



# Improved Running Time Analysis of the (1+1)-ES on the Sphere Function

Wu Jiang<sup>1(✉)</sup>, Chao Qian<sup>1</sup>, and Ke Tang<sup>2</sup>

<sup>1</sup> Anhui Province Key Lab of Big Data Analysis and Application,  
University of Science and Technology of China, Hefei 230027, China  
{jw1992, chaoqian}@mail.ustc.edu.cn

<sup>2</sup> Shenzhen Key Lab of Computational Intelligence,  
Southern University of Science and Technology, Shenzhen 518055, China  
tangk3@sustc.edu.cn

**Abstract.** During the last two decades, much progress has been achieved on the running time analysis (one essential theoretical aspect) of evolutionary algorithms (EAs). However, most of them focused on discrete optimization, and the theoretical understanding is largely insufficient for continuous optimization. The few studies on evolutionary continuous optimization mainly analyzed the running time of the (1+1)-ES with Gaussian and uniform mutation operators solving the sphere function, the known bounds of which are, however, quite loose compared with the empirical observations. In this paper, we significantly improve their lower bound, i.e., from  $\Omega(n)$  to  $\Omega(e^{cn})$ . Then, we study the effectiveness of 1/5-rule, a widely used self-adaptive strategy, for continuous EAs using uniform mutation operator for the first time. We prove that for the (1+1)-ES with uniform mutation operator solving the sphere function, using 1/5-rule can reduce the running time from exponential to polynomial.

**Keywords:** Running time analysis · Continuous optimization  
Evolution strategies

## 1 Introduction

Evolutionary algorithms (EAs) are widely used in real-world applications and have achieved great success in solving both continuous and discrete optimization problems. As a class of general-purpose optimization algorithms, EAs are designed to search for the optimum without the information of gradients or Hessian matrix. Thus, they have been regarded as a major approach when the optimization problems are non-differential, multi-modal, or black-box. However, a great gap lies between the numerous practice and weak theoretical foundation. Moreover, most of the theoretical works focus on the discrete domain, such as pseudo-Boolean functions [4, 10–12, 14, 15], while only a few works deal with continuous optimization problems [1, 7, 8].

Note that the theoretical analysis of EAs in discrete domain cannot be directly extended to continuous domain. In the following, we briefly summarize the difference between continuous EAs and discrete EAs.

- Optimization problems they are designed to solve originally. Continuous EAs are designed to solve continuous optimization problems, while discrete EAs are more often used to solve discrete optimization problems, such as combinatorial optimization problems, etc.
- Coding they use to represent the individuals. Individuals are represented by a real vector in continuous EAs in contrast to a binary vector in discrete EAs.
- Mutation operators. In discrete EAs, an offspring solution is generated by flipping the bits of the parent solution. In continuous EAs, we first sample a stochastic vector from a given distribution. A new candidate solution is generated by adding the preceding stochastic vector to the solution.
- Stopping criteria. The searching space is infinite and uncountable in continuous domain, which means the probability of finding the exact optimum is 0. Thus, we focus on the number of steps that are needed to halve the approximation error, i.e., the distance from the optimum in continuous domain. While in discrete domain, the evolution would not stop until the exact optimum or an acceptable suboptimal solution is found.

Most of the previous theoretical works in continuous EAs analyze the expected running time of evolution strategies (ES) on the sphere function. Jägersküpfer [7] proved that the (1+1)-ES with Gaussian mutation operator has a polynomial lower bound  $\Omega(n)$  on certain function scenarios (including the sphere function), and derived an upper bound  $O(n)$  when 1/5-rule is introduced. These two results together confirmed the effectiveness of 1/5-rule. Moreover, Akimoto et al. proved that the expected running time of the (1+1)-ES with Gaussian mutation operator adapted by 1/5-rule on the sphere function is  $\Theta((\log \frac{d}{\epsilon})n)$ , where  $d$  denotes the distance of the initial individual to the global optimum and  $\epsilon$  is the parameter of the stopping criteria [2]. Uniform mutation operator was then analyzed by Agapie et al., with a lower bound  $\Omega(n)$  and an upper bound  $O(e^{cn})$  [1].

