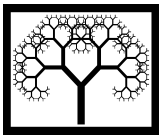
Paper 62



Proceedings of the Eleventh International Conference
on Computational Structures Technology,
B.H.V. Topping, (Editor),
Civil-Comp Press, Stirlingshire, Scotland

Deterministic Global Optimization Methods for Solving Engineering Problems

D.E. Kvasov[1,2] and Y.D. Sergeyev[1,2]
[1] Department of Electronics, Computer and System Sciences
University of Calabria, Rende (CS), Italy
[2] Software Department
N.I. Lobachevsky State University, Nizhny Novgorod, Russia

Abstract

In many important practical problems, some decision should be made by finding the global optimum of a multiextremal objective function subject to a set of constraints. Frequently, especially in engineering applications, the functions involved in optimization process are black-box with unknown analytical representations and hard to evaluate. Such computationally challenging decision-making problems often cannot be solved by traditional optimization techniques based on strong suppositions about the problem (convexity, differentiability, etc.). In this paper, some innovative and powerful approaches developed by the authors to construct efficient numerical methods for solving the mentioned problems are presented.

Keywords: global optimization, black-box functions, derivative-free methods, Lipschitz condition, applied problems.

1 Introduction

Numerical approaches to efficiently find optimal parameters of mathematical models arising from different real-life problems are nowadays becoming more and more significant in industrial processes. Optimization models characterized by the functions with several local optima (typically, their number is unknown and can be very high) have a particular importance for practical applications. When the best set of parameters should be determined for these multiextremal models, traditional local optimization techniques can be insufficient and, therefore, global optimization methods are used. Moreover, the objective functions and constraints to be examined are often black-box and hard to evaluate functions with unknown analytical representations. For example, their values can be obtained by executing some computationally expensive simulation, by performing a set of experiments, and so on. Such a kind of functions

is frequently met in various fields of human activity (as, e.g., automatics and robotics, structural optimization, safety verification problems, engineering design, network and transportation problems, mechanical design, chemistry and molecular biology, economics and finance, data classification, etc.) and corresponds to computationally challenging global optimization problems, being actively studied around the world (see, e.g., [1–12] and the references given therein).

To illustrate real world global optimization problems we deal with, let us consider the field of geomechanics and geophysics where one has to work with different mechanical-mathematical optimization models. Generally, these models are very complex since they can involve multidimensional linear or nonlinear partial differential equations with multiple contact boundaries, regions with sharp changes in functions values, ill-posedness, and so on. Knowledge of the properties and types of geological rocks lying at a depth of several kilometers is of great interest, e.g., for prospecting seismology, which determines the location of oil fields by means of acoustic waves. Its techniques allow one both to avoid the costly exploration methods (e.g., well drilling) and to accelerate the process of pinpointing oil resources. Among these techniques, numerical methods for solving inverse problems are of fundamental importance in prospecting seismology. They aim at estimating parameters of the Earth's structure and material properties (e.g., location of the inhomogeneities as cracks/cavities in the crust) based on data measured on the surface.

To give a concrete example of the resulting global optimization problem, a simplified version of the prospecting seismology inverse problem can be taken: namely, let us suppose that there is a fluid-filled crack of a given length located in the host rock with known elastic properties (see, e.g., [13]). Then, the vector x of unknown parameters defining the region geometry contains only two components: the depth of the crack occurrence h, $h_1 \le h \le h_2$, and the crack inclination angle α, $\alpha_1 \le \alpha \le \alpha_2$ (h_1, h_2, α_1, and α_2 are known constants).

One of the peculiarities of the stated problem is that information can be obtained only from experimental measurements with the usage of acoustic sounding (see, e.g., [14]). A number of seismic detectors are located at points d_i on the surface of the Earth, which record the vertical components $\tilde{V}_y(d_i, t_j)$ of particles velocity in the reflected wave at time instances t_j. We look for such a value of x that best fits the numerically simulated response $V_y(x, d_i, t_j)$ to a measured one. Computational simulation can be performed by some numerical integration algorithm: for example, the grid-characteristic method (see, e.g., [15, 16]) can be used for this scope, thus taking into account the physical features of the problem and allowing one to set correctly the boundary and contact conditions.

Hereby, this particular problem can be formulated as the following least squares optimization problem (see, e.g., [11, 12, 17–20]):

$$\min f(x), \quad x \in D = [h_1, h_2] \times [\alpha_1, \alpha_2], \tag{1}$$

$$f(x) = \sum_i \sum_j [V_y(x, d_i, t_j) - \tilde{V}_y(d_i, t_j)]^2. \quad (2)$$

Function (2) is essentially multiextremal, it has no analytical representation and its evaluation (sometimes called trial) is associated with performing computationally expensive numerical experiments. Therefore, the usage of fast and robust global optimization methods aimed at tackling this class of complex multiextremal problems is required for solving efficiently the problem (1)–(2).

Because of the huge computational costs involved, typically a small number of functions evaluations are available for a decision-maker (engineer, physicist, chemist, economist, etc.) when optimizing such costly functions. Thus, the main goal is to develop fast global optimization methods that produce acceptable solutions with a limited number of functions evaluations. But to obtain this goal, there are still a lot of difficulties that are mainly related either to the lack of information about the objective function (and constraints, if any) or to the impossibility to adequately represent the (poor) available information about the functions.

For example, gradient-based algorithms (see, e.g., [2,4,7]) cannot be used in many applications because black-box functions are often non-differentiable or derivatives are not available and their finite-difference approximations are too expensive to obtain. Automatic differentiation (see, e.g., [21]), as well as interval-based approaches (see, e.g., [22,23]), cannot be appropriately used in cases of black-box functions when their source codes are not available. Probably, a simple alternative could be the usage of the so-called direct (or derivative-free) search methods (see, e.g., [24–28]), frequently used now for solving engineering design problems (see, e.g., the DIRECT method [2, 24, 29], the response surface, or surrogate model methods [30, 31], etc.). But unfortunately (see, e.g., [32–34]), these methods either are designed to find only stationary points or can require too high computational effort for their work.

Therefore, solving the described global optimization problems is actually a real challenge both for theoretical and applied scientists. In this context, deterministic global optimization is a well developed mathematical theory which has many important applications (see, e.g., [2, 4, 8, 9, 11, 35, 36]). One of its main advantages is the possibility to obtain guaranteed estimations of global solutions and to demonstrate (under certain analytical conditions) rigorous global convergence properties. However, the currently available deterministic models can still require too large number of functions evaluations to obtain adequately good solutions for these problems.

