Chapter 2
One-Dimensional Algorithms and Their Acceleration

It is common sense to take a method and try it. If it fails, admit it frankly and try another. But above all, try something.

Franklin D. Roosevelt

2.1 One-Dimensional Lipschitz Global Optimization

In order to give an insight to the class of geometric Lipschitz global optimization (LGO) methods, one-dimensional problems will be considered in this chapter. In global optimization, these problems play a very important role both in the theory and practice and, therefore, they were intensively studied in the last decades (see, e.g., [51, 82, 102, 148, 242, 290, 315, 323, 326, 348]). In fact, on the one hand, theoretical analysis of one-dimensional problems is quite useful since mathematical approaches developed to solve them very often can be generalized to the multidimensional case by numerous schemes (see, e.g., [58, 148, 150, 154, 176, 187, 207, 227, 234, 242, 278, 290, 323, 348]). On the other hand, there exist a large number of real-life applications where it is necessary to solve these problems (see, e.g., [227, 242, 244, 257, 290, 323, 348]).

Let us consider, for example, the following common problem in electronic measurements and electrical engineering. There exists a device whose behavior depends on a characteristic $f(x)$, $x \in [a, b]$, where the function $f(x)$ may be, for instance, an electrical signal obtained by a complex computer aided simulation over a time interval $[a, b]$ (see the function graph shown by thick line in Fig. 2.1). The function $f(x)$ is often multiextremal and Lipschitzian (it can be also differentiable with the Lipschitz first derivative). The device works correctly while $f(x) > 0$. Of course, at the initial moment $x = a$ we have $f(a) > 0$. It is necessary to describe the performance of the device over the time interval $[a, b]$ either by determining the point x^* such that

Y.D. Sergeyev and D.E. Kvasov, *Deterministic Global Optimization*,
SpringerBriefs in Optimization, DOI 10.1007/978-1-4939-7199-2_2

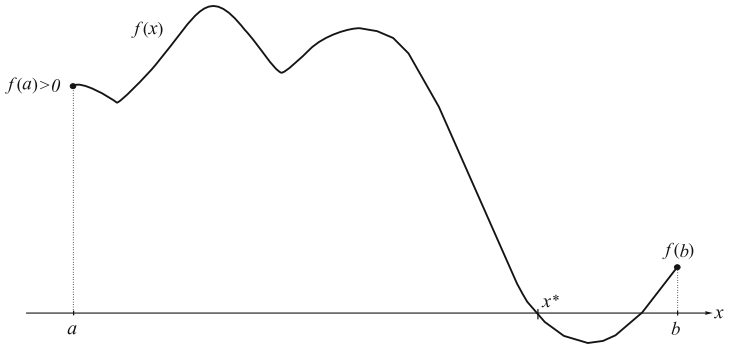

Fig. 2.1 The problem arising in electrical engineering of finding the minimal root of an equation $f(x) = 0$ with multiextremal non-differentiable *left* part

$$f(x^*) = 0, \quad f(x) > 0, \quad x \in [a, x^*), \quad x^* \in (a, b], \tag{2.1}$$

or by demonstrating that x^* satisfying (2.1) does not exist in $[a, b]$. In the latter case the device works correctly for the whole time period; thus, any information about the global minimum of $f(x)$ could be useful in practice to measure the device reliability.

This problem is equivalent to the problem of finding the minimal root (the first root from the left) of the equation $f(x) = 0$, $x \in [a, b]$, in the presence of certain initial conditions and can be reformulated as a global optimization problem. There is a simple approach to solve this problem based on a grid technique. It produces a dense mesh starting from the left margin of the interval and going on by a small step till the signal becomes less than zero. For an acquired signal, the determination of the first zero crossing point by this technique is rather slow especially if the search accuracy is high. Since the objective function $f(x)$ is multiextremal (see Fig. 2.1) the problem is even more difficult because many roots can exist in $[a, b]$ and, therefore, classical local root finding techniques can be inappropriate and application of global methods become desirable.

A box-constrained Lipschitz global optimization problem (1.1), (1.3), (1.4) can be stated in its one-dimensional version as follows. Given a small positive constant ε, it is required to find an ε-approximation of the global minimum point (*global minimizer*) x^* of a multiextremal, black-box (and, often, hard to evaluate) objective function $f(x)$ over a closed interval $[a, b]$:

$$f^* = f(x^*) = \min f(x), \quad x \in [a, b]. \tag{2.2}$$

It can be supposed either that the objective function $f(x)$ is not necessarily differentiable and satisfies the Lipschitz condition with an (unknown) Lipschitz constant L, $0 < L < \infty$,

$$|f(x') - f(x'')| \leq L|x' - x''|, \quad x', x'' \in [a, b], \tag{2.3}$$

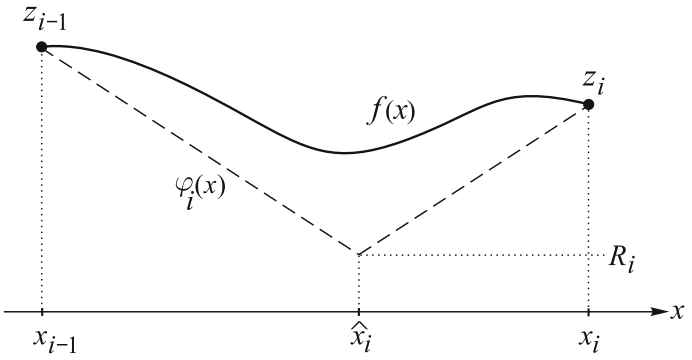

Fig. 2.2 Geometric interpretation of the Lipschitz condition

or that $f(x)$ is differentiable with the first derivative $f'(x)$ being itself multiextremal black-box Lipschitz function with an (unknown) Lipschitz constant $K, 0 < K < \infty$,

$$|f'(x') - f'(x'')| \leq K|x' - x''|, \quad x', x'' \in [a, b]. \tag{2.4}$$

As it was mentioned in Sect. 1.1, both the problems (2.2),(2.3) and (2.2),(2.4) are very frequent in practice (see, e.g., [10, 51, 102, 103, 137, 148, 227, 229, 242, 244, 255, 323, 348]).

The subsequent sections of this chapter will be dedicated to the geometric methods based on the geometric interpretation of the Lipschitz condition (2.3) or (2.4) for solving both the problems. Let us consider first the problem (2.2)–(2.3) and suppose that the objective function $f(x)$ from (2.2)–(2.3) has been evaluated at two points x_{i-1} and x_i of the search interval $[a, b]$ with the corresponding function values $z_{i-1} = f(x_{i-1})$ and $z_i = f(x_i)$ (see the function graph shown by thick line in Fig. 2.2). Then, the following inequality is satisfied over $[x_{i-1}, x_i]$:

$$f(x) \geq \phi_i(x),$$

where $\phi_i(x)$ is a piecewise linear function called *minorant* or *lower bounding function* (its graph is drawn by dashed line in Fig. 2.2),

$$\phi_i(x) = \max\{z_{i-1} - L(x - x_{i-1}), \ z_i + L(x - x_i)\}, \quad x \in [x_{i-1}, x_i]. \tag{2.5}$$

The minimal value of $\phi_i(x)$ over $[x_{i-1}, x_i]$ is therefore a lower bound of $f(x)$ over this interval. It is called *characteristic* (see [133, 135, 323]) of the interval $[x_{i-1}, x_i]$ and is calculated as follows:

$$R_i = R_{[x_{i-1}, x_i]} = \phi_i(\hat{x}_i) = \frac{z_{i-1} + z_i}{2} - L\frac{x_i - x_{i-1}}{2}, \tag{2.6}$$

where the point $\hat{x}_i$ (see Fig. 2.2) is

Table 2.1 Some geometric LGO methods characterized by different ways of estimating the Lipschitz constant

