RESEARCH ARTICLE | SEPTEMBER 01 2023

Possible extensions to the DIRECT global optimization algorithm based on space-filling and diagonal curves FREE

Dmitri E. Kvasov ✉; Daniela Lera; Yaroslav D. Sergeyev

AIP Conference Proceedings 2849, 450033 (2023)
https://doi.org/10.1063/5.0162069

Possible Extensions to the DIRECT Global Optimization Algorithm Based on Space-Filling and Diagonal Curves

Dmitri E. Kvasov[1,2,a)], Daniela Lera[3,b)] and Yaroslav D. Sergeyev[1,2,4,c)]

[1]DIMES, University of Calabria, via P. Bucci 41C, 87036 – Rende (CS), Italy
[2]IITMM, Lobachevsky State University of Nizhni Novgorod, 603950 – Nizhni Novgorod, Russia
[3]DMI, University of Cagliari, via Ospedale 72, 09124 – Cagliari, Italy
[4]ICAR-CNR, via P. Bucci, 8-9C, 87036 – Rende (CS), Italy

[a)]Corresponding author: kvadim@dimes.unical.it
[b)]lera@unica.it
[c)]yaro@dimes.unical.it

Abstract. In this paper, the Lipschitz global optimization problem is considered both in the cases of non-differentiable and differentiable objective functions over hyperintervals. It is shown that space-filling and diagonal curves can be successfully used to extend promising one-dimensional methods to the multidimensional case. In particular, several DIRECT-based algorithms using Peano-Hilbert space-filling curves and adaptive diagonal curves are surveyed.

PROBLEM STATEMENT

The following global optimization problem is considered in this paper: it is required to approximate the global solution ϕ^* of the objective function $\phi(x)$ and the respective global minimizer x^* such that:

$$\phi^* = \phi(x^*) = \min \phi(x), \quad x \in \Delta \subset \mathbb{R}^n. \tag{1}$$

We suppose here that the search domain Δ is a hyperinterval $\Delta = [a, b] \subset \mathbb{R}^n$ and the function $\phi(x)$ is black-box, multiextremal, hard to evaluate, and satisfies the Lipschitz condition (with an unknown Lipschitz constant L, $0 < L < \infty$). If $\phi(x)$ is differentiable, its gradient $\nabla\phi(x)$ is assumed to be computationally expensive Lipschitz vector-function (with an unknown Lipschitz constant K, $0 < K < \infty$).

This problem statement (known as Lipschitz global optimization problem) is very frequent in practice (see, e. g., [1–8]) and, therefore, numerical methods are being intensively developed (see, e. g., [9–24]) to address this challenging problem efficiently (usually, the less function evaluations are required to obtain an accepted solution the better is the method).

PEANO-TYPE AND DIAGONAL CURVES IN LIPSCHITZ GLOBAL OPTIMIZATION

One of the promising approaches to deal with multidimensional Lipschitz global optimization problems consists of extending efficient one-dimensional methods to the multidimensional case, for example, by means of space-filling curves (introduced by Peano and further by Hilbert, see, e. g., [25, 26] for details) that pass through the points of the search domain. Different techniques for the construction of approximations to Peano-Hilbert curves, their theoretical study, and implementations are reported, e. g., in [8, 20, 27, 28].

Given a single-valued Peano-Hilbert curve $x(t)$ continuously mapping the unit segment $[0, 1]$ onto the search hyperinterval $\Delta \subset \mathbb{R}^n$, it can be proved (see, e. g., [8, 27]) that

$$\min_{x\in\Delta} \phi(x) = \min_{t\in[0,1]} \phi(x(t)), \tag{2}$$

International Conference of Numerical Analysis and Applied Mathematics ICNAAM 2021
AIP Conf. Proc. 2849, 450033-1–450033-4; https://doi.org/10.1063/5.0162069
Published by AIP Publishing. 978-0-7354-4589-5/$30.00

