Stopping Rules for Box-Constrained Stochastic Global Optimization

I. E. Lagaris* and 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. We report comparative results of the application of these rules on a host of test problems in the framework of two stochastic global optimization methods.

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

*Corresponding author. Email: lagaris@cs.uoi.gr

dwell further on this, instead we refer to the article by Pardalos et al (2000). The problem we are interested in, may be described as:

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

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^*.

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

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 spend a huge number of local searches to detect that all minima have been found. The **"region of attraction"** or the **"basin"** 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 $m(A)$ indicate 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 $m(A_i) \neq 0$.

The main disadvantage of *Multistart* is that the local search may be initiated from points belonging to the same basin of attraction, leading to the discovery of the same minimizer over and over, wasting computational resources. A way of curing that, would be to design algorithms capable of recognizing if a point belongs to the basin of an already found minimizer and hence avoid starting a local search from there. This however can be achieved only within a probabilistic framework, since otherwise it would mean the apriori knowledge of the different regions of attraction. In this line the clustering approach aims to form clusters of points that belong to the same region of attraction and then start a local search from only one point of each cluster. Several stopping rules have been developed in the past, most of them based on Bayesian considerations (Zieliński (1981); Boender and Rinnooy Kan (1987); Betrò and Schoen (1992); Boender and Romeijn (1995)) and they have been successfully used in practical applications. A review that analyzes the topic of stopping rules is given by *Törn* and *Žilinskas* in their book (*Törn* and *Žilinskas* (1987)). We refer also to Hart (Hart (1998)) 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 2 of the *Multistart* based methods.

2 The Double-Box Stopping Rule

This termination rule is tightly coupled to the sampling scheme. In the conducted experiments we need to know the uncovered portion of the search domain in order to prevent the local search procedure from performing unnecessary computations. A relative measure for the region that remains uncovered is given by:

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

where w is the number of the discovered so far local minima. The quantity $\frac{m(A_i)}{m(S)}$ can be approximated asymptotically by the fraction $\frac{L_i}{L}$ where L_i is the number of local searches that have ended up to the local minimum x_i^* and L is the total number of performed local searches. An approximation for U may then be given by:

$$U \simeq \tilde{U} = 1 - \sum_{i=1}^{w} \frac{L_i}{L} \tag{7}$$

Unfortunately in every iteration the quantity $\sum_{i=1}^{w} \frac{L_i}{L}$ has the value of 1, and as a consequence the uncovered space can not be estimated with the above procedure. A larger box S_2 is constructed that contains S and such that $m(S_2) = 2 \times m(S)$. We consider that a unique (fake) local minimum is located in $A_0 = S_2 - S$. The volume of the region of attraction of the fake minimum is $m(A_0) = m(S)$. Hence we obtain:

