INCREASING POPULATION $(\mu + \lambda)$-CMA-ES
WITH CENTER AND ELITISM ($IPOP!+$)

Mark Wineberg and Samuel Opawale

University of Guelph
School of Computer Science
50 Stone Road East Guelph, Ontario, N1G 2W1
Canada
mwineber@uoguelph.ca, sopawale@uoguelph.ca

Abstract: *Elitism has previously been introduced to the CMA-ES family of algorithms, where the "," selection operator is replaced by the "+" selection operator. Here we investigate in detailed the addition of elitism to IPOP. Furthermore, a new selection operator was added: the "!" operator (pronounced "bang" or "here"). This operator includes the results of ES recombination into the population for selection, unmodified by mutation, and evaluated separately. From the analysis, we noticed a remarkable improvement in the behavior of IPOP with or without elitism. Only one function (Levy) proved difficult when elitism was used. Under close examination, it was determined that for this function, the population under elitism converges prematurely, and stalled out. Currently we do not know what is the cause of this difference. Perhaps in the future this effect could be avoided or detected and remedial measures applied.*

Keywords: *CMA-ES, center, elitism, restart, selection, IPOP, IPOP+, IPOP!, IPOP!+.*

1 Introduction

Covariance matrix Adaptation evolution strategies (CMA-ES) over the years has proven to be one of the most powerful evolutionary algorithm(EA) for real value optimization [5]. As proposed in 1996 by Hansen and Ostermeier, the idea of CMA-ES is to have a system that enables an evolution strategies (ES) algorihm to adapt to the correct scaling of a given problem and also ensure invariance with respect to any rotation of the fitness function [5].

Researchers over the years have worked on CMA-ES and created several variants of the algorithm over the years. Some of the variant include: LS (Least-Square)-CMA-ES proposed in 2004 [3], that uses the idea of increasing the performance of the algorithm by tuning the covariance matrix of the algorithm. Local Restart (LR-CMA-ES) [1] and IPOP [1] immediately followed in 2005 with the idea of introducing a restart to the algorithm when stagnation was detected with increasing population size on the restart. Over the years people continued to improve on the basic CMA-ES algorithm many of which can be found in [4]. Recently, Van Rijn isolated each feature that have been previously proposed and performed a massive study combining the features in various combination against a varied test width to see what combination of features work well together and if any dominate [9]. In doing so, Van Rijn added the ES style "+" selection operator which he called elitism [8, 9] as one of the various features. While $(\mu + \lambda)$-ES style selection is obviously not novel, applying it to CMA-ES was. Traditionally, (μ, λ)-ES style selection has been used with some variants using the $(1 + 1)$-ES style hill climber as well as $(1 + \lambda)$ [6, 7]. To our knowledge, extending $(1 + 1)$-ES style selection to $(\mu + \lambda)$ was never previously done. Of all the features studied by Van Rijn, the ones that most consistently conferred improvement was elitism.

In this paper, we will be further investigating elitism with IPOP's restart which we called "IPOP+" . Also, we shall introduce a new minor variant to elitism: taking the center computed by CMA-ES upon which the offspring are computed, evaluating it and adding it into the mix for selection as a parent, when this feature is used alone we will refer to it as IPOP! and when used in combination with elitism, we will refer to it as IPOP!+.

This paper is structured as follows: section II provides a detailed information about the systems from CMA-ES to IPOP, elitism(IPOP+), Center(IPOP!) and combined (IPOP!+). Section III shows the first part of the experimental designed adopted in this paper while the results were discussed in section IV. A second phase of experiment follow, investigating a particular phenomenon observed in the first round of experiment, section V details the experimental set up for thi phase and section VI present the results and discussion.

2 Systems

Increasing Population CMA-ES (IPOP) is an extension of Local Restart CMA-ES (LR- CMA-ES) [1, 2]. Whenever a run of the (μ_W, λ) - CMA-ES is terminated due to one of the five to six local stagnation criterion [2], the population size is increased by a factor η for the next independent run of the (μ_W, λ) - CMA-ES restarted from the beginning. Elitism(IPOP+) is a version of IPOP that uses a $(\mu_W + \lambda)$ CMA-ES [8, 9]. Van Rijn found that, adding elitism to the system often shows a massive improvement over IPOP when tested on many test functions.

