Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematics and Computation 197 (2008) 622–632
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

Stopping rules for box-constrained stochastic global optimization

I.E. Lagaris, I.G. Tsoulos *

Department of Computer Science, University of Ioannina, P.O. Box 1186, Ioannina 45110, Greece

Abstract

We present three new stopping rules for *Multistart* based methods. The first uses a device that enables the determination of the coverage of the bounded search domain. The second is based on the comparison of asymptotic expectation values of observable quantities to the actually measured ones. The third offers a probabilistic estimate for the number of local minima inside the search domain. Their performance is tested and compared to that of other widely used rules on a host of test problems in the framework of *Multistart*.

Keywords: Stochastic global optimization; Multistart; Stopping rules

1. Introduction

The task of locating all the local minima of a continuous function inside a box-bounded domain, is frequently required in several scientific as well as practical problems. We will not dwell further on this, instead we refer to the article by [8]. The problem we are interested in, may be described as

$$\text{Given an objective function } f(x), x \in S \subset R^n, \text{find all its local minimizers } x_i^* \in S. \tag{1}$$

S will be considered herein to be a rectangular hyperbox in N dimensions. We limit our consideration to problems with a finite number of local minima. This is a convenient hypothesis as far as the implementation is concerned. We are interested in stochastic methods based on *Multistart*, a brief review of which follows.

The multistart algorithm

Step 0: Set $i = 0$ and $X^* = \emptyset$
Step 1: Sample x at random from S
Step 2: Apply a deterministic local search procedure (LS) starting at x and concluding at a local minimum x^*.

* Corresponding author.
E-mail address: itsoulos@cs.uoi.gr (I.G. Tsoulos).

0096-3003/$ - see front matter
doi:10.1016/j.amc.2007.08.001

Step 3: Check if a new minimum is discovered
If $x^* \notin X^*$ then
increment: $i \leftarrow i + 1$
set: $x_i^* = x^*$
add: $X^* \leftarrow X^* \cup \{x_i^*\}$
Endif
Step 4: If a stopping rule applies, STOP
Step 5: Go to Step-1

It would be helpful at this point to state a few definitions and terms to be used in the rest of the article. The "*region of attraction*" of a local minimum associated with a deterministic local search procedure LS is defined as

$$A_i \equiv \{x : x \in S, \mathrm{LS}(x) = x_i^*\}, \tag{2}$$

where LS(x) is the minimizer returned when the local search procedure LS is started at point x. If S contains a total of w local minima, from the definition above follows:

$$\cup_{i=1}^{w} A_i = S. \tag{3}$$

Let $m(A)$ stand for the *Lebesgue measure* of $A \subseteq R^n$. Since the regions of attraction for deterministic local searches do not overlap, i.e. $A_i \cap A_j = \emptyset$ for $i \neq j$, then from Eq. (3) one obtains:

$$m(S) = \sum_{i=1}^{w} m(A_i). \tag{4}$$

If a point in S is sampled from a uniform distribution, the apriori probability π_i that it is contained in A_i is given by $\pi_i = \frac{m(A_i)}{m(S)}$. If K points are sampled from S, the apriori probability that at least one point is contained in A_i is given by

$$1 - \left(1 - \frac{m(A_i)}{m(S)}\right)^K = 1 - (1 - \pi_i)^K. \tag{5}$$

From the above we infer that for large enough K, this probability tends to one, i.e. it becomes "asymptotically certain" that at least one sampled point will be found to belong to A_i. This holds $\forall A_i$, with $m(A_i) \neq 0$.

Good stopping rules are important and should combine reliability and economy. A reliable rule is one that stops only when all minima have been collected with certainty. An economical rule is one that does not waste a large number of local searches to detect that all minima have been found. Several stopping rules have been developed in the past, most of them based on Bayesian considerations [9,5,4,6] and they have been successfully used in practical applications. A review analyzing the topic of stopping rules is given in the book by Törn and Žilinskas [3]. We refer also to Hart [2] noting however that his stopping rules aim to terminate the search as soon as possible once the global minimum is found and they are not designed for the retrieval of all the local minima. We present three different stopping rules. In Section 2, a rule that relies on a coverage argument is presented. In Section 3, a rule based on the comparison of asymptotic to measured values of observable quantities is developed, and in Section 4, a probabilistic approach is employed to estimate the expected number of minimizers. We report in Section 5, results of numerical experiments in conjunction with the *Multistart* method.