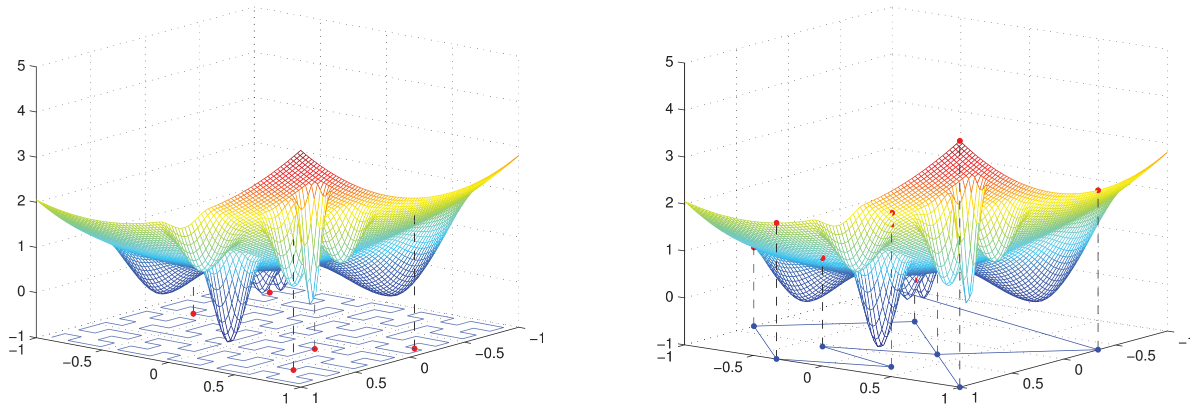

FIGURE 1. The usage of Peano-Hilbert space-filling curves (*left figure*) and adaptive diagonal curves (*right figure*) for extending one-dimensional methods to solve multidimensional problems

and, therefore, the stated multidimensional problem is reduced to one dimension (see the left part of Fig. 1). When the multidimensional function $\phi(x)$ from (1) is Lipschitz with a constant L, the one-dimensional function $\phi(x(t))$, $t \in [0, 1]$, is Hölderian over $[0, 1]$ with the exponent n^{-1} and the coefficient $2L\sqrt{n+3}$ (see [8, 27] for details). Although the Lipschitz assumption does not hold for the one-dimensional function $\phi(x(t))$, this function satisfies the more general Hölder condition that allows one to generalize Lipschitz one-dimensional methods to optimization of Hölder functions (see, e. g., in [8, 20, 27, 29–32]).

Another interesting methodology for extending one-dimensional algorithms to higher dimensions is the diagonal approach (see, e. g., [18] for details). In the diagonal algorithms, the hyperinterval Δ from (1) is adaptively partitioned into smaller hyperintervals and the objective function $\phi(x)$ is evaluated only at two vertices corresponding to the main diagonal of the obtained hyperintervals. The non-redundant diagonal partition strategy introduced in [33] allows one to construct efficient multidimensional global optimization methods. A general diagonal scheme (see the right part of Fig. 1) from [18, 34, 35] takes advantage of both the ideas of the diagonal approach and Peano-Hilbert space-filling curves. Different multidimensional diagonal algorithms for solving Lipschitz global optimization problems have been proposed and their convergence properties have been analyzed (see, e. g., [34, 36, 37]).

We will illustrate the usage of the above-mentioned Peano-type and diagonal curves to construct multidimensional Lipschitz global optimization methods. As a basic one-dimensional algorithm to be extended at higher dimensions, the DIRECT algorithm [38, 39] will be considered due to its versatility and popularity in the global optimization community. It is characterized by the multiple estimates of the Lipschitz constant L for the objective function from (1), as reported in the next Section.

MULTIPLE ESTIMATES OF THE LIPSCHITZ CONSTANTS

In Lipschitz global optimization, the influence of the Lipschitz constant estimates on the convergence speed of global optimization algorithms should be carefully taken into account. As pointed out, e. g., in [18, 40] either the Lipschitz constant is known and an algorithm is developed correspondingly, or it is not known but there exist sufficiently many settings of parameters of the algorithm under examination ensuring its convergence. Thus, estimating the Lipschitz constants (for the objective function and its derivative, if available) is among the main aspects to be considered in the context of the stated problem (1).

In particular, this constant can be a priori given or its adaptive local/global estimates can be computed during the search (see, e. g., [18] for reviewing various techniques proposed to obtain estimates of the Lipschitz constants). Jones et al. [38] have proposed to use multiple estimates (from zero to infinity) of the Lipschitz constant for $\phi(x)$ in divide-the-best Lipschitz global optimization algorithms (see, e. g., [18, 41]) thus giving rise to a fruitful research direction in this field. The basic idea of DIRECT (in its one-dimensional variant) is to select for partitioning and evaluating the function $\phi(x)$ the so-called potentially optimal intervals, i. e., intervals where $\phi(x)$ can obtain an improvement w.r.t. an estimate of the Lipschitz constant (clearly, there can be several of such intervals corresponding to various estimates of the Lipschitz constant). To do this, each interval $[a_i, b_i]$ of the adaptive partition of $[a, b]$ is represented graphically on a two-dimensional diagram as a dot with horizontal coordinate $d_i = (b_i - a_i)/2$ and vertical coordinate $\phi(d_i)$. Potentially optimal intervals correspond to the dots located on the lower right convex hull of all current dots.

Despite the existence of many DIRECT-based algorithms, there was not known a way to use multiple estimates of the Lipschitz constants for Hölderian objective functions or for differentiable ones until publications [42] and [43], respectively, by the authors of the present extended abstract.