In this paper, we first prove that the general lower bound on the running time of the (1+1)-ES with uniform mutation operator inside a hypersphere is  $\Omega(e^{cn})$  which largely improves the known bound  $\Omega(n)$  which was given in [1]. Second, we derive a similar result when the mutation operator is replaced by Gaussian mutation, i.e.  $\Omega(e^{cn})$ , in contrast to  $\Omega(n)$  given by Jägersküpfer in [7]. Third, we study the effectiveness of 1/5-rule for the (1+1)-ES using uniform mutation inside a hypersphere for the first time. We prove that the running time of the (1+1)-ES using uniform mutation inside a hypersphere has polynomial upper bound  $O(n)$  after the incorporation of 1/5-rule.

The rest of the paper is organized as follows. Section 2 introduces some preliminaries, including the studied problem and algorithm, and also the analysis tools that will be used. Section 3 derives the exponential lower bound of the (1+1)-ES using uniform and Gaussian mutation operators. Section 4 analyzes the effectiveness of 1/5-rule for the (1+1)-ES using uniform mutation operator inside a hypersphere. Section 5 concludes the paper.

## 2 Preliminaries

In this section, we first introduce the problem and algorithm studied in this paper, respectively. Then, we present the analysis tools that we use throughout the paper.

### 2.1 Problem

We consider minimizing the sphere function in this paper. Its formulation is as follows,

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

The sphere function has two properties:

- A minimum exists.
- $f(x) < f(y)$  if  $|x| < |y|$  for any two points  $x$  and  $y$ , which means that a mutant closer to the optimum is always accepted.

### 2.2 (1+1)-ES and 1/5-Rule

Jägersküpfer analyzed the time complexity of the  $(1 + \lambda)$ -ES on the sphere function and proved that the offspring population helps to reduce the running time of ES [8]. However, parent population does not help for the hypersphere function, and even makes the expected running time larger [6]. Our theoretical analysis will focus on the (1+1)-ES, which maintains only one parent solution and one offspring solution in each generation. The (1+1)-ES replaces the current individual by a better candidate offspring, and discards a worse one in every iteration. The process of the (1+1)-ES using uniform mutation inside a sphere is as follows,

*Initialization:* Create a random initial individual  $\xi_0$  which satisfies  $d(\xi_0, o) = R$ , where  $R$  denotes a positive constant and  $o$  is the global optimum.

*Mutation:* Create a new search point  $\xi' = \xi_t + \sigma m$ , where  $\sigma$  is scaling factor and  $m$  is a  $n$ -dimensional vector that uniformly distributed in a sphere of fixed radius  $r$  (or sampled from  $n$ -variate standard Gaussian distribution).

*Evaluation:* If  $f(\xi') \leq f(\xi_t)$ ,  $\xi_{t+1} = \xi'$ , else  $\xi_{t+1} = \xi_t$ .

*Stopping criteria:*  $f(\xi_t) \leq R/2$ .

Among all the self-adaptive strategies for the coefficients of mutation operators in real-world applications and research areas, 1/5-rule is the most commonly used one [3, 13]. This heuristic strategy has shown excellent performance by controlling the mutation strength of ES. The 1/5-rule works as follows: the optimization process is observed for  $n$  steps without changing  $\sigma$ ; if more than one fifth of the steps in this observed phase have been successful,  $\sigma$  is doubled; otherwise,  $\sigma$  is halved.