2. The double-box stopping rule

The covered portion of the search domain is a key element in preventing waistfull applications of the local search procedure. A relative measure for the region that has been covered is given by

$$C = \sum_{i=1}^{w} \frac{m(A_i)}{m(S)}, \tag{6}$$

where w is the number of the local minima discovered so far. The rule would then instruct to stop further searching when $C \to 1$.

The quantity $\frac{m(A_i)}{m(S)}$ is not known and generally cannot be calculated, however asymptotically it can be approximated by the fraction $\frac{L_i}{L}$, where L_i is the number of points, started from which, the local search led to the local minimum x_i^*, and $L = \sum_{i=1}^{w} L_i$, is the total number of sampled points (or equivalently, the total number of local search applications). An approximation for C may then be given by

$$C \simeq \widetilde{C} = \sum_{i=1}^{w} \frac{L_i}{L}. \tag{7}$$

However the quantity $\sum_{i=1}^{w} \frac{L_i}{L}$ is by definition equal to 1, and as a consequence the covered space can not be estimated by the above procedure. To circumvent this, a larger box S_2 is constructed that contains S and such that $m(S_2) = 2 \times m(S)$. At every iteration, 1 point in S is collected, by sampling uniformly from S_2 and rejecting points not contained in S. Let the number of points that belong to $A_0 \equiv S_2 - S$ be denoted by L_0. The total number of sampled points is then given by $L = L_0 + \sum_{i=1}^{w} L_i$ and the relative coverage may be rewritten as

$$C = \frac{\sum_{i=1}^{w} m(A_i)}{m(S)} = 2 \sum_{i=1}^{w} \frac{m(A_i)}{m(S_2)}. \tag{8}$$

The quantity $\frac{m(A_i)}{m(S_2)}$ asymptotically is approximated by $\frac{L_i}{L}$, leading to

$$C \simeq \widetilde{C} = 2 \sum_{i=1}^{w} \frac{L_i}{L}. \tag{9}$$

After k iterations, let the accumulated number of points sampled from S_2 be M_k, k of which are contained in S. The quantity then: $\delta_k \equiv \frac{k}{M_k}$ has an expectation value $\langle \delta \rangle_k = \frac{1}{k} \sum_{i=1}^{k} \delta_i$ that asymptotically, i.e. for large k, tends to $\frac{m(S)}{m(S_2)} = \frac{1}{2}$.

The variance is given by $\sigma_k^2(\delta) = \langle \delta^2 \rangle_k - \langle \delta \rangle_k^2$ and tends to zero as $k \to \infty$. This is a smoother quantity than $\langle \delta \rangle_k$ (see Fig. 1), and hence better suited for a termination criterion. We permit iterating without finding new minima until $\sigma^2(\delta) < p\sigma_{\text{last}}^2(\delta)$, where $\sigma_{\text{last}}(\delta)$ is the standard deviation at the iteration during which the most recent minimum was found, and $p \in (0, 1)$ is a parameter that controls the compromise between an exhaustive search ($p \to 0$) and a search optimized for speed ($p \to 1$).

In Table 1 we list the results from the application of the double-box termination rule and the Multistart method in a series of test problems for different values of the parameter p. As p increases the method becomes

