Random Dynamics Optimum Tracking with Evolution Strategies

Dirk V. Arnold and Hans-Georg Beyer

Department of Computer Science XI
University of Dortmund
44221 Dortmund, Germany
{arnold,beyer}@ls11.cs.uni-dortmund.de

Abstract. Dynamic optimization is frequently cited as a prime application area for evolutionary algorithms. In contrast to static optimization, the objective in dynamic optimization is to continuously adapt the solution to a changing environment – a task that evolutionary algorithms are believed to be good at. At the time being, however, almost all knowledge with regard to the performance of evolutionary algorithms in dynamic environments is of an empirical nature. In this paper, tools devised originally for the analysis in static environments are applied to study the performance of a popular type of recombinative evolution strategy with cumulative mutation strength adaptation on a dynamic problem. With relatively little effort, scaling laws that quite accurately describe the behavior of the strategy and that greatly contribute to its understanding are derived and their implications are discussed.

1 Introduction

Dynamic optimization is distinguished from static optimization in that in the former, the objective function is not constant but varies with time. Dynamic optimization problems arise in many areas of engineering and computer science, as for example in the determination of optimal control policies or in connection with online job scheduling where new jobs arrive in the course of the optimization. While the goal in static optimization is to locate an optimal solution rapidly and accurately, the objective in dynamic optimization is to track a moving target as closely as possible. Strategies for dynamic optimization thus need to continuously adapt to changes in the environment.

In enumerations of potential domains of application of evolutionary algorithms, dynamic optimization often takes one of the top spots. At the time being, almost all knowledge with regard to the capabilities of evolutionary algorithms in dynamic environments is based on empirical observations. An extensive survey of the literature of the field along with a collection of benchmark functions and a discussion of methods that have been proposed to improve the performance of evolutionary algorithms in dynamic environments has been compiled by Branke [7]. Angeline [1] compares empirically the tracking performance of an evolutionary algorithm employing a form of mutative self-adaptation with that of a strategy using a simple heuristic for mutation strength adaptation. The fitness environment considered is a three-dimensional, spherically symmetric objective function that is shifted periodically either in a random fashion or on a linear or a spherical path. Angeline observes that the self-adaptation mechanism is not without problems in the dynamic case. In that same fitness environment, Bäck [4] compares different variants of mutative self-adaptation and presents evidence that seems to indicate that the lognormal self-adaptation used in evolution strategies performs better than the variant of self-adaptation commonly used in evolutionary programming. Salomon and Eggenberger [12] compare the performance of evolution strategies with that of a breeder genetic algorithm on the sphere, an ellipsoid, and Rastrigin's function, where the coordinates are shifted by a constant increment in every time step. The search space dimensionalities they consider for the sphere are $N = 10$ and $N = 30$. Without quantifying the term, they find that the sensitivity to the particular implementation of the strategy and to its parameter values is much lower for the tracking task than it is in a static environment. Without providing details, they also report to have observed that recombination is not beneficial for tracking a moving target. Weicker and Weicker [13] contrast self-adaptation of a single mutation strength with that of N mutation strengths and adaptation of the full mutation covariance matrix and find that in more rapidly changing environments, the adaptation of more than a single mutation strength becomes unreliable. Finally, Droste [8] presents a first rigorous analysis of the performance of a $(1+1)$-strategy on a discrete, dynamic objective function. However, focus in that paper is not on the tracking behavior of the strategy but rather on the expected time required to first reach the optimum.

While useful for providing the reader with a first idea of the capabilities and the limitations of evolutionary algorithms in dynamic environments, the results of empirical studies are not always easy to interpret. If one strategy variant is observed to perform better than another, it is often not obvious what the reasons for the performance advantage of the former strategy are. Furthermore, it is unclear whether the observed results generalize to other settings of the parameters of the strategies or of the fitness environments. Moreover, it is not clear how difficult the task of tracking is, and how to interpret the quality of the results. Many of the experiments have been conducted with large populations in low-dimensional search spaces – a case that should arguably be comparatively easy to handle. It would therefore be desirable to have scaling laws that describe the influence of the parameters of the strategies and of the fitness environment. Not only would such scaling laws yield an improved understanding of the behavior of the strategies and their operators and parameters, but they would also allow for the analytical calculation of optimal strategy parameter values and for the comparison of the performance of different strategy variants.