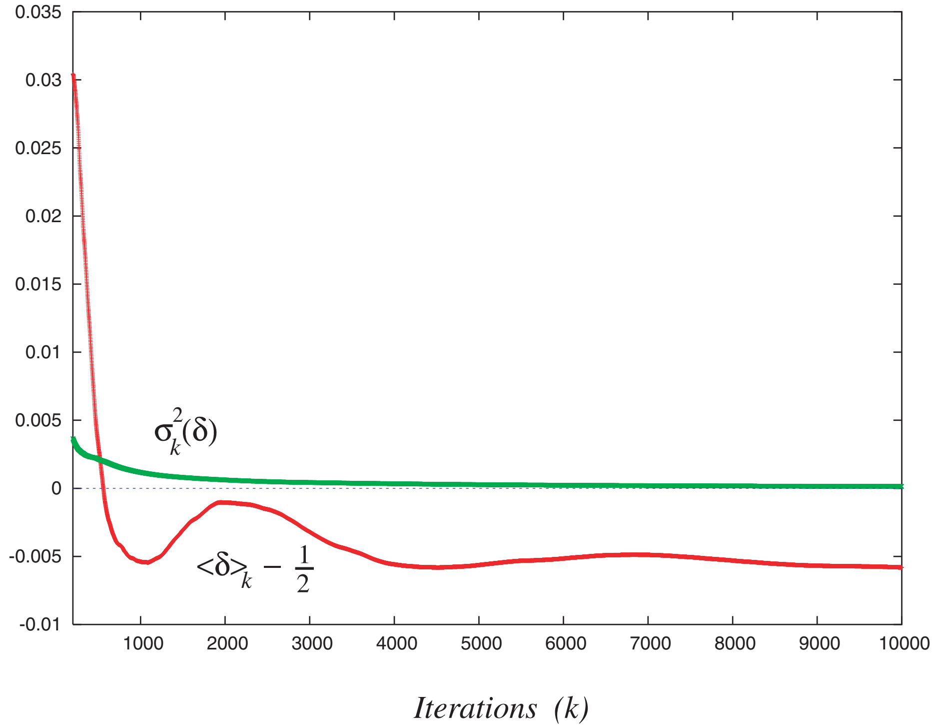
$$U = 1 - \frac{\sum_{i=0}^{w} m(A_i)}{m(S_2)} = 1 - \frac{m(A_0)}{m(S_2)} - \frac{\sum_{i=1}^{w} m(A_i)}{m(S_2)} = \frac{1}{2} - \frac{\sum_{i=1}^{w} m(A_i)}{m(S_2)} \tag{8}$$

The quantity $\frac{m(A_i)}{m(S_2)}$ is approximated again by the fraction $\frac{L_i}{L}$, where now L stands for the total number of local searches started in S_2 and hence U may be approximated by:

$$U \simeq \tilde{U} = \frac{1}{2} - \sum_{i=1}^{w} \frac{L_i}{L} \tag{9}$$

At every iteration, N points in S are collected, by sampling uniformly from S_2 and rejecting points not contained in S. (Note that for *Multistart* $N = 1$). After k iterations, let the accumulated number of points sampled from S_2 be M_k, kN of which are contained in S. The quantity then: $\delta_k \equiv \frac{kN}{M_k}$ has an expectation value $< \delta >_k = \frac{1}{k} \sum_{i=1}^{k} \delta_i$ that asymptotically, i.e.

Figure 1: Plots of $< \delta >_k - \frac{1}{2}$ and $\sigma_k^2(\delta)$ versus k



for large k, tends to $\frac{m(S)}{m(S_2)} = \frac{1}{2}$. The variance is given by $\sigma_k^2(\delta) =< \delta^2 >_k - < \delta >_k^2$ tends to zero as $k \to \infty$. This is a smoother quantity than $< \delta >_k$ (see figure 1) and hence better suited for a termination criterion. We permit iterating without finding new minima until $\sigma^2(\delta) < p\sigma_{last}^2(\delta)$, where $\sigma_{last}(\delta)$ is the standard deviation at the iteration during which the most recent minimum was found and $p \in (0, 1)$, 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 faster, but some local minima may not be found. The suggested value for general use is $p = 0.5$. Hence the algorithm may be stated as :

1. Initially set $\alpha = 0$

Table 1: Multistart with Double Box rule for a set of p-values.

	$p = 0.3$		$p = 0.5$		$p = 0.7$		$p = 0.9$	
FUNCTION	MIN	FC	MIN	FC	MIN	FC	MIN	FC
SHUBERT	400	1150243	400	577738	400	322447	395	139768
GKLS(3,30)	30	961269	29	302583	23	41026	15	3920
RASTRIGIN	49	50384	49	19593	49	13581	49	10034
Test2N(5)	32	78090	32	30607	32	20870	32	13462
Test2N(6)	64	85380	64	34840	64	22535	64	15393
Guilin(20,100)	100	3405112	100	1906288	100	854511	71	79331
Shekel10	10	93666	10	36838	10	23780	10	15976

2. Sample N points as described above.

3. Calculate $\sigma^2(\delta)$

4. Apply an iteration of a *Multistart*–like algorithm, neglecting its first step and using instead the N points already sampled.

