Review response for Benchmarking SHADE algorithm enhanced with model based optimization on the BBOB noiseless testbed

Thank you for accepting our paper. Below are details on how the requests of reviewers have been included in the final version.

Reviewer 1.

Comment:

1. The description of the algorithm is a bit short and does not allow to reimplement it. Please give more details about the algorithm:
    a. what happens if not enough or too many solutions have been evaluated for the system of equalities to have a unique solution in the model fitting?

Response:
We have presented this cases in the section 3.2.1 added in the revised version of the paper. For now we detect convexity only for simple model (and in case of concave shape of square function along one of the dimensions we select an appropriate boundary point). For the full square model there is simply a possibility of sampling poor function value and "loosing" this sample. However, please note that even adding such a behavior of the model achieved an improvement.

    b. what exactly do the procedures $SelectIndividualsForModelApplication$, $GetSampleAsModelFunctionStationaryPoint$, and $GetSampleAsInSHADEAndStoreShadeFactors$ do?
    c. a few more details about the concrete stopping criteria used in the experiments will be useful as well. No improvement evaluations, Values convergence, and Locations convergence values alone are not sufficient to be able to reproduce/reimplement in my viewpoint.
    d. where in Algorithm 2 is the crossover and mutation from DE?

Response:

We have enhanced section 3.3 Proposed algorithm with the method description and restart criteria definition. Crossover and mutation from DE are performed within $GetSampleAsInSHADEAndStoreShadeFactors$ function.

Comment:

2. Both SHADE-LM and SHADE-LM-POP4-to-10 appear in the figures with the same name. Renaming the folders before the postprocessing allows to change that. Please consider.
    Response:
    We have renamed the algorithms on the plots to SL-10 and SL-4-10, but kept the old name in the body of the paper as it was more meaningful.

Comment:

3. Which of the four available bbob data sets by Tanabe and Fukunaga did you use in your paper? I couldn't find it in the text.
    Response:
    It is R-SHADE-10e5, we have update the experiment section with that information.

Comment: Minor:

- line 3 of algorithm: add blanks before and after "to"
- please correct "as given in the 1" on page 2
- legend and plots in Figure 1 are not consistent in terms of markers and colors for the algorithms
- twice the same words on page 2: "… gradually increases its population from $4D$to $10D$ gradually increasing it…"

Response:

Thank you for the above comments, we have corrected our mistakes.

Reviewer 2.

Comment:

1. I feel somewhat confusing that a hybrid of DE and quadratic modeling uses some kind of PSO framework. If this is not important for the method itself, I would remove it from the description.

Response:

It was an important step in the design process of the algorithm. The whole idea of the seamless/transparent hybridization came from swarm-like perception of the population based methods: that each particle can have its own sampling logic, and as long as it has the same "interface" the other individuals do not need to know that logic.

We have enhanced that motivation in the Introduction section of the article and better explained this design concept in the GAPSO Framework section.

Comment:

2. Since the paper proposes a novel algorithm which is not described elsewhere, I would prefer much more detailed description of the algorithm. I do not have the feeling that I would be able to reimplement the algorithm just on the basis of the article.

Response:

We have enhanced section 3.3 Proposed algorithm with the method description and restart criteria definition.

Comment:

3. It would be also nice to discuss various design choices made by the authors. E.g., couldn't it happen that the model learnt from data is not convex? What is the point of adding a stationary point of the model in that case? Or do you detect the situation somehow?

Response:

We have presented this cases in the section 3.2.1 added in the revised version of the paper. For now we detect convexity only for simple model (and in case of concave shape of square function along one of the dimensions we select an appropriate boundary point). For the full square model there is simply a possibility of sampling poor function value and "loosing" this sample. However, please note that even adding such a behavior of the model achieved an improvement.

Comment:

4. How do the functions SelectIndividualsForModelApplication and GetSampleAsModelFunctionStationaryPoint actually work? How do you select the points used to train the quadratic model?

Response:

SelectIndividualsForModelApplication chooses which individuals of the current population will be sampled as model optimum, while the rest will be sampled as in SHADE algorithm. The models are trained/fitted on all of the samples forming the current population, hence the models may vary even within the same iteration, as the population is constantly updated if a better replacement for an individual is found. Section 3.3. Proposed algorithm has been enhanced with more detailed explanation of the algorithm.

Comment:

5. How do you decide whether you train model (3) or model (4)? What is the threshold on the number of available data points?

Response:

The threshold comes directly from the minimum number of samples necessary for a given model. 2D+1 for simple model, (D^2 + 3D)/2 + 1 for full model. The additional explanation has been given in section 3.2. Model based optimizers.

Comment: Typos:

- In sec. 4, first sentence: "… given in the 1."
- In sec. 6, in "10^4 times D" and "10^6 times D", you probably forgot the backslash before "times": --> "10^4\times D".

Response:

Thank you for the above comments, we have corrected our mistakes.

Reviewer 3.

Comment:

1. While model building is a valid approach to try to improve algorithms, the paper gives relatively little insight into the motivation of the specific combination of model building and differential evolution and on specific decisions taken.

Response:

The algorithm idea was of the seamless/transparent hybridization, and it came from swarm-like perception of the population based methods: that each particle can have its own sampling logic, and as long as it has the same "interface" the other individuals do not need to know that logic.

We have enhanced that motivation in the Introduction section of the article and better explained this design concept in the GAPSO Framework section.

Comment:

2. Also very few conclusions from the observations and results are drawn.

Response:
We have provided additional conclusions for excellent performance on f1 and f5 and weaker performance on f21 – f24 group of functions.

Comment: An article seems missing in the following phrases:

- "Single iteration of"
- "Model is fitted on the samples"
- "the coefficients of simple square model"
- "taken from square function model optimum"

Response:

We have enhanced the whole 3.2. section discussing the utilization of models and added a 3.2.1 subsection discussing handling of the special cases.

Comment: An article seems missing in the following phrases:

- "with the settings as given in the 1" should be "in Table 1"
- "maximum budget equal to 10^6D) function evaluations" the closing brace should be removed
- "$timesD$" should read "x D" (twice)

Response:

Thank you for the above comments, we have corrected our mistakes.