Center(IPOP!) is a new approach introduced here. Where the center is added to the generated offspring before truncation. This is basically taking the results of the ES recombination used in evaluating it without adding mutation, and directly including those members into the population to be selected from as well. Since CMA-ES uses weighted recombination to produce a single individual, the center, we call this "! selection", where "!" is pronounced "here" as we are including the location of the center which is where the population currently resides; "it is here". Finally, elitism and center are combined producing the IPOP!+ system. The idea of this is to combine the effective power of elitism and center in a single system to see their effect on different test functions.

The performance of IPOP, IPOP !, IPOP+ and IPOP!+ is investigated on variety of test functions: Ackley, Rastrigin, Griewank, Levy, and elliptical, each having very different properties from unimodal to multimodal with simple to complex structure.

3 Experimental Design - Part 1: Main Effects

For our initial experimental design, we performed a fully factorial design which includes the following factors: five fitness functions (Ackley, Rastrigin, Levy, Griewank and elliptical), two number of dimensions (25, 50) against the four different systems either with or without elitism or with or without centre. Each treatment was repeated 60 times.

In order to use the μ parents within the selection routine, care has to be taken when updating the CMA-ES central model. Van Rijn [8] does not mention how he accounted for this. Here, when computing the model, if a member of the new μ - selected population comes originally from the parent, the error term from the center of the model used to generate that member is recomputed through measuring the Euclidean distance from that member to the center of the current model, the same as for the offspring(not the Euclidean distance from the center that had generated the parental member which may have been from the previous generation or perhaps even earlier if that member is specially fit and had survived across multiple generations). The center has a Euclidean distance to the model center of $\vec{0}$ which is what is used when updating the model.

The fitness functions used are now detailed:

1. Elliptical:

$$f(x) = \sum_{i=1}^{d} \sum_{j=1}^{i} x_j^2 \tag{1}$$

2. Ackley:

$$f(x) = -a \exp(-b\sqrt{\frac{1}{d} \sum_{i=1}^{d} x_i^2}) - \exp(\frac{1}{d} \sum_{i=1}^{d} cos(cx_i)) + a + exp(1) \tag{2}$$

$$where \ \ a = 20, \ b = 0.2 \ and \ c = 2\pi$$

3. Griewank:

$$f(x) = \sum_{i=1}^{d} \frac{x_i^2}{4000} - \pi_{i=1}^{d} cos(\frac{x_i}{\sqrt{i}}) + 1 \tag{3}$$

4. Rastrigin:

$$f(x) = 10d + \sum_{i=1}^{d} [x_i^2 - 10cos(2\pi x_i)] \tag{4}$$

5. Levy:

$$f(x) = sin^2(\pi\omega_1) + \sum_{i=1}^{d-1}(\omega_i - 1)^2[1 + 10sin^2(\pi\omega_i + 1)]$$
$$+ (\omega_d - 1)^2[1 + sin^2(2\pi\omega_d)] \quad (5)$$
$$where \quad \omega_i = 1 + \frac{x_i - 1}{4}, \quad for \; all \quad i = 1, ..., d$$

The starting location of the center for the algorithm is randomly chosen point of $\sqrt{(d)}$ unit Euclidean distance from known optimum. If the solution is $\vec{0}$, many researchers will choose the point $\vec{1}$ has the starting point. Here we follow this pattern, instead of starting at $\vec{1}$, we start from the same Euclidean distance which is $\sqrt{(d)}$ but in a random direction to eliminate any bias. In all cases, the initial $\sigma = 0$.

For all systems, the default values used for IPOP, as presented in [2], were followed, which sets the initial population size λ to $\lambda = 4 + \lfloor 3 \log d \rfloor$, and the population size update $\eta = 2$.

The run is terminated if the fitness evaluation of the best member of the population comes within ϵ of the fitness of the optimum, where ϵ is set to 10^{-5}, or 1,000,000 evaluation has been performed, which ever come first.

4   Results and Discussion - Part 1

To analyze the results of the factorial design detailed above, a multi-way ANOVA was used. A normal distribution plot was performed on the residuals from the ANOVA, and it was determined to be highly non-normal. The typical technique for performing a non-parametric ANOVA to handle non-normal data is either Kruskal-Wallis or Friedman. Unfortunately, both only perform a one way and not a multi-way ANOVA as needed here. As indicated in [10], a Box-Cox transform can often turn non-normal EC results to become normally distributed. The Box-Cox transform is:

$y' = \frac{y^p - 1}{p}$, where y is the number of evaluations a run took before terminating (either by finding a solution or hitting maximum number of evaluations), and p is set by the researcher.