DIRECT-BASED METHODS USING PEANO-TYPE AND DIAGONAL CURVES

Let us conclude this short paper by giving some suggestions on how to use the ideas of DIRECT in the framework of multidimensional algorithms based on Peano-type and diagonal curves (see, e. g., [20, 37]).

In the case of the usage of Peano-Hilbert space-filling curves as the basis to construct multidimensional global optimization methods starting from the one-dimensional DIRECT algorithm, one needs to overcome the difficulties with locating potentially optimal intervals for the Hölder objective functions. This non-trivial task can be solved (see [20, 42]) by introducing, in the graphical representation of intervals, the Hölderian metric instead of the Euclidean one. In particular, it is proposed in [42] to represent each interval $[a_i, b_i]$ by a dot with horizontal coordinate $\bar{d}_i = |b_i - a_i|^{1/n}$ and vertical coordinate $\phi(c_i)$ with $c_i = 0.5(a_i + b_i)$. The values R_i called characteristics (see, e. g., [41]) and being lower bounds for the Hölderian (with the Hölder constant H) function $\phi(x)$ over each interval $[a_i, b_i]$,

$$R_i = \phi(c_i) - \hat{H}\bar{d}_i, \quad \hat{H} > H,$$

can so be obtained. Thus, the introduced Hölderian metric allows one to avoid the non-linearity in determining the lower right convex hull of dots representing intervals $[a_i, b_i]$ and to extend one-dimensional DIRECT-based methods to higher dimensions over space-filling curves. Theoretical and experimental results obtained in [42] (see also [20, 44, 45]) opened up a new interesting research direction in Lipschitz and Hölder global optimization.

Similar difficulties with the DIRECT-inspired representation of intervals arise in the case of differentiable objective functions when at a point $x \in \Delta$ the information about both $\phi(x)$ and $\nabla\phi(x)$ is available. To use multiple estimates of the Lipschitz constant K for $\nabla\phi(x)$, a non-linear (with a quadratic distance as horizontal coordinate) bi-dimensional diagram representing dots of the current partition of the search domain has been proposed for the first time in [43]. This innovative idea allowed the authors to extend the ideas of DIRECT to multidimensional diagonal algorithms for differentiable optimization where characteristics R_i over hyperintervals $[a_i, b_i] \subset \Delta$,

$$R_i = \Phi_i - 0.5\hat{K}\|b_i - a_i\|^2 \quad \hat{K} > K,$$

can be found graphically in a simple way (the vertical coordinate Φ_i can easily be obtained as shown, e. g., in [46]). This contribution provided the global optimization researchers with another powerful framework to develop efficient global optimization methods when the information about gradients of the objective function is available.

ACKNOWLEDGMENTS

Dmitri E. Kvasov is grateful for the support provided by the Russian Science Foundation, project No. 21-11-00204. Daniela Lera thanks INdAM-GNCS Project 2020 “Numerical algorithms in optimization, ODEs, and applications”.

REFERENCES

[1] J. W. Gillard and D. E. Kvasov, Statistics and Its Interface **10**, 59–70 (2016).
[2] J. W. Gillard and A. Zhigljavsky, J. Global Optim. **57**, 733–751 (2013).
[3] D. E. Kvasov and Y. D. Sergeyev, Automat. Remote Control **74**, 1435–1448 (2013).
[4] D. E. Kvasov and Y. D. Sergeyev, Advances in Engineering Software **80**, 58–66 (2015).
[5] D. Lera, M. Posypkin, and Y. D. Sergeyev, Appl. Math. Comput. **390** (2021), article 125660.
[6] Y. D. Sergeyev and V. A. Grishagin, J. Comput. Anal. Appl. **3**, 123–145 (2001).
[7] Y. D. Sergeyev, A. Candelieri, D. E. Kvasov, and R. Perego, Soft Comput. **24**, 17715–17735 (2020).

