E-04	2.88.E-04	9.32.E-04	8.11.E-04
			k_{fin}	371	383	299	478	448	232	293	402	271	302	347.9
			Function call times	1855	1915	1495	2390	2240	1160	1465	2010	1355	1510	1739.5
$n=10$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	3.72.E-07	9.31.E-07	5.21.E-07	2.33.E-07	7.80.E-07	6.06.E-07	5.18.E-07	5.18.E-07	2.19.E-07	1.91.E-07	4.89.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.74.E-04	9.99.E-04	9.84.E-04	9.11.E-04	9.98.E-04	9.70.E-04	9.98.E-04	9.94.E-04	9.90.E-04	8.65.E-04	9.68.E-04
			k_{fin}	4522	31263	7909	4633	15312	3119	24996	2114	4382	1696	9994.6
			Function call times	9044	62526	15818	9266	30624	6238	49992	4228	8764	3392	19989.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	5.10.E-07	9.01.E-07	4.84.E-07	5.11.E-07	7.74.E-07	4.56.E-07	5.28.E-07	1.11.E-07	2.21.E-07	2.69.E-07	4.76.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.87.E-04	9.94.E-04	9.87.E-04	9.45.E-04	9.85.E-04	9.88.E-04	9.99.E-04	6.30.E-04	9.34.E-04	9.07.E-04	9.36.E-04
			k_{fin}	3953	25746	7109	4304	15033	3327	24900	2074	4702	1576	9272.4
			Function call times	7906	51492	14218	8608	30066	6654	49800	4148	9404	3152	18544.8
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	4.62.E-07	2.48.E-07	1.93.E-07	2.88.E-07	6.23.E-07	3.42.E-07	4.86.E-08	8.45.E-08	1.76.E-07	4.14.E-08	2.51.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.76.E-04	8.61.E-04	6.90.E-04	9.60.E-04	9.80.E-04	9.84.E-04	3.62.E-04	6.00.E-04	6.39.E-04	5.01.E-04	7.55.E-04
			k_{fin}	1654	1431	1494	1560	2064	1504	1400	1687	1336	1181	1531.1
			Function call times	8270	7155	7470	7800	10320	7520	7000	8435	6680	5905	7655.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.15.E-07	5.48.E-08	2.24.E-07	8.14.E-08	5.78.E-07	2.32.E-07	2.89.E-07	1.87.E-07	1.47.E-07	9.37.E-08	2.10.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.75.E-04	5.84.E-04	8.50.E-04	8.38.E-04	9.87.E-04	8.59.E-04	8.84.E-04	7.50.E-04	8.22.E-04	6.88.E-04	8.24.E-04
			k_{fin}	1013	1044	1144	1091	1618	1244	824	1007	1018	986	1098.9
			Function call times	5065	5220	5720	5455	8090	6220	4120	5035	5090	4930	5494.5

Now, with a proof of concept, we show spiral trajectories of (1) with the mentioned settings in Fig. 5. It should be noted that $r = h$ and, to make the setting of (41) satisfy the setting of the spiral model (1), we put $m = 2$, $\boldsymbol{x}_1(0) = \boldsymbol{x}(0)$, and $\boldsymbol{x}_2(0) = \boldsymbol{x}^\star$.

B. Numerical Verifications

1) Conditions: The test objective functions are the five types of continuously differentiable functions shown in Appendix B: 1) sphere function; 2) Schwefel function; 3) 2^n minima function; 4) levy function; and 5) translated sphere function. To check the convergence and characters of Algorithm 2 under different conditions, we set the number of search points to $m = 2$ and 5 and the number of dimensions to $n = 5$ and 10, where the termination criteria are $\|\nabla f(\boldsymbol{x}^\star(k))\| < 10^{-3}$ or $k = 10^8$.

For each condition, we run the algorithm ten times starting with different initial points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ placed at random in $[-5, 5]^n$. The resultant initial points $x_{i,j}(0)$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ for the gth $(g = 1, \ldots, 10)$ trial are shown in Table I where each number in the "Trial g" column corresponds to the gth trial number, each number in the "search point i" column represents the ith search point number, and each number in the "vector element j" row is the jth element number of each search point vector.