### 2.3 Analysis Tools

It is not hard to notice the common property between the evolving of population in ES and a Markov Chain  $\{\xi_t\}_{t \geq 0}$ , since the future state/population is independent of the past states/populations and only effected by the current state/population. However, it is nearly impossible to construct a transition probability matrix for ES. Thus, a new effective model, renewal process, is introduced for ES. For each  $t = 0, 1, 2, \dots$ , let  $\xi_t$  be the individual of  $t$ -th generation in ES and  $d(\xi_t)$  be the distance from  $\xi_t$  to the optimum 0. Define the stochastic process  $\{\rho_t\}_{t \geq 0}$  by  $\rho_t = d(\xi_t) - d(\xi_{t+1})$ , then  $\{\rho_t\}_{t \geq 0}$  is a renewal process.

Since ES is a class of stochastic algorithms, the first hitting time is a random variable, either. We take the expectation of first hitting time (FHT) as the measure of time complexity of ES. The definition of FHT is as follows,

**First Hitting Time.** Assume that the initial state set is  $\chi_I$ , and the final state set is  $\chi^*$ , the first hitting time  $T$  is defined as  $T = \min\{t | \xi_t \in \chi^* \wedge \xi_0 \in \chi_I\}$ . In our paper, particularly  $\chi_I = \{\xi \in \mathbb{R}^n | d(\xi) = R\}$  and  $\chi^* = \{\xi \in \mathbb{R}^n | d(\xi) < R/2\}$ .

Several analysis tools have been proposed with different ideas and principles in discrete domain. All of these methods have shown great power in the running time analysis of EAs. Drift analysis, constructs a distance function and derives the expected running time by bounding the one-step progress [4]. Fitness level method cuts the search space into different sets or levels based on the fitness values and derives the expected running time by bounding the transition probability of the solutions moving from the current level to other levels [14]. Switch analysis compares the evolutionary process with another one which has been well analyzed and derives the expected running time by estimating the one-step differences of the two [15]. However, these three effective methods cannot be applied to continuous optimization directly since they are all based on Markov Chain with finite states.

Similar to the drift analysis theorem, which was first introduced to analyze the running time of EAs by He and Yao [4], the analysis approach for ES introduced by Jägerskupper also focuses on the difference of the two adjacent generations, defined as progress rate  $\rho_t$  [7].

**Progress Rate.** Assume that the  $t$ -th generation is  $\xi_t$ , the distance of  $\xi_t$  from the optimizer is  $d(\xi_t)$ , then  $\rho_t = d(\xi_t) - d(\xi_{t+1})$  is called the progress rate of the  $t$ -th generation.

**Theorem 1 (Lower bound theorem [7]).** Let  $\rho_1, \rho_2, \dots$  denote random variables with bounded range and  $T$  be the random variable defined by  $T = \min\{t | \sum_{i=0}^t \rho_i \geq g\}$  for a given  $g > 0$ . If  $E(T)$  exists and  $E(\rho_i | T \geq i) \leq u$ , then  $E(T) \geq g/u$ .

Theorem 1 reveals that if the progress rate has an upper bound, we could derive a lower bound for the expected running time.

### 3 Exponential Lower Bound of the (1+1)-ES

In this section, we prove that the expected running time of the (1+1)-ES using uniform mutation operator inside a hypersphere and Gaussian mutation operator has exponential lower bound when used for minimizing the sphere function, presented in Theorems 2 and 3, respectively.

#### 3.1 Uniform Mutation Inside a Hypersphere

Lower bound of the (1+1)-ES minimizing the sphere function derived by Agapie et al. is not tight enough, which could be testified by simulation experiments [1]. The super-polynomial bound by Huang et al., however, is concerned with uniform mutation inside a hypercube and based on a very strong assumption on the initialization and mutation operator [5]. In this subsection, we give out the lower bound which is tighter than the state-of-the-art bound. Besides, our method is more general, which has no assumption on the coefficients of the mutation operator and initialization.

