Applying Memetic algorithm with Improved L-SHADE and Local Search Pool for the 100-digit challenge on Single Objective Numerical Optimization

1st Daniel Molina
DASCI Andalusian Institute of Data Science and Computational Intelligence, University of Granada
Granada, Spain
dmolina@decsai.ugr.es

2nd Francisco Herrera
DASCI Andalusian Institute of Data Science and Computational Intelligence, University of Granada
Granada, Spain
herrera@decsai.ugr.es

Abstract—In this paper, we have proposed a new optimization algorithm, Memetic improved L-SHADE with a local search pool, MiLSHADE-LSP, a memetic algorithm that combines an improved L-SHADE with a local search pool. Improved L-SHADE modifies several important parameters during the run to encourage exploration in initial stages and to focus later the search around the most promising solutions. The local search pool is responsible to continuously improve the best solutions. MiLSHADE-LSP uses a pool of two different local search, LS, methods, the Broyden-Fletcher-Goldfarb-Shanno method with limited memory, L-BFGS-B, and the Solis-Wets algorithms, with an adaptive mechanism to choose which one of them is applied in each iteration selecting which had obtained a greater improvement last time it was applied. In order to avoid waste LS applications, the proposed algorithm stores a list of individuals that were not previously improved by each LS method. It also includes a restart mechanism to explore new areas when the search is stuck, restarting the population but maintaining the best found solution, and resetting the LS Pool parameters. In the experimental section we have tested and analyzed MiLSHADE-LSP using the proposed benchmark for the competition 100-digit challenge on Single Objective Numerical Optimization, obtaining that the LS Pool improves the algorithm, both achieving more optima and with a better performance. Results obtained show that MiLSHADE-LSP is a very competitive algorithm.

Index Terms—Continuous optimization, global optimization, memetic algorithm, single objective numerical optimization, numerical optimization, differential evolution.

I. INTRODUCTION

There are many optimization problems in which there is not any suitable and exact technique to solve it in a reasonable time. In these problems, Evolutionary Algorithms, EA [1], have arisen as a very good alternative, because they are able to obtain acceptable results with limited resources.

In order to obtain an adequate search, it is important to achieve a good trade-off between a exploration of all domain search and a exploitation of the most promising solutions. For doing that, hybrid meta-heuristics, like Memetic Algorithms, MA [2], [3], which combine different algorithms as components, are widely used. Initially, MAs were the combination of a population-based algorithm with a local search improvement, but nowadays there are more complex combination of different algorithms. Also, the convenience of using a pool of different local search methods has been proved [4].

Another important factor that can contribute to a good balance is focused on exploring in the early stages, and later, as the search progresses, on seeking more in depth around the best solutions. One option is to use initially a greater population to maintain a higher diversity at initial stages, and during the run to decrease the population size, as L-SHADE [5] does. Another strategy is to adapt several parameters during the search to enforce exploitation as the search evolves, as the improved L-SHADE [6].

In this work, we propose a new memetic algorithm, memetic improved L-SHADE with a local search Pool, MiLSHADE-LSP, which combines the continuous exploitation by using a local search pool with an EA, improved L-SHADE, iL-SHADE, that enforces the exploitation during the search. The main features of our proposal are:

- A decreasing population size, to enforce diversity at the beginning of the search, and focusing the search around most promising solutions. If the population prematurely converges, it is restarted.
- In each iteration, we are going to improve a promising solution by one local search, LS, method.
- The behavior of LS method strongly depends on the problem, so the proposal has a pool of LS methods, and it chooses each time which LS method is going to be applied.
- It stores when a solution was not previously improved by a LS method, to avoid applying it again, wasting resources.

We are going to test our proposal with the 100-digit challenge on single objective numerical optimization [7]. One interesting feature of this benchmark is that integrates difficult problems, and they are evaluated both the number of right
digits achieved and the number of evaluations required to obtain them. Thus, although it could be considered a benchmark focused in the accuracy, the convergence speed of the algorithm is also taken in account. Results obtained show that proposed algorithm is very competitive.

This paper has the following structure. In Section II, we are going to describe the proposal in detail. In Section III, the experimental environment is introduced, showing the results obtained by the proposal, and analyzing the results. Finally, in Section IV, the main conclusions and several future works are summarized.