Table 1: Analysis of the IPOP experiment using ANOVA on a reduced model after the Box-Cox transformation has been applied. Here fn = function, dim = dimensionality, cnr = include center (!), emn = elitism (+). Notice that there are not only main effects, but interaction effects as well.

Factors	df	Sums of Squares	Mean Square	F-ratio	P-value
Intercept	1	77740.9	77740.9	3383200	< 0.0001
fn	4	403.211	100.803	4386.8	< 0.0001
dim	1	72.4584	72.4584	3153.3	< 0.0001
fn*dim	4	7.3598	1.83995	80.073	< 0.0001
elm	1	39.2938	39.2938	1710	< 0.0001
fn*elm	4	467.946	116.987	5091.1	< 0.0001
fn*dim*elm	4	6.2681	1.56702	68.195	< 0.0001
cnr	1	25.1699	25.1699	1095.4	< 0.0001
fn*cnr	4	4.94701	1.23675	53.822	< 0.0001
elm*cnr	1	9.45385	9.45385	411.42	< 0.0001
fn*elm*cnr	4	3.90802	0.977005	42.518	< 0.0001
Error	2371	54.4821	0.022979		
Total	2399	1094.5			

Following this procedure, we manually explored a variety of Box-Cox p settings to maximize the r^2 measure that evaluates the fit of a regression line through the normal-distribution plot. We found that a Box-Cox p setting of -0.092 produced the most normal result with an $r^2 = 97.4\%$, which is highly normally distributed.

The results of the ANOVA can be seen in Table 1. Here we can see that the system does behave differently under center and elitism than without them (as well as showing interaction effects). To see which settings performed better pair-wise comparisons were done using Student T tests based on the ANOVA model, with Sheffe post-hoc correction applied. The results can be seen in Table 2.

Table 2: The average number of evaluations, with confidence intervals, to find the optimum within epsilon = e-5 for the different experimental setups (with and without elitism, with and without the center added, across the five different fitness functions). Note: the number of evaluations is averaged after the Box-Cox transform has been applied, and the reported after the average has been converted back using the inverse transform.

Fn	Elitism	Center	Lower Bound	Average	Upper Bound
ackley	FALSE	FALSE	8,266.18	9,679.22	11,366.90
ackley	FALSE	TRUE	5,963.97	7,108.36	8,503.02
ackley	TRUE	FALSE	5,222.75	6,325.20	7,693.45
ackley	TRUE	TRUE	1,543.33	1,865.16	2,263.62
elliptical	FALSE	FALSE	6,887.59	8,234.41	9,881.50
elliptical	FALSE	TRUE	5,937.65	7,074.95	8,460.52
elliptical	TRUE	FALSE	3,695.74	4,613.24	5,792.01
elliptical	TRUE	TRUE	1,339.07	1,669.46	2,093.34
griewank	FALSE	FALSE	4,753.68	5,643.88	6,723.97
griewank	FALSE	TRUE	3,370.30	4,070.75	4,937.41
griewank	TRUE	FALSE	2,790.86	3,458.76	4,309.83
griewank	TRUE	TRUE	1,133.21	1,377.24	1,681.31
levy	FALSE	FALSE	8,247.14	10,186.33	12,647.77
levy	FALSE	TRUE	6,866.96	8,523.09	10,637.02
levy	TRUE	FALSE	1,001,909.80	1,002,366.87	1,002,824.18
levy	TRUE	TRUE	1,001,796.49	1,002,268.74	1,002,741.24
rastrigin	FALSE	FALSE	618,434.42	693,976.01	779,950.74
rastrigin	FALSE	TRUE	474,151.95	558,152.24	659,084.66
rastrigin	TRUE	FALSE	5,007.26	6,186.97	7,685.05
rastrigin	TRUE	TRUE	1,858.64	2,308.27	2,882.58

Please note that, while the ANOVA and summary information (such as the means for the various factor levels and treatments) are performed on the transformed data, when reporting the results, the inverse transform must be applied to obtain the "average" number of evaluations performed for that factor setting. Confidence intervals can be produced in the same manner by taking the standard error, multiplying it by the appropriate T-score. For our confidence intervals, we chose a 99% confidence level with the appropriate Bonferroni post-hoc correction for all possible comparisons between treatment means that could be done using the results found in Table 2. These comparisons, using confidence interval overlap, most match the more accurate p-values produced using the pair-wise comparisons from the ANOVA, which is not shown here for space consideration; if different we will make note of it when important.