**Theorem 2.** For the (1+1)-ES minimizing the sphere function in  $n$ -dimensional Euclidian space using uniform mutation in a hypersphere and elitist selection, the expected number of generations  $T$  to half the approximation error is  $\Omega(e^{cn})$ .

**Proof.** Denote the global optimum as  $o$  and the current individual as  $c$ , i.e.  $\xi_t$ . Supposing that the distance between  $o$  and  $c$  is  $d(o, c) = R$  and the radius of the mutation operator is  $r$ . Due to symmetry, we assume that  $c$  is on axes  $ox_1$ .

As shown in Fig. 1. Intersection of fitness sphere and mutation sphere, we call the left sphere  $S_o = \{\xi \in \mathbb{R}^n | d(\xi, o) \leq R\}$  fitness sphere, since the all points inside  $S_o$  are better than the current population  $c$ . Accordingly, we call the right sphere  $S_c = \{\xi \in \mathbb{R}^n | d(\xi, c) \leq r\}$  mutation sphere, since all the candidate offsprings scatter in it.  $A$  is denoted as the intersection of the two spheres.

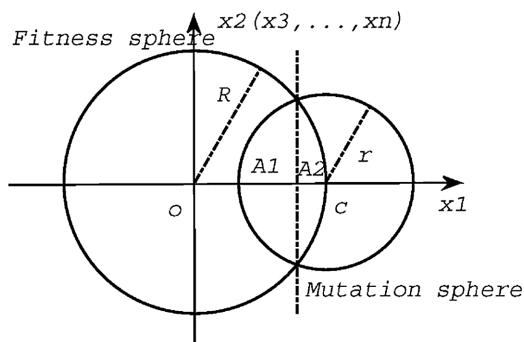


Fig. 1. Intersection of fitness sphere and mutation sphere

According to the definition of progress rate, we get

$$E(\rho_t) = \frac{1}{C_n r^n} \int (R - d(\xi_{t+1})) d\xi_{t+1} \leq \frac{1}{C_n r^n} \cdot R \cdot V(A) \tag{2}$$

Where  $C_n r^n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} r^n$  is the volume of the  $n$ -dimensional sphere with radius  $r$  and  $V(A)$  is the volume of the intersection of the two hyperspheres. We begin our deduction in an algebraic way. By solving the equations below

$$\begin{cases} \sum_{i=1}^n x_i^2 = R^2 \\ (x_1 - R)^2 + \sum_{i=2}^n x_i^2 = r^2 \end{cases} \tag{3}$$

We get

$$x_1 = R - \frac{r^2}{2R} \tag{4}$$

Which is the  $x_1$ -coordinate of the hyperplane  $P$ .

1.  $r > R$ . In this case, we have

$$V(A) \leq V(S_o) = C_n R^n = C_n r^n \cdot O(e^{c_1 \cdot n}) \tag{5}$$

Where  $c_1 < 0$  is a negative constant.

2.  $r \leq R$ . In this case, the intersection  $A$  can be viewed as two spherical caps.

$$\begin{aligned} V(A) &= V(A_1) + V(A_2) \\ &= C_n r^n \cdot \frac{1}{2} I_{\left(1 - \frac{r^2}{4R^2}\right)} \left(\frac{n+1}{2}, \frac{1}{2}\right) + C_n R^n \frac{1}{2} I_{\left(\frac{r^2}{R^2} - \frac{r^4}{4R^4}\right)} \left(\frac{n+1}{2}, \frac{1}{2}\right) \\ &< C_n r^n \cdot \frac{\frac{R}{r} \left(1 - \frac{r^2}{4R^2}\right)^{\frac{n+1}{2}}}{\frac{n+1}{2} \cdot B\left(\frac{n+1}{2}, \frac{1}{2}\right)} + C_n R^n \cdot \frac{\frac{R^2}{2R^2 - r^2} \left(\frac{r^2}{R^2} - \frac{r^4}{4R^4}\right)^{\frac{n+1}{2}}}{\frac{n+1}{2} \cdot B\left(\frac{n+1}{2}, \frac{1}{2}\right)} \\ &= C_n r^n \cdot \frac{\frac{2R^3}{r(2R^2 - r^2)} \left(1 - \frac{r^2}{4R^2}\right)^{\frac{n+1}{2}}}{\frac{n+1}{2} \cdot B\left(\frac{n+1}{2}, \frac{1}{2}\right)} \\ &= C_n r^n \cdot O(e^{c_2 \cdot n}) \end{aligned}$$