In the realm of evolution strategies, much work has been done towards deriving such scaling laws in simple static environments. Many of the tools developed and the main results can be found in the monographs by Beyer [6] and by Rechenberg [11]. In the present paper, we will see that the tools developed for static environments can be applied to the analysis of evolution strategies for dynamic optimization in real-valued search spaces with relatively little effort. In particular, the tracking performance of the $(\mu/\mu, \lambda)$-ES with cumulative mutation strength adaptation is studied analytically for a spherically symmetric objective function the center of which is shifted randomly in every time step. For that purpose, in Sec. 2 we briefly introduce and motivate the choice of strategy and fitness environment. In Sec. 3, we analyze the behavior of the strategy in the environment thus introduced for fixed mutation strength. In Sec. 4, the performance of the mutation strength adaptation scheme is investigated. Finally, Sec. 5 concludes with a summary and a discussion of directions for future research.

2 Preliminaries

In all of what follows, we assume real-valued objective functions $\mathbb{R}^N \to \mathbb{R}$. The $(\mu/\mu, \lambda)$-ES with isotropic normal mutations in every time step generates $\lambda > \mu$ offspring candidate solutions from a population of μ parents and subsequently replaces the parental population by the μ best of the offspring. Generation of an offspring candidate solution consists in adding a vector $\sigma\mathbf{z}$, where $\mathbf{z}$ consists of independent, standard normally distributed components, to the centroid of the parental population. The standard deviation σ of the components of vector $\sigma\mathbf{z}$ is referred to as the *mutation strength*, vector $\mathbf{z}$ as the *mutation vector*. The average of those mutation vectors that correspond to offspring candidate solutions that are selected to form the population of the next time step is the *progress vector* $\langle\mathbf{z}\rangle$. Note that due to the definition of global intermediate recombination, $\sigma\langle\mathbf{z}\rangle$ connects consecutive centroids of the population. The choice of strategy is motivated both by the fact that it is relatively amenable to mathematical analysis and by its proven good performance in static settings.

No evolution strategy in real-valued search spaces is complete without a mutation strength adaptation mechanism. It is necessary for the mutation strength to be adapted continuously to fit the local characteristics of the objective function. Two mechanisms that are commonly used for the adaptation of the mutation strength are *mutative self-adaptation* and *cumulative mutation strength adaptation*. While the former is the more popular approach counting the number of publications, we choose to analyze cumulative mutation strength adaptation as it is known that mutative self-adaptation is unable to make full use of the genetic repair effect in combination with global intermediate recombination. Also note that while the original algorithm by Hansen and Ostermeier [9, 10] adapts the entire mutation covariance matrix, the variant considered here uses isotropic mutations and therefore only a single mutation strength.

The cumulative mutation strength adaptation mechanism relies on the conjecture that if the mutation strength is below its optimal value consecutive steps of the strategy tend to be parallel, and if the mutation strength is too high consecutive steps tend to be antiparallel. For optimally adapted mutation strength, the steps taken by the evolution strategy are uncorrelated. So as to be able to reliably detect parallel or antiparallel correlations between successive steps, information from a number of time steps needs to be accumulated. For the $(\mu/\mu, \lambda)$-ES, the *accumulated progress vector* $\mathbf{s}$ is defined by $\mathbf{s}^{(0)} = \mathbf{0}$ and the recursive relationship

$$\mathbf{s}^{(t+1)} = (1-c)\mathbf{s}^{(t)} + \sqrt{c(2-c)}\sqrt{\mu}\langle \mathbf{z} \rangle^{(t)}, \tag{1}$$

where c is a constant determining how far back the "memory" of the accumulation process reaches. The mutation strength is updated according to

$$\sigma^{(t+1)} = \sigma^{(t)} \exp\left(\frac{\|\mathbf{s}^{(t+1)}\|^2 - N}{2DN}\right), \tag{2}$$

where D denotes a damping constant. The term N in the numerator of the argument to the exponential function is the mean squared length of the accumulated progress vector if consecutive progress vectors are stochastically independent. The constants c and D are set to $1/\sqrt{N}$ and $\sqrt{N}$, respectively, according to recommendations made by Hansen [9].