The composite rotation matrix $R(\theta)$ that satisfies condition 1 was set by (39), which results in (40), for each n, as shown in Table II. The step rate h that satisfies condition 3 was set by (42) with parameter ω taking the two patterns shown in Table II. We used (41) to check whether

TABLE VI
RESULTS FOR THE 2^n MINIMA FUNCTION

Setup n	Setup m	Setup ω	Checked Items	Trial Number 1	2	3	4	5	6	7	8	9	10	10 Trials' Mean
$n=5$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-3.07.E+02	-3.07.E+02	-3.63.E+02	-2.50.E+02	-3.07.E+02	-3.07.E+02	-3.35.E+02	-3.63.E+02	-3.35.E+02	-3.07.E+02	-3.18.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	7.39.E-04	7.82.E-04	9.84.E-04	9.71.E-04	8.68.E-04	9.99.E-04	8.62.E-04	7.53.E-04	9.46.E-04	7.84.E-04	8.69.E-04
			k_{fin}	775	768	3517	1697	518	2870	937	529	956	609	1317.6
			Function call times	1550	1536	7034	3394	1036	5740	1874	1058	1912	1218	2635.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-3.07.E+02	-3.35.E+02	-3.63.E+02	-2.50.E+02	-3.35.E+02	-3.63.E+02	-3.35.E+02	-3.63.E+02	-3.35.E+02	-3.35.E+02	-3.32.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	7.25.E-04	6.81.E-04	9.99.E-04	9.80.E-04	8.76.E-04	9.97.E-04	9.48.E-04	9.48.E-04	7.14.E-04	9.66.E-04	8.84.E-04
			k_{fin}	582	442	3650	2418	517	3914	621	426	785	462	1381.7
			Function call times	1164	884	7300	4836	1034	7828	1242	852	1570	924	2763.4
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-3.07.E+02	-3.07.E+02	-3.07.E+02	-3.92.E+02	-3.35.E+02	-3.35.E+02	-3.35.E+02	-3.63.E+02	-3.63.E+02	-3.63.E+02	-3.41.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	8.85.E-04	8.54.E-04	6.13.E-04	5.64.E-04	9.78.E-04	7.44.E-04	7.51.E-04	9.00.E-04	9.65.E-04	9.60.E-04	8.22.E-04
			k_{fin}	479	909	593	1074	737	576	651	735	603	529	688.6
			Function call times	2395	4545	2965	5370	3685	2880	3255	3675	3015	2645	3443
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-3.07.E+02	-3.07.E+02	-3.35.E+02	-3.63.E+02	-3.35.E+02	-3.35.E+02	-3.63.E+02	-3.63.E+02	-3.63.E+02	-3.35.E+02	-3.41.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.83.E-04	9.74.E-04	9.50.E-04	6.49.E-04	6.33.E-04	8.07.E-04	8.01.E-04	5.96.E-04	6.62.E-04	8.67.E-04	7.92.E-04
			k_{fin}	503	470	496	446	367	416	421	510	430	425	448.4
			Function call times	2515	2350	2480	2230	1835	2080	2105	2550	2150	2125	2242.0
$n=10$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-6.42.E+02	-6.42.E+02	-6.99.E+02	-6.14.E+02	-6.99.E+02	-6.14.E+02	-6.70.E+02	-6.99.E+02	-6.42.E+02	-6.14.E+02	-6.53.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	8.25.E-04	9.62.E-04	6.08.E-04	9.17.E-04	6.68.E-04	9.36.E-04	9.77.E-04	9.54.E-04	9.84.E-04	9.78.E-04	8.81.E-04
			k_{fin}	1708	2317	2053	2107	2109	2052	2932	1833	2385	2438	2193.4
			Function call times	3416	4634	4106	4214	4218	4104	5864	3666	4770	4876	4386.8
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-6.70.E+02	-7.27.E+02	-6.70.E+02	-5.57.E+02	-5.85.E+02	-6.14.E+02	-7.27.E+02	-6.99.E+02	-6.42.E+02	-6.99.E+02	-6.59.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.76.E-04	8.86.E-04	9.23.E-04	6.29.E-04	6.56.E-04	8.74.E-04	9.95.E-04	9.65.E-04	9.88.E-04	7.79.E-04	8.67.E-04
			k_{fin}	1625	1461	1448	1727	1773	1533	2974	1489	2320	1680	1803
			Function call times	3250	2922	2896	3454	3546	3066	5948	2978	4640	3360	3606
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-6.42.E+02	-6.42.E+02	-6.42.E+02	-6.99.E+02	-6.99.E+02	-6.70.E+02	-7.83.E+02	-6.70.E+02	-6.42.E+02	-5.85.E+02	-6.67.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	8.82.E-04	8.11.E-04	9.92.E-04	6.81.E-04	5.51.E-04	9.73.E-04	6.74.E-04	9.71.E-04	9.24.E-04	7.95.E-04	8.25.E-04
			k_{fin}	2124	2321	2236	1515	1998	1516	1558	1522	2234	1867	1889.1
			Function call times	10620	11605	11180	7575	9990	7580	7790	7610	11170	9335	9445.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	-6.70.E+02	-6.42.E+02	-6.14.E+02	-6.42.E+02	-6.99.E+02	-6.14.E+02	-6.14.E+02	-6.70.E+02	-6.70.E+02	-7.27.E+02	-6.56.E+02
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	8.96.E-04	9.23.E-04	8.51.E-04	9.10.E-04	9.48.E-04	8.47.E-04	9.34.E-04	8.24.E-04	9.28.E-04	9.80.E-04	9.04.E-04
			k_{fin}	1492	1539	1353	1367	1587	1418	1464	1525	1577	1286	1460.8
			Function call times	7460	7695	6765	6835	7935	7090	7320	7625	7885	6430	7304