Where  $c_2 < 0$  is a negative constant. The second equation derived by applying the formula of volume of spherical cap  $V(C_{a,b}) = C_n a^n \cdot \frac{1}{2} I_{\frac{2ab-a^2}{a^2}}(\frac{n+1}{2}, \frac{1}{2})$  by Li, where  $a$  is the radius of the hypersphere that the spherical cap is cut from and  $b$  is the height of the spherical cap [9]. The first inequality is based on

$$\begin{aligned}
 I_x\left(\frac{n+1}{2}, \frac{1}{2}\right) &\equiv \frac{B(x; \frac{n+1}{2}, \frac{1}{2})}{B(\frac{n+1}{2}, \frac{1}{2})} = \frac{\int_0^x t^{\frac{n-1}{2}}(1-t)^{-\frac{1}{2}} dt}{B(\frac{n+1}{2}, \frac{1}{2})} \\
 &< \frac{(1-x)^{-\frac{1}{2}} \int_0^x t^{\frac{n-1}{2}} dt}{B(\frac{n+1}{2}, \frac{1}{2})} = \frac{(1-x)^{-\frac{1}{2}} x^{\frac{n+1}{2}}}{\frac{n+1}{2} \cdot B(\frac{n+1}{2}, \frac{1}{2})}
 \end{aligned}
 \tag{6}$$

The last equation is based on

$$B\left(\frac{n+1}{2}, \frac{1}{2}\right) = \frac{\Gamma(\frac{n+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n}{2} + 1)} = \Theta\left(\left(\frac{n}{2} + 1\right)^{-\frac{1}{2}}\right) = \Theta\left(n^{-\frac{1}{2}}\right)
 \tag{7}$$

Summing up the two cases, we get

$$E(\rho_t) = \frac{1}{C_n r^n} \int (R - d(\xi_{t+1})) d\xi_{t+1} \leq \frac{1}{C_n r^n} \cdot R \cdot V(A) = O(e^{c_3 \cdot n})
 \tag{8}$$

Where  $c_3 = \max\{c_1, c_2\} < 0$  is a negative constant. By lower bound theorem, we have

$$E(T) \geq \frac{R - R/2}{O(e^{c_3 \cdot n})} = \Omega(e^{cn})
 \tag{9}$$

Where  $c = -c_3 > 0$  is a positive constant.

### 3.2 Gaussian Mutation

The analysis by Huang et al. for Gaussian mutation is based on a very strong assumption on the initialization [5]. In this section, we give out a lower bound which is much tighter than the bound  $\Omega(n)$  given by Jägersküpfer [7]. Moreover, our analysis is more general than [5] since we have no assumption on the initialization.

**Theorem 3.** For the (1+1)-ES minimizing the sphere function in  $n$ -dimensional Euclidian space using Gaussian mutation and elitist selection, the expected number of generations  $T$  to half the approximation error is  $\Omega(e^{cn})$ .

**Proof.** Denote the global optimum as  $o$  and the current individual as  $c$ , i.e.  $\xi_t$ . Supposing that the distance between  $o$  and  $c$  is  $d(o, c) = R$ . Due to symmetry, we assume that  $c$  is on axes  $ox_1$ . Similarly, we call the left sphere  $S_o = \{\xi \in \mathbb{R}^n \mid d(\xi, o) \leq R\}$  fitness sphere, since the all points inside  $S_o$  are better than the current population  $c$ .