On a further note: if you look carefully, you may notice that the mean does not fall in the exact center of the confidence interval. This is expected when doing a Box-Cox transform with a p setting far from one. The p setting we used is close to 0, which produces a geometric mean and not an arithmetic mean that would produce symmetrical +- error bounds for the confidence intervals.

When observing the results in Table 2, many observations of interest become readily apparent in no particular order, note the following:

1. The effects of center (IPOP!): IPOP! shows improved performance over IPOP on Ackley, Rastrigin, Elliptical, Griewank but not on levy where it slightly hurts rather than helps (as determined by the p-values produced by the ANOVA). It found its solution within its maximum evaluation cut-off.

2. The effects of Elitism (IPOP+): On Ackley, elliptical, Rastrigin, and Griewank IPOP+ shows massive improvement over IPOP and some improvements over IPOP! but perform worse on Levy.

3. The effects of Elitism and Center(IPOP!+): IPOP!+ shows tremendous improvement over IPOP, IPOP! and IPOP!+ with Ackley, Rastrigin, Griewank and elliptical. It converges extraordinarily fast and often finds the solution only using many orders of magnitude fewer evaluations when compare to others. Interestingly, IPOP!+ performs worst on Levy especially when compared with the original IPOP. The presence of elitism seems to be hurting Levy fitness function. Our second set of experiments are designed to investigate this effect.
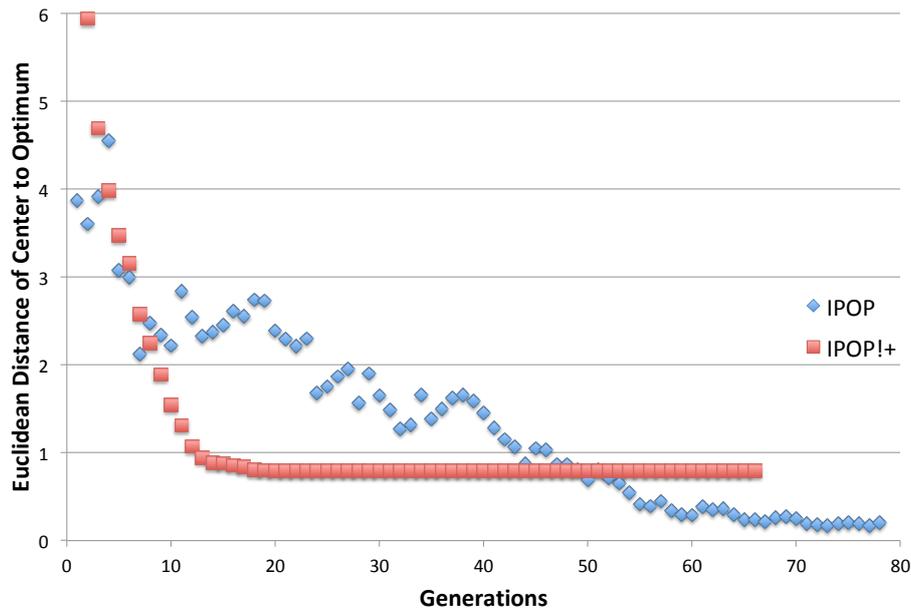
Figure 1: The Euclicean distance to the optimum of the best member from IPOP and IPOP!+, recorded generation by generation, solving the Levy function. It can be easily seen that, after making quick progress, IPOP!+ stalls by generation 15.

5 Experimental Design - Part 2: Why is Levy different?

The question we are trying to answer in this section is what is making Levy so different from other fitness function? What is making it performing badly on IPOP+ and IPOP!+? What factor is responsible for its nice performance on IPOP? These and many more questions will be answered in this section.

