Derivative-Free Local Tuning and Local Improvement Techniques Embedded in the Univariate Global Optimization

Yaroslav D. Sergeyev   Marat S. Mukhametzhanov   Dmitri E. Kvasov
Daniela Lera

Abstract

Geometric and information frameworks for constructing global optimization algorithms are considered, and several new ideas to speed up the search are proposed. The accelerated global optimization methods automatically realize a local behavior in the promising subregions without the necessity to stop the global optimization procedure. Moreover, all the trials executed during the local phases are used also in the course of the global ones. The resulting geometric and information global optimization methods have a similar structure and a smart mixture of new, and traditional computational steps leads to 22 different global optimization algorithms. All of them are studied and numerically compared on three test sets including 120 benchmark functions and 4 applied problems.

1 Introduction

In this paper, black-box Lipschitz global optimization problems are considered in their univariate statement. Problems of this kind attract a great attention of the global optimization community. This happens because, first, there exists a huge number of real-life applications where it is necessary to solve univariate global optimization problems (see, e. g., [15,18,19,23,43,55]). This kind of problems is often encountered in scientific and engineering applications (see, e. g., [5, 24, 28, 33, 44, 47, 53]), and, in particular, in statistical applications (see, e. g., [1, 2, 10, 13, 14, 57]) or in electrical engineering optimization problems (see, e. g., [3, 6, 29, 41, 42, 54]). On the other hand, it is important to study one-dimensional methods because they can be successfully generalized in several ways. For instance, they can be extended to the multi-dimensional case by numerous schemes (see, for example, one-point based, diagonal, simplicial, space-filling curves, and other popular approaches in [4, 27, 30–32, 49]). Another possible generalization consists in developing methods for solving problems, where the first derivative of the objective function satisfies also the Lipschitz condition with an unknown constant (see, e. g., [12, 21, 26, 38, 45, 46]).

In the seventies of the XXth century two algorithms for solving the above mentioned problems have been proposed in [35, 52]. The first method was introduced by Piyavskij (see also [50]) by using geometric ideas (based on the Lipschitz condition) and an a priori given overestimate of the Lipschitz constant for

the objective function. The method [35] constructs a piecewise linear auxiliary function, being a minorant for the objective function, that is adaptively improved during the search. The latter algorithm [51, 52] was introduced by Strongin, who developed a statistical model that allowed him to calculate probabilities of locating global minimizers within each of the subintervals of the search interval taken into consideration. Moreover, this model provided a dynamically computed estimate of the Lipschitz constant during the process of optimization. Both the methods became sources of multiple generalizations and improvements (see, e. g., [7, 30, 32, 49, 56]) giving rise to classes of *geometric* and *information* global optimization methods.

Very often in global optimization (see, e. g., [8, 17, 58]) local techniques are used to accelerate the global search, and frequently global and local searches are realized by different methods having completely alien structures. Such a combination introduces at least two inconveniences. First, evaluations of the objective function (called hereinafter *trials*), executed by a local search procedure, are not used typically in the subsequent phases of the global search. Results of only some of these trials (for instance, the current best found value) can be used, and the other ones are not taken into consideration. Second, there arises the necessity to introduce both a rule that stops the global phase and starts the local one, and a rule that stops the local phase and decides whether it is necessary to re-start the global search. Clearly, a premature stop of a global phase of the search can lead to the loss of the global solution, while a late stop of the global phase can slow down the search.

In this paper, both frameworks, geometric and information, are taken into consideration and a number of derivative-free techniques that were proposed to accelerate the global search are studied and compared. Several new ideas that can be used to speed up the search both in the framework of geometric and information algorithms are introduced. All the acceleration techniques have the advantage to get over both the difficulties mentioned above, namely:

– the accelerated global optimization methods automatically realize a local behavior in the promising subregions without the necessity to stop the global optimization procedure;

– all the trials executed during the local phases are used also in the course of the global ones.

It should be emphasized that the resulting geometric and information global optimization methods have a similar structure, and a smart mixture of new and traditional computational steps leads to 22 different global optimization algorithms. All of them are studied and compared on three sets of tests: the widely used set of 20 test functions taken from [16] and listed in Appendix A; 100 randomly generated functions from [34]; four functions arising in practical problems (see Appendix B).

2 Acceleration Techniques

The considered global optimization problem can be formulated as follows:

$$f^* := f(x^*) = \min f(x), \ \ x \in [a, b], \tag{1}$$

where the function $f(x)$ satisfies the Lipschitz condition over the interval $[a, b]$:

$$|f(x_1) - f(x_2)| \le L|x_1 - x_2|, \ \ x_1, x_2 \in [a, b], \tag{2}$$

with the Lipschitz constant $L, 0 < L < \infty$. It is supposed that the objective function $f(x)$ can be multiextremal, non-differentiable; black-box; with an unknown Lipschitz constant L; and evaluation of $f(x)$ even at one point is a time-consuming operation.
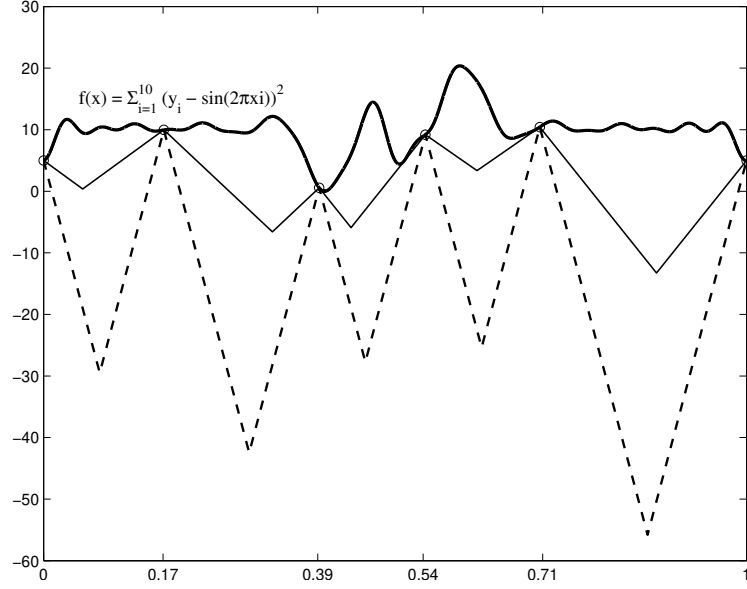
Figure 1 An auxiliary function (solid thin line) and a minorant function (dashed line) for a Lipschitz function $f(x)$ over $[a, b]$, constructed by using estimates of local Lipschitz constants and by using the global Lipschitz constant, respectively (trial values are circled).

As mentioned in Introduction, the geometric and information frameworks are taken into consideration in this paper. The original geometric and information methods, apart the origins of their models, have the following important difference. Piyavskij's method requires for its correct work an overestimate of the value L, that usually is hard to get in practice. In contrast, the information method of Strongin adaptively estimates L during the search. As it was shown in [36, 37] for both the methods, these two strategies for obtaining the Lipschitz information can be substituted by the so-called "local tuning approach". In fact, the original methods of Piyavskij and Strongin use estimates of the global constant L during their work (the term "global" means that the same estimate is used over the whole interval $[a, b]$). However, the global estimate can provide a poor information about the behavior of the objective function $f(x)$ over every small subinterval $[x_{i-1}, x_i] \subset [a, b]$. In fact, when the local Lipschitz constant related to the interval $[x_{i-1}, x_i]$ is significantly smaller than the global constant L, then the methods using only this global constant or its estimate can work slowly over such an interval (see, e. g., [22, 37, 44, 54]).