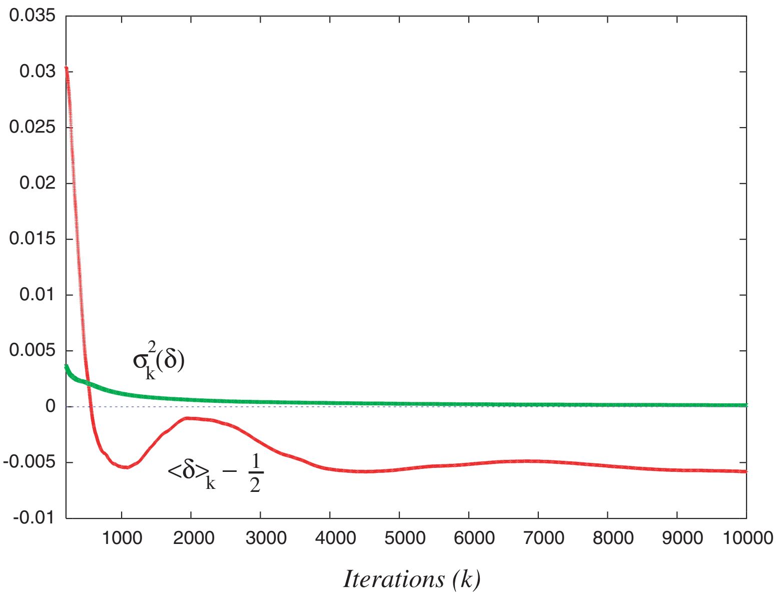

Fig. 1. Plots of $\langle \delta \rangle_k - \frac{1}{2}$ and $\sigma_k^2(\delta)$ versus k.

Table 1
Multistart with double-box rule for a set of p-values

FUNCTION	$p = 0.3$		$p = 0.5$		$p = 0.7$		$p = 0.9$	
	MIN	FC	MIN	FC	MIN	FC	MIN	FC
SHUBERT	400	1,150,243	400	577,738	400	322,447	395	139,768
GKLS(3,30)	30	961,269	29	302,583	23	41,026	15	3920
RASTRIGIN	49	50,384	49	19,593	49	13,581	49	10,034
Test2N(5)	32	78,090	32	30,607	32	20,870	32	13,462
Test2N(6)	64	85,380	64	34,840	64	22,535	64	15,393
GUILIN(20,100)	100	3,405,112	100	1,906,288	100	854,511	71	79,331
SHEKEL10	10	93,666	10	36,838	10	23,780	10	15,976

faster, but some local minima may be missed. The suggested value for general use is $p = 0.5$. Hence the algorithm may be stated as

1. Initially set $\alpha = 0$.
2. Sample from S_2 until a point falls in S as described above.
3. Calculate $\sigma^2(\delta)$.
4. Apply an iteration of Multistart (i.e. steps 2 and 3).
5. If a new minimum is found, set: $\alpha = p\sigma^2(\delta)$ and repeat from step 2.
6. STOP if $\sigma^2(\delta) < \alpha$, otherwise repeat from step 2.

3. The observables stopping rule

We have developed a scheme based on probabilistic estimates for the number of times each of the minima is being rediscovered by the local search. Let $L_1, L_2, \ldots, L_w$ be the number of local searches that ended-up to the local minima $x_1^*, x_2^*, \ldots, x_w^*$ (indexed in order of their appearance). Let $m(A_1), m(A_2), \ldots, m(A_w)$ be the measures of the corresponding regions of attraction, and let $m(S)$, be the measure of the bounded domain S. x_1^* is discovered for the first time with one application of the local search. Let n_2 be the number of the subsequent applications of the local search procedure spent, until x_2^* is discovered for the first time. Similarly denote by $n_3, n_4, \ldots, n_w$ the incremental number of local search applications to discover $x_3^*, x_4^*, \ldots, x_w^*$, i.e., x_2^* is found after $1 + n_2$ local searches, x_3^* after $1 + n_2 + n_3$, etc. $n_2, n_3, \ldots$ are counted during the execution of the algorithm, i.e. they are observable quantities. Considering the above and taking into account that we sample points using a uniform distribution, the expected number $L_J^{(w)}$ of local search applications that have ended-up to x_J^* at the time when the wth minimum is discovered for the first time, is given by

$$L_J^{(w)} = L_J^{(w-1)} + (n_w - 1)\frac{m(A_J)}{m(S)}. \tag{10}$$

The apriori probability that a local search procedure starting from a point sampled at random, concludes to the local minimum x_J^* is given by the ratio $m(A_J)/m(S)$, while the posteriori probability (observed frequency) is correspondingly given by $L_J/\sum_{i=1}^{w} L_i$. On the asymptotic limit the posteriori reaches the apriori probability, which implies $m(A_i)/m(A_j) = L_i/L_j$, which in turn permits substituting in Eq. (10) L_i in place of $m(A_i)$ leading to

