A Surrogate Model Assisted (1+1)-ES with Increased Exploitation of the Model

Jingyun Yang and Dirk V. Arnold
Faculty of Computer Science, Dalhousie University
Halifax, Nova Scotia, Canada B3H 4R2
Jingyun.Yang@dal.ca|dirk@cs.dal.ca

ABSTRACT

In black-box optimization, surrogate models can be employed in two fundamentally different ways. They can either be used to provide inexpensive but inaccurate assessments of the quality of candidate solutions generated by the black-box optimization algorithm; or they can be used in the process of generating those candidate solutions themselves. The latter approach more fully exploits the model, but may be more prone to fall prey to the model's inaccuracy. This paper examines the effect of the degree of exploitation of the surrogate model in the context of a simple $(1 + 1)$-ES. We derive the potential gain from more fully exploiting surrogate models using a simple modelling technique for the error resulting from the use of surrogate models. We then observe the effects of increased exploitation in an evolution strategy employing local Gaussian process surrogate models. We show that while the benefits predicted under the simple error model cannot be fully realized, the potential gain from increased exploitation of the surrogate model can nonetheless be significant.

CCS CONCEPTS

•Mathematics of computing → Bio-inspired optimization;
•Computing methodologies → Continuous space search;

KEYWORDS

Stochastic black-box optimization; evolution strategy; surrogate modelling; Gaussian process

ACM Reference format:
Jingyun Yang and Dirk V. Arnold. 2019. A Surrogate Model Assisted (1+1)-ES with Increased Exploitation of the Model. In *Proceedings of GECCO '19, Prague, Czech Republic, July 13-17, 2019,* 3 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Surrogate modelling techniques are commonly used when solving black-box optimization problems where the evaluation of the objective function is expensive. Models are built based on information gained through evaluation of the objective function in previous iterations. These models can then be used as inaccurate but inexpensive surrogates for the true objective function. Jin [4] and Loshchilov [8] provide surveys covering the use of surrogate modelling techniques in evolutionary computation.

Black-box optimization algorithms can employ surrogate models in fundamentally different ways. One approach is to leave the generation of candidate solutions under the control of the black-box optimization algorithm. The surrogate model is used in place of the true objective function in order to avoid the costly evaluation of likely poor candidate solutions. In the context of evolutionary algorithms, this strategy is sometimes referred to as preselection. An example of an algorithm that implements this approach is the Local Meta-Model Covariance Matrix Adaptation Evolution Strategy (lmm-CMA-ES) by Kern et al. [6].

A second approach is to involve the surrogate model in the generation of candidate solutions. That is, some approach of finding near optimal solutions of the surrogate model is used in order to determine candidate solutions to be evaluated by the true objective function. This approach is employed for example in the Gaussian Process Optimization Procedure (GPOP) by Büche et al. [3].

Arguably, the latter approach more fully exploits the surrogate models.

- There are at least two approaches to using surrogate models: either the models are used to replace evaluations using the true objective function (Kern, and, Koumoutsakos), or the models are used to recommend the next point to sample (Büche et al.).
- If Jingyun's observations are clear, then the problem is that the simple model for surrogate models loses validity as λ increases. Observations aren't random. The algorithm is mislead systematically. Compare the number of successes expected based on the analysis with those actually observed.
- Compared to the PPSN paper, focus here is on a more focused selection of directions. There, random directions are evaluated using the surrogate model. Here, recombination and selection are used to potentially generate better directions before a candidate solution is evaluated using the objective function.
- A nice consequence of analyzing algorithms on the quadratic sphere is that a well-studied baseline exists; Rechenberg
- How much of a speed-up is possible?
- The inclusion or covariance matrix adaptation is orthogonal to what is studied here.
- [5, 7]

The remainder of this paper is organized as follows. Section 2 reviews related work and introduces a surrogate model assisted $(1+$

GECCO '19, Prague, Czech Republic

978-x-xxxx-xxxx-x/YY/MM…$15.00
DOI: 10.1145/nnnnnnn.nnnnnnn

Generate step vector $\mathbf{z} \in \mathbb{R}^n$ and let $\mathbf{y} = \mathbf{x} + \sigma\mathbf{z}$.
Evaluate $\mathbf{y}$ using the surrogate model, yielding $f_\epsilon(\mathbf{y})$.
if $f_\epsilon(\mathbf{y}) \geq f(\mathbf{x})$ **then**
    Let $\sigma \leftarrow \sigma\,\mathrm{e}^{-c_1/D}$.
else
    Evaluate $\mathbf{y}$ using the objective function, yielding $f(\mathbf{y})$.
    Update the surrogate model.
    if $f(\mathbf{y}) \geq f(\mathbf{x})$ **then**
        Let $\sigma \leftarrow \sigma\,\mathrm{e}^{-c_2/D}$.
    else
        Let $\mathbf{x} \leftarrow \mathbf{y}$ and $\sigma \leftarrow \sigma\,\mathrm{e}^{c_3/D}$.
    end if
end if

Figure 1: Single iteration of the surrogate model assisted $(1+1)$**-ES by Kayhani and Arnold [5].**

1)-ES with variable exploitation of the model. Section 3 studies the performance of that algorithm on spherically symmetric objective functions by assuming a simple model for the error resulting from the use of surrogate models. Section 4 employs Gaussian process surrogate models and applies the algorithm to a wider range of test functions. Section 5 concludes with a brief summary and proposed future work.

2 BACKGROUND

- [3, 6, 9]

3 ANALYSIS

This section generalizes the analysis of Kayhani and Arnold [5] by considering step vectors from distributions other than a standard Gaussian distribution. Specifically, we consider the case that the step generated in Line

- derive

$$p_{\text{eval}} = \int_{-\infty}^{\infty} p_1(z)\Phi\left(\frac{\sigma^* z - {\sigma^*}^2/2\mu}{\sigma_\epsilon^*}\right)\mathrm{d}z$$

and

$$\mathrm{E}[\varDelta] = \int_{-\sigma^*/2\mu}^{\infty}\left(\sigma^* z - \frac{{\sigma^*}^2}{2\mu}\right)p_1(z)\Phi\left(\frac{\sigma^* z - {\sigma^*}^2/2\mu}{\sigma_\epsilon^*}\right)\mathrm{d}z$$

- give the exact distribution of z_1; then provide moments, referring to [1, 2]
- Gram-Charlier expansion:

$$p_1(z) = \frac{1}{\sqrt{2\pi\kappa_2}}\exp\left(-\frac{(z-\kappa_1)^2}{2\kappa_2}\right)\left[1 + \frac{\gamma_1}{3!}\mathrm{He}_3\left(\frac{z-\kappa_1}{\sqrt{\kappa_2}}\right) + \ldots\right]$$

where $\gamma_1 = \kappa_3/\kappa_2^{3/2}, \ldots$, and $\mathrm{He}_k(\cdot)$ is the kth Hermite polynomial
- for $\mu = \lambda = 1$, Arash's results are recovered
- use heatmaps to illustrate expected quality gain as a function of evaluation and false positive rates

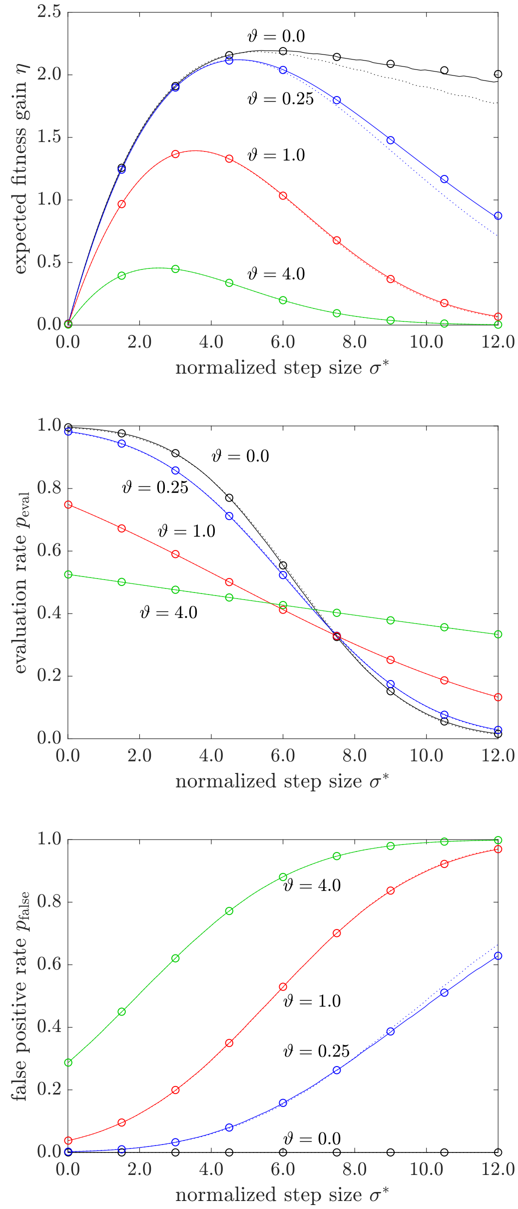

Figure 2: Expected fitness gain, evaluation rate, and false positive rate of the surrogate model assisted (1+1)-ES with $(3/3, 10)$ **preselection plotted against the normalized step size. The dashed and solid lines represent the results from Eqs. (??), (??), and (??) considering approximations with terms up to the quadratic and cubic ones in the Gra-Charlier expansion, respectively**

ACKNOWLEDGEMENTS

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

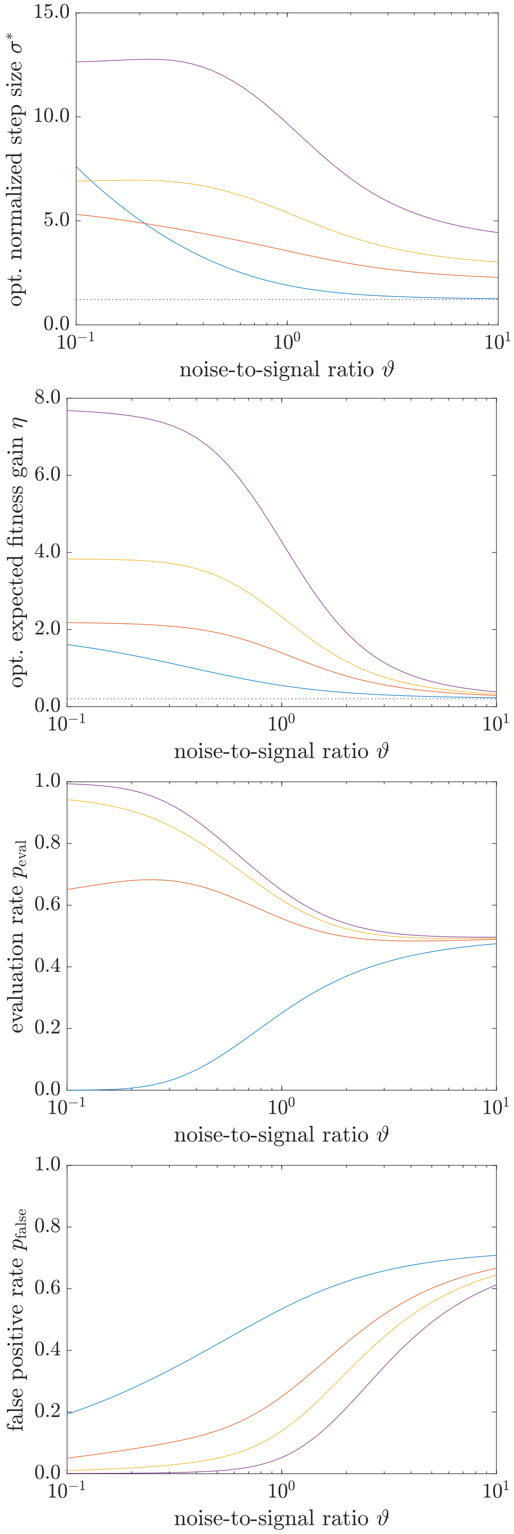

Figure 3

REFERENCES

[1] D. V. Arnold. 2002. *Noisy Optimization with Evolution Strategies*. Kluwer Academic Publishers.

[2] D. V. Arnold and H.-G. Beyer. 2005. Expected sample moments of concomitants of selected order statistics. *Statistics and Computing* 15, 3 (2005), 241–250.

[3] D. Büche, N. N. Schraudolph, and P. Koumoutsakos. 2005. Accelerating evolutionary algorithms with Gaussian process fitness function models. *IEEE Transactions on Systems, Man, and Cybernetics, Part C* 35, 2 (2005), 183–194.

[4] Y. Jin. 2011. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation* 1, 2 (2011), 61–70.

[5] A. Kayhani and D. V. Arnold. 2018. Design of a surrogate model assisted (1 + 1)-ES. In *Parallel Problem Solving from Nature — PPSN XV*, A. Auger and others (Eds.). Springer Verlag, 16–28.

[6] S. Kern, N. Hansen, and P. Koumoutsakos. 2006. Local meta-models for optimization using evolution strategies. In *Parallel Problem Solving from Nature — PPSN IX*, T. P. Runarsson and others (Eds.). Springer Verlag, 939–948.

[7] O. Kramer. 2016. *Machine Learning for Evolution Strategies*. Springer Verlag.

[8] I. Loshchilov. 2013. *Surrogate-Assisted Evolutionary Algorithms*. Ph.D. Dissertation. Université Paris Sud – Paris XI.

[9] A. Stuart and J. K. Ord. 1994. *Kendall's Advanced Theory of Statistics* (sixth ed.). Vol. I: Distribution Theory. Wiley.