Analysis of a Meta-ES on a Conically Constrained Problem

Michael Hellwig
Vorarlberg University of Applied Sciences
Dornbirn, Austria
michael.hellwig@fhv.at

Hans-Georg Beyer
Vorarlberg University of Applied Sciences
Dornbirn, Austria
hans-georg.beyer@fhv.at

ABSTRACT

The paper presents the theoretical performance analysis of a hierarchical Evolution Strategy (meta-ES) variant for mutation strength control on a conically constrained problem. Infeasible offspring are repaired by projection onto the boundary of the feasibility region. Closed-form approximations are used for the one-generation progress of the lower-level evolution strategy. An interval that brackets the expected progress over a single isolation period of the meta-ES is derived. Approximate deterministic evolution equations are obtained that characterize the upper-level strategy dynamics. It is shown that the dynamical behavior of the meta-ES is determined by the choice of the mutation strength control parameter. The obtained theoretical results are compared to experiments for assessing the approximation quality.

CCS CONCEPTS

• **Theory of computation → Bio-inspired optimization**; **Theory of randomized search heuristics**;

KEYWORDS

Constrained Optimization, Meta Evolution Strategies, Analysis, Mutation Strength Control

ACM Reference Format:
Michael Hellwig and Hans-Georg Beyer. 2019. Analysis of a Meta-ES on a Conically Constrained Problem. In *Proceedings of Genetic and Evolutionary Computation Conference (GECCO '19).* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3321707.3321824

1 INTRODUCTION

Among the class of Evolution Strategies (ES), hierarchical Evolution Strategies, or Meta-Evolution Strategies (meta-ES), represent an alternative approach for strategy parameter control [12, 13, 19]. A crucial property of an ES is the ability to properly control its internal mutation strength, i.e. the standard deviation of the offspring distribution that can roughly be associated with the length of a search-steps. In the context of mutation strength control on the Ellipsoid model, a simple meta-ES variant revealed a performance in between those of the σ-Self-Adaptation ES (σSA-ES) [17] and the Cumulative Step-Size Adaptation ES (CSA-ES) [16], but enhanced robustness with respect to strategy parameter variations [11]. Due to theoretical results on the Sharp Ridge test function which shares similar features with simple constrained problems [7], meta-ES approaches are considered successful in constrained settings. Hence, this paper investigates the ability of meta-ES to successfully control the mutation strength in a constrained setting.

Recently, the design and analysis of ES applicable to constrained optimization problems are of growing interest. Theoretical analyses reveal a deeper understanding of the ES working principles and their potential on such problems. Knowledge gained from theory yields information about what kind of problems are suitably solved by ES and guidelines for applying ES to constrained real-world problems. Theoretical investigations gain insights into algorithm behavior and allow for recommendations of strategy parameters. Further, theory supports the design of novel algorithmic ideas.

For the analysis of ES variants with different constraint handling mechanisms on problems with a single linear constraint, refer to [1, 3]. The study [10] revealed that the methods that repair infeasible solutions by projection are beneficial in such situations. Another idea for constraint handling is based on augmented Lagrangian constraint handling [4–6]. These investigations make use of a Markov chain model to analyze the one-generation behavior in the presence of a single or multiple linear constraints. A conically constrained problem is considered in [2] for the analysis of the $(1, \lambda)$-ES using a death penalty approach. Adding self-adaptive mutation strength control, the dynamical behavior of the $(1, \lambda)$-σSA-ES is studied in [20] by making use of the dynamical systems approach [14] on the same problem. That work was extended to recombinative σSA-ES in [22]. In [21] the analysis approach is transferred to mutation strength control by cumulative step-size adaptation (CSA). The aim of this paper is to analyze the performance of the non-recombinative $[1, 2(1, \lambda)^{\gamma}]$-meta-ES on the conically constrained problem. To keep the analysis simpler, the paper confines itself to isolation periods of length $\gamma = 1$. Hence, the present analysis represents the first step towards a complete comprehension of mutation strength control in the respective constrained environment.

The meta-ES variant is presented in Sec. 2. In Sec. 3 a description of the constrained problem and the repair mechanism for infeasible candidate solutions is provided. The theoretical analysis of the dynamical meta-ES behavior is going to be conducted in two steps. At first, considering a constant mutation strength, the microscopic dynamics of the lower-level (inner) ESs are examined in Sec. 4. The results are then used to obtain theoretical predictions for the upper-level strategy in Sec. 5. A description of the meta-ES dynamics for isolation periods of length one is provided. The results are then used to estimate the expected convergence rate of the strategy. All analysis steps are substantiated by comparisons with experimental results obtained from real ES runs. Section 6 concludes the paper with a discussion of the theoretical results.

GECCO '19, July13–17, 2019, Prague, Czech Republic

ACM ISBN 978-1-4503-6111-8/19/07…$15.00
https://doi.org/10.1145/3321707.3321824

Algorithm 1 The standard $(1, \lambda)$-ES with constant mutation strength and repair.

1: Initialize: $\sigma, \mathbf{y}, \lambda, \gamma$;
2: $g \leftarrow 0$;
3: **while** $g < \gamma$ **do**
4: **for** $i \leftarrow 1$ to λ **do**
5: $\mathbf{y}_i \leftarrow \mathbf{y}^{(g)} + \sigma\mathcal{N}(0, \mathbf{I})$; ▷ Mutation
6: $\mathbf{y}_i \leftarrow$Repair$(\mathbf{y}_i)$ ▷ Repair by *Projection*
7: $F_i \leftarrow F(\mathbf{y}_i)$; ▷ Evaluation
8: **end for**
9: $\mathbf{y}^{(g+1)} \leftarrow \mathbf{y}_{1;\lambda}$; ▷ Selection
10: $g \leftarrow g + 1$;
11: **end while**
12: **return** $[\mathbf{y}, F(\mathbf{y})]$;

2 THE META-ES ALGORITHM

This section is concerned with the description of the $[1, 2(1, \lambda)^\gamma]$-meta-ES which is investigated in the conically constrained setting. The meta-ES consists of two hierarchical levels. On the lower level, two similar ES evolve for a fixed number of generations γ (the isolation time) with constant strategy parameter. A single ES on the upper level evaluates the lower-level performance and updates the lower-level strategy parameters accordingly. The lower-level strategies are represented by the standard $(1, \lambda)$-ES. [8]. The corresponding pseudo code is presented in Alg. 1.

After initialization, the $(1, \lambda)$-Evolution Strategy generates a population of λ offspring. Therefore, the algorithm successively samples vectors $z \sim \mathcal{N}(0, \mathbf{I})$ from a multivariate standard normal distribution in line 5, scales these vectors with σ, and adds the product to the present $\mathbf{y}$. In line 6, the offspring is computed by a repair procedure if need be (otherwise it is just a copy of $\mathbf{y}_i$ in line 5). The parameter σ represents the standard deviation of the offspring distribution around the parental centroid. The parameter σ is commonly referred to as mutation strength of the Evolution Strategy. Having generated a feasible offspring candidate solution $\mathbf{y}_i$, its objective function value is evaluated in line 7.

The constrained problem under investigation assumes unrelaxable constraints of known analytical structure only, i.e. the objective function cannot be evaluated for an infeasible candidate solution.

Algorithm 2 The pseudo code of the $[1, 2(1, \lambda)^\gamma]$-meta-ES. The Code of the inner ES is displayed in Alg. 1.

