Anne Auger and Nikolaus Hansen

A Restart CMA Evolution Strategy With Increasing Population Size

Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2005, pp. 1769-1776

ERRATUM: end of Section 2, stopping criterium "noeffectcoord": "Stop if adding 0.2-standard deviation in each coordinate does change..." should be "Stop if adding 0.2-standard deviation in any coordinate does not change..."

A Restart CMA Evolution Strategy With Increasing Population Size

Anne Auger
CoLab Computational Laboratory,
ETH Zürich, Switzerland
Anne.Auger@inf.ethz.ch

Nikolaus Hansen
CSE Lab,
ETH Zürich, Switzerland
Nikolaus.Hansen@inf.ethz.ch

Abstract- In this paper we introduce a restart-CMA-evolution strategy, where the population size is increased for each restart (IPOP). By increasing the population size the search characteristic becomes more global after each restart. The IPOP-CMA-ES is evaluated on the test suit of 25 functions designed for the special session on real-parameter optimization of CEC 2005. Its performance is compared to a local restart strategy with constant small population size. On unimodal functions the performance is similar. On multi-modal functions the local restart strategy significantly outperforms IPOP in 4 test cases whereas IPOP performs significantly better in 29 out of 60 tested cases.

1 Introduction

The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [5, 7, 4] is an evolution strategy that adapts the full covariance matrix of a normal search (mutation) distribution. Compared to many other evolutionary algorithms, an important property of the CMA-ES is its invariance against linear transformations of the search space: the CMA-ES exhibits the same performances on a given objective function $f : x \in \mathbb{R}^n \mapsto f(x) \in \mathbb{R}$, where $n \in \mathbb{N}$, and on the same function where a linear transformation is applied, *i.e.* $f_R : x \in \mathbb{R}^n \mapsto f(Rx) \in \mathbb{R}$ where R denotes a full rank linear transformation. This is true only if a corresponding transformation of the strategy (distribution) parameters is made. In practice this transformation is learned by the CMA algorithm.

The CMA-ES efficiently minimizes unimodal objective functions and is in particular superior on ill-conditioned and non-separable problems [7, 6]. In [3], Hansen and Kern show that increasing the population size improves the performance of the CMA-ES on multi-modal functions. Consequently, they suggest a CMA-ES restart strategy with successively increasing population size. Such an algorithm, referred to as IPOP-CMA-ES in the following, is introduced and investigated here for optimizing the test suit of the CEC special session on real-parameter optimization [8].

The remainder of this paper is organized as follows: Section 2 presents the main lines of the algorithm. Section 3 explains the experimental procedure. Section 4 presents and comments the experimental results. The IPOP-CMA-ES is compared to a restart strategy with constant population size.

2 The restart CMA-ES

The (μ_W, λ)-CMA-ES In this paper we use the (μ_W, λ)-CMA-ES thoroughly described in [3]. We sum up the general principle of the algorithm in the following and refer to [3] for the details.

For generation $g+1$, λ offspring are sampled independently according to a multi-variate normal distribution

$$\vec{x}_k^{(g+1)} \sim \mathcal{N}\left(\langle\vec{x}\rangle_{\mathrm{W}}^{(g)}, (\sigma^{(g)})^2 \boldsymbol{C}^{(g)}\right) \text{ for } k = 1, \ldots, \lambda$$

where $\mathcal{N}(\vec{m}, \boldsymbol{C})$ denotes a normally distributed random vector with mean $\vec{m}$ and covariance matrix $\boldsymbol{C}$. The μ best offspring are recombined into the new mean value $\langle\vec{x}\rangle_{\mathrm{W}}^{(g+1)} = \sum_{i=1}^{\mu} w_i \vec{x}_{i:\lambda}^{(g+1)}$, where the positive weights $w_i \in \mathbb{R}$ sum to one. The equations for updating the remaining parameters of the normal distribution are given in [3]: Eqs. 2 and 3 for the covariance matrix $\boldsymbol{C}$, Eqs. 4 and 5 for the step-size σ (cumulative step-size adaptation / path length control). On convex-quadratic functions, the adaptation mechanisms for σ and $\boldsymbol{C}$ allow to achieve log-linear convergence after an adaptation time which can scale between 0 and n^2. The default strategy parameters are given in [3, Eqs. 6–8]. Only $\langle\vec{x}\rangle_{\mathrm{W}}^{(0)}$ and $\sigma^{(0)}$ have to be set depending on the problem.[1]

The default population size prescribed for the (μ_W, λ)-CMA-ES grows with $\log n$ and equals to $\lambda = 10, 14, 15$ for $n = 10, 30, 50$. On multi-modal functions the optimal population size λ can be considerably greater than the default population size [3].

The restart (μ_W, λ)-CMA-ES with increasing population (IPOP-CMA-ES) For the restart strategy the (μ_W, λ)-CMA-ES is stopped whenever one stopping criterion as described below is met, and an independent restart is launched with the population size increased by a factor of 2. For all parameters of the (μ_W, λ)-CMA-ES the default values are used (see [3]) except for the population size, starting from the default value but then repeatedly increased. To our intuition, for the increasing factor, values between 1.5 and 5 could be reasonable. Preliminary empirical investigations on the Rastrigin function reveal similar performance for factors between 2 and 3. We chose 2, to conduct a larger number of restarts per run.

To decide when to restart the following (default stopping) criteria do apply.[2]

[1] A more elaborated algorithm description can be accessed via http://www.bionik.tu-berlin.de/user/niko/cmatutorial.pdf.

[2] These stopping criteria were developed before the benchmark function

- Stop if the range of the best objective function values of the last $10 + \lceil 30n/\lambda \rceil$ generations is zero (`equalfunvalhist`), or the range of these function values and all function values of the recent generation is below `Tolfun`$= 10^{-12}$.
- Stop if the standard deviation of the normal distribution is smaller than `TolX` in all coordinates and $\sigma\vec{p}_c$ (the evolution path from Eq. 2 in [3]) is smaller than `TolX` in all components. We set `TolX`$= 10^{-12}\sigma^{(0)}$.
- Stop if adding a 0.1-standard deviation vector in a principal axis direction of $\boldsymbol{C}^{(g)}$ does not change $\langle\vec{x}\rangle_{\mathrm{W}}^{(g)}$ (`noeffectaxis`)[3]
- Stop if adding 0.2-standard deviation in each coordinate does change $\langle\vec{x}\rangle_{\mathrm{W}}^{(g)}$ (`noeffectcoord`).
- Stop if the condition number of the covariance matrix exceeds 10^{14} (`conditioncov`).

The distributions of the starting points $\langle\vec{x}\rangle_{\mathrm{W}}^{(0)}$ and the initial step-sizes $\sigma^{(0)}$ are derived from the problem dependent initial search region and their setting is described in the next section, as well as the overall stopping criteria for the IPOP-CMA-ES.

3 Experimental procedure

The restart-(μ_W, λ)-CMA-ES with increasing population size, IPOP-CMA-ES, has been investigated on the 25 test functions described in [8] for dimension 10, 30 and 50. For each function a bounded subset $[A, B]^n$ of $\mathbb{R}^n$ is prescribed. For each restart the initial point $\langle\vec{x}\rangle_{\mathrm{W}}^{(0)}$ is sampled uniformly within this subset and the initial step-size $\sigma^{(0)}$ is equal to $(B - A)/2$. The stopping criteria before to restart are described in the last section. The overall stopping criteria for the algorithm prescribed in [8] are: stop after $n \times 10^4$ function evaluations or stop if the objective function error value is below 10^{-8}.

The boundary handling is done according to the standard implementation of CMA-ES and consists in penalizing the individuals in the infeasible region.[4] For each test function, 25 runs are performed. All performance criteria are evaluated based on the same runs. In particular, the times when to measure the objective function error value (namely at 10^3, 10^4, 10^5 function evaluations) were not used as input parameter to the algorithm (*e.g.*, to set the maximum number of function evaluations to adjust an annealing rate).

Test functions The complete definition of the test suit is available in [8]. The definition of functions 1–12 is based on classical benchmark functions, that we will refer in the sequel also by their name. Functions 1–5 are unimodal and functions 6–12 are multi-modal. Functions 13–25 result from the composition of several functions. To prevent exploitation of symmetry of the search space and of the typical zero value associated with the global optimum, the local optimum is shifted to a value different from zero and the function values of the global optima are non-zero.

Success performances Evaluating and comparing the performances of different algorithms on multi-modal problems implies to take into account that some algorithms may have a small probability of success but converge fast whereas others may have a larger probability of success but be slower.