Lipschitz constant estimate	Objective function		
	Non-differentiable	Differentiable	
		Non-smooth minorants	Smooth minorants
A priori	[81, 140, 246, 306]	[13, 34, 116]	[200, 271, 276]
Multiple	[107, 141, 154, 289]	[177]	—
Global	[150, 242, 315, 323]	[116, 270, 276]	[271, 276, 323]
Local	[272, 273, 290, 323]	[270, 275, 323]	[271, 276, 290]

$$\hat{x}_i = \frac{x_{i-1} + x_i}{2} - \frac{z_i - z_{i-1}}{2L}. \tag{2.7}$$

In the next two sections, several geometric methods for solving both the problems (2.2), (2.3) and (2.2), (2.4) will be considered by classifying them on the way of obtaining the Lipschitz information (a priori given, multiple, global, and local estimates of the Lipschitz constant). References to some algorithms from each group are reported in Table 2.1 where in the case of the differentiable objective function (2.2), (2.4) the geometric methods are further differentiated by the type of the used minorant function that can be either non-smooth or smooth. Notice that methods working with smooth minorants and using multiple estimates of the Lipschitz constants for derivatives have not been proposed yet. The choice of the algorithms is explained mainly by the following two aspects. First, they are sufficiently representative for demonstrating various approaches used in the literature to estimate the Lipschitz constant. Second, they manifest a good performance on various sets of test and practical functions from the literature, and, therefore, are often chosen as worthy candidates for multidimensional extensions.

2.2 Geometric LGO Methods for Non-differentiable Functions

One of the first and well studied methods for solving the one-dimensional LGO problem (2.2), (2.3) is the already mentioned Piyavskij–Shubert method (see [245, 246, 306]), discussed in a vast literature (see, e.g., the survey [140] and references given therein). Various modifications of this algorithm have been also proposed (see, e.g., [22, 82, 140, 150, 266, 304, 328, 333]). We start this section from a detailed description of this algorithm due to its methodological importance for the further analysis of geometric LGO approaches.

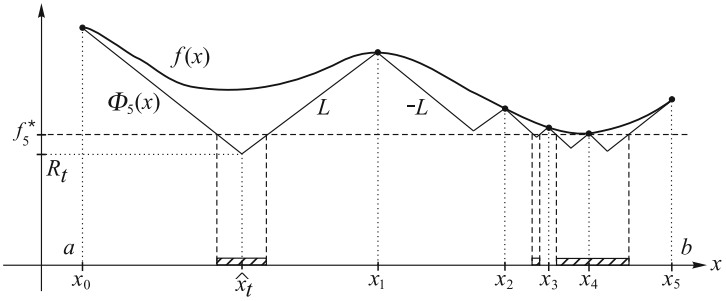

Fig. 2.3 The Lipschitzian objective function $f(x)$ (with the Lipschitz constant L) and its minorant $\Phi_5(x)$; sub-intervals which can contain the global minimizer of $f(x)$ during the work of the Piyavskij–Shubert algorithm are drawn by hatching

The Piyavskij–Shubert method is a sequential algorithm that uses a priori given information about the Lipschitz constant (namely, a given value of the Lipschitz constant or its overestimate) to adaptively construct minorant auxiliary functions for the objective function $f(x)$ as mentioned in the introductory Chap. 1. At the first iteration $k = 1$ of the algorithm, initial trials are performed at several points $x_0, x_1, \ldots, x_{n(1)}$ (usually, the left and right margins of the search interval $[a, b]$ are chosen for this scope with $n(1) = 1$; hereinafter, only the variant $n(1) = 1$ will be considered). At each successive iteration $k > 1$, trial points x_i (in the order of growth of their coordinates) and the corresponding function values $z_i = f(x_i)$, $0 \le i \le k$, are taken and the current lower bounding function $\Phi_k(x)$ is constructed as the union of minorants $\phi_i(x)$ from (2.5) over all sub-intervals $[x_{i-1}, x_i]$, $1 \le i \le k$. Then, a sub-interval with the minimal characteristic R_t is determined (see Fig. 2.3) taking into account (2.6). Finally, a new trial point $\hat{x}_t$ is calculated on this sub-interval by formula (2.7). In such a way, a new information about the objective function is acquired by evaluating $f(\hat{x}_t)$, the piecewise linear minorant function is updated, whereas the value

$$x_k^* = \arg\min_{0 \le i \le k} z_i$$

is chosen as a new approximation of the global minimizer from (2.2), with

$$f_k^* = \min_{0 \le i \le k} z_i$$

being the current approximation of the global minimum f^*. The trial points are reordered and this iterative process is repeated until some stopping criterion is verified (e.g., until the sub-interval with a new trial point becomes sufficiently small).

In Fig. 2.3, an example of the lower bounding function $\Phi_k(x)$ constructed after six function trials (i.e., after $k = 5$ iterations) is shown by continuous thin line. The black dots on the objective function graph (thick line) indicate function values at the ordered trial points x_0, x_1, …, x_5. The next, i.e., the seventh, trial will be performed at the

point $\hat{x}_t = \hat{x}_1$ (see (2.7)). After executing this new trial, the lower bounding function will be reconstructed (precisely, $\Phi_6(x)$ will differ from $\Phi_5(x)$ over the sub-interval $[x_0, x_1]$ in Fig. 2.3), thus making better the current approximation of the problem solution. Note that some sub-intervals (namely, those for which the characteristic value is greater than the estimate f_k^* of the global minimum value; see, for instance, the sub-interval $[x_1, x_2]$ in Fig. 2.3) can be eliminated in order to reduce the region where x^* can be located. More precisely, the global minimizer x^* can be found only within the set $X^*(k)$ defined as

$$X^*(k) = \{x \in [a, b] : \Phi_k(x) \leq f_k^*\}.$$

In Fig. 2.3, the sub-intervals forming the set $X^*(k)$ are indicated by hatching drawn over the axis x.

As observed, e.g., in [63, 152, 326], this algorithm is *optimal in one step*, i.e., the choice of the current evaluation point ensures the maximal improvement of the lower bound of the global minimum value of $f(x)$ with respect to a number of reasonable criteria (see, e.g., [150, 323, 326, 348] for the related discussions on optimality principles in Lipschitz global optimization). As many other geometric LGO methods, the Piyavskij–Shubert algorithm can be viewed as a branch-and-bound algorithm (see, e.g., [145, 150]) or, more generally, as a *divide-the-best algorithm* studied in [277]. Within this theoretical framework, the convergence of the sequence of trial points generated by the method to the global minimizers only (*global convergence*) can be easily established (see, e.g., [246, 277, 290]).

It should be noticed that branch-and-bound ideas were used also in the powerful method of non-uniform coverings proposed in [81, 82] (see also [87]) for functions with a priori given Lipschitz constants (see, e.g., [86, 88, 249] for modern high-performance realizations of this approach).

An interesting variation of the Piyavskij–Shubert algorithm has been proposed in [154] where the DIRECT method has been introduced. This algorithm iteratively selects several sub-intervals of the search region $[a, b]$ for partitioning and subdivides each of them into thirds with subsequent evaluations of the objective function $f(x)$ at the central points of new sub-intervals. The selection procedure is based on estimates of the lower bounds of $f(x)$ over sub-intervals by using a set of possible values for the Lipschitz constant, going from zero to infinity. In terms of the geometric approach, it is possible to say that all admissible minorant functions (in this case, they are piecewise linear discontinuous functions) are examined during the current iteration of the algorithm without constructing a specific one.

In Fig. 2.4, an example of a partition of $[a, b]$ into 5 sub-intervals performed by the DIRECT method is represented. The objective function $f(x)$ has been evaluated at the central points c_1, c_2, c_3, c_4, and c_5 of the corresponding sub-intervals (trial points are indicated by black dots in Fig. 2.4). Notice that at the first iteration the interval $[a, b]$ has been partitioned into three equal sub-intervals and the first three trials have been executed at the points c_3, c_1, and c_5. Then, the central sub-interval among the three ones has been divided into thirds and two new trials have been performed at the central points c_2 and c_4 of the left and the right of these thirds, respectively.