1: Initialize: $\sigma^{(0)}, \mathbf{y}^{(0)}, \lambda, \gamma$;
2: $t \leftarrow 0$;
3: **repeat**
4: $\sigma_1 \leftarrow \sigma^{(t)}\alpha$;
5: $\sigma_2 \leftarrow \sigma^{(t)}/\alpha$;
6: $[\mathbf{y_1}, F(\mathbf{y}_1)] \leftarrow \mathrm{ES}(\lambda, \gamma, \sigma_1, \mathbf{y}^{(t)})$;
7: $[\mathbf{y_2}, F(\mathbf{y}_2)] \leftarrow \mathrm{ES}(\lambda, \gamma, \sigma_2, \mathbf{y}^{(t)})$;
8: $\sigma^{(t+1)} \leftarrow \sigma_{1;2}$;
9: $\mathbf{y}^{(t+1)} \leftarrow \mathbf{y_{1;2}}$;
10: $t \leftarrow t + 1$;
11: **until** termination condition

For this reason, the ES needs to apply a constraint handling technique that ensures the feasibility of the whole offspring population in each generation. The Repair subroutine in line 6 checks the feasibility and repairs infeasible solutions by projecting them onto the boundary of the feasible region. We refer to this method as *projection*. Having generated a population of λ feasible candidate solutions, the best candidate (w.r.t. the function values F_i) is selected. Note that the notation $1; \lambda$ refers to the best out of λ offspring of a certain generation. The main loop of Alg. 1 is iterated until the predefined isolation time γ is exceeded. The algorithm returns the best solution $\mathbf{y}$ of the final population together with the associated objective function value $F(\mathbf{y})$.

An internal mutation strength control mechanism (alike σSA or CSA) is omitted in Alg. 1. Instead, we add an upper-level ES that governs the mutation strength of the $(1, \lambda)$-ES. Note that the upper-level strategy considered in this report does only control the mutation strength σ of two inner strategies and keeps all other strategy parameters fixed, respectively. Refer to Alg. 2 for the pseudo code of the upper-level strategy.

The $[1, 2(1, \lambda)^\gamma]$-meta-ES represents a rather simple variant of a hierarchically organized Evolution Strategy. It employs two inner $(1, \lambda)$-ES which evolve from the same initial search space parameter vector $\mathbf{y_p}$ but with different mutation strength values $\sigma_1 = \sigma_p\alpha$ and $\sigma_2 = \sigma_p/\alpha$ with $\alpha > 1$. The mutation strength control is performed in between two consecutive isolation periods t and $t + 1$ of length γ. The best performing inner ES (with respect to the returned objective function value) passes its mutation strength $\sigma_{1;2}$ as well as the returned parameter vector $\mathbf{y_{1;2}}$ to the next iteration, or isolation period, respectively. Algorithm 2 terminates after reaching a maximal number of isolation periods or another stopping criterion (e.g. a predefined target precision).

3 PROBLEM FORMULATION

Consider the linear constrained optimization problem

$$\begin{aligned} \min_{\mathbf{y}\in\mathbb{R}^N} \; & F(\mathbf{y}) = cy_1 \\ \text{s.t. } & y_1^2 - \xi\sum_{i=2}^{N} y_i^2 \geq 0 \\ & y_1 \geq 0 \end{aligned} \tag{1}$$

with parameters $c > 0$, $\xi > 0$, linear objective function $F(\mathbf{y}) = cy_1$ and a quadratic constraint. Without loss of generality, the parameter c is assumed to be equal to one, i.e. $c = 1$. The two constraints in problem 1 define a cone-shaped feasible set in the search space. By construction, the cone is symmetrically placed around the first coordinate axis of the $\mathbb{R}^N$. Due to this symmetry, each point in the search space can be uniquely described by its distance x from 0 in direction of the cone axis y_1 and its perpendicular distance r from the cone axis. We refer to this representation as the $(x, r)^\top$-space. Note that due to the isotropy of the mutations used in the meta-ES, the coordinate system can w.l.o.g. be rotated. Thus, a particular candidate solution's distance from the cone axis coincides with the second component of this rotated coordinate system, i.e. $(\tilde{x}, \tilde{r})^\top$ corresponds to $(\tilde{x}, \tilde{r}, 0, \ldots, 0)^\top$.

A two-dimensional ($N = 2$) illustration of problem 1 is provided in Fig. 1. Since each point has a positive distance from the cone

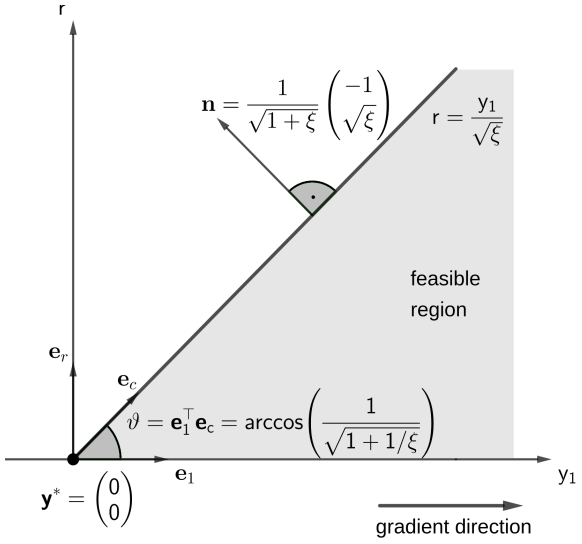

Figure 1: Two-dimensional illustration of the conically constrained optimization problem (1).

axis, only half of the cone is displayed. The cone boundary follows the equation $r = y_1/\sqrt{\xi}$. In the $(x, r)^\top$-space, it can be equivalently represented by the unit normal vector $\mathbf{n}$ of the constraint boundary

$$\mathbf{n}^\top \begin{pmatrix} x \\ r \end{pmatrix} = \frac{1}{\sqrt{1+\xi}} \begin{pmatrix} -1 \\ \sqrt{\xi} \end{pmatrix}^\top \begin{pmatrix} x \\ r \end{pmatrix} = 0. \tag{2}$$

The parameter $\xi > 0$ governs the opening angle of the cone which is decreasing with growing ξ. For $\xi = 1$ the cone exhibits an opening angle of 45° with the cone axis.

In order to deal with the constraint functions under consideration, there exist a number of different constraint handling techniques [15]. The following analysis confines itself to repair by projection. In case of the conically constrained problem, the projection operator that maps infeasible candidate solutions back onto the boundary of the feasible region is given by

$$\mathbf{y}_l \leftarrow \begin{cases} (\mathbf{e}_c^\top \mathbf{y}_l)\mathbf{e}_c, & \text{if } \mathbf{e}_c^\top \mathbf{y}_l > 0 \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

In this representation, the vector

$$\mathbf{e}_c = \frac{1}{\sqrt{1+1/\xi}} \left(\mathbf{e}_1 + \frac{\mathbf{e}_r}{\sqrt{\xi}} \right) \tag{4}$$

denotes the unit vector in direction of the cone boundary, $\mathbf{e}_1$ is the unit vector in direction of the y_1 axis, and $\mathbf{e}_r$ refers to the unit vector in r direction.