In Fig. 1, an example of the auxiliary function for a Lipschitz function $f(x)$ over $[a, b]$ constructed by using estimations of local Lipschitz constants over subintervals of $[a, b]$ is shown by a solid thin line; a minorant function for $f(x)$ over $[a, b]$ constructed by using an overestimate of the global Lipschitz constant is represented by a dashed line. Note that the former piecewise function estimates the behavior of $f(x)$ over $[a, b]$ more accurately than the latter one, especially over subintervals where the corresponding local Lipschitz constants are smaller than the global one.

The *local tuning technique* proposed in [36, 37] adaptively estimates *local* Lipschitz constants at different subintervals of the search region during the course of the optimization process. Estimates l_i of local Lipschitz

constants L_i are computed for each interval $[x_{i-1}, x_i]$, $i = 2, ..., k$, as follows:

$$l_i = r \cdot \max\{\lambda_i, \gamma_i, \xi\}, \tag{3}$$

where

$$\lambda_i = \max\{H_{i-1}, H_i, H_{i+1}\}, \quad i = 2, ..., k, \tag{4}$$

$$H_i = \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}}, \quad i = 2, ..., k, \tag{5}$$

$$H^k = \max\{H_i : i = 2, ..., k\}. \tag{6}$$

Here, $z_i = f(x_i), i = 1, ..., k$, i.e., values of the objective function calculated at the previous iterations at the trial points $x_i, i = 1, ..., k$, (when $i = 2$ and $i = k$ only H_2, H_3, and H_{k-1}, H_k, should be considered, respectively). The value γ_i is calculated as follows:

$$\gamma_i = H^k \frac{(x_i - x_{i-1})}{X^{max}}, \tag{7}$$

with H^k from (6) and

$$X^{max} = \max\{x_i - x_{i-1} : i = 2, ..., k\}. \tag{8}$$

Let us give an explanation of these formulae. The parameter $\xi > 0$ from (3) is a small number, that is required for a correct work of the local tuning at initial steps of optimization, where it can happen that $\max\{\lambda_i, \gamma_i\} = 0$; $r > 1$ is the reliability parameter. The two components, λ_i and γ_i, are the main players in (3). They take into account, respectively, the local and the global information obtained during the previous iterations. When the interval $[x_{i-1}, x_i]$ is large, the local information represented by λ_i can be not reliable and the global part γ_i has a decisive influence on l_i thanks to (3) and (7). In this case $\gamma_i \to H^k$, namely, it tends to the estimate of the global Lipschitz constant L. In contrast, when $[x_{i-1}, x_i]$ is small, then the local information becomes relevant, the estimate γ_i is small for small intervals (see (7)), and the local component λ_i assumes the key role. Thus, the local tuning technique automatically balances the global and the local information available at the current iteration. It has been proved for a number of global optimization algorithms that the usage of the local tuning can accelerate the search significantly (see [20, 37, 38, 40, 44, 47, 48, 54]). This local tuning strategy will be called *"Maximum" Local Tuning* hereinafter.

Recently, a new local tuning strategy called hereinafter *"Additive" Local Tuning* has been proposed in [9, 11, 53] for certain information algorithms. It proposes to use the following additive convolution instead of (3):

$$l_i = r \cdot \max\{\frac{1}{2}(\lambda_i + \gamma_i), \xi\}, \tag{9}$$

where r, ξ, λ_i, and γ_i have the same meaning as in (3). The first numerical examples executed in [9, 53] have shown a very promising performance of the "Additive" Local Tuning. These results induced us to execute in the present paper a broad experimental testing and a theoretical analysis of the "Additive" Local Tuning. In particular, geometric methods using this technique are proposed here (remind that the authors of [9, 53] have introduced it in the framework of information methods only). During our study some features suggesting a careful usage of this technique have been discovered, especially in cases where it is applied to geometric global optimization methods.

In order to start our analysis of the "Additive" Local Tuning, let us remind (see, e. g., [33, 35, 44, 49, 52, 54]) that in both the geometric and the information univariate algorithms, an interval $[x_{t-1}, x_t]$ is chosen in a

certain way at the $(k + 1)$-th iteration of the optimization process and a new trial point, x^{k+1}, where the $(k + 1)$-th evaluation of $f(x)$ is executed, is computed as follows:

$$x^{k+1} = \frac{x_t + x_{t-1}}{2} - \frac{z_t - z_{t-1}}{2l_t}. \tag{10}$$

For a correct work of this kind of algorithms it is necessary that x^{k+1} is such that $x^{k+1} \in]x_{t-1}, x_t[$. It is easy to see that the necessary condition for this inclusion is $l_t > H_t$, where H_t is calculated following (5). Notice that l_t is obtained by using (9), where the sum of two addends plays the leading role. Since the estimate γ_i is calculated as shown in (7), it can be very small for small intervals, creating so the possibility of occurrence of the situation $l_t \leq H_t$ and, consequently, $x^{k+1} \notin]x_{t-1}, x_t[$. Obviously, by increasing the value of the parameter r this situation can be easily avoided and the method should be re-started. In fact, in information algorithms where $r \geq 2$ is usually used, this risk is less pronounced, while in geometric methods where $r > 1$ is applied it becomes more probable. On the other hand, it is well known in Lipschitz global optimization (see, e. g., [33, 44, 49, 54]) that increasing the parameter r can slow down the search. In order to understand better the functioning of the "Additive" Local Tuning, it is broadly tested in Section 4 together with other competitors.

The analysis provided above shows that the usage of the "Additive" Local Tuning can become tricky in some cases. In order to avoid the necessity to check the satisfaction of the condition $x^{k+1} \in]x_{t-1}, x_t[$ at each iteration, we propose a new strategy called the *"Maximum-Additive" Local Tuning* where, on the one hand, this condition is satisfied automatically and, on the other hand, advantages of both the local tuning techniques described above are incorporated in the unique strategy. This local tuning strategy calculates the estimate l_i of the local Lipschitz constants as follows:

$$l_i = r \cdot \max\{H_i, \frac{1}{2}(\lambda_i + \gamma_i), \xi\}, \tag{11}$$

where r, ξ, H_i, λ_i, and γ_i have the usual meaning. It can be seen from (11) that this strategy both maintains the additive character of the convolution and satisfies condition $l_i > H_i$. The latter condition provides that in case the interval $[x_{i-1}, x_i]$ is chosen for subdivision (i. e., $t := i$ is assigned), the new trial point x^{k+1} will belong to $]x_{t-1}, x_t[$. Notice that in (11) the equal usage of the local and global estimate is applied. Obviously, a more general scheme similar to (9) and (11) can be used, where $\frac{1}{2}$ is substituted by different weights for the estimates λ_i and γ_i, for example, as follows:

$$l_i = r \cdot \max\{H_i, \frac{\lambda_i}{r} + \frac{r-1}{r}\gamma_i, \xi\}$$

(r, H_i, λ_i, γ_i, and ξ are as in (11)).

Let us now present another acceleration idea, that will be introduced formally in next Section. It consists in the following observation, related to global optimization problems with a fixed budget of possible evaluations of the objective function $f(x)$, i. e., when only, for instance, 100 or $1\,000\,000$ evaluations of $f(x)$ are allowed. In these problems, it is necessary to obtain the best possible value of $f(x)$ as soon as possible. Suppose that f_k^* is the best value (the record value) obtained after k iterations. If a new value $f(x^{k+1}) < f_k^*$ has been obtained, then it can make sense to try to improve this value locally, instead of continuing the usual global search phase. As was already mentioned, traditional methods stop the global procedure and start a local descent: trials executed during this local phase are not then used by the global search since the local method has usually a completely different nature.

