

Structural Bias in Differential Evolution: a preliminary study

Fabio Caraffini^{1,a)} and Anna V. Kononova^{2,3,b)}

¹*Institute of Artificial Intelligence, School of Computer Science and Informatics,
De Montfort University, Leicester LE1 9BH, UK*

²*Department of Computer Science, Heriot-Watt University, Edinburgh EH14 4AS, UK*

³*School of Communication, Media and Information Technology, Hogeschool Rotterdam, The Netherlands*

^{a)}Corresponding author: fabio.caraffini@dmu.ac.uk

^{b)}anna.kononova@gmail.com

Abstract. This paper extends the study of structural bias in popular metaheuristic global optimisation methods. Previously, it has been shown that both Genetic Algorithms and Particle Swarm Optimisation suffer from such bias. This means that difficulties already posed for a structurally biased algorithm by the fitness landscape itself are further unnecessarily exacerbated by the unexpected oversampling of some regions of the search space and avoidance of the others, to potential great detriment of the overall optimisation performance. Such bias is inherent in the core design of the algorithm. After careful examination, the authors conclude that some variants of Differential Evolution are not free of the structural bias. However, investigation suggests that the mechanisms of the formation of structural bias in Differential Evolution is different and can be balanced through a more careful design.

INTRODUCTION

Many general purpose metaheuristics for global optimisation are nowadays available to the practitioners. However, their algorithmic design process has been done in a partially blind manner, largely due to a major lack of knowledge on the actual dynamics inside the algorithm, which makes it very difficult to pick and properly tune the most appropriate problem-specific optimiser. Consequently, the “off-the-shelf” algorithms often fail at producing results of satisfactory quality in most real-world scenarios. To shed light on this problem, researchers have started studying algorithmic behaviour of the popular metaheuristics and have found biases. see e.g [1, 2], preventing algorithms from exploring all parts of the search space to equal extent.

Recently, the research in [3] has unveiled the presence of the so-called structural bias in population-based metaheuristics and has contributed both to the theoretical and empirical understanding of this phenomenon. It has been theoretically shown, under some hypotheses, that the structural bias correlates with the population size and, as common denominator for most population-based algorithms for real-valued global optimisation, it is amplified proportionally to the population size. Such conclusion have also been empirically confirmed for two instances of Particle Swarm Optimisation (PSO) and Genetic Algorithm (GA), as in [4] and [5], thus questioning common belief that a large population size is beneficial for a number of reasons, e.g. a summary is given in [6], and is a key to tackling large-scale problems efficiently [7]. Moreover, a graphical approach to visualising the bias, as a non-uniform clustering of the population over time, has been developed. This visual approach has turned out to be extremely informative and highly recommendable, to both users and designers, prior to finalising the parametric and design configuration of any population-based optimisation algorithm.

OBJECTIVES

Differential Evolution (DE) is a powerful yet simple metaheuristic for global real-valued optimisation which only requires three parameters to function efficiently [8]: the scale factor $F \in [0, 2]$, the crossover ratio $CR \in [0, 1]$ and the

population size NP . However, its performance heavily depends on the choice of these values. While research literature is rich of studies on setting F and CR , see e.g. [7, 9, 10, 11], significantly less information is available on setting the value of NP . A promising way to fill this gap in knowledge is presented in [3], which provides the practitioners with a method to evaluate the impact of the population size in terms of the structural bias. To the best of our knowledge, this has never been done for the DE.

In this preliminary study, the authors focus their attention on 4 widely used DE schemes. By using rigorous DE nomenclature, they can be referred to as “DE/rand/1/bin”, “DE/rand/1/exp”, “DE/current-to-best/1/bin” and “DE/current-to-best/1/exp”. For the sake of clarity, let us remind the reader that “rand” indicates the original random mutation strategy in [8], while “bin” and “exp” are the binomial and the exponential crossover, respectively. The second mutation, “current-to-best” has been introduced later on as an improvement upon the original mutation. More details can be found in [12].

The objectives of this paper are to search for structural biases in the aforementioned algorithms, with the goal of having a better undersigning of the extent of structural bias present in the DE and investigate how it manifests itself under different combinations of mutation and crossover operators. It is our intention to extend this work, in the near future, to other DE schemes and study their structural bias as a function of different parameter settings.

METHODS

An algorithm is said to possess structural bias when it is unable to explore all areas of the search space to the equal extent, regardless of the fitness function [3]. The best testbed to identify such bias, in terms of its effects on the distribution of the final best solutions over multiple runs, is the function

$$f_0 : [0, 1]^n \subset \mathbb{R}^n \rightarrow [0, 1], \quad f_0(x) = \text{Uniform}(0, 1)$$

for which, as rigorously explained in [3], an ideal unbiased algorithm should return a uniform distribution of the best final solutions, over a series of independent runs. Thus, if displayed by means of “parallel coordinates”, i.e. coordinates of a vector in an n -dimensional space are marked in n equally spaced parallel lines (see [3, 13] for better explanation), the resulting figure would show points homogeneously filling the entire interval in each dimension. Conversely, in the presence of a strong structural bias, clusters will appear as best solutions would tend to accumulate in one or more segments of each parallel line.

For optimisation problems in bounded domains, it is common to apply strategies to correct the solutions generated outside the search domain during the optimisation process. To demonstrate clearly that structural bias can manifest itself regardless of the correction strategy, two different approaches were used here. First, we considered a simple saturation scheme, i.e. solutions outside the optimisation domain are clipped to the boundaries. Second, the geometrical transformation introduced in [12], i.e. toroidal correction strategy. Their graphical representation is shown in Fig. 1.

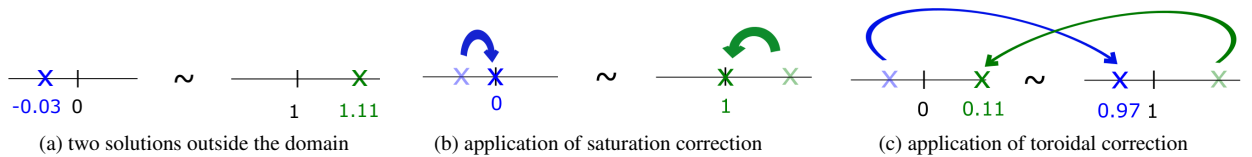


FIGURE 1. Schematic explanation of correction strategies for domain $[0, 1]$

The four preselected DE schemes (DE/rand/1/bin, DE/rand/1/exp, DE/current-to-best/1/bin, and DE/current-to-best/1/exp) were run on f_0 for the $n = 30$ dimensional case with population sizes of $NP \in \{5, 20, 100\}$ and multiple combinations of the F and CR parameter values, i.e. $F \in \{0.05, 0.2, 0.4, 0.7, 0.9\}$ and $CR \in \{0.05, 0.4, 0.7, 0.9, 0.99\}$, in order to find which F - CR pair should be used to produce the reliable results. The values of $F = 0.1$ and $CR = 0.2$ were then chosen as they implied the least number of corrections required during the optimisation process of f_0 . It has to be mentioned that keeping the number of corrections low is a key in this study since correction strategies can add undesired side effects and prevent the true understanding of the nature of the structural bias.

Experiments for all DE variants considered for this study have been executed with the computational budget of $10000n = 300000$ function calls repeated over 50 runs. Three increasing population sizes has been considered,

i.e. $NP \in \{5, 20, 100\}$ while F , CR and dimensionality n have been kept constant (0.1, 0.2 and 30, respectively). Moreover, Results shown below are cover “saturation” and “toroidal” options correction strategies. The distribution of the positions of the best solution from each run has been investigated.

RESULTS

Surprisingly, from the four DE variants considered in this study only DE/current-to-rand/1/bin displayed a clear structural bias. To counterpose the results, Fig. 2 show positions of final best solutions from series of 50 runs of DE/current-to-rand/1/bin and DE/rand/1/bin, with saturation and toroidal correction respectively, all with NP set to 100. Figs. 2(a), 2(b) clearly show underexplored areas in the corners of the search domain. Meanwhile, Figs. 2(c), 2(d) contain much “denser” distribution of points covering the intervals in all dimensions, thus demonstrating less structural bias. Results for the remaining configurations of DE perturbation schemes, population sizes, and correction strategies follow similar patterns regardless of the population size. Due to the limited space available in this publication, they have been made available online [14].

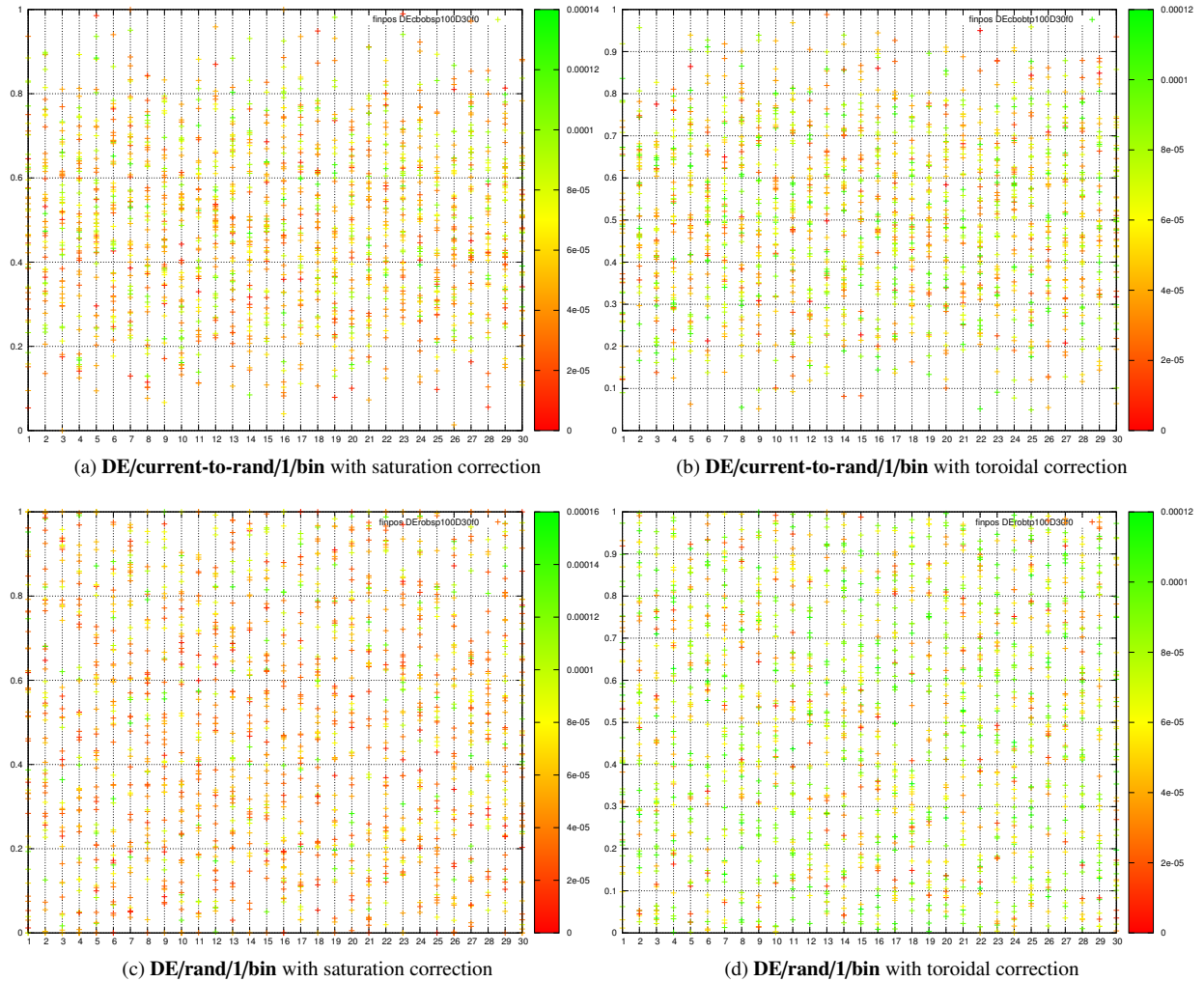


FIGURE 2. Positions in parallel coordinates of final best solutions on f_0 in 30 dimensions from variants of DE with different correction strategies with population size $NP = 100$, with 50 runs for each. Horizontal axis shows dimension, vertical axis shows position of the best solution in this dimension, fitness value is shown on colour scale. Clear structural bias is present for DE/current-to-rand/1/bin, meanwhile no structural bias is evident for DE/rand/1/bin.

From the results above, it becomes evident that DE/current-to-best/1/bin tends to favour specific regions - it performs better on problems whose minima are located in the centre of the search space regardless of the population size and correction strategy. This suggests that the triggering mechanism for structural bias in DE differs to the one found in [3], i.e. the larger population size, for GA and PSO. Furthermore, it must be noted that DE/current-to-rand/1/exp unexpectedly appears to be less biased than DE/current-to-rand/1/exp. Our interpretation of this discrepancy is that a different number of exchanged components, i.e. higher in binomial crossover and lower in exponential crossover - also depending on n in the exponential scheme [15] - due to the use of a fixed CR value might be responsible for such behaviour. Further experimentations are currently being carried out and will be reported later.

CONCLUSIONS AND THE WAY FORWARD

The classic “rand/1” Differential Evolution framework turns out to be a well designed and robust stochastic optimisation method as it does not present a clear structural bias for any choice of crossover and correction operators considered. The same cannot be said for the “current-to-best/1” scheme, for which clear structural bias arises if the binomial crossover is employed. Conversely, the strength of the structural bias appears to be mitigated by the use of exponential crossover.

It is current opinion of the authors that presence of strong structural bias can be viewed as consequence of disbalance in exploration and exploitation. Such algorithms with disrupted balance are no longer capable to carry out the search at both the global and local levels [16]. Similar considerations are known to be investigated in divide-the-best methods and adaptive diagonal partition strategy by Sergeyev for Lipschitz global optimisation¹ [17, 18].

Results presented in this paper have given rise to several new research questions. The authors intend on extending this investigation in the near future. More popular DE schemes are already under test, with a wider range of correction strategies (including penalty functions and other methods), and it is our intention to study the impact of the F and CR parameters.

REFERENCES

- [1] W. Spears, D. Green, and S. Diana F., *International Journal of Swarm Intelligence Research (IJSIR)* **1**, 34–5711 (2010).
- [2] M. Davarynejad, J. van den Berg, and J. Rezaei, *Information Sciences* **278**, 802 – 821 (2014).
- [3] A. V. Kononova, D. W. Corne, P. D. Wilde, V. Shneer, and F. Caraffini, *Information Sciences* **298**, 468–490 (2015).
- [4] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Neural Networks, 1995. Proceedings., IEEE International Conference on*, Vol. 4 (1995), pp. 1942–1948 vol.4.
- [5] M. D. Vose, *The simple genetic algorithm: foundations and theory*, Vol. 12 (MIT press, 1999).
- [6] A. Prügel-Bennett, *IEEE Transactions on Evolutionary Computation* **14**, 500–517 (2010).
- [7] J. Liu and J. Lampinen, *Soft Computing* **9**, 448–462 (2005).
- [8] R. Storn and K. Price, “Differential evolution - a simple and efficient adaptive scheme for global optimization over continuous spaces,” *Tech. Rep. TR-95-012 (ICSI)*, (1995).
- [9] D. Zaharie, “Critical values for control parameters of differential evolution algorithm,” in *Proceedings MENDEL* (2002), pp. 62–67.
- [10] G. Iacca, F. Caraffini, and F. Neri, “Continuous parameter pools in ensemble differential evolution,” in *2015 IEEE Symposium Series on Computational Intelligence* (2015), pp. 1529–1536.
- [11] J. Cheng, G. Zhang, F. Caraffini, and F. Neri, *Integrated Computer-Aided Engineering* **22**, 103–107 (2015).
- [12] K. V. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization* (Springer, 2005).
- [13] A. Inselberg, *The Visual Computer* **1**, 69–91Aug (1985).
- [14] F. Caraffini and A. V. Kononova, *Structural bias in differential evolution: a preliminary study - extended results*, www.tech.dmu.ac.uk/~fcaraf00/BIAS/GOW18_BIAS_DE_RESULTS.pdf (2018).
- [15] G. Iacca, F. Caraffini, and F. Neri, *Journal of Computer Science and technology* **27**, 1056–1076 (2012).
- [16] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, *Journal of Optimization Theory and Applications* **79**, 157–181 (1993).
- [17] Y. D. Sergeyev, *Optimization* **44**, 303–325 (1998).
- [18] D. E. Kvasov and Y. D. Sergeyev, *Automation and Remote Control* **74**, 1435–1448 (2013).

¹We would like to thank the anonymous Reviewer 2 for hinting towards this direction.