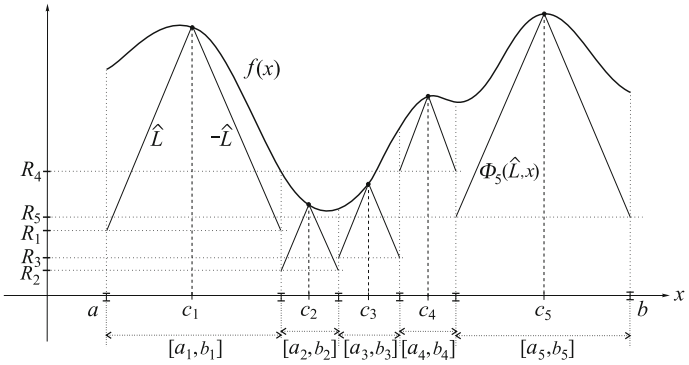

Fig. 2.4 Lower bounds R_i of $f(x)$ over sub-intervals $[a_i, b_i]$ corresponding to a particular estimate $\hat{L}$ of the Lipschitz constant

Given an overestimate $\hat{L}$ of the Lipschitz constant from (2.3), the objective function is bounded from below over $[a, b]$ by a piecewise linear discontinuous minorant function $\Phi_5(\hat{L}, x)$ (see Fig. 2.4; on the vertical axis the sub-intervals characteristics R_i, $1 \leq i \leq 5$, are indicated).

For the known overestimate $\hat{L}$ of the Lipschitz constant, new trials should be executed (similarly to the Piyavskij–Shubert algorithm) within the sub-interval $[a_2, b_2]$ having the minimal characteristic R_2 in order to obtain an improvement of the current estimate of the global minimum. But since in practical applications the exact Lipschitz constant L (or its overestimate) is often unknown, it is not possible to indicate with certainty that exactly this sub-interval is the most promising and it should be subdivided. In fact, the sub-interval $[a_2, b_2]$ in Fig. 2.4 has the smallest lower bound of $f(x)$ with respect to the estimate $\hat{L}$ of the Lipschitz constant. However, if a higher estimate $\tilde{L} \gg \hat{L}$ is taken, the slope of the lines in Fig. 2.4 increases and the sub-interval $[a_1, b_1]$ becomes preferable to all others since the lower bound of $f(x)$ over this sub-interval becomes the smallest one with respect to this new estimate $\tilde{L}$.

Jones et al. have proposed in [154] to use various estimates of the Lipschitz constant from zero to infinity at each iteration of their DIRECT algorithm. This corresponds to examination of all possible slopes $\hat{L}$ when auxiliary functions are considered (which in this case are not always minorants for $f(x)$) and lower bounds are calculated. Such a consideration leads to the basic idea of the DIRECT: to select for partitioning and sampling (i.e., performing the function trials) the so-called *potentially optimal sub-intervals*, i.e., sub-intervals over which $f(x)$ could have the best improvement with respect to a particular estimate of the Lipschitz constant. This is done by representing each sub-interval $[a_i, b_i]$ of the current partition of $[a, b]$ as a dot in a two-dimensional diagram with horizontal coordinate $c_i = (b_i - a_i)/2$ and vertical coordinate $f(c_i)$. It is then possible to prove that dots representing some potentially optimal sub-intervals are located on the lower right convex hull of all the dots corresponding to the sub-intervals (see [154]).

Thus, the DIRECT method is essentially the Piyavskij–Shubert algorithm modified to use a center-sampling strategy and to subdivide all potentially optimal subintervals related to different estimates of the Lipschitz constant L. Since during the search it uses a set of possible estimates of L and does not use its single overestimate, only the so-called *everywhere dense convergence* (i.e., convergence of the sequence of trial points to any point of the search interval) can be established for this method. It should be also noticed that it is difficult to apply for the DIRECT some meaningful stopping criterion, such as, for example, stopping on achieving a desired accuracy in solution. Nevertheless, due to its relative simplicity and a satisfactory performance on several test functions and applied problems, the DIRECT has been widely adopted in practical applications (see, e.g., [21, 30, 44, 57, 104, 141, 215, 338]) and has attracted the attention of many researchers (for its theoretical and experimental analysis and several modifications see, e.g., [53, 57, 93, 107, 128, 141, 154, 179, 189, 194, 197, 232, 289]).

As already observed, an assumption that the objective function satisfies the Lipschitz condition raises a question of estimating the corresponding Lipschitz constant. Strongin's algorithm (see, e.g., [315, 323]) answers this question by an adaptive estimation of the Lipschitz constant during the global search. It has a good convergence rate as compared to a number of other global optimization methods using only values of the objective function during the search (see, e.g., [132, 315, 323]). Therefore, this method has been often chosen as a good candidate for multidimensional extensions (see, e.g., [102, 176, 186, 216, 242, 286, 290, 323]).

Formally, this algorithm belongs to the class of the so-called *information-statistical algorithms* (or, just, *information algorithms*). The information approach originated in the works [220, 313] (see also [315, 323]) and, together with the Piyavskij–Shubert algorithm and non-uniform covering methods, it has consolidated foundations of the Lipschitz global optimization. The main idea of this approach is to apply the theory of random functions to building a mathematical representation of an available (certain or uncertain) a priori information on the objective function behavior. A systematic approach to the description of some uncertain information on this behavior is to accept that the unknown black-box function to be minimized is a sample of some known random function. Generally, to provide an efficient analytical technique for deriving some estimates of the global optimum with a finite number of trials, i.e., for obtaining by Bayesian reasoning some conditional (with respect to the trials performed) estimations, the random function should have some special structure. It is then possible to deduce the decision rules for performing new trials as some optimal decision functions (see, e.g., [29, 138, 155, 210, 212, 315, 323, 329, 348–350, 356, 357]).

In the Strongin algorithm, a global estimate H^k of the Lipschitz constant L is adaptively calculated by using the obtained results $z_i = f(x_i)$ of the function trials at the ordered trial points x_i, $0 \le i \le k$,

$$H^k = H(k) = r \cdot \max_{1 \le i \le k} \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}}, \tag{2.8}$$

where $r > 1$ is the algorithm parameter. It is possible to prove that for each fixed Lipschitzian function $f(x)$ there exists a value r^* such that taking $r \geq r^*$ guarantees convergence of the sequence of trial points generated by the method to the global minimizers of $f(x)$ only (see [315, 323]).

The Strongin algorithm can be considered in the framework of divide-the-best algorithms (see [277]). As demonstrated, e.g., in [216, 323], there is a firm relation between the information and geometric approaches. In fact, the characteristics R_i of the Strongin information algorithm associated with each sub-interval $[x_{i-1}, x_i]$, $1 \leq i \leq k$, of the search interval $[a, b]$ (see Fig. 2.2) can be rewritten (see [216, 323]) in a form similar to that of the Piyavskij–Shubert algorithm. This interesting fact allows one to interpret the Strongin method as a geometric algorithm adaptively constructing auxiliary piecewise linear functions during its work.

The usage of the global information only about behavior of the objective function, like the just mentioned global estimates of the Lipschitz constant, can lead to a slow convergence of algorithms to global minimizers. One of the traditional ways of overcoming this difficulty (see, e.g., [11, 148, 150, 194] and references given therein) recommends stopping the global procedure and switching to a local minimization method in order to improve the current solution and to accelerate the search during its final phase. Applying this technique can result in some problems related to the combination of global and local phases. The main problem is determining the moment to stop the global procedure and to start the local one. A premature arrest can provoke the loss of the global solution whereas a late one can slow down the search.

For example, it is well known that the DIRECT method balances global and local information during its work. However, the local phase is too pronounced in this balancing. The DIRECT executes too many function trials in attraction regions of local optima and, therefore, manifests a slow convergence to the global minimizers when the objective function has many local minimizers. In [289], a new geometric algorithm inspired by the DIRECT ideas has been proposed to solve difficult multiextremal LGO problems. To accomplish this task, a two-phase approach consisting of explicitly defined global and local phases has been incorporated in this method providing so a faster convergence to the global minimizers. Another type of local improvement strategy has been introduced in [186, 187, 301, 303]. This technique 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.