Here, we propose two *local improvement techniques*, the *"optimistic"* and the *"pessimistic"* one, that perform the local improvement within the global optimization scheme. The optimistic method alternates

the local steps with the global ones and, if during the local descent a new promising local minimizer is not found, then the global method stops when a local stopping rule is satisfied. The pessimistic strategy does the same until the satisfaction of the required accuracy on the local phase, and then switches to the global phase where the trials performed during the local phase are also taken into consideration.

3   Numerical Methods and Their Convergence Study

All the methods described in this Section have a similar structure and belong to the class of "Divide the Best" global optimization algorithms introduced in [39] (see also [44]; for methods using the "Additive" Local Tuning this holds if the parameter r is such that $l_{i(k)} > rH_{i(k)}$ for all i and k). The algorithms differ in the following:

– methods are either geometric or information;

– methods differ in the way the Lipschitz information is used: an a priori estimate, a global estimate, and a local tuning;

– in cases where a local tuning is applied, methods use 3 different strategies: Maximum, Additive, and Maximum-Additive;

– in cases where a local improvement is applied, methods use either the optimistic or the pessimistic strategy.

Let us describe the General Scheme (GS) of the methods used in this work. A concrete algorithm will be obtained by specifying one of the possible implementations of Steps 2–4 in this GS.

Step 0. *Initialization.* Execute first two trials at the points a and b, i.e., $x^1 := a$, $z^1 := f(a)$ and $x^2 := b$, $z^2 := f(b)$. Set the iteration counter $k := 2$.

Let $flag$ be the local improvement switch to alternate global search and local improvement procedures; set its initial value $flag := 0$. Let $i_{\min}$ be an index (being constantly updated during the search) of the current record point, i.e., $z_{i_{\min}} = f(x_{i_{\min}}) \leq f(x_i)$, $i = 1, \ldots, k$ (if the current minimal value is attained at several trial points, then the smallest index is accepted as $i_{\min}$).

Suppose that $k \geq 2$ iterations of the algorithm have already been executed. The iteration $k+1$ consists of the following steps.

Step 1. *Reordering.* Reorder the points $x^1, \ldots, x^k$ (and the corresponding function values $z^1, \ldots, z^k$) of previous trials by subscripts so that

$$a = x_1 < \ldots < x_k = b, \quad z_i = f(x_i), \; i = 1, \ldots, k.$$

Step 2. *Estimates of the Lipschitz constant.* Calculate the current estimates l_i of the Lipschitz constant for each subinterval $[x_{i-1}, x_i]$, $i = 2, \ldots, k$, in one of the following ways.

Step 2.1. *A priori given estimate.* Take an a priori given estimate $\hat{L}$ of the Lipschitz constant for the whole interval $[a, b]$, i.e., set $l_i := \hat{L}$.

Step 2.2. *Global estimate.* Set $l_i := r \cdot \max\{H^k, \xi\}$, where r and ξ are two parameters with $r > 1$ and ξ sufficiently small, H^k is from (6).

Step 2.3. *"Maximum" Local Tuning.* Set l_i following (3).

6

Step 2.4. *"Additive" Local Tuning.* Set l_i following (9).

Step 2.5. *"Maximum-Additive" Local Tuning.* Set l_i following (11).

Step 3. *Calculation of characteristics.* Compute for each subinterval $[x_{i-1}, x_i]$, $i = 2, \ldots, k$, its characteristic R_i by using one of the following rules.

Step 3.1. *Geometric methods.*

$$R_i = \frac{z_i + z_{i-1}}{2} - l_i \frac{x_i - x_{i-1}}{2}.$$

Step 3.2. *Information methods.*

$$R_i = 2(z_i + z_{i-1}) - l_i(x_i - x_{i-1}) - \frac{(z_i - z_{i-1})^2}{l_i(x_i - x_{i-1})}.$$

Step 4. *Subinterval selection.* Determine subinterval $[x_{t-1}, x_t]$, $t = t(k)$, for performing next trial by using one of the following rules.

Step 4.1. *Global phase.* Select the subinterval $[x_{t-1}, x_t]$ corresponding to the minimal characteristic, i. e., such that $t = \arg\min_{i=2,\ldots,k} R_i$.

Steps 4.2–4.3. *Local improvement.*

     if $flag = 1$, **then** *(perform local improvement)*

         if $z^k = z_{i_{\min}}$, **then** $t = \arg\min\{R_i : i \in \{i_{\min} + 1, i_{\min}\}\}$;

         else alternate the choice of subinterval between $[x_{i_{\min}}, x_{i_{\min}+1}]$ and $[x_{i_{\min}-1}, x_{i_{\min}}]$ starting from the right subinterval $[x_{i_{\min}}, x_{i_{\min}+1}]$.

         end if

     else $t = \arg\min_{i=2,\ldots,k} R_i$ *(do not perform local improvement at the current iteration).*

     end if

The subsequent part of this Step differs for two local improvement techniques.

Step 4.2. *Pessimistic local improvement.*

     if $flag = 1$ and

$$x_t - x_{t-1} \leq \delta, \tag{12}$$

where $\delta > 0$ is the local search accuracy,

     then $t = \arg\min_{i=2,\ldots,k} R_i$ *(local improvement is not performed since the local search accuracy has been achieved).*

     end if

Set $flag := NOT(flag)$ *(switch the local/global flag).*

Step 4.3. *Optimistic local improvement.*

Set $flag := NOT(flag)$ *(switch the local/global flag: the accuracy of local search is not separately checked in this strategy).*

Step 5. *Global stopping criterion.* **If**

$$x_t - x_{t-1} \leq \varepsilon, \tag{13}$$

where $\varepsilon > 0$ is a given accuracy of the global search, **then Stop** and take as an estimate of the global minimum f^* the value $f_k^* = \min_{i=1,\ldots,k}\{z_i\}$ obtained at a point $x_k^* = \arg\min_{i=1,\ldots,k}\{z_i\}$.

Otherwise, go to Step 6.

Step 6. *New trial.* Execute next trial at the point x^{k+1} from (10): $z^{k+1} := f(x^{k+1})$. Increase the iteration counter $k := k + 1$, and go to Step 1.

Table 1 Description of the methods considered in the paper, the signs "+" show a combination of implementations of Steps 2–4 in the GS for each method.

Method	Step2					Step3		Step4		
	2.1	2.2	2.3	2.4	2.5	3.1	3.2	4.1	4.2	4.3
Geom-AL	+					+		+		
Geom-GL		+				+		+		
Geom-LTM			+			+		+		
Geom-LTA				+		+		+		
Geom-LTMA					+	+		+		
Geom-LTIMP			+			+			+	
Geom-LTIAP				+		+			+	
Geom-LTIMAP					+	+			+	
Geom-LTIMO			+			+				+
Geom-LTIAO				+		+				+
Geom-LTIMAO					+	+				+
Inf-AL	+						+	+		
Inf-GL		+					+	+		
Inf-LTM			+				+	+		
Inf-LTA				+			+	+		
Inf-LTMA					+		+	+		
Inf-LTIMP			+				+		+	
Inf-LTIAP				+			+		+	
Inf-LTIMAP					+		+		+	
Inf-LTIMO			+				+			+
Inf-LTIAO				+			+			+
Inf-LTIMAO					+		+			+

All the Lipschitz global optimization methods considered in the paper are summarized in Table 1, from which concrete implementations of Steps 2–4 in the GS can be individuated. As shown experimentally in Section 4, the methods using an a priori given estimate of the Lipschitz constant or its global estimate lose, as a rule, in comparison with methods using local tuning techniques, in terms of the trials performed to approximate the global solutions to problems. Therefore, local improvement accelerations (Steps 4.2–4.3 of the GS) were implemented for methods using local tuning strategies only. In what follows, the methods from Table 1 are furthermore specified (for the methods known in the literature, the respective references are provided).