II. PROPOSAL: MEMETIC WITH IMPROVED L-SHADE

In this paper, we propose a memetic algorithm with improved L-SHADE and a LS pool, MiLSHADE-LSP. The main features of this proposal are:

- The use of an advanced EA, the improved L-SHADE [6], responsible of the global exploration. This EA enforces global exploration at the beginning, focusing exploration around the most promising solutions during the run.
- A restart mechanism when the population prematurely converges.
- A LS Pool that contains different and complementary LS methods. The criterion used is to chose each time the algorithm whose last application had obtained a better relative improvement.
- A memory keeping, after the LS application, the solutions which were previous selected without a significant improvement. This memory is used to avoid applying a LS method many times without improvement.

In the following, we are going to describe the global scheme of the proposal, and later its components, the evolutionary algorithm responsible of the global exploration, the local search mechanism, and the restart mechanism.

A. Memetic framework

Algorithm 1 shows the global scheme of the proposal. It can be seen that in each iteration the population evolves as expected by the iLSHADE (lines 1-8).

When a new population is generated, one LS method from the LS Pool is selected, and the solution to apply it is chosen from the p best solutions. If there is no solution available, no LS is applied. The relative improvement is calculated, and it will be used for future selections (lines 17-24).

After the LS application, it is decided if the algorithm must be restarted, in that case the population is randomly reset, maintaining the current best solution in the next population (lines 25-27).

B. Improved L-SHADE

The exploratory algorithm is the improved L-SHADE [6] without any change. This algorithm has the following main features:

- A very advanced self-adaptation of the DE parameters, CR and F, allowing a good adaptation to each problem.

The only required parameters are related to the population size.

- The mutation operator uses previous solutions stored in an archive, increasing the diversity of the new solutions.
- The mutation operator is biased towards not always selecting the best solution. Instead, it randomly selects among the best p solutions.

In order to enforce more exploration at initial stages, and increasing selective pressure later, different parameters change during the run, according to the ratio of evaluations $r = \frac{\text{numevals}}{\text{maxevals}}$:

- The population size is linearly reduced during the run: \( \text{p} = \text{p}_{\text{initial}} - r \cdot (\text{p}_{\text{final}} - \text{p}_{\text{initial}}) \).

When the population is reduced, only the individuals with better fitness are maintained. This reduction produces a greater diversity in the initial population, to enforce exploration, leaving later in the population only the most promising solutions to focus the search around them.

- The parameter $p$ used to indicate the number of best solutions considered for the mutation also is reduced: \( p = \text{p}_{\text{initial}} + r \cdot (\text{p}_{\text{final}} - \text{p}_{\text{initial}}) \).

Thus, in initial stages, more promising solutions are used to guide the search, and later only a few of them are used,

Algorithm 1 MiLSHADE-LSP scheme

1: \text{p} = \text{p}_{\text{initial}}
2: \text{population} \leftarrow \text{random(dim, p)}
3: \text{pop_fitness} \leftarrow \text{fitness_function}(\text{p})
4: \text{current_best} = \min(\text{pop_fitness})
5: \text{current_best} = \text{LS(initial_solution)}
6: \text{while} \text{totalevals} < \text{maxevals} \text{do}
7: \text{population} \leftarrow \text{iLSHADE}\_\text{iteration}(\text{population}, \text{p})
8: \text{totalevals} \leftarrow \text{totalevals} + \text{p}
9: \text{pop_fitness} \leftarrow \text{fitness}(\text{p})
10: \text{current_best} = \min(\text{pop_fitness})
11: \text{Update parameter p}
12: \text{Update population size}
13: \text{if new p < previous p} \text{then}
14: \text{population} \leftarrow \text{best individuals from population}
15: \text{end if}
16: \text{Choose LS method from LS pool to apply}
17: \text{solution}_\text{LS} \leftarrow \text{One of p best solutions not local optima}
18: \text{previous} \leftarrow \text{fitness}(\text{solution}_\text{LS})
19: \text{solution}_\text{LS}, \text{evals}_\text{LS} \leftarrow \text{LS(solution, bounds)}
20: \text{totalevals} \leftarrow \text{totalevals} + \text{evals}_\text{LS}
21: \text{new_fit} \leftarrow \text{fitness(solution)}
22: \text{improvement} \leftarrow (\text{previous} - \text{new_fit})/\text{previous}
23: \text{Update improvement history of applying LS}
24: \text{if Must restart} \text{then}
25: \text{Restart maintaining current best individual}
26: \text{end if}
27: \text{end while}
the best ones, to enforce selective pressure.
- The allowed values for CR and F depend on the number of iterations, by the following constraints:

\[
\begin{align*}
\text{CR, F constraints} & : \\
& \begin{cases}
CR \geq 0.5, F \leq 0.7 & \text{if } r \leq 0.25 \\
CR \geq 0.25, F \leq 0.8 & \text{if } r \leq 0.5
\end{cases}
\]

Thus, in initial stages, at least the half of variables for solutions are mutated, and later, a smaller ratio of variables can be mutated.