4 THE INNER DYNAMICS

For a description of the meta-ES, the one-generation progress of the inner $(1, \lambda)$-ES with fixed mutation strength $\sigma^{(g)} = \sigma = const.$ must be considered, first. The starting point is the progress rate theory obtained in the context of self-adaptive ES [20]. In the same way, the following analysis assumes that fluctuation terms can be neglected in the asymptotic limit case $N \to \infty$. Hence, the transition from generation g to generation $g + 1$ can be described by the system of deterministic evolution equations

$$x^{(g+1)} = x^{(g)} \left(1 - \varphi_x^{(g)*}/N \right) \tag{5}$$

$$r^{(g+1)} = r^{(g)} \left(1 - \varphi_r^{(g)*}/N \right), \tag{6}$$

with normalized progress rates φ_x^* and φ_r^* in x, and in r, direction, respectively. The normalized progress rates strongly depend on the probability of whether a generated offspring candidate solution is initially feasible or infeasible (and needs to be repaired). To this end, the paper [20] derived the probability P_{fea} that an initially feasible candidate solution is generated by the ES. In the asymptotical limit of large N, the probability is determined by the state variables of generation g as

$$P_{\text{fea}} \simeq \Phi \left[\frac{1}{\sigma^{(g)}} \left(\frac{x^{(g)}}{\sqrt{\xi}} - \bar{r} \right) \right] \text{ with } \bar{r} = r^{(g)} \sqrt{1 + \frac{\sigma^{(g)*2}}{N} \left(1 - \frac{1}{N} \right)}, \tag{7}$$

where $\bar{r}$ is the expected value of the r normal approximation and Φ is the cumulative distribution function of a standard normal variate. The probability that the offspring needs to be repaired is $P_{\text{inf}} = 1 - P_{\text{fea}}$. The overall expected normalized progress φ_x^* in direction of the cone axis x is then obtained by combining the progress rates calculated for both extreme cases and weighting them with the respective probabilities [20]

$$\begin{aligned} \varphi_x^{(g)*} &:= \varphi_x^*(x^{(g)}, r^{(g)}, \sigma^{(g)}) = \mathrm{E}\left[x^{(g)} - x^{(g+1)} | x^{(g)}, r^{(g)}, \sigma^{(g)} \right] \\ &\approx P_{\text{fea}} \left[\frac{r^{(g)}}{x^{(g)}} \sigma^{(g)*} c_{1,\lambda} \right] + P_{\text{inf}} \cdot \varphi_{x_{\text{inf}}}^{(g)*}, \end{aligned} \tag{8}$$

with

$$\begin{aligned} \varphi_{x_{\text{inf}}}^{(g)*} =& \frac{N}{1+\xi} \left(1 - \frac{\sqrt{\xi} r^{(g)}}{x^{(g)}} \sqrt{1 + \frac{\sigma^{(g)*2}}{N}} \right) \\ &+ \frac{\xi}{1+\xi} \frac{r^{(g)}}{x^{(g)}} \sigma^{(g)*} c_{1,\lambda} \sqrt{1 + \frac{1}{\xi} \frac{1 + \frac{\sigma^{(g)*2}}{2N}}{1 + \frac{\sigma^{(g)*2}}{N}}} \end{aligned} \tag{9}$$

Here, the term $c_{1,\lambda}$ refers to the progress coefficient introduced in [18]. It is only depending on the offspring population size λ. Furthermore, the normalized progress rate φ_r^* in r direction reads [20]

$$\begin{aligned} \varphi_r^{(g)*} &:= \varphi_r^*(x^{(g)}, r^{(g)}, \sigma^{(g)}) = \mathrm{E}\left[r^{(g)} - r^{(g+1)} | x^{(g)}, r^{(g)}, \sigma^{(g)} \right] \\ &\approx P_{\text{fea}} N \left(1 - \sqrt{1 + \frac{\sigma^{(g)*2}}{N}} \right) + P_{\text{inf}} N \left(1 - \frac{x^{(g)}}{\sqrt{\xi} r^{(g)}} \left(1 - \frac{\varphi_{x_{\text{inf}}}^{(g)*}}{N} \right) \right). \end{aligned} \tag{10}$$

In both progress representations, the term $\sigma^{(g)*} = \sigma N / r^{(g)}$ represents the normalized mutation strength of the strategy. The use of these progress rates is justified by iterating the system of evolution equations (5), (6), and comparing the results to experimental runs of the inner $(1, \lambda)$-ES that operates with fixed mutation strength σ, refer to Fig. 2. All runs are initialized close to the cone axis $\mathbf{y}_{\text{init}} = (\sqrt{N}, 10^{-3}, 0, \ldots, 0)^\top$ with $\sigma = 1 = const.$ to ensure feasibility and to demonstrate the good agreement even in the transient phase of the dynamics. Different values of the cone parameter ξ are considered.

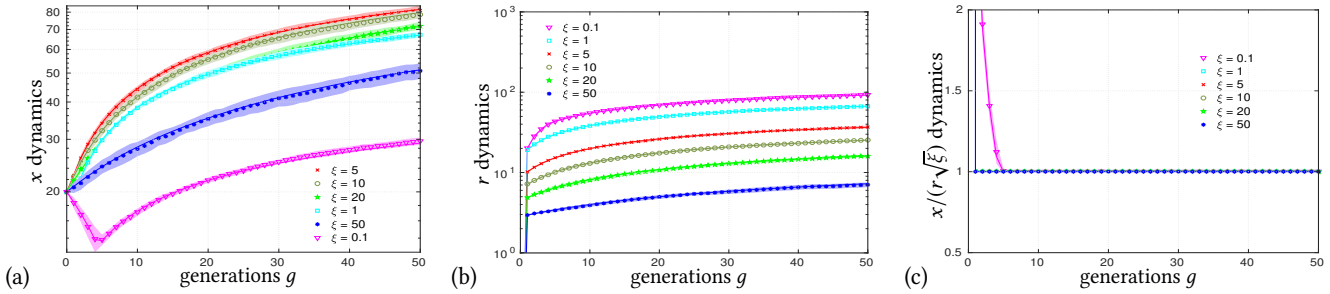

Figure 2: Comparison of the experimental $(1, 10)$-ES dynamics (data points displayed by the 'x' markers) and the theoretical dynamics that are obtained by iterating the evolutionary system (5), (6) (solid lines). The predictions use the progress rate representations from Eq. (8) and (10). All dynamics are obtained for constant mutation strength $\sigma = 1$ in $N = 400$. The experimental runs are averaged over 20 independent algorithm runs and the standard deviation is indicated by use of the error band plots.

Assuming that the ES operates sufficiently close to the constraint boundary, the transient phase is almost skipped and the steady-state behavior is observable comparably quickly. In its steady-state, the ES is likely to generate infeasible solutions that are then repaired by the projection operator. Hence, $P_{\text{inf}} \approx 1$ can be assumed and the corresponding progress will be determined by Eq. (9) (x-direction) and the second addend of Eq. (10) (r-direction) alone.

The progress rate (9) can be further simplified. In the asymptotic limit of $N \to \infty$, the quadratic normalized mutation strength ${\sigma^{(g)*}}^2$ is supposed to be considerably smaller than the search space dimension N, i.e. ${\sigma^{(g)*}}^2 \ll N$. Accordingly, the second square root in Eq. (9) is asymptotically equal to

$$\sqrt{1 + \frac{1}{\xi}\frac{1 + \frac{{\sigma^{(g)*}}^2}{2N}}{1 + \frac{{\sigma^{(g)*}}^2}{N}}} \simeq \sqrt{1 + \frac{1}{\xi}}, \tag{11}$$

and the normalized progress rate (9) (conditional on $P_{\text{fea}} \approx 0$) in x-direction becomes

$$\varphi_x^{(g)*} = \frac{N}{1+\xi}\left(1 - \frac{\sqrt{\xi}r^{(g)}}{x^{(g)}}\sqrt{1 + \frac{{\sigma^{(g)*}}^2}{N}}\right) + \frac{c_{1,\lambda}\sigma^{(g)*}}{\sqrt{1+\xi}}\frac{\sqrt{\xi}r^{(g)}}{x^{(g)}}. \tag{12}$$