To evaluate the performances of an algorithm **A** the following success performance criterion has been defined in [8]

$$\widehat{SP1} = \frac{\widehat{\mathbb{E}(T_A^s)}}{\widehat{p_{\mathrm{s}}}} \ , \qquad (1)$$

where $\widehat{\mathbb{E}(T_A^s)}$ is an estimator of the expected number of objective function evaluations during a successful run of **A**,

$$\widehat{\mathbb{E}(T_A^s)} = \frac{\text{Nbr. evaluations in all successful runs}}{\text{Nbr. successful runs}}$$

and $\widehat{p_{\mathrm{s}}}$ an estimator of the probability of success[5] of **A**

$$\widehat{p_{\mathrm{s}}} = \frac{\text{Nbr. successful runs}}{\text{Nbr. runs}} \ .$$

In [1] we have shown that this performance criterion is a particular case of a more general criterion. More precisely, consider T the random variable measuring the overall running time (or number of function evaluations) before a success criterion is met by independent restarts of **A**, then the expectation of T is equal to

$$\mathbb{E}(T) = \Big(\frac{1 - p_{\mathrm{s}}}{p_{\mathrm{s}}}\Big)\mathbb{E}(T_A^{us}) + \mathbb{E}(T_A^s) \qquad (2)$$

where p_{s} is the probability of success of **A**, and $\mathbb{E}(T_A^{us})$ and $\mathbb{E}(T_A^s)$ are the expected number of function evaluations for unsuccessful and successful runs, respectively (see [1, Eq. 2]). For $\mathbb{E}(T_A^{us}) = \mathbb{E}(T_A^s)$, Eq. 2 simplifies to $\mathbb{E}(T) = \mathbb{E}(T_A^s)/p_{\mathrm{s}}$. Therefore $\widehat{SP1}$ estimates the expected running time T under this particular assumption.

Taking now into account that the algorithm investigated in this paper is a restart strategy and that the maximum number of function evaluations allowed is $FE_{\max} = n \times 10^4$, we can derive another success performance criterion (see [1] for the details). Indeed for a restart strategy it is reasonable to assume that a run is unsuccessful only because it reaches $FE_{\max}$ evaluations. Therefore Eq. 2 simplifies into

$$\mathrm{SP2} = \Big(\frac{1 - p_{\mathrm{s}}}{p_{\mathrm{s}}}\Big)FE_{\max} + \mathbb{E}(T_A^s) \qquad (3)$$

for which the estimator we will use is

$$\widehat{\mathrm{SP2}} = \Big(\frac{1 - \widehat{p_{\mathrm{s}}}}{\widehat{p_{\mathrm{s}}}}\Big)FE_{\max} + \widehat{\mathbb{E}(T_A^s)} \ . \qquad (4)$$

suit used in this paper was assembled.

[3] More formally, stop if $\langle\vec{x}\rangle_{\mathrm{W}}^{(g)}$ equals to $\langle\vec{x}\rangle_{\mathrm{W}}^{(g)} + 0.1\,\sigma^{(g)}\sqrt{\lambda_i}\vec{u}_i$, where $i = (g \mod n) + 1$, and λ_i and $\vec{u}_i$ are respectively the ith eigenvalue and eigenvector of $\boldsymbol{C}^{(g)}$, with $\|\vec{u}_i\| = 1$.

[4] For details refer to the used MATLAB code, cmaes.m, Version 2.35, see http://www.bionik.tu-berlin.de/user/niko/formersoftwareversions.html

[5] If $\widehat{p_{\mathrm{s}}}$ is zero $\widehat{SP1}$ is not computed.

Table 1: Measured CPU-seconds, according to [8], using MATLAB 7.0.1, Red Hat Linux 2.4, 1GByte RAM, Pentium 4 3GHz processor. Time T2 is the CPU-time for running the IPOP-CMA-ES until 2×10^5 function evaluations on function 3. For $n = 30$ the IPOP-CMA-ES needs on average 0.12 CPU-milliseconds per function evaluation. The strategy internal time consumption scales with $\mathcal{O}(n^2)$. The large values for T1 reflect the large number of objective functions calls, while for T2 a complete, eventually large, population is evaluated (serially) within a single function call. Running the same code using MATLAB 6.5.0, Windows XP, 512MByte, 2.4GHz, increases T0 by more than a factor of ten, whereas T1 and T2 increase by less than a factor of two

	T0	T1	T2
$n = 10$	0.4s	32s	17s
$n = 30$	0.4s	41s	24s
$n = 50$	0.4s	49s	56s

We also derived the variance of SP2

$$\mathrm{var}(\mathrm{SP2}) = \left(\frac{1 - p_\mathrm{s}}{p_\mathrm{s}^2}\right)(FE_{\max})^2 + \mathrm{var}(T_A^s) \ , \quad (5)$$

and consider here its following estimator

$$\widehat{\mathrm{var}(\mathrm{SP2})} = \left(\frac{1 - \widehat{p}_\mathrm{s}}{\widehat{p}_\mathrm{s}^{\,2}}\right)(FE_{\max})^2 + \widehat{\mathrm{var}(T_A^s)}. \quad (6)$$

where $\widehat{\mathrm{var}(T_A^s)}$ is an estimator for the variance of the number of objective function evaluations during a successful run of **A**.

4 Results

Figure 1 presents the convergence graphs associated to the median value at each generation. Steps that repeatingly occur in the graphs, most prominent for function 12, indicate that the larger population sizes achieve better results. The observed maximal final population size is $\lambda = 640, 448, 480$, which means $2^6, 2^5, 2^5$ times $\lambda_\mathrm{start} = 10, 14, 15$, for $n = 10, 30, 50$, respectively.

According to the requirements, Table 1 reports CPU-time measurements, Table 2 gives the number of function evaluations to reach the success criterion (if successful), the success rate, and the success performances as defined in the previous section. The objective function error values after 10^3, 10^4, 10^5 and $n \times 10^4$ function evaluations are presented in Table 5, 6 and 7.

On the ill-conditioned and non-separable function 3, the performance is exceptionally good whereas many evolutionary algorithms fail to locate the optimum of this convex-quadratic function. Invariance of the CMA-ES against orthogonal transformations leads to similar performance on functions 9 and 10 which are respectively the Rastrigin function and the rotated Rastrigin function. If the maximum number of function evaluations is increased on Rastrigin function the success performance values improve (*e.g.* on function 10 for $n = 30$, by a factor of almost three, given $FE_{\max} = 3 \times n \times 10^4$). On function 8, the Ackley function composed with a linear transformation, the success rate is zero. This result seems to contradict previous results on the Ackley function [3], but it can be explained by the different scaling that is applied here: outside the usual bounds ($[-32, 32]^n$) the Ackley function is almost flat and the composition with a linear transformation with a condition number of 100 brings this flat region into the search space. Therefore the function looks like a needle in the haystack problem. In case of the noisy Schwefel Problem (function 4) we observe premature convergence for dimension 30 and 50 due to a too fast decrease of the step-size.

Comparison with a pure local restart (LR) strategy To evaluate the impact of the increase of the population after each restart, we compare the results presented here with a pure restart strategy, *i.e.* where the population is kept fixed after each restart of the (μ_W, λ)-CMA-ES, presented in [1] and referred to as LR-CMA-ES in the following. Moreover, to stress the local search characteristics, in LR-CMA-ES the initial step-size equals to $0.5 \times 10^{-2}(B - A)$. Even though such a small initial step-size suggests that the strategy operates as a pure local search, the step-size adaptation mechanism allows yet to search more globally than a pure local search method. In fact, in some cases the step-size adaptation increases the step-size by a factor close to 10^2. The detailed results for this latter approach are presented in [1].

We test the statistical significance of the differences observed on the data. Table 3 reports the result of the Kolmogorov-Smirnov test for comparing the distributions of the objective function values from Tables 5, 6, and 7 compared to [1]. Table 4 reports results for the two-sided Wilcoxon rank sum test for equal medians.

From the convergence plots for $n = 30$ in Fig. 1 we observe remarkable differences in the convergence graphs compared to [1] on functions 4, 9–12, 17, and 24. All these differences are highly significant and favor IPOP-CMA-ES.

The effect of the different initial step-size is significantly visible after 10^3 function evaluations in all functions at least for some dimension (Table 4), except for the Ackley function 8. Unexpectedly, in most cases the larger initial step-size of IPOP-CMA-ES leads to a better function value distribution even after 10^3 function evaluations.

According to Table 4, disregarding the initial differences after a small number of function evaluations, IPOP-CMA-ES significantly outperforms LR-CMA-ES on the functions 4, 9–12, 14, 16, 17, 22, and 24. Contrariwise, on functions 13 ($n = 10$), 21 ($n = 10; 50$), and 23 ($n = 50$) LR-CMA-ES significantly outperforms IPOP-CMA-ES. For function 21 the diversity of the solutions found in the local-restart approach suggests that the small population size is favorable, because it allows to converge faster to a local minimum and therefore allows to do more restarts and visit more different local optima before reaching $n \times 10^4$ function evaluations.