The *sphere model* is the set of all functions $f : \mathbb{R}^N \to \mathbb{R}$ with

$$f(\mathbf{x}) = g(\|\hat{\mathbf{x}} - \mathbf{x}\|),$$

where $g : \mathbb{R} \to \mathbb{R}$ is a strictly monotonic function of the distance $R = \|\hat{\mathbf{x}} - \mathbf{x}\|$ of a candidate solution $\mathbf{x}$ from the target $\hat{\mathbf{x}}$. The sphere model usually serves as a model for fitness landscapes at a stage where the population of candidate solutions is in relatively close proximity to the target and is most often studied in the limit of very high search space dimensionality. So as to study the tracking behavior of evolutionary algorithms, several authors ([1, 4, 12]) have added a dynamic component to the sphere model by stipulating that the target $\hat{\mathbf{x}}$ varies with time. Several modes of motion of the target are conceivable and have been explored empirically. Examples include random motion, linear motion, and circular motion in search space. For the present paper, we restrict ourselves to considering random motion and assume that the target at time step $t+1$ is given by $\hat{\mathbf{x}}^{(t+1)} = \hat{\mathbf{x}}^{(t)} + \delta\hat{\mathbf{z}}$, where vector $\hat{\mathbf{z}}$ consists of N independent, standard normally distributed components. The standard deviation δ is a measure for the speed of the target. The same model has been considered by Angeline [1], Bäck [4], and Branke [7].

3 Dynamic Sphere

The analysis of the behavior of evolution strategies on the sphere model is greatly simplified by the symmetries inherent in both the strategies and the environment. Following an idea introduced by both Beyer [5] and Rechenberg [11], a vector $\mathbf{z}$ originating at search space location $\mathbf{x}$ can be written as the sum of two vectors $\mathbf{z}_A$ and $\mathbf{z}_B$, where $\mathbf{z}_A$ is parallel to $\hat{\mathbf{x}} - \mathbf{x}$ and $\mathbf{z}_B$ is in the hyperplane normal to that. We will refer to $\mathbf{z}_A$ and $\mathbf{z}_B$ as the *central* and the *lateral components* of vector $\mathbf{z}$, respectively. We define the *signed length* z_A of the central component

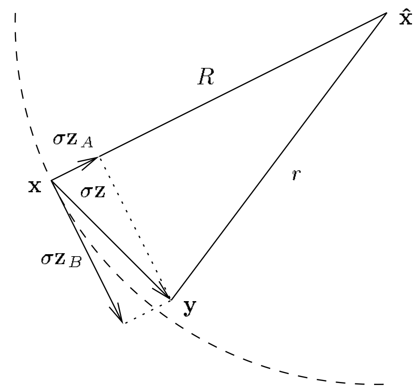

Fig. 1. *Decomposition of a vector* $\mathbf{z}$ *into central component* $\mathbf{z}_A$ *and lateral component* $\mathbf{z}_B$ *for the sphere model. Vector* $\mathbf{z}_A$ *is parallel to* $\hat{\mathbf{x}} - \mathbf{x}$*, vector* $\mathbf{z}_B$ *is in the hyperplane perpendicular to that. The starting and end points,* $\mathbf{x}$ *and* $\mathbf{y} = \mathbf{x} + \sigma\mathbf{z}$*, of vector* $\sigma\mathbf{z}$ *are at distances* R *and* r *from the target* $\hat{\mathbf{x}}$*, respectively.*

of vector $\mathbf{z}$ to equal $||\mathbf{z}_A||$ if $\mathbf{z}_A$ points towards the target and to equal $-||\mathbf{z}_A||$ if $\mathbf{z}_A$ points away from it. Fig. 1 illustrates the decomposition.

