



*Research article*

## The first hitting time analysis of evolutionary algorithms based on renewal process

Zhensheng Zhou<sup>1,2</sup>, Lin Wang<sup>2,\*</sup>, Xue Zou<sup>1,2</sup>, Fei Wang<sup>3</sup>, Zaijun Zhang<sup>4,5</sup> and Xiaobo Yan<sup>2,\*</sup>

<sup>1</sup> School of Data Sciences and Information Engineering, Guizhou Minzu University, Guiyang 550025, China

<sup>2</sup> Guizhou Key Laboratory of Pattern Recognition and Intelligent System, Guiyang 550025, China

<sup>3</sup> College of Big Data and Information Engineering, Guiyang Institute of Humanities and Technology, Guiyang 550025, China

<sup>4</sup> School of Mathematics and Statistics, Qiannan Normal University for Nationalities, Duyun 558000, China

<sup>5</sup> Key Laboratory of Industrial Automation and Machine Vision of Qiannan, Duyun 558000, China

\* **Correspondence:** Email: wanglin@gzmu.edu.cn, xbYan@gzmu.edu.cn.

**Abstract:** Running time analysis of evolutionary algorithms for continuous optimization is one research challenge in the field of evolutionary algorithms (EAs). However, the theoretical analysis results have rarely been applied to evolutionary algorithms for continuous optimization in practice, let alone their variants for evolution strategy. In this paper, we regarded the first hitting time of evolution strategy as the stopping time of the renewal process on the basis of the renewal process and in combination with Wald's inequality and stopping time theory. Afterwards, to demonstrate the application of the proposed model in the first hitting time analysis of  $(1 + 1)$  ES, we analyzed it with different mutation operators on the sphere function. First, we significantly improved the lower bound on the first hitting time of  $(1 + 1)$  ES with a uniform mutation operator, i.e., from  $\Omega(n)$  to  $\Omega(e^n)$ . Next,  $O(n^2\sqrt{n})$  was the upper bound on the first hitting time of  $(1 + 1)$  ES with a Gaussian mutation operator from the initial distance  $R$  to half of the initial distance  $R/2$ . The numerical experimental results showed that the theoretical calculation was consistent with the actual running time, which provides a novel method for analyzing the first hitting time of EAs.

**Keywords:** evolutionary algorithms for continuous optimization; renewal process; stopping time theory; first hitting time; sphere function

---

## 1. Introduction

Evolutionary algorithms (EAs) are a class of adaptive, global search algorithms inspired by natural evolutionary processes, mainly including genetic algorithms (GA), evolution strategy (ES), genetic programming (GP), and evolution programming (EP) [1]. EAs have been widely applied to the practical application domains, such as black box optimization and combinatorial optimization, yet a number of theoretical analysis results are relatively limited. As a commonly used indicator to measure the computational time, the first hitting time is the number of generations when EAs find the optimal solution or a satisfactory solution for the first time [2]. Moreover, the expected first hitting time is the average number of evaluations of the fitness function when EAs obtain the optimal or satisfactory solution for the first time, which evaluates the performance of EAs [3].

In the past few years, most theoretical approaches of the first hitting time have focused on discrete search spaces. Drift analysis, combined with Markov chains in the analysis of computation time complexity of EAs, was first introduced by He and Yao [4]. As drift analysis was being developed, multiplicative drift [5,6] and variable drift [7–9] have shown that the lower bound of the first hitting time is stronger than its upper bound. Sarker et al. [10] considered the first hitting time of EAs as a random variable and derived its time on a pseudo-Boolean function. Yu et al. [11] proposed a switch analysis method for estimating the time bound of EAs by combining two Markov chains, one of which served as a reference chain. Qian et al. [12] transformed the bit-by-bit noise into one-bit noise and discussed the first hitting time of EAs on the OneMax and LeadingOnes functions. Wegener et al. [13] analyzed the running time of EAs on the pseudo-Boolean function using tail inequalities. In these works, however, there have been few theoretical analyses in EAs for continuous optimization. Huang et al. [14] combined statistical methods with the average gain model to obtain an empirical distribution function for the expected gain and derived the upper bound closed expression. Feng et al. [15] established an equivalent relationship to analyze the effect of selection operators on the computation time of EAs. To generalize the drift analysis model, Morinaga et al. [16] introduced  $\sigma$ -algebraic flow stochastic process into the model. Jägersküpfer [17] combined drift analysis with Markov chains, treating drift as a Markov chain, and analyzed the first hitting time of EAs on the linear function. Many important theoretical analysis results of them primarily focus on simplified versions rather than the  $(1 + 1)$  ES, which stems from the difficulties posed by population-based characteristics and the intricate adaptive selection strategies involved.

Agapie et al. [18] combined the transition kernel with the success region where each generation of  $(1 + 1)$  ES arrives and obtained the cumulative distribution function of success probability, which was estimated as the upper and lower bounds of the first hitting time for  $(1 + 1)$  ES with a uniform mutation operator on the sphere function. However, when the gap between the population fitness value and the optimal fitness value was substantial, they approximated the intersection of the fitness sphere and the mutation sphere as half of the mutation sphere. These approximations resulted in an inflated expected progress rate, ultimately causing significant looseness between the lower bound of the first hitting time of  $(1 + 1)$  ES on the sphere function and its actual running time. Morinaga et al. [19] proved that the expected hitting time of  $(1 + 1)$  ES on the convex quadratic functions was  $\Theta(\exp(-1/d))$ . Zhang et al. [20] analyzed the upper bound of the expected hitting time of  $(1 + 1)$  ES on the sphere function by combining martingale theory with the Lebesgue-dominated convergence theorem, but they did not provide a corresponding theoretical lower bound analysis. Agapie et al. [21] combined the reduction of distance from the  $t-1$  generation population to the optimal solution with differential equations and analyzed the lower bound of the hitting time of  $(1 + 1)$  ES with a Gaussian mutation operator on the sphere function

as  $\Omega(n \cdot \log(\frac{R_0}{\varepsilon}))$ . Doerr et al. [22] established a dynamic drift model to analyze the relationship between the evolutionary strategy parameter selection and the first hitting time and obtained the expected first hitting time  $\Theta(n\lambda / \log \lambda + n / \log n)$  on the OneMax function. Akimoto et al. [23] improved additive drift by treating the progress rate as a monotonic decreasing function, and analyzed that the convergence rate of the 1/5 success rule (1 + 1) ES solving sphere function was  $\Theta(1/d)$ . Jägersküpfer [24] hypothesized that individuals were generated through anisotropic mutation and further obtained the asymptotical running time of (1 + 1) ES with a Gaussian mutation operator solving sphere function. Although many achievements have been made in the analysis of the first hitting time of (1 + 1) ES on the sphere function, few theoretical research results pay attention to its upper bound with a Gaussian mutation operator. Meanwhile, due to the simplified relationship between the sphere objective function and the mutation sphere, its lower bound with a uniform mutation operator is not tighter.

The main innovations and contributions of this paper can be outlined as follows:

1) The proposed method is based on a non-negative progress rate stochastic process, which is conducive to separating specific cases and algorithms, making the model more general.

2) The intersection of the fitness sphere and the mutation sphere is considered the success region where each generation of (1 + 1) ES arrives, and its lower bound with a uniform mutation operator is obtained for a tighter first hitting time.

3) The effectiveness of the 1/5 success rule of the (1 + 1) ES is verified, and its upper bound with a Gaussian mutation operator is obtained.

The structure of this paper is as follows. Some preliminary knowledge on the renewal theorem and stopping time is introduced, and then the renewal model is proposed in Section 2. The lower bound of the expected first hitting time of (1 + 1) ES with a uniform mutation operator is analyzed in Section 3. The upper bound of the expected first hitting time of (1 + 1) ES with a Gaussian mutation operator is analyzed in Section 4. Numerical experiments and results are reported in Section 5. Finally, Section 6 draws the conclusion.

## 2. Preliminaries

In this section, we introduce the general process framework of (1 + 1) ES, renewal process, stopping time, and mathematical modeling of the algorithm.

### 2.1. Problem description and algorithm introduction

We consider minimizing the sphere function in this paper. Its corresponding mathematical expression is as follows:

$$f(x) = \sqrt{\sum_{i=1}^n x_i^2}, \quad (1)$$

$$s.t. x_i \in \mathbb{Q} \subseteq \mathbb{R}^n,$$

where  $f(x)$  is the fitness function and  $\mathbb{Q}$  is the search space. Let  $\xi^t = \{\xi_1^t, \xi_2^t, \dots, \xi_n^t\} \in \mathbb{Q}$  be the  $t$ -generation population of EAs and  $n$  represents the problem size. The process description of (1 + 1) ES can be described as follows [23]:

**Algorithm 1:** (1 + 1) ES with 1/5 success rule

1. **Input:** Initialize solution  $\xi_0$  and step size  $\sigma_0$  ;
2. **while**  $t = 1, 2, \dots$ , until some stopping criterion is fulfilled **do**
3. **mutation:**  $\xi_t := \xi_{t-1} + \sigma_{t+G} \cdot \Delta_t$  ;
4. **evaluate** solution  $\xi_t$  ;
5. **selection:** if  $f(\xi_{t-1}) < f(\xi_t)$ , then
6.  $\xi_t := \xi_{t-1}$  ;
7.  $\sigma_t := \sigma_{t-1} \cdot A$  ;
8. **else**
9.  $\xi_t := \min\{\xi_t, \xi_{t-1}\}$  ;
10. **end while**
11. **output** the satisfactory solution.