$$L_J^{(w)} = L_J^{(w-1)} + (n_w - 1)\frac{L_J}{\sum_{i=1}^{w} L_i} = L_J^{(w-1)} + (n_w - 1)\frac{L_J}{\sum_{i=1}^{w} n_i} \tag{11}$$

with $n_1 = 1$, $J \leqslant w - 1$ and $L_w^{(w)} = 1$. Now consider that after having found w minima, an additional number of K local searches are performed without discovering any new minima. We denote by $\mathscr{L}_J^{(w)}(K)$ the expected number of times the Jth minimum is found at that moment. One readily obtains

$$\mathscr{L}_J^{(w)}(K) = \mathscr{L}_J^{(w)}(K-1) + \frac{L_J}{K + \sum_{i=1}^{w} n_i} \tag{12}$$

with $\mathscr{L}_J^{(w)}(0) = L_J^{(w)}$.

The quantity

$$E_2(w,K) \equiv \frac{1}{w}\sum_{J=1}^{w}\left(\frac{\mathscr{L}_J^{(w)}(K) - L_J}{\sum_{l=1}^{w} L_l}\right)^2 \tag{13}$$

tends to zero asymptotically, hence a criterion based on the variance $\sigma^2(E_2)$ may be stated as

$$\text{Stop if } \sigma^2(E_2) < p\sigma^2_{\text{last}}(E_2),$$

where $\sigma^2_{\text{last}}(E_2)$ is the variance of E_2 calculated at the time when the last minimum was retrieved. The value of the parameter p has the same justification as in the Double-Box rule and the suggested value is again $p = 0.5$, although the user may choose to modify it according to his needs.

4. The expected minimizers stopping rule

This technique is based on estimating the expected number of existing minima of the objective function in the specified domain. The search stops when the number of recovered minima, matches this estimate. Note that the estimate is updated iteratively as the algorithm proceeds. Let P^l_m denote the probability that after m draws, l minima have been discovered. Here by "draw" we mean the application of a local search, initiated from a point sampled from the uniform distribution. Let also π_k denote the probability that with a single draw the minimum located at x^*_k is found. This probability is apriori equal to $\pi_k = \frac{m(A_k)}{m(S)}$. The P^l_m probability can be recursively calculated by

$$P^l_m = \left(1 - \sum_{i=1}^{l-1}\pi_i\right)P^{l-1}_{m-1} + \left(\sum_{i=1}^{l}\pi_i\right)P^l_{m-1}. \tag{14}$$

Note that $P^0_1 = 0$, and $P^1_1 = 1$. Also $P^l_m = 0$ if $l\ m$, $P^0_m = 0 \forall m \geqslant 1$. The rational for the derivation of Eq. (14) is as follows. The probability that at the mth draw l minima are recovered, is connected with the probabilities at the level of the $(m-1)$th draw, that either $l-1$ minima are found (and the lth is found at the next, i.e. the mth, draw) or l minima are found (and no new minimum is found at the mth draw). The quantity $\sum_{i=1}^{l}\pi_i$ is the probability that one of the l minima is found in a single draw, likewise the quantity $1 - \sum_{i=1}^{l-1}\pi_i$ is the probability that none of the $l-1$ minima is found in a single draw. Combining these observations the recursion above is readily verified. Since P^l_m denote probabilities they ought obey the closure:

$$\sum_{l=1}^{m} P^l_m = 1. \tag{15}$$

To prove the above let us define the quantity $s_l = \sum_{i=1}^{l}\pi_i$. Perform a summation over l on both sides of Eq. (14) and obtain:

$$\sum_{l=1}^{m} P^l_m = \sum_{l=1}^{m} P^{l-1}_{m-1} - \sum_{l=1}^{m} s_{l-1}P^{l-1}_{m-1} + \sum_{l=1}^{m} s_l P^l_{m-1}. \tag{16}$$