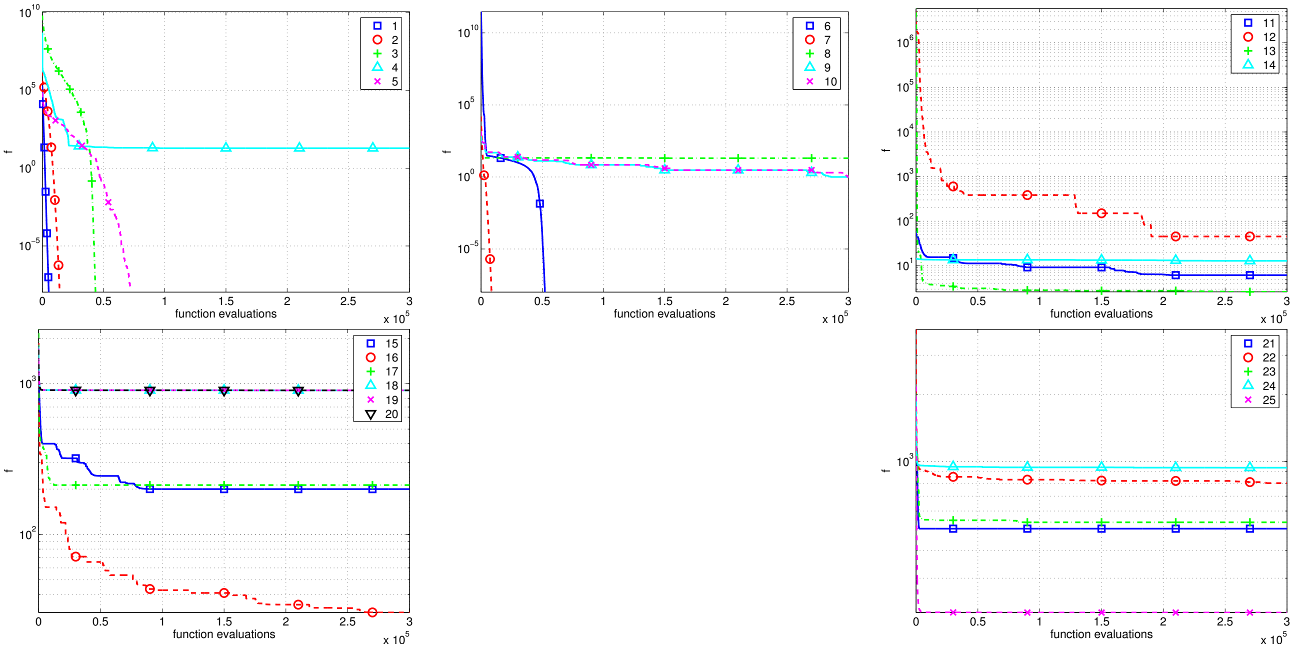

Figure 1: Best objective function error value (log scale) versus number of function evaluations for the 25 problems in dimension 30. For each run the best individual found until the current generation is considered and shown is the median value of 25 runs at each generation. The respective problem number is given in the legend.

Table 2: For the 25 problems in dimension $n = 10$, $n = 30$, and $n = 50$, number of function evaluations (min, 7^{th}, median, 19^{th}, maximum, mean and standard deviation) needed to reach the neighborhood of the global optimum with the objective function error value (accuracy) as given in the Tol column. A run is successful if it reaches Tol before $n \times 10^4$ function evaluations. For functions 13 to 25 none of the runs reach the given accuracy. Success rate (p_s) and success performance SP1 and SP2 as defined in Eq. 1 and Eq. 4. Standard deviation for SP2 as defined in Eq. 6.

	Prob.	Tol	min	7^{th}	median	19^{th}	max	mean	std	p_s	SP1	SP2	std(SP2)
$n = 10$	1	1e−6	1.44e+3	1.58e+3	1.63e+3	1.65e+3	1.71e+3	1.61e+3	6.14e+1	1.00	1.61e+3	1.61e+3	6.14e+1
	2	1e−6	2.20e+3	2.33e+3	2.35e+3	2.44e+3	2.60e+3	2.38e+3	1.06e+2	1.00	2.38e+3	2.38e+3	1.06e+2
	3	1e−6	5.84e+3	6.31e+3	6.51e+3	6.71e+3	7.20e+3	6.50e+3	2.92e+2	1.00	6.50e+3	6.50e+3	2.92e+2
	4	1e−6	2.52e+3	2.82e+3	2.88e+3	3.03e+3	3.22e+3	2.90e+3	1.68e+2	1.00	2.90e+3	2.90e+3	1.68e+2
	5	1e−6	5.36e+3	5.63e+3	5.83e+3	5.97e+3	6.72e+3	5.85e+3	2.89e+2	1.00	5.85e+3	5.85e+3	2.89e+2
	6	1e−2	5.67e+3	7.08e+3	8.55e+3	1.37e+4	2.26e+4	1.08e+4	5.00e+3	1.00	1.08e+4	1.08e+4	5.00e+3
	7	1e−2	1.49e+3	1.80e+3	5.83e+3	6.24e+3	1.33e+4	4.67e+3	2.83e+3	1.00	4.67e+3	4.67e+3	2.83e+3
	8	1e−2	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	9	1e−2	2.33e+4	4.62e+4	7.85e+4	8.31e+4	-	5.75e+4	2.11e+4	0.76	7.57e+4	8.91e+4	6.78e+4
	10	1e−2	2.68e+4	4.74e+4	5.15e+4	7.84e+4	-	5.98e+4	1.81e+4	0.92	6.50e+4	6.85e+4	3.57e+4
	11	1e−2	3.05e+4	-	-	-	-	6.31e+4	2.56e+4	0.24	2.63e+5	3.80e+5	3.64e+5
	12	1e−2	2.37e+3	3.92e+3	3.10e+4	7.18e+4	-	2.88e+4	2.78e+4	0.88	3.27e+4	4.24e+4	4.82e+4

	Prob.	Tol	min	7^{th}	median	19^{th}	max	mean	std	p_s	SP1	SP2	std(SP2)
$n = 30$	1	1e−6	4.15e+3	4.42e+3	4.50e+3	4.58e+3	4.72e+3	4.50e+3	1.33e+2	1.00	4.50e+3	4.50e+3	1.33e+2
	2	1e−6	1.20e+4	1.28e+4	1.31e+4	1.32e+4	1.36e+4	1.30e+4	3.52e+2	1.00	1.30e+4	1.30e+4	3.52e+2
	3	1e−6	4.15e+4	4.23e+4	4.27e+4	4.30e+4	4.42e+4	4.27e+4	6.06e+2	1.00	4.27e+4	4.27e+4	6.06e+2
	4	1e−6	1.94e+4	2.71e+4	-	-	-	2.36e+4	4.79e+3	0.40	5.90e+4	4.74e+5	5.81e+5
	5	1e−6	1.91e+4	5.74e+4	6.83e+4	7.62e+4	1.03e+5	6.59e+4	1.85e+4	1.00	6.59e+4	6.59e+4	1.85e+4
	6	1e−2	3.76e+4	4.47e+4	4.83e+4	5.82e+4	1.55e+5	6.00e+4	2.81e+4	1.00	6.00e+4	6.00e+4	2.81e+4
	7	1e−2	4.12e+3	4.82e+3	4.97e+3	5.23e+3	1.99e+4	6.11e+3	4.02e+3	1.00	6.11e+3	6.11e+3	4.02e+3
	8	1e−2	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	9	1e−2	2.75e+5	2.91e+5	-	-	-	2.85e+5	6.87e+3	0.36	7.90e+5	8.18e+5	6.67e+5
	10	1e−2	2.87e+5	-	-	-	-	2.90e+5	2.44e+3	0.12	2.42e+6	2.49e+6	2.35e+6
	11	1e−2	1.99e+5	-	-	-	-	1.99e+5	0.00e+0	0.04	4.98e+6	7.40e+6	7.35e+6
	12	1e−2	1.67e+4	1.38e+5	-	-	-	7.19e+4	7.54e+4	0.32	2.25e+5	7.09e+5	7.77e+5

	Prob.	Tol	min	7^{th}	median	19^{th}	max	mean	std	p_s	SP1	SP2	std(SP2)
$n = 50$	1	1e−6	6.54e+3	6.80e+3	6.89e+3	6.96e+3	7.13e+3	6.88e+3	1.42e+2	1.00	6.88e+3	6.88e+3	1.42e+2
	2	1e−6	3.00e+4	3.09e+4	3.11e+4	3.15e+4	3.29e+4	3.13e+4	6.55e+2	1.00	3.13e+4	3.13e+4	6.55e+2
	3	1e−6	1.15e+5	1.16e+5	1.17e+5	1.17e+5	1.18e+5	1.17e+5	6.77e+2	1.00	1.17e+5	1.17e+5	6.77e+2
	4	1e−6	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	5	1e−6	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	6	1e−2	1.15e+5	1.23e+5	1.36e+5	1.50e+5	3.59e+5	1.58e+5	6.68e+4	1.00	1.58e+5	1.58e+5	6.68e+4
	7	1e−2	7.32e+3	7.67e+3	8.00e+3	8.22e+3	1.01e+4	8.03e+3	5.56e+2	1.00	8.03e+3	8.03e+3	5.56e+2
	8	1e−2	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	9	1e−2	4.06e+5	4.64e+5	-	-	-	4.35e+5	2.22e+4	0.28	1.55e+6	1.72e+6	1.52e+6
	10	1e−2	4.32e+5	-	-	-	-	4.52e+5	2.00e+4	0.12	3.76e+6	4.12e+6	3.91e+6
	11	1e−2	-	-	-	-	-	-	-	0.00	-	0.00e+0	-
	12	1e−2	-	-	-	-	-	-	-	0.00	-	0.00e+0	-

Table 3: Results of the Kolmogorov-Smirnov test between IPOP-CMA-ES and LR-CMA-ES for dimensions 10, 30, and 50. For each test problem (column) and for each number of function evaluations (FES) being equal to 10^3, 10^4, 10^5 and $n \times 10^5$ (row), the hypothesis is tested whether the measured objective function value distributions are identical for the IPOP-CMA-ES and the LR-CMA-ES. Given is the negative base ten logarithm of the p-value. That is, for the upper left entry we have a significance level of $p = 10^{-5.8}$. The 5%-significance level using the (most conservative) Bonferroni correction for multiple testing evaluates to $-\log_{10}(0.05/375) \approx 3.9$; that is, entries larger than 4 can be well regarded as statistically significant