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 $L_1, L_2, \cdots, L_w$ be the number of local searches that end–up to the local minima $x_1^*, x_2^*, \cdots, x_w^*$ that are indexed in order of their appearance. Let $m(A_1), m(A_2), \cdots, m(A_w)$ 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 $n_3, n_4, \cdots, n_w$ the incremental number of local search applications to discover $x_3^*, x_4^*, \cdots, x_w^*$. 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 w^{th} 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:

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

with $n_1 = 1, \quad J \leq w - 1$ 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 $\mathcal{L}_J^{(w)}(K)$ the expected number of times the J^{th} minimum is found at that moment. One readily obtains:

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

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

The quantity

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

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

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

where $\sigma^2_{last}(E_2)$ 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_m^l 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 $\pi_k = \dfrac{m(A_k)}{m(S)}$. The P_m^l probability can be recursively calculated by:

$$P_m^l = \left(1 - \sum_{i=1}^{l-1}\pi_i\right)P_{m-1}^{l-1} + \left(\sum_{i=1}^{l}\pi_i\right)P_{m-1}^{l} \tag{14}$$

Note that $P_1^0 = 0$, and $P_1^1 = 1$. Also $P_m^l = 0$ if $l > m$, $P_m^0 = 0$, $\forall m \geq 1$. The rational for the derivation of eq. (14) is as follows. The probability that at the m^{th} 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 l^{th} is found at the next, i.e. the m^{th}, draw) or l minima are found (and no new minimum is found at the m^{th} 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_m^l denote probabilities they ought obey the closure:

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

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:

$$\sum_{l=1}^{m} P_m^l = \sum_{l=1}^{m} P_{m-1}^{l-1} - \sum_{l=1}^{m} s_{l-1} P_{m-1}^{l-1} + \sum_{l=1}^{m} s_l P_{m-1}^{l} \tag{16}$$

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

$$\sum_{l=1}^{m} P_m^l = \sum_{l=1}^{m-1} P_{m-1}^l = \ldots = \sum_{l=1}^{m-k} P_{m-k}^l = \sum_{l=1}^{1} P_1^l = P_1^1 = 1$$

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

$$< L >_m \equiv \sum_{l=1}^{m} l P_m^l$$

and its variance by:

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

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:

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

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 $< L >_m$ 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** $\sigma^2(L)_m < p\sigma^2(L)_{last}$, where again $\sigma^2(L)_{last}$ 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 (Zieliński (1981)):

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

The corresponding rule is then:

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

ϵ being a small positive number. In our experiments we used $\epsilon = 0.001$. Boender and Rinnooy Kan (1987) 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:

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

In another rule (Boender and Romeijn (1995)) 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 (Powell (1989)). We report the average number of the local minima recovered, as well as the mean number of functional evaluations. In tables 4 and 5 results are presented for two global optimization methods, namely Multistart and TMLSL (Ali and Storey (1994)). We used a set of 21 test functions that cover a wide spectrum of cases, i.e. functions of lower or higher dimensionality, with a few or with many local minima, with narrow or wide basins of attraction etc. These test functions are described in the appendix in an effort to make the article as self contained as possible. 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 equations (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 equation (25) is rather impractical, as can be readily verified by inspecting tables 2 and 3. Note the excessive number of function calls even for $\tau = 0.7$ (a value that is too low) and also the premature termination for $\tau = 0.8$ in the case of TMLSL (table 3). Hence this rule is not included in tables 4 and 5, where the complete set of the test functions is used.

Table 2: Multistart with eq. (25) rule.

	τ=0.7		τ=0.8		τ=0.9	
FUNCTION	MIN	FC	MIN	FC	MIN	FC
RASTRIGIN	49	168103	49	268721	49	568843
SHUBERT	400	11248711	400	17983401	400	38083156
GKLS(3,30)	18	10615	24	27910	28	77326
GUILIN(10,200)	200	6627109	200	10589110	200	22429999

Table 3: TMLSL with eq. (25) rule.