In [272, 273] (see also [290, 301, 323]), another fruitful approach (the so-called *local tuning approach*) which allows global optimization algorithms to tune their behavior to the shape of the objective function at different sub-intervals has been proposed. The main idea behind this approach lies in the adaptive balancing of local and global information obtained during the search for every sub-interval $[x_{i-1}, x_i]$, $1 \leq i \leq k$, formed by trial points x_i. When a sub-interval $[x_{i-1}, x_i]$ is narrow, the local information obtained within the near vicinity of the trial points x_{i-1} and x_i has the major influence on the method behavior. In this case, the results of trials executed at points lying far from the interval $[x_{i-1}, x_i]$ are less significant for the method. In

contrast, working with a wide sub-interval the method takes into consideration the global search information obtained from the whole search interval.

Both the comparison and the balancing of global and local data are effected by estimating local Lipschitz constants l_i for each sub-interval $[x_{i-1}, x_i], 1 < i \leq k$, as follows:

$$l_i = r \cdot \max\{\lambda_i, \gamma_i, \xi\}, \tag{2.9}$$

where

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

$$H_i = \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}}, \quad i = 2, \ldots, k, \tag{2.10}$$

$$H^k = \max\{H_i : i = 2, \ldots, k\}. \tag{2.11}$$

Here, $z_i = f(x_i)$, $i = 1, \ldots, k$, i.e., values of the objective function calculated at the previous iterations at the trial points x_i, $i = 1, \ldots, k$, (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:

$$\gamma_i = H^k \frac{(x_i - x_{i-1})}{X^{max}}, \tag{2.12}$$

with H^k from (2.11) and

$$X^{\max} = \max\{x_i - x_{i-1} : \ i = 2, \ldots, k\}. \tag{2.13}$$

The value H^k is an estimate of the global Lipschitz constant L over the interval $[a, b]$. The estimate l_i of the local Lipschitz constant L_i over an interval $[x_{i-1}, x_i]$ contains the following two fundamental parts: λ_i, which accounts for local properties, and γ_i, which accounts for global ones. If the interval $[x_{i-1}, x_i]$ is large, the global part increases, because in this case local information may not be reliable. In the opposite case, for small intervals $[x_{i-1}, x_i]$, the global part decreases, because local information is of the major importance and the global one loses its influence. Thus, at every sub-interval a balancing of local and global information is performed automatically.

Similarly to the Piyavskij–Shubert method, this algorithm constructs auxiliary functions approximating the objective function. Although these functions are not always minorants for $f(x)$ over the whole search interval, they are iteratively improved during the search in order to obtain appropriate bounds for the global minimum from (2.2). In Fig. 2.5, an example of the auxiliary function $\hat{\Phi}_k(x)$ for a

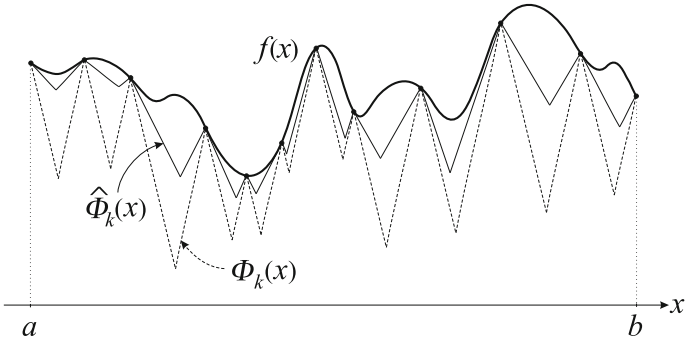

Fig. 2.5 An auxiliary function $\hat{\Phi}_k(x)$ (*solid thin line*) and a lower bounding function $\Phi_k(x)$ (*dashed line*) for a Lipschitz function $f(x)$ over $[a, b]$, constructed by using local Lipschitz estimates and by using the global Lipschitz constant, respectively

Lipschitz function $f(x)$ over $[a, b]$ constructed by using estimates of local Lipschitz constants over sub-intervals of $[a, b]$ is shown by a solid thin line; a lower bounding function $\Phi_k(x)$ for $f(x)$ over $[a, b]$ constructed by using an overestimate of the global Lipschitz constant is represented by a dashed line. It can be seen that $\hat{\Phi}_k(x)$ estimates the behavior of $f(x)$ over $[a, b]$ more accurately than $\Phi_k(x)$, especially over sub-intervals where the corresponding local Lipschitz constants are significantly smaller than the global one.

The local tuning approach enjoys the following properties (see [272, 273, 323]):

(1) the problem of determining the moment when to stop the global procedure does not arise because the local information is taken into consideration throughout the whole duration of the global search;
(2) the local information is taken into account not only in the neighborhood of a global minimizer but also over the whole search interval, thus allowing an additional acceleration of the global search;
(3) in order to guarantee convergence to the global minimizer x^* from (2.2) it is not necessary to know the exact Lipschitz constant over the whole search interval; on the contrary, only an overestimate of a local Lipschitz constant at a neighborhood of x^* is needed;
(4) geometric local tuning algorithms can be successfully parallelized (see, e.g., [284–286, 323]) and easily extended to the multidimensional case (see, e.g., [176, 186, 272, 290, 292, 303, 323]).

These advantages allow one to adopt the local tuning approach for an efficient solving different univariate and multidimensional LGO problems (see, e.g., [175, 186, 216, 272, 273, 276, 278, 290, 292, 303, 323]).

To conclude our presentation of the Lipschitz geometric ideas for solving the LGO problem (2.2)–(2.3), let us return to the problem of finding the minimal root of an equation with multiextremal non-differentiable left part (see (2.1) and Fig. 2.1 in Chap. 1). A possible fast and efficient algorithm for solving this important practical

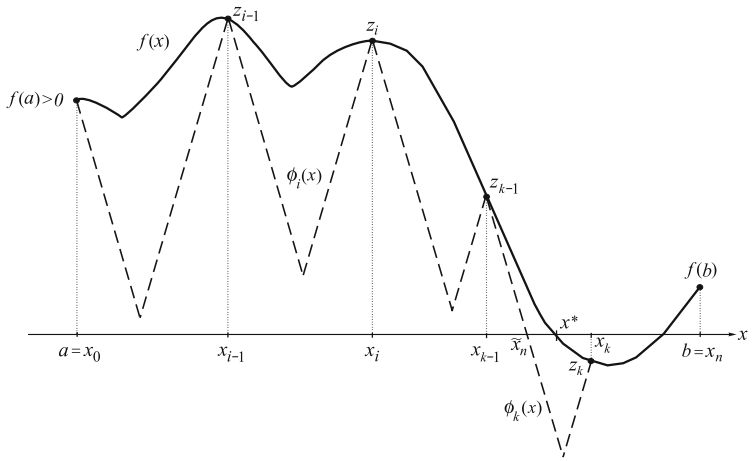

Fig. 2.6 Finding the minimal root of equation $f(x) = 0$ with multiextremal non-differentiable *left* part by a geometric method

problem can be developed as follows (see, e.g., [290, 323]). Let us suppose that the objective function $f(x)$ has been already evaluated at some trial points x_i, $0 \le i \le n$, and $z_i = f(x_i)$ (see Fig. 2.6). For every interval $[x_{i-1}, x_i]$, $1 \le i \le n$, a piecewise linear function $\phi_i(x)$ is constructed (it is drawn by dashed line in Fig. 2.6) by using the Lipschitz information in such a way that $\phi_i(x) \le f(x)$, $x \in [x_{i-1}, x_i]$. By knowing the structure of the auxiliary function $\phi_i(x)$, $1 \le i \le n$, it is possible to determine the minimal index $k \ge 1$, such that the equation $\phi_k(x) = 0$ has a solution (point $\tilde{x}_n$ in Fig. 2.6) over $[x_{k-1}, x_k]$. Adaptively improving the set of functions $\phi_i(x)$, $1 \le i \le n$, by adding new trial points $\tilde{x}_n$, $n > 1$, we improve both our lower approximation of $f(x)$ and the current solution to the problem. In this manner, such a geometric method either finds the minimal root of the equation $f(x) = 0$ or determines the global minimizer of $f(x)$ (in the case when the equation under consideration has no roots on the given interval). Its performance is significantly faster in comparison with the methods traditionally used by engineers for solving this problem. The usage of the local tuning technique on the behavior of $f(x)$ allows one to obtain a further acceleration of the search (see, e.g., [218, 290, 323]).