According to the definition of progress rate, we get

$$E(\rho_t) = \int \varphi(\xi_{t+1})(R - d(\xi_{t+1}))d\xi_{t+1} \leq \int \varphi(\xi_{t+1})Rd\xi_{t+1} \tag{10}$$

Where  $\varphi(\cdot)$  is the probability density function of the  $n$ -variate standard Gaussian distribution. Since the  $\varphi(\xi_{t+1}) \leq \max\{\varphi(m)|m \in \mathbb{R}^n\} = (2\pi)^{-n/2}$ , we have

$$E(\rho_t) \leq \int (2\pi)^{-\frac{n}{2}}Rd\xi_{t+1} = (2\pi)^{-\frac{n}{2}} \cdot R \cdot V(S_o) = (2\pi)^{-\frac{n}{2}} \cdot \frac{\pi^{n/2}R^{n+1}}{\Gamma(n/2 + 1)} \tag{11}$$

Where  $V(S_o)$  is the volume of the fitness sphere  $S_o$ . Since  $R/2 \leq d(\xi_i) \leq R$  for  $\forall i \in [0, T]$ , we have  $E(\rho_t) \leq \frac{R^{n+1}}{\Gamma(n/2 + 1) \cdot 2^{n/2}}$ . By lower bound theorem, we have

$$E(T) \geq \frac{R - \frac{R}{2}}{\frac{R^{n+1}}{\Gamma(\frac{n}{2} + 1) \cdot 2^{\frac{n}{2}}}} = \frac{2^{\frac{n}{2}-1}}{R^n} \cdot \Gamma\left(\frac{n}{2} + 1\right) = \Omega(e^{cn}) \tag{12}$$

Where  $c > 0$  is a positive constant, since  $R > 0$  is a positive constant.

#### 4 On the Effectiveness of 1/5-Rule

Jägersküpper proved that the upper bound of the (1+1)-ES minimize the sphere function using Gaussian mutations adapted by 1/5-rule is  $O(n)$  [7]. In this section, we study the effectiveness of 1/5-rule for the (1+1)-ES using uniform mutation operator.

**Lemma 1** [7]. For a  $n$ -dimensional vector  $m$  with each component independently standard normal distributed, the expectation  $l_E = E(|m|)$  exists, and  $P\{|m| - l_E| \geq \delta \cdot l_E\} \leq \frac{\delta^{-2}}{2n-1}$ .

Similar result holds for uniform mutation inside a hypersphere. Lemma 2 reveals that the length of the  $n$ -dimensional vector uniformly sampled from the unit hypersphere is  $\Theta(1)$  w.o.p., with overwhelming probability.

**Lemma 2.** For a  $n$ -dimensional vector  $m$  uniformly distributed in a unit hypersphere, the expectation  $l_E = E(|m|)$  exists, and  $P\{||m| - l_E| \geq \delta \cdot l_E\} \leq \frac{\delta^{-2}}{n^2 + 2n}$ .

**Proof.** According to the definition of expectation in mathematics, we have

$$E(|m|) = \int_0^1 rA_n(r) \cdot \frac{1}{V_n(1)} dr = \frac{n}{n+1} \tag{13}$$



Where  $V_n(1) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$  is the volume of the unit hypersphere and  $A_n(r) = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})} \cdot r^{n-1}$  is the surface area of the  $n$ -dimensional hypersphere of radius  $r$ . Similarly, we have

$$\mathbb{E}\left(|m|^2\right) = \int_0^1 r^2 A_n(r) \cdot \frac{1}{V_n(1)} dr = \frac{n}{n+2} \quad (14)$$

Consequently, we get

$$\text{Var}(|m|) = \mathbb{E}\left(|m|^2\right) - \mathbb{E}(|m|)^2 = \frac{n}{(n+1)^2(n+2)} \quad (15)$$

Thus, we get

$$\mathbb{P}\{|m| - l_E \geq \delta \cdot l_E\} \leq \frac{\delta^{-2}}{n^2 + 2n} \quad (16)$$

by applying the Chebyshev's inequality.