In this paper, Algorithm 1 includes the use of mutation and selection.  $\Delta_t$  uses a uniform mutation operator or Gaussian mutation operator; election is elite selection, that is, one individual is selected from the parent and offspring individuals as the population individual of the next generation. Within a certain range of iteration times  $G$ , the successful iteration times  $G'$  is the iteration number that satisfies the condition of  $f(\xi_t) < f(\xi_{t-1})$ . Let  $p_s = \frac{G'}{G}$ , if  $p_s > 1/5$ , the mutation strength reduces, i.e.,  $A = \alpha$  ; if  $p_s < 1/5$ , the mutation strength increases, i.e.,  $A = \alpha^{-1/4}$ . Otherwise, it remains unchanged. Generally speaking,  $0.85 \leq \alpha < 1$  [25]. However, the main results presented in this paper, i.e., Theorem 3, are independent of such a specific implementation and can be applied to almost any stochastic search algorithm.

## 2.2. A renewal model of EAs

We consider the gradual process of EAs from their initial position to a satisfactory solution as a renewal process to establish a renewal model, which is different from the theoretical results on intelligent optimization algorithms [26,27]. Before establishing the model, the following theorem needs to be given.

**Definition 1 (fitness difference function).** Let  $d(\xi_{opt}^t)$ ,  $t > 0$  be a random variable, then  $d(\xi_{opt}^t) = f(\xi_{opt}^t) - f_{opt}$  is called the fitness difference function, where  $\xi_{opt}^t$  represents the optimal individual of the  $t$  generation, and  $f_{opt}$  represents the target fitness value, which measures the distance of the  $t$  generation population to the optimal solution.

**Definition 2 (progress rate) [28].** During the running of EAs, the progress rate represents the decrease in the optimal distance between the  $t-1$  to  $t$  generation populations, i.e.,  $X_t = d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)$ ,  $t=1, 2, \dots$ , and quality gain represents the decrease in fitness between  $t-1$  generation and  $t$  generation populations, which reflects the convergence time of EAs. The larger the progress rate, the faster the distance reduction from the optimal solution changes, and the faster the convergence time. It is obvious that  $\{X_t\}_{t=0}^{\infty}$  is a non-negative stochastic process, and the definition is as follows.

**Definition 3 (renewal process of progress rate).** Let  $\{X_t\}_{t=0}^{\infty}$  be a non-negative stochastic process. If there exists  $Y_t = \sum_{n=1}^t X_n = d(\xi_0) - d(\xi_{opt}^t)$  for any  $t=1, 2, \dots$ , then  $Y_t$  is referred to as the renewal process of the progress rate of EAs in the  $t$  generation.

In stochastic algorithms, the first hitting time is an effective way to characterize its performance, i.e., the number of generations when EAs find a satisfactory solution for the first time. It can be described by the stopping time of the renewal process, and the definition of stopping time is defined.

**Definition 4 (stopping time of renewal process).** Let  $\{X_t\}_{t=0}^{\infty}$  be a non-negative stochastic process, if  $T_{\tau} = \{\tau | \sum_{k=1}^{\tau-1} X_k < d(\xi_0) - \varepsilon, \sum_{k=\tau}^{\infty} X_k \geq d(\xi_0) - \varepsilon\}$  is satisfied, then  $T_{\tau}$  is described as a stopping time of  $X_t$ , where  $\tau$  is a non-negative random variable.

The stopping time of the renewal process is an important concept in stochastic processes, which is described as independent of the state of the progress rate stochastic process  $\{X_t\}_{t=T_{\tau}}^{\infty}$  after time  $T_{\tau}$  and related to the state of the progress rate stochastic process  $\{X_t\}_{t=0}^{T_{\tau}}$  before that time.

**Theorem 1 (Wald's equation).** If  $\{X_t\}_{t=0}^{\infty}$  are random variables with independent and identically distributed  $X$  having finite expectations  $E(X) < \infty$ , and  $T_{\tau}$  is a stopping time for  $\{X_t\}_{t=0}^{\infty}$  such that  $E(T_{\tau}) < \infty$ , then

$$E\left(\sum_{t=1}^{T_{\tau}} X_t\right) = E(T_{\tau})E(X). \quad (2)$$

*Proof.* Let:  $I_t = \begin{cases} 1, & \text{if } T_{\tau} \leq t; \\ 0, & \text{if } T_{\tau} > t. \end{cases}$  Because  $\{I_t = 1\} = \{T_{\tau} \leq t\} = \bigcup_{k=1}^t \{T_{\tau} = k\}$  and the progress rate stochastic process  $\{X_t\}_{t=T_{\tau}}^{\infty}$  are independent of each other, it can be concluded that  $\{I_t = 0\}$  and  $\{X_t\}_{t=T_{\tau}}^{\infty}$  are also independent of each other. Therefore,

$$E\left(\sum_{t=1}^{T_{\tau}} X_t\right) = E\left(\sum_{t=1}^{\infty} X_t I_t\right) = \sum_{t=1}^{\infty} E(X_t)E(I_t) = E(X) \sum_{t=1}^{\infty} P(T_{\tau} \geq t) = E(X)E(T_{\tau}).$$

In Theorem 1, it holds true that the expectation of the renewal process of the progress rate is equal to the expected first hitting time multiplied by the expected progress rate.

**Inference 1 (Wald's inequation).** According to Theorem 1, in the running of EAs, if  $T_{\tau}$  is the stopping time of  $\{X_t\}_{t=0}^{\infty}$  and there are two functions  $M_1$  and  $M_2$  that make  $M_1 < E(X_t) < M_2$  hold, then  $E\left(\sum_{t=1}^{T_{\tau}} X_t\right)$  satisfies:

$$M_1 E(T_{\tau}) < E\left(\sum_{t=1}^{T_{\tau}} X_t\right) < M_2 E(T_{\tau}). \quad (3)$$

We can derive the renewal theory theorem for the expected first hitting time based on Inference 1, as shown in Theorem 2.

**Theorem 2 (renewal theorem).** If  $\{X_t\}_{t=0}^{\infty}$  are random variables with independent and identically distributed  $X$  having  $M_1 < E(X) < M_2$  and  $T_{\tau}$  is a stopping time for  $\{X_t\}_{t=0}^{\infty}$  such that  $E(T_{\tau}) < \infty$ , then

$$\frac{1}{M_2} < \frac{E(T_{\tau})}{d_{T_{\tau}}} < \frac{1}{M_1}. \quad (4)$$

*Proof.* For the lower bound proof, the stopping time  $T_{\tau} = \{\tau | \sum_{k=1}^{\tau-1} X_k < d(\xi_0) - \varepsilon, \sum_{k=\tau}^{\infty} X_k \geq d(\xi_0) - \varepsilon\}$  in Definition 4 is equivalent to  $T_{\tau} = \{\tau | \sum_{k=1}^{\tau} X_k > d(\xi_0) - \varepsilon, \sum_{k=1}^{\tau-1} X_k \leq d(\xi_0) - \varepsilon\}$ , it follows from Inference 1 that the equation  $d_{T_{\tau}} < E\left(\sum_{n=1}^{T_{\tau}} X_n\right) < M_2 E(T_{\tau})$  holds. Thereby,  $\liminf_{d_{T_{\tau}} \rightarrow \infty} \frac{E(T_{\tau})}{d_{T_{\tau}}} > \frac{1}{M_2}$ . Now, we prove its

upper bound. Let us fix a constant  $a$  and define another renewal process, i.e.,  $X_t^r = \begin{cases} X_t, & \text{if } X_t \leq a; \\ a, & \text{if } X_t > a. \end{cases}$ ,

$\mu_a = \min\{M_1, \inf E(X_t^r)\}$ . Since  $0 < X_t^r < X_t$  and  $0 < \mu_a < E(X_t^r) < M_2$ , for all  $t$ , it follows that  $E(\sum_{n=1}^{T_r} X_n) < E(\sum_{n=1}^{T_t} X_n)$ , i.e.,  $E(T_r) > E(T_t)$ . Due to Inference 1, this yields  $\mu_a E(T_t) < \mu_a E(T_r) < E(\sum_{n=1}^{T_r} X_n) < d_{T_r} + a$ , i.e.,  $\limsup_{d_{T_r} \rightarrow \infty} \frac{E(T_r)}{d_{T_r}} < \frac{1}{M_1}$ , thus  $\frac{1}{M_2} < \frac{E(T_r)}{d_{T_r}} < \frac{1}{M_1}$ .