2.3 Geometric LGO Methods for Differentiable Functions with the Lipschitz First Derivatives

The restriction of the class of the objective functions (i.e., the examination of the LGO problem (2.2), (2.4) rather than the problem (2.2), (2.3)) opens new opportunities for developing efficient geometric LGO methods. In fact, if at each point $x \in [a, b]$ it is possible to evaluate both the objective function $f(x)$ and its first derivative $f'(x)$,

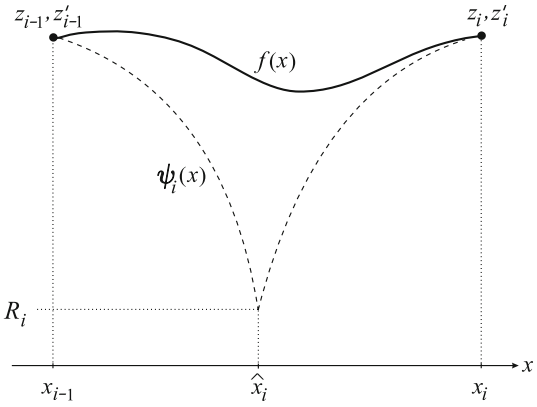

Fig. 2.7 Non-smooth piecewise quadratic auxiliary function $\psi_i(x)$ which can be constructed over a sub-interval $[x_{i-1}, x_i]$ for a function $f(x)$ with the Lipschitz first derivative

this gives the opportunity to obtain more information about the problem (especially, regarding its local properties represented by the derivative values). The usage of this information allows one to construct auxiliary functions that fit closely the objective function and, therefore, to accelerate the global search.

The geometric approach for solving the LGO problem (2.2), (2.4) has obtained a strong impact to expansion after publication of the papers [34, 116]. In these articles, non-smooth piecewise quadratic minorants have been used to approximate the behavior of the objective function $f(x)$ from (2.2) by using the Lipschitz condition (2.4).

In Fig. 2.7, an example of such a lower bounding function $\psi_i(x)$ over a sub-interval $[x_{i-1}, x_i]$ is shown. Both the objective function $f(x)$ from (2.2) and its first derivative $f'(x)$ satisfying the Lipschitz from (2.4) have been evaluated at two points x_{i-1} and x_i of the search interval $[a, b]$ with the corresponding values $z_{i-1} = f(x_{i-1})$, $z'_{i-1} = f'(x_{i-1})$ and $z_i = f(x_i)$, $z'_i = f'(x_i)$. If an overestimate $m \geq K$ is known, the lower bounding function $\psi_i(x)$ for $f(x)$ over $[x_{i-1}, x_i]$ can be constructed using the Lipschitz condition (2.4) and the Taylor formula for $f(x)$ as follows:

$$\psi_i(x) = \max \{ z_{i-1} + z'_{i-1}(x - x_{i-1}) - 0.5m(x - x_{i-1})^2, \\ z_i - z'_i(x_i - x) - 0.5m(x_i - x)^2 \}. \tag{2.14}$$

The lower bound value R_i for $f(x)$ over $[x_{i-1}, x_i]$ (the sub-interval characteristic) is shown in Fig. 2.7. It can be explicitly calculated in a way similar to (2.6) and, as a rule, it results to be closer to the minimal value of $f(x)$ over $[x_{i-1}, x_i]$ than the piecewise linear minorants (2.5). This can accelerate the global search being a natural consequence of the availability and the usage of more complete information about the function in the LGO problem (2.2), (2.4) with respect to the problem (2.2), (2.3).

As in the case of the LGO problem (2.2), (2.3), various approaches can be used to obtain an estimate m of the Lipschitz constant K for constructing auxiliary functions over the whole interval $[a, b]$ as the union of functions $\psi_i(x)$ from (2.14). For example, in [34] (see also [14, 15]), the constant K from (2.4) is supposed to be a

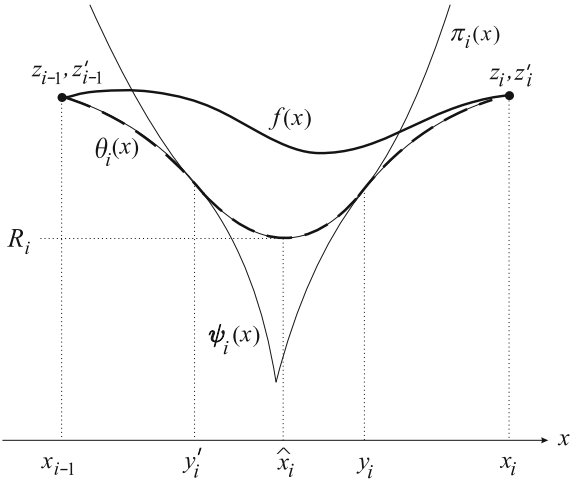

Fig. 2.8 Smooth piecewise quadratic auxiliary function $\theta_i(x)$ (*dashed line*) for the objective function $f(x)$ (*thick line*) with the Lipschitz first derivative over $[x_{i-1}, x_i]$

priori known. In [116] (see also [323]), an adaptive global estimate m of the constant K during the function minimization is proposed.

In [270, 275], it was shown that the Lipschitz constant K can be estimated more accurately over the whole interval $[a, b]$ in comparison with [116] and that the local tuning approach can be used in a similar to (2.9) manner, thus providing a significant acceleration of the global search (see also [290, 323]). It is important to emphasize that in order to ensure the convergence to the global minimizer x^* from (2.2), it is not necessary to estimate correctly the global Lipschitz constant K (it may be underestimated) during the execution of the local tuning algorithm but it is sufficient to have an overestimate only of the local Lipschitz constant over a sub-interval containing the point x^*.

It is evident from Fig. 2.7 that the auxiliary functions based on (2.14) are not smooth at the points $\hat{x}_i$ in spite of the smoothness of the objective function $f(x)$ over $[a, b]$. In [271, 276], it has been demonstrated how to obtain smooth auxiliary functions making them closer to $f(x)$ than those shown in Fig. 2.7 and, therefore, accelerating the global search (see also [15, 119, 188, 200, 320] where similar constructions are discussed). A general scheme describing the methods using smooth bounding procedures has been introduced in [271, 276] with several approaches for the Lipschitz constant estimation (a priori given, global, and local estimates were considered). Let us present this scheme.

The construction of a smooth auxiliary function over $[x_{i-1}, x_i]$ is based on the following considerations. The objective function $f(x)$ should be strictly above the function $\psi_i(x)$ for all $x \in (y'_i, y_i)$ (see Fig. 2.8) because due to (2.4) its curvature is bounded by a parabola

$$\pi_i(x) = 0.5mx^2 + b_i x + c_i,$$

where the unknowns y_i', y_i, b_i, and c_i can be determined by solving the following system of equations:

$$\begin{cases} \psi_i(y_i') = \pi_i(y_i'), \\ \psi_i(y_i) = \pi_i(y_i), \\ \psi_i'(y_i') = \pi_i'(y_i'), \\ \psi_i'(y_i) = \pi_i'(y_i). \end{cases}$$

Here the first equation provides the coincidence of $\psi_i(x)$ and $\pi_i(x)$ at the point y_i' and the third one provides the coincidence of their derivatives $\psi_i'(x)$ and $\pi_i'(x)$ at the same point. The second and fourth equations provide the fulfillment of analogous conditions at the point y_i.