Using elementary geometry and denoting the respective distances of $\mathbf{x}$ and $\mathbf{y} = \mathbf{x} + \sigma\mathbf{z}$ from the target by R and r, it is easily seen that

$$\begin{aligned} r^2 &= (R - \sigma z_A)^2 + \sigma^2 ||\mathbf{z}_B||^2 \\ &= R^2 \left(1 - \frac{2\sigma}{R} z_A + \frac{\sigma^2}{R^2} ||\mathbf{z}||^2\right). \end{aligned}$$

Moreover, we will see that in high-dimensional search spaces, progress of the strategies requires $\sigma \ll R$, making it possible to expand the term in parentheses around unity and to cut off after the linear term, yielding

$$r = R - \sigma z_A + \frac{\sigma^2}{2R} ||\mathbf{z}||^2 + \ldots \tag{3}$$

for the distance of $\mathbf{y}$ from the target.

Let us first consider the case that $\mathbf{z}$ is a mutation vector. Then, as mutations are isotropic, we can without loss of generality assume that $z_A = z_1$ and $\mathbf{z}_B = (0, z_2, \ldots, z_N)^{\mathrm{T}}$, where the z_i, $i = 1, \ldots, N$, are independently drawn from a standardized normal distribution. The squared length $||\mathbf{z}_B||^2$ of the lateral component is the sum of squares of $N-1$ terms and as such χ^2_{N-1}-distributed. Recall that the χ^2_{N-1}-distribution has mean $N-1$ and standard deviation $\sqrt{2(N-1)}$. With increasing search space dimensionality, the influence of the central component $\mathbf{z}_A$ on the squared length $||\mathbf{z}||^2$ of a mutation vector decreases more and more. In the limit of infinite search space dimensionality, the quotient $(||\mathbf{z}||^2 - N)/N$ tends to zero, and thus $||\mathbf{z}||^2$ is well approximated by N. Therefore, according to Eq. (3), the difference in distance from the target of the centroid of the parental population and an offspring candidate solution has normal distribution with mean $-\sigma^2 N/2R$ and with variance σ^2. Note that it is only the central component of a mutation vector that determines the fitness of the offspring candidate solution; in the limit of infinite search space dimensionality, the contribution of the lateral component is the same for all offspring and thus selectively neutral. This has been shown more formally using an approach based on characteristic functions in [2].

For the sphere model, selection ensures that those offspring candidate solutions with the smallest values of r form the population of the next time step. Recombination averages those candidate solutions that are selected. As the lateral components of the mutation vectors are selectively neutral in that they contribute a constant to the fitness of the offspring they generate, they are independent. The lateral component of the progress vector is thus a vector in the plane perpendicular to the direction to the target and of squared length N/μ, where the reduction in squared length by the factor μ is due to the independence of the vectors being averaged. According to Eq. (3), the lateral component of the progress vector increases the distance to the target by $\sigma^2 N/2\mu R$. The reduction of that term by the factor μ in the denominator has been termed the *genetic repair effect*.

The central component of the progress vector is the average of those μ of the λ mutation vectors that have the largest signed lengths of the central components. The expected signed length of the central component of the progress vector is thus $c_{\mu/\mu,\lambda}$, where $c_{\mu/\mu,\lambda} \geq 0$ is the expected average of the first μ order statistics of a sample of λ standard normally distributed, independent random variables. An integral expression for $c_{\mu/\mu,\lambda}$ has been derived in [6]. According to Eq. (3), the central component of the progress vector reduces the distance to the target by $\sigma c_{\mu/\mu,\lambda}$. Note that while the *gain term* resulting from the central components is linear in σ, the *loss term* resulting from the lateral components grows quadratically. The range of useful mutation strengths is thus limited, requiring $\sigma \ll R$ as stipulated above.

So far, we have not considered the motion of the target. For the dynamic sphere, in every time step, the target takes a random step $\delta\hat{\mathbf{z}}$, where $\hat{\mathbf{z}}$ consists of N independent, standard normally distributed components. In effect, that is the same as the centroid of the population taking a random step $-\delta\hat{\mathbf{z}}$. The consequences of such a step have been investigated above: it leads to a increase in distance from the target by $\delta^2 N/2R$. Thus, the expected distance between the target and the centroid of the population at the next time step is

$$\mathrm{E}\left[R^{(t+1)}\right] = R^{(t)} - \sigma c_{\mu/\mu,\lambda} + \frac{\sigma^2 N}{2\mu R} + \frac{\delta^2 N}{2R}, \tag{4}$$

where the second and third terms on the right hand side result from the central and lateral components of the progress vector and the fourth summand is a consequence of the motion of the target.