1. **Geom-AL**: Piyavskij's method with the a priori given Lipschitz constant (see [35, 50] and [22, 44] for generalizations and discussions): GS with Step 2.1, Step 3.1 and Step 4.1.

2. **Geom-GL**: Geometric method with the global estimate of the Lipschitz constant (see [44]): GS with Step 2.2, Step 3.1 and Step 4.1.

3. **Geom-LTM**: Geometric method with the "Maximum" Local Tuning (see [37, 44, 54]): GS with Step 2.3, Step 3.1 and Step 4.1.

4. **Geom-LTA**: Geometric method with the "Additive" Local Tuning: GS with Step 2.4, Step 3.1 and Step 4.1.

5. **Geom-LTMA**: Geometric method with the "Maximum-Additive" Local Tuning: GS with Step 2.5, Step 3.1 and Step 4.1.

6. **Geom-LTIMP**: Geometric method with the "Maximum" Local Tuning and the pessimistic strategy of the local improvement (see [26, 44]): GS with Step 2.3, Step 3.1 and Step 4.2.

7. **Geom-LTIAP**: Geometric method with the "Additive" Local Tuning and the pessimistic strategy of the local improvement: GS with Step 2.4, Step 3.1 and Step 4.2.

8. **Geom-LTIMAP**: Geometric method with the "Maximum-Additive" Local Tuning and the pessimistic strategy of the local improvement: GS with Step 2.5, Step 3.1 and Step 4.2.

9. **Geom-LTIMO**: Geometric method with the "Maximum" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.3, Step 3.1 and Step 4.3.

10. **Geom-LTIAO**: Geometric method with the "Additive" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.4, Step 3.1 and Step 4.3.

11. **Geom-LTIMAO**: Geometric method with the "Maximum-Additive" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.5, Step 3.1 and Step 4.3.

12. **Inf-AL**: Information method with the a priori given Lipschitz constant (see [44]): GS with Step 2.1, Step 3.2 and Step 4.1.

13. **Inf-GL**: Strongin's information-statistical method with the global estimate of the Lipschitz constant (see [51, 52, 54]): GS with Step 2.2, Step 3.2 and Step 4.1.

14. **Inf-LTM**: Information method with the "Maximum" Local Tuning (see [36, 49, 54]): GS with Step 2.3, Step 3.2 and Step 4.1.

15. **Inf-LTA**: Information method with the "Additive" Local Tuning (see [9, 53]): GS with Step 2.4, Step 3.2 and Step 4.1.

16. **Inf-LTMA**: Information method with the "Maximum-Additive" Local Tuning: GS with Step 2.5, Step 3.2 and Step 4.1.

17. **Inf-LTIMP**: Information method with the "Maximum" Local Tuning and the pessimistic strategy of the local improvement [25, 49]: GS with Step 2.3, Step 3.2 and Step 4.2.

18. **Inf-LTIAP**: Information method with the "Additive" Local Tuning and the pessimistic strategy of the local improvement: GS with Step 2.4, Step 3.2 and Step 4.2.

19. **Inf-LTIMAP**: Information method with the "Maximum-Additive" Local Tuning and the pessimistic strategy of the local improvement: GS with Step 2.5, Step 3.2 and Step 4.2.

20. **Inf-LTIMO**: Information method with the "Maximum" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.3, Step 3.2 and Step 4.3.

21. **Inf-LTIAO**: Information method with the "Additive" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.4, Step 3.2 and Step 4.3.

22. **Inf-LTIMAO**: Information method with the "Maximum-Additive" Local Tuning and the optimistic strategy of the local improvement: GS with Step 2.5, Step 3.2 and Step 4.3.

Let us spend a few words regarding convergence of the methods belonging to the GS. To do this, we study an infinite trial sequence $\{x^k\}$ generated by an algorithm belonging to the general scheme GS for solving the problem (1), (2) with $\delta = 0$ from (12) and $\varepsilon = 0$ from (13).

Assume that the objective function $f(x)$ satisfies the Lipschitz condition (2) with a finite constant $L > 0$, and let x' be any limit point of $\{x^k\}$ generated by an algorithm belonging to the GS, that does not use the "Additive" Local Tuning and works with one of the estimates (3), (6), (11). Then, the following assertions hold:

1. if $x' \in \,]a, b[$ then convergence to x' is bilateral, i.e., there exist two infinite subsequences of $\{x^k\}$ converging to x': one from the left, the other from the right;

2. $f(x^k) \geq f(x')$, for all trial points x^k, $k \geq 1$;

3. if there exists another limit point $x'' \neq x'$, then $f(x'') = f(x')$;

4. if the function $f(x)$ has a finite number of local minima in $[a, b]$, then the point x' is locally optimal;

5. *(Sufficient conditions for convergence to a global minimizer).* Let x^* be a global minimizer of $f(x)$. If there exists an iteration number k^* such that for all $k > k^*$, then the inequality

$$l_{j(k)} > L_{j(k)} \tag{14}$$

holds, where $L_{j(k)}$ is the Lipschitz constant for the interval $[x_{j(k)-1}, x_{j(k)}]$ containing x^*, and $l_{j(k)}$ is its estimate. Then, the set of limit points of the sequence $\{x^k\}$ coincides with the set of global minimizers of the function $f(x)$.

Proof. Since all the methods mentioned in the Theorem belong to the "Divide the Best" class of global optimization algorithms introduced in [39], the proofs of assertions 1–5 can be easily obtained as particular cases of the respective proofs in [39, 44].

Assertions *1–5* hold for methods belonging to the GS and using the "Additive" Local Tuning if the condition $l_{i(k)} > rH_{i(k)}$ is fulfilled for all i and k. *Proof.* Fulfillment of the condition $l_{i(k)} > rH_{i(k)}$ ensures that: (i) each new trial point x^{k+1} belongs to the interval (x_{t-1}, x_t) chosen for partitioning; (ii) the distances $x^{k+1} - x_{t-1}$ and $x_t - x^{k+1}$ are finite. The fulfillment of these two conditions implies that the methods belong to the class of "Divide the Best" global optimization algorithms and, therefore, proofs of assertions 1–5 can be easily obtained as particular cases of the respective proofs in [39, 44].

Note that in practice, since both ε and δ assume finite positive values, methods using the optimistic local improvement can miss the global optimum and stop in the δ-neighbourhood of a local minimizer (see Step 4 of the GS).

Next Theorem ensures existence of the values of the reliability parameter r satisfying condition (14), providing so the fact that all global minimizers of $f(x)$ will be determined by the proposed methods without using the a priori known Lipschitz constant.

For any function $f(x)$ satisfying the Lipschitz condition *(2)* with $L < \infty$ and for methods belonging to the GS and using one of the estimates *(3), (6), (9), (11)* there exists a value r^* such that, for all $r > r^*$, condition *(14)* holds. *Proof.* It follows from, $(3), (6), (9), (11)$, and the finiteness of $\xi > 0$ that approximations of the Lipschitz constant l_i in the methods belonging to the GS are always positive. Since L in (2) is finite and any positive value of the parameter r can be chosen in $(3), (6), (9), (11)$, it follows that there exists an r^* such that condition (14) will be satisfied for all global minimizers for $r > r^*$.

4   Numerical Experiments

Seven series of numerical experiments were executed on the following three sets of test functions to compare 22 global optimization methods described in Section 3:

1. the widely used set of 20 test functions from [16] reported in Appendix A;

2. 100 randomly generated Pintér's functions from [34];

3. four functions originated from practical problems (see Appendix B): first two problems are from [44, page 113] and the other two functions from [14] (see also [13]).