TABLE VII
RESULTS FOR THE LEVY FUNCTION

Setup n	Setup m	Setup ω	Checked Items	Trial Number 1	2	3	4	5	6	7	8	9	10	10 Trials' Mean
$n=5$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	4.36.E+00	5.60.E+00	3.70.E-07	3.74.E+00	7.89.E+01	3.82.E-07	3.87.E-07	1.12.E+01	3.58.E-07	3.55.E-07	1.04.E+01
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.71.E-04	9.79.E-04	9.98.E-04	9.67.E-04	8.22.E-04	9.97.E-04	9.90.E-04	8.55.E-04	9.97.E-04	9.84.E-04	9.56.E-04
			k_{fin}	779	9251	84712	62271	4666	60976	7793	729	15624	10897	25769.8
			Function call times	1558	18502	169424	124542	9332	121952	15586	1458	31248	21794	51539.6
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	4.36.E+00	9.95.E+00	3.62.E-07	6.22.E-01	7.14.E+01	3.84.E-07	5.60.E+00	3.74.E-07	1.50.E+01	2.62.E-07	1.07.E+01
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	8.65.E-04	9.77.E-04	9.97.E-04	9.98.E-04	9.89.E-04	9.99.E-04	9.99.E-04	9.71.E-04	8.83.E-04	9.40.E-04	9.62.E-04
			k_{fin}	802	21342	73662	31781	6259	72071	64940	2584	1300	2435	27717.6
			Function call times	1604	42684	147324	63562	12518	144142	129880	5168	2600	4870	55435.2
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	6.22.E-01	2.49.E+00	1.24.E+00	2.39.E-07	1.24.E+00	2.49.E+00	6.22.E-01	7.42.E-08	1.41.E-07	6.52.E-09	8.71.E-01
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	6.62.E-04	9.91.E-04	7.43.E-04	8.67.E-04	6.66.E-04	5.63.E-04	7.43.E-04	6.91.E-04	9.64.E-04	1.61.E-04	7.05.E-04
			k_{fin}	1659	3762	1314	464	1226	579	1279	627	469	514	1189.3
			Function call times	8295	18810	6570	2320	6130	2895	6395	3135	2345	2570	5946.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	4.99.E+00	9.95.E+00	1.35.E-08	1.59.E-07	3.56.E-07	2.49.E+00	1.24.E+00	5.89.E+01	5.62.E+00	4.92.E-08	8.32.E+00
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	7.90.E-04	8.45.E-04	4.81.E-04	6.93.E-04	9.87.E-04	9.25.E-04	7.44.E-04	6.99.E-04	9.19.E-04	7.07.E-04	7.79.E-04
			k_{fin}	7162	4283	508	568	773	326	1397	483	3255	545	1930.0
			Function call times	35810	21415	2540	2840	3865	1630	6985	2415	16275	2725	9650.0
$n=10$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	7.21.E-07	3.05.E+01	7.44.E-07	2.80.E+01	7.65.E-07	3.11.E+00	7.43.E-07	7.68.E-07	7.82.E+00	1.44.E+01	8.38.E+00
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.64.E-04	9.23.E-04	9.95.E-04	9.93.E-04	9.82.E-04	9.53.E-04	9.69.E-04	9.87.E-04	9.90.E-04	9.85.E-04	9.74.E-04
			k_{fin}	519642	6814	24314	266620	343592	16083	649358	180738	573240	134919	271532
			Function call times	1039284	13628	48628	533240	687184	32166	1298716	361476	1146480	269838	543064
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.43.E+01	7.52.E-07	3.11.E+00	3.42.E+00	4.69.E+00	5.91.E+00	4.25.E+01	7.54.E-07	7.62.E-07	7.37.E-07	7.40.E+00
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.97.E-04	9.82.E-04	9.82.E-04	9.33.E-04	9.92.E-04	9.98.E-04	9.61.E-04	9.74.E-04	9.95.E-04	9.73.E-04	9.79.E-04
			k_{fin}	225673	11856	88620	39882	97585	56784	11715	164707	50829	1201195	194884.6
			Function call times	451346	23712	177240	79764	195170	113568	23430	329414	101658	2402390	389769.2
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	6.57.E+00	3.11.E-01	1.43.E+01	3.11.E-01	1.24.E+00	3.51.E+01	1.78.E-07	3.57.E+01	6.22.E-01	3.11.E-01	9.45.E+00
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	9.25.E-04	7.60.E-04	9.99.E-04	7.01.E-04	9.81.E-04	9.64.E-04	9.81.E-04	9.18.E-04	9.31.E-04	8.38.E-04	9.00.E-04
			k_{fin}	21782	34250	52645	2009	4901	80170	2076	54154	4689	2392	25906.8
			Function call times	108910	171250	263225	10045	24505	400850	10380	270770	23445	11960	129534
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.07.E-07	2.51.E-07	2.71.E+01	2.72.E+01	3.11.E-01	2.66.E+01	2.57.E-07	5.79.E+01	6.26.E+00	1.40.E+01	1.59.E+01
			$\lVert\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\rVert$	4.06.E-04	7.82.E-04	9.90.E-04	9.74.E-04	9.83.E-04	9.47.E-04	5.74.E-04	8.33.E-04	9.78.E-04	8.40.E-04	8.31.E-04
			k_{fin}	1322	26161	126761	66016	6958	35910	10168	1628	16475	17310	30870.9
			Function call times	6610	130805	633805	330080	34790	179550	50840	8140	82375	86550	154354.5