For more details, you can consult [6].

C. Local Search Application

The proposal, for each iteration, after the steps of mutation, crossover, selection, and population reduction (see previous subsection), applies a LS method to one solution, to try to improve even more its fitness (lines 17-23 in the Algorithm 1).

This process is done using a LS pool composed by different LS methods and feedback information about previous applications. In this proposal, the LS methods used are the well-known Solis-Wets method [8], that applies small changes using a normal distribution with an adaptive variance, and the classic L-BFGS-B [9] that uses an approximation of the gradient to improve the search. However, the LS methods could be changed without problem, the LS Pool is flexible enough to change both the LS methods considered and its number.

In the following, we indicate the steps of the LS phase:
- a) Selection of the LS method: Initially, every one of the available LS methods are selected.
- b) Selection of the solution to improve: It is selected one of the \( p \) (parameter already used by iL-SHADE) best solutions in the population. The idea is to select one solution that was not previously applied by the same LS method obtaining an insufficient improvement (using a threshold value, \( \text{Threshold}_{LS} \)). For doing that, for each LS Method it is kept a list of solutions which were previous selected without obtained enough improvement in the LS application.
- c) Applying the LS method to the selected individual: In that step, if a solution was chosen, the selected LS method is applied to that individual during \( LS_{eval} \) evaluations. For Solis-Wets' algorithm, an initial step size of \( I_{step} \) is applied.

\[ d) \text{Calculating the improvement and updating the decision parameters:} \text{After each LS method, it is calculated } \text{Improvement}_{LS} \text{ obtained by that LS following Equation 1.} \text{Improvement}_{LS} \text{ is used to update the previous value for the executed LS method.} \]

If the relative improvement is lower than a \( \text{Threshold}_{LS} \) value, solutions before and after the LS method are kept into the LS pool to avoid selecting them again (for the same LS method).

D. Restart technique

The final steps in each iteration decide if the population should be restarted (lines 25-27 in Algorithm 1).

- It is considered that the population must be restarted when the difference in fitness between the best and the worst solution in the population is lower than a \( \text{Threshold}_{Restart} \). value.
- The restart implies that, at exception of the current best solution, all the other solutions in the populations are randomly generated, and their fitness calculated again. The restart takes in account the current population size (see subsection II-B).
- Additionally, not only the population is restarted, the \( LS_{Pool} \) parameters and memories are also updated (its history of last improvements, and the memory of solutions are also removed).

III. EXPERIMENTAL SECTION

A. 100-digit benchmark

The algorithm is tested using the benchmark, proposed by the 100-digit competition\(^1\) [7].

This benchmark is composed by 10 different functions, and the goal is to achieve the optimum value (1.0 in all functions) with at least 10 decimal places. The 10 decimal places for the 10 problems are which give it the name 100-digit challenge.

The source code of the benchmark is available in C++ and Matlab\(^2\), that we have used from Python with our own package, freely available by us including source code at https://github.com/dmolina/cec2019comp100digit.

The functions have several search domains (although the majority have range [-100, 100]), and different dimension (9 for \( f_1 \), 16 for \( f_2 \), 18 for \( f_3 \), and 10 for the others).

Experiments have been carried out with the following process:

- For each function, the algorithm is run 50 consecutive trials, each one with a different initial population randomly generated.
- It is counted the total number of correct digits of the 25 trials that give the best results.
- The number of evaluations for each level of digit accuracy is registered.

\( ^1\)http://cec2019.org/programs/competitions.html\(^1\)#ecce-06
• The score for that function is the average number of correct digits in the best 25 trials, and the total score is obtained.

At different of the majority of benchmarks, none stopping criterion is given, it is up to the researcher to decide when the algorithm should stop. That flexibility may suggest that MiLSHADE-LSP, that uses the ratio of evaluations inside its strategy, could not be a good option when an arbitrary maxevals is used. However, as we will see later, it is not true, because we obtain very good results.

In the comparisons carried out, we analyze both the final scores (total and by functions) and the required number of evaluations to reach the number of decimal digits.

You can consult [7] to get more details about the benchmark and the experimental conditions.

B. Tuning of Parameters

Although it is acceptable to tune 2 parameter(s) independently for each problem, in MiLSHADE-LSP we are going to analyze results with and without the tuning. This double analysis is carried out because we want also to observe the behavior of our proposal considering each problem as a black-box one. The only information used is the dimension value (to define the initial parameter values), which is a parameter known for every real-world problem, and the ratio of evaluations (used to adapt both the population size and the p parameter used in the mutation method).