Our general set up for this design is in two phases: The first phase is to concentrate on levy (where elitism hurts) and rastrigin which is difficult for IPOP but very easy for IPOP!+. In order to investigate what is really going on in the algorithm, we looked at the systems generation by generation with 3 reps per settings in other to get a clear view of how IPOP, IPOP!, IPOP+, and IPOP!+ are behaving on Levy.

The second phase of our design, we call "Shadowing". A "Shadow" is a system different from the one being run, yet applied to the population at each generation to produce what that system "thinks" should be the population and/or model for the next generation, although those values are discarded each time. Two systems were used: IPOP and IPOP!+ with the idea of computing the offspring and center independent of the system's run in order to get the full knowledge of the type of offspring and center IPOP is giving us on Levy compared to when the elitism is on IPOP!+. We do not use "full shadowing" as we only look at the center and the offspring that the shadow produces and do not look at the update of the paths step-size and covariance matrices.

5.1 Measures

The performance measures were based on the max, min, weighted upper and lower quartile, and weighted median as well as the weighted average across the selected population with the weighting used being same as used by CMA-ES and IPOP.

Also, we look at the measure of condition of C, Model step-size, center value of the model (weighted average), fitness of center, fitness of offspring etc.

6 Results and Discussion - Part 2

6.1 Results

The results can be seen in Table 3 - 5 and Figure 1. The first two tables shows the generation by generation results of IPOP!+ on Rastrigin and Levy respectively. Here, we track the breakdown of where the members from the μ-size post-selection population originate from. Do they come from the parent, from the offspring or was that member the center? The first panel looks at the percentages of the selected population that comes from the three sources, the second panel looks at the relative rank of those members within the population and the third panel reflects the fitness of those members. If a "-" is present it indicates no members was selected

Table 3: Generation-by-generation monitoring of IPOP!+ in a sample run when solving the Rastrigin problem.

| | | | | Rank | | | | | | | Fitness | | | | | | |
| | | | | Parents | | | Center | Offspring | | | Parents | | | Center | Offspring | | |
gen	Parents	Center	Offspring	Max	Avg	Min	Rank	Max	Avg	Min	Max	Avg	Min	Fitness	Max	Avg	Min
1	67%	0%	33%	6	3.8	1	-	4	3.00	2	190.41	166.52	138.31	-	168.6	157.30	145.99
2	50%	17%	33%	5	2.7	1	4	6	4.50	3	160.15	148.15	138.31	158.98	162.7	159.40	156.10
3	50%	0%	50%	6	4.0	1	-	4	3.00	2	156.10	146.80	138.31	-	145.8	142.90	139.14
4	100%	0%	0%	6	3.5	1	-	-	-	-	156.10	144.85	138.31	-	-	-	-
5	67%	0%	33%	6	4.0	2	-	4	2.50	1	145.84	141.75	138.31	-	143.1	139.23	135.39
6	67%	17%	17%	6	4.3	2	1	3	3.00	3	143.08	138.98	135.39	114.44	136.4	136.35	136.35
7	100%	0%	0%	6	3.5	1	-	-	-	-	143.08	134.45	114.44	-	-	-	-
8	50%	17%	33%	4	2.7	1	5	6	4.00	2	136.35	128.73	114.44	136.96	137.2	133.34	129.47
9	33%	17%	50%	5	3.5	2	3	6	3.67	1	129.47	121.96	114.44	116.69	132.9	118.61	98.14
10	67%	17%	17%	6	3.5	1	2	5	5.00	5	124.80	113.52	98.14	111.89	123.7	123.73	123.73
11	50%	17%	33%	6	4.0	2	1	5	4.00	3	114.44	108.16	98.14	97.51	113.7	109.96	106.18
12	83%	0%	17%	6	3.4	1	-	4	4.00	4	113.74	105.49	97.51	-	111.5	111.51	111.51
13	83%	17%	0%	6	4.0	2	1	-	-	-	111.89	105.05	97.51	88.35	-	-	-
14	67%	17%	17%	6	3.8	2	1	5	5.00	5	106.18	97.55	88.35	83.96	105.2	105.19	105.19
15	83%	17%	0%	6	4.0	2	1	-	-	-	105.19	94.63	83.96	81.91	-	-	-
16	83%	17%	0%	6	3.6	1	3	-	-	-	98.14	89.98	81.91	85.64	-	-	-
17	67%	0%	33%	5	3.5	2	-	6	3.5	1	88.35	84.97	81.91	-	95.7	82.83	69.95
18	83%	17%	0%	6	3.8	1	2	-	-	-	88.35	81.97	69.95	70.46	-	-	-
19	83%	17%	0%	6	3.6	1	3	-	-	-	85.64	78.39	69.95	73.62	-	-	-
20	83%	17%	0%	6	4.0	2	1	-	-	-	83.96	75.98	69.95	66.50	-	-	-
21	83%	17%	0%	6	4	2	1	-	-	-	81.91	72.49	66.50	61.51	-	-	-
22	83%	17%	0%	6	4	2	1	-	-	-	73.62	68.41	61.51	49.7621	-	-	-
23	83%	17%	0%	6	4	2	1	-	-	-	70.46	63.64	49.76	36.2422	-	-	-
24	83%	17%	0%	6	4	2	1	-	-	-	69.95	56.79	36.24	23.8376	-	-	-
25	83%	17%	0%	6	4	2	1	-	-	-	66.50	47.57	23.84	14.1191	-	-	-
26	67%	17%	17%	6	3.75	2	1	5	5	5	49.76	30.99	14.12	7.8881	47.1	47.12	47.12
27	83%	17%	0%	6	4	2	1	-	-	-	47.12	25.84	7.89	4.15444	-	-	-
28	83%	17%	0%	6	4	2	1	-	-	-	36.24	17.25	4.15	2.10166	-	-	-
29	83%	17%	0%	6	4	2	1	-	-	-	23.84	10.42	2.10	1.07529	-	-	-
30	83%	17%	0%	6	4	2	1	-	-	-	14.12	5.87	1.08	0.51968	-	-	-
31	83%	17%	0%	6	4	2	1	-	-	-	7.89	3.15	0.52	0.2401	-	-	-
32	83%	17%	0%	6	4	2	1	-	-	-	4.15	1.62	0.24	0.10665	-	-	-
33	83%	17%	0%	6	4	2	1	-	-	-	2.10	0.81	0.11	0.04563	-	-	-
34	83%	17%	0%	6	4	2	1	-	-	-	1.08	0.40	0.05	0.01881	-	-	-
35	83%	17%	0%	6	4	2	1	-	-	-	0.52	0.19	0.02	0.00746	-	-	-