the initial search points $\boldsymbol{x}_i(0)$ $(i = 1, \ldots, m)$ in Table I satisfy condition 2 in each trial. The results shown in Table III confirm that all initial points lie within condition (41), i.e., each rank value based on the left-hand side of (41) equals the dimension n.

Remark 2: Structurally, Algorithm 2 does not need any gradient of the objective function. However, to verify Theorem 2 of the main result appropriately, this experiment needs to calculate the gradient as a termination criterion. In practice, the termination criterion should be replaced with other criteria without gradients.

2) Results and Discussion: The results of the numerical experiments for the five objective functions are shown in Tables IV–VIII. Each table shows the objective function value $f(\boldsymbol{x}^\star(k_{\text{fin}}))$, the gradient norm $\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$, and the objective function call times at the k_{fin} iteration when the algorithm terminated based on the termination criteria $\|\nabla f(\boldsymbol{x}^\star(k))\| < 10^{-3}$ or $k = 10^8$. Their values are shown for each condition composed of setups n, m, and ω for each trial using the different initial search points setting $x_{i,j}(0)$ shown in Table I. Furthermore, each mean value of 10 trials is shown in the rightmost column.

Note that the Translated sphere function's perturbations $a_i (i = 1, \ldots, n)$ in Appendix B, which were randomly placed in each trial, are shown in Table IX.

From the results we can confirm that the gradient norm reached the termination criterion $\|\nabla f(\boldsymbol{x}^\star(k))\| < 10^{-3}$ and the algorithm arrived at a stationary point in all cases. The established theory is verified numerically.

In addition to the main purpose, these results indicate the following aspects based on various perspective.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VIII
RESULTS FOR THE TRANSLATED SPHERE FUNCTION

