COAP 2019 Best Paper Prize: Paper of S. Gratton, C. W. Royer, L. N. Vicente, and Z. Zhang

Each year, the editorial board of Computational Optimization and Applications selects a paper from the preceding year's publications for the Best Paper Award. Two of the 92 papers published by the journal in 2019 tied for the award. This article highlights the research related to the award winning paper of Serge Gratton (University of Toulouse), Clément W. Royer (Université Paris Dauphine-PSL), Luis Nunes Vicente (Lehigh University), and Zaikun Zhang (Hong Kong Polytechnic University) whose award winning paper "Direct search based on probabilistic feasible descent for bound and linearly constrained problems" was published in volume 72, pages 525–559.

Their paper [6] describes a direct-search scheme for solving linearly constrained optimization problems in which the derivatives of the objective are not available for algorithmic use. This derivative-free algorithm relies on randomly generated directions and is analyzed from a probabilistic viewpoint, leading to complexity guarantees for both deterministic and probabilistic versions of the method. These bounds suggest that strategies which incorporate randomness may prove more efficient, which is confirmed by the practical experiments.

The motivation for this algorithm was twofold. First, following recent developments in nonconvex optimization, complexity results have become increasingly popular in derivative-free optimization [8]. Most results, including some developed by the authors, focused on the unconstrained case, and were particularly concerned with bounding the number of evaluations of the objective function. This represents a key issue in computational derivative-free optimization, where the function values typically come from complex simulators that are expensive to evaluate. One goal of this work was to provide evaluation complexity bounds for direct search applied to linearly constrained problems, which could be informative about the evaluation cost of such methods.

Secondly, recent developments in the unconstrained, derivative-free setting showed the interest of incorporating randomness within algorithmic properties. This idea was first presented in the context of derivative-free trust-region methods [2], together with a convergence theory based on martingale-type arguments. For direct search, the authors of [6] not only established convergence by using randomly generated directions [4], but went one step further and derived high probability complexity bounds on the number of iterations and function evaluations required to reach approximate stationarity. The underlying argument was general enough to apply to other schemes such as trust region (which were studied in a subsequent work [5]). For the class of direct-search methods, it was shown that as few as two directions randomly generated on the unit sphere were sufficient to yield almost-sure convergence, as well as global rates with overwhelming probability. Moreover, the evaluation complexity bounds showed an improvement of order n, from $\mathcal{O}(n^2)$ (using deterministic directions in a positive spanning set) to $\mathcal{O}(n)$ (using randomized directions), where n is the dimension of the problem. In the paper [3] it was shown that the bound $\mathcal{O}(n^2)$ for deterministic directions could not be further improved, which made the findings even more relevant in an algorithmic context highly dependent on the problem size. Once the unconstrained case was thoroughly studied, it became natural to focus on constrained problems to see whether the evaluation cost of direct-search methods could be improved through randomness.

In an unconstrained setting, direct-search methods proceed by sampling the space in suitably chosen sets of directions, called polling sets. To establish global convergence, one of those directions must be able to produce descent, i.e., decrease the objective function. In the presence of constraints, the polling vectors must conform to the geometry of the feasible set. A number of strategies for generating such directions have been proposed in the literature [7], that mainly revolve around computing positive generators of certain (tangent) cones as polling sets, and this process can lead to many function evaluations being performed at every iteration, especially in the presence of linear inequality constraints. The first contribution of the awarded work was the identification of a feasible descent property, that must be satisfied in a deterministic fashion for a deterministic scheme to converge. In the unconstrained case, it is required that one of the polling directions makes an acute angle with the negative gradient; in presence of linear constraints, this property can be generalized by using the tangent cone of the current approximate active constraints. In that sense, the directions are required to be close to the projection of the negative gradient onto such a cone. The authors leveraged this property to derive a complexity analysis of a generic deterministic direct-search scheme, which was a problem left open in the literature. The generic analysis covered all possible combinations of bounds, linear equality constraints, and linear inequality constraints. Two cases of particular interest were linear equality constraints, for which the complexity depended on the dimension n and the number of constraints, and bound constraints, for which the dependency was similar to the unconstrained case, due to the use of coordinate directions.

The paper then turned to the probabilistic setting, replacing the feasible descent property by a probabilistic equivalent. By assuming that the polling sets satisfied this property with a given probability, complexity bounds were established. Special attention was paid to the case of linear equality constraints and bound constraints; when the problem was locally unconstrained in some directions forming a subspace, only $\mathcal{O}(1)$ directions were required to guarantee feasible descent in this subspace, exactly matching the unconstrained case. This approach proved particularly useful for constraints such as bounds that may not necessarily be active at a given iteration. In addition to improving the complexity results, this approach turned out to be beneficial in practice.