Geometric and information methods with and without the local improvement techniques (optimistic and pessimistic) were tested in these experimental series. In all the experiments, the accuracy of the global search was chosen as $\varepsilon = 10^{-5}(b-a)$, where $[a, b]$ is the search interval. The accuracy of the local search was set as $\delta = \varepsilon$ in the algorithms with the local improvement. Results of numerical experiments are reported

in Tables 2–9, where the number of function trials executed until the satisfaction of the stopping rule is presented for each considered method (the best results for the methods within the same class are shown in bold).

Table 2 Number of trials performed by the considered geometric methods without the local improvement on 20 tests from Appendix A.

#	Geom-AL	Geom-GL	Geom-LTM	Geom-LTA	Geom-LTMA
1	595	446	50	44	**35**
2	457	373	49	52	**39**
3	577	522	176	202	**84**
4	1177	1235	57	73	**47**
5	383	444	57	65	**43**
6	301	299	70	73	**50**
7	575	402	53	51	**41**
8	485	481	164	183	**82**
9	469	358	55	57	**41**
10	571	481	55	58	**42**
11	1099	1192	100	109	**78**
12	993	1029	93	96	**68**
13	2833	2174	93	88	**68**
14	379	303	56	60	**39**
15	2513	1651	89	118	**72**
16	2855	2442	102	120	**83**
17	2109	1437	125	171	**122**
18	849	749	55	58	**41**
19	499	377	49	47	**39**
20	1017	166	53	58	**40**
Avg	1036.80	828.05	80.05	89.15	**57.70**

The first series of numerical experiments was carried out with geometric and information algorithms without the local improvement on 20 test functions given in Appendix A. Parameters of the geometric methods Geom-AL, Geom-GL, Geom-LTM, Geom-LTA, and Geom-LTMA were chosen as follows. For the method Geom-AL, the estimates of the Lipschitz constants were computed as the maximum between the values calculated as relative differences on 10^{-7}-grid and the values given in [16]. For the methods Geom-GL, Geom-LTM, and Geom-LTMA, the reliability parameter $r = 1.1$ was used as recommended in [44]. The technical parameter $\xi = 10^{-8}$ was used for all the methods with the local tuning (Geom-LTM, Geom-LTA, and Geom-LTMA). For the method Geom-LTA, the parameter r was increased with the step equal to 0.1, starting from $r = 1.1$ until all 20 test problems were solved (i. e., for all the problems the algorithm stopped in the ε-neighbourhood of a global minimizer). This situation happened for $r = 1.8$: the corresponding results are shown in the column Geom-LTA of Table 2.

As can be seen from Table 2, the performance of the method Geom-LTMA was better with respect to the other geometric algorithms tested. The experiments also showed that the additive convolution (Geom-LTA) did not guarantee the proximity of the found solution to the global minimum with the common value $r = 1.1$. With an increased value of the reliability parameter r, the average number of trials performed by this method on 20 tests was also slightly worse than that of the method with the maximum convolution (Geom-LTM), but better than the averages of the methods using global estimates of the Lipschitz constants (Geom-AL and Geom-GL).

Results of numerical experiments with information methods without the local improvement techniques (methods Inf-AL, Inf-GL, Inf-LTM, Inf-LTA, and Inf-LTMA) on the same 20 tests from Appendix A are shown in Table 3. Parameters of the information methods were chosen as follows. The estimates of the Lipschitz constants for the method Inf-AL were the same as for the method Geom-AL. The reliability parameter $r = 2$ was used in the methods Inf-GL, Inf-LTM, and Inf-LTMA, as recommended in [44, 52, 54].

Table 3 Number of trials performed by the considered information methods without the local improvement on 20 tests from Appendix A.

#	Inf-AL	Inf-GL	Inf-LTM	Inf-LTA	Inf-LTMA
1	422	501	46	35	**32**
2	323	373	47	38	**36**
3	390	504	173	72	**56**
4	833	1076	51	56	**47**
5	269	334	59	47	**37**
6	208	239	65	46	**45**
7	403	318	49	38	**37**
8	157	477	163	113	**63**
9	329	339	54	48	**42**
10	406	435	51	42	**38**
11	773	1153	95	78	**75**
12	706	918	88	71	**64**
13	2012	1351	54	54	**51**
14	264	349	55	44	**38**
15	1778	1893	81	82	**71**
16	2023	1592	71	67	**64**
17	1489	1484	128	121	**105**
18	601	684	52	**43**	**43**
19	352	336	44	34	**33**
20	681	171	55	**39**	**39**
Avg	720.95	726.35	74.05	58.40	**50.80**

Table 4 Results of numerical experiments with the considered geometric and information methods without the local improvement on 100 Pintér's test functions from [34].

Method	Average	StDev	Method	Average	StDev
Geom-AL	1080.24	91.17	Inf-AL	750.03	66.23
Geom-GL	502.17	148.25	Inf-GL	423.19	109.26
Geom-LTM	58.96	9.92	Inf-LTM	52.13	5.61
Geom-LTA	70.48	17.15	Inf-LTA	**36.47**	6.58
Geom-LTMA	**42.34**	6.63	Inf-LTMA	38.10	5.96

For all the information methods with the local tuning techniques (Inf-LTM, Inf-LTA, and Inf-LTMA), the value $\xi = 10^{-8}$ was used. For the method Inf-LTA, the parameter r was increased (starting from $r = 2$) up to the value $r = 2.3$ when all 20 test problems were solved.

As can be seen from Table 3, the performance of the method Inf-LTMA was better (as also verified for its geometric counterpart) with respect to the other information algorithms tested. The experiments also showed that the average number of trials performed by the Inf-LTA method with $r = 2.3$ on 20 tests was better than that of the method with the maximum convolution (Inf-LTM).

The second series of experiments (see Table 4) was executed on the class of 100 Pintér's test functions from [34] with all geometric and information algorithms without the local improvement (i. e., all the methods used in the first series of experiments). Parameters of the methods Geom-AL, Geom-GL, Geom-LTM, Geom-LTMA, and Inf-AL, Inf-GL, Inf-LTM and Inf-LTMA were the same as in the first experimental series ($r = 1.1$ for all the geometric methods and $r = 2$ for the information methods). The reliability parameter for the method Geom-LTA was increased from $r = 1.1$ to $r = 1.8$ (when all 100 problems were solved). All the information methods were able to solve all 100 test problems with $r = 2$ (see Table 4). The average performance of the Geom-LTMA and the Inf-LTA methods was the best among the other considered geometric and information algorithms, respectively.

Table 5 Number of trials performed by the considered geometric and information methods without the local improvement on four test functions from Appendix B.

Method	Test problem				Average
	1	2	3	4	
Geom-AL	37	395	261	332	256.25
Geom-GL	39	388	216	307	237.50
Geom-LTM	37	54	59	232	95.50
Geom-LTA	74	58	68	204	101.00
Geom-LTMA	**33**	**39**	**48**	**137**	**64.25**
Inf-AL	**12**	278	180	187	164.25
Inf-GL	35	333	215	229	203.00
Inf-LTM	25	53	56	212	86.50
Inf-LTA	19	**35**	**40**	165	64.75
Inf-LTMA	24	**35**	**40**	**122**	**55.25**

Table 6 Number of trials performed by the considered geometric and information methods with the *optimistic* local improvement on 20 tests from Appendix A.