The dynamics presented in Fig. 2 indicate that the $(1, \lambda)$-ES, see Alg. 1, approaches a constant steady-state ratio of the $x^{(g)}$ and the $r^{(g)}$ states. This ratio is proportional to the square root of cone parameter ξ. The geometrical interpretation is that the ES moves towards the optimizer along the cone boundary in its steady-state. As a result, the constant quotient $x^{(g)}/(r^{(g)}\sqrt{\xi}) = 1$ can be inserted into the progress rate representations yielding

$$\varphi_x^{(g)*}(\sigma^{(g)*}) = \frac{N}{1+\xi}\left(1 - \sqrt{1 + \frac{{\sigma^{(g)*}}^2}{N}}\right) + \frac{c_{1,\lambda}\sigma^{(g)*}}{\sqrt{1+\xi}}, \tag{13}$$

as well as

$$\varphi_r^{(g)*}(\sigma^{(g)*}) = N\left(1 - \left(1 - \frac{\varphi_x^{(g)*}(\sigma^{(g)*})}{N}\right)\right) = \varphi_x^{(g)*}(\sigma^{(g)*}). \tag{14}$$

The very good approximation quality of these equations is illustrated in Fig. 3. All dynamics are obtained for constant mutation strength $\sigma = 1$ in $N = 400$. The starting point is located on the cone boundary, $\mathbf{y}_{\text{init}} = (\sqrt{N}, \sqrt{N/\xi}, 0, \ldots, 0)^{\top}$, to illustrate the steady state behavior. Infeasible candidate solutions are repaired by use of the projection operator (3).

Figure 4 illustrates the corresponding $1, 2[(1, 10)]^1$-meta-ES dynamics. To this end, the iterative dynamics are obtained by iterating the evolutionary system (5), (6) and updating the mutation strength according to Alg. 2 after each isolation period of $\gamma = 1$.

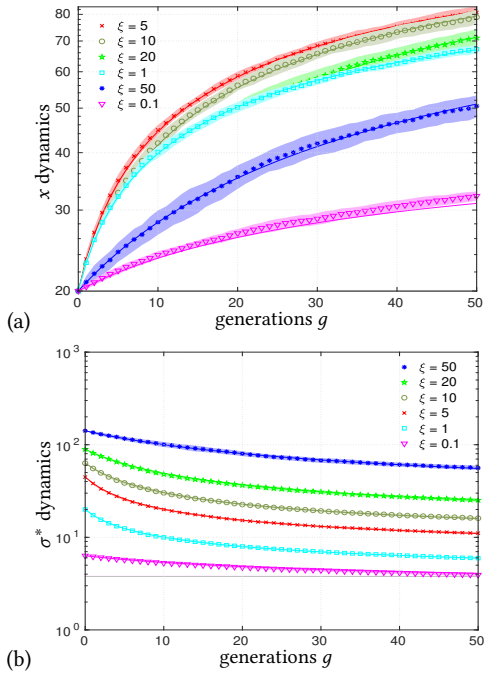

Figure 3: Averaged over 20 runs, the experimental $(1, 10)$-ES results are given by the data points. The standard deviation is indicated by use of the error band plots. The theoretical dynamics are illustrated by solid lines that are obtained by iterating system (5), (6). The theoretical predictions use the progress rates (13), and (14), respectively.

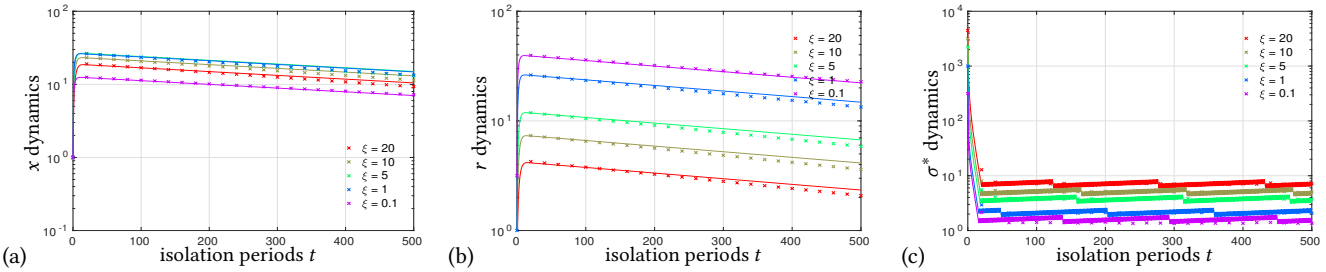

Figure 4: Comparison of the experimental $[1, 2(1, 10)]^1$-meta-ES dynamics (data points displayed by the 'x' markers) and the theoretical dynamics that are obtained by iterating system (5), (6) and by making use of the progress rates in Eq. (13), and (14), respectively. All dynamics are obtained by making use of $\alpha = 1.2$ in $N = 1000$. The experimental runs are averaged over 1000 independent meta-ES algorithm runs to dampen the fluctuations that occur due to the σ update of the upper-level ES.

5 THE OUTER DYNAMICS

Based on the considerations above, the next step is concerned with the analysis of the behavior of the non-recombinative $[1, 2(1, \lambda)^\gamma]$-meta-ES in Alg. 2 that controls the mutation strength, cf. Sec. 2.

The mutation strength control within Alg. 2 is governed by the difference of the expected fitness values of the candidate solutions returned by the inner ESs. Considering the objective of (1), it is evident that the expected fitness change from generation g to $g + 1$ is determined by the progress rate in direction of the cone axis (8). Assuming that the meta-ES is already operating in the vicinity of the cone boundary (where $P_{\text{inf}} \approx 1$ holds), the progress rate approximation (13) qualifies to describe the expected fitness at the end of each isolation period.

This analysis is confined to isolation periods of length $\gamma = 1$, i.e. the inner ESs are compared (and updated) after a single generation. The transition of a candidate solution before isolation $\mathbf{y}_1^{(t)} = \mathbf{y}_1^{(g_0)}$ to the candidate solution after the isolation period $\mathbf{y}^{(t+1)} = \mathbf{y}_1^{(g_0+\gamma)}$ is then given by the evolution equations (5), (6). Consequently, the expected fitness values returned by the inner ESs after isolation period $t + 1$ are given as

$$\begin{aligned} F_1^{(t+1)} &= F(\mathbf{y}_1^{(t+1)}) = x_1^{(t+1)} = x^{(t)} \left(1 - \varphi_x^{*(t)}(\alpha\sigma^*)/N\right) \\ F_2^{(t+1)} &= F(\mathbf{y}_2^{(t+1)}) = x_2^{(t+1)} = x^{(t)} \left(1 - \varphi_x^{*(t)}(\sigma^*/\alpha)/N\right). \end{aligned} \tag{15}$$

Their difference $D^{(t+1)} := F_1^{(t+1)} - F_2^{(t+1)}$ determines the mutation strength adaptation in the meta-ES strategy. Accordingly, the mutation strength σ control rule becomes

$$\begin{aligned} \sigma^{(t+1)} &= \alpha\sigma^{(t)} && \text{if} \quad D^{(t+1)} < 0, \\ \sigma^{(t+1)} &= \sigma^{(t)}/\alpha && \text{if} \quad D^{(t+1)} > 0. \end{aligned} \tag{16}$$

Reconsidering the normalized progress rate representation (13) yields the difference $D^{(t+1)}$ of the fitness values as

$$\begin{aligned} D^{(t+1)} = \frac{x^{(t)}}{1+\xi} \Bigg(&\sqrt{1 + \frac{\alpha^2 \sigma^{*(t)^2}}{N}} - \frac{c_{1,\lambda}\alpha\sigma^{*(t)}\sqrt{1+\xi}}{N} \\ &- \sqrt{1 + \frac{\sigma^{*(t)^2}}{\alpha^2 N}} + \frac{c_{1,\lambda}\sigma^{*(t)}\sqrt{1+\xi}}{\alpha N} \Bigg). \end{aligned} \tag{17}$$