Stochastic approaches (see, e.g., [2, 4, 6, 7, 12, 37]) can often deal with the stated problems in a simpler manner than the deterministic algorithms (being also suitable for the problems where the evaluations of the functions are corrupted by noise). However, there can be difficulties with these methods, as well (e.g., in studying their convergence properties). Several restarts can also be involved, requiring even more functions evaluations. Moreover, solutions found by many stochastic algorithms (especially, by heuristic methods like evolutionary algorithms, simulated annealing, etc.; see, e.g., [5, 7, 38, 39]) can be only local solutions to the problems, far from global

ones. This can preclude such methods from their usage in practice. That is why we concentrate, hereafter, on deterministic approaches.

The possibility to outperform the 'brute-force computation' techniques in solving global optimization problems is fundamentally based on the availability of some realistic a priori assumptions characterizing the objective function and eventual constraints (see, e.g., [4, 7–9, 11, 12]). They serve as mathematical tools for obtaining estimates of the global solution related to a finite number of function trials and, therefore, play a crucial role in the construction of any efficient global search algorithm. As observed, e.g., in [4, 40], if no particular assumptions are made on the objective function and constraints, any finite number of function evaluations cannot guarantee getting close to the global minimum value, since this function may have very high and narrow peaks.

One of the natural and powerful (from both the theoretical and the applied points of view) assumptions on the global optimization problem is that the objective function (and constraints) have bounded slopes. In other words, any limited change in the object parameters yields limited changes in the characteristics of the objective performance. This assumption can be justified by the fact that in technical systems the energy of change is always limited (see the related discussion in [11]). One of the most popular mathematical formulations of this property is the Lipschitz continuity condition, which assumes that the difference (in the sense of a chosen norm) of any two function values is majorized by the difference of the corresponding function arguments, multiplied by a positive factor $L < \infty$. In this case, the function is said to be Lipschitz and the corresponding factor L is said to be the Lipschitz constant. The problem involving Lipschitz functions (the objective function and constraints) is said to be the Lipschitz global optimization problem (see, e.g., [4, 8–12, 41, 42] and the references given therein).

The Lipschitz continuity assumption, being quite realistic for many practical black-box problems, is also an effective tool for obtaining accurate global optimum estimates after performing a limited number of functions evaluations. It is used by the authors to develop new efficient and reliable deterministic methods for solving multidimensional constrained global optimization problems from different real-life applied areas (as, e.g., the problem (1)–(2)), which are characterized by black-box multiextremal and hard to evaluate functions.

In the next Section, the Lipschitz global optimization problem is formally stated and examined more in detail. Some new approaches proposed by the authors to construct efficient numerical methods for solving the mentioned problems are briefly presented in Section 3. Section 4 concludes the paper by presenting some results of numerical experiments.

2 Lipschitz global optimization problem

A general Lipschitz global optimization problem can be formalized as follows (see, e.g., [8, 9, 11, 12, 42]):

$$f^* = f(x^*) = \min f(x), \quad x \in \Omega \subset \mathbb{R}^N, \tag{3}$$

where Ω is a bounded set defined as

$$\Omega = \{x \in D : \phi_i(x) \leq 0, 1 \leq i \leq p\}, \tag{4}$$

$$D = [a, b] = \{x \in \mathbb{R}^N : a(j) \leq x(j) \leq b(j), 1 \leq j \leq N\}, \quad a, b \in \mathbb{R}^N, \tag{5}$$

with N being the problem dimension. In (3)–(5), the objective function $f(x)$ and the constraints $\phi_i(x)$, $1 \leq i \leq p$, are multiextremal, non necessarily differentiable, black-box and hard to evaluate functions that satisfy the Lipschitz condition over the search hyperinterval D:

$$|f(x') - f(x'')| \leq L\|x' - x''\|, \quad x', x'' \in D, \tag{6}$$

$$|\phi_i(x') - \phi_i(x'')| \leq L_i\|x' - x''\|, \quad x', x'' \in D, \quad 1 \leq i \leq p, \tag{7}$$

where $\|\cdot\|$ denotes, usually, the Euclidean norm, L and L_i, $1 \leq i \leq p$, are the (unknown) Lipschitz constants such that $0 < L < \infty$, $0 < L_i < \infty$, $1 \leq i \leq p$. If $p = 0$ in (4), the problem is said to be box-constrained.

The admissible region Ω can consist of disjoint, non-convex subregions because of the multiextremality of the constraints $\phi_i(x)$. Moreover, these constraints can be partially defined, i.e., a constraint $\phi_{i+1}(x)$ (or the objective function $f(x)$) can be defined only over subregions where $\phi_i(x) \leq 0$, $1 \leq i \leq p$ (see, e.g., [10, 11] for more details and applied examples).

The problem (3), (5), (6) with a differentiable objective function having the Lipschitz (with an unknown Lipschitz constant) gradient $f'(x)$ (which could be itself a multiextremal black-box function) is sometimes included in the same class of Lipschitz global optimization problems (see, e.g., the references given in [11, 34, 43]).

As evidenced, e.g., in [10, 11], it is not easy to manage multiextremal constraints (4) within the context of Lipschitz global optimization. For example, the traditional penalty approach (see, e.g., the references in [2, 4, 27]) can lead to extremely high Lipschitz constants, thus forcing degeneration of the methods. In this connection, a promising approach called the index scheme (see, e.g., [11, 44–46]) can be applied. It does not introduce additional variables and/or parameters by opposition as, e.g., many traditional penalty approaches do, and reduces the general constrained problem (3)–(7) to a box-constrained discontinuous one.

Therefore, in order to give an insight into the principal ideas of the authors' techniques for solving the stated problem, the box-constrained Lipschitz global optimization problem (3), (5), (6) will be considered in the following.

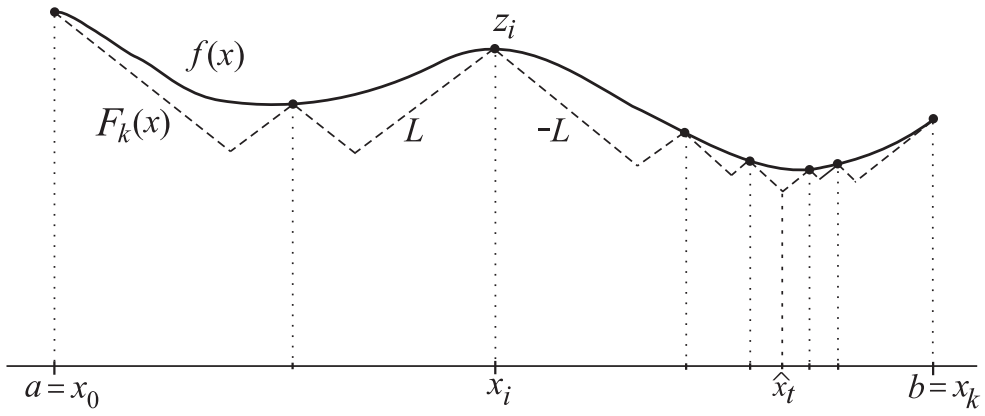


Figure 1: Lower bounding function $F_k(x)$ (dashed line) constructed for a Lipschitz function $f(x)$ (solid line) over $[a, b]$ after having performed $k + 1$ function trials (in this case, $k = 7$).