$n = 10$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	5.8	2.0	2.0	11.5	5.8	0.9	8.8	0.2	8.8	9.7	3.4	1.6	0.1	5.8	5.8	2.0	8.8	0.9	0.6	0.6	1.2	2.0	2.0	7.3	9.7
1e4	0.1	0.4	0.4	9.7	0.1	0.2	0.6	1.6	11.5	11.5	5.8	5.2	0.9	8.0	2.0	3.9	6.5	2.9	1.2	0.9	0.6	2.4	0.9	5.2	1.2
1e5	0.1	0.4	0.4	6.5	0.1	0.2	2.0	0.0	11.5	11.5	7.3	2.9	4.5	10.6	1.2	8.0	9.7	3.4	3.9	3.4	6.5	6.5	2.9	5.2	1.6

$n = 30$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	9.7	11.5	10.6	11.5	10.6	10.6	11.5	0.9	10.6	11.5	3.9	9.7	0.9	8.8	2.4	2.9	11.5	11.5	10.6	11.5	7.3	8.8	10.6	11.5	11.5
1e4	0.0	2.9	7.3	9.7	0.4	0.9	1.6	0.4	11.5	11.5	7.3	3.4	0.4	8.8	1.2	0.2	10.6	8.8	9.7	10.6	0.1	5.8	0.2	8.0	4.5
1e5	0.0	2.0	0.1	8.0	0.4	0.1	1.2	3.4	11.5	11.5	8.0	6.5	0.2	7.3	2.9	5.2	7.3	9.7	11.5	9.7	0.0	10.6	0.1	8.0	2.4
3e5	0.0	2.0	0.1	8.0	0.4	0.4	1.2	0.9	11.5	11.5	9.7	8.0	1.6	8.0	2.4	8.0	7.3	7.3	9.7	7.3	0.1	9.7	1.2	8.0	2.4

$n = 50$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	11.5	11.5	11.5	11.5	5.8	10.6	11.5	0.4	11.5	11.5	5.2	11.5	4.5	8.8	2.0	8.8	11.5	11.5	10.6	11.5	11.5	9.7	11.5	11.5	11.5
1e4	0.9	8.8	11.5	11.5	1.2	0.6	5.8	0.1	11.5	11.5	10.6	3.4	0.4	8.0	0.9	2.0	11.5	4.5	8.0	6.5	11.5	9.7	3.4	3.4	11.5
1e5	0.9	0.2	3.4	5.8	1.2	0.2	0.6	0.1	11.5	11.5	11.5	8.0	0.9	9.7	5.8	5.2	10.6	5.8	5.8	8.0	11.5	11.5	3.4	2.4	1.2
5e5	0.9	0.2	0.1	2.0	0.6	0.4	0.6	1.2	11.5	11.5	11.5	7.3	0.1	7.3	5.2	9.7	10.6	3.4	2.4	2.9	11.5	11.5	10.6	2.4	1.6

Table 4: Results of the two-sided Wilcoxon rank sum test for median function values between IPOP-CMA-ES and LR-CMA-ES for dimensions 10 , 30, and 50. For each test problem (column) and for each number of function evaluations (FES) being equal to 10^3, 10^4, 10^5, and $n \times 10^5$ (row), the hypothesis is tested whether the median for IPOP-CMA-ES and LR-CMA-ES are the same. Given is the negative base ten logarithm of the p-value. That is, for the upper left entry we have a significance level of $p = 10^{-6.7} \approx 2 \times 10^{-7}$. The plus indicates that the median was lower (better) for LR-CMA-ES, the star denotes cases where this is statistically significant. The 5%-significance level using the (most conservative) Bonferroni correction for multiple testing evaluates to $-\log_{10}(0.05/375) \approx 3.9$; that is, entries larger than 4 can be well regarded as statistically significant.

$n = 10$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	6.7	3.2	$_+$2.1	8.8	6.2	0.8	7.1	$_+$0.0	7.6	8.0	3.7	0.6	$_+$0.0	6.3	4.6	1.9	8.3	0.7	0.6	1.4	0.8	2.8	$_+$0.8	4.7	7.6
1e4	$_+$0.0	0.6	0.4	8.5	0.0	$_+$0.4	1.4	$_+$1.7	8.9	8.9	5.3	4.0	$_+$0.6	6.7	2.5	4.3	6.5	2.6	1.5	1.0	0.0	1.8	$_+$0.6	6.0	1.1
1e5	$_+$0.0	0.6	0.4	6.9	0.0	0.4	2.3	0.5	8.9	8.8	6.6	3.4	$^*_+$4.4	8.6	0.3	7.0	8.1	4.1	4.6	4.7	$^*_+$5.4	5.7	1.6	5.6	2.0

$n = 30$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	8.5	$^*_+$8.8	$^*_+$8.7	8.8	8.5	8.8	8.8	$_+$1.3	8.3	8.8	$^*_+$4.3	$^*_+$8.0	1.0	6.7	2.8	3.8	8.8	8.8	8.8	8.8	8.0	7.7	8.1	8.8	8.8
1e4	0.0	3.2	$^*_+$7.5	8.5	0.3	0.4	2.1	$_+$0.3	8.8	8.8	6.7	3.0	$_+$0.3	6.7	2.1	0.3	8.4	8.2	7.8	8.6	0.9	5.4	$_+$0.0	4.5	3.9
1e5	0.0	2.4	$_+$0.5	7.4	0.2	0.2	2.3	$_+$3.4	8.9	8.9	7.9	5.8	$_+$0.2	6.1	2.6	3.3	7.1	6.4	8.8	6.4	0.5	8.8	0.4	4.5	2.5
3e5	0.0	2.4	$_+$0.5	7.4	0.2	0.6	2.3	2.6	8.8	8.9	8.3	6.4	$_+$1.5	7.5	0.2	5.5	7.2	3.5	6.4	3.5	1.4	8.6	1.2	4.5	2.5

$n = 50$

FES	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	8.8	$^*_+$8.8	$^*_+$8.8	8.8	6.7	8.6	8.8	$_+$0.4	8.8	8.8	$^*_+$4.1	$^*_+$8.8	$^*_+$4.1	8.1	$_+$0.2	8.4	8.8	8.8	8.7	8.8	8.8	8.6	8.8	8.8	8.8
1e4	0.7	$^*_+$8.2	$^*_+$8.8	8.8	$_+$1.0	0.4	5.0	$_+$0.4	8.8	8.8	7.9	1.0	$_+$0.3	7.2	$_+$0.8	1.0	8.8	3.7	6.3	4.6	$^*_+$9.6	8.5	$_+$0.3	0.4	8.8
1e5	0.7	0.3	$_+$3.1	4.9	1.1	0.0	0.5	$_+$0.1	8.8	8.9	8.8	5.9	0.5	7.4	6.1	3.9	8.6	2.5	2.2	4.4	$^*_+$10.0	8.8	$_+$0.3	$_+$0.0	2.0
5e5	0.7	0.3	$_+$0.3	2.6	$_+$0.5	$_+$0.3	0.5	2.0	8.9	8.8	8.8	6.3	$_+$0.1	7.4	2.0	7.1	8.6	$_+$0.3	0.1	0.1	$^*_+$10.0	8.8	$^*_+$7.6	$_+$0.0	2.1

5 Summary and conclusions

In this paper we have investigated the IPOP-CMA-ES, a restart CMA-ES with a population size successively increased by a factor of two. The algorithm has been tested on 25 benchmark functions and compared with a pure restart CMA-ES with constant population size where the initial step-size is small. As expected, increasing the population size is usually the better choice, and IPOP-CMA-ES significantly outperforms the pure restart approach on many multi-modal problems. On unimodal non-noisy functions no relevant differences are observed and IPOP-CMA-ES performs as well as the standard (μ_W, λ)-CMA-ES.

To draw a final conclusion we emphasize two important aspects that have to be taken into account when judging the performance of any search algorithm. First, as for the standard (μ_W, λ)-CMA-ES, the IPOP-CMA-ES is **quasi parameter free**.[6] In the presented experiments only the ini-

[6] Remark that the number of parameters in the description of an algorithm is somewhat arbitrary: the more general the description, the more parameters appear. Therefore, the existence or absence of parameters in the algorithm description cannot have influence on the assessment of the number of parameters that need to be (empirically or heuristically) determined each time the algorithm is applied. For the CMA-ES, strategy parameters have been chosen in advance, based on principle algorithmic considerations and in-depth empirical investigations on a few simple test functions. To our experience the strategy parameters (*e.g.* a learning rate, a time horizon, or a damping factor) mostly depend on algorithmic internal considerations and on the search space dimension, and to a much lesser extend on the specific objective function the algorithm is applied to. Nevertheless, it is possible to improve the performance by tuning strategy parameters and stopping criteria on most (all?) functions.

Table 5: Best objective function error value reached after 10^3, 10^4 and 10^5 function evaluations (FES) respectively (rows) on the 25 test problems (columns) in dimension $n = 10$. Given are minimum, 7^{th}, median, 19^{th}, and maximum value, as well as mean value and standard deviation from 25 runs. A run is stopped whenever the objective function error value drops below 10^{-8} and its final value is used for all larger FES.