from the source that generation.

The figures show the fitness of the best member in the selected population or the Euclidean distance of that members solutions to the optimum.

6.2   Discussion and Conclusion

From Figure 1 we can see that IPOP precede to the solution fairly rapidly obtaining near optimal values by generation 20 with continuing improvement until a solution within ϵ was found by generation 68. In contrast, IPOP!+ seems to stall out early around generation 12.

The explanation therefore can be seen by comparing Table 3 to Table 4, for Rastrigin in Table 3, at the beginning of the run, the parent and the offspring can be seen has been selected with approximately equal frequency while the center rarely is selected. By generation 10 however, the center starts to always be present usually rank first. Meanwhile the offspring, are now never chosen. This in effect allows IPOP!+ to become a hill climber once the population has owned in on a good location which can account for its massive increase in speed.

However, when looking at table IPOP!+ is performing on Levy with different behavior from Rastrigin. By generation 13, the center quickly dominates and its ranked 1 very early but by generation 14, it's no longer chosen. The percentage of parent start out roughly equal to that of the offspring but begin to climb until by generation 21 only parents are selected, this has completely stalled the process. No offspring produced are of any value and the model become unchanging.

However, when we looked at the step-size σ being produced, as well as the condition of the covariance matrix C, both became unstable when the population stalled. For smaller population, the step-size would decrease as expected, while for larger populations, which would be obtained after doubling over a few restarts, the step-size was seen to climb, with faster rates for larger populations (which would account for no offspring ever being produced becoming better than any parent). For small populations size, the condition of the covariance matrix would grow from 1 to around 60. For large populations, the covariance condition rose and then started to oscillate between 30 and 300. We are still investigating what properties of Levy, in combination with the CMA-ES algorithm, cause these effects.