Once a valid estimate of the Lipschitz constant is known and some function trials are performed, the Lipschitz condition (6) allows us to easily find the lower bounds of a Lipschitz function at different subregions of the search domain D from (5). Let us consider, for the sake of example, a one-dimensional objective function $f(x)$ defined over an interval $[a, b]$ (see Figure 1) that satisfies the Lipschitz condition (6) with a known Lipschitz constant L. If the function values z_i have been obtained at points x_i, $0 \leq i \leq k$ (see black dots on the objective function graph in Figure 1), the following inequality is satisfied over $[a, b]$:

$$f(x) \geq F_k(x) = \max_{0 \leq i \leq k} \{z_i - L|x - x_i|\}, \tag{8}$$

where $F_k(x)$ is a piecewise linear function (called lower bounding or minorant function, see, e.g., [10, 11, 41]; its graph is drawn by dashed line in Figure 1).

A method (e.g., the Piyavskij–Shubert method being one of the first methods in Lipschitz global optimization, see [4, 10, 11, 41, 47]), using in its work this simple but efficient geometric interpretation, iteratively constructs an auxiliary function which bounds the objective function $f(x)$ from below and evaluates $f(x)$ at a point ($\hat{x}_t$ in Figure 1) corresponding to a minimum of the bounding function. This point is easy to find (see, e.g., [4, 10, 11, 41]). The methods of this type form the class of the so-called geometric algorithms that are based on constructing, updating, and improving auxiliary piecewise functions built by using an estimate of the Lipschitz constant L. As shown, e.g., in [9, 11], there exists a strong relationship between the geometric approach and another possible technique for solving the stated problem—the so-called information-statistical approach (see, e.g., [11,48] and also [6,12] for other probabilistic techniques). Together with the geometric ideas of the Piyavskij-Shubert method, it has consolidated foundations of the Lipschitz global optimization.

In order to develop Lipschitz global optimization methods, the Lipschitz constant L from (6) should be estimated. It can be done in several ways. For example, the

Lipschitz constant can be given a priori (see, e.g., [4, 47, 49]). More practical approaches are based on an adaptive estimation of L in the course of the search: such algorithms can use either an adaptive global estimate of the Lipschitz constant (see, e.g., [8, 11, 48, 50]) valid for the whole search domain D, or adaptive local estimates L_i valid only for some subregions $D_i \subset D$ (see, e.g., [9, 11, 51–53]). Estimating local Lipschitz constants during the work of a global optimization algorithm allows one to significantly accelerate the global search. Balancing between local and global information must be performed in an appropriate way (see, e.g., [9, 11, 51]) since an unjustified usage of local information can lead to the loss of the global solution (see, e.g., [40]). Finally, multiple estimates of L can be also used (see, e.g., [9, 24, 29, 32, 54]). We would like to emphasize here that either the Lipschitz constant is given and an algorithm is developed correspondingly, or it is not known but there exist a sufficiently large number of parameters of the considered algorithm ensuring its convergence (convergence properties of the Lipschitz global optimization methods are thoroughly examined, e.g., in [9, 11, 48, 55]).

Considering both the theoretical generality and the application diffusion of the Lipschitz global optimization problem (3), (5), (6), it is used by the authors to mathematically model various real-life decision-making problems (see [9, 11, 13, 56–58]).

3 Some new deterministic approaches in Lipschitz global optimization

In this Section, some innovative deterministic approaches developed by the authors for constructing efficient global optimization techniques are briefly presented. The consolidated success of these ideas, confirmed by important international publications and presentations around the world, allows the authors' group, on the one hand, to develop new optimization approaches over a solid scientific basis, thus eliminating the theoretical faults risks, and, on the other hand, to tackle difficult black-box practical optimization problems (e.g., from control theory, environmental sciences and geological mechanics, electrical engineering and telecommunications, gravitational physics, etc.) with more efficiency with respect to traditionally used techniques.

3.1 Algorithmic framework: 'Divide-the-Best' scheme

Many global optimization algorithms (of both deterministic and stochastic types) have a similar structure. Therefore, several attempts aiming to construct a general framework for describing computational schemes and providing their convergence conditions in a unified manner have been made (see, e.g., [2, 4, 8, 59]). One of the more flexible and robust among such unifying schemes is the 'Divide-the-Best' approach (see [9, 55]), which generalizes both the schemes of adaptive partition [8] and characteristic [9, 11, 59] algorithms, widely used for describing and studying numerical global optimization methods.

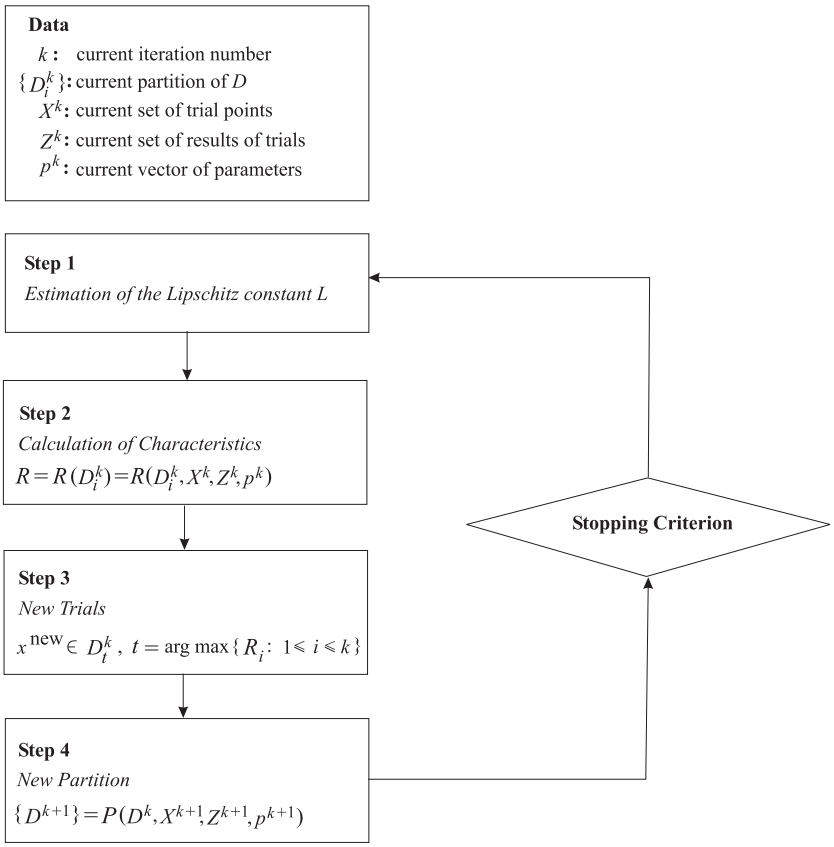


Figure 2: Flow chart of 'Divide-the-Best' scheme.