Note that since $P^0_{m-1} = 0$ and $P^m_{m-1} = 0$ the last two sums in Eq. (16) cancel, and hence we get: $\sum_{l=1}^{m} P^l_m = \sum_{l=1}^{m-1} P^l_{m-1}$. This step can be repeated to show that

$$\sum_{l=1}^{m} P^l_m = \sum_{l=1}^{m-1} P^l_{m-1} = \cdots = \sum_{l=1}^{m-k} P^l_{m-k} = \sum_{l=1}^{1} P^l_1 = P^1_1 = 1.$$

The expected number of minima after m draws is then given by

$$\langle L \rangle_m \equiv \sum_{l=1}^{m} l P^l_m$$

and its variance by

$$\sigma^2(L)_m = \sum_{l=1}^{m} l^2 P^l_m - \left(\sum_{l=1}^{m} l P^l_m\right)^2. \tag{17}$$

The quantities π_i are unknown apriori and need to be estimated. Naturally the estimation will improve as the number of draws grows. A plausible estimate $\pi_i^{(m)}$ for approximating π_i after m draws, may be given by

$$\pi_i^{(m)} \equiv \frac{L_i^{(m)}}{m} \to \frac{m(A_i)}{m(S)} = \pi_i, \tag{18}$$

where $L_i^{(m)}$ is the number of times the minimizer x_i^* is found after m draws. Hence Eq. (14) is modified and reads:

$$P_m^l = \left(1 - \sum_{i=1}^{l-1} \pi_i^{(m-1)}\right) P_{m-1}^{l-1} + \left(\sum_{i=1}^{l} \pi_i^{(m-1)}\right) P_{m-1}^l. \tag{19}$$

The expectation $\langle L \rangle_m$ tends to w asymptotically. Hence a criterion based on the variance $\sigma^2(L)_m$, that asymptotically tends to zero, may be proper. Consequently, the rule may be stated as: **Stop if** $\sigma^2(L)_m < p\sigma^2(L)_{\text{last}}$, where again $\sigma^2(L)_{\text{last}}$ is the variance at the time when the last minimum was found and the parameter p is used in the same manner as before. The suggested value for p is again $p = 0.5$.

5. Computational experiments

We compare the new stopping rules proposed in the present article to three established rules that have been successfully used in a host of applications. If by w we denote the number of recovered local minima after having performed t local search procedures, then the estimate of the fraction of the uncovered space is given by [9]

$$P(w) = \frac{w(w+1)}{t(t-1)}. \tag{20}$$

The corresponding rule is then

$$\text{Stop when } P(w) \leqslant \epsilon, \tag{21}$$

ϵ being a small positive number. In our experiments we used $\epsilon = 0.001$. Ref. [5] showed that the estimated number of local minima is given by

$$w_{\text{est}} = \frac{w(t-1)}{t-w-2} \tag{22}$$

and the associated rule becomes

$$\text{Stop when } w_{\text{est}} - w \leqslant \frac{1}{2}. \tag{23}$$

In another rule [6] the probability that all local minima have been observed is given by

$$\prod_{i=1}^{w} \left(\frac{t-1-i}{t-1+i}\right) \tag{24}$$

leading to the rule:

$$\text{Stop when } \prod_{i=1}^{w} \left(\frac{t-1-i}{t-1+i}\right) > \tau, \tag{25}$$

τ tends to 1 from below.

Every experiment represents 100 runs, each with different seed for the random number generator. The local search procedure used is a BFGS version due to Powell [1]. We report the average number of the local minima recovered, as well as the mean number of functional evaluations. In Table 3 results are presented Multistart. We used a set of 21 test functions that cover a wide spectrum of cases, i.e. lower and higher dimensionality, small and large number of local minima, with narrow and wide basins of attraction, etc. These test functions are described in Appendix in an effort to make the article as self contained as possible.