	τ=0.7		τ=0.8		τ=0.9	
FUNCTION	MIN	FC	MIN	FC	MIN	FC
RASTRIGIN	32	7273	46	21954	49	58984
SHUBERT	369	502570	400	874785	400	1855688
GKLS(3,30)	7	751	10	1906	23	14753
GUILIN(10,200)	199	2501449	200	4029885	200	8580472

As we can observe from tables 4 and 5 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 in the case of Multistart a greater number of minimizers, while for TMLSL this is evened out.

Table 4: Multistart

	PCOV		KAN		DOUBLE		OBS		EXPM	
FUNCTION	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**
CAMEL	6	5642	6	2549	6	5503	6	2720	6	2916
RASTRIGIN	49	38104	49	121182	49	19593	49	13342	49	9007
SHUBERT	400	316640	400	8034563	400	577738	400	369958	400	212353
HANSEN	527	426056	527	14220225	527	612015	527	391597	527	240092
GRIEWANK2	528	565932	529	18941546	529	1765175	528	996188	527	449090
GKLS(3,30)	16	5286	13	4249	29	302853	23	84291	25	96260
GKLS(3,100)	34	11464	61	97124	97	7492103	94	5658721	92	3416276
GKLS(4,100)	20	6010	12	7816	95	8629052	73	5290564	93	6358587
GUILIN(10,200)	191	354650	200	4736609	200	3351391	200	2178890	199	1136783
GUILIN(20,100)	96	263869	100	1760826	100	1906288	100	973307	99	655374
Test2N(4)	16	17373	16	18716	16	19424	16	5296	16	3970
Test2N(5)	32	37639	32	78931	32	30607	32	10700	32	7707
Test2n(6)	64	81893	64	336353	64	34840	64	27679	64	18367
Test2n(7)	128	175850	128	1435579	128	117953	128	70370	128	41981
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	22128	5	6430	5	8850
SHEKEL7	7	11742	6	5485	7	30702	7	7581	7	10914
SHEKEL10	10	16020	10	10611	10	36838	9	9812	10	12751

Table 5: TMLSL

	PCOV		KAN		DOUBLE		OBS		EXPM	
FUNCTION	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**	**MIN**	**FC**
CAMEL	4	406	4	404	6	3415	6	7325	6	1815
RASTRIGIN	31	1821	15	1614	49	24123	49	12815	49	9199
SHUBERT	52	2971	236	226048	400	358623	400	321031	400	275773
HANSEN	66	3837	361	451662	527	811679	527	781209	527	501844
GRIEWANK2	295	56348	48	298174	528	1929165	527	1602467	526	591899
GKLS(3,30)	8	908	7	624	28	191867	28	364849	29	483227
GKLS(3,100)	10	1171	7	727	91	6683608	90	2535356	98	32034155
GKLS(4,100)	2	310	2	310	78	12269342	72	11199668	93	70398347
GUILIN(10,200)	139	76928	181	1600657	200	3470206	200	2650971	199	1752886
GUILIN(20,100)	74	72794	99	734645	100	1915478	100	1333144	99	867312
Test2N(4)	12	874	8	508	16	3821	16	2807	16	4162
Test2N(5)	32	3798	32	7750	32	8953	32	6634	32	6092
Test2N(6)	20	1513	11	1184	64	54555	64	34649	64	27500
Test2N(7)	104	10776	128	73947	128	88750	128	69371	128	52248
GOLDSTEIN	4	583	3	587	4	4606	4	11709	4	2522
BRANIN	3	354	3	359	3	1007	3	182919	3	1227
HARTMAN3	3	1567	3	1567	3	2014	3	18271	3	3857
HARTMAN6	2	685	2	677	2	912	2	87124	2	2094
SHEKEL5	4	446	4	446	5	6434	5	90356	5	3310
SHEKEL7	4	436	4	434	6	14617	7	161969	6	4363
SHEKEL10	5	412	5	406	9	18447	10	78645	9	10288

6 Conclusions

We presented three new stopping rules for use in conjunction with *Multistart*–based methods 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 (Wiberg and Boyd (1972); Saunders (1987, 1990)). Devising new rules and adapting the present ones to other stochastic global optimization methods is within our interests and currently under investigation.