In the running of EAs, due to the difficulty in finding the optimal solution, when the fitness difference function value of the parent individual to the optimal solution is equal to  $\varepsilon$ , the algorithm terminates, that is, it reaches a specific stopping distance of  $Y_t = E(\sum_{i=1}^{T_r} X_i) = d(\xi_0) - \varepsilon$ , and then Theorem 3 represents the stochastic process model of the expected first hitting time analysis.

**Theorem 3 (renewal model).** According to Theorem 2, in the running of EAs, if  $T_r$  is the stopping time of  $\{X_t\}_{t=0}^{\infty}$  and there exist two functions  $M_1$  and  $M_2$ , such that  $M_1 < E(X_t) < M_2$ , when EAs pass through the interval  $[\varepsilon, d(\xi_0)]$ , then it holds for  $T_r$  that

$$\frac{d(\xi_0) - \varepsilon}{M_2} < E(T_r) < \frac{d(\xi_0) - \varepsilon}{M_1}. \quad (5)$$

Theorem 3 can be obtained from the proof of the additive drift theorem of Kötzing [29], which is a continuous space adaptation of the original discrete space drift result of He and Yao [30]. In Theorem 3, the upper and lower bounds on the expected progress rate vary as a function of  $M_1$  and  $M_2$ , which result in tighter upper and lower bounds on the first hitting time. Sections 3 and 4 analyze the expected first hitting time on the sphere function using the uniform and Gaussian mutation operators as a basis.

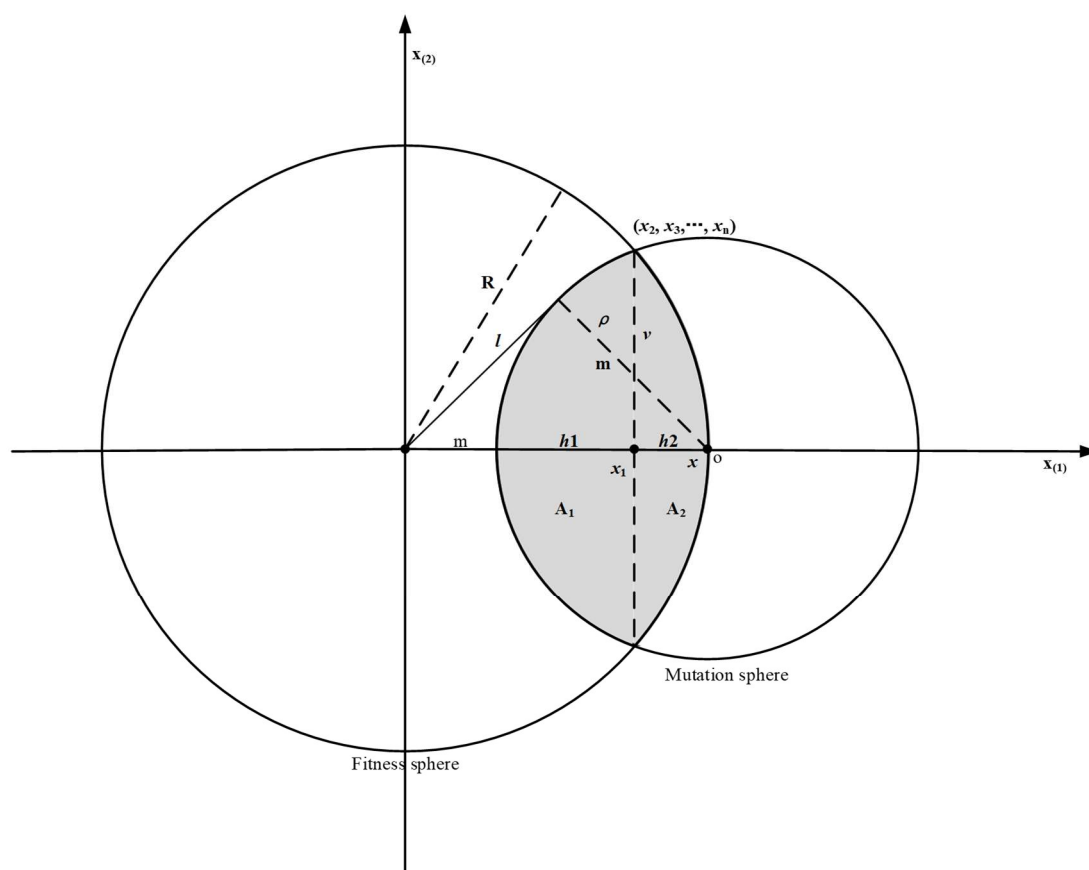
### 3. Sphere function and uniform distribution

In this section, we prove that there is an exponential lower bound on the expected first hitting time when the (1 + 1) ES with a uniform mutation operator is used to solve the sphere function minimization problem, which will be given in Theorem 4.

**Theorem 4.** In the running of the (1 + 1) ES with elitist selection, when the radius of the sphere function of the  $n$ -dimensional Euclidean space goes from fitness difference function value  $R/2$  to  $\varepsilon$ , then the lower bound of the first hitting time  $T_r$  is satisfied:

$$E(T_r) > \Omega(e^{cn}). \quad (6)$$

Due to the symmetry of the sphere, we assume that the current individual is located on the  $ox$  axis, the distance from the center of the fitness hypersphere is  $R$ , the state of the individual is  $o$ , and the  $n$  dimensional sphere with  $o$  as the center and  $\rho$  as the radius is the mutation sphere  $E_1$ , which contains the next generation of individuals that are likely to reach the set. At the same time, fitness hypersphere  $E_2$  contains all possible solution space sets of the next-generation individual better than the current individual, as shown in Figure 1. However, during the running of (1 + 1) ES, not all the next-generation individuals generated by the mutation operator are superior to the previous-generation individuals. Therefore, we regard the fitness spherical cap  $A_2$  and mutation spherical cap  $A_1$  as the reached region  $S_0 = E_1 \cap E_2 = A_1 \cup A_2$  by the next generation individuals, and the two spherical caps can be regarded as the cut-off of the same hyperplane  $P$ , i.e., the success region. The region is irregular and the calculation of the volume is relatively complicated, so we will calculate the volume of the spherical cap by means of the spherical sector and the spherical cone. Next, we will give some important conclusions.



**Figure 1.** Success region of elitist ES on sphere.

### 3.1. The uniform distribution inside the sphere

Suppose that in the  $n$ -dimensional sphere, the uniform mutation operator is uniformly distributed in the sphere with radius  $\rho$ , then, its corresponding volume is denoted as

$$Vol_n(\rho) = \frac{\pi^{\frac{n}{2}} \rho^n}{\Gamma(\frac{n}{2} + 1)}, D_n(\rho) = \{\mathbf{x} \in R^n, \|\mathbf{x}\| < \rho\}, \quad (7)$$

and the probability density function of the sphere with radius  $\rho$  is

$$f(\mathbf{x}) = \frac{1}{Vol_n(\rho)} = \frac{\Gamma(\frac{n}{2} + 1)}{\pi^{\frac{n}{2}} \rho^n} I_{D_n(\rho)}(\mathbf{x}), \quad (8)$$

where  $I_{D_n(\rho)}(\mathbf{x})$  is the indicator function. When the uniform mutation operator is evenly distributed in the radius  $\rho$  sphere, the corresponding value is 1; otherwise, the value is 0. For more details, see [31].

**Lemma 1.** Let the radius of the hypersphere containing the spherical cap be  $r$  and its height be  $h$ . Then, the volume of the spherical sector is

$$Vol_n^s(r) = \frac{1}{n} S_{cap} r, \quad (9)$$

where  $S_{cap}$  is the area of the spherical cap and  $n$  represents the dimension of the hypersphere.

**Lemma 2.** Let the radius of the hypersphere containing the spherical cap be  $r$  and its height be  $h$ ,  $n > 3$ . Then, the area of the spherical cap is

$$S_{cap} = \begin{cases} \frac{2\pi^{\frac{n-1}{2}} r^{n-2} (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!} \left[ h - (r-h) \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} \left( \frac{2h}{r} - \frac{h^2}{r^2} \right)^m \right], & n \in \text{even}; \\ \left( \pi^{\frac{n+1}{2}} r^{n-1} - \pi^{\frac{n-1}{2}} r^{n-1} \arccos\left(\frac{\sqrt{2rh-h^2}}{r}\right) - \pi^{\frac{n-1}{2}} r^{n-2} (r-h) \times \right. \\ \left. \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left( \frac{\sqrt{2rh-h^2}}{r} \right)^{2m-1} \right) \frac{2}{\Gamma(\frac{n-1}{2}) (n-2)!!}, & n \in \text{odd}. \end{cases} \quad (10)$$

**Lemma 3.** Let the radius of the bottom surface of the spherical cap be  $r_{cap}$  and its height be  $h$ . Then, the volume of the spherical cone is

$$Vol_n^c(r_{cap}) = \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (r_{cap})^{n-1} (r-h). \quad (11)$$