Intuitively, we partition a run of the (1+1)-ES adapted by 1/5-rule into phases each of which lasts  $n$  steps. Thus, in each phase, the radius of the mutation sphere is fixed. Denote the radius and the scaling factor in the  $t$ -th phase as  $r_t$  and  $\delta_t$ . Jägersküpper find the relation between the radius of the mutation sphere and the progress rate, for Gaussian mutation though, but could be extended to uniform mutation similarly.

**Lemma 3** [7]. Let the (1+1)-ES minimize the sphere function using Gaussian mutations adapted by the 1/5-rule and elitist selection. Then

1. If  $r_t = \Theta(d(\xi_t/\sqrt{n}))$ , then w.o.p.  $d(\xi_{t+1}) = d(\xi_t) - \Omega(d(\xi_t))$ , i.e., the distance to the optimum is reduced by a constant fraction in the  $t$ -th phase.
2. If  $\delta$  is doubled after the  $t$ -th phase, then  $r_t = O(d(\xi_t/\sqrt{n}))$  w.o.p.;
3. If  $\delta$  is halved after the  $t$ -th phase, then  $r_{t+1} = \Omega(d(\xi_{t+1}/\sqrt{n}))$  w.o.p.;

Since the proof of Lemma 3 is merely based on the isotropy of Gaussian mutation and Lemma 1, we could derive the following result by simply applying the isotropy of uniform mutation and Lemma 2.

**Lemma 4.** The (1+1)-ES minimize the sphere function using uniform mutations inside a hypersphere adapted by the 1/5-rule and elitist selection. Then

1. If  $r_t = \Theta(d(\xi_t/\sqrt{n}))$ , then w.o.p.  $d(\xi_{t+1}) = d(\xi_t) - \Omega(d(\xi_t))$ , i.e., the distance to the optimum is reduced by a constant fraction in the  $t$ -th phase.
2. If  $\delta$  is doubled after the  $t$ -th phase, then  $r_t = O(d(\xi_t/\sqrt{n}))$  w.o.p.;
3. If  $\delta$  is halved after the  $t$ -th phase, then  $r_{t+1} = \Omega(d(\xi_{t+1}/\sqrt{n}))$  w.o.p.;

**Theorem 4** [7]. Let the (1+1)-ES minimize the sphere function in  $n$ -dimensional Euclidian space using Gaussian mutation adapted by 1/5-rule and elitist selection. Given that the initialization ensures  $d(\xi_0)/\delta_0 = \Theta(n)$ , the expected number of generations  $T$  to half the approximation error is w.o.p.  $O(n)$ .

Since the proof of Theorem 4 is merely based on the isotropy of Gaussian mutation and Lemma 3, we could derive the following result by simply applying the isotropy of uniform mutation and Lemma 4.

**Theorem 5.** Let the (1+1)-ES minimize the sphere function in  $n$ -dimensional Euclidian space using uniform mutation in a hypersphere adapted by 1/5-rule and elitist selection. Given that the initialization ensures  $d(\xi_0)/\delta_0 = \Theta(\sqrt{n})$ , the expected number of generations  $T$  to half the approximation error is w.o.p.  $O(n)$ .

## 5 Conclusion

In this paper, we first derive a tighter lower bound, i.e.,  $\Omega(e^{cn})$ , for (1+1)-ES using uniform mutation inside a hypersphere on the sphere function in contrast to the state-of-the-art bound  $\Omega(n)$ . Second, we extend the result to the case when the mutation operator is Gaussian mutation. Third, we study the effectiveness of 1/5-rule for (1+1)-ES using uniform mutation inside a hypersphere for the first time, and prove that the incorporation of 1/5-rule reduces the time complexity from exponential to polynomial.