In this scheme (the flow chart of its generic iteration is reported in Figure 2), given a vector p of the method parameters, an adaptive partition of the admissible region D from (5) into a collection $\{D_i^k\}$ of the finite number of robust subsets D_i^k is considered at each iteration k. The 'merit' (called characteristic) R_i of each subset (see Step 2 in Figure 2) for performing a subsequent, more detailed, investigation (see Steps 3 and 4 in Figure 2) is estimated on the basis of the obtained information X^k, Z^k about the objective function. Several strategies (mainly, in the context of the geometric approach) for selection of a subset for further partitioning (see Step 3 in Figure 2) and for performing this partitioning (by means of an operator P, see Step 4 in Figure 2) are proposed by the authors from a general viewpoint and successfully used for solving practical applications (see, e.g., the references in [9, 10]).

Convergence properties of the 'Divide-the-Best' family for different types of characteristic values and partition operators are studied in [9, 55]. Great attention is given to situations when conditions of global (local) convergence are satisfied not in the whole search domain D, but only in its subregion (or a set of subregions). This can correspond, for example, to Lipschitz global optimization algorithms that work underestimating the Lipschitz constant or which are oriented on using local information in subregions of D (see, e.g., [9, 11, 55]).

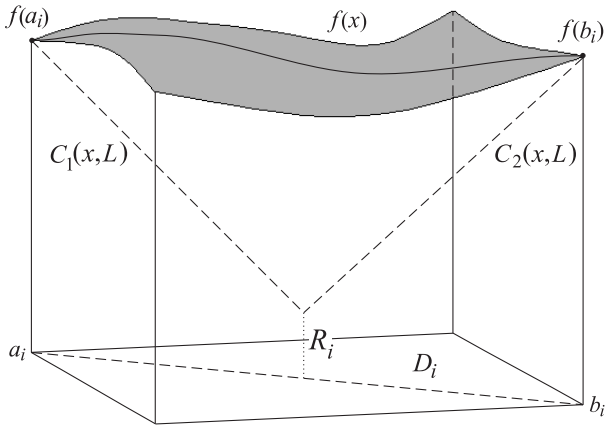


Figure 3: Lower bounding a Lipschitz function $f(x)$ over a hyperinterval D_i in a diagonal algorithm.

3.2 Efficient partition strategy

Regarding the partitioning strategies (partitioning operator P on Step 4 in Figure 2), the main attention of the authors is focused on the diagonal partition strategies (see the references in [8–10, 32]).

In this approach, the initial hyperinterval D from (5) is partitioned into a set of smaller hyperintervals, the objective function is evaluated only at two vertices corresponding to the main diagonal of hyperintervals of the current partition of D (see, e.g., points a_i and b_i of a hyperinterval D_i in Figure 3), and the results of these evaluations are used to select a hyperinterval for the further subdivision. The diagonal approach has a number of attractive theoretical properties and has proved to be efficient in solving applied problems.

First, it allows one to easily perform an extension of efficient one-dimensional global optimization algorithms to the multidimensional case (see, e.g., [9, 10, 32]). In fact, in order to calculate the characteristic R_i of a multidimensional subregion D_i, some one-dimensional characteristics can be used as prototypes. After an appropriate transformation they can be applied to the one-dimensional segment being the main diagonal $[a_i, b_i]$ of the hyperinterval D_i (see Lipschitz-based lower bounding functions C_1 and C_2 in Figure 3).

Second, the diagonal approach is close from the computational point of view to one of the simplest strategies—centre-sampling technique (see, e.g., [25, 29, 42, 49, 54]—but at the same time, the objective function is evaluated at two points of each subregion, providing in this way more information about the function over the subregion than centre-sampling methods.

Different exploration techniques based on various diagonal adaptive partition strategies are analyzed, e.g., in [9, 10, 60]. It is demonstrated that partition strategies traditionally used in the framework of the diagonal approach do not fulfil the requirements

of computational efficiency because of the execution of many redundant trials. Such a redundancy slows down significantly the global search in the case of costly functions.

A new efficient diagonal partition strategy is therefore proposed in [9, 60], that allows one to avoid such a computational redundancy of traditional diagonal schemes. In contrast to these schemes, the new strategy produces regular meshes of the function evaluation points in such a way that one vertex where $f(x)$ is evaluated can belong to several hyperintervals (up to 2^N, N is the problem dimension from (5)). Thus, the time-consuming procedure of the function evaluations is replaced by a significantly faster operation of reading (up to 2^N times) the function values obtained at the previous iterations and saved in a special database (see, e.g., [61, 62]). Hence, the new partition strategy considerably speeds up the search and also leads to saving computer memory. It is particularly important that the advantages of the new strategy become more pronounced when the problem dimension N increases (see, e.g., [9, 32, 50]).

A new scheme for creating fast Lipschitz global optimization algorithms is, thus, introduced by the authors. It relies on the new diagonal partition strategy allowing an efficient extension of popular one-dimensional Lipschitz global optimization algorithms to the multidimensional case. In a sense, this scheme combines the ideas of the diagonal approach and Peano space-filling curves (see, e.g., [11, 48, 63]. Innovative multidimensional diagonal algorithms for solving Lipschitz global optimization problems, based on different ways for obtaining the Lipschitz information and developed in the framework of the efficient diagonal scheme, are proposed by the authors and their convergence properties are analyzed, e.g., in [9, 32, 50].

3.3 Balancing local and global information

Is well known (see, e.g., [2, 4, 9, 11, 40]) that the usage of the only global information on the objective function and constraints during optimization can lead to a slow convergence of algorithms to global minimizers. Therefore, particular attention is paid by the authors to the usage of a local information in global optimization methods, as well. One of the traditional ways in this context (see, e.g., [2, 4, 7]) recommends stopping the global procedure and switching to a local optimization method in order to improve the solution and to accelerate the search during its final phase. Unfortunately, applying this technique can lead to some problems related to the combination of global and local phases, the main problem being that of determining when to stop the global procedure and start the local one. A premature arrest can provoke the loss of the global solution whereas a late one can slow down the search.