Thus, once the values y_i', y_i, b_i, and c_i are determined (see [271, 276, 290, 323] for detailed discussions describing how to get explicit formulae for them), it may be concluded that the following function

$$\theta_i(x) = \begin{cases} \psi_i(x), \ x \in [x_{i-1}, y_i') \cup [y_i, x_i], \\ \pi_i(x), \ x \in [y_i', y_i), \end{cases}$$

is a smooth piecewise quadratic auxiliary function for $f(x)$ over $[x_{i-1}, x_i]$ (see Fig. 2.8), i.e., that the first derivative $\theta_i'(x)$ exists over $x \in [x_{i-1}, x_i]$ and

$$f(x) \geq \theta_i(x), \quad x \in [x_{i-1}, x_i].$$

As demonstrated also by extensive numerical experiments (see, e.g., [45, 276, 290, 295, 323]), the performance of geometric methods for solving the LGO problem (2.2), (2.4) with smooth auxiliary functions overcomes that of geometric methods with non-smooth minorants. The usage of the local tuning technique, in its turn, ensures the further speed up of the methods, especially when a high accuracy of the problem solution is required.

It is worthy to mention that geometric methods based on construction of the Lipschitz piecewise quadratic auxiliary functions can be also applied to solving efficiently the problem of finding the minimal root of an equation with multiextremal differentiable left part, discussed above (see, e.g., [158, 282, 290, 323]).

Up to now, geometric methods for solving the LGO problem (2.2), (2.4) that use in their work an a priori given estimate of K from (2.4), its adaptive global estimate or adaptive estimates of local Lipschitz constants have been considered. Algorithms working with multiple estimates of the Lipschitz constant for $f'(x)$ chosen from a set of possible values were not known until 2009 (see [177]) in spite of the fact that a geometric method working in this way with the Lipschitz objective functions (the DIRECT method described in the previous section) has been proposed in 1993 (see [154]). The main obstacle in implementing such an algorithm for differentiable objective functions was a lack of an efficient procedure for determining sub-intervals to perform new trials. A geometric method resolving this problem in a simple way and evolving the DIRECT ideas to the case of the objective function having the Lipschitz first derivative has been introduced and studied in [177].

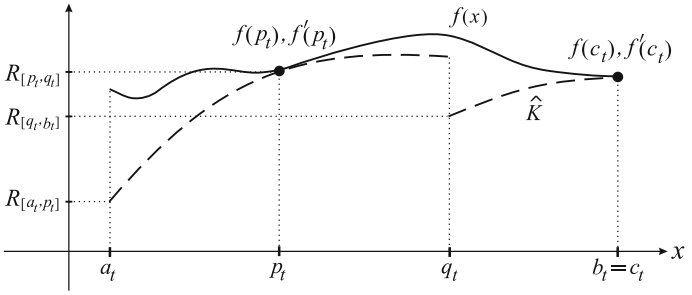

Fig. 2.9 Subdivision of a sub-interval $[a_t, b_t]$ in the situation where $f(x)$ and $f'(x)$ are evaluated at the point p_t given that they have been previously evaluated at the point b_t; a discontinuous piecewise quadratic auxiliary function for $f(x)$ is drawn by a *dashed line*

In this algorithm, the partition of the search interval $[a, b]$ into sub-intervals $[a_i, b_i]$ is iteratively performed by subdividing a selected sub-interval $[a_t, b_t]$ (see Fig. 2.9) into three equal parts of the length $(b_t - a_t)/3$, i.e.,

$$[a_t, b_t] = [a_t, p_t] \cup [p_t, q_t] \cup [q_t, b_t],$$
$$p_t = a_t + (b_t - a_t)/3, \quad q_t = b_t - (b_t - a_t)/3.$$

A new trial is carried out either at the point p_t (if both the objective function $f(x)$ and its first derivative $f'(x)$ have been evaluated over the sub-interval $[a_t, b_t]$ at the point b_t, see Fig. 2.9), or at the point q_t (if $f(x)$ and $f'(x)$ have been evaluated at the point a_t). Notice that an efficient partition strategy is adopted here since each selected sub-interval is subdivided by only one new trial point into three new sub-intervals. Remind that in a center-sampling partition strategy, in order to have three new sub-intervals, it is necessary to perform the function evaluations at two new points.

A series of non-smooth (discontinuous) piecewise quadratic functions corresponding to different estimates $\hat{K}$ of the Lipschitz constants K from (2.4) is then taken into account to find approximations $R_i = R_{[a_i,b_i]}$ of the lower bounds of $f(x)$ over sub-intervals $[a_i, b_i]$ (see Fig. 2.9); given an estimate $\hat{K}$, a lower bound R_i of the function values over the sub-interval $[a_i, b_i]$ can be calculated as

$$R_i = f(c_i) \pm f'(c_i)(b_i - a_i) - 0.5\hat{K}(b_i - a_i)^2, \tag{2.15}$$

where the sign '$-$' is used in the case of the right-end function evaluation, i.e., $c_i = b_i$, and the sign '$+$' is used in the case of the left-end function evaluation, i.e., $c_i = a_i$ (see Fig. 2.9).

In the DIRECT method for solving the problem (2.2), (2.3), the potentially optimal sub-intervals are determined as candidates for partitioning at each iteration by varying estimates $\hat{L}$ of the Lipschitz constant L from zero to infinity. The determination of these sub-intervals is a relatively simple technical task. In fact, it is sufficient

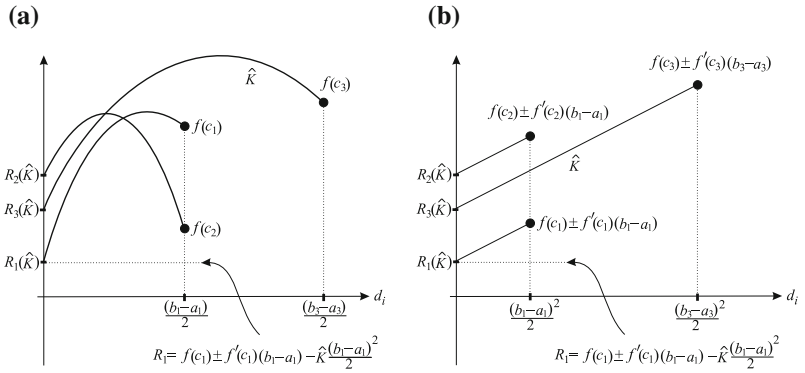

Fig. 2.10 Different ways of a graphical representation of sub-intervals in geometric methods for solving the LGO problem (2.2),(2.4) based on multiple estimates of the Lipschitz constant (2.4) with non-smooth piecewise quadratic auxiliary functions

to represent each sub-interval $[a_i, b_i]$ as a dot in a two-dimensional diagram with horizontal coordinate $c_i = (b_i - a_i)/2$ and vertical coordinate $f(c_i)$ and to locate the dots on a lower right convex hull of all the dots.

When solving the problem (2.2), (2.4), the same operation gives troubles due to the nonlinear part in (2.15). The difficulties in establishing the relation of domination (in terms of the lower bounds R_i) between sub-intervals of a current partition generated by the method are illustrated in Fig. 2.10a. Here, the sub-intervals $[a_1, b_1]$ and $[a_3, b_3]$ are the so-called *non-dominated sub-intervals* (i.e., sub-intervals having the smallest lower bound for some particular estimate of the Lipschitz constant for $f'(x)$; see [177] for details); but it is impossible to determine this fact from the diagram in Fig. 2.10a.

An efficient solution to this inconvenience is given by a new diagram in Fig. 2.10b where the intersection of the line with a slope $\hat{K}$ passed through any dot representing a sub-interval and the vertical coordinate axis gives us the lower bound (2.15) of $f(x)$ over the corresponding sub-interval. Thus, the procedure of selecting sub-intervals to be partitioned becomes simple in the case of differentiable objective functions too (the two non-dominated sub-intervals have in Fig. 2.10b coordinates $(0.5(b_1 - a_1), f(c_1) \pm f'(c_1)(b_1 - c_1))$ and $(0.5(b_3 - a_3), f(c_3) \pm f'(c_3)(b_3 - c_3))$, respectively) and another practical geometric method for solving the LGO problem (2.2), (2.4) is also available (see [177]). The usage of derivatives allows one to obtain, as it is expected, an acceleration in comparison with the DIRECT method.