Setup n	Setup m	Setup ω	Checked Items	Trial Number 1	2	3	4	5	6	7	8	9	10	10 Trials' Mean
$n=5$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.98.E-07	2.26.E-07	2.49.E-07	2.04.E-07	1.92.E-07	2.48.E-07	2.35.E-07	2.24.E-07	1.94.E-07	3.62.E-08	2.01.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.89.E-04	9.51.E-04	9.98.E-04	9.04.E-04	8.77.E-04	9.95.E-04	9.69.E-04	9.46.E-04	8.80.E-04	3.80.E-04	8.79.E-04
			k_{fin}	657	526	2470	1559	425	5946	619	339	416	489	1344.6
			Function call times	1314	1052	4940	3118	850	11892	1238	678	832	978	2689.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.42.E-07	2.19.E-07	2.37.E-07	2.25.E-07	2.39.E-07	2.48.E-07	2.36.E-07	6.55.E-08	2.37.E-07	2.49.E-07	2.20.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.84.E-04	9.37.E-04	9.73.E-04	9.48.E-04	9.78.E-04	9.96.E-04	9.71.E-04	5.12.E-04	9.73.E-04	9.99.E-04	9.27.E-04
			k_{fin}	582	317	2479	1496	370	5765	514	404	519	328	1277.4
			Function call times	1164	634	4958	2992	740	11530	1028	808	1038	656	2554.8
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.49.E-07	3.34.E-08	1.22.E-07	1.61.E-07	2.39.E-07	2.37.E-07	2.02.E-07	2.48.E-07	2.01.E-07	2.33.E-07	1.93.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	9.98.E-04	3.66.E-04	6.98.E-04	8.03.E-04	9.78.E-04	9.74.E-04	9.00.E-04	9.97.E-04	8.96.E-04	9.66.E-04	8.57.E-04
			k_{fin}	581	566	520	406	439	458	473	485	616	481	502.5
			Function call times	2905	2830	2600	2030	2195	2290	2365	2425	3080	2405	2512.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.00.E-07	2.26.E-07	1.06.E-07	2.43.E-07	1.87.E-07	1.46.E-07	2.10.E-07	1.53.E-07	2.11.E-07	2.00.E-07	1.88.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.93.E-04	9.52.E-04	6.50.E-04	9.87.E-04	8.66.E-04	7.65.E-04	9.17.E-04	7.83.E-04	9.18.E-04	8.94.E-04	8.62.E-04
			k_{fin}	425	408	316	395	293	337	307	221	368	375	344.5
			Function call times	2125	2040	1580	1975	1465	1685	1535	1105	1840	1875	1722.5
$n=10$	$m=2$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.02.E-07	2.27.E-07	2.22.E-07	2.40.E-07	1.87.E-07	1.87.E-07	2.24.E-07	2.37.E-07	2.08.E-07	1.88.E-07	2.12.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.99.E-04	9.54.E-04	9.42.E-04	9.80.E-04	8.65.E-04	8.65.E-04	9.46.E-04	9.74.E-04	9.12.E-04	8.67.E-04	9.20.E-04
			k_{fin}	1769	1440	1836	1674	1868	1756	2402	1498	2047	1416	1770.6
			Function call times	3538	2880	3672	3348	3736	3512	4804	2996	4094	2832	3541.2
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.97.E-07	2.16.E-07	2.10.E-07	1.89.E-07	2.23.E-07	2.36.E-07	2.27.E-07	2.08.E-07	2.35.E-07	2.26.E-07	2.17.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.87.E-04	9.30.E-04	9.16.E-04	8.69.E-04	9.45.E-04	9.71.E-04	9.53.E-04	9.12.E-04	9.70.E-04	9.50.E-04	9.30.E-04
			k_{fin}	1552	1388	1313	1557	1352	1202	2069	1029	1817	1283	1456.2
			Function call times	3104	2776	2626	3114	2704	2404	4138	2058	3634	2566	2912.4
	$m=5$	$\omega=\frac{1}{2}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	2.00.E-07	1.62.E-07	2.47.E-07	2.09.E-07	1.97.E-07	2.21.E-07	2.20.E-07	2.40.E-07	1.33.E-07	1.81.E-07	2.01.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.95.E-04	8.05.E-04	9.94.E-04	9.13.E-04	8.87.E-04	9.39.E-04	9.38.E-04	9.81.E-04	7.29.E-04	8.50.E-04	8.93.E-04
			k_{fin}	1304	1244	1615	1108	1105	1458	1243	1215	1673	1426	1339.1
			Function call times	6520	6220	8075	5540	5525	7290	6215	6075	8365	7130	6695.5
		$\omega=\frac{1}{5}$	$f(\boldsymbol{x}^\star(k_{\text{fin}}))$	1.73.E-07	2.34.E-07	2.12.E-07	2.24.E-07	2.45.E-07	2.47.E-07	2.29.E-07	1.89.E-07	2.17.E-07	1.84.E-07	2.15.E-07
			$\|\nabla f(\boldsymbol{x}^\star(k_{\text{fin}}))\|$	8.32.E-04	9.68.E-04	9.21.E-04	9.47.E-04	9.89.E-04	9.93.E-04	9.57.E-04	8.70.E-04	9.32.E-04	8.58.E-04	9.27.E-04
			k_{fin}	1024	1123	832	1120	1121	1168	974	1362	1329	1614	1166.7
			Function call times	5120	5615	4160	5600	5605	5840	4870	6810	6645	8070	5833.5