Theoretical and experimental results obtained by the authors (see, e.g., [9, 11, 51–53]) confirm that more fruitful approaches can be considered. The first one is the so-called local tuning approach [51] allowing global optimization algorithms to tune their behaviour to the shape of the functions at different parts of the search domain by estimating the local Lipschitz constants.

In fact, the Lipschitz constant L has a significant influence on the convergence speed of the Lipschitz global optimization algorithms and the problem of its specify-

ing is of the great importance. Accepting, for instance, too high a value of L for a concrete objective function means assuming that the function has complicated structure with sharp peaks and narrow attraction regions of minimizers within the whole admissible region. Thus, if the value of L does not correspond to the real behaviour of the objective function, it can lead to a slow convergence of the algorithm to the global minimizer. Global optimization algorithms using in their work a global estimate of L (or some value of L given a priori) do not take into account local information about behaviour of the objective function over every small subregion of D. Therefore, estimating local Lipschitz constants allows one to significantly accelerate the global search (see, e.g., [9, 11, 52, 53]).

The second technique regards a continual local improvement of the current best solution incorporated in a global search procedure (see, e.g., [9, 32, 64]). Particularly, it forces the global optimization method to make a local improvement of the best approximation of the global minimum immediately after a new approximation better than the current one is found. These techniques become even more efficient when information about the objective function derivatives is available (see, e.g., [34, 43]).

3.4 Computational aspects

The advent of parallel computers has created conditions for the elaboration of methods which can accelerate the finding of a solution to many applied problems (see, e.g., the references in [2, 11])). In the case of black-box global optimization, the usage of distributed computation is extremely attractive because the solution to these problems is a very time consuming process: the time taken to evaluate the functions at a point is long and the number of evaluations needed is high even when the most efficient numerical techniques are used.

As a rule, when parallel methods are proposed, either the problem to be solved or some sequential method can have an inherent parallelism and can be used as the source of elaboration. Since the first type of parallelism depends greatly on the specific nature of the problem being solved and must be defined separately for every single case, attention of the authors' group is concentrated on the second way of parallelization (see, e.g., the references in [11]).

The constructed parallel algorithms perform several functions evaluations simultaneously at every iteration: one evaluation at each of the processors of a multiprocessor system the algorithm is implemented on. Such a procedure allows us to accelerate the search of global minima of costly black-box functions. Incomplete account of information about evaluation results, due both to the inter-processors communications and to an eventually heterogeneous loading of processors, can bring up the situation when the parallel algorithm produces the evaluation points in a more dense way in comparison with its sequential prototype, i.e., it can generates redundant evaluations. The new concept of non-redundant parallelism is, therefore, particularly considered by the authors' group during construction and implementation of global search algorithms on high-performance multiprocessor systems (see, e.g., [11, 59, 65]).

A sequential method taken as the basis for parallelization must have elevated estimates of the convergence speed for the class of problems under examination, because there is no sense in using a parallel method if a more rapid sequential one already exists. That is why new sequential algorithms are parallelized only after their extensive study and thorough testing on benchmarks and real-life problems.

For this purpose, a set of test functions is usually taken, problems from this set are solved by the algorithms to be compared, and a conclusion about the efficiency of these algorithms is made on the basis of the obtained numerical results. This approach, being an important instrument for acquiring knowledge about the existing and new global optimization algorithms, presents at the same time some limitations since the conclusions made can be valid only for the selected functions, and their propagation to a more wide set of functions requires particular caution. Testing an algorithm on a relatively large set of test functions can, in a sense, diminish these limitations, but it needs, among other things, the coding of the functions, and it is a tedious and time-consuming job. Therefore, the global optimizers are very interested in simple and powerful software tools realizing test problems. As observed, e.g., in [9,11,12], a well designed testing framework is of the primary importance in identifying the merits of each algorithm and implementation.

To tackle the problem of testing global optimization algorithms systematically, the GKLS-generator described in [66] is proposed by the authors' group. The generator produces several classes of multidimensional and multiextremal test functions with known local and global minima. Each test class provided by the generator includes 100 functions. By changing the user-defined parameters, classes with different properties can be created. For example, fixed dimension of the functions and number of local minima, a more difficult class can be created either by shrinking the attraction region of the global minimizer, or by moving the global minimizer closer to the domain boundary.

The generator is available on the ACM Collected Algorithms (CALGO) database (the CALGO is part of a family of publications produced by the Association for Computing Machinery) and it is also downloadable for free from `http:\\wwwinfo.deis.unical.it\~yaro\GKLS.html`. It has already been downloaded by companies and research organizations from more than 40 countries of the world.

4 Conclusion

To conclude, we would like to report some numerical results obtained by using a new Lipschitz global optimization method proposed by the authors in [32]. In developing this method for solving the problem (3), (5), (6), techniques from the previous Section have been applied. Particularly, it is a multidimensional 'Divide-the-Best' global optimization method that uses in its work multiple estimates of the Lipschitz constant and based on efficient diagonal partitions.

Numerical results performed on the GKLS-generator to compare this new algo-

rithm with two algorithms belonging to the same class of methods for solving the problem (3), (5), (6) — the DIRECT algorithm from [29] and its locally-biased modification DIRECTl from [54] — are presented here, as described in [32]. As known, both of these methods are widely used in solving practical engineering problems (see, e.g., the references in [2, 24, 32]).

Eight GKLS classes of continuously differentiable test functions of dimensions $N = 2$, 3, 4, and 5 have been used. For each dimension, both a 'hard' and a 'simple' classes have been considered. The difficulty of a class was increased either by decreasing the radius of the attraction region of the global minimizer, or by decreasing the distance from the global minimizer x^* to the domain boundaries.

The global minimizer $x^* \in D$ was considered to be found when the algorithm generated a trial point x' inside a hypercube with a vertex x^* and the volume smaller than the volume of the initial hypercube $D = [a, b]$ multiplied by an accuracy coefficient Δ, $0 < \Delta \leq 1$, i.e.,

$$|x'(j) - x^*(j)| \leq \sqrt[N]{\Delta}(b(j) - a(j)) \tag{9}$$

for all i, $1 \leq j \leq N$, where N is from (5). The algorithm stopped either when the maximal number of trials equal to 1 000 000 was reached, or when condition (9) was satisfied.