Notice that the development of a geometric method constructing smooth auxiliary functions with multiple estimates of the Lipschitz constant (2.4) for derivatives still remains an open problem.

2.4 Acceleration Techniques Embedded in the Univariate Global Optimization

Very often in global optimization 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, function trials executed by a local search procedure are not used typically in the subsequent phases of the global search. Only some results of these trials (for instance, the current best found value) can be used, and the other ones are not taken into consideration by the global search method. 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 the 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 section, both frameworks, geometric and information, are taken into consideration and several derivative-free techniques proposed to accelerate the global search are studied and compared. Some promising 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 in the course of the global phases, as well.

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 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 [140]; 100 randomly generated functions from [243]; and four functions arising in practical problems [120, 290].

As mentioned in the previous sections (see formulae (2.9)–(2.13)), the *local tuning technique* proposed in [272, 273] adaptively estimates *local* Lipschitz constants at different sub-intervals of the search region during the course of the optimization process. The two components, λ_i and γ_i, are the main players in (2.9). 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 (2.9) and (2.12). 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 narrow, then the local information becomes relevant, the estimate γ_i becomes small (see (2.12)), and the local component λ_i assumes the key role. Thus, the local tuning technique automat-

ically 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 [175, 273, 276, 281, 287, 290, 300, 323]). 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 [115, 119, 320] for certain information algorithms. It proposes to use the following additive convolution instead of (2.9):

$$l_i = r \cdot \max\{\frac{1}{2}(\lambda_i + \gamma_i), \xi\}, \tag{2.16}$$

where r, ξ, λ_i, and γ_i have the same meaning as in (2.9). The first numerical examples executed in [115, 320] have shown a very promising performance of the "Additive" Local Tuning. These results induced us to execute in [301] a broad experimental testing and a theoretical analysis of the "Additive" Local Tuning. In particular, geometric methods using this technique were proposed in [301] (remind that the authors of [115, 320] have introduced (2.16) 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., [242, 246, 290, 303, 315, 323]) 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:

$$x^{k+1} = \frac{x_t + x_{t-1}}{2} - \frac{z_t - z_{t-1}}{2l_t}. \tag{2.17}$$

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 as in (2.9). Thus, if the estimate l_t is obtained by using (2.16), then the sum of the two addends (λ_i and γ_i) plays the leading role. Since the estimate γ_i is calculated as shown in (2.12), it can be very small for small intervals, creating so the possibility of occurrence of the situation $l_t \leq H_t$, leading to $x^{k+1} \notin (x_{t-1}, x_t)$ that provokes an incorrect work of the algorithm using (2.16). Obviously, by increasing the value of the parameter r this unhappy situation can be easily avoided but 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., [242, 290, 303, 323]) 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 the numerical part of this chapter together with other competitors (see Sect. 2.5).

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, a new strategy called the "*Maximum-Additive*" *Local Tuning* has been proposed in [301] 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:

$$l_i = r \cdot \max\{H_i, \frac{1}{2}(\lambda_i + \gamma_i), \xi\}, \tag{2.18}$$

where r, ξ, H_i, λ_i, and γ_i have the usual meaning (see (2.9)–(2.13)). It can be seen from (2.18) 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 the open interval (x_{t-1}, x_t). Notice that in (2.18) the equal usage of the local and global estimate is applied. Obviously, a more general scheme similar to (2.16) and (2.18) can be used, where $\frac{1}{2}$ is substituted by different weights for the estimates λ_i and γ_i, for example, as follows:

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

where r, H_i, λ_i, γ_i, and ξ are as in (2.18).

Let us now present another acceleration idea from [301]. It consists of 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 evaluations of $f(x)$ are allowed (see problem (P2) in the introductory Chap. 1). 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.

In order to overcome this drawback, two *local improvement techniques*, the "*optimistic*" and the "*pessimistic*" ones, that perform the local improvement within the global optimization scheme have been studied in [301]. 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.

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 [277] (see also [290]; 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 section. 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., $z_{i_{\min}} = f(x_{i_{\min}}) \le f(x_i)$, $i = 1, \ldots, k$ (if the current minimal value is attained at several trial points, then the smallest index is accepted as $i_{\min}$).
Suppose that $k \ge 2$ 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

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

Step 2. *Estimates of the Lipschitz constant*. Calculate the current estimates l_i of the Lipschitz constant for each sub-interval $[x_{i-1}, x_i]$, $i = 2, \ldots, k$, 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 (2.11).
Step 2.3. "*Maximum*" *Local Tuning*. Set l_i following (2.9).
Step 2.4. "*Additive*" *Local Tuning*. Set l_i following (2.16).
Step 2.5. "*Maximum-Additive*" *Local Tuning*. Set l_i following (2.18).

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

Step 3.1. *Geometric methods*.

$$R_i = \frac{z_i + z_{i-1}}{2} - l_i \frac{x_i - x_{i-1}}{2}. \tag{2.19}$$

Step 3.2. *Information methods*.

$$R_i = 2(z_i + z_{i-1}) - l_i(x_i - x_{i-1}) - \frac{(z_i - z_{i-1})^2}{l_i(x_i - x_{i-1})}. \tag{2.20}$$

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

Step 4.1. *Global phase*. Select the sub-interval $[x_{t-1}, x_t]$ 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 sub-interval between $[x_{i_{\min}}, x_{i_{\min}+1}]$ and $[x_{i_{\min}-1}, x_{i_{\min}}]$ starting from the right sub-interval $[x_{i_{\min}}, x_{i_{\min}+1}]$.
- **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

$$x_t - x_{t-1} \leq \delta, \tag{2.21}$$

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**

$$x_t - x_{t-1} \leq \varepsilon, \tag{2.22}$$

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 (2.17): $z^{k+1} := f(x^{k+1})$. Increase the iteration counter $k := k + 1$, and go to Step 1.

All the Lipschitz global optimization methods considered in this section are summarized in Table 2.2, from which concrete implementations of Steps 2–4 in the GS can be individuated. As shown experimentally in the numerical part of this section, the methods using an a priori given estimate of the Lipschitz constant or its global estimate loss, as a rule, in comparison with the 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 2.2 are furthermore specified (for the methods known in the literature, the respective references are provided).

1. **Geom-AL**: Piyavskij–Shubert's method with the a priori given Lipschitz constant (see [246, 306] and [180, 290] 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 [290]): GS with Step 2.2, Step 3.1, and Step 4.1.
3. **Geom-LTM**: Geometric method with the "Maximum" Local Tuning (see [273, 290, 323]): 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 [188, 290]): 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.

Table 2.2 Description of the methods considered in this section, 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					+		+			+

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 [290]): 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 [314, 315, 323]): GS with Step 2.2, Step 3.2, and Step 4.1.

14. **Inf-LTM**: Information method with the "Maximum" Local Tuning (see [272, 303, 323]): GS with Step 2.3, Step 3.2, and Step 4.1.

15. **Inf-LTA**: Information method with the "Additive" Local Tuning (see [115, 320]): 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 [186, 303]: 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 (2.2), (2.3) with $\delta = 0$ from (2.21) and $\varepsilon = 0$ from (2.22).

Theorem 2.1 *Assume that the objective function $f(x)$ satisfies the Lipschitz condition (2.3) 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 (2.9), (2.11), (2.18). 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 $f(x)$ and an iteration number k^* exist such that for all $k > k^*$ the inequality*

$$l_{j(k)} > L_{j(k)} \tag{2.23}$$

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 [277], the proofs of assertions 1–5 can be easily obtained as particular cases of the respective proofs in [277, 290]. □

Corollary 2.1 *Assertions* 1–5 *hold for methods belonging to the GS and using the "Additive" Local Tuning if the condition* $l_{i(k)} > r H_{i(k)}$ *is fulfilled for all* i *and* k.

Proof Fulfillment of the condition $l_{i(k)} > r H_{i(k)}$ 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 [277, 290]. □

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-neighborhood of a local minimizer (see Step 4 of the GS).

Next Theorem ensures existence of the values of the reliability parameter r satisfying condition (2.23), 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.