Finally, looking at Table 5, we tracked, the mirror of IPOP!+ while running IPOP on Levy. What we see is that early on the center while ranked highly bounces around quite a bit by generation 40, it has settled into the

Table 4: Generation-by-generation monitoring of IPOP!+ in a sample run when solving the Levy problem

	Percentage			Rank Parents (Rank)			Center	Offspring (Rank)			Fitness Parents (Fit)			Center	Offspring (Fit)		
gen	Parents	Center	Offspring	Max	Avg	Min	Rank	Max	Avg	Min	Max	Avg	Min	Fitness	Max	Avg	Min
1	50%	4%	46%	23	12.3	1	13	24	12.64	2	7.95	6.37	3.73	6.69	8.25	6.52	5.20
2	33%	4%	63%	22	14.4	6	7	24	11.87	1	5.91	5.28	3.73	3.86	5.94	4.77	2.73
3	29%	4%	67%	17	12.3	6	1	24	13.31	2	3.86	3.45	2.73	1.89	4.69	3.56	2.35
4	75%	4%	21%	24	12.9	2	1	22	13.20	3	3.98	3.14	1.89	1.13	3.74	3.08	1.96
5	63%	4%	33%	24	12.4	2	1	23	14.13	5	3.47	2.64	1.13	1.08	3.47	2.81	2.01
6	75%	4%	21%	24	13.9	2	1	14	9.80	4	3.15	2.41	1.08	0.91	2.50	2.21	1.62
7	63%	4%	33%	23	12.7	2	1	24	13.50	5	2.51	1.97	0.91	0.76	2.58	2.12	1.46
8	58%	4%	38%	24	13.4	2	1	21	12.33	3	2.25	1.66	0.76	0.61	2.19	1.61	0.88
9	63%	4%	33%	24	11.5	2	1	23	15.88	9	1.89	1.28	0.61	0.55	1.85	1.55	1.14
10	63%	4%	33%	24	12.6	1	3	21	13.50	5	1.55	1.11	0.55	0.61	1.42	1.20	0.87
11	58%	4%	38%	23	11.3	1	2	24	15.56	5	1.30	0.91	0.55	0.61	1.31	1.07	0.71
12	67%	4%	29%	24	11.6	1	6	23	15.57	8	1.07	0.84	0.55	0.76	1.07	0.93	0.78
13	67%	4%	29%	23	11.3	1	19	24	14.29	5	0.93	0.78	0.55	0.89	0.94	0.84	0.66
14	75%	0%	25%	24	12.2	1	-	17	13.33	7	0.89	0.77	0.55	-	0.85	0.80	0.75
15	83%	0%	17%	23	12.8	1	-	24	11.25	5	0.88	0.75	0.55	-	0.88	0.73	0.61
16	83%	0%	17%	23	11.6	1	-	24	17	8	0.85	0.72	0.55	-	0.86	0.79	0.68
17	83%	0%	17%	22	11.7	1	-	24	16.5	8	0.80	0.71	0.55	-	0.84	0.76	0.68
18	96%	0%	4%	24	12.6	1	-	10	10	10	0.81	0.72	0.55	-	0.70	0.70	0.70
19	96%	0%	4%	24	12.7	1	-	8	8	8	0.80	0.71	0.55	-	0.68	0.68	0.68
20	92%	0%	8%	24	12.2	1	-	16	15.5	15	0.80	0.70	0.55	-	0.75	0.75	0.75
21	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
22	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
23	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
24	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
25	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
26	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
27	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
28	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
29	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
30	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
31	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
32	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
33	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
34	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-
35	100%	0%	0%	24	12.5	1	-	-	-	-	0.80	0.71	0.55	-	-	-	-

top ranked position among potential offspring. This means that, if moved into a good location IPOP!+ starts to be perform similarly to what we saw with Rastrigin. This provides hope that a technique that combines the two properties, the exploration ability of IPOP with the exploitation capability of IPOP!+ could be in the future be developed.

References

[1] Auger, A., Hansen, N.: Performance evaluation of an advanced local search evolutionary algorithm. In: Evolutionary Computation, 2005. The 2005 IEEE Congress on, vol. 2, pp. 1777–1784. IEEE (2005)

[2] Auger, A., Hansen, N.: A restart cma evolution strategy with increasing population size. In: Evolutionary Computation, 2005. The 2005 IEEE Congress on, vol. 2, pp. 1769–1776. IEEE (2005)

[3] Auger, A., Schoenauer, M., Vanhaecke, N.: Ls-cma-es: A second-order algorithm for covariance matrix adaptation. In: International Conference on Parallel Problem Solving from Nature, pp. 182–191. Springer (2004)