After initialization effects have faded, the distance from the target of the centroid of the population has a time-invariant distribution. A good approximation

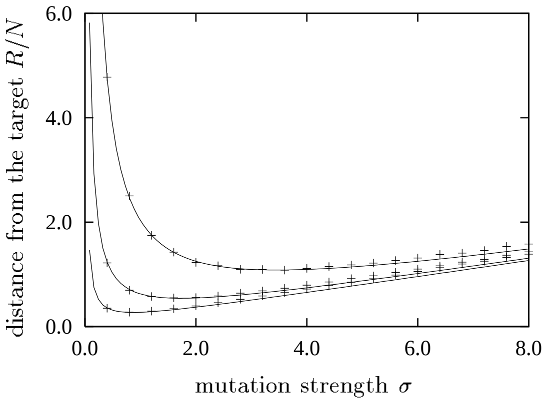

Fig. 2. *Distance from the target R/N for a $(3/3, 10)$-ES with, from bottom to top, $\delta = 0.5$, 1.0, and 2.0. Search space dimensionality is $N = 40$. The crosses are measured values, the lines reflect predictions from Eq.* (5).

to the mean of that distribution is obtained from Eq. (4) by ignoring fluctuations and simply demanding that $\mathrm{E}[R^{(t+1)}] = R^{(t)} = R$. Solving for R yields

$$R = \frac{N}{2c_{\mu/\mu,\lambda}} \left(\frac{\sigma}{\mu} + \frac{\delta^2}{\sigma} \right). \tag{5}$$

Fig. 2 shows that the resulting predictions are quite accurate. The accuracy increases further with increasing search space dimensionality. It can be seen that it somewhat decreases with increasing mutation strength as some of the assumptions made in the derivation of Eq. (3) are violated, but is very good in the range of mutation strengths where R is small.

Using Eq. (5), the optimal mutation strength can be found by computing the derivative with respect to σ and finding a root thereof. From

$$\frac{\partial R}{\partial \sigma} = \frac{N}{2c_{\mu/\mu,\lambda}} \left(\frac{1}{\mu} - \frac{\delta^2}{\sigma^2} \right)$$

it follows that $\sigma = \sqrt{\mu}\delta$ is optimal, and Eq. (5) shows that the corresponding distance from the target is

$$R = \frac{N\delta}{\sqrt{\mu} c_{\mu/\mu,\lambda}}. \tag{6}$$

That is, the distance from the target increases linearly with the speed of the target and can be decreased by increasing the population size. This decrease is due to the use of global intermediate recombination. To the reader acquainted with the theory of evolution strategies on the sphere model, it will not come as a surprise that from $\sigma = \sqrt{\mu}\delta$ with Eq. (6) it follows that the optimal mutation strength is $\sigma = \mu c_{\mu/\mu,\lambda} R/N$. This relationship has long been known to hold for the static sphere model. We now know that it holds as well when tracking a moving target with random dynamics.

4 Cumulative Mutation Strength Adaptation

Clearly, when using a mutation strength adaptation mechanism, there is no guarantee that the mutation strength that the strategy realizes is optimal; indeed, it often is not. The performance of the cumulative mutation strength adaptation mechanism on the static sphere model in the presence of noise has recently been analyzed in [3]. It has been found that even in the absence of noise, the mutation strength that the adaptation mechanism realizes is suboptimal due to the fact that the optimum is approached and that the approach is not instantaneous, leading to the strategy always trailing behind. The approach pursued in [3] can be adapted to the dynamic sphere model as follows. Note that due to space limitations, the derivation necessarily needs to remain sketchy at times.