**Theorem 5.** Let the radius of the hypersphere containing the spherical cap be  $r$ , its radius of the bottom surface be  $r_{cap}$ , and its height be  $h$ . Then, the volume of the spherical cap is

$$Vol = Vol_n^s(r) - Vol_n^c(r_{cap}) = \frac{1}{n} S_{cap} r - \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (r_{cap})^{n-1} (r-h). \quad (12)$$

To calculate the volume of the fitness spherical cap, it is necessary to calculate not only the area  $S_{fcap}$  of the fitness spherical cap, but also the bottom radius  $r_{fcap}$ , its heights  $h_2$  and  $h_1$ , and the simultaneous equations:

$$\begin{cases} \sqrt{\sum_{i=1}^n x_i^2} = R; \\ \sqrt{(x_1 - R)^2 + \sum_{i=2}^n x_i^2} = \rho. \end{cases} \quad (13)$$

We get it by solving

$$x_1 = R \left( 1 - \frac{1}{2} \left( \frac{\rho}{R} \right)^2 \right), \quad (14)$$



where  $x_1$  denotes the abscissa of the common hyperplane of the fitness and mutation spherical cap. Let the heights of the corresponding spherical caps be  $h_2$  and  $h_1$ . The association with Eq (14) yields

$$\begin{cases} h_2 = R - x_1 = \frac{\rho^2}{2R}; \\ h_1 = \rho - (R - x_1) = \rho(1 - \frac{\rho}{2R}). \end{cases} \quad (15)$$

From Eq (15), we can obtain  $h_2$  and  $h_1$ . Next, we compute the bottom radius of the common hyperplane of the fitness and mutation spherical cap, and it is useful to assume that the height of the spherical cap is  $0 < h_1, h_2 < R$ . Then, the bottom radius  $r_{fcap}$  is

$$r_{cap} = \sqrt{\rho^2 - (R - x_1)^2} = \rho \sqrt{1 - \frac{\rho^2}{4R^2}}. \quad (16)$$

### 3.2. The volume of the fitness spherical cap

Before calculating the volume of the fitness spherical cap, its area is calculated without loss of generality. Assume that the radius of the mutation sphere where the fitness sphere cap intersects with the mutation sphere is  $\rho$ , the radius of the fitness sphere is  $R$ , its heights are  $h_2 < R$  and  $n > 3$ , and the bottom radius is  $r_{cap}$ ; then the volume of the fitness spherical cap is discussed in two cases:

i) When  $n$  is odd:

$$S_{fcap} = \frac{2\pi^{\frac{n-1}{2}} R^{n-2} (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!} \left[ \frac{\rho^2}{2R} - (R - \frac{\rho^2}{2R}) \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} (\frac{\rho^2}{R^2} - \frac{\rho^4}{4R^4})^m \right]. \quad (17)$$

We substitute Eqs (14)–(16) into  $Vol = \frac{1}{n} S_{fcap} R - \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (r_{cap})^{n-1} (R - h_2)$  to obtain the

volume of the fitness spherical cap as:

$$\begin{aligned} Vol_n^{fcap} &= \frac{1}{n} \frac{2\pi^{\frac{n-1}{2}} R^{n-2} (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!} \left[ \frac{\rho^2}{2R} - (R - \frac{\rho^2}{2R}) \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} (\frac{\rho^2}{R^2} - \frac{\rho^4}{4R^4})^m \right] R - \\ &\frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (\sqrt{\rho^2 - \frac{\rho^4}{4R^2}})^{n-1} (R - \frac{\rho^2}{2R}) \\ &< \frac{1}{n} \frac{\pi^{\frac{n-1}{2}} \rho^2 R^{n-2} (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!}. \end{aligned} \quad (18)$$

ii) When  $n$  is even:

$$S_{fcap} = \left[ \pi^{\frac{n+1}{2}} R^{n-1} - \pi^{\frac{n-1}{2}} R^{n-1} \arccos\left(\sqrt{\frac{\rho^2 - \rho^4}{4R^2}}\right) - \pi^{\frac{n-1}{2}} R^{n-2} \left(\frac{2R - \rho^2}{2R^2}\right) \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left(\sqrt{\frac{\rho^2 - \rho^4}{4R^2}}\right)^{2m-1} \right] \frac{2}{\Gamma\left(\frac{n-1}{2}\right)} \frac{(n-3)!!}{(n-2)!!}.$$

Similarly, we substitute Eqs (14)–(16) into Eq (12) to obtain the volume of the fitness spherical cap as:

$$\begin{aligned} Vol_n^{fcap} &= \frac{1}{n} \frac{2}{\Gamma\left(\frac{n-1}{2}\right)} \frac{(n-3)!!}{(n-2)!!} \left[ \pi^{\frac{n+1}{2}} R^{n-1} - \pi^{\frac{n-1}{2}} R^{n-1} \arccos\left(\sqrt{\frac{\rho^2 - \rho^4}{4R^2}}\right) - \right. \\ &\quad \left. \pi^{\frac{n-1}{2}} R^{n-2} \left(\frac{2R - \rho^2}{2R^2}\right) \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left(\sqrt{\frac{\rho^2 - \rho^4}{4R^2}}\right)^{2m-1} \right] R - \\ &\quad \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2} + 1\right)} \left(\sqrt{\rho^2 - \frac{\rho^4}{4R^2}}\right)^{n-1} \left(R - \frac{\rho^2}{2R}\right) \\ &< \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}} R^n (n-3)!!}{\Gamma\left(\frac{n-1}{2}\right) (n-2)!!}. \end{aligned} \quad (19)$$

### 3.3. The volume of the mutation spherical cap

Assume that the radius of the fitness sphere where the mutation spherical cap intersects with the fitness sphere is  $R$ , the radius of the mutation sphere is  $\rho$ , its heights are  $h_1 < \rho$  and  $n > 3$ , and the bottom radius is  $r_{cap}$ . Then the volume of the mutation spherical cap is discussed in two cases:

i) When  $n$  is odd:

$$\begin{aligned} S_{mcap} &= \frac{2\pi^{\frac{n-1}{2}} \rho^{n-2} (n-3)!!}{\Gamma\left(\frac{n-1}{2}\right) (n-2)!!} \left[ \rho - \frac{\rho^2}{2R} - \left(\rho - \rho + \frac{\rho^2}{2R}\right) \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} \left(\frac{2(\rho - \frac{\rho^2}{2R})}{\rho} - \frac{(\rho - \frac{\rho^2}{2R})^2}{\rho^2}\right)^m \right] \\ &= \frac{2\pi^{\frac{n-1}{2}} \rho^{n-2} (n-3)!!}{\Gamma\left(\frac{n-1}{2}\right) (n-2)!!} \left[ \rho - \frac{\rho^2}{2R} - \frac{\rho^2}{2R} \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} \left(1 - \frac{\rho^2}{4R^2}\right)^m \right]. \end{aligned} \quad (20)$$

Inserting Eqs (15), (16), and (20) in  $Vol = \frac{1}{n} S_{mcap} \rho - \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2} + 1\right)} (r_{cap})^{n-1} (\rho - h_1)$ , we obtain the

volume of the mutation spherical cap as:

$$\begin{aligned}
Vol_n^{mcap} &= \frac{1}{n} \frac{2\pi^{\frac{n-1}{2}} \rho^{n-2}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left[ \rho - \frac{\rho^2}{2R} - \frac{\rho^2}{2R} \sum_{m=1}^{\frac{n-3}{2}} \frac{(2m-1)!!}{(2m)!!} \left(1 - \frac{\rho^2}{4R^2}\right)^m \right] \rho - \\
&\frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (\rho \sqrt{1 - \frac{\rho^2}{4R^2}})^{n-1} (\rho - (\rho - \frac{\rho^2}{2R})) \\
&< \frac{1}{n} \frac{2\pi^{\frac{n-1}{2}} \rho^{n-1}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} (\rho - \frac{\rho^2}{2R}).
\end{aligned} \tag{21}$$

ii) When  $n$  is even:

$$\begin{aligned}
S_{mcap} &= \frac{2}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left[ \pi^{\frac{n+1}{2}} \rho^{n-1} - \pi^{\frac{n-1}{2}} \rho^{n-1} \arccos\left(\frac{\sqrt{2\rho(\rho - \frac{\rho^2}{2R}) - (\rho - \frac{\rho^2}{2R})^2}}{\rho}\right) - \right. \\
&\left. \pi^{\frac{n-1}{2}} \rho^{n-2} (\rho - (\rho - \frac{\rho^2}{2R})) \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left(\frac{\sqrt{2\rho(\rho - \frac{\rho^2}{2R}) - (\rho - \frac{\rho^2}{2R})^2}}{\rho}\right)^{2m-1} \right] \\
&= \frac{2}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left[ \pi^{\frac{n+1}{2}} \rho^{n-1} - \pi^{\frac{n-1}{2}} \rho^{n-1} \arccos\left(\frac{\sqrt{4R^2 - 4\rho R - \rho^2}}{2R}\right) - \right. \\
&\left. \pi^{\frac{n-1}{2}} \frac{\rho^n}{2R} \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left(\frac{\sqrt{4R^2 - 4\rho R - \rho^2}}{2R}\right)^{2m-1} \right].
\end{aligned} \tag{22}$$