#	Geom LTIMO	Geom LTIAO	Geom LTIMAO	Geom LTIAMAO	Inf LTIMO	Inf LTIAO	Inf LTIMAO	Inf LTIAMAO
1	45	41	**35**	43	47	**35**	37	**35**
2	47	49	**35**	45	45	**37**	41	**37**
3	49	45	**39**	69	55	**45**	51	**45**
4	47	53	**43**	55	**49**	53	53	53
5	55	49	**47**	**47**	51	47	47	**45**
6	51	49	**45**	49	47	**43**	47	47
7	45	45	**39**	49	49	**37**	39	41
8	37	41	**35**	45	**41**	45	47	43
9	49	51	**41**	51	51	51	**40**	48
10	47	49	**41**	49	51	**43**	43	43
11	49	53	**45**	53	**55**	59	55	**51**
12	43	53	**35**	49	53	67	**45**	67
13	**51**	53	57	**51**	41	51	55	51
14	45	45	**43**	43	49	**43**	45	45
15	**45**	57	47	57	**45**	55	53	53
16	**49**	55	53	55	**47**	49	53	55
17	93	**53**	95	55	59	55	**53**	53
18	45	47	**37**	55	49	**41**	44	**41**
19	45	43	**35**	43	46	**33**	35	**33**
20	43	45	**37**	43	49	**35**	39	40
Avg	49	48.8	**44.20**	50.3	48.95	46.2	**46.10**	46.3

The third series of the experiments (see Table 5) was carried out on four applied test problems from [44] and [14] (see Appendix B). All the methods without the local improvement used in the previous two series of experiments (Geom-AL, Geom-GL, Geom-LTM, Geom-LTA, Geom-LTMA and Inf-AL, Inf-GL, Inf-LTM, Inf-LTA, Inf-LTMA) were tested and all the parameters for these methods were the same as above, except the reliability parameters of the methods Geom-LTA and Inf-LTA. Particularly, the applied problem 4 was not solved by the Geom-LTA method with $r = 1.1$. With the increased value $r = 1.8$, the obtained results (reported in Table 5) of this geometric method were worse than the results of the other geometric methods with the local tuning (Geom-LTM and Geom-LTMA). The method Inf-LTA solved all the four applied problems also with a higher value $r = 2.3$ and was outrun by the Inf-LTMA method (the latter one produced the best average result among all the competitors; see Table 5).

13

Table 7 Number of trials performed by the considered geometric and information methods with the *pessimistic* local improvement on 20 tests from Appendix A.

#	Geom LTIMP	Geom LTIAP	Geom LTIMAP	Geom LTIAMAP	Inf LTIMP	Inf LTIAP	Inf LTIMAP	Inf LTIAMAP
1	49	46	**36**	37	47	38	35	**34**
2	49	50	**38**	49	47	37	**35**	36
3	165	212	**111**	200	177	**56**	57	**56**
4	56	73	**47**	69	51	56	**46**	46
5	63	66	**48**	65	57	47	**38**	38
6	70	71	**51**	58	64	46	**45**	45
7	54	53	**41**	44	51	39	**38**	39
8	157	182	**81**	184	163	116	99	**71**
9	53	57	**43**	54	52	52	43	**42**
10	56	59	**42**	54	52	43	**39**	40
11	100	114	**77**	99	95	78	**72**	**72**
12	93	97	**69**	86	87	73	**64**	64
13	97	86	**68**	72	55	52	50	**46**
14	58	197	**43**	59	60	46	42	**37**
15	79	120	**76**	101	79	82	70	**53**
16	97	115	**81**	92	71	66	60	**57**
17	140	189	**139**	156	127	129	**100**	102
18	55	60	**42**	58	51	**42**	**42**	**42**
19	52	50	**36**	43	46	33	32	**31**
20	54	56	**40**	47	51	**37**	40	**37**
Avg	79.85	97.65	**60.45**	81.35	74.15	58.4	52.35	**49.4**

In the following several series of experiments, the local improvement techniques were compared on the same three sets of test functions. In the fourth series (see Table 6), six methods (geometric and information) with the optimistic local improvement (methods Geom-LTIMO, Geom-LTIAO, Geom-LTIMAO and Inf-LTIMO, Inf-LTIAO and Inf-LTIMAO) were compared on the class of 20 test functions from [16] (see Appendix A). The reliability parameter $r = 1.1$ was used for the methods Geom-LTIMO and Geom-LTIMAO and $r = 2$ was used for the method Inf-LTIMO. For the method Geom-LTIAO r was increased to 1.6, for the method Geom-LTIAMAO r was increased to 1.8, and for the methods Inf-LTIMAO, Inf-LTIAO, and Inf-LTIAMAO r was increased to 2.3. As can be seen from Table 6, the best average result among all the algorithms was shown by the method Geom-LTIMAO (while the Inf-LTIMAO was the best in average among the considered information methods).

In the fifth series of experiments, six methods (geometric and information) using the pessimistic local improvement were compared on the same 20 test functions. The obtained results are presented in Table 7. The usual values $r = 1.1$ and $r = 2$ were used for the geometric (Geom-LTIMP and Geom-LTIMAP) and the information (Inf-LTIMP, Inf-LTIMAP, and Inf-LTIAMAP) methods, respectively. The values of the reliability parameter ensuring the solution to all the test problems in the case of methods Geom-LTIAP, Inf-LTIAP, and Geom-LTIAMAP were set as $r = 1.8$, $r = 2.3$, and $r = 1.6$ respectively. As can be seen from Table 7, the "Maximum" and the "Maximum-Additive" local tuning techniques were more stable, and generally allowed us to find the global solution for all test problems without increasing r. Moreover, the methods Geom-LTIMAP and Inf-LTIMAP showed the best performance with respect to the other techniques in the same geometric and information classes, respectively.

In the sixth series of experiments, the local improvement techniques were compared on the class of 100 Pintér's functions. The obtained results are presented in Table 8. The values of the reliability parameter r for all the methods were increased, starting from $r = 1.1$ for the geometric methods and $r = 2$ for the information methods, until all 100 problems from the class were solved. It can be seen from Table 8, that the best average number of trials for both the optimistic and pessimistic strategies was almost the same (36.90 and 37.21 in the case of information methods and 45.76 and 48.24 in the case of geometric methods, for the

Table 8 Results of numerical experiments with the considered geometric and information methods with the local improvement techniques on 100 Pintér's test functions from [34].

Optimistic strategy				Pessimistic strategy			
Method	r	Average	StDev	Method	r	Average	StDev
Geom-LTIMO	1.3	49.52	4.28	Geom-LTIMP	1.1	66.44	21.63
Geom-LTIAO	1.9	48.32	5.02	Geom-LTIAP	1.8	93.92	197.61
Geom-LTIMAO	1.4	**45.76**	5.83	Geom-LTIMAP	1.1	**48.24**	14.12
Geom-LTIAMAO	1.8	48.26	4.94	Geom-LTIAMAP	1.8	71.33	19.42
Inf-LTIMO	2.0	48.31	4.29	Inf-LTIMP	2.0	53.06	7.54
Inf-LTIAO	2.1	**36.90**	5.91	Inf-LTIAP	2.0	**37.21**	7.25
Inf-LTIMAO	2.0	38.24	6.36	Inf-LTIMAP	2.0	39.06	6.84
Inf-LTIAMAO	2.0	36.95	7.11	Inf-LTIAMAP	2.0	37.72	7.55

Table 9 Number of trials performed by the considered geometric and information methods with the local improvement techniques on four test functions from Appendix B.