Table 2
Multistart with Eq. (25) rule

FUNCTION	$\tau = 0.7$		$\tau = 0.8$		$\tau = 0.9$	
	MIN	FC	MIN	FC	MIN	FC
RASTRIGIN	49	168,103	49	268,721	49	568,843
SHUBERT	400	11,248,711	400	17,983,401	400	38,083,156
GKLS(3, 30)	18	10,615	24	27,910	28	77,326
GUILIN(10, 200)	200	6,627,109	200	10,589,110	200	22,429,999

Table 3
Multistart

FUNCTION	PCOV		KAN		DOUBLE		OBS		EXPM	
	MIN	FC	MIN	FC	MIN	FC	MIN	FC	MIN	FC
CAMEL	6	5642	6	2549	6	5503	6	2720	6	2916
RASTRIGIN	49	38,104	49	121,182	49	19,593	49	13,342	49	9007
SHUBERT	400	316,640	400	8,034,563	400	577,738	400	369,958	400	212,353
Hansen	527	426,056	527	1,422,0225	527	612,015	527	391,597	527	240,092
GRIEWANK2	528	565932	529	18,941,546	529	1,765,175	528	996,188	527	449,090
GKLS(3, 30)	16	5286	13	4249	29	302,853	23	84,291	25	96,260
GKLS(3, 100)	34	11,464	61	97,124	97	7,492,103	94	5,658,721	92	3,416,276
GKLS(4, 100)	20	6010	12	7816	95	8,629,052	73	5,290,564	93	6,358,587
GUILIN(10, 200)	191	354,650	200	4,736,609	200	3,351,391	200	2,178,890	199	1,136,783
GUILIN(20, 100)	96	263,869	100	1,760,826	100	1,906,288	100	973,307	99	655,374
Test2N(4)	16	17,373	16	18,716	16	19,424	16	5296	16	3970
Test2N(5)	32	37,639	32	78,931	32	30,607	32	10,700	32	7707
Test2n(6)	64	81,893	64	336,353	64	34,840	64	27,679	64	18,367
Test2n(7)	128	175,850	128	1,435,579	128	117,953	128	70,370	128	41,981
GOLDSTEIN	4	5906	4	3812	4	5391	4	3842	4	3850
BRANIN	3	2173	3	1782	3	1856	3	1782	3	1782
HARTMAN3	3	3348	3	2750	3	3509	3	2778	3	2772
HARTMAN6	2	3919	2	3851	2	3903	2	3907	2	3851
SHEKEL5	5	8720	5	4733	5	22,128	5	6430	5	8850
SHEKEL7	7	11,742	6	5485	7	30,702	7	7581	7	10,914
SHEKEL10	10	16,020	10	10,611	10	36,838	9	9812	10	12,751

Columns labeled as FUNCTION, MIN, FC list the function name, the number of recovered minimizers and the number of function calls. The labels PCOV and KAN refer to the stopping rules given in Eqs. (21) and (23), while the labels DOUBLE, OBS and EXPM to the proposed rules in an obvious correspondence.

Experiments have indicated that the rule in Eq. (25) is rather impractical, as can be readily verified by inspecting Table 2. Note the excessive number of function calls even for $\tau = 0.7$ (a value that is too low). Hence this rule is not included in Table 3, where the complete set of the test functions is used. As we can observe from Table 3 the new rules in most cases perform better, requiring fewer functional evaluations. However in the case of functions such as CAMEL, GOLDSTEIN, SHEKEL, HARTMAN, where only a few minima exist, the rules PCOV and KAN have a small advantage. Among the new rules there is not a clear winner, although EXPM seems to perform marginally better than the other two in terms of function evaluations. The rule DOUBLE seems to be more exhaustive and retrieves a greater number of minimizers.

6. Conclusions

We presented three new stopping rules for use in conjunction with *Multistart* for global optimization. These rules, although quite different in nature, perform similarly and significantly better than other rules that have been widely used in practice. The comparison does not render a clear winner among them, hence the one that is more conveniently integrated with the global optimization method of choice may be used. Efficient stopping rules are important especially for problems where the number of minima is large and the objective function expensive. Such problems occur frequently in molecular physics, chemistry and biology where the interest is in collecting stable molecular conformations that correspond to local minimizers of the steric energy function [10–12]. Devising new rules and adapting the present ones to other stochastic global optimization methods is within our interests and currently under investigation.