Inserting Eqs (15), (16), and (22) in  $Vol = \frac{1}{n} S_{mcap} \rho - \frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (r_{cap})^{n-1} (\rho - h_1)$ , we obtain the

volume of the mutation spherical cap as:

$$\begin{aligned}
Vol_n^{mcap} &= \frac{1}{n} \frac{2}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left[ \pi^{\frac{n+1}{2}} \rho^{n-1} - \pi^{\frac{n-1}{2}} \rho^{n-1} \arccos\left(\frac{\sqrt{4R^2 - 4\rho R - \rho^2}}{2R}\right) - \right. \\
&\left. \pi^{\frac{n-1}{2}} \frac{\rho^n}{2R} \sum_{m=1}^{\frac{n-2}{2}} \frac{(2m-2)!!}{(2m-1)!!} \left(\frac{\sqrt{4R^2 - 4\rho R - \rho^2}}{2R}\right)^{2m-1} \right] \rho - \\
&\frac{1}{n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2} + 1)} (\rho \sqrt{1 - \frac{\rho^2}{4R^2}})^{n-1} (\rho - (\rho - \frac{\rho^2}{2R})) \\
&< \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}} \rho^n}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!}.
\end{aligned} \tag{23}$$

Since the expected progress rate probability density function is not easy to accurately calculate [32], it is difficult to calculate the expected first hitting time of the (1 + 1) ES on the sphere function. Therefore, we estimate by calculating the sum of the volume of the mutation spherical cap and the volume of the fitness spherical cap, which provides a theoretical basis for improving the estimation of the lower bound of the first hitting time of (1 + 1) ES on the sphere function.

### 3.4. Expected first hitting time of (1 + 1) ES

The position of the next generation of individuals generated by (1 + 1) ES during the running of the sphere function is the gray part of Figure 1, i.e., the intersection of the irregular successful regions is the union of spherical caps  $A_1$  and  $A_2$  :

$$\begin{cases} D = \{(x, v) \in A_1 \cup A_2\} \\ A_2 = \{(x, v) \mid R - \rho^2 / 2R \leq x \leq \rho, 0 \leq v \leq \rho^2 - x^2\} . \\ A_1 = \{(x, v) \mid 0 \leq x \leq R - \rho^2 / 2R, 0 \leq v \leq 2Rx - x^2\} \end{cases}$$

By Definition 2, we get the expected progress rate:

$$\begin{aligned} E(X_t) &= E(d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)) \\ &= E(R - d(\xi_{opt}^t)) \\ &= \int_{\xi_{opt}^t \in D} f(\mathbf{x})(R - d(\xi_{opt}^t)) d\xi_{opt}^t \\ &= \iint_D Rf(x, v) dv dx - \iint_D f(x, v) \sqrt{(R-x)^2 + \|v\|^2} dv dx \\ &= \frac{\Gamma(\frac{n}{2} + 1)R}{\pi^{\frac{n}{2}} \rho^n} \iint_D 1 dv dx - \iint_D \frac{\Gamma(\frac{n}{2} + 1)}{\pi^{\frac{n}{2}} \rho^n} \sqrt{(R-x)^2 + \|v\|^2} dv dx. \end{aligned} \quad (24)$$

i) When  $n$  is even:

$$\begin{aligned} Prob_{A_1 \cup A_2} &= Vol_n^{fcap} + Vol_n^{mcap} \\ &= \frac{1}{n} \frac{\pi^{\frac{n-1}{2}} \rho^2 R^{n-2}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} + \frac{1}{n} \frac{2\pi^{\frac{n-1}{2}} \rho^{n-1}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left(\rho - \frac{\rho^2}{2R}\right) \\ &= \frac{1}{n} \frac{\pi^{\frac{n-1}{2}} \rho^2 (R^{n-1} + 2\rho^{n-2} - \rho^{n-1})}{R\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!}. \end{aligned} \quad (25)$$

According to Eqs (24) and (25), we can obtain:

$$\begin{aligned}
E(X_t) &= \frac{\Gamma(\frac{n}{2}+1)R}{\pi^{\frac{n}{2}}\rho^n} \iint_D 1 \, dvdx - \iint_D \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx \\
&< \frac{1}{n} \frac{\Gamma(\frac{n}{2}+1)\pi^{\frac{1}{2}}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left(\frac{R^{n-1}}{\rho^{n-2}} + 2 - \rho\right) - \\
&\frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \left( \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\rho^2-x^2} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx + \right. \\
&\left. \int_0^{R-\frac{\rho^2}{2R}} \int_0^{2Rx-x^2} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx \right) \\
&= \frac{1}{n} \frac{\Gamma(\frac{n}{2}+1)\pi^{\frac{1}{2}}}{\Gamma(\frac{n-1}{2})} \frac{(n-3)!!}{(n-2)!!} \left(\frac{R^{n-1}}{\rho^{n-2}} + 2 - \rho\right) - \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} (G_1 + G_2).
\end{aligned} \tag{26}$$

In order to calculate  $G_1$ , we reduce the dimension of the integral to make  $w = \|v\|^2$ . Thus, we obtain:

$$\begin{aligned}
G_1 &= \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\rho^2-x^2} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx \\
&= \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\rho^2-x^2} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} w^{\frac{n-1}{2}-1} \sqrt{(R-x)^2 + w} \, dw dx.
\end{aligned} \tag{27}$$

Because the integrand function of this integral contains the radical  $\sqrt{(R-x)^2 + w}$ , it is difficult to solve the integral. By taking into account the idea of substitution that regards  $(R-x)^2$  as the common factor of  $(R-x)^2 + w$  under the radical sign, where  $w = u(R-x)^2$ , then  $dw = (R-x)^2 du$ ,

$$\sqrt{(R-x)^2 + w} = (R-x)\sqrt{1+u} \quad \text{and} \quad \sqrt{1+u} = 1 + \frac{1}{2}u + \frac{1}{8}u^2 + \dots + \frac{\frac{1}{2}(-\frac{1}{2})\dots(\frac{1}{2}-m+1)}{m!} u^m + R_m(u).$$

When  $w = 0$ ,  $u = 0$ ; when  $w = \rho^2 - x^2$ ,  $u = \frac{\rho^2 - x^2}{(R-x)^2}$ . Hence, we obtain:

$$\begin{aligned}
G_1 &= \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\frac{\rho^2-x^2}{(R-x)^2}} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} u^{\frac{n-1}{2}-1} (R-x)^{n-3} \sqrt{(R-x)^2 + u(R-x)^2} (R-x)^2 du dx \\
&= \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\frac{\rho^2-x^2}{(R-x)^2}} u^{\frac{n-1}{2}-1} (R-x)^n \sqrt{1+u} du dx \\
&\approx \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} \int_{R-\frac{\rho^2}{2R}}^{\rho} \int_0^{\frac{\rho^2-x^2}{(R-x)^2}} u^{\frac{n-1}{2}-1} (R-x)^n (1 + \frac{1}{2}u + \frac{1}{8}u^2) du dx \\
&= \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} \frac{(\rho^2-x^2)^{\frac{n-1}{2}}}{(R-x)^{n-1}} \left[ \frac{2}{n-1} + \frac{1}{n+1} \frac{\rho^2-x^2}{(R-x)^2} \right. \\
&\quad \left. + \frac{1}{4(n+3)} \frac{(\rho^2-x^2)^2}{(R-x)^4} \right] \times \frac{1}{n+1} [(R-\rho + \frac{\rho^2}{2R})^{n+1} - (R-\rho)^{n+1}].
\end{aligned} \tag{28}$$

Similarly,

$$\begin{aligned}
G_2 &= \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} \frac{(2Rx-x^2)^{\frac{n-1}{2}}}{(R-x)^{n-1}} \left[ \frac{2}{n-1} + \frac{1}{n+1} \frac{2Rx-x^2}{(R-x)^2} \right. \\
&\quad \left. + \frac{1}{4(n+3)} \frac{(2Rx-x^2)^2}{(R-x)^4} \right] \times \frac{1}{n+1} [R^{n+1} - (\frac{\rho^2}{2R})^{n+1}].
\end{aligned} \tag{29}$$

According to Eqs (26)–(29), we can draw the following conclusions:

$$E(X_i) < \Omega\left(\left(\frac{R}{\rho}\right)^{n-2}\right).$$

ii) When  $n$  is odd:

$$\begin{aligned}
Prob_{A_1 \cup A_2} &= Vol_n^{fcap} + Vol_n^{mcap} \\
&= \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}} R^n (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!} + \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}} \rho^n (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!} \\
&= \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}} (R^n + \rho^n) (n-3)!!}{\Gamma(\frac{n-1}{2}) (n-2)!!}.
\end{aligned} \tag{30}$$