As mutation vectors and progress vectors, the accumulated progress vector $\mathbf{s}$ can be written as the sum of its central and lateral components, $\mathbf{s}_A$ and $\mathbf{s}_B$. We write s_A for the signed length of the central component and as in Sec. 3 assume without loss of generality that at time t, the direction to the target is such that $s_A = s_1$. Using Eq. (1) and for notational simplicity writing $\mathbf{z}$ instead of $\langle\mathbf{z}\rangle$ for the progress vector, we have

$$\begin{aligned}
\|\mathbf{s}^{(t+1)}\|^2 &= \sum_{i=1}^{N}\left((1-c)s_i^{(t)} + \sqrt{c(2-c)}\sqrt{\mu}z_i^{(t)}\right)^2 \\
&= (1-c)^2\|\mathbf{s}^{(t)}\|^2 + 2(1-c)\sqrt{c(2-c)\mu}\sum_{i=1}^{N} s_i^{(t)}z_i^{(t)} + c(2-c)\mu\|\mathbf{z}^{(t)}\|^2.
\end{aligned}$$

The signed length of the central component of a vector equals the inner product of that vector with a vector of length unity from the centroid of the population to the target. Therefore, writing $\mathbf{x}$ for the centroid of the population and using Eq. (1), we have

$$\begin{aligned}
s_A^{(t+1)} &= \left((1-c)\mathbf{s}^{(t)} + \sqrt{c(2-c)}\sqrt{\mu}\mathbf{z}^{(t)}\right)^{\mathrm{T}} \frac{\hat{\mathbf{x}}^{(t)} + \delta\hat{\mathbf{z}}^{(t)} - \mathbf{x}^{(t)} - \sigma^{(t)}\mathbf{z}^{(t)}}{R^{(t+1)}} \\
&= \frac{R^{(t)}}{R^{(t+1)}}\left[(1-c)\left(s_A^{(t)} + \frac{\delta}{R^{(t)}}\sum_{i=1}^{N} s_i^{(t)}\hat{z}_i^{(t)} - \frac{\sigma^{(t)}}{R^{(t)}}\sum_{i=1}^{N} s_i^{(t)}z_i^{(t)}\right)\right. \\
&\qquad \left. +\sqrt{c(2-c)}\sqrt{\mu}\left(z_A^{(t)} + \frac{\delta}{R^{(t)}}\sum_{i=1}^{N}\hat{z}_i^{(t)}z_i^{(t)} - \frac{\sigma^{(t)}}{R^{(t)}}\|\mathbf{z}^{(t)}\|^2\right)\right].
\end{aligned}$$

These two equations together with Eq. (2) decribe the development from one time step to the next of the squared length of the accumulated progress vector, the signed length of its central component, and the mutation strength. In analogy to the way we proceeded in Sec. 3, we ignore fluctuations and assume that all quantities can be replaced by their expected values. Omitting time superscripts, for the stationary case it follows that

$$\|\mathbf{s}\|^2 = (1-c)^2\|\mathbf{s}\|^2 + 2(1-c)\sqrt{c(2-c)}\sqrt{\mu}s_A c_{\mu/\mu,\lambda} + c(2-c)N, \tag{7}$$

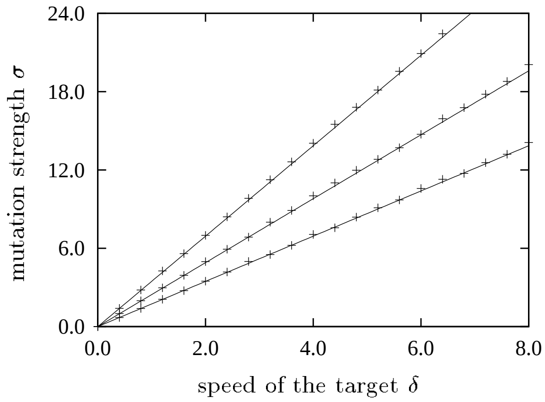

Fig. 3. *Mutation strength σ as a function of the speed of the target δ for, from bottom to top, a $(3/3, 10)$-ES, a $(6/6, 20)$-ES, and a $(12/12, 40)$-ES. Search space dimensionality is $N = 40$. The crosses are measured values realized by cumulative mutation strength adaptation, the lines reflect optimal values $\sigma = \sqrt{\mu}\delta$.*

where we have used the fact that $\mathrm{E}[z_i] = 0$ for $i \neq 1$ due to symmetry reasons, that

$$s_A = (1-c)s_A + \sqrt{c(2-c)}\sqrt{\mu}\left(c_{\mu/\mu,\lambda} - \frac{\sigma N}{\mu R}\right), \tag{8}$$

where $\sigma s_A c_{\mu/\mu,\lambda}/R$ has been neglected as compared to s_A as $\sigma \ll R$, and that