	Method	r	Test problem				Average
			1	2	3	4	
Optimistic LI	GeomLTIMO	6.5	59	55	63	79	64.00
	Geom-LTIAO	1.8	55	49	49	**41**	48.50
	GeomLTIMAO	6.9	63	59	71	75	67.00
	GeomLTIAMAO	1.3	**19**	**29**	**45**	51	**36.00**
	Inf-LTIMO	6.5	49	55	67	77	62.00
	Inf-LTIAO	9.4	47	55	71	71	61.00
	Inf-LTIMAO	8.0	55	55	71	73	63.50
	Inf-LTIAMAO	8.0	47	55	71	71	61.00
Pessimistic LI	Geom-LTIMP	1.1	39	64	71	228	100.50
	Geom-LTIAP	1.8	243	102	63	1254	415.50
	Geom-LTIMAP	1.1	31	46	51	**106**	**58.50**
	Geom-LTIAMAP	1.3	20	**29**	48	159	64.00
	Inf-LTIMP	2.0	25	52	58	185	80.00
	Inf-LTIAP	2.3	**18**	36	49	174	69.25
	Inf-LTIMAP	2.0	24	35	**43**	134	59.00
	Inf-LTIAMAP	2.4	22	39	51	147	64.75

optimistic and for the pessimistic strategies, respectively). However, the pessimistic strategy seemed to be more stable since its reliability parameter (needed to solve all the problems) generally remained smaller than that of the optimistic strategy. In average, the Geom-LTMA and the Inf-LTA methods was the best among the other considered geometric and information algorithms, respectively.

Finally, the last, seventh, series of the experiments (see Table 9) was executed on the class of four applied test problems from Appendix B, where in the third column the values of the reliability parameter used to solve all the problems are also indicated. Again, as in the previous experiments, the pessimistic local improvement strategy seemed to be more stable in the case of this test set, since the optimistic strategy required a significant increase of the parameter r to determine global minimizers of these applied problems (although the best average value obtained by the optimistic Geom-LTIAO method was smaller that that of the best pessimistic Geom-LTIMAP method, see the last column in Table 9).

5   Conclusions

Univariate derivative-free global optimization has been considered in the paper, and several numerical methods belonging to the geometric and information classes of algorithms have been proposed and analyzed. New acceleration techniques to speed up the global search have been introduced. They can be used in both the geometric and information frameworks. All of the considered methods automatically switch from the global optimization to the local one and back, avoiding so the necessity to stop the global phase manually. An original mixture of new and traditional computational steps has allowed the authors to construct 22 different global optimization algorithms having, however, a similar structure. As shown, 9 instances of this mixture can lead to known global optimization methods, and the remaining 13 methods described in the paper are new. All of them have been studied theoretically and numerically compared on 124 theoretical and applied benchmark tests. It has been shown that the introduced acceleration techniques allowed the global optimization methods to significantly speed up the search with respect to some known algorithms.

Appendix A

Table 10 Twenty test functions from [16] used in numerical experiments.

#	Function $f(x)$	$[a, b]$	L	Global Minimizers
1	$\frac{1}{6}x^6 - \frac{52}{25}x^5 + \frac{39}{80}x^4 + \frac{71}{10}x^3 - \frac{79}{20}x^2 - x + \frac{1}{10}$	$[-1.5, 11]$	13870	$x^* = 10$
2	$\sin x + \sin \frac{10x}{3}$	$[2.7, 7.5]$	4.3	$x^* = 5.146$
3	$-\sum_{k=1}^{5} k \sin[(k+1)x + k]$	$[-10, 10]$	68.5	$x_1^* = -0.491,$ $x_2^* = -6.775,$ $x_3^* = 5.792$
4	$-(16x^2 - 24x + 5)\,e^{-x}$	$[1.9, 3.9]$	3.0	$x^* = 2.868$
5	$(3x - 1.4)\sin 18x$	$[0, 1.2]$	36.0	$x^* = 0.966$
6	$-(x + \sin x)\,e^{-x^2}$	$[-10, 10]$	2.5	$x^* = 0.680$
7	$\sin x + \sin \frac{10x}{3} + \ln x - 0.84x + 3$	$[2.7, 7.5]$	6.0	$x^* = 5.200$
8	$-\sum_{k=1}^{5} k \cos[(k+1)x + k]$	$[-10, 10]$	69.5	$x_1^* = -0.800,$ $x_2^* = -7.084,$ $x_3^* = 5.483$
9	$\sin x + \sin \frac{2x}{3}$	$[3.1, 20.4]$	1.7	$x^* = 17.040$
10	$-x \sin x$	$[0, 10]$	11.0	$x^* = 7.979$
11	$2\cos x + \cos 2x$	$[-1.57, 6.28]$	3.6	$x_1^* = 4.189,$ $x_2^* = 2.094$
12	$\sin^3 x + \cos^3 x$	$[0, 6.28]$	2.2	$x_1^* = 4.712,$ $x_2^* = 3.142$
13	$-x^{2/3} + (x^2 - 1)^{1/3}$	$[0.001, 0.99]$	8.5	$x^* = 0.707$
14	$-e^{-x}\sin 2\pi x$	$[0, 4]$	6.5	$x^* = 0.225$
15	$\frac{x^2 - 5x + 6}{x^2 + 1}$	$[-5, 5]$	6.5	$x^* = 2.414$
16	$2(x - 3)^2 + e^{0.5x^2}$	$[-3, 3]$	294.1	$x^* = 1.591$
17	$x^6 - 15x^4 + 27x^2 + 250$	$[-4, 4]$	2520.0	$x_1^* = -3,$ $x_2^* = 3$
18	$\begin{cases} (x-2)^2, & x \le 3 \\ 2\ln(x-2) + 1, & x > 3 \end{cases}$	$[0, 6]$	4.0	$x^* = 2$
19	$-x + \sin 3x - 1$	$[0, 6.5]$	4.1	$x^* = 5.873$
20	$(\sin x - x)\,e^{-x^2}$	$[-10, 10]$	1.3	$x^* = 1.195$

Appendix B

Table 11 Four one-dimensional test functions taken from practical applications.

#	Function $f(x)$	$[a,b]$	L	Global Minimizers	Ref.
1	$\|\sin^3(x) + \cos^3(x)\| + 0.1$	$[0, 2\pi]$	2.2	$x_1^* = 2.356,$ $x_2^* = 5.498$	[44]
2	$x\|\sin(x)\| + 6$	$[-10, 10]$	9.7	$x^* = -7.979$	[44]
3	$\sum_{i=1}^{10}(y_i^{(1)} - \sin(2\pi xi))^2$	$[0, 1]$	432.0	$x^* = 0.4$	[14]
4	$\sum_{i=1}^{100}(y_i^{(2)} - \sin(2\pi xi))^2$	$[0, 1]$	28690.8	$x^* = 0.4$	[14]

References

[1] J. M. Calvin, Y. Chen, and A. Žilinskas. An adaptive univariate global optimization algorithm and its convergence rate for twice continuously differentiable functions. *J. Optimization Theory and Applications*, 155(2):628–636, 2012.

[2] J. M. Calvin and A. Žilinskas. One-dimensional global optimization for observations with noise. *Computers and Mathematics with Applications*, 50(1–2):157–169, 2005.

[3] L. G. Casado, I. García, and Y. D. Sergeyev. Interval algorithms for finding the minimal root in a set of multiextremal non-differentiable one-dimensional functions. *SIAM J. Scientific Computing*, 24(2):359–376, 2002.

[4] T. Csendes, editor. *Developments in Reliable Computing*. Kluwer Academic Publishers, Dordrecht, 1999.

[5] P. Daponte, D. Grimaldi, A. Molinaro, and Y. D. Sergeyev. An algorithm for finding the zero-crossing of time signals with Lipschitzean derivatives. *Measurement*, 16(1):37–49, 1995.

[6] P. Daponte, D. Grimaldi, A. Molinaro, and Y. D. Sergeyev. Fast detection of the first zero-crossing in a measurement signal set. *Measurement*, 19(1):29–39, 1996.