[4] Back, T.: Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford university press (1996)

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

[6] Igel, C., Hansen, N., Roth, S.: Covariance matrix adaptation for multi-objective optimization. Evolutionary computation **15**(1), 1–28 (2007)

[7] Igel, C., Suttorp, T., Hansen, N.: A computational efficient covariance matrix update and a (1+ 1)-cma for evolution strategies. In: Proceedings of the 8th annual conference on Genetic and evolutionary computation, pp. 453–460. ACM (2006)

[8] van Rijn, S., Wang, H., van Leeuwen, M., Bäck, T.: Evolving the structure of evolution strategies. In: Computational Intelligence (SSCI), 2016 IEEE Symposium Series on, pp. 1–8. IEEE (2016)

Table 5: Monitoring IPOP!+ shadowing IPOP in a sample run when solving the Levy problem - average ranks among the selected members that come from the parents, center and offspring respectively.

gen	Parents	Center	Offspring	gen	Parents	Center	Offspring	gen	Parents	Center	Offspring
1	4.0	NaN	1.00	36	3.5	2	4.00	71	4.0	1	4.00
2	6.0	3	3.00	37	4.3	1	3.50	72	4.3	1	3.50
3	3.0	5	3.50	38	4.0	1	4.00	73	4.0	1	4.00
4	3.5	6	2.67	39	3.0	1	4.25	74	3.0	1	4.67
5	2.0	1	4.50	40	4.0	1	4.00	75	4.0	1	4.00
6	5.5	2	2.67	41	4.5	1	3.67	76	5.0	1	3.75
7	3.7	1	4.50	42	4.0	1	4.00	77	5.0	1	3.33
8	6.0	2	3.25	43	4.0	1	4.00	78	4.3	1	3.50
9	3.0	1	4.67	44	4.3	1	3.00	79	4.0	1	4.00
10	5.0	2	3.00	45	3.0	1	4.25	80	4.0	1	4.00
11	4.0	1	4.00	46	5.0	1	3.75	81	4.3	1	3.50
12	6.0	1	3.50	47	4.5	1	2.00	82	3.5	1	4.33
13	3.0	1	4.67	48	4.0	1	4.00	83	3.0	1	4.25
14	3.0	1	5.50	49	3.5	2	4	84	4.0	1	4
15	5.0	2	3.5	50	4.0	1	4	85	3.0	2	4
16	4.7	2	2.5	51	4.3	1	3.5	86	2.5	2	4.7
17	3.3	2	4.5	52	3.0	1	4.7	87	4.0	1	4
18	5.0	1	2.5	53	5.0	1	3.3	88	4.0	1	4
19	3.5	1	4.3	54	4.5	1	3.7	89	4	1	4
20	3.7	2	4	55	5	1	3.3	90	6	1	3.5
21	6	4	2.75	56	2.5	1	5	91	4	1	4
22	3.75	1	5	57	4	1	4	92	3.5	1	4.3
23	4.25	1	3	58	5.5	1	3	93	5	1	2.5
24	4.5	2	3.3	59	3	1	4.7	94	5	1	3.3
25	2	3	4	60	4.25	1	3	95	5.5	1	3
26	4	2	3.5	61	5	1	3.75	96	3.5	1	4.3
27	3.5	2	4	62	4.5	1	2	97	5	2	3
28	4.3	1	3.5	63	2	1	4.5	98	4	1	4
29	4	1	4	64	4	1	4	99	4.7	1	3
30	5.5	2	2.7	65	2	1	4.5	100	4.5	1	3.7
31	3.3	3	4	66	4.5	1	3.7	101	3	1	4.7
32	4	2	3.5	67	4	1	4	102	4	1	4
33	4	2	3	68	3.75	1	5	103	4	1	4
34	3	1	4.25	69	6	1	3.5	104	3.7	1	4.5
35	3	2	4	70	3.5	1	4.33	105	5.5	1	3.0

[9] van Rijn, S., Wang, H., van Stein, B., Bäck, T.: Algorithm configuration data mining for cma evolution strategies. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 737–744. ACM (2017)

[10] Wineberg, L.: Reexpressing problematic optimization data. In: Proceedings of the 2017 Annual Conference on Genetic and Evolutionary Computation, pp. 897–904. ACM (2017)