$$\sigma = \sigma \exp\left(\frac{||\mathbf{s}||^2 - N}{2DN}\right). \tag{9}$$

Solving Eqs. (7), (8), and (9) for the mutation strength yields $\sigma = \mu c_{\mu/\mu,\lambda} R/N$. Fig. 3 demonstrates the accuracy of the result. The resulting distance R to the target is given by Eq. (6). Comparison with the optimal mutation strength $\sigma = \sqrt{\mu}\delta$ derived in Sec. 3 shows that cumulative mutation strength adaptation achieves optimal performance on the sphere model with random dynamics of the target.

5 Conclusions

In this paper, we have studied the tracking performance of the $(\mu/\mu, \lambda)$-ES on a variant of the sphere model with random dynamics of the target. A scaling law that describes the dependence of the distance from the target on the mutation strength has been found and solved analytically for the optimal mutation strength. It has then been shown that cumulative mutation strength adaptation works perfectly in that the optimal mutation strength is realized by the algorithm. Thus, our results do not support the observation made by Branke [7] that in dynamic optimization, evolutionary algorithms lose the diversity necessary for exploring the search space.

A goal of future research is the analysis of the performance of the $(\mu/\mu, \lambda)$-ES on a variant of the sphere model where the motion of the target is deterministic and linear. Linear motion of the target can be expected to introduce correlations in the sequence of steps to be taken and thus to lead to qualitatively new insights in the performance of cumulative mutation strength adaptation.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants Be1578/4-2 and Be1578/6-3. The publication of the work was also supported by the DFG as part of the Collaborative Research Center "Computational Intelligence" (SFB 531). Hans-Georg Beyer is a Heisenberg Fellow of the DFG.

References

1. P. J. Angeline. Tracking extrema in dynamic environments. In *Proc. of the Sixth International Conference on Evolutionary Programming*, pages 335–345. Springer Verlag, Heidelberg, 1997.
2. D. V. Arnold and H.-G. Beyer. Local performance of the $(\mu/\mu_I, \lambda)$-ES in a noisy environment. In W. N. Martin and W. M. Spears, editors, *Foundations of Genetic Algorithms 6*, pages 127–141. Morgan-Kaufmann, San Francisco, 2000.
3. D. V. Arnold and H.-G. Beyer. An analysis of cumulative mutation stength adaptation. In preparation, 2002.
4. T. Bäck. On the behavior of evolutionary algorithms in dynamic environments. In *Proc. of the 1998 International Conference on Evolutionary Computation*, pages 446–451. IEEE Press, Piscataway, NJ, 1998.
5. H.-G. Beyer. Toward a theory of evolution strategies: Some asymptotical results from the $(1 \stackrel{+}{,} \lambda)$-theory. *Evolutionary Computation*, 1(2):165–188, 1993.
6. H.-G. Beyer. *The Theory of Evolution Strategies*. Natural Computing Series. Springer Verlag, Heidelberg, 2001.
7. J. Branke. *Evolutionary Optimization in Dynamic Environments*. Genetic Algorithms and Evolutionary Computation Series. Kluwer Academic Publishers, Dordrecht, 2001.
8. S. Droste. Analysis of the $(1+1)$ EA for a dynamically changing objective function. In *Proc. of the 2002 Congress on Evolutionary Computation*. IEEE Press, Piscataway, NJ, 2002.
9. N. Hansen. *Verallgemeinerte individuelle Schrittweitenregelung in der Evolutionsstrategie*. Mensch & Buch Verlag, Berlin, 1998.
10. N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.
11. I. Rechenberg. *Evolutionsstrategie '94*. Frommann-Holzboog, Stuttgart, 1994.
12. R. Salomon and P. Eggenberger. Adaptation on the evolutionary time scale: A working hypothesis and basic experiments. In *Proc. of the Third Conference on Artificial Evolution*, pages 251–262. Springer Verlag, Heidelberg, 1997.

13. K. Weicker and N. Weicker. On evolution strategy optimization in dynamic environments. In *Proc. of the 1999 Congress on Evolutionary Computation*, pages 2039–2046. IEEE Press, Piscataway, NJ, 1999.