A simplification of this expression can be obtained by reusing the assumption $\sigma^{(g)*^2} \ll N$ and by expanding $\sqrt{1 + \sigma^{*^2}/N}$ into its Taylor series around zero (and neglecting higher order terms). One obtains

$$D^{(t+1)} = \frac{c_{1,\lambda}\sigma^{*(t)}x^{(t)}(\alpha^2 - 1)}{\alpha N\sqrt{1+\xi}} \left(\frac{\sigma^{*(t)}}{2c_{1,\lambda}\sqrt{1+\xi}} \left(\alpha + \frac{1}{\alpha}\right) - 1 \right). \tag{18}$$

Since $\alpha > 1$, the sign of the difference depends on the discriminant

$$\Delta(\sigma^{*(t)}) := \left(\frac{\sigma^{*(t)}}{2c_{1,\lambda}\sqrt{1+\xi}} \left(\alpha + \frac{1}{\alpha}\right) - 1 \right), \tag{19}$$

and the evolution equation (16) of the mutation strength becomes

$$\sigma^{(t+1)} = \sigma^{(t)} \alpha^{-\text{sign}\left(\Delta(\sigma^{*(t)})\right)}. \tag{20}$$

Taking into account the normalization $\sigma^{*(t)} = \sigma^{(t)} N/r^{(t)}$, Eq. (20) yields the evolution equation of the normalized mutation strength

$$\frac{\sigma^{*(t+1)} r^{(t+1)}}{N} = \frac{\sigma^{*(t)} r^{(t)}}{N} \alpha^{-\text{sign}\left(\Delta(\sigma^{*(t)})\right)}. \tag{21}$$

By multiplication with $N/r^{(t+1)}$ and by taking into account the evolution equation (6) together with the progress rate approximation (14), Eq. (21) reads

$$\sigma^{*(t+1)} = \sigma^{*(t)} \frac{\alpha^{-\text{sign}\left(\Delta(\sigma^{*(t)})\right)}}{1 - \frac{1}{N}\varphi_x^*\left(\sigma^{*(t)}\alpha^{-\text{sign}\left(\Delta(\sigma^{*(t)})\right)}\right)}. \tag{22}$$

Equation (22) represents a complicated iterative mapping of the form $\sigma^{*(t+1)} = f_\sigma(\sigma^{*(t)}; \alpha, N)$. This kind of recurrence equation does not allow for the straight forward analysis of parameter influences on the normalized mutation strength dynamics. It can exhibit different qualitative dynamics: stable fixed points, limit cycles, or chaotic behaviors. Studying the qualitative behavior in Fig. 5, a point of discontinuity can be observed. It corresponds to a change in the σ adaptation pattern of the meta-ES and can be calculated as the zero of the discriminant function $\Delta(\sigma_0^*) = 0$. We obtain

$$\sigma_0^* = 2c_{1,\lambda}\sqrt{1+\xi}\frac{\alpha}{1+\alpha^2}. \tag{23}$$

The point of discontinuity depends on the cone parameter as well as on the mutation strength control parameter α. Different dynamics can be observed depending on α: If the fixed point ${\sigma_f}^*$ of the iterative mapping (22) is at σ_0^*, it cannot be a stable attractor, cf.

Fig. 5. Due to $\mathrm{d}f_\sigma(\sigma^*)/\mathrm{d}\sigma^* > 1$, the other fixed point at $\sigma^* = 0$ is also unstable. Hence, the case in Fig. 5(b) displays a limit cycle.

Being interested in finding the influence factors that realize a continuous mutation strength reduction, consider the case where the fixed point of (22) satisfies ${\sigma_f}^* > \sigma_0^*$, i.e. $\Delta({\sigma_f}^*) > 0$. Considering the fixed point condition $\sigma^{*(t+1)} = \sigma^{*(t)}$, Eq. (22) becomes

$$1 = \frac{1}{\alpha}\frac{1}{1-\frac{1}{N}\varphi_x^*\left(\sigma^{*(t)}/\alpha\right)}. \tag{24}$$

After straight forward transformations, and by using the linear part of the Taylor expansion of $\sqrt{1+{\sigma^*}^2/N}$ around zero, one obtains a quadratic equation in ${\sigma_f}^*$

$${\sigma_f^*}^2 - 2\alpha c_{1,\lambda}\sqrt{1+\xi}\sigma_f^* + 2(1+\xi)N(\alpha-1)\alpha = 0. \tag{25}$$

Hence, its solution reads

$$\sigma_{f\pm}^* = \alpha c_{1,\lambda}\sqrt{1+\xi}\left(1 \pm \sqrt{1-\frac{2N}{c_{1,\lambda}^2}\frac{\alpha-1}{\alpha}}\right). \tag{26}$$

Depending on the magnitude of the square root in Eq. (26), the solution ${\sigma_{f-}}^*$ may be of relevance or not. If $\sigma_{f-}^* < \sigma_0^*$, it is not a solution of the iterative mapping because it violates the assumption $\Delta(\sigma_f^*) > 0$. Accordingly, one obtains the fixed point as

$$\sigma_{f+}^* = \alpha c_{1,\lambda}\sqrt{1+\xi}\left(1 + \sqrt{1-\frac{2N}{c_{1,\lambda}^2}\frac{\alpha-1}{\alpha}}\right). \tag{27}$$

Otherwise, if $\sigma_{f-}^* > \sigma_0^*$ holds, two fixed points must be considered. In this situation, the one which satisfies $\mathrm{d}f_\sigma(\sigma^*)/\mathrm{d}\sigma^* < 1$ turns out to be the stable attractor.

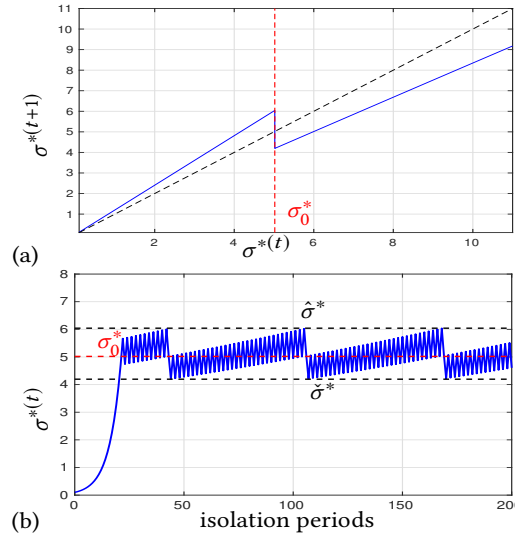

Figure 5: On the σ^* dynamics of the $[1, 2(1,10)^1]$-meta-ES for $N = 400$, $\xi = 10$. An unstable fixed point of Eq. (22) is displayed in (a). This corresponds to the limit cycle illustrated in (b). The point of discontinuity σ_0^* is indicated by the dashed red line. For $\hat{\sigma}$ and $\check{\sigma}$ refer to Eqs. (29) and (30).