By Eqs (24) and (30), we can draw the following conclusions:

$$\begin{aligned}
E(X_t) &= \frac{\Gamma(\frac{n}{2}+1)R}{\pi^{\frac{n}{2}}\rho^n} \iint_D 1 \, dvdx - \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \iint_D \sqrt{(R-x)^2 + \|v\|^2} \, dvdx \\
&< \frac{\Gamma(\frac{n}{2}+1)R}{\pi^{\frac{n}{2}}\rho^n} \frac{1}{n} \frac{2\pi^{\frac{n+1}{2}}(R^n + \rho^n)(n-3)!!}{\Gamma(\frac{n-1}{2})(n-2)!!} \\
&\quad \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \int_0^{R-\frac{\rho^2}{2R}} \int_0^{2Rx-x^2} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx \\
&= \frac{1}{n} \frac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{n-1}{2})} R \left( \left(\frac{R}{\rho}\right)^n + 1 \right) \frac{(n-3)!!}{(n-2)!!} - \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \int_0^{R-\frac{\rho^2}{2R}} \int_0^{2Rx-x^2} \sqrt{(R-x)^2 + \|v\|^2} \, dvdx.
\end{aligned} \tag{31}$$

Also, from Eqs (27)–(29), the following asymptotic expressions can be approximated:

$$\begin{aligned}
E(X_t) &< \frac{1}{n} \frac{\Gamma(\frac{n}{2}+1)}{\Gamma(\frac{n-1}{2})} R \left( \left(\frac{R}{\rho}\right)^n + 1 \right) \frac{(n-3)!!}{(n-2)!!} - \frac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\rho^n} \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})} \frac{(\rho^2 - x^2)^{\frac{n-1}{2}}}{(R-x)^{n-1}} \left( \frac{2}{n-1} + \right. \\
&\quad \left. \frac{1}{n+1} \frac{\rho^2 - x^2}{(R-x)^2} + \frac{1}{4(n+3)} \frac{(\rho^2 - x^2)^2}{(R-x)^4} \right) \times \frac{1}{n+1} \left( \left(R - \rho + \frac{\rho^2}{2R}\right)^{n+1} - (R - \rho)^{n+1} \right) + \\
&\quad \frac{(2Rx - x^2)^{\frac{n-1}{2}}}{(R-x)^{n-1}} \left( \frac{2}{n-1} + \frac{1}{n+1} \frac{2Rx - x^2}{(R-x)^2} + \frac{1}{4(n+3)} \frac{(2Rx - x^2)^2}{(R-x)^4} \right) \times \frac{1}{n+1} \left( R^{n+1} - \left(\frac{\rho^2}{2R}\right)^{n+1} \right) \\
&= O\left(\left(\frac{R}{\rho}\right)^n\right).
\end{aligned} \tag{32}$$

Obviously, when the dimension of the sphere function is odd or even, we find that the upper bound on the expected progress rate is exponential. By Theorem 3, the expected first hitting time  $E(T_\tau)$  from the fitness difference function value  $R$  to  $R/2 - \varepsilon$  in the running of the sphere function with a uniform mutation operator is:

$$E(T_\tau) > \frac{R - R/2 + \varepsilon}{O\left(\left(\frac{R}{\rho}\right)^n\right)} = \Omega\left(\left(\frac{\rho}{R}\right)^n\right).$$

Let  $c = \ln\left(\frac{\rho}{R}\right)$ , thus  $\left(\frac{\rho}{R}\right)^n = e^c$ , then we can get that  $E(T_\tau) > \Omega(e^{cn})$  is the expected first hitting time of the elitist selection (1 + 1) ES with a uniform mutation operator solving sphere function. Therefore, Theorem 4 holds.

Theorem 4 has significantly improved the current lower bound  $\Omega(n)$ , which has a significant effect on reducing the time complexity of analyzing the 1/5 success rule of (1 + 1) ES.

#### 4. The sphere function and Gaussian mutation operator

Without loss of generality, the Gaussian mutation operator is different from the uniform mutation operator. It is assumed that independent offspring individuals are generated through the mutation vector  $\mathbf{m} = \sigma_n N(0, I)$  and obey Gaussian distribution (see Figure 1).

**Theorem 6.** Let  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  be an individual from the population  $\sigma_n N(0, I)$ . Then

$$k^2 = x_1^2 + x_2^2 + \dots + x_n^2, \quad (33)$$

which obeys the chi-square distribution with degree of freedom  $n$ , and is denoted as  $k^2 \sim \chi^2(n)$ . The probability density function of  $k = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$  is

$$p(y) = \frac{2y^{n-1}}{\sigma^n 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \exp\left(-\frac{y^2}{2\sigma^2}\right). \quad (34)$$

These results are a simplified form of [32]. For a more detailed proof, see [33]. Its mathematical expectations are as follows:

$$k := E(y) = \int_0^{\infty} y p(y) dy = \int_0^{\infty} y \frac{2y^{n-1}}{\sigma^n 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \sqrt{2}\sigma. \quad (35)$$

When  $n \rightarrow \infty$ , then

$$k = \lim_{n \rightarrow \infty} E(y) = \sqrt{n}\sigma. \quad (36)$$

The validity of relation  $l^2 = k^2 - m^2 + (R - m)^2$  is evident based on the observations from Figure 1. Since offspring individuals  $\xi_{opt}^t$  are generated by the mutation vector  $\mathbf{m} = \sigma_n N(0, I)$  and  $d(\xi_{opt}^{t-1})$  is a constant,  $d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)$  is a random variable that depends on the mutation vector, and the specific form of  $d(\xi_{opt}^t)$  is as follows [33]:

$$\begin{aligned} d(\xi_{opt}^t) &= \sqrt{k^2 - m^2 + (R - m)^2} \\ &= R \sqrt{\frac{m^2}{R} + 1 - \frac{2m}{R}}. \end{aligned} \quad (37)$$

The  $d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)$  probability density function is

$$p(m) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{m^2}{2\sigma^2}\right). \quad (38)$$

Thus, we can get the expected progress rate  $E(X_t)$ :



$$\begin{aligned}
E(X_t) &= E(d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)) \\
&= E(R - d(\xi_{opt}^t)) \\
&= \int_{m1}^{m2} (R - R\sqrt{\frac{m^2}{R} + 1 - \frac{2m}{R}}) \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{m^2}{2\sigma^2}) dm,
\end{aligned} \tag{39}$$

where  $m1 = \frac{k^2}{2R}$  and  $m2 = k = \sqrt{n}\sigma$ .

By normalizing the expected progress rate  $E(X_t)$  and the mutation strength  $\sigma$  [32]:

$$E^*(X_t) = \frac{n}{R} E(X_t), \sigma^* = \frac{n}{R} \sigma. \tag{40}$$

And we can get the normalized expected progress rate  $E^*(X_t)$ :

$$\begin{aligned}
E^*(X_t) &= E^*(d(\xi_{opt}^{t-1}) - d(\xi_{opt}^t)) \\
&= \int_{\sigma^*/2}^{1/\sqrt{n}} \frac{\sigma^*}{\sqrt{2\pi}} (t - \frac{\sigma^*}{2}) \exp(-\frac{t^2}{2}) dt \\
&\approx \frac{\sigma^*}{\sqrt{2\pi}} \int_{\sigma^*/2}^{1/\sqrt{n}} (t - \frac{\sigma^*}{2}) (\frac{t^2}{2} - 1 + \frac{t^4}{8}) dt \\
&> \frac{\sigma^*}{\sqrt{2\pi}} ((\frac{\sigma^*}{80n^2\sqrt{n}} + \frac{1}{8n^2} + \frac{\sigma^*}{2\sqrt{n}}) - (\frac{1}{40}(\frac{\sigma^*}{2})^6 + \frac{1}{8}(\frac{\sigma^*}{2})^4 + (\frac{\sigma^*}{2})^2)) \\
&= \Omega(\frac{1}{n^2\sqrt{n}}).
\end{aligned} \tag{41}$$

Since (1 + 1) ES is adopted as the elitist selection strategy,  $\frac{R}{2} - \varepsilon < E^*(X_t) < E^*(X_0) < R$  holds for  $\forall t \in (0, \tau)$ . By Theorem 3, when running (1 + 1) ES on the sphere function, the first hitting time  $E(T_\tau)$  from the fitness difference function value  $\frac{R}{2}$  to  $\varepsilon$  with a Gaussian mutation operator holds:

$$E(T_\tau) < \frac{R - (\frac{R}{2} - \varepsilon)}{\Omega(\frac{1}{n^2\sqrt{n}})} = O(n^2\sqrt{n}), \tag{42}$$

which implies that we obtain the upper bound of the first hitting time. The above analysis shows that we can consider that each individual follows an independent isotropic distribution. Jägerskupper [24] adopted a similar approach to analyzing the first hitting time of (1 + 1) ES on the sphere function and obtained the lower bound of the first hitting time, while the upper bound was not given the relevant theoretical analysis. It is shown that the proposed method is valuable for the first hitting time analysis of population-based algorithms.