FES	Prob.	1	2	3	4	5	6	7	8	9	10	11	12
1e3	min	1.88e−3	6.22e+0	1.07e+6	8.86e+1	6.71e+0	1.12e+1	6.51e−1	2.05e+1	5.02e+0	6.11e+0	2.33e+0	3.26e+1
	7^{th}	7.56e−3	2.40e+1	5.73e+6	5.81e+2	1.25e+1	1.27e+2	7.28e−1	2.07e+1	1.43e+1	3.63e+1	6.54e+0	2.82e+2
	med.	1.34e−2	4.35e+1	1.14e+7	1.58e+3	2.16e+1	1.91e+3	9.06e−1	2.08e+1	2.75e+1	4.56e+1	8.21e+0	2.28e+3
	19^{th}	2.46e−2	8.36e+1	2.39e+7	4.28e+3	3.76e+1	3.15e+3	9.74e−1	2.08e+1	4.70e+1	5.14e+1	9.35e+0	8.37e+4
	max	4.68e−2	1.70e+2	5.67e+7	1.75e+4	6.71e+1	9.89e+5	1.11e+0	2.09e+1	5.91e+1	6.37e+1	1.13e+1	1.29e+5
	mean	1.70e−2	5.83e+1	1.68e+7	3.00e+3	2.81e+1	4.30e+4	8.69e−1	2.08e+1	3.07e+1	4.17e+1	7.97e+0	3.59e+4
	std	1.20e−2	4.74e+1	1.59e+7	3.83e+3	1.85e+1	1.97e+5	1.38e−1	9.93e−2	1.72e+1	1.57e+1	2.22e+0	4.95e+4
1e4	min	1.84e−9	2.21e−9	2.21e−9	1.71e−9	2.46e−9	3.29e−9	9.31e−10	2.03e+1	1.99e+0	2.98e+0	3.38e−1	3.42e−9
	7^{th}	3.75e−9	3.27e−9	4.61e−9	3.85e−9	5.02e−9	4.49e−9	2.81e−9	2.05e+1	4.97e+0	4.97e+0	2.79e+0	6.66e−9
	med.	5.65e−9	4.53e−9	5.51e−9	4.78e−9	6.33e−9	7.37e−9	5.46e−9	2.06e+1	5.97e+0	6.96e+0	4.65e+0	1.88e+1
	19^{th}	6.42e−9	5.71e−9	6.58e−9	6.46e−9	8.60e−9	3.99e+0	7.77e−9	2.06e+1	7.96e+0	7.96e+0	5.72e+0	1.35e+3
	max	9.34e−9	7.67e−9	9.66e−9	7.80e−9	9.84e−9	2.63e+2	1.48e−2	2.07e+1	1.09e+1	1.59e+1	7.87e+0	6.34e+4
	mean	5.20e−9	4.70e−9	5.60e−9	5.02e−9	6.58e−9	1.17e+1	2.27e−3	2.05e+1	6.21e+0	7.16e+0	4.42e+0	2.98e+3
	std	1.94e−9	1.56e−9	1.93e−9	1.71e−9	2.17e−9	5.24e+1	4.32e−3	8.62e−2	2.10e+0	3.12e+0	2.10e+0	1.26e+4
1e5	min	1.84e−9	2.21e−9	2.21e−9	1.71e−9	2.46e−9	1.44e−9	6.22e−10	2.00e+1	1.52e−10	1.50e−10	5.27e−10	1.08e−9
	7^{th}	3.75e−9	3.27e−9	4.61e−9	3.85e−9	5.02e−9	3.81e−9	1.65e−9	2.00e+1	3.46e−10	3.34e−10	3.48e−2	2.81e−9
	med.	5.65e−9	4.53e−9	5.51e−9	4.78e−9	6.33e−9	4.69e−9	2.84e−9	2.00e+1	6.14e−10	5.64e−10	6.34e−1	3.89e−9
	19^{th}	6.42e−9	5.71e−9	6.58e−9	6.46e−9	8.60e−9	5.67e−9	5.46e−9	2.00e+1	3.50e−9	1.08e−9	1.64e+0	5.94e−9
	max	9.34e−9	7.67e−9	9.66e−9	7.80e−9	9.84e−9	8.13e−9	7.77e−9	2.00e+1	9.95e−1	9.95e−1	3.19e+0	7.12e+2
	mean	5.20e−9	4.70e−9	5.60e−9	5.02e−9	6.58e−9	4.87e−9	3.31e−9	2.00e+1	2.39e−1	7.96e−2	9.34e−1	2.93e+1
	std	1.94e−9	1.56e−9	1.93e−9	1.71e−9	2.17e−9	1.66e−9	2.02e−9	3.89e−3	4.34e−1	2.75e−1	9.00e−1	1.42e+2

FES	Prob.	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	min	2.26e+0	4.07e+0	2.21e+2	1.12e+2	1.43e+2	3.04e+2	3.04e+2	3.01e+2	5.00e+2	7.86e+2	5.59e+2	2.00e+2	3.86e+2
	7^{th}	3.35e+0	4.19e+0	4.03e+2	1.73e+2	1.90e+2	7.33e+2	8.01e+2	7.85e+2	5.01e+2	7.97e+2	1.09e+3	2.15e+2	3.94e+2
	med.	3.75e+0	4.29e+0	4.16e+2	2.07e+2	2.37e+2	8.09e+2	9.65e+2	8.03e+2	8.87e+2	8.02e+2	1.26e+3	5.10e+2	3.99e+2
	19^{th}	4.51e+0	4.37e+0	4.41e+2	2.38e+2	2.69e+2	1.02e+3	1.02e+3	9.83e+2	1.25e+3	8.57e+2	1.27e+3	9.76e+2	4.03e+2
	max	5.28e+0	4.52e+0	5.89e+2	4.05e+2	1.01e+3	1.06e+3	1.06e+3	1.07e+3	1.27e+3	9.49e+2	1.28e+3	9.94e+2	5.05e+2
	mean	3.84e+0	4.28e+0	4.12e+2	2.19e+2	2.74e+2	8.43e+2	8.86e+2	8.50e+2	9.05e+2	8.29e+2	1.08e+3	5.89e+2	4.06e+2
	std	8.05e−1	1.33e−1	7.98e+1	7.15e+1	1.73e+2	2.14e+2	1.97e+2	1.62e+2	3.40e+2	4.81e+1	2.94e+2	3.31e+2	2.99e+1
1e4	min	4.94e−1	2.95e+0	1.69e+2	7.95e+1	1.17e+2	3.00e+2	3.00e+2	3.00e+2	5.00e+2	7.41e+2	5.59e+2	2.00e+2	3.74e+2
	7^{th}	7.50e−1	3.75e+0	2.00e+2	1.02e+2	1.28e+2	3.00e+2	3.57e+2	5.47e+2	5.00e+2	7.58e+2	5.59e+2	2.00e+2	3.80e+2
	med.	9.77e−1	3.99e+0	3.00e+2	1.09e+2	1.38e+2	5.28e+2	8.00e+2	7.22e+2	5.00e+2	7.66e+2	1.23e+3	2.00e+2	3.82e+2
	19^{th}	1.11e+0	4.08e+0	4.00e+2	1.16e+2	1.70e+2	8.00e+2	8.08e+2	8.00e+2	8.56e+2	8.00e+2	1.27e+3	2.00e+2	3.84e+2
	max	1.84e+0	4.51e+0	5.00e+2	1.69e+2	9.99e+2	1.03e+3	1.03e+3	9.90e+2	1.25e+3	8.98e+2	1.28e+3	9.72e+2	3.89e+2
	mean	9.71e−1	3.91e+0	2.99e+2	1.11e+2	1.95e+2	6.02e+2	6.75e+2	6.85e+2	7.05e+2	7.81e+2	9.49e+2	3.04e+2	3.82e+2
	std	3.18e−1	3.05e−1	1.02e+2	1.66e+1	1.85e+2	2.76e+2	2.52e+2	2.26e+2	3.00e+2	3.95e+1	3.52e+2	2.58e+2	3.42e+0
1e5	min	4.07e−1	2.08e+0	1.69e+2	7.92e+1	9.48e+1	3.00e+2	3.00e+2	3.00e+2	5.00e+2	7.16e+2	5.59e+2	2.00e+2	3.70e+2
	7^{th}	6.44e−1	2.75e+0	2.00e+2	9.05e+1	1.13e+2	3.00e+2	3.00e+2	3.00e+2	5.00e+2	7.23e+2	5.59e+2	2.00e+2	3.71e+2
	med.	6.82e−1	3.00e+0	2.00e+2	9.13e+1	1.22e+2	3.00e+2	3.00e+2	3.00e+2	5.00e+2	7.30e+2	5.59e+2	2.00e+2	3.74e+2
	19^{th}	7.61e−1	3.28e+0	2.00e+2	9.25e+1	1.30e+2	3.00e+2	3.00e+2	3.00e+2	5.00e+2	7.32e+2	5.59e+2	2.00e+2	3.76e+2
	max	1.05e+0	3.51e+0	4.00e+2	9.68e+1	2.04e+2	7.18e+2	7.69e+2	3.00e+2	5.00e+2	7.44e+2	5.59e+2	2.00e+2	3.79e+2
	mean	6.96e−1	3.01e+0	2.28e+2	9.13e+1	1.23e+2	3.32e+2	3.26e+2	3.00e+2	5.00e+2	7.29e+2	5.59e+2	2.00e+2	3.74e+2
	std	1.50e−1	3.49e−1	6.80e+1	3.49e+0	2.09e+1	1.12e+2	9.93e+1	0.00e+0	3.48e−13	6.86e+0	3.24e−11	2.29e−6	3.22e+0

tial search region is defined problem dependent; no (further) adjustment of parameters is conducted. Second, the IPOP-CMA-ES has several **invariance properties** [2], like invariance against order preserving transformations of the objective function value and invariance against linear transformations of the search space [7]. Invariance properties are highly desirable, because they imply uniform performance on classes of functions and therefore allow for generalization of the empirical results. We believe that parameter tuning and invariances need to be carefully regarded for a conclusive performance evaluation of search algorithms.

