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 offers a probabilistic estimate for the number of local minima inside the search domain. The third is based on the comparison of asymptotic expectation values of observable quantities to the actually measured ones. We report comparative results of the application of these rules in a host of test problems and various stochastic global optimization methods.

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 Pardalos et al Pardalos et al (2000). The problem we are interested in, may be described as:

$$\text{Find all } x_i^* \in S \subset R^n \text{ that satisfy:}$$
$$x_i^* = \arg\min_{x \in S_i} f(x), \quad S_i = S \cap \{x, |x - x_i^*| < \epsilon\} \tag{1}$$

S will be considered herein to be a rectangular hyperbox in N dimensions. This is not an essential hypothesis, however a convenient one as far as the implementation is concerned. We are interested in stochastic methods based on *Multistart*, a brief review of which follows.

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

The Multistart Algorithm

Step–0: Set $i = 0$ and $X^* = \emptyset$

Step–1: Sample x at random from S

Step–2: Apply a local search procedure L, starting at x and concluding at a local minimum x^*

Step–3: Check if a new minimum is recovered
If $x^* \ni X^*$ 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 local search procedure L is defined as:

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

where $L(x)$ is the minimizer returned when the local search procedure L 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 by definition the regions of attraction 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 p_i that it is contained in A_i is given by $p_i = \frac{m(A_i)}{m(S)}$. If K points are sampled from S, the apriori probability that at least one point is contained in A_i is given by:

$$1 - (1 - \frac{m(A_i)}{m(S)})^K = 1 - (1 - p_i)^K \tag{5}$$

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 simple *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 so 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 Kan (1987); Betrò and Schoen (1992); Boender and Romeijn (1995)) and they have been succesfully used in practical applications. 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 conjuction with various *Multistart* based methods.

2 The Double-Box Stopping Rule

This termination rule is tightly coupled to the sampling scheme. A larger box S_2 is constructed that contains S and such that $m(S_2) = 2 \times m(S)$. 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 simple *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 $< \delta >= \frac{1}{k} \sum_{i=1}^{k} \delta_i$ that asymptotically, i.e. for large k, tends to $\frac{m(S)}{