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**.

   $f(x) = Gkls(x, n, w)$, is a function with w local minima, described in Gaviano et al (2003).

   $x \in [-1, 1]^n, n \in [2, 100]$. 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 6 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 $-5 \le x_1 \le 10,\ 0 \le x_2 \le 15$. The function has 3 minima in the specified range.

10. **Goldstein & Price**

$$\begin{aligned} f(x) &= [1 + (x_1 + x_2 + 1)^2 \\ &\quad (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \times \end{aligned}$$

$$[30 + (2x_1 - 3x_2)^2$$

$$(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$

The function has 4 local minima in the range $[-2, 2]^2$.

11. **Hartman3**

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

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 3 minima in the specified range.

12. **Hartman6**

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

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 2 local minima in the specified range.

13. **Shekel-5**.

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

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 5 local minima in the specified range.

14. **Shekel-7**.

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

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 7 local minima in the specified range.

15. **Shekel-10**.

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

where: $m = 10, A = \begin{bmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 5 & 3 & 3 \\ 8 & 1 & 8 & 1 \\ 6 & 2 & 6 & 2 \\ 7 & 3.6 & 7 & 3.6 \end{bmatrix} \quad c = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.4 \\ 0.4 \\ 0.6 \\ 0.3 \\ 0.7 \\ 0.5 \\ 0.5 \end{bmatrix}$

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

References

Powell, M.J.D. (1989) *A Tolerant Algorithm for Linearly Constrained Optimization Calculations*, Mathematical Programming **45**, 547.

Hart, W.E. (1998) *Sequential stopping rules for random optimization methods with applications to multistart local search*, Siam J. Optim. **9**, 270-290.

Törn A. and *Žilinskas* A. (1987),*Global Optimization, volume 350 of Lecture Notes in Computer Science*, Springer, Heidelberg.

Ali, M.M. and Storey, C.(1994), *Topographical Multilevel Single Linkage*, Journal of Global Optimization **5**, 349-358.

Betrò B., Schoen F. (1992), *Optimal and sub-optimal stopping rules for the Multistart algorithm in global optimization*, Mathematical Programming **57**, 445-458

Boender C.G.E., Kan Rinnooy A.H.G. (1987), *Bayesian stopping rules for multistart global optimization methods*, Mathematical Programming **37**, 59-80

Boender C.G.E. and Romeijn Edwin H. (1995), *Stochastic Methods*, in *Handbook of Global Optimization* (Horst, R. and Pardalos, P. M. eds.), Kluwer, Dordrecht, 829-871

Gaviano, M. and Ksasov, D.E. and Lera, D. and Sergeyev, Y.D.(2003), *Software for generation of classes of test functions with known local and global minima for global optimization*, ACM Trans. Math. Softw. **29**, 469-480

Pardalos Panos M., Romeijn Edwin H., Tuy Hoang (2000), *Recent developments and trends in global optimization*, Journal of Computational and Applied Mathematics **124**, 209-228

Zieliński Ryszard (1981), *A statistical estimate of the structure of multiextremal problems*, Mathematical Programming **21**, 348-356

Wiberg K. M. and Boyd R. H. (1972), *Application of Strain Energy Minimization to the Dynamics of Conformational Changes*, J. Am. Chem.Soc., **94**, 8426.

Saunders M. (1987), *Stochastic Exploration of Molecular Mechanics Energy Surfaces. Hunting for the Global Minimum*, J. Am. Chem. Soc., **109**, 3150-3152.

Saunders M., Houk K.N., Wu Yun-Dong, Still W. C., Lipton M., Chang G., Guida W.G. (1990), *Conformations of Cycloheptadecane. A comparison of Methods for Conformational Searching*, J. Am. Chem. Soc. **112**, 1419-1427.