[8] R. G. Strongin and Y. D. Sergeyev, *Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms* (Kluwer Academic Publishers, Dordrecht, 2000).
[9] F. Archetti and A. Candelieri, *Bayesian Optimization and Data Science*, SpringerBriefs in Optimization (Springer, New York, 2019).
[10] K. A. Barkalov and R. G. Strongin, J. Glob. Optim. **71**, 21–36 (2018).
[11] V. Gergel, K. Barkalov, and A. Sysoev, Numer. Algebra Contr. Optim. **8**, 47–62 (2018).
[12] V. Gergel, V. Grishagin, and R. Israfilov, "Adaptive dimensionality reduction in multiobjective optimization with multiextremal criteria," in *Machine Learning, Optimization, and Data Science (LOD 2018)*, LNCS, Vol. 11331, edited by G. Nicosia et al. (2019), pp. 129–140.
[13] V. A. Grishagin and R. A. Israfilov, "Global search acceleration in the nested optimization scheme," in *AIP Conference Proceedings*, Vol. 1738 (2016) p. 400010, doi: 10.1063/1.4952198.
[14] V. A. Grishagin, R. A. Israfilov, and Y. D. Sergeyev, Appl. Math. Comput. **318**, 270–280 (2018).
[15] R. Paulavičius and J. Žilinskas, *Simplicial Global Optimization*, SpringerBriefs in Optimization (Springer, New York, 2014).
[16] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, and J. Žilinskas, J. Global Optim. **59**, 545–567 (2014).
[17] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, and J. Žilinskas, Expert Systems with Applications **144** (2020), article 113052.
[18] Y. D. Sergeyev and D. E. Kvasov, *Deterministic Global Optimization: An Introduction to the Diagonal Approach*, SpringerBriefs in Optimization (Springer, New York, 2017).
[19] Y. D. Sergeyev, D. E. Kvasov, and F. M. H. Khalaf, Optimization Letters **1**, 85–99 (2007).
[20] Y. D. Sergeyev, R. G. Strongin, and D. Lera, *Introduction to Global Optimization Exploiting Space-Filling Curves*, SpringerBriefs in Optimization (Springer, New York, 2013).
[21] Y. D. Sergeyev, D. E. Kvasov, and M. S. Mukhametzhanov, Sci. Rep. **8** (2018), article 453.
[22] Y. D. Sergeyev, M. C. Nasso, M. S. Mukhametzhanov, and D. E. Kvasov, J. Comput. Appl. Math. **383** (2021), article 113134.
[23] A. Zhigljavsky and A. Žilinskas, *Bayesian and High-Dimensional Global Optimization*, SpringerBriefs in Optimization (Springer, New York, 2021).
[24] A. Žilinskas and A. Zhigljavsky, Informatica **27**, 229–256 (2016).
[25] H. Sagan, *Space–Filling Curves* (Springer–Verlag, New York, 1994).
[26] R. G. Strongin, in *Encyclopedia of Optimization*, Vol. 2, edited by C. A. Floudas and P. M. Pardalos (Kluwer Academic Publishers, Dordrecht, 2001), pp. 345–350.
[27] R. G. Strongin, *Numerical Methods in Multiextremal Problems (Information-Statistical Algorithms)* (Nauka, Moscow, 1978) in Russian.
[28] R. G. Strongin and Y. D. Sergeyev, Parallel Computing **18**, 1259–1273 (1992).
[29] V. A. Grishagin, R. A. Israfilov, and Y. D. Sergeyev, "Comparative efficiency of dimensionality reduction schemes in global optimization," in *AIP Conference Proceedings*, Vol. 1776 (2016) p. 060011, doi: 10.1063/1.4965345.
[30] D. Lera and Y. D. Sergeyev, BIT **42**, 119–133 (2002).
[31] D. Lera and Y. D. Sergeyev, Appl. Numer. Math. **60**, 115–129 (2010).
[32] R. J. Vanderbei, J. Global Optim. **14**, 205–216 (1999).
[33] Y. D. Sergeyev, J. Optim. Theory Appl. **107**, 145–168 (2000).
[34] D. E. Kvasov and Y. D. Sergeyev, Comput. Math. Math. Phys. **43**, 40–56 (2003).
[35] D. E. Kvasov, 4OR – Quart. J. Oper. Res. **6**, 403–406 (2008).
[36] D. E. Kvasov, C. Pizzuti, and Y. D. Sergeyev, Numer. Math. **94**, 93–106 (2003).
[37] Y. D. Sergeyev and D. E. Kvasov, SIAM J. Optim. **16**, 910–937 (2006).
[38] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, J. Optim. Theory Appl. **79**, 157–181 (1993).
[39] D. R. Jones and J. R. R. A. Martins, J. Global Optim. **79**, 521–566 (2021).
[40] D. E. Kvasov and Y. D. Sergeyev, Numer. Algebra Contr. Optim. **2**, 69–90 (2012).
[41] V. A. Grishagin, Y. D. Sergeyev, and R. G. Strongin, J. Global Optim. **10**, 185–206 (1997).
[42] D. Lera and Y. D. Sergeyev, Commun. Nonlinear Sci. Numer. Simul. **23**, 328–342 (2015).
[43] D. E. Kvasov and Y. D. Sergeyev, Optim. Lett. **3**, 303–318 (2009).
[44] D. Lera and Y. D. Sergeyev, "Space-filling curves and multiple estimates of Hölder constants in derivative-free global optimization," in *AIP Conference Proceedings*, Vol. 1738 (2016) p. 400008, doi: 10.1063/1.4952196.
[45] D. Lera and Y. D. Sergeyev, J. Glob. Optim. **71**, 193–211 (2018).
[46] D. E. Kvasov and Y. D. Sergeyev, J. Comput. Appl. Math. **236**, 4042–4054 (2012).