Convergent Derivative-Free Optimization in Mixed-Integer Domains

Juan J Torres[1], Giacomo Nannicini[2], Emiliano Traversi[1], Roberto Wolfler Calvo[1]

[1] Université Sorbonne Paris Nord, Sorbonne Paris Cité, LIPN, CNRS, (UMR 7030), France
{torresfigueroa,traversi,wolfler}@lipn.univ-paris13.fr
[2] IBM T.J. Watson Research Center, NY, U.S.A.
nannicini@us.ibm.com

Mots-clés : *Derivative-free optimization, Mixed-integer programming, Surrogate-based optimization, Trust-region method*

1 Introduction

In this project we propose an algorithm for the solution of the following problem :

$$\min_{x,y} f(x,y) \tag{1a}$$

$$x_{lb} \leq x \leq x_{ub}, \quad y_{lb} \leq y \leq y_{ub} \tag{1b}$$

$$x \in \mathbb{R}^{n_1}, \quad y \in \mathbb{Z}^{n_2} \tag{1c}$$

where $f : [\mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}] \to \mathbb{R}$ is a mixed-integer black-box function. Black-box functions are often expensive-to-evaluate and do not present analytical form, which means that no gradient nor second-order information can be used to optimize them. Black-box problems arise in several settings, such as medical imaging, engineering design, operations research and financial applications, among others.

Different methodologies have been developed to address the complexity of black-box problems, aiming to find near-optimal solutions while performing a reduced number of objective function evaluations to keep the optimization time acceptable. Such methodologies often include the use (and combination) of heuristics, direct-search algorithms, model based (or surrogate) approximation and randomized search. Derivative-free optimization is a natural technique to solve black-box optimization problems, as it is intrinsically designed to avoid the computation of derivatives.

Black-box problems often include non-relaxable discrete variables. The presence of discrete (binary, integer or categorical) variables add additional challenges to the ones faced in the optimization of mixed-integer functions. Methodologies for the solution of such problems are often based on the generalization of continuous derivative-free techniques, for example :

1. Alternate continuous and integer search, via direct search and local search, respectively [1].
2. Local quadratic surrogate approximation based on trust-region methods [2].
3. Global surrogate approximation via Radial Basis Functions (RBF) models [3], [4], [5].

A crucial issue faced while addressing mixed integer-derivative free problems is the definition of a mixed-integer minimizer, and a proper convergence criterion. Newby and Ali [2] identify three possible definitions : *separate-local minimum*, *stronger-local minimum* and *combined-local minimum*. The three concepts deal with the degree of exploration of different integer manifolds around a tentative solution. Such conditions are not often met and algorithm termination is commonly addressed via objective improvement and distance criteria.

In this project we develop a convergent algorithm that tackles several issues. We propose a trust-region based method which uses a tailored mixed-quadratic approximation that enables the efficient reuse of previously sampled points and establish global convergence into (some form of) a local minimum. The approximation is based on the assumption of dealing with a structured objective named Locally Quadratic Mixed-Integer (LMQI) function, for which the notions of continuous *fully-linear/quadratic models* could be extended, facilitating the convergence analysis, even when such assumptions are not hold. In addition, LQMI approximations can be obtained in a modular manner, providing flexibility in the surrogate model computation and maintenance, using efficient methods based on continuous derivative-free optimization.

To asses the rate of convergence and the quality of the solutions obtained by three different variants of our proposed methodology, we tested several instances with and without LQMI structure. The results are compared with the ones obtained from *pyNomad*, a python interface of NOMAD [6]. NOMAD is a derivative-free solver which implements a generalization of the Mesh Adaptive Direct Search (MADS) algorithm [7], designed to solve a broad class of constrained nonlinear optimization problems which may include integer, and, in cases categorical variables.

To benchmark the behavior of the solution approaches we use the scheme of *performance* and *data* profiles [8]. In the derivative-free setting, a solver $s \in S$ is said to be convergent for problem $p \in P$, if it is able to achieve a specific fraction of the best objective reduction obtained for p. *Performance profiles* are useful to identify how well a solver performs relative to other solvers in S for the set of instances in P. The *performance* profile of solver s represent the cumulative distribution of probability of *performance ratio* $r_{s,p} := t_{s,p}/ \min_{\hat{s} \in S}\{t_{\hat{s},p}\}$, where $t_{s,p}$ is the number of function evaluations required to satisfy the convergence test. Larger values of $t_{s,p}$ denote worse performance. On the other hand, *data* profiles are handy to identify solver performance in cases when the number of function evaluations is limited. A *data* profile represents the percentage of instances that solver s can solve with κ *simplex* gradients $t_{s,p}/(n_{p,1} + n_{p,2} + 1)$.

The next two figures show the performance of NOMAD and the three variants of the algorithm in 260 LQMI instances and 290 general MINLP test functions, including hard to solve non-smooth instances.