![Table I: Fixed parameter values](image)

<table>
<thead>
<tr>
<th>Component</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>iL-SHADE</td>
<td>Initial popsize</td>
<td>12 ∙ dim</td>
</tr>
<tr>
<td></td>
<td>Final popsize</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>from 0.2 to 0.1</td>
</tr>
<tr>
<td>LS Pool</td>
<td>LS Methods</td>
<td>Solis Wets, L-BFGS-B</td>
</tr>
<tr>
<td></td>
<td>LS eval</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>ThresholdLS</td>
<td>10−10</td>
</tr>
<tr>
<td>Restart</td>
<td>ThresholdRestart</td>
<td>10−8</td>
</tr>
</tbody>
</table>

**TABLE II: Default values for parameters used for tune**

The proposal has the fixed parameter values indicated in Table I. The majority of parameters values used are the recommended values by the authors of iL-SHADE [6], the only additional parameters are \( LS_{eval} \) and \( I_{Step} \). In Table II, there are the default parameter values for parameters used in the tuned version. The maximum evaluation number could be a high number, but the majority of optima are obtained in a small percentage of the evaluations.

C. Results

In this section, we analyze the results obtained by MiLSHADE-LSP. First, we are going to observe if the changes made to iL-SHADE (mainly the introduction of the LS improvement) improve the results. Later, we are going to analyze the influence of the maxevals parameter. Finally, we are going to analyze the results obtained with and without tuned parameters.

a) Improvement from non-memetic algorithm (iL-SHADE): In order to study the convenience of the memetic version of the iL-SHADE, first we are going to show the results, both of original iL-SHADE and the memetic version proposed in this work. The values of all shared parameters are the same.

Tables VIII and IX show the obtained results, as it was asked for the competition, for original iL-SHADE and MiLSHADE-LSP, respectively. These tables show, from the best 25 of the 50 runs, the number of times that each number of correct digits is achieved.

![Table III: Score in original iL-SHADE vs non-tuned MiLSHADE-LSP](image)

<table>
<thead>
<tr>
<th>Function</th>
<th>iL-SHADE</th>
<th>MiLSHADE-LSP</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_1 )</td>
<td>2894</td>
<td>307</td>
<td>10%</td>
</tr>
<tr>
<td>( f_2 )</td>
<td>173041</td>
<td>287444</td>
<td>166%</td>
</tr>
<tr>
<td>( f_3 )</td>
<td>1757291</td>
<td>72308</td>
<td>4%</td>
</tr>
<tr>
<td>( f_5 )</td>
<td>68928</td>
<td>9819</td>
<td>14%</td>
</tr>
<tr>
<td>Total</td>
<td>2002154</td>
<td>369878</td>
<td>18%</td>
</tr>
</tbody>
</table>

**TABLE IV: Ratio of evaluations iL-SHADE vs MiLSHADE-LSP in optimized functions**

b) Influence of the maxevals parameter: MiLSHADE-LSP has a series of parameters whose values strongly depend on the current ratio of evaluations done. Thus, in our proposal, changing maxevals (the maximum evaluation number) not
only involves a greater number of evaluations, allowing to increase the accuracy of solutions, but also it could affect the search balance. A higher maxevals value gives more resources to achieve more accuracy in the search, but, at the same time, encourages a greater global exploration, that could be excessive to obtain good results.

Table V shows the scores by function and total for different values of maxevals \((10 \cdot 10^4 \cdot \text{dim}, 25 \cdot 10^4 \cdot \text{dim}, \text{and } 50 \cdot 10^4 \cdot \text{dim})\). It can be observed that increasing the number of evaluations produces a better score, but the improvement is lower than expected, only two more points with more than double of evaluations. However, when maxevals increases even more, results are worse, maybe due to a worse trade-off between exploration and exploitation. Thus, the fact that MiLSHADE-LSP has adaptive variables related to the ratio of evaluations, becomes our algorithm very sensitive to the maxevals parameter, as expected.