## 5. Numerical experiments and result analysis

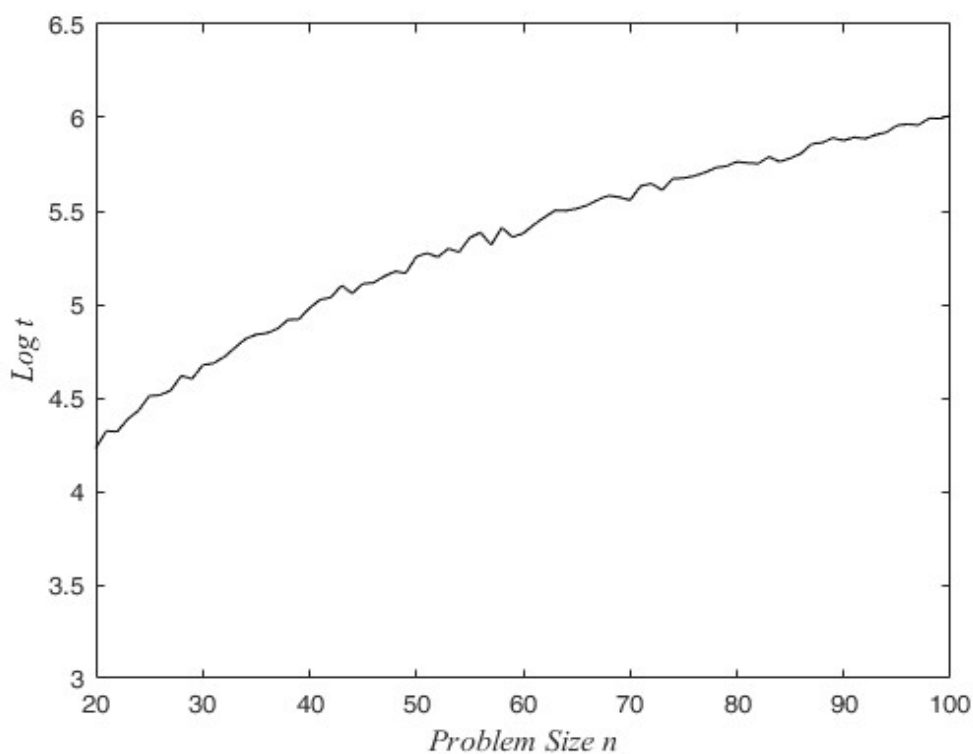
The purpose of the experiments in this section is to verify the upper and lower bounds on the theoretical accuracy of the expected first hitting time for  $(1 + 1)$  ES with uniform and Gaussian mutation operators solving sphere functions.

### 5.1. Experimental parameter setting

In this section, we mainly set the following experimental parameters:  $R = d(\xi_0) = (x_1, x_2, \dots, x_n) = (5, 0, \dots, 0) = 5$ ,  $\rho = 1$ ,  $\varepsilon = 0.001$ , and  $n \in [1, 100]$ . For each given problem size  $n$ , we conduct 100 runs for Algorithm 1 on the sphere function. We define  $T_{ri}$  to denote the  $i$ -th run first hitting time of Algorithm 1 and  $E(T_r) = \frac{\sum_{i=1}^{100} T_{ri}}{100}$  to denote the actual running average first hitting time, i.e.,  $E(T_r) = t$ .

### 5.2. $(1 + 1)$ ES with a uniform mutation operator

Theorem 4 gives the expected first hitting time of  $(1 + 1)$  ES with a uniform mutation operator from the previous  $\Omega(n)$  to  $\Omega(e^n)$ . We arbitrarily take  $n \in [20, 100]$  to verify that its lower bound is exponentially related to the problem size  $n$ . To verify the exponential time of the first hitting time, we only need to verify the linear relationship between  $\log t$  and the problem size  $n$ , thus verifying our theoretical results. MATLAB simulation results are shown in Figure 2.

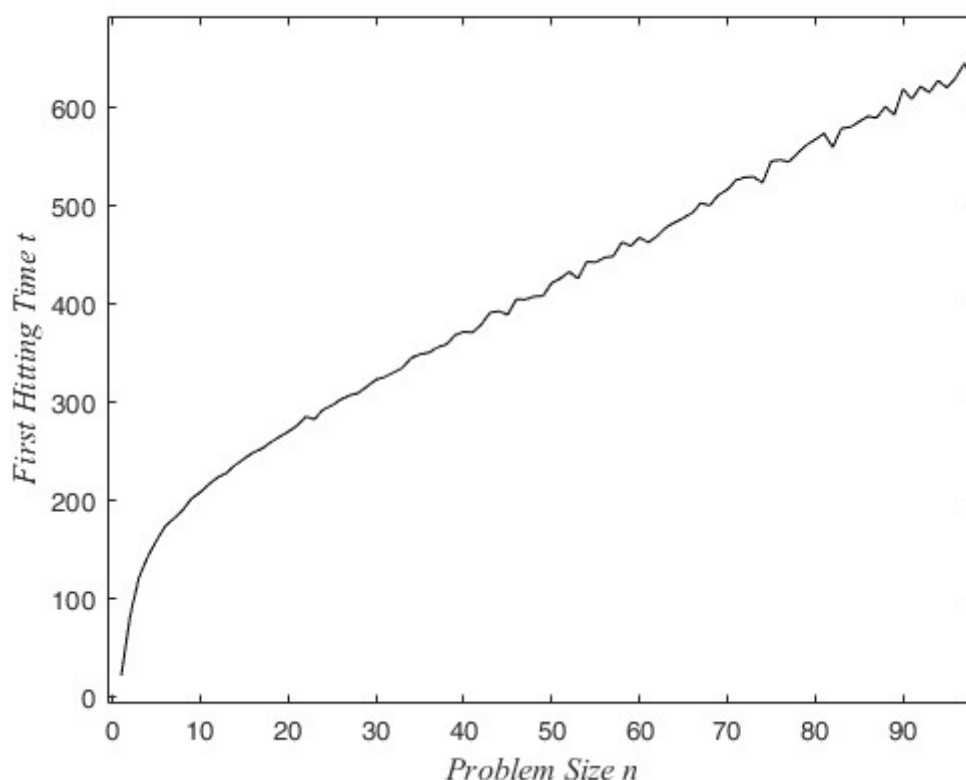


**Figure 2.** Problem size  $n$  and  $\log t$  of  $(1 + 1)$  ES with a uniform mutation operator.

We can see from Figure 2 that  $\log t$  increases with increasing problem size  $n$ , showing a linear relationship. Consequently, we have derived a tighter lower bound for the time complexity, which is consistent with the theoretical lower bound for the analysis time.

### 5.3. $(1 + 1)$ ES with a Gaussian mutation operator

In Section 4,  $O(n^2\sqrt{n})$  is given as the upper bound of the expected first hitting time of  $(1 + 1)$  ES with a Gaussian mutation operator. The first hitting time of a polynomial is acceptable in real-world optimization problems. MATLAB simulation results are shown in Figure 3.



**Figure 3.** Problem size  $n$  and the first hitting time of  $(1 + 1)$  ES with a Gaussian mutation operator.

We can see from Figure 3 that when about  $n < 20$ , the first hitting time increases as problem size  $n$  increases, which is faster than that of about  $n > 20$ . We think that the first hitting time is a polynomial when  $n$  is large enough.

## 6. Conclusions

In this paper, it is difficult to estimate the upper and lower bounds of  $(1 + 1)$  ES in the first time-solving sphere function, and the renewal model is proposed to solve it. The proposed method does not aim to directly determine the first hitting time, but views the individual population as a special renewal process, by introducing the stopping time of the special renewal process and combining Wald's inequality with the renewal theorem. The model is independent of the specific implementation of the

algorithm and presents a gradual convergence process of EAs on a continuous search space, showing the gradual convergence process of  $(1 + 1)$  ES in a continuous search space. To verify the validity of the proposed method, we analyze the expected first hitting time of  $(1 + 1)$  ES on the sphere function. We obtain the lower bound of exponential time  $\Omega(e^{cn})$  with a uniform operator  $(1 + 1)$  ES and the upper bound of polynomial time  $O(n^2\sqrt{n})$  with a uniform operator  $(1 + 1)$  ES. Furthermore, the proposed method can be applicable to other intelligent optimization algorithms, such as differential evolution algorithm and particle swarm optimization.

This will also be further investigated in our future work. Based on the progress rate, statistical methods will be used to estimate the first hitting time of EAs, especially in the running time analysis with recombination, which will be very challenging. In addition, based on the discussed upper bound on the running time of  $(1 + 1)$  ES with a Gaussian mutation operator, another direction for further research is to analyze the lower bound on the running time of  $(1 + 1)$  ES with a Gaussian mutation operator. A closed interval can be constructed by upper and lower bounds to estimate the first hitting time of EAs, and then the theoretical results can be extended to the practical application of EAs.

### Use of AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

### Acknowledgments

The authors extend their sincere gratitude to the anonymous reviewers for their helpful comments and insightful suggestions. This work was supported by the Science and Technology Plan Project of Guizhou Province (QKH-Foundation-ZK[2022] General 550, QKH-Platform Talent-ZCKJ[2021]007), the Youth Science and Technology Talent Growth Project of the Guizhou Provincial Education Department (QJH-KY[2021]115, QJH-KY[2022]177), the Open Project of the Guizhou Key Laboratory of Pattern Recognition and Intelligent System (GZMUKL[2022]KF05), the Ministry of Education's Industry-University Collaborative Education Project (221001766110209), and the Research Fund Project of the Guiyang Institute of Humanities and Technology (2022RWJS034).