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

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

[9] A. V. Gergel, V. A. Grishagin, and R. G. Strongin. Development of the parallel adaptive multistep reduction method. *Vestnik of Lobachevsky State University of Nizhni Novgorod*, 6(1):216–222, 2013. (In Russian).

[10] V. P. Gergel. A global search algorithm using derivatives. In Yu. I. Neymark, editor, *Systems Dynamics and Optimization*, pages 161–178. N. Novgorod University Press, 1992. (In Russian).

[11] V. P. Gergel, V. A. Grishagin, and R. A. Israfilov. Local tuning in nested scheme of global optimization. *Procedia Computer Science*, 51:865–874, 2015. (International Conference on Computational Science ICCS 2015 – Computational Science at the Gates of Nature).

[12] V. P. Gergel and Ya. D. Sergeyev. Sequential and parallel algorithms for global minimizing functions with Lipschitzian derivatives. *Comput. Math. Appl.*, 37(4–5):163–179, 1999.

[13] J. W. Gillard and D. E. Kvasov. Lipschitz optimization methods for fitting a sum of damped sinusoids to a series of observations. *Stat. Interface*, 2015. Submitted.

[14] J. W. Gillard and A. A. Zhigljavsky. Optimization challenges in the structured low rank approximation problem. *J. Global Optim.*, 57(3):733–751, 2013.

[15] K. Hamacher. On stochastic global optimization of one-dimensional functions. *Physica A: Statistical Mechanics and its Applications*, 354:547–557, 2005.

[16] P. Hansen and B. Jaumard. Lipschitz optimization. In R. Horst and P. M. Pardalos, editors, *Handbook of Global Optimization*, volume 1, pages 407–493. Kluwer Academic Publishers, Dordrecht, 1995.

[17] R. Horst and P. M. Pardalos, editors. *Handbook of Global Optimization*, volume 1. Kluwer Academic Publishers, Dordrecht, 1995.

[18] D. E. Johnson. *Introduction to Filter Theory*. Prentice Hall Inc., New Jersey, 1976.

[19] D. E. Kvasov, D. Menniti, A. Pinnarelli, Y. D. Sergeyev, and 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.

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

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

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

[23] D. E. Kvasov and Y. D. Sergeyev. Lipschitz global optimization methods in control problems. *Automation and Remote Control*, 74(9):1435–1448, 2013.

[24] D. E. Kvasov and Y. D. Sergeyev. Deterministic approaches for solving practical black-box global optimization problems. *Advances in Engineering Software*, 80:58–66, 2015.

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

[26] D. Lera and Y. D. Sergeyev. Acceleration of univariate global optimization algorithms working with Lipschitz functions and Lipschitz first derivatives. *SIAM J. Optimization*, 23(1):508–529, 2013.

[27] D. Lera and Y. D. Sergeyev. Deterministic global optimization using space-filling curves and multiple estimates of Lipschitz and Hölder constants. *Communications in Nonlinear Science and Numerical Simulation*, 23:328–342, 2015.

[28] G. Liuzzi, S. Lucidi, V. Piccialli, and A. Sotgiu. A magnetic resonance device designed via global optimization techniques. *Math. Program.*, 101(2):339–364, 2004.

[29] A. Molinaro and Y. D. Sergeyev. Finding the minimal root of an equation with the multiextremal and nondifferentiable left-hand part. *Numerical Algorithms*, 28(1–4):255–272, 2001.

[30] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, and J. Žilinskas. Globally-biased DISIMPL algorithm for expensive global optimization. *J. Global Optim.*, 59(2-3):545–567, 2014.

[31] R. Paulavičius, J. Žilinskas, and A. Grothey. Investigation of selection strategies in branch and bound algorithm with simplicial partitions and combination of Lipschitz bounds. *Optimization Letters*, 4(2):173–183, 2010.

[32] R. Paulavičius and J. Žilinskas. *Simplicial Global Optimization*. SpringerBriefs in Optimization. Springer, New York, 2014.

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

[34] J. D. Pintér. Global optimization: software, test problems, and applications. In P. M. Pardalos and H. E. Romeijn, editors, *Handbook of Global Optimization*, volume 2, pages 515–569. Kluwer Academic Publishers, Dordrecht, 2002.

[35] S. A. Piyavskij. An algorithm for finding the absolute extremum of a function. *USSR Comput. Math. Math. Phys.*, 12(4):57–67, 1972. (In Russian: Zh. Vychisl. Mat. Mat. Fiz., 12(4) (1972), pp. 888–896).

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

[37] Y. D. Sergeyev. A one-dimensional deterministic global minimization algorithm. *Comput. Math. Math. Phys.*, 35(5):705–717, 1995.

[38] Y. D. Sergeyev. Global one-dimensional optimization using smooth auxiliary functions. *Mathematical Programming*, 81(1):127–146, 1998.

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

[40] Y. D. Sergeyev. Univariate global optimization with multiextremal non-differentiable constraints without penalty functions. *Computational Optimization and Applications*, 34(2):229–248, 2006.

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

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

[43] Y. D. Sergeyev and V. A. Grishagin. A parallel algorithm for finding the global minimum of univariate functions. *J. Optimization Theory and Applications*, 80(3):513–536, 1994.

[44] Y. D. Sergeyev and D. E. Kvasov. *Diagonal Global Optimization Methods*. Fizmatlit, Moscow, 2008. (In Russian).

[45] Y. D. Sergeyev and D. E. Kvasov. A deterministic global optimization using smooth diagonal auxiliary functions. *Communications in Nonlinear Science and Numerical Simulation*, 21:99–111, 2015.

[46] Y. D. Sergeyev and D. E. Kvasov. On deterministic diagonal methods for solving global optimization problems with Lipschitz gradients. In A. Migdalas and A. Karakitsiou, editors, *Optimization, Control, and Applications in the Information Age*, volume 130 of *Springer Proceedings in Matematics and Statistics*, pages 319–337. Springer, Switzerland, 2015.

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

[48] Y. D. Sergeyev and D. L. Markin. An algorithm for solving global optimization problems with nonlinear constraints. *J. Global Optim.*, 7(4):407–419, 1995.

[49] Y. D. Sergeyev, R. G. Strongin, and D. Lera. *Introduction to Global Optimization Exploiting Space-Filling Curves*. SpringerBriefs in Optimization. Springer, New York, 2013.

[50] B. O. Shubert. A sequential method seeking the global maximum of a function. *SIAM J. Numer. Anal.*, 9(3):379–388, 1972.

[51] R. G. Strongin. On the convergence of an algorithm for finding a global extremum. *Engineering Cybernetics*, 11:549–555, 1973.

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

[53] R. G. Strongin, V. P. Gergel, V. A. Grishagin, and K. A. Barkalov. *Parallel Computing for Global Optimization Problems*. Moscow University Press, Moscow, 2013. (In Russian).

[54] R. G. Strongin and Y. D. Sergeyev. *Global Optimization with Non-Convex Constraints: Sequential and Parallel Algorithms*. Kluwer Academic Publishers, Dordrecht, 2000. (2nd ed., 2012; 3rd ed., 2014, Springer, New York).

[55] A. Žilinskas. Optimization of one-dimensional multimodal functions: Algorithm AS 133. *Applied Statistics*, 23:367–375, 1978.

[56] A. Žilinskas. On similarities between two models of global optimization: Statistical models and radial basis functions. *J. Global Optim.*, 48(1):173–182, 2010.

[57] A. Žilinskas. A one-step worst-case optimal algorithm for bi-objective univariate optimization. *Optimization Letters*, 8(7):1945–1960, 2014.

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