Theorem 2.2 *For any function* $f(x)$ *satisfying the Lipschitz condition (2.3) with* $0 < L < \infty$ *and for methods belonging to the GS and using one of the estimates (2.9), (2.11), (2.16), (2.18) there exists a value* r^* *such that, for all* $r > r^*$, *condition (2.23) holds.*

Proof It follows from, (2.9), (2.11), (2.16), (2.18), 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 (2.3) is finite and any positive value of the parameter r can be chosen in (2.9), (2.11), (2.16), (2.18), it follows that there exists an r^* such that condition (2.23) will be satisfied for all global minimizers for $r > r^*$. □

2.5 Numerical Illustrations

Seven series of numerical experiments were executed in [301] (see also [299]) on the following three sets of test functions to compare the described 22 global optimization methods:

1. the widely used set of 20 test functions from [140];
2. 100 randomly generated Pintér's functions from [243];
3. four functions originated from practical problems: first two problems are from [290, page 113] and the other two functions from [123] (see also [120, 296]).

Table 2.3 Number of trials performed by the considered geometric methods without the local improvement on 20 tests from [140]

#	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**

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 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9 and 2.10, 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).

The first series of numerical experiments was carried out with geometric and information algorithms without the local improvement on 20 test functions from [140]. 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 [140]. For the methods Geom-GL, Geom-LTM, and Geom-LTMA, the reliability parameter

Table 2.4 Number of trials performed by the considered information methods without the local improvement on 20 tests from [140]

#	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 2.5 Results of numerical experiments with the considered geometric and information methods without the local improvement on 100 Pintér's test functions from [243]

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

$r = 1.1$ was used as recommended in [290]. 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-neighborhood of a global minimizer).

Table 2.6 Number of trials performed by the considered geometric and information methods without the local improvement on four applied test functions

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 2.7 Number of trials performed by the considered geometric and information methods with the *optimistic* local improvement on 20 tests from [140]

#	Geom LTIMO	Geom LTIAO	Geom LTIMAO	Inf LTIMO	Inf LTIAO	Inf LTIMAO
1	45	41	**35**	47	**35**	37
2	47	49	**35**	45	**37**	41
3	49	45	**39**	55	**45**	51
4	47	53	**43**	**49**	53	53
5	55	49	**47**	51	**47**	**47**
6	51	49	**45**	47	**43**	47
7	45	45	**39**	49	**37**	39
8	37	41	**35**	**41**	45	47
9	49	51	**41**	51	51	**40**
10	47	49	**41**	51	**43**	**43**
11	49	53	**45**	**55**	59	**55**
12	43	53	**35**	53	67	**45**
13	**51**	53	57	**41**	51	55
14	45	45	**43**	49	**43**	45
15	**45**	57	47	**45**	55	53
16	**49**	55	53	**47**	49	53
17	93	**53**	95	59	55	**53**
18	45	47	**37**	49	**41**	44
19	45	43	**35**	46	**33**	35
20	43	45	**37**	49	**35**	39
Avg	49.00	48.80	**44.20**	48.95	46.20	**46.10**

Table 2.8 Number of trials performed by the considered geometric and information methods with the *pessimistic* local improvement on 20 tests from [140]

#	Geom LTIMP	Geom LTIAP	Geom LTIMAP	Inf LTIMP	Inf LTIAP	Inf LTIMAP
1	49	46	**36**	47	38	**35**
2	49	50	**38**	47	37	**35**
3	165	212	**111**	177	**56**	57
4	56	73	**47**	51	56	**46**
5	63	66	**48**	57	47	**38**
6	70	71	**51**	64	46	**45**
7	54	53	**41**	51	39	**38**
8	157	182	**81**	163	116	**99**
9	53	57	**43**	52	52	**43**
10	56	59	**42**	52	43	**39**
11	100	114	**77**	95	78	**72**
12	93	97	**69**	87	73	**64**
13	97	86	**68**	55	52	**50**
14	58	197	**43**	60	46	**42**
15	79	120	**76**	79	82	**70**
16	97	115	**81**	71	66	**60**
17	140	189	**139**	127	129	**100**
18	55	60	**42**	51	**42**	**42**
19	52	50	**36**	46	33	**32**
20	54	56	**40**	51	**37**	40
Avg	79.85	97.65	**60.45**	74.15	58.40	**52.35**

This situation has happened for $r = 1.8$: the corresponding results are shown in the column Geom-LTA of Table 2.3.

As can be seen from Table 2.3, the performance of the method Geom-LTMA was better than the behavior of 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 $r = 1.1$ used by the other tested methods. 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 [140] are shown in Table 2.4. 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

Table 2.9 Results of numerical experiments with the considered geometric and information methods with the local improvement techniques on 100 Pintér's test functions from [243]

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
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

Table 2.10 Number of trials performed by the considered geometric and information methods with the local improvement techniques on four applied test functions

	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
	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
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**
	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

parameter $r = 2$ was used in the methods Inf-GL, Inf-LTM, and Inf-LTMA, as recommended in [290, 315, 323]. 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 2.4, 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 have also shown 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 2.5) was executed on the class of 100 Pintér's test functions from [243] 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 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 again from $r = 1.1$ to $r = 1.8$ (when all 100 problems have been solved). All the information methods were able to solve all 100 test problems with $r = 2$ (see Table 2.5). The average performance of the Geom-LTMA and the Inf-LTA methods was the best among the other considered geometric and information algorithms, respectively.

The third series of the experiments (see Table 2.6) was carried out on four applied test problems from [123, 290]. All the methods without the local improvement used in the previous two series of experiments 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 2.6) 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 has solved all the four applied problems also with a higher value $r = 2.3$ and was outrun by the Inf-LTMA method.

In the following series of experiments, the local improvement techniques were compared on the same three sets of test functions. In the fourth series (see Table 2.7), 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 [140]. 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, and for the methods Inf-LTIMAO and Inf-LTIAO r was increased to 2.3. As can be seen from Table 2.7, the best average result 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 2.8. The usual values $r = 1.1$ and $r = 2$ were used for the geometric (Geom-LTIMP and Geom-LTIMAP) and the information (Inf-LTIMP and Inf-LTIMAP) methods, respectively. The values of the reliability parameter ensuring the solution to all the test problems in the case of methods Geom-LTIAP and Inf-LTIAP were set as $r = 1.8$ and $r = 2.3$, respectively. As can be seen from Table 2.8, 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. The methods Geom-LTIMAP and Inf-LTIMAP showed the best performance with respect to the other geometric and information techniques, 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 2.9. 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 meth-

ods, until all 100 problems from the class were solved. It can be seen from Table 2.9, 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 optimistic and for the pessimistic strategies, respectively). However, the pessimistic strategy seemed to be more stable since its reliability parameter (providing solutions to all the problems) generally remained smaller than that of the optimistic strategy. In average, the Geom-LTMA and the Inf-LTA methods were the best among the other considered geometric and information algorithms, respectively.

Finally, the last, seventh, series of the experiments (see Table 2.10) was executed on the class of four applied test problems, 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 than that of the best pessimistic Geom-LTIMAP method).

Let us summarize the results of this section. Univariate derivative-free global optimization has been considered and several numerical methods belonging to the geometric and information classes of algorithms have been proposed and analyzed. Acceleration techniques (see, e.g., [301]) to speed up the global search have been described. 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 in [301] to construct 22 different global optimization algorithms having, however, a similar structure. As was shown, 9 instances of this mixture can lead to well-known global optimization methods, and the remaining 13 methods tested in the present section were first introduced in [301]. All of them have been studied theoretically in the previous section and numerically compared on 124 theoretical and applied benchmark tests. It has been shown that the introduced acceleration techniques allow the global optimization methods to significantly speed up the search with respect to some known algorithms.

http://www.springer.com/978-1-4939-7197-8

Deterministic Global Optimization
An Introduction to the Diagonal Approach
Sergeyev, Y.D.; Kvasov, D.E.
2017, X, 136 p. 39 illus., Softcover
ISBN: 978-1-4939-7197-8