In view of the high computational complexity of each trial of the objective function, the methods were compared in terms of the number of evaluations of $f(x)$ required to satisfy condition (9). The number of hyperintervals generated until condition (9) is satisfied, was taken as the second criterion for comparison of the methods. This number reflects indirectly degree of qualitative examination of D during the search for a global minimum (see, e.g., [9, 32, 50]).

Results of numerical experiments with eight GKLS tests classes are reported in Tables 1–2. These tables show, respectively, the maximal number of trials and the corresponding number of generated hyperintervals required for satisfying condition (9) for a half of the functions of a particular class (columns "50%") and for all 100 function of the class (columns "100%"). The notation "$>$ 1 000 000 (k)" means that after 1 000 000 trials the method under consideration was not able to solve k problems.

Note that on a half of test functions from each class (which were simple for each method with respect to the other functions of the class) the new algorithm manifested a good performance with respect to DIRECT and DIRECTl in terms of the number of generated trial points (see Table 1). When all functions were taken in consideration, the number of trials produced by the new algorithm was significantly fewer in comparison with two other methods (see columns "100%" of Table 1), providing at the same time a good examination of the admissible region (see Table 2).

As it can be seen from Tables 1 – 2, the new method [32] demonstrates a quite satisfactory performance with respect to popular DIRECT [29] and DIRECTl [54] methods when multidimensional functions with a really complex structure are minimized.

This method not only has manifested a high performance on a large set of tests, but has been also successfully applied for solving real world global optimization prob-

Table 1: Number of trial points for 800 GKLS test functions.

N	Δ	Class	50%			100%		
			DIRECT	DIRECTl	New	DIRECT	DIRECTl	New
2	10^{-4}	simple	111	152	166	1159	2318	403
2	10^{-4}	hard	1062	1328	613	3201	3414	1809
3	10^{-6}	simple	386	591	615	12507	13309	2506
3	10^{-6}	hard	1749	1967	1743	>1000000 (4)	29233	6006
4	10^{-6}	simple	4805	7194	4098	>1000000 (4)	118744	14520
4	10^{-6}	hard	16114	33147	15064	>1000000 (7)	287857	42649
5	10^{-7}	simple	1660	9246	3854	>1000000 (1)	178217	33533
5	10^{-7}	hard	55092	126304	24616	>1000000 (16)	>1000000 (4)	93745

Table 2: Number of hyperintervals for 800 GKLS test functions.

N	Δ	Class	50%			100%		
			DIRECT	DIRECTl	New	DIRECT	DIRECTl	New
2	10^{-4}	simple	111	152	269	1159	2318	685
2	10^{-4}	hard	1062	1328	1075	3201	3414	3307
3	10^{-6}	simple	386	591	1545	12507	13309	6815
3	10^{-6}	hard	1749	1967	5005	>1000000	29233	17555
4	10^{-6}	simple	4805	7194	15145	>1000000	118744	73037
4	10^{-6}	hard	16114	33147	68111	>1000000	287857	211973
5	10^{-7}	simple	1660	9246	21377	>1000000	178217	206323
5	10^{-7}	hard	55092	126304	177927	>1000000	>1000000	735945

lems. For example, its application to a control theory problem has been considered in [56]. This problem regards global tuning of fuzzy power system stabilizers present in a multi-machine power system in order to damp the power system oscillations. Power system stabilizers with conventional industry structure are extensively used in modern power systems as an efficient means of damping power. Traditionally their parameters are determined by local tuning procedure based on a single-machine infinite-bus system in which the effects of inter-machine and inter-area dynamics are usually ignored. Heuristic methods (like genetic algorithms) are usually used for their optimizing that often leads to very rough solutions (see, e.g., the references in [56]). To improve overall system dynamic performance, novel global optimization techniques have been therefore applied by the authors' group in [56].

In Figure 4, the graph that illustrates the best solution (the axis of ordinates) obtained by a particular genetic algorithm (often used by engineers from the control field) and by the method [32] after a number of simulations (the axis of abscissas) is reported. It can be seen that the global optimization method proposed by the authors spent more function trials (namely, 284 trials) than the genetic algorithm at the initial iterations. This phase corresponds to the initial exploration of the search domain and it is necessary for all global optimization techniques. On the initial phase of the work (less than 300 trials) the genetic algorithm has found local solutions to the problem better than those found by the new method, but far from the final global solution

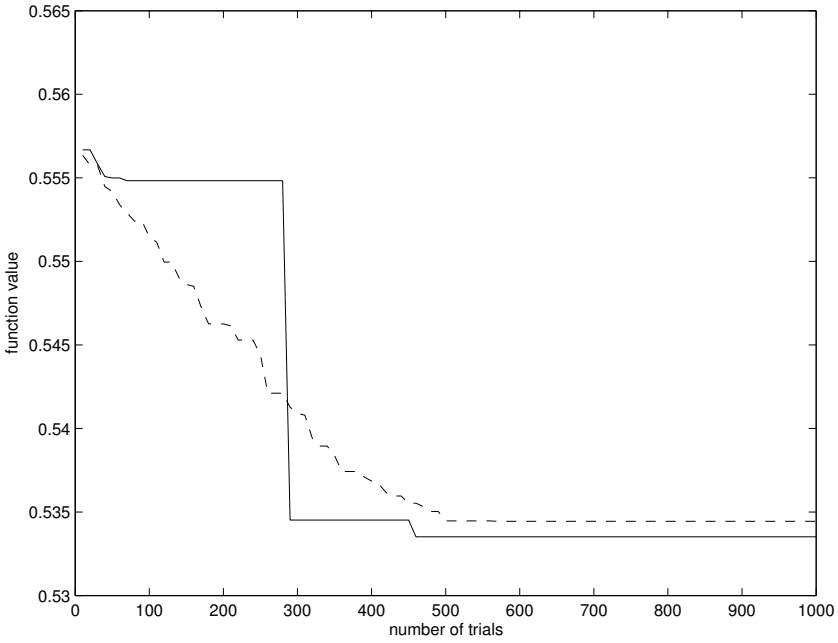


Figure 4: Solutions to the problem of global tuning fuzzy power system stabilizers (see [56]) obtained by applying the method [32] based on the authors' techniques (solid line) and by a traditionally used genetic approach (dashed line).

($f^* \approx 0.533$). However, it is more important and should be underlined that the new method has determined a solution to the problem very close to the global optimal one (as demonstrated in [56]) in almost half of the simulations with respect to the genetic algorithm (284 trials for the new method and 500 for the genetic algorithm). Moreover, it has found an attraction region of a new minimizer with a much better solution to the problem (see the graph jump in Figure 4 around 450 trials) than that found by the genetic approach. Thus, when a reasonable limit of function trials is given, the considered method [32] can determine a good estimate of the global solution to the studied control theory problem faster than the traditionally used genetic techniques.