The goal of the numerical experiments was analyze the efficiency of the algorithmic approach that, although possibly incorporating random components, rely upon deterministic polling techniques to ensure convergence. Compared to the built-in MATLAB `patternsearch` subroutine [9], the new algorithm was particularly efficient on problems with a relatively large number of bound constraints and/or linear equality constraints, where the subspace technique led to fewer evaluations being performed. Comparisons were also performed with more mature direct-search solvers for linearly constrained problems such as PSwarm [10] and `NOMAD` [1]. Despite the relatively straightforward implementation, the new algorithm provided encouraging results, and was particularly competitive on problems involving linear equality constraints.

Since the publication of this paper in *Computational Optimization and Applications*, the associated code, called `DSPFD`, has been made publicly available at the following location:

`https://www.lamsade.dauphine.fr/~croyer/codes/dspfd_sources.zip`

It can handle both unconstrained and linearly constrained minimization formulations, and only requires access to function objective values. Extensions to nonlinear constraints remains a challenging topic for further investigation, and one that has been relatively less explored in derivative-free optimization. Probabilistic techniques may prove helpful in this setting to mitigate the cost of possibly expensive algorithmic frameworks without compromising their theoretical properties.

References

[1] C. Audet, S. Le Digabel, and C. Tribes. NOMAD user guide. Technical Report G-2009-37, Les cahiers du GERAD, 2009.

[2] A. S. Bandeira, K. Scheinberg, and L. N. Vicente. Convergence of trust-region methods based on probabilistic models. *SIAM J. Optim.*, 24:1238–1264, 2014.

[3] M. Dodangeh, L. N. Vicente, and Z. Zhang. On the optimal order of worst case complexity of direct search. *Optim. Lett.*, 10:699–708, 2016.

[4] S. Gratton, C. W. Royer, L. N. Vicente, and Z. Zhang. Direct search based on probabilistic descent. *SIAM J. Optim.*, 25:1515–1541, 2015.

[5] S. Gratton, C. W. Royer, L. N. Vicente, and Z. Zhang. Complexity and global rates of trust-region methods based on probabilistic models. *IMA J. Numer. Anal.*, 38:1579–1597, 2018.

[6] S. Gratton, C. W. Royer, L. N. Vicente, and Z. Zhang. Direct search based on probabilistic feasible descent for bound and linearly constrained problems. *Comput. Optim. Appl.*, 72:525–559, 2019.

[7] T. G. Kolda, R. M. Lewis, and V. Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev.*, 45:385–482, 2003.

[8] J. Larson, M. Menickelly, and S. M. Wild. Derivative-free optimization methods. *Acta Numer.*, 28:287–404, 2019.

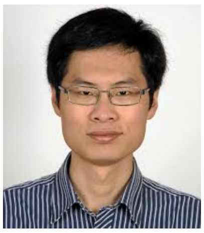
[9] The Mathworks, Inc. *Global Optimization Toolbox User's Guide,version 3.3*, October 2014.

[10] A. I. F. Vaz and L. N. Vicente. PSwarm: A hybrid solver for linearly constrained global derivative-free optimization. *Optim. Methods Softw.*, 24:669–685, 2009.

Serge Gratton is a chair professor on Data Assimilation and Machine Learning at ANITI (Artificial and Natural Intelligence Toulouse Institute). Full professor at ENSEEIHT, University of Toulouse, France, since 2009, he is also the head of the Parallel Algorithms and Optimization Team at IRIT, Toulouse. He has held researcher positions in CERFACS, both as a senior researcher and as a group leader, and in the French Spatial agency (CNES), as research engineer. He received his Ph.D. from the University of Toulouse in 1998.

Clément W. Royer is an associate professor of computer science at Université Paris Dauphine-PSL, France. From 2016 to 2019, he was a postdoctoral research associate at the Wisconsin Institute of Discovery, University of Wisconsin-Madison, USA. He received his Ph.D. from the University of Toulouse, France in 2016.

Luis Nunes Vicente is the Timothy J. Wilmott '80 Endowed Chair Professor and Department Chair of the Department of Industrial and Systems Engineering of Lehigh University. He received his Ph.D. from Rice University in 1996. From 1996 to 2018, he was a faculty member at the Department of Mathematics of the University of Coimbra, Portugal. He held visiting positions at the IBM Watson Research Center, CERFACS, INP-Toulouse, and the Universities of Minnesota, New York, Paul Verlaine, Sapienza, and Rice.



Zaikun Zhang is an assistant professor in the Department of Applied Mathematics of Hong Kong Polytechnic University, Hong Kong, China. He was a postdoctoral researcher in IRIT, France, from 2014 to 2016, and at the University of Coimbra, Portugal, from 2012 to 2014. He received his Ph.D. from the Chinese Academy of Sciences, China in 2012.