However, the limiting case for the existence/non-existence of a stable fixed point, is governed by the radicand in (27). The transition point α_0 can be computed as the root of the radicand with respect to α

$$\alpha_0 = 1 + \frac{c_{1,\lambda}^2}{2N - c_{1,\lambda}^2} = \frac{2N}{2N - c_{1,\lambda}^2}. \tag{28}$$

This expression emphasizes the impact of the search space dimension N, and the offspring population size λ, on the mutation strength control parameter α. For given N and λ, a stable fixed point can only be expected for mutation strength control parameters $\alpha < \alpha_0$. If the condition $\alpha < \alpha_0$, and thus $\Delta(\sigma_f^*) > 0$, holds, the meta-ES is expected to gradually decrease the mutation strength σ with each isolation period (cf. Eq. (20)). On the other hand, if $\alpha > \alpha_0$ holds, the meta-ES will not approach a stable fixed point. Instead, the σ^* dynamics will fluctuate around the point of discontinuity σ_0^* leading to both σ-decreasing and σ-increasing steps.

In the latter case, the range of normalized mutation strength values observed in the limit cycle can be bracketed. The σ^* values are confined in the interval between the left-sided and right-sided limit of Eq. (22) at σ_0^* (see Fig. 5(a)). With $\varphi_+^* = \varphi_x^*(\alpha\sigma_0^*)$ as well as $\varphi_-^* = \varphi_x^*(\sigma_0^*/\alpha)$ from Eq. (13), one obtains

$$\begin{aligned}\hat{\sigma}^* &:= \lim_{\sigma^*\to\sigma_0^*-0}\frac{\alpha\sigma^*}{1-\frac{1}{N}\varphi_x^*(\alpha\sigma^*)} \\ &= \frac{\sigma_0^*\alpha}{1-\frac{1}{N}\varphi_+^*} = 2c_{1,\lambda}\sqrt{1+\xi}\frac{\alpha^2}{1+\alpha^2}\frac{1}{1-\frac{1}{N}\varphi_+^*},\end{aligned} \tag{29}$$

and

$$\begin{aligned}\check{\sigma}^* &:= \lim_{\sigma^*\to\sigma_0^*+0}\frac{\sigma^*/\alpha}{1-\frac{1}{N}\varphi_x^*(\sigma^*/\alpha)} \\ &= \frac{\sigma_0^*}{\alpha(1-\frac{1}{N}\varphi_-^*)} = 2c_{1,\lambda}\sqrt{1+\xi}\frac{1}{1+\alpha^2}\frac{1}{1-\frac{1}{N}\varphi_-^*},\end{aligned} \tag{30}$$

where

$$\varphi_+^* = \varphi_-^* = \frac{2c_{1,\lambda}^2\alpha^2}{(1+\alpha^2)^2} \tag{31}$$

can be obtained by Taylor expansion of $\sqrt{1+{\sigma^*}^2/N}$ around zero in Eq. (13). Provided that $\alpha > \alpha_0$, the expected value of the normalized mutation strength is bracketed in

$$\sigma^* \in [\check{\sigma}^*, \hat{\sigma}^*] = [1, \alpha^2]\cdot\frac{2c_{1,\lambda}\sqrt{1+\xi}}{1+\alpha^2}\frac{1}{1-\frac{1}{N}\varphi_+^*}. \tag{32}$$

Considering the σ^* interval (32), the expected meta-ES progress in x direction $\tilde{\varphi}_x^*$ will be contained in the interval

$$\tilde{\varphi}_x^* \in \left[\varphi_x^*(\check{\sigma}^*), \varphi_x^*(\sigma_{\mathrm{opt}}^*)\right] = \left[\left(\frac{2\alpha}{(1+\alpha^2)}\right)^2, 1\right]\cdot\hat{\varphi}_x^*, \tag{33}$$

where $\hat{\varphi}_x^* = c_{1,\lambda}^2/2$ represents the maximal progress which is derived from the progress rate (13) in the asymptotical limit case ($N \to \infty$). Figure 6 illustrates the limits derived in Eq. (32) and presents the influence of α on the normalized steady state mutation strength. To this end, the theoretical predictions are compared with experimental meta-ES algorithm runs on problem (1). Single runs of the $[1, 2(1,\lambda)^1]$-meta-ES have been executed for different α

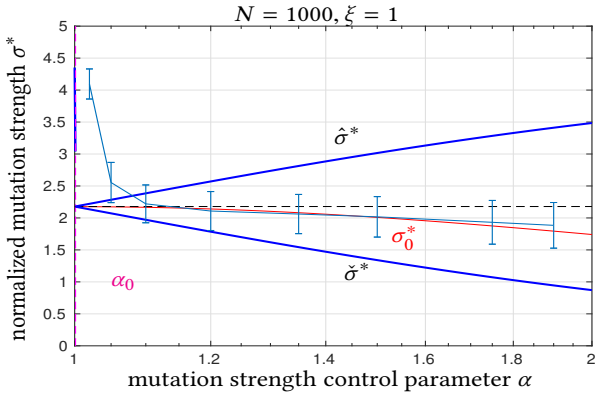

Figure 6: The σ^* distribution of the $[1, 2(1, 10)^1]$-meta-ES for different choices of α. The experimental results displayed by the error bar plot represent mean and standard deviation of the normalized mutation strength measured during the last 90% of 10^4 isolation periods of length $\gamma = 1$ (obtained from 50 independent meta-ES runs). The lower and upper limit, $\hat{\sigma}$ and $\check{\sigma}$, are displayed by the solid blue lines, see Eqs. (29), (30). The solid red line shows the location of σ_0^* within these bounds. The dashed vertical line close to $\alpha = 1$ indicates the position of α_0 (in this case $\alpha_0 \approx 1.001$), while the dashed horizontal line displays $\sigma^*_{\text{opt}} \simeq c_{1,\lambda}\sqrt{1+\xi}$.

values ($\alpha \in \{1.02, 1.05, 1.1, 1.2, 1.35, 1.5, 1.7, 1.9\}$). The empirically obtained data are represented by the error bar plot. They display the mean and standard deviation of the normalized mutation strength values of the last 90% of 10^4 isolation periods.

The dynamical meta-ES behavior is characterized by the above derivations. However, the question of how to choose the control parameter α in the conically constrained setting (1) still needs to be discussed. The σ^* adaptation should be fast and reliable while it would be desirable to generate the best average performance (i.e., the progress towards the optimizer) possible. Choosing $\alpha < \alpha_0$ will drive the meta-ES to approach the stable fixed point (27) of the iterative mapping (22). Yet, it is obvious that this fixed point cannot be shifted towards the normalized mutation strength value $\sigma^*_{\text{opt}} = c_{1,\lambda}\sqrt{1+\xi}$ that ensures maximal progress with respect to Eq. (13) (since $\alpha > 1$ must hold). Consequently, choosing $\alpha < \alpha_0$ will result in suboptimal progress. Further, such α values that satisfy $\alpha < \alpha_0$ turn out to be very close to 1 (for large N) in the case of the $[1, 2(1, \lambda)^1]$-meta-ES (see Fig. 6). Hence, the σ dynamics are only gradually decreasing in this case. Further, it is shown in [9] (Section 5.2.2) that the $[1, 2(\mu/\mu_I, \lambda)^1]$-meta-ES dynamics are subject to large fluctuations that impede the prediction quality of the deterministic evolution equations for very small α values. The best average performance of the meta-ES is realized for $\alpha > \alpha_0$, i.e. with normalized mutation strength oscillating around the point of discontinuity σ_0^*. The point of discontinuity σ_0^* always resides within the limits of the normalized mutation strength (32). In Fig. 6, it is represented by the solid red line. According to Eq. (23), σ_0^* depends on the population size λ, the cone parameter ξ, as well as on the mutation strength control parameter α. With growing α values, the point of discontinuity will slowly decrease. Yet, moderate choices of α yield σ_0^* values in close vicinity of the optimal normalized mutation strength value $\sigma^*_{\text{opt}} = c_{1,\lambda}\sqrt{1+\xi}$.