Appendix A. Test functions

We list the test functions used in our experiments, the associated search domains and the number of the existing local minima

1. Rastrigin
   $f(x) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2)$,
   $x \in [-1,1]^2$ with 49 local minima.
2. Shubert
   $f(x) = -\sum_{i=1}^{2}\sum_{j=1}^{5} j\{\sin[(j+1)x_i] + 1\}$,
   $x \in [-10,10]^2$ with 400 local minima.
3. GKLS
   $f(x) = \mathrm{GKLS}(x,n,w)$, is a function with w local minima, described in [7].
   $x \in [-1,1]^n, n \in [2,100]$. In our experiments we considered the following cases:
   (a) $n = 3,\ w = 30$.
   (b) $n = 3,\ w = 100$.
   (c) $n = 4,\ w = 100$.
4. Guilin Hills
   $f(x) = 3 + \sum_{i=1}^{n} c_i \frac{x_i+9}{x_i+10} \sin\left(\frac{\pi}{1-x_i+1/(2k_i)}\right)$,
   $x \in [0,1]^n$, $c_i > 0$, and k_i are positive integers. This function has $\prod_{i=1}^{n} k_i$ minima. In our experiments we chose $n = 10$ and $n = 20$ and arranged k_i so that the number of minima is 200 and 100, respectively.
5. Griewank #2
   $f(x) = 1 + \frac{1}{200}\sum_{i=1}^{2} x_i^2 - \prod_{i=1}^{2} \frac{\cos(x_i)}{\sqrt{(i)}}$,
   $x \in [-100,100]^2$ with 529 minima.
6. Hansen
   $f(x) = \sum_{i=1}^{5} i\cos[(i-1)x_1 + i]\sum_{j=1}^{5} j\cos[(j+1)x_2 + j]$,
   $x \in [-10,10]^2$ with 527 minima.
7. Camel
   $f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$,
   $x \in [-5,5]^2$ with six minima.
8. Test2N
   $$f(x) = \frac{1}{2}\sum_{i=1}^{n} x_i^4 - 16x_i^2 + 5x_i$$
   with $x \in [-5,5]^n$. The function has 2^n local minima in the specified range. In our experiments we have used the values $n = 4,5,6,7$. These cases are denoted by Test2N(4), Test2N(5), Test2N(6) and Test2N(7), respectively.
9. Branin
   $f(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10$ with $-5 \leqslant x_1 \leqslant 10,\ 0 \leqslant x_2 \leqslant 15$. The function has three minima in the specified range.
10. Goldstein and Price
   $$f(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)][30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$
   The function has four local minima in the range $[-2,2]^2$.

11. Hartman3

$$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{3} a_{ij}(x_j - p_{ij})^2\right)$$

with $x \in [0,1]^3$ and

$$a = \begin{pmatrix} 3 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3 & 10 & 30 \\ 0.1 & 10 & 35 \end{pmatrix}$$

and

$$c = \begin{pmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{pmatrix}$$

and

$$p = \begin{pmatrix} 0.3689 & 0.117 & 0.2673 \\ 0.4699 & 0.4387 & 0.747 \\ 0.1091 & 0.8732 & 0.5547 \\ 0.03815 & 0.5743 & 0.8828 \end{pmatrix}.$$

The function has three minima in the specified range.

12. Hartman6

$$f(x) = -\sum_{i=1}^{4} c_i \exp\left(-\sum_{j=1}^{6} a_{ij}(x_j - p_{ij})^2\right)$$

with $x \in [0,1]^6$ and

$$a = \begin{pmatrix} 10 & 3 & 17 & 3.5 & 1.7 & 8 \\ 0.05 & 10 & 17 & 0.1 & 8 & 14 \\ 3 & 3.5 & 1.7 & 10 & 17 & 8 \\ 17 & 8 & 0.05 & 10 & 0.1 & 14 \end{pmatrix}$$

and

$$c = \begin{pmatrix} 1 \\ 1.2 \\ 3 \\ 3.2 \end{pmatrix}$$

and

$$p = \begin{pmatrix} 0.1312 & 0.1696 & 0.5569 & 0.0124 & 0.8283 & 0.5886 \\ 0.2329 & 0.4135 & 0.8307 & 0.3736 & 0.1004 & 0.9991 \\ 0.2348 & 0.1451 & 0.3522 & 0.2883 & 0.3047 & 0.6650 \\ 0.4047 & 0.8828 & 0.8732 & 0.5743 & 0.1091 & 0.0381 \end{pmatrix}.$$

The function has two local minima in the specified range.