Uniform mutation operator that distributed in a hypercube is also used by some researchers. It does not belong to isotropic distribution like uniform distribution inside a hypersphere do. The effectiveness of 1/5-rule on it has not been studied yet. We leave it as our future work.

## References

1. Agapie, A., Agapie, M., Rudolph, G., Zbaganu, G.: Convergence of evolutionary algorithms on the  $n$ -dimensional continuous space. *IEEE Trans. Cybern.* **43**(5), 1462–1472 (2013)
2. Akimoto, Y., Auger, A., Glasmachers, T.: Drift theory in continuous search spaces: expected hitting time of the (1+1)-ES with 1/5 success rule. [arXiv:1802.03209](https://arxiv.org/abs/1802.03209) (2018)
3. Beyer, H.-G.: *The Theory of Evolution Strategies*. Springer, New York (2001). <https://doi.org/10.1007/978-3-662-04378-3>
4. He, J., Yao, X.: A study of drift analysis for estimating computation time of evolutionary algorithms. *Nat. Comput.* **3**(1), 21–35 (2004)
5. Huang, H., Xu, W., Zhang, Y., Lin, Z., Hao, Z.: Runtime analysis for continuous (1+1) evolutionary algorithm based on average gain model. *SCIENTIA SINICA Informationis* **44**(6), 811–824 (2014)
6. Jägersküpfer, J., Witt, C.: Rigorous runtime analysis of a  $(\mu + 1)$  ES for the sphere function. In: *7th International Proceedings of Genetic and Evolutionary Conference*, Washington, D. C., pp. 849–856. ACM (2005)
7. Jägersküpfer, J.: Algorithmic analysis of a basic evolutionary algorithm for continuous optimization. *Theor. Comput. Sci.* **379**(3), 329–347 (2007)
8. Jägersküpfer, J.: Probabilistic runtime analysis of  $(1 +, \lambda)$  ES using isotropic mutations. In: *8th International Proceedings of the Genetic and Evolutionary Computation Conference*, Seattle, WA, USA, pp. 461–468. ACM (2006)
9. Li, S.: Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Stat.* **4**(1), 66–70 (2011)

10. Neumann, F.: Expected runtimes of a simple evolutionary algorithm for the multi-objective minimum spanning tree problem. *Eur. J. Oper. Res.* **181**(3), 1620–1629 (2007)
11. Qian, C., Bian, C., Jiang, W., Tang, K.: Running time analysis of the (1+1)-EA for onemax and leadingones under bitwise noise. In: 19th International Proceedings of the Genetic and Evolutionary Computation Conference, Berlin, German, pp. 1399–1406. ACM (2017)
12. Qian, C., Yu, Y., Jin, Y., Zhou, Z.-H.: On the effectiveness of sampling for evolutionary optimization in noisy environments. In: Bartz-Beielstein, T., Branke, J., Filipič, B., Smith, J. (eds.) PPSN 2014. LNCS, vol. 8672, pp. 302–311. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-10762-2\\_30](https://doi.org/10.1007/978-3-319-10762-2_30)
13. Tang, K., Yang, P., Yao, X.: Negatively correlated search. *IEEE J. Sel. Areas Commun.* **34**(3), 542–550 (2016)
14. Wegener, I.: Methods for the analysis of evolutionary algorithms on pseudo-Boolean functions. In: Sarker, R., Mohammadian, M., Yao, X. (eds.) *Evolutionary Optimization*, pp. 349–369. Springer, Boston (2003). [https://doi.org/10.1007/0-306-48041-7\\_14](https://doi.org/10.1007/0-306-48041-7_14)
15. Yu, Y., Qian, C., Zhou, Z.-H.: Switch analysis for running time analysis of evolutionary algorithms. *IEEE Trans. Evol. Comput.* **19**(6), 777–792 (2015)