### Conflict of interest

The authors declare that there are no conflicts of interest.

### References

1. T. Bäck, H. P. Schwefel, An overview of evolutionary algorithms for parameter optimization, *Evol. Comput.*, **1** (1993), 1–23. <https://doi.org/10.1162/evco.1993.1.1.1>
2. J. He, X. Yao, Average drift analysis and population scalability, *IEEE Trans. Evol. Comput.*, **21** (2016), 426–439. <https://doi.org/10.1109/tevc.2016.2608420>
3. X. Yao, Unpacking and understanding evolutionary algorithm, in *IEEE World Congress on Computational Intelligence*, Berlin, Springer Berlin Heidelberg, (2012), 60–76. [https://doi.org/10.1007/978-3-642-30687-7\\_4](https://doi.org/10.1007/978-3-642-30687-7_4)
4. J. He, X. Yao, Towards an analytic framework for analysing the computation time of evolutionary algorithms, *Artif. Intell.*, **145** (2003), 59–97. [https://doi.org/10.1016/s0004-3702\(02\)00381-8](https://doi.org/10.1016/s0004-3702(02)00381-8)

5. C. Witt, Tight bounds on the optimization time of a randomized search heuristic on linear functions, *Comb. Probab. Comput.*, **22** (2013), 294–318. <https://doi.org/10.1017/s0963548312000600>
6. B. Doerr, T. Kötzing, J. A. G. Lagodzinski, J. Lengler, The impact of lexicographic parsimony pressure for ORDER/MAJORIT on the run time, *Theor. Comput. Sci.*, **816** (2020), 144–168. <https://doi.org/10.1016/j.tcs.2020.01.011>
7. B. Doerr, C. Doerr, J. Yang, Optimal parameter choices via precise black-box analysis, in *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, New York: ACM, (2016), 1123–1130. <https://doi.org/10.1145/2908812.2908950>
8. P. K. Lehre, C. Witt, Tail bounds on hitting times of randomized search heuristics using variable drift analysis, *Comb. Probab. Comput.*, **30** (2020), 550–569. <https://doi.org/10.1017/s0963548320000565>
9. C. Gießen, C. Witt, Optimal mutation rates for the  $(1 + \lambda)$  EA on OneMax through asymptotically tight drift analysis, *Algorithmica*, **80** (2017), 1710–1731. <https://doi.org/10.1007/s00453-017-0360-y>
10. R. Sarker, M. Mohammadian, X. Yao, I. Wegener, *Methods for the Analysis of Evolutionary Algorithms on Pseudo-Boolean Functions*, Springer US, New York, USA, (2002), 349–369. <https://doi.org/10.17877/DE290R-15272>
11. Y. Yu, C. Qian, Z. H. Zhou, Switch analysis for running time analysis of evolutionary algorithms, *IEEE Trans. Evol. Comput.*, **19** (2015), 777–792. <https://doi.org/10.1109/tevc.2014.2378891>
12. C. Qian, C. Bian, W. Jiang, K. Tang, Running time analysis of the  $(1 + 1)$ -EA for OneMax and LeadingOnes under bit-wise noise, *Algorithmica*, **81** (2018), 749–795. <https://doi.org/10.1007/s00453-018-0488-4>
13. I. Wegener, *Methods for the Analysis of Evolutionary Algorithms on Pseudo-Boolean Functions Evolutionary Optimization*, Springer US, New York, USA, **48** (2003), 349–369. [https://doi.org/10.1007/0-306-48041-7\\_14](https://doi.org/10.1007/0-306-48041-7_14)
14. H. Huang, J. P. Su, Y. S. Zhang, Z. F. Hao, An experimental method to estimate running time of evolutionary algorithms for continuous optimization, *IEEE Trans. Evol. Comput.*, **24** (2020), 275–289. <https://doi.org/10.1109/tevc.2019.2921547>
15. F. J. Feng, H. Huang, Y. S. Zhang, Z. F. Hao, Comparative analysis for first hitting time of evolutionary algorithms based on equal-in-time model, *Chin. J. Comput.*, **42** (2019), 2297–2308. <https://doi.org/10.11897/SP.J.1016.2019.02297>
16. D. Morinaga, K. Fukuchi, J. Sakuma, Y. Akimoto, Convergence rate of the  $(1 + 1)$  ES on locally strongly convex and Lipschitz smooth functions, *IEEE Trans. Evol. Comput.*, 2023. <https://doi.org/10.1109/tevc.2023.3266955>
17. J. Jägersküpper, Combining Markov-chain analysis and drift analysis: The  $(1 + 1)$  evolutionary algorithm on linear functions reloaded, *Algorithmica*, **59** (2010), 409–424. <https://doi.org/10.1007/s00453-010-9396-y>
18. A. Agapie, M. Agapie, G. Rudolph, G. Zbaganu, Convergence of evolutionary algorithms on the  $n$ -dimensional continuous space, *IEEE Trans. Cybern.*, **43** (2013), 1462–1472. <https://doi.org/10.1109/TCYB.2013.2257748>
19. D. Morinaga, K. Fukuchi, J. Sakuma, Y. Akimoto, Convergence rate of the  $(1 + 1)$  evolution strategy with success-based step-size adaptation on convex quadratic functions, in *Proceedings of the Genetic and Evolutionary Computation Conference*, (2021), 1169–1177. <https://doi.org/10.1145/3449639.3459289>

20. Y. S. Zhang, H. Huang, Z. F. Hao, G. W. Hu, First hitting time analysis of continuous evolutionary algorithms based on average gain, *Chin. J. Comput.*, **42** (2019), 624–635. <https://doi.org/10.11897/SP.J.1016.2019.00624>
21. A. Agapie, Spherical distributions used in evolutionary algorithms, *Mathematics*, **9** (2021), 2227–7390. <https://doi.org/10.3390/math9233098>
22. B. Doerr, C. Witt, J. Yang, Running time analysis for self-adaptive mutation rates, in *Proceedings of the Genetic and Evolutionary Computation Conference*, (2018), 1475–1482. <https://doi.org/10.1007/S00453-020-00726-2>
23. Y. Akimoto, A. Anne, G. Tobias, Drift theory in continuous search spaces: expected hitting time of the  $(1 + 1)$  ES with  $1/5$  success rule, in *Proceedings of the Genetic and Evolutionary Computation Conference*, (2018), 801–808. <https://doi.org/10.1145/3205455.3205606>
24. J. Jägersküpfer, Lower bounds for randomized direct search with isotropic sampling, *Oper. Res. Lett.*, **36** (2008), 327–332. <https://doi.org/10.1016/j.orl.2007.10.003>
25. H. G. Beyer, H. P. Schwefel, Evolution strategies—a comprehensive introduction, *Nat. Comput.*, **1** (2002), 3–52. <https://doi.org/10.1023/a:1015059928466>
26. F. Neumann, C. Witt, Runtime Analysis of the  $(1 + 1)$  EA on weighted Sums of transformed linear functions, in *International Conference on Parallel Problem Solving from Nature*, Cham: Springer International Publishing, (2022), 542–554. [https://doi.org/10.1007/978-3-642-30687-7\\_4](https://doi.org/10.1007/978-3-642-30687-7_4)
27. B. Doerr, T. Kötzing, Lower bounds from fitness levels made easy, in *Proceedings of the Genetic and Evolutionary Computation Conference*, New York: ACM, (2021), 1142–1150. <https://doi.org/10.1145/3449639.3459352>
28. H. G. Beyer, *The Theory of Evolution Strategies*, Springer Science & Business Media, Berlin, Germany, 2001. [https://doi.org/10.1007/978-3-662-04448-3\\_6](https://doi.org/10.1007/978-3-662-04448-3_6)
29. T. Kötzing, Concentration of first hitting times under additive drift, *Algorithmica*, **75** (2016), 490–506.
30. J. He, X. Yao, Drift analysis and average time complexity of evolutionary algorithms, *Artif. Intell.*, **127** (2001), 57–85. [https://doi.org/10.1016/s0004-3702\(01\)00058-3](https://doi.org/10.1016/s0004-3702(01)00058-3)
31. K. W. Fang, *Symmetric Multivariate and Related Distributions*, CRC Press, London, UK, 2018. <https://doi.org/10.1201/9781351077040>
32. A. Agapie, O. Solomon, L. Bădin, Theory of  $(1 + 1)$  ES on sphere revisited, *IEEE Trans. Evol. Comput.*, **27** (2023), 938–948. <https://doi.org/10.1109/tevc.2022.3217524>
33. H. G. Beyer, Toward a theory of evolution strategies: Some asymptotical results from the  $(1 + \lambda)$ -theory, *Evol. Comput.*, **1** (1993), 165–188. <https://doi.org/10.1162/evco.1993.1.2.165>



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>).