13. Shekel-5

$$f(x) = -\sum_{i=1}^{5} \frac{1}{(x-a_i)(x-a_i)^{\mathrm{T}} + c_i}$$

with $x \in [0, 10]^4$ and

$$a = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \end{pmatrix}$$

and

$$c = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix}.$$

The function has five local minima in the specified range.

14. Shekel-7

$$f(x) = -\sum_{i=1}^{7} \frac{1}{(x-a_i)(x-a_i)^{\mathrm{T}} + c_i}$$

with $x \in [0, 10]^4$ and

$$a = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 3 & 5 & 3 \end{pmatrix}$$

and

$$c = \begin{pmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \\ 0.6 \\ 0.3 \end{pmatrix}.$$

The function has seven local minima in the specified range.

15. Shekel-10

$$f(x) = -\sum_{i=1}^{m} \left(\frac{1}{(x - A_i)(x - A_i)^{\mathrm{T}} + c_i} \right),$$

$$\text{where } m = 10, A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 5 & 3 & 3 \\ 8 & 1 & 8 & 1 \\ 6 & 2 & 6 & 2 \\ 7 & 3.6 & 7 & 3.6 \end{bmatrix}, c = \begin{bmatrix} 0.1 \\ 2 \\ 2 \\ 0.4 \\ 0.4 \\ 0.6 \\ 0.3 \\ 0.7 \\ 0.5 \\ 0.5 \end{bmatrix} \; x \in [0, 10]^4 \text{ with 10 minima.}$$

References

[1] M.J.D. Powell, A tolerant algorithm for linearly constrained optimization calculations, Math. Program. 45 (1989) 547.
[2] W.E. Hart, Sequential stopping rules for random optimization methods with applications to multistart local search, Siam J. Optim. 9 (1998) 270–290.
[3] A. Törn, A. Žilinskas, Global Optimization, Lecture Notes in Computer Science, 350, Springer, Heidelberg, 1987.
[4] B. Betrò, F. Schoen, Optimal and sub-optimal stopping rules for the multistart algorithm in global optimization, Math. Program. 57 (1992) 445–458.
[5] C.G.E. Boender, A.H.G. Kan Rinnooy, Bayesian stopping rules for multistart global optimization methods, Math. Program. 37 (1987) 59–80.
[6] C.G.E. Boender, H. Romeijn Edwin, Stochastic methods, in: R. Horst, P.M. Pardalos (Eds.), Handbook of Global Optimization, Kluwer, Dordrecht, 1995, pp. 829–871.
[7] M. Gaviano, D.E. Ksasov, D. Lera, Y.D. Sergeyev, Software for generation of classes of test functions with known local and global minima for global optimization, ACM Trans. Math. Softw. 29 (2003) 469–480.
[8] M. Pardalos Panos, H. Romeijn Edwin, Hoang Tuy, Recent developments and trends in global optimization, J. Comput. Appl. Math. 124 (2000) 209–228.
[9] Zieliński Ryszard, A statistical estimate of the structure of multiextremal problems, Math. Program. 21 (1981) 348–356.
[10] K.M. Wiberg, R.H. Boyd, Application of strain energy minimization to the dynamics of conformational changes, J. Am. Chem. Soc. 94 (1972) 8426.
[11] M. Saunders, Stochastic exploration of molecular mechanics energy surfaces. hunting for the global minimum, J. Am. Chem. Soc. 109 (1987) 3150–3152.
[12] M. Saunders, K.N. Houk, Yun-Dong Wu, W.C. Still, M. Lipton, G. Chang, W.G. Guida, Conformations of cycloheptadecane. A comparison of methods for conformational searching, J. Am. Chem. Soc. 112 (1990) 1419–1427.