TABLE IX
TRANSLATED SPHERE FUNCTION'S PERTURBATIONS $a_i (i = 1, \ldots, n)$ RANDOMLY PLACED IN $(-5, 5)$ FOR EACH TRIAL

Trial No.	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1	1.44318	-2.92258	-1.88898	0.94896	-4.14484	4.63089	-4.62261	-3.93238	-4.69459	-3.17078
2	-1.21391	-1.98754	4.2338	-2.37788	-2.37518	0.46806	3.85168	1.53757	2.44074	-2.60068
3	3.1158	-0.29077	-0.69793	1.02843	3.01015	0.21136	4.13287	-0.05826	0.00022	3.86512
4	0.32826	-2.69512	-3.15184	2.11216	-4.7078	-2.68406	2.96184	2.79052	-0.20078	-4.71326
5	-1.49273	3.44309	4.04881	-2.78253	4.28854	-0.11102	-4.01288	2.15037	4.04722	-0.10099
6	4.39002	-3.05236	4.79748	-3.82582	2.30331	1.2406	-2.38129	4.03721	1.09867	-3.32073
7	3.75943	-2.74078	-0.6113	-2.03324	-0.11391	1.79136	-1.64643	3.90923	1.17666	4.78681
8	0.50156	-3.29292	-3.88881	-1.81222	0.78525	-1.04485	1.79728	-1.65837	3.59442	2.12694
9	1.22475	-2.72336	-2.41935	-0.75833	-2.62716	-1.32563	-3.63447	1.98746	3.05489	0.00472
10	0.87045	-0.64301	-0.9128	0.07858	-0.41151	4.87982	2.21227	-3.0219	0.76722	-0.28912

1) *Dimension Number n:* Comparing the mean values of k_{fin} between $n = 5$ and $n = 10$ for each condition in each table, we observe that $n = 10$ makes their values bigger than $n = 5$ for each case. We can analyze this reason in $\sigma > 0$ in (14) in Theorem 1's proof as follows. From (10) of Theorem 1, σ is the minimum angle among all angles between the spiral descent directions and the orthogonal directions of the gradients at the centers for all iterations. This means that making σ small enables the spiral descent directions to approach their orthogonal directions, which can cause slow convergence. Therefore, due to the definition $\sigma = 1/(\zeta_{\max}\sqrt{n})$ of (14), we can understand that increasing n makes σ decrease and causes slower convergence.
2) *Search Points Number m:* Comparing the mean values of the function call times between $m = 2$ and $m = 5$ for each condition in Tables IV and V, we observe that for the sphere function, $m = 2$ makes the values smaller than $m = 5$. For the Schwefel function, $m = 5$ makes the values much smaller than $m = 2$. Thus, considering that the two functions are unimodal with unique stationary points, we understand that the effect of m depends on the structure of each function.
3) *Setting Parameter w:* Comparing the mean values of k_{fin} between $w = 1/2$ and $w = 1/5$ for each condition in each table, we could not observe significant differences for each case.
4) *Function Structure:* To investigate the effects of function structures, we first checked the results in Table IV for the sphere function and Table VI for the 2^n minima function, where the degree of the nonlinearity of the 2^n minima function is four more than two of the sphere function. It was confirmed that each mean value of k_{fin} in the 2^n minima is much greater than each one in another function. Second, we also checked the results in Table IV for the sphere function and Table VIII for the Translated sphere function, where both degrees of nonlinearity are the same. The mean values of k_{fin} for each condition are similar. From the two comparisons, we can understand that the degree of nonlinearity of the objective function can affect convergence speed.

VII. CONCLUSION

In this paper, we proposed the conditions and settings under which the SPO algorithm converges to a stationary point. These cover the composite rotation matrix, the step rate, and the initial placement of search points, which characterize each spiral trajectory. Their effectiveness was mathematically proved and numerically verified.

The SPO algorithm has been studied as a nature-inspired metaheuristic that aims to find a better approximated solution within a limited number of iterations specified by the user. In this paper, we showed that this algorithm can also be considered as a strict direct search method to find a stationary point from the proposed settings. This demonstrates the versatility of the SPO algorithm only by changing the conditions of the parameter settings.

This is possibly the first time convergence of a nature-inspired continuous algorithm to a stationary point has been proved. Initially, the spiral phenomena were intuitively considered as nature phenomena appropriate for metaheuristics considering that the behavior have both diversification and intensification. Currently, the spiral phenomena have been considered theoretically as a natural phenomenon appropriate for optimization. Accordingly, this convergence search is conducted using exclusively the simple dynamical system of a spiral model.

Based on these results, future research on the SPO algorithm should aim to enhance its global search performance and efficiency.

Appendix A

Lemma A [19]: For any $\boldsymbol{y} \in \mathbb{Y} = \{\boldsymbol{\xi} \in \mathbb{R}^n \mid \|\boldsymbol{\xi}\| = 1\}$, the following holds:

$$\max_{1\leqq j\leqq n}\left\{\left|\boldsymbol{y}^\top \boldsymbol{e}_j\right|\right\} \geqq \frac{1}{\sqrt{n}} \tag{43}$$

where $\boldsymbol{e}_j$ is a unit vector of which jth element is 1.