<table>
<thead>
<tr>
<th>Function</th>
<th>Maxevals</th>
<th>Maxevals</th>
<th>(In %)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10 \cdot 10^4 \cdot \text{dim}</td>
<td>25 \cdot 10^4 \cdot \text{dim}</td>
<td>50 \cdot 10^4 \cdot \text{dim}</td>
</tr>
<tr>
<td>(f_1)</td>
<td>307</td>
<td>307</td>
<td>100%</td>
</tr>
<tr>
<td>(f_2)</td>
<td>287444</td>
<td>276423</td>
<td>103%</td>
</tr>
<tr>
<td>(f_3)</td>
<td>72308</td>
<td>65767</td>
<td>109%</td>
</tr>
<tr>
<td>(f_4)</td>
<td>988132</td>
<td>2285193</td>
<td>43%</td>
</tr>
<tr>
<td>(f_5)</td>
<td>9812</td>
<td>8313</td>
<td>118%</td>
</tr>
<tr>
<td>(f_6)</td>
<td>125047</td>
<td>183564</td>
<td>64%</td>
</tr>
<tr>
<td>(f_{10})</td>
<td>926102</td>
<td>2139961</td>
<td>43%</td>
</tr>
<tr>
<td>Total</td>
<td>2140640</td>
<td>6804199</td>
<td>35%</td>
</tr>
</tbody>
</table>

Table VI: Ratio of evaluations in optimized runs functions for different maxevals

In order to study the speed of convergence, Table VI shows the number of evaluations when the optima are calculated with different maxevals. The third column is the relative difference in evaluations to achieve the optima between lower maximum evaluation number \((10 \cdot 10^4 \cdot \text{dim})\) and with the greater evaluation number \((25 \cdot 10^4 \cdot \text{dim})\). A value lower than 100% means that the optima can be obtained with less evaluations when maxevals is lower. On the contrary, a value higher than 100% means that it is more efficient a greater maxevals. While in functions \(f_1, f_2, f_3\) and \(f_5\), the number of evaluations are lightly reduced when maxevals increases, in functions \(f_4, f_6, \text{and } f_{10}\), the number of evaluations greatly increases with a higher maxevals. Although a higher maxevals increases lightly the score, makes the convergence slower. Thus, maxevals = \(10 \cdot 10^4 \cdot \text{dim}\) could be considered the most adequate value if speed is important.

\(c)\) Tuned vs non-tuned MiLSHADE-LSP: The competition 100-digit allows researcher to tune 2 parameters independently for each problem. Table VII shows the used tuned parameters in our case, maxevals and \(I_{\text{step}}\). With this tuning we obtain the results shown in Table X. It can be seen that tuned version achieves a total score of 60.72, more than 7 points more than the non-tuned version. In particular, it is obtained the optima in all cases in one more function, \(f_{10}\). Also, it is obtained a significant improvement in functions \(f_4, f_6\) (it obtains the optimum 13 times instead of 5 times), and \(f_5\) (it obtains two right digits 7 times instead of only once).

To summarize, the LS pool improves both the results and the performance. Moreover, an adequate maxevals value is needed to obtained a good trade-off between exploration and exploitation. Finally, tuning maxevals and \(I_{\text{step}}\) parameters by function improves significantly the accuracy. In general, the proposed algorithm, MiLSHADE-LSP, obtains very good results in the 100-digit challenge.

### IV. Conclusions

In this paper, we have proposed a new optimization algorithm, MiLSHADE-LSP, a memetic algorithm which combines an adaptive differential evolution algorithm that modifies its behavior during the run, with a LS Pool that continuously improves the results. MiLSHADE-LSP uses a pool of two very different LS methods, the L-BFGS-B and the Solis-Wets algorithms. A memory is used to adapt which LS method is selected in each iteration. The DE algorithm, iL-SHADE, is continuously applied to the same population among iterations, to introduce the performance obtained by the LS to guide the search. Also, it has a restart mechanism to explore new areas when the search is stuck, restarting the population, and resetting the LS Pool parameters.

In the experimental section, we have tested and analyzed MiLSHADE-LSP using the proposed benchmark for the competition 100-digit challenge on Single Objective Numerical Optimization. First, we have compared the contribution of the memetic proposal over the original iL-SHADE, obtaining that it contributes to improve not only the results, but also...
its performance. Later, we have analyzed the behavior of the proposal when the maximum evaluation number is increased. We have obtained that, due to the adaptive parameters, a greater maxevals could getting worse results. Also, even when a greater maxevals increased lightly the results in score, the speed to reach the optimum was reduced in several functions. Finally, as it was allowed in the competition, we have tuned two parameters independently by each problem, maxevals and the initial step size, $I_{\text{Step}}$. We have observed that tuned-version got better results, obtaining a total score of more than 60. Thus, MiLSHADE-LSP have achieved very good results in the benchmark for the competition.

As a future work, we are going to try to adapt the $I_{\text{Step}}$, and improve other components, like the restart mechanism, to improve even more the results.

ACKNOWLEDGMENTS

This work was supported by grants from the Spanish Ministry of Science (TIN2014-57481-C2-2-R, TIN2016-8113-R, TIN2017-83132-C2-2-R and TIN2017-89517-P).

REFERENCES