Assume that the meta-ES on average operates with σ^* in close proximity to the point of discontinuity σ_0^* in its limit cycle. This allows for modeling the steady state dynamics of the normalized mutation strength by a mean value $\overline{\sigma^*_{ss}}$ and corresponding fluctuation parts $\epsilon_{\sigma^*_{ss}}$ as $\sigma^*_{ss} = \overline{\sigma^*_{ss}} + \epsilon_{\sigma^*_{ss}}$. In case that the mean value dynamics of this equation are sufficiently characterized by σ_0^* one can approximate the steady state behavior of the meta-ES. Omitting of the fluctuation term, by using Eq. (23) one obtains

$$\sigma^*_{ss} \simeq \sigma_0^* = 2c_{1,\lambda}\sqrt{1+\xi}\,\frac{\alpha}{1+\alpha^2}. \tag{34}$$

Hence, the steady state fitness dynamics of the $[1, 2(1, \lambda)^1]$-meta-ES can be derived by making use of the progress rate (13). Ignoring the fluctuations, the fitness change from one isolation period to the next is determined by

$$x^{(t+1)} \simeq x^{(t)}\left(1 - \varphi^*_x(\sigma_0^*)/N\right). \tag{35}$$

Division by $x^{(t)}$, inserting (34) into Eq. (13), and taking into account the limit $t \to \infty$, yields the $[1, 2(1, \lambda)^1]$-meta-ES convergence rate $\nu := \lim_{t\to\infty} x^{(t+1)}/x^{(t)} = 1 - \varphi^*_x(\sigma_0^*)/N < 1$ for mutation strength control parameters $\alpha > \alpha_0$. Considering (13), it reads

$$\nu \simeq \frac{1}{1+\xi}\left(\xi + \sqrt{1 + \frac{4c^2_{1,\lambda}\alpha^2(1+\xi)}{N(1+\alpha^2)^2}}\right) - \frac{2c^2_{1,\lambda}\alpha}{N(1+\alpha^2)}. \tag{36}$$

Making use of the Taylor expansion of the square root ($N \to \infty$), the asymptotical convergence rate becomes

$$\nu \simeq 1 - \frac{2c^2_{1,\lambda}\alpha}{N(1+\alpha^2)}\left(1 - \frac{\alpha}{1+\alpha^2}\right). \tag{37}$$

Figure 7 illustrates the convergence rates predicted by Eq. (37) for different values of α in dimension $N = 400$ and $N = 1000$. The results are compared to experimental convergence rates obtained by measuring $\nu = 1 - \varphi^*_x(\sigma_0^*)/N$ over the last 50% of 10^4 isolation periods and averaging over 100 independent meta-ES runs with fixed normalized mutation strength $\sigma^*_{ss} = \sigma_0$. The experimental data are displayed by the error bar plot while the theoretical derivations are presented by the solid lines. The dashed and the dotted black lines display the expected minimal and maximal convergence rates according to Eq. (33). One observes that the predicted convergence rates do slightly deviate from the theoretical predictions (37). These deviations that are reduced with growing search space dimension N are discussed in Sec. 6.

The representation of the fitness dynamics provides an estimate for the expected running time needed to reach a fitness improvement of a factor of $2^{-\beta}$. Denoting the number of necessary isolation periods as T, one obtains

$$2^{-\beta} = \frac{x^{(t+T)}}{x^{(t)}} \simeq \nu^T. \tag{38}$$

Taking the logarithm and considering (13) together with Eq. (34) results in $-\beta\log(2) \simeq T\log(\nu)$. Straight forward transformations

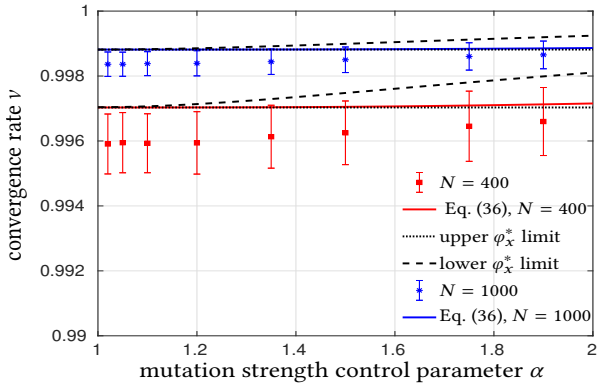

Figure 7: Comparison of the convergence rate prediction with experimental measurements from real meta-ES runs.

yield the expected number of isolation periods to reach an improvement of the factor 2^{-b} as

$$T \simeq -\beta \log(2)/\log(\nu). \tag{39}$$

By expanding $\log(\nu)$ in the Taylor series around zero ($N \to \infty$), its asymptotically exact approximation is

$$\log(\nu) \simeq -\frac{2c_{1,\lambda}^2 \alpha}{N(1+\alpha^2)}\left(1 - \frac{\alpha}{1+\alpha^2}\right). \tag{40}$$

Consequently, the insertion of (40) in Eq. (39) yields

$$T \simeq \beta \log(2) \frac{N(1+\alpha^2)^2}{2c_{1,\lambda}^2 \alpha \left(\alpha^2 - \alpha + 1\right)}. \tag{41}$$

The influence of α on the approximated expected running time T in terms of isolation periods of the $[1, 2(1,\lambda)^2]$-meta-ES is illustrated in Fig. 8 for search space dimensions $N = 400$, $N = 1000$, and $N = 2000$. All cases consider the running time until an improvement by a factor of $2^{-\beta}$ with $\beta = 2$ is realized. The experimental results obtained in 100 independent meta-ES runs are represented by the error bar plots. Similar to Fig. 7, the experimental dynamics do achieve slightly better running times than expected.

6 DISCUSSION AND OUTLOOK

This paper investigates the behavior of the $[1, 2(1,\lambda)^1]$-meta-ES in the conically constrained environment (1). To this end, the inner ES dynamics (Alg. 1) are described by the progress rate theory obtained in the context of the analysis of σSA-ES [20]. Assuming that the meta-ES is predominantly generating infeasible candidate solutions that need to be repaired (cf. (3)) with probability $P_{\text{inf}} \approx 1$ in its steady state, the analysis is extended to the upper-level strategy (Alg. 2). The evolution equations of σ, as well as σ^*, are derived and the mutation strength control behavior of the meta-ES is characterized with respect to the choice of the control parameter $\alpha > 1$. Choosing α smaller than the obtained threshold α_0 realizes a continuous mutation strength reduction. On the downside, the α_0 value turns out to be very small when considering non-recombinative inner ESs. Thus the corresponding σ adaptation is very slow and the ensuing dynamics are accompanied with extremely large fluctuations that are due to aggravated discriminability of the inner strategies [9].

For $\alpha > \alpha_0$, the σ dynamics of the meta-ES are oscillating around the point of discontinuity σ_0^*. In this case, the expected progress per isolation period of length $\gamma = 1$ can be approximated. This allows for the derivation of the dynamical long-term behavior of the meta-ES with respect to its convergence rate ν and the expected runtime T.

Taking into account the convergence rate ν in Fig. 7, one observes that the experimental dynamics slightly deviate from the theoretical predictions. It turns out that the theoretical results are too pessimistic especially in small search space dimensions. An explanation of these deviations is the negligence of the first addend of progress rate (8), i.e. the assumption of $P_{\text{fea}} \approx 0$. In fact, the probability to generate feasible candidate solutions that do not need to be repaired is approaching zero with growing N. However, in finite dimensions there is always a non-zero probability to find so-called initially feasible candidate solutions. Consequently, Eq. (8) includes an additional positive contribution to the expected progress that is omitted in the representation of Eq. (13), and in the subsequent meta-ES analysis in Sec. 5. Hence, the predictions may be regarded to represent the worst case behavior.