Proof: See [19, Lemma 3.1]. ∎

Mean-Value Theorem [31]: Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable on a convex set $\mathbb{D} \subset \mathbb{R}^n$. Then, for any two points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{D}$, there is $\lambda \in (0, 1)$ such that

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) = < \nabla f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} >.$$

Cauchy–Schwarz Inequality: For any two vectors $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^n$, the following holds: $|\boldsymbol{x}^\top \boldsymbol{y}| \leqq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$.

Appendix B

Used continuous differentiable test functions are defined with properties as follows.

1) *Sphere Function:*
   a) *Definition:* $f(\boldsymbol{x}) = \sum_{i=1}^n x_i^2$.
   b) *Opt. Solution:* $\boldsymbol{x}^* = [0, 0, \ldots, 0]^\top$.
   c) *Opt. Value:* $f(\boldsymbol{x}^*) = 0$.
   d) *Properties:* Unimodality, separability.
2) *Schwefel Function:*
   a) *Definition:* $f(\boldsymbol{x}) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$.
   b) *Opt. Solution:* $\boldsymbol{x}^* = [0, 0, \ldots, 0]^\top$.
   c) *Opt. Value:* $f(\boldsymbol{x}^*) = 0$.
   d) *Properties:* Unimodality, nonseparability.
3) *2^n Minima Function:*
   a) *Definition:* $f(\boldsymbol{x}) = \sum_{i=1}^n x_i^4 - 16x_i^2 + 5x_i$.
   b) *Opt. Solution:* $\boldsymbol{x}^* \approx [-2.9, -2.9, \ldots, -2.9]^\top$.
   c) *Opt. Value:* $f(\boldsymbol{x}^*) \approx -78n$.
   d) *Properties:* Multimodality, separability.
4) *Levy Function:*
   a) *Definition:* $f(\boldsymbol{x}) = \pi/n[\sum_{i=1}^{n-1}\{(x_i - 1)^2(1 + 10\sin^2(\pi x_{i+1}))\} + 10\sin^2(\pi x_1) + (x_n - 1)^2]$.
   b) *Opt. Solution:* $\boldsymbol{x}^* = [1, 1, \ldots, 1]^\top$.
   c) *Opt. Value:* $f(\boldsymbol{x}^*) = 0$.
   d) *Properties:* Multimodality, nonseparability.
5) *Translated Sphere Function:*
   a) *Definition:* $f(\boldsymbol{x}) = \sum_{i=1}^n (x_i - a_i)^2$ where a_i is a uniform random real number in $(-5, 5)$.
   b) *Opt. Solution:* $\boldsymbol{x}^* = [a_1, a_2, \ldots, a_n]^\top$.
   c) *Opt. Value:* $f(\boldsymbol{x}^*) = 0$.
   d) *Properties:* Unimodality, separability.

References

[1] K. Tamura and K. Yasuda, "Primary study of spiral dynamics inspired optimization," *IEEJ Trans. Elect. Electron. Eng.*, vol. 6, pp. 98–100, Feb. 2011.
[2] K. Tamura and K. Yasuda, "Spiral dynamics inspired optimization," *J. Adv. Comput. Intell. Inf.*, vol. 15, no. 8, pp. 1116–1122, Oct. 2011.
[3] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, 2nd ed. Frome, U.K.: Luniver Press, 2010.
[4] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, WA, Australia, 1995, pp. 1942–1948.
[5] M. Clerc and J. Kennedy, "The particle swarm: Explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, Feb. 2002.
[6] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," *J. Glob. Optim.*, vol. 39, no. 3, pp. 459–471, Nov. 2007.
[7] B. Akay and D. Karaboga, "A modified Artificial Bee Colony algorithm for real-parameter optimization," *Inf. Sci.*, no. 192, pp. 120–142, Jun. 2012.
[8] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *Proc. NaBIC*, Coimbatore, India, 2009, pp. 210–214.
[9] M. R. AlRashidi and M. E. El-Hawary, "A survey of particle swarm optimization applications in electric power systems," *IEEE Trans. Evol. Comput.*, vol. 13, no. 4, pp. 913–918, Aug. 2009.
[10] N. Todorovic and S. Petrovic, "Bee colony optimization algorithm for nurse rostering," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 2, pp. 467–473, Mar. 2013.
[11] W. Han *et al.*, "Cuckoo search and particle filter-based inversing approach to estimating defects via magnetic flux leakage signals," *IEEE Trans. Magn.*, vol. 52, no. 4, pp. 1–11, Apr. 2016.
[12] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 256–279, Jun. 2004.
[13] H. Chen *et al.*, "Artificial bee colony optimizer based on bee life-cycle for stationary and dynamic optimization," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 2, pp. 327–346, Feb. 2017.
[14] W. Yuan, Y. Liu, H. Wang, and Y. Cao, "A geometric structure-based particle swarm optimization algorithm for multiobjective problems," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.
[15] R. M. Lewis, V. Torczon, and M. W. Trosset, "Direct search methods: Then and now," *J. Comput. Appl. Math.*, vol. 124, nos. 1–2, pp. 191–207, Dec. 2000.
[16] R. Hooke and T. A. Jeeves, "'Direct Search' solution of numerical and statistical and problems," *J. ACM*, vol. 8, no. 2, pp. 212–229, 1961.
[17] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.
[18] V. Torczon, "On the convergence of the multidirectional search algorithm," *SIAM J. Optim.*, vol. 1, no. 1, pp. 123–145, Feb. 1991.
[19] V. Torczon, "On the convergence of pattern search algorithms," *SIAM J. Optim.*, vol. 7, no. 1, pp. 1–25, Feb. 1997.
[20] K. Tamura and K. Yasuda, "Quantitative analysis based tuning law for convergence rate of spiral optimization," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Seoul, South Korea, 2012, pp. 767–772.