Altogether, results of this paper together with previous results for the CMA-ES with large population size [3] give evidence that IPOP-CMA-ES is a remarkably well performing search algorithm for non-linear, non-convex, nonseparable global optimization problems.

Acknowledgments

The authors are extremely grateful to Stefan Kern for his considerable contribution to this paper. [redacted]

Bibliography

[1] A. Auger and N. Hansen. Performance evaluation of an advanced local search evolutionary algorithm. In *Proceedings of the IEEE Congress on Evolutionary Computation*, 2005.

[2] N. Hansen. Invariance, self-adaptation and correlated mutations in evolution strategies. In M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J.J. Merelo, and H.-P. Schwefel, editors, *Parallel Problem Solving from Nature - PPSN VI*, pages 355–364. Springer, 2000.

[3] N. Hansen and S. Kern. Evaluating the CMA evolution strategy on multimodal test functions. In Xin Yao et al., editors, *Parallel Problem Solving from Nature - PPSN VIII, LNCS 3242*, pages 282–291. Springer, 2004.

[4] N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evol. Comput.*, 11(1):1–18, 2003.

Table 6: Best objective function error values reached in dimension $n = 30$, see caption of Table 5 for details

FES	Prob.	1	2	3	4	5	6	7	8	9	10	11	12
1e3	min	4.49e+2	1.12e+5	3.84e+8	6.13e+5	6.21e+3	3.26e+6	4.10e+1	2.12e+1	2.19e+2	2.43e+2	4.17e+1	4.23e+5
	7^{th}	5.48e+2	1.73e+5	8.00e+8	1.12e+6	9.57e+3	7.31e+6	9.39e+1	2.12e+1	2.45e+2	2.65e+2	4.43e+1	1.62e+6
	med.	7.40e+2	2.35e+5	1.00e+9	1.44e+6	1.09e+4	1.23e+7	1.20e+2	2.12e+1	2.50e+2	2.74e+2	4.55e+1	1.75e+6
	19^{th}	1.04e+3	2.94e+5	1.38e+9	1.87e+6	1.24e+4	1.99e+7	1.59e+2	2.13e+1	2.66e+2	2.88e+2	4.71e+1	1.90e+6
	max	1.61e+3	3.83e+5	2.07e+9	3.29e+6	1.42e+4	6.81e+7	3.26e+2	2.13e+1	2.87e+2	3.08e+2	4.83e+1	2.16e+6
	mean	8.16e+2	2.39e+5	1.07e+9	1.55e+6	1.07e+4	1.77e+7	1.39e+2	2.12e+1	2.53e+2	2.77e+2	4.54e+1	1.67e+6
	std	3.01e+2	7.80e+4	4.43e+8	6.15e+5	2.13e+3	1.62e+7	7.17e+1	4.35e−2	1.65e+1	1.90e+1	1.76e+0	4.06e+5
1e4	min	3.98e−9	2.29e−3	1.24e+6	4.88e+2	5.00e−2	1.77e+1	3.93e−9	2.10e+1	2.39e+1	3.08e+1	7.44e+0	1.34e+1
	7^{th}	4.70e−9	1.60e−2	3.41e+6	1.46e+3	1.00e+3	2.28e+1	4.85e−9	2.11e+1	4.28e+1	4.38e+1	1.39e+1	6.25e+2
	med.	5.20e−9	2.57e−2	4.90e+6	3.51e+3	1.32e+3	2.58e+1	5.69e−9	2.11e+1	4.88e+1	5.27e+1	1.57e+1	3.02e+3
	19^{th}	6.10e−9	3.99e−2	8.21e+6	5.18e+4	2.04e+3	2.22e+2	6.95e−9	2.11e+1	5.47e+1	5.87e+1	1.85e+1	1.96e+4
	max	7.51e−9	7.49e−2	1.42e+7	2.88e+5	3.20e+3	2.66e+3	2.46e−2	2.12e+1	7.96e+1	8.26e+1	2.28e+1	1.71e+6
	mean	5.42e−9	2.73e−2	6.11e+6	4.26e+4	1.51e+3	4.60e+2	1.77e−3	2.11e+1	4.78e+1	5.14e+1	1.58e+1	2.51e+5
	std	9.80e−10	1.79e−2	3.79e+6	7.43e+4	8.82e+2	8.29e+2	5.52e−3	4.04e−2	1.15e+1	1.25e+1	3.97e+0	5.77e+5
1e5	min	3.98e−9	4.48e−9	4.07e−9	6.06e−9	7.15e−9	4.05e−9	1.76e−9	2.00e+1	2.98e+0	9.95e−1	7.43e−2	4.27e−9
	7^{th}	4.70e−9	5.59e−9	4.78e−9	8.75e−9	8.06e−9	5.31e−9	4.59e−9	2.02e+1	4.97e+0	5.97e+0	7.23e+0	6.01e−2
	med.	5.20e−9	6.13e−9	5.44e−9	1.93e+1	8.61e−9	6.32e−9	5.41e−9	2.09e+1	6.96e+0	6.96e+0	9.23e+0	3.85e+2
	19^{th}	6.10e−9	6.85e−9	6.16e−9	2.72e+3	9.34e−9	7.52e−9	6.17e−9	2.10e+1	8.95e+0	8.95e+0	1.13e+1	1.57e+3
	max	7.51e−9	8.41e−9	8.66e−9	1.57e+5	2.51e−6	3.99e+0	7.81e−9	2.11e+1	1.19e+1	1.09e+1	1.39e+1	1.37e+6
	mean	5.42e−9	6.22e−9	5.55e−9	1.27e+4	1.08e−7	4.78e−1	5.31e−9	2.07e+1	6.89e+0	6.96e+0	9.10e+0	5.95e+4
	std	9.80e−10	8.95e−10	1.09e−9	3.59e+4	4.99e−7	1.32e+0	1.41e−9	4.28e−1	2.22e+0	2.45e+0	3.10e+0	2.74e+5
3e5	min	3.98e−9	4.48e−9	4.07e−9	6.06e−9	7.15e−9	3.27e−9	1.76e−9	2.00e+1	4.35e−6	3.08e−4	8.27e−10	3.79e−9
	7^{th}	4.70e−9	5.59e−9	4.78e−9	8.75e−9	8.06e−9	4.63e−9	4.59e−9	2.00e+1	4.74e−4	9.95e−1	3.53e+0	6.43e−9
	med.	5.20e−9	6.13e−9	5.44e−9	1.93e+1	8.61e−9	5.76e−9	5.41e−9	2.00e+1	9.95e−1	1.00e+0	6.18e+0	4.54e+1
	19^{th}	6.10e−9	6.85e−9	6.16e−9	2.72e+3	9.20e−9	7.06e−9	6.17e−9	2.00e+1	1.99e+0	2.17e+0	7.20e+0	8.16e+2
	max	7.51e−9	8.41e−9	8.66e−9	1.31e+5	9.96e−9	9.25e−9	7.81e−9	2.10e+1	4.97e+0	4.59e+0	1.14e+1	1.10e+6
	mean	5.42e−9	6.22e−9	5.55e−9	1.11e+4	8.62e−9	5.90e−9	5.31e−9	2.01e+1	9.38e−1	1.65e+0	5.48e+0	4.43e+4
	std	9.80e−10	8.95e−10	1.09e−9	3.02e+4	8.53e−10	1.61e−9	1.41e−9	2.79e−1	1.18e+0	1.35e+0	3.13e+0	2.19e+5