Yet, the contribution of the typical meta-ES fluctuations that impede the prediction quality of the deterministic evolution equations [9] to these deviations needs to be examined more closely. To this end, additional investigations of the $[1, 2(\mu/\mu_I, \lambda)^\gamma]$-meta-ES need to be conducted considering recombination for $\mu > 1$, and larger isolation periods $\gamma > 1$, respectively. Both parameters are expected to reduce the impact of the fluctuations and to stabilize the long-term dynamics. First experiments indicate that larger values of μ, or γ, result in a rise of the α_0 threshold that determines the meta-ES characteristics. However, an extension of the analysis to these cases poses novel difficulties that need to be managed first.

ACKNOWLEDGMENTS

This work was supported by the Austrian Science Fund FWF under grant P29651-N32.

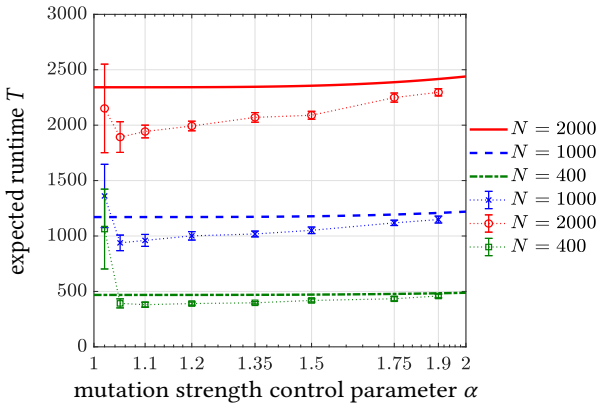

Figure 8: The expected runtime of the $[1, 2(1, 10)^1]$-meta-ES on problem (1) with $\xi = 1$. The predictions of Eq. (41) are displayed by the dotted, dashed, and solid, red lines. The data points are obtained by averaging over the experimental measurements of 100 independent meta-ES runs.

REFERENCES

[1] Dirk V. Arnold. 2012. On the behaviour of the $(1, \lambda)$-σSA-ES for a constrained linear problem. In *Parallel Problem Solving from Nature, PPSN 2012*, C. A. Coello Coello et al. (Eds.). Springer-Verlag Berlin Heidelberg, 82–91. https://doi.org/10.1007/978-3-642-32937-1_9

[2] Dirk V. Arnold. 2013. On the Behaviour of the $(1, \lambda)$-ES for a Conically Constrained Problem. In *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation (GECCO'13)*. ACM, 423–430. https://doi.org/10.1145/2463372.2463426

[3] Dirk V. Arnold. 2013. Resampling versus repair in evolution strategies applied to a constrained linear problem. *Evolutionary Computation* 21, 3 (2013), 389–411.

[4] Dirk V. Arnold and Jeremy Porter. 2015. Towards an augmented lagrangian constraint handling approach for the (1+1)-ES. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation, GECCO'15*. ACM, 249–256. https://doi.org/10.1145/2739480.2754813

[5] Asma Atamna, Anne Auger, and Nikolaus Hansen. 2016. Augmented lagrangian constraint handling for CMA-ES - case of a single linear constraint. In *International Conference on Parallel Problem Solving from Nature, PPSN*. Springer, 181–191.

[6] Asma Atamna, Anne Auger, and Nikolaus Hansen. 2017. Linearly convergent evolution strategies via augmented lagrangian constraint handling. In *Proceedings of the 14th ACM/SIGEVO Conference on Foundations of Genetic Algorithms*. ACM, 149–161.

[7] Hans-Georg Beyer and Michael Hellwig. 2012. Mutation strength control by Meta-ES on the sharp ridge. In *Genetic and Evolutionary Computation Conference, GECCO '12, Philadelphia, PA, USA, July 7-11*. 305–312. https://doi.org/10.1145/2330163.2330208

[8] Hans-Georg Beyer and Hans-Peter Schwefel. 2002. Evolution Strategies - A comprehensive introduction. *Natural Computing* 1, 1 (2002), 3–52.

[9] Michael Hellwig. 2017. *Analysis of Mutation Strength Adaptation within Evolution Strategies on the Ellipsoid Model and Methods for the Treatment of Fitness Noise*. Dissertation. Open Access Repositorium der Universität Ulm, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany. https://doi.org/10.18725/OPARU-4253

[10] Michael Hellwig and Dirk V. Arnold. 2016. Comparison of Constraint Handling Mechanisms for the $(1, \lambda)$-ES on a Simple Constrained Problem. *Evolutionary Computation* 24, 1 (2016), 1–23. https://doi.org/10.1162/EVCO_a_00139

[11] Michael Hellwig and Hans-Georg Beyer. 2015. Mutation strength control via meta evolution strategies on the ellipsoid model. *Theoretical Computer Science* 623 (2015), 160–179. https://doi.org/10.1016/j.tcs.2015.12.011

[12] Michael Herdy. 1992. Reproductive Isolation as Strategy Parameter in Hierarchally Organized Evolution Strategies.. In *PPSN* (2002-01-03), Reinhard Männer and Bernard Manderick (Eds.). Elsevier, 209–220. http://dblp.uni-trier.de/db/conf/ppsn/ppsn1992.html#Herdy92

[13] Michael Herdy. 1993. The Number of Offspring As Strategy Parameter in Hierarchically Organized Evolution Strategies. *SIGBIO Newsl.* 13, 2 (June 1993), 2–9. https://doi.org/10.1145/163428.163431

[14] Silja Meyer-Nieberg and Hans-Georg Beyer. 2012. The Dynamical Systems Approach – Progress Measures and Convergence Properties. In *Handbook of natural computing*. Springer, 741–814.

[15] Efren Mezura-Montes and Carlos A. Coello Coello. 2011. Constraint-handling in nature-inspired numerical optimization: past, present and future. *Swarm and Evolutionary Computation* 1, 4 (2011), 173–194.

[16] Andreas Ostermeier, Andreas Gawelczyk, and Nikolaus Hansen. 1994. Step-size adaptation based on non-local use of selection information. In *Parallel Problem Solving from Nature — PPSN III*, Yuval Davidor, Hans-Paul Schwefel, and Reinhard Männer (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 189–198.

[17] Ingo Rechenberg. 1973. Evolutionsstrategie – Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. (1973).

[18] Ingo Rechenberg. 1984. The Evolution Strategy. A Mathematical Model of Darwinian Evolution. Springer.

[19] Ingo Rechenberg. 1994. *Evolutionsstrategie 94*. Werkstatt Bionik und Evolutionstechnik, Vol. 1. Frommann-Holzboog, Stuttgart.

[20] Patrick Spettel and Hans-Georg Beyer. 2018. Analysis of the $(1, \lambda)$-σ-Self-Adaptation Evolution Strategy with repair by projection applied to a conically constrained problem. *Theoretical Computer Science* (2018). https://doi.org/10.1016/j.tcs.2018.10.036

[21] Patrick Spettel and Hans-Georg Beyer. 2018. Analysis of the $(\mu/\mu_I, \lambda)$-CSA-ES with Repair by Projection Applied to a Conically Constrained Problem. *arXiv preprint* (2018). https://arxiv.org/abs/1901.07871

[22] Patrick Spettel and Hans-Georg Beyer. 2018. Analysis of the $(\mu/\mu_I, \lambda)$-σ-Self-Adaptation Evolution Strategy with Repair by Projection Applied to a Conically Constrained Problem. *arXiv preprint* (2018). https://arxiv.org/1812.06300