Therefore, new global optimization techniques briefly presented in this paper can provide the scientists and engineers with comprehensive and powerful tools for successful solving challenging decision-making problems from different real-life application areas, which are characterized by black-box multiextremal and hard to evaluate functions.

Acknowledgements

The authors work was supported by the grant 1960.2012.9 awarded by the President of the Russian Federation for the leading research groups, as well as by the grant 11-01-00682-a awarded by the Russian Foundation for Fundamental Research. The au-

thors thank also the Italian Inter-University Consortium for the Application of Super-Computing for Universities and Research (CASPUR), project "Advanced techniques for global optimization: Numerical methods and applications" in the framework of the "HPC Grant 2011 on GPU cluster". The participation of the first author at the CST2012 has been supported by the Italian National Group for Scientific Computation of the National Institute for Advanced Mathematics 'F. Severi' (INdAM–GNCS).

References

[1] C. Audet, P. Hansen, G. Savard (Editors), *Essays and Surveys in Global Optimization*, GERAD 25th Anniversary. Springer, New York, 2005.

[2] C.A. Floudas, P.M. Pardalos (Editors), *Encyclopedia of Optimization (6 Volumes)*, Springer, 2nd edition, 2009.

[3] I.E. Grossmann (Editor), *Global Optimization in Engineering Design*, Kluwer Academic Publishers, Dordrecht, 1996.

[4] R. Horst, P.M. Pardalos (Editors), *Handbook of Global Optimization*, Volume 1, Kluwer Academic Publishers, Dordrecht, 1995.

[5] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer–Verlag, Berlin, 3rd edition, 1996.

[6] J. Mockus, *A Set of Examples of Global and Discrete Optimization: Applications of Bayesian Heuristic Approach*, Kluwer Academic Publishers, Dordrecht, 2000.

[7] P.M. Pardalos, H.E. Romeijn (Editors), *Handbook of Global Optimization*, Volume 2, Kluwer Academic Publishers, Dordrecht, 2002.

[8] J.D. Pintér, *Global Optimization in Action (Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications)*, Kluwer Academic Publishers, Dordrecht, 1996.

[9] Y.D. Sergeyev, D.E. Kvasov, *Diagonal Global Optimization Methods*, FizMatLit, Moscow, 2008, MAIK 2008 Prize for the best scientific monograph published in Russian.

[10] Y.D. Sergeyev, D.E. Kvasov, "Lipschitz global optimization", in J.J. Cochran (Editor), *Wiley Encyclopedia of Operations Research and Management Science*, Volume 4, pages 2812–2828. Wiley, New York, 2011.

[11] R.G. Strongin, Y.D. Sergeyev, *Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms*, Kluwer Academic Publishers, Dordrecht, 2000.

[12] A.A. Zhigljavsky, A. Žilinskas, *Stochastic Global Optimization*, Springer, New York, 2008.

[13] V.I. Golubev, D.E. Kvasov, I.E. Kvasov, "Identification of seismogeological cracks location by using numerical global optimization methods", in *Proc. of the 53rd MIPT Scientific Conference "Recent Advances in Basic and Applied Sciences"*, Volume VII (2), pages 20–22. MIPT Press, Moscow, November 2010.

[14] V.B. Leviant, I.B. Petrov, F.B. Chelnokov, I.Y. Antonova, "Nature of the scattered seismic response from zones of random clusters of cavities and fractures in

a massive rock", *Geophysical Prospecting*, 55(4): 507–524, 2007.

[15] I.B. Petrov, A.S. Kholodov, "Numerical investigation of certain dynamical problems of the mechanics of a deformable solid body by the grid-characteristic method", *USSR Comput. Math. Math. Phys.*, 24(3): 61–73, 1984.

[16] I.E. Kvasov, I.B. Petrov, "High-performance computer simulation of wave processes in geological media in seismic exploration", *Comput. Math. Math. Phys.*, 52(2): 302–313, 2012.

[17] L. Ljung, *System Identification: Theory for the User*, PTR Prentice Hall, Upper Saddle River, N.J., 2nd edition, 1999.

[18] K. Schittkowski, *Numerical Data Fitting in Dynamical Systems: A Practical Introduction with Applications and Software*, Kluwer Academic Publishers, Dordrecht, 2002.

[19] M.K. Sen, P.L. Stoffa, *Global Optimization Methods in Geophysical Inversion*, Elsevier Science B.V., Amsterdam, 1995.

[20] G.E. Stavroulakis, *Inverse and Crack Identification Problems in Engineering Mechanics*, Kluwer Academic Publishers, Dordrecht, 2001.

[21] G. Corliss, C. Faure, A. Griewank, L. Hascoet, U. Naumann (Editors), *Automatic Differentiation of Algorithms: From Simulation to Optimization*, Springer, New York, 2002.

[22] T. Csendes (Editor), *Developments in Reliable Computing*, Kluwer Academic Publishers, Dordrecht, 2000.

[23] R.B. Kearfott, *Rigorous Global Search: Continuous Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[24] C.T. Kelley, *Iterative Methods for Optimization*, SIAM Publications, Philadelphia, 1999.

[25] A.R. Conn, K. Scheinberg, L.N. Vicente, *Introduction to Derivative-Free Optimization*, SIAM, Philadelphia, USA, 2009.

[26] T.G. Kolda, R.M. Lewis, V. Torczon, "Optimization by direct search: New perspectives on some classical and modern methods", *SIAM Rev.*, 45(3): 385–482, 2003.

[27] G. Liuzzi, S. Lucidi, M. Sciandrone, "Sequential penalty derivative-free methods for nonlinear constrained optimization", *SIAM J. Optim.*, 20(5): 2614–2635, 2010.

[28] A.L. Custódio, J.F. Aguilar Madeira, A.I.F. Vaz, L.N. Vicente, "Direct multisearch for multiobjective optimization", *SIAM J. Optim.*, 21(3): 1109–1140, 2011.

[29] D.R. Jones, C.D. Perttunen, B.E. Stuckman, "Lipschitzian optimization without the Lipschitz constant", *J. Optim. Theory Appl.*, 79(1): 157–181, 1993.

[30] D.R. Jones, M. Schonlau, W.J. Welch, "Efficient global optimization of expensive black-box functions", *J. Global Optim.*, 13(4): 455–492, 1998.

[31] R.G. Regis, C.A. Shoemaker, "Constrained global optimization of expensive black box functions using radial basis functions", *J. Global Optim.*, 31(1): 153–171, 2005.

[32] Y.D. Sergeyev, D.E. Kvasov, "Global search based on efficient diagonal parti-

tions and a set of Lipschitz constants", *SIAM J. Optim.*, 16(3): 910–937, 2006.