[21] K. Tamura and K. Yasuda, "A stability analysis based parameter setting method for spiral optimization," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Manchester, U.K., 2013, pp. 3909–3914.

[22] K. Tamura and K. Yasuda, "A parameter setting method for spiral optimization from stability analysis of dynamic equilibrium point," *SICE J. Control Meas. Syst. Integr.*, vol. 7, no. 3, pp. 173–182, May 2014.

[23] K. Tamura and K. Yasuda, "Spiral optimization algorithm using periodic descent directions," *SICE J. Control Meas. Syst. Integr.*, vol. 9, no. 3, pp. 134–143, May 2016.

[24] K. Suzuki, K. Tamura, and K. Yasuda, "Study on cluster-structured spiral optimization," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, San Diego, CA, USA, 2014, pp. 94–99.

[25] A. N. K. Nasir and M. O. Tokhi, "An improved spiral dynamic optimization algorithm with engineering application," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 6, pp. 943–954, Jun. 2015.

[26] A. N. K. Nasir, R. M. T. R. Ismail, and M. O. Tokhi, "Adaptive spiral dynamics metaheuristic algorithm for global optimisation with application to modelling of a flexible system," *Appl. Math. Model.*, vol. 40, nos. 9–10, pp. 5442–5461, May 2016.

[27] A. N. K. Nasir and M. O. Tokhi, "A novel hybrid spiral-dynamics randomchemotaxis optimization algorithm with application to modelling of a flexible robot manipulator," in *Proc. 16th Int. Conf. Climbing Walking Robots (CLAWAR)*, Sydney, NSW, Australia, 2013, pp. 667–674.

[28] A. Ouadi, H. Bentarzi, and A. Recioui, "Optimal multiobjective design of digital filters using spiral optimization technique," *Springerplus*, vol. 2, no. 461, pp. 697–707, Sep. 2013.

[29] L. Benasla, A. Belmadani, and M. Rahli, "Spiral optimization algorithm for solving combined economic and emission dispatch," *Int. J. Elect. Power Energy Syst.*, vol. 62, pp. 163–174, Nov. 2014.

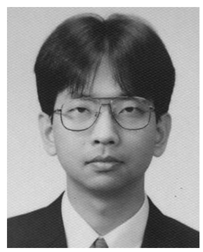
[30] K. A. Sidarto and A. Kania, "Finding all solutions of systems of nonlinear equations using spiral dynamics inspired optimization with clustering," *J. Adv. Comput. Intell. Inf.*, vol. 19, no. 5, pp. 697–707, Sep. 2015.

[31] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. New York, NY, USA: Academic, 1980.

[32] J. E. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Classics in Applied Mathematics). Philadelphia, PA, USA: SIAM, 1996.

Kenichi Tamura (A'08) received the B.E. degree in system design engineering and the M.E. and Ph.D. degrees in integrated design engineering from Keio University, Yokohama, Japan, in 2003, 2005, and 2008, respectively.

Since 2008, he has been with Tokyo Metropolitan University, Tokyo, Japan, where he is currently an Assistant Professor with the Department of Electrical and Electronic Engineering. His current research interests include black-box optimization, direct search theory, and control theory.



Keiichiro Yasuda (M'99) received the B.E. degree from Tokyo Metropolitan University, Tokyo, Japan, in 1984, and the M.E. and Ph.D. degrees from Hokkaido University, Sapporo, Japan, in 1986 and 1989, respectively, all in electrical engineering.

Since 1989, he has been with Tokyo Metropolitan University, where he is currently a Professor with the Department of Electrical and Electronic Engineering. His current research interests include systems optimization and power systems engineering.