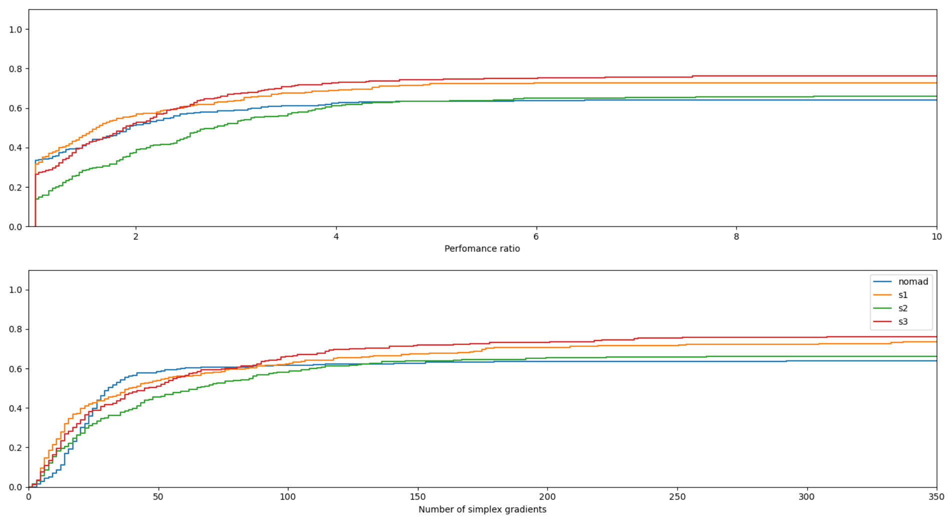

FIG. 1 – Performance and data profiles in LQMI instances with 99 % of objective reduction

Figure 1 shows that the three variants of the algorithm obtain convergence faster and better quality solution than NOMAD in LQMI test instances, which include ellipsoidal, quadratic and maximum over quadratic functions. Figure 2 shows a better performance of NOMAD in instances for which the objective function does not exhibit LQMI structure ; nonetheless, with sufficient number of function evaluations all three algorithm variants are able to solve a large percentage of non-LQMI instances. Current research is centered in the improvement and acceleration of the methodology in a more general class of MINLP functions.

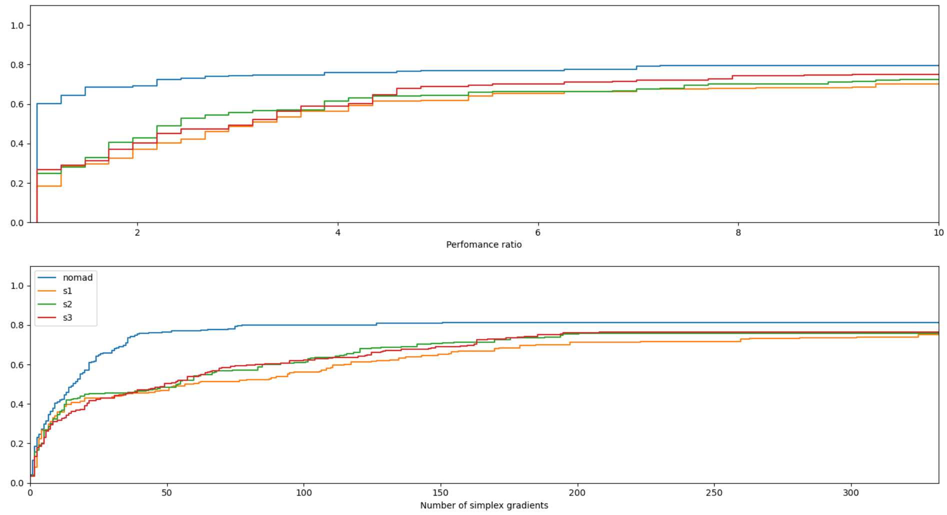

FIG. 2 – Performance and data profiles in MINLP instances with 99 % of objective reduction

Références

[1] G. Liuzzi, S. Lucidi, and F. Rinaldi, “Derivative-free methods for bound constrained mixed-integer optimization,” *Computational Optimization and Applications*, vol. 53, no. 2, pp. 505–526, 2012.

[2] E. Newby and M. M. Ali, “A trust-region-based derivative free algorithm for mixed integer programming,” *Computational Optimization and Applications*, vol. 60, no. 1, pp. 199–229, 2015.

[3] K. Rashid, S. Ambani, and E. Cetinkaya, “An adaptive multiquadric radial basis function method for expensive black-box mixed-integer nonlinear constrained optimization,” *Engineering Optimization*, vol. 45, no. 2, pp. 185–206, 2013.

[4] A. Costa and G. Nannicini, “Rbfopt : an open-source library for black-box optimization with costly function evaluations,” *Mathematical Programming Computation*, vol. 10, no. 4, pp. 597–629, 2018.

[5] K. Holmström, N.-H. Quttineh, and M. M. Edvall, “An adaptive radial basis algorithm (arbf) for expensive black-box mixed-integer constrained global optimization,” *Optimization and Engineering*, vol. 9, no. 4, pp. 311–339, 2008.

[6] S. Le Digabel, “Algorithm 909 : Nomad : Nonlinear optimization with the mads algorithm,” *ACM Trans. Math. Softw.*, vol. 37, Feb. 2011.

[7] C. Audet and J. E. Dennis Jr, “Mesh adaptive direct search algorithms for constrained optimization,” *SIAM Journal on optimization*, vol. 17, no. 1, pp. 188–217, 2006.

[8] J. J. Moré and S. M. Wild, “Benchmarking derivative-free optimization algorithms,” *SIAM Journal on Optimization*, vol. 20, no. 1, pp. 172–191, 2009.