[33] D. Di Serafino, G. Liuzzi, V. Piccialli, F. Riccio, G. Toraldo, "A modified DIviding RECTangles algorithm for a problem in astrophysics", *J. Optim. Theory Appl.*, 151(1): 175–190, 2011.

[34] D.E. Kvasov, Y.D. Sergeyev, "Lipschitz gradients for global optimization in a one-point-based partitioning scheme", *J. Comput. Appl. Math.*, 236(16): 4042–4054, 2012.

[35] C.A. Floudas, *Deterministic Global Optimization: Theory, Algorithms, and Applications*, Kluwer Academic Publishers, Dordrecht, 2000.

[36] M. Tawarmalani, N.V. Sahinidis, *Convexification and Global Optimization in Continuous and Mixed–Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*, Kluwer Academic Publishers, Dordrecht, 2002.

[37] A.A. Zhigljavsky, *Theory of Global Random Search*, Kluwer Academic Publishers, Dordrecht, 1991.

[38] J.H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, USA, 1975.

[39] J.J. Schneider, S. Kirkpatrick, *Stochastic Optimization*, Springer, Berlin, 2006.

[40] C.P. Stephens, W. Baritompa, "Global optimization requires global information", *J. Optim. Theory Appl.*, 96(3): 575–588, 1998.

[41] D.E. Kvasov, Y.D. Sergeyev, "Univariate geometric Lipschitz global optimization algorithms", *Numer. Algebra Contr. Optim.*, 2(1): 69–90, 2012.

[42] Y.G. Evtushenko, *Numerical Optimization Techniques*, Translations Series in Mathematics and Engineering. Springer–Verlag, Berlin, 1985.

[43] D.E. Kvasov, Y.D. Sergeyev, "A univariate global search working with a set of Lipschitz constants for the first derivative", *Optim. Lett.*, 3(2): 303–318, 2009.

[44] K.A. Barkalov, R.G. Strongin, "A global optimization technique with an adaptive order of checking for constraints", *Comput. Math. Math. Phys.*, 42(9): 1289–1300, 2002.

[45] Y.D. Sergeyev, D. Famularo, P. Pugliese, "Index Branch-and-Bound Algorithm for Lipschitz univariate global optimization with multiextremal constraints", *J. Global Optim.*, 21(3): 317–341, 2001.

[46] Y.D. Sergeyev, D.E. Kvasov, F.M.H. Khalaf, "A one-dimensional local tuning algorithm for solving GO problems with partially defined constraints", *Optim. Lett.*, 1(1): 85–99, 2007.

[47] Y.G. Evtushenko, "Numerical methods for finding global extrema (Case of a non-uniform mesh)", *USSR Comput. Math. Math. Phys.*, 11(6): 38–54, 1971.

[48] R.G. Strongin, *Numerical Methods in Multiextremal Problems (Information-Statistical Algorithms)*, Nauka, Moscow, 1978, In Russian.

[49] Y.G. Evtushenko, M.A. Posypkin, "An application of the nonuniform covering method to global optimization of mixed integer nonlinear problems", *Comput. Math. Math. Phys.*, 51(8): 1286–1298, 2011.

[50] D.E. Kvasov, Y.D. Sergeyev, "Multidimensional global optimization algorithm based on adaptive diagonal curves", *Comput. Math. Math. Phys.*, 43(1): 40–56, 2003.

[51] Y.D. Sergeyev, "An information global optimization algorithm with local tuning", *SIAM J. Optim.*, 5(4): 858–870, 1995.

[52] Y.D. Sergeyev, "Univariate global optimization with multiextremal nondifferentiable constraints without penalty functions", *Comput. Optim. Appl.*, 34 (2): 229–248, 2006.

[53] D.E. Kvasov, C. Pizzuti, Y.D. Sergeyev, "Local tuning and partition strategies for diagonal GO methods", *Numer. Math.*, 94(1): 93–106, 2003.

[54] J.M. Gablonsky, C.T. Kelley, "A locally-biased form of the DIRECT algorithm", *J. Global Optim.*, 21(1): 27–37, 2001.

[55] Y.D. Sergeyev, "On convergence of "Divide the Best" global optimization algorithms", *Optimization*, 44(3): 303–325, 1998.

[56] D.E. Kvasov, D. Menniti, A. Pinnarelli, Y.D. Sergeyev, N. Sorrentino, "Tuning fuzzy power-system stabilizers in multi-machine systems by global optimization algorithms based on efficient domain partitions", *Electr. Power Syst. Res.*, 78(7): 1217–1229, 2008.

[57] Y.D. Sergeyev, P. Daponte, D. Grimaldi, A. Molinaro, "Two methods for solving optimization problems arising in electronic measurements and electrical engineering", *SIAM J. Optim.*, 10(1): 1–21, 1999.

[58] D. Famularo, P. Pugliese, Y.D. Sergeyev, "A global optimization technique for checking parametric robustness", *Automatica*, 35(9): 1605–1611, 1999.

[59] V.A. Grishagin, Y.D. Sergeyev, R.G. Strongin, "Parallel characteristic algorithms for solving problems of global optimization", *J. Global Optim.*, 10(2): 185–206, 1997.

[60] Y.D. Sergeyev, "An efficient strategy for adaptive partition of N-dimensional intervals in the framework of diagonal algorithms", *J. Optim. Theory Appl.*, 107 (1): 145–168, 2000.

[61] D.E. Kvasov, "Multidimensional Lipschitz global optimization based on efficient diagonal partitions", *4OR – Quart. J. Oper. Res.*, 6(4): 403–406, 2008.

[62] D.E. Kvasov, "Diagonal numerical methods for solving Lipschitz global optimization problems", *Bollettino U.M.I.*, I (Serie IX)(3): 857–871, 2008.

[63] D. Lera, Y.D. Sergeyev, "Lipschitz and Hölder global optimization using space-filling curves", *Appl. Numer. Math.*, 60(1–2): 115–129, 2010.

[64] D. Lera, Y.D. Sergeyev, "An information global minimization algorithm using the local improvement technique", *J. Global Optim.*, 48(1): 99–112, 2010.

[65] R.G. Strongin, Y.D. Sergeyev, "Global optimization: Fractal approach and non-redundant parallelism", *J. Global Optim.*, 27(1): 25–50, 2003.

[66] M. Gaviano, D.E. Kvasov, D. Lera, Y.D. Sergeyev, "Algorithm 829: Software for generation of classes of test functions with known local and global minima for global optimization", *ACM Trans. Math. Software*, 29(4): 469–480, 2003.