FES	Prob.	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	min	3.05e+1	1.37e+1	4.93e+2	2.75e+2	3.10e+2	9.18e+2	9.16e+2	9.17e+2	6.70e+2	1.00e+3	8.14e+2	9.58e+2	2.71e+2
	7^{th}	5.45e+1	1.42e+1	5.66e+2	3.18e+2	4.02e+2	9.29e+2	9.33e+2	9.34e+2	8.63e+2	1.03e+3	9.26e+2	9.77e+2	2.87e+2
	med.	7.36e+1	1.42e+1	6.93e+2	3.57e+2	4.58e+2	9.48e+2	9.45e+2	9.47e+2	9.60e+2	1.06e+3	1.10e+3	9.89e+2	2.95e+2
	19^{th}	1.25e+2	1.43e+1	7.37e+2	4.54e+2	5.38e+2	9.56e+2	9.65e+2	9.60e+2	1.01e+3	1.11e+3	1.11e+3	9.95e+2	3.08e+2
	max	4.98e+2	1.44e+1	8.51e+2	6.01e+2	6.83e+2	9.89e+2	1.03e+3	1.00e+3	1.11e+3	1.21e+3	1.13e+3	1.24e+3	3.80e+2
	mean	1.14e+2	1.42e+1	6.69e+2	3.75e+2	4.79e+2	9.45e+2	9.51e+2	9.50e+2	9.44e+2	1.08e+3	1.03e+3	9.97e+2	3.05e+2
	std	1.07e+2	1.59e−1	1.15e+2	7.97e+1	1.14e+2	1.81e+1	2.47e+1	2.14e+1	1.08e+2	5.78e+1	1.05e+2	5.34e+1	2.75e+1
1e4	min	2.46e+0	1.34e+1	2.08e+2	5.75e+1	6.88e+1	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.63e+2	5.34e+2	2.00e+2	2.10e+2
	7^{th}	3.39e+0	1.37e+1	3.26e+2	7.13e+1	1.59e+2	9.05e+2	9.06e+2	9.04e+2	5.00e+2	8.93e+2	5.35e+2	9.54e+2	2.11e+2
	med.	3.87e+0	1.38e+1	4.00e+2	1.52e+2	2.17e+2	9.07e+2	9.07e+2	9.05e+2	5.00e+2	9.01e+2	5.48e+2	9.56e+2	2.11e+2
	19^{th}	4.10e+0	1.38e+1	4.16e+2	4.00e+2	4.68e+2	9.10e+2	9.08e+2	9.07e+2	5.00e+2	9.10e+2	7.02e+2	9.60e+2	2.12e+2
	max	5.62e+0	1.40e+1	5.53e+2	5.00e+2	6.08e+2	9.14e+2	9.21e+2	9.11e+2	1.09e+3	9.30e+2	1.10e+3	9.64e+2	2.15e+2
	mean	3.80e+0	1.38e+1	3.87e+2	1.96e+2	3.00e+2	9.08e+2	9.08e+2	9.06e+2	5.47e+2	9.00e+2	6.92e+2	9.26e+2	2.11e+2
	std	7.27e−1	1.48e−1	8.48e+1	1.45e+2	1.94e+2	2.76e+0	3.62e+0	2.28e+0	1.63e+2	1.57e+1	2.38e+2	1.51e+2	1.18e+0
1e5	min	2.43e+0	1.27e+1	2.00e+2	2.69e+1	6.67e+1	9.03e+2	9.03e+2	9.03e+2	5.00e+2	7.97e+2	5.34e+2	2.00e+2	2.10e+2
	7^{th}	2.69e+0	1.32e+1	2.00e+2	3.43e+1	1.57e+2	9.03e+2	9.04e+2	9.04e+2	5.00e+2	8.14e+2	5.34e+2	9.38e+2	2.10e+2
	med.	2.83e+0	1.36e+1	2.00e+2	4.27e+1	2.13e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.30e+2	5.34e+2	9.42e+2	2.10e+2
	19^{th}	2.98e+0	1.37e+1	2.22e+2	6.41e+1	4.68e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.40e+2	5.35e+2	9.48e+2	2.11e+2
	max	3.67e+0	1.40e+1	3.20e+2	1.28e+2	5.95e+2	9.07e+2	9.06e+2	9.06e+2	5.00e+2	8.51e+2	1.10e+3	9.56e+2	2.14e+2
	mean	2.89e+0	1.35e+1	2.25e+2	5.34e+1	2.92e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.27e+2	5.82e+2	9.13e+2	2.11e+2
	std	3.59e−1	3.17e−1	4.10e+1	2.85e+1	1.94e+2	7.73e−1	6.07e−1	6.09e−1	1.19e−13	1.72e+1	1.55e+2	1.49e+2	9.27e−1
3e5	min	1.10e+0	1.18e+1	2.00e+2	1.53e+1	6.66e+1	9.03e+2	9.03e+2	9.03e+2	5.00e+2	7.67e+2	5.34e+2	2.00e+2	2.10e+2
	7^{th}	2.44e+0	1.27e+1	2.00e+2	2.50e+1	1.57e+2	9.03e+2	9.03e+2	9.03e+2	5.00e+2	7.88e+2	5.34e+2	9.35e+2	2.10e+2
	med.	2.61e+0	1.29e+1	2.00e+2	3.04e+1	2.13e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.00e+2	5.34e+2	9.39e+2	2.10e+2
	19^{th}	2.76e+0	1.31e+1	2.00e+2	3.49e+1	4.68e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.18e+2	5.34e+2	9.45e+2	2.11e+2
	max	3.20e+0	1.37e+1	3.00e+2	1.08e+2	5.95e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.42e+2	5.34e+2	9.56e+2	2.14e+2
	mean	2.49e+0	1.29e+1	2.08e+2	3.50e+1	2.91e+2	9.04e+2	9.04e+2	9.04e+2	5.00e+2	8.03e+2	5.34e+2	9.10e+2	2.11e+2
	std	5.13e−1	4.19e−1	2.75e+1	2.04e+1	1.93e+2	2.88e−1	2.71e−1	2.48e−1	1.31e−13	1.86e+1	2.22e−4	1.48e+2	9.21e−1

[5] N. Hansen and A. Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In *Proceedings of the 1996 IEEE Conference on Evolutionary Computation (ICEC '96)*, pages 312–317, 1996.

[6] N. Hansen and A. Ostermeier. Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: The $(\mu/\mu_I, \lambda)$-CMA-ES. In *Proceedings of the 5th European Congresson Intelligent Techniques and Soft Computing*, pages 650–654, 1997.

[7] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.*, 9(2):159–195, 2001.

[8] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari. Problem definitions and evaluation criteria for the cec 2005 special session on real-parameter optimization. Technical report, Nanyang Technological University, Singapore, May 2005. http://www.ntu.edu.sg/home/EPNSugan.

Table 7: Best objective function error values reached in dimension $n = 50$, see caption of Table 5 for details

FES	Prob.	1	2	3	4	5	6	7	8	9	10	11	12
1e3	min	5.76e+3	7.22e+5	1.90e+9	2.78e+6	2.24e+4	4.63e+8	4.95e+2	2.12e+1	4.61e+2	5.21e+2	7.46e+1	6.55e+6
	7^{th}	9.19e+3	8.48e+5	2.83e+9	3.29e+6	2.56e+4	1.03e+9	6.25e+2	2.13e+1	5.26e+2	5.73e+2	7.97e+1	7.76e+6
	med.	1.13e+4	9.50e+5	3.59e+9	4.10e+6	2.71e+4	1.47e+9	7.15e+2	2.13e+1	5.38e+2	5.93e+2	8.06e+1	8.07e+6
	19^{th}	1.27e+4	1.05e+6	4.72e+9	4.54e+6	3.01e+4	2.22e+9	8.12e+2	2.14e+1	5.59e+2	6.26e+2	8.20e+1	8.49e+6
	max	1.88e+4	1.53e+6	6.36e+9	7.33e+6	3.51e+4	3.61e+9	1.26e+3	2.14e+1	5.99e+2	6.82e+2	8.33e+1	8.99e+6
	mean	1.12e+4	9.90e+5	3.80e+9	4.16e+6	2.77e+4	1.71e+9	7.38e+2	2.13e+1	5.41e+2	5.94e+2	8.04e+1	8.04e+6
	std	2.96e+3	2.11e+5	1.34e+9	1.09e+6	3.41e+3	9.04e+8	1.70e+2	4.40e−2	3.06e+1	3.87e+1	2.22e+0	6.10e+5
1e4	min	4.51e−9	8.26e+3	3.56e+7	4.04e+5	4.13e+3	4.34e+1	6.47e−5	2.12e+1	6.96e+1	6.47e+1	2.57e+1	9.49e+3
	7^{th}	5.35e−9	1.13e+4	5.27e+7	9.41e+5	5.41e+3	4.60e+1	1.23e−4	2.12e+1	8.95e+1	8.76e+1	2.90e+1	2.16e+4
	med.	5.63e−9	1.39e+4	6.44e+7	1.30e+6	6.57e+3	4.33e+2	2.07e−4	2.13e+1	9.75e+1	1.00e+2	3.08e+1	5.13e+4
	19^{th}	6.68e−9	2.10e+4	8.70e+7	1.44e+6	7.46e+3	6.30e+3	3.07e−4	2.13e+1	1.16e+2	1.19e+2	3.46e+1	6.32e+6
	max	7.28e−9	5.78e+4	2.17e+8	2.85e+6	9.55e+3	3.17e+4	1.00e−2	2.13e+1	1.67e+2	1.82e+2	4.00e+1	7.69e+6
	mean	5.87e−9	1.81e+4	7.67e+7	1.35e+6	6.51e+3	6.52e+3	8.96e−4	2.13e+1	1.04e+2	1.05e+2	3.19e+1	2.25e+6
	std	8.59e−10	1.13e+4	4.23e+7	5.62e+5	1.48e+3	1.11e+4	2.41e−3	4.15e−2	2.20e+1	2.78e+1	4.10e+0	3.29e+6
1e5	min	4.51e−9	6.35e−9	5.12e−2	4.64e+4	5.52e+2	3.17e+0	3.59e−9	2.00e+1	1.01e+1	1.12e+1	1.52e+1	9.67e+0
	7^{th}	5.35e−9	7.44e−9	1.00e+1	3.77e+5	7.49e+2	6.86e+0	6.94e−9	2.11e+1	1.73e+1	2.19e+1	1.77e+1	1.49e+3
	med.	5.63e−9	8.01e−9	3.10e+1	6.02e+5	1.10e+3	9.29e+0	7.41e−9	2.12e+1	2.15e+1	2.69e+1	2.17e+1	2.55e+3
	19^{th}	6.68e−9	8.17e−9	6.15e+1	8.63e+5	1.52e+3	1.63e+1	7.78e−9	2.12e+1	2.39e+1	3.18e+1	2.47e+1	8.57e+3
	max	7.28e−9	9.02e−9	2.76e+2	1.46e+6	2.72e+3	4.72e+2	8.65e−9	2.13e+1	4.28e+1	3.58e+1	2.69e+1	6.41e+6
	mean	5.87e−9	7.86e−9	4.84e+1	6.16e+5	1.21e+3	3.74e+1	7.22e−9	2.11e+1	2.19e+1	2.61e+1	2.12e+1	4.84e+5
	std	8.59e−10	7.24e−10	6.25e+1	3.61e+5	5.68e+2	9.92e+1	1.03e−9	3.30e−1	7.60e+0	6.42e+0	3.83e+0	1.66e+6
5e5	min	4.51e−9	6.35e−9	5.02e−9	3.48e+4	1.09e−5	4.28e−9	3.59e−9	2.00e+1	3.08e−10	1.72e−8	5.85e+0	9.67e+0
	7^{th}	5.35e−9	7.44e−9	5.65e−9	2.57e+5	3.80e−3	6.41e−9	6.94e−9	2.00e+1	6.51e−6	9.95e−1	8.93e+0	5.36e+2
	med.	5.63e−9	8.01e−9	6.19e−9	4.27e+5	5.70e−1	7.28e−9	7.41e−9	2.00e+1	9.95e−1	9.97e−1	1.21e+1	2.36e+3
	19^{th}	6.68e−9	8.17e−9	6.53e−9	6.43e+5	3.61e+0	7.76e−9	7.78e−9	2.00e+1	1.99e+0	1.99e+0	1.40e+1	8.38e+3
	max	7.28e−9	9.02e−9	8.06e−9	1.24e+6	1.35e+1	9.27e−9	8.65e−9	2.11e+1	2.99e+0	5.97e+0	1.68e+1	5.57e+6
	mean	5.87e−9	7.86e−9	6.14e−9	4.68e+5	2.85e+0	7.13e−9	7.22e−9	2.01e+1	1.39e+0	1.72e+0	1.17e+1	2.27e+5
	std	8.59e−10	7.24e−10	6.86e−10	3.11e+5	4.32e+0	1.11e−9	1.03e−9	3.25e−1	1.11e+0	1.42e+0	3.14e+0	1.11e+6

FES	Prob.	13	14	15	16	17	18	19	20	21	22	23	24	25
1e3	min	5.38e+3	2.37e+1	5.73e+2	3.73e+2	4.49e+2	1.01e+3	1.00e+3	9.69e+2	1.04e+3	1.09e+3	1.04e+3	1.03e+3	5.19e+2
	7^{th}	4.10e+4	2.39e+1	8.39e+2	4.33e+2	5.06e+2	1.04e+3	1.03e+3	1.01e+3	1.05e+3	1.15e+3	1.05e+3	1.06e+3	6.42e+2
	med.	7.53e+4	2.41e+1	8.54e+2	4.78e+2	5.73e+2	1.06e+3	1.08e+3	1.04e+3	1.05e+3	1.18e+3	1.06e+3	1.09e+3	6.73e+2
	19^{th}	1.13e+5	2.42e+1	9.23e+2	5.22e+2	6.27e+2	1.08e+3	1.10e+3	1.07e+3	1.06e+3	1.23e+3	1.06e+3	1.11e+3	7.20e+2
	max	3.14e+5	2.43e+1	9.49e+2	6.52e+2	7.48e+2	1.15e+3	1.18e+3	1.14e+3	1.07e+3	1.48e+3	1.08e+3	1.29e+3	1.16e+3
	mean	9.25e+4	2.41e+1	8.49e+2	4.81e+2	5.78e+2	1.07e+3	1.08e+3	1.04e+3	1.05e+3	1.20e+3	1.06e+3	1.09e+3	7.05e+2
	std	7.50e+4	2.00e−1	9.85e+1	7.05e+1	8.45e+1	3.36e+1	4.97e+1	4.68e+1	8.55e+0	8.58e+1	9.73e+0	5.12e+1	1.27e+2
1e4	min	5.11e+0	2.32e+1	3.32e+2	7.22e+1	1.05e+2	9.13e+2	9.16e+2	9.13e+2	1.00e+3	9.38e+2	1.01e+3	2.00e+2	2.15e+2
	7^{th}	6.69e+0	2.36e+1	3.75e+2	1.01e+2	1.70e+2	9.17e+2	9.20e+2	9.18e+2	1.01e+3	9.58e+2	1.02e+3	9.97e+2	2.17e+2
	med.	8.07e+0	2.37e+1	4.14e+2	1.23e+2	2.72e+2	9.24e+2	9.22e+2	9.23e+2	1.01e+3	9.67e+2	1.02e+3	1.00e+3	2.18e+2
	19^{th}	8.77e+0	2.38e+1	4.62e+2	1.83e+2	4.40e+2	9.29e+2	9.23e+2	9.25e+2	1.01e+3	9.77e+2	1.02e+3	1.01e+3	2.18e+2
	max	2.33e+1	2.40e+1	5.23e+2	5.00e+2	6.26e+2	9.52e+2	9.32e+2	9.41e+2	1.02e+3	9.88e+2	1.02e+3	1.02e+3	2.22e+2
	mean	8.68e+0	2.37e+1	4.22e+2	1.73e+2	3.08e+2	9.25e+2	9.22e+2	9.22e+2	1.01e+3	9.67e+2	1.02e+3	9.72e+2	2.18e+2
	std	3.86e+0	2.10e−1	5.79e+1	1.21e+2	1.55e+2	8.82e+0	4.27e+0	6.47e+0	3.33e+0	1.34e+1	2.87e+0	1.61e+2	1.37e+0
1e5	min	4.11e+0	2.13e+1	2.00e+2	3.24e+1	9.79e+1	9.12e+2	9.12e+2	9.13e+2	1.00e+3	8.12e+2	1.01e+3	2.00e+2	2.14e+2
	7^{th}	4.99e+0	2.30e+1	2.10e+2	4.09e+1	1.23e+2	9.14e+2	9.15e+2	9.14e+2	1.00e+3	8.34e+2	1.02e+3	9.86e+2	2.15e+2
	med.	5.40e+0	2.31e+1	2.33e+2	4.50e+1	1.61e+2	9.16e+2	9.17e+2	9.15e+2	1.01e+3	8.51e+2	1.02e+3	9.90e+2	2.16e+2
	19^{th}	5.72e+0	2.32e+1	2.77e+2	6.49e+1	2.69e+2	9.18e+2	9.18e+2	9.17e+2	1.01e+3	8.62e+2	1.02e+3	9.95e+2	2.16e+2
	max	7.29e+0	2.37e+1	3.48e+2	1.63e+2	5.92e+2	9.20e+2	9.21e+2	9.20e+2	1.01e+3	8.84e+2	1.02e+3	1.01e+3	2.17e+2
	mean	5.48e+0	2.30e+1	2.50e+2	6.50e+1	2.36e+2	9.16e+2	9.17e+2	9.16e+2	1.01e+3	8.50e+2	1.02e+3	9.60e+2	2.16e+2
	std	7.44e−1	5.11e−1	4.92e+1	3.89e+1	1.59e+2	2.39e+0	2.43e+0	2.15e+0	1.80e+0	1.89e+1	3.19e+0	1.59e+2	9.20e−1
5e5	min	2.94e+0	2.13e+1	2.00e+2	1.26e+1	9.79e+1	9.11e+2	9.11e+2	9.11e+2	1.00e+3	7.95e+2	1.01e+3	2.00e+2	2.14e+2
	7^{th}	4.25e+0	2.26e+1	2.00e+2	1.72e+1	1.23e+2	9.12e+2	9.12e+2	9.12e+2	1.00e+3	7.99e+2	1.01e+3	9.79e+2	2.15e+2
	med.	4.71e+0	2.30e+1	2.00e+2	2.15e+1	1.61e+2	9.13e+2	9.12e+2	9.12e+2	1.00e+3	8.03e+2	1.01e+3	9.86e+2	2.15e+2
	19^{th}	4.89e+0	2.32e+1	2.00e+2	4.14e+1	2.69e+2	9.13e+2	9.13e+2	9.13e+2	1.00e+3	8.08e+2	1.01e+3	9.92e+2	2.16e+2
	max	5.37e+0	2.37e+1	3.00e+2	1.32e+2	5.92e+2	9.15e+2	9.15e+2	9.13e+2	1.00e+3	8.39e+2	1.02e+3	1.00e+3	2.17e+2
	mean	4.59e+0	2.29e+1	2.04e+2	3.09e+1	2.34e+2	9.13e+2	9.12e+2	9.12e+2	1.00e+3	8.05e+2	1.01e+3	9.55e+2	2.15e+2
	std	5.15e−1	5.78e−1	2.00e+1	2.53e+1	1.58e+2	8.42e−1	7.26e−1	5.05e−1	8.19e−1	8.86e+0	1.86e+0	1.58e+2	9.07e−1