PhD position: Stochastic and surrogate-assisted multi-objective optimization
in collaboration with INRIA (Dimo Brockhoff, Nikolaus Hansen), CNRS (Rodolphe Le Riche) and STORENGY (Frédéric Huguet)

Context

Optimization problems with multiple objective functions are pervasive in practice. They are also key problems at Storengy, mainly for gas storage reservoir history matching studies.
The objective functions are typically: the bottom well pressure, the gas-water interface position in wells and the water production. The reservoir simulator can be seen as an expensive high-dimensional simulator which internal parameters need to be adapted in order to match simulator output and measured history data. Due to the number of conflicting objective functions, the solution of the problem is not unique and we need to find a set of solutions to explain the uncertainty on parameters and measurements.
Recently, it turned out that a reformulation of 2-objective optimization problems as a single-objective problem, optimizing the quality of the entire solution set evaluated so far, can be solved efficiently by single-objective solvers such as Quasi-Newton BFGS or the well-known Covariance Matrix Adaptation Evolution Strategy (CMA-ES, [1]). The usage of the latter resulted in the so-called COMO-CMA-ES [3]. Theoretical results in combination with numerical experiments show that for two objective functions and convex quadratic objective functions, convergence towards the theoretical optimum should occur [3]. However, the fact that objective functions can come from expensive numerical simulations reduces the amount of available function evaluations to a few hundred or thousand evaluations and might not allow a good convergence. In this budget range surrogate based algorithms such as classical trust-region methods or Bayesian optimization approaches are expected to outperform stochastic algorithms like CMA-ES. At the same time, the combination of CMA-ES with Bayesian approaches can help to investigate and manage the uncertainties present in the problems at Storengy. As importantly, the current COMO-CMA-ES does not allow to exploit highly parallel computations, another important aspect when dealing with expensive objective functions.

Goal

The goal of this thesis project is to extend the available COMO-CMA-ES algorithm towards expensive optimization by using surrogate models in order to save true function evaluations and manage high dimensional problems. A first approach for single-objective optimization problems, which is a portfolio algorithm combining a model-based version of CMA-ES with a classical solver such as SLSQP, has significantly reduced the amount of necessary function evaluations to reach a given target [2]. In this thesis, we would like to follow a similar line for the multiobjective COMO-CMA-ES algorithm. The COMO-CMA-ES builds a single-objective (dynamic) objective function to be optimized to find multiple near-optimal solutions. In this context, the approach for using surrogate models to solve multiobjective problems will be to employ a single surrogate model for the single-objective function, optimized within the COMO-CMA-ES instead of previous approaches that build a surrogate model for each objective function separately, see for example [4,5]. The advantage of a single surrogate is to capture in a unique model the compromise that underlies multi-objective decision problems. It furthermore reduces the internal complexity of the algorithm if only a single surrogate has to be learned. The option of a single surrogate will be compared to the more classical approach with a surrogate per optimization criterion.

Challenges

The COMO-like aggregation of the objectives, the augmentation of CMA-ES by surrogates and the distribution of optimization algorithms are research directions with already proven benefits [6,7]. The intersection of these principles raises clear hopes for further progress, but important scientific questions need to be addressed first.

- Kriging is a candidate surrogate that naturally accounts for the uncertainties on the true functions. CMA-ES constructs an uncertainty model in the form of a multinormal law in the space of design variables. It is still required to properly understand the relationship between the two models from a Bayesian perspective.
- Second, the single-objective reformulation within COMO-CMA-ES is a dynamically changing function and we need to address the question of how to integrate/exploit this aspect in the model building and how to combine it with a parallelized implementation.
- Third, which type of surrogate model shall we use (for example a simple and quick linear/quadratic one vs. a more elaborate but also more expensive to fit model like kriging)? If we use multiple models, when shall we switch among them? From the perspective of the search space, to tackle the dimensionality issue, how can we reduce dynamically the dimension of the surrogate model?

Collaboration team

- The thesis will be coadvised between the INRIA RandOpt team (Dimo Brockhoff and Anne Auger), CNRS LIMOS at Mines Saint-Etienne (Rodolphe Le Riche) and the Storengy company (Frédéric Huguet).
- The PhD is part of a larger consortium project called CIROQUO ("Consortium Industrie Recherche pour l'Optimisation et la QUantification d'incertitude pour les données Onéreuses") with 6 industrial and 6 academic partners carrying out research in statistical modeling and optimization. This guarantees numerous scientific interactions during the course of the PhD.

Candidate Profile

- Probability/statistics/operational research student, with a master degree or equivalent
- Good mastery of the foundations of statistical learning and optimization
- Ease in scientific programming, with a good knowledge of R, Python

Practical arrangements

- Start: fall 2021
- Location: Paris Ile de France (INRIA Saclay) or Mines Saint-Etienne depending on the candidate
- Salary: around 2000 Euros/month

Contacts

- dimo.brockhoff@inria.fr
- frederic.huguet@storengy.com
- leriche@emse.fr

References

[1] Hansen, Nikolaus, and Andreas Ostermeier. "Completely derandomized self-adaptation in evolution strategies." Evolutionary computation 9.2 (2001): 159-195.

[2] Hansen, Nikolaus. "A global surrogate assisted CMA-ES." Proceedings of the Genetic and Evolutionary Computation Conference. 2019.

[3] Touré, Cheikh, et al. "Uncrowded hypervolume improvement: COMO-CMA-ES and the sofomore framework." Proceedings of the Genetic and Evolutionary Computation Conference. 2019.
[4] Tabatabaei, Mohammad, et al. "A survey on handling computationally expensive multiobjective optimization problems using surrogates: non-nature inspired methods." Structural and Multidisciplinary Optimization 52.1 (2015): 1-25.
[5] David Gaudrie, Rodolphe Le Riche, Victor Picheny, Benoît Enaux and Vincent Herbert, "Targeting Solutions in Bayesian Multi-Objective Optimization: Sequential and Batch Versions", Annals of Mathematics and Artificial Intelligence, doi:10.1007/s10472-019-09644-8, vol.88, no.1, 2020, pp.187-212.
[6] L. Bajer, Z. Pitra, J. Repický, and M. Holena, "Gaussian Process Surrogate Models for the CMA Evolution Strategy", Evolutionary Computation, 27:665–697, 2019.
[7] Janusevskis, J., Le Riche, R., Ginsbourger, D. and R. Girdziusas, "Expected improvements for the asynchronous parallel global optimization of expensive functions : potentials and challenges", LNCS 7219, Springer Verlag, ISBN: 978-3-642-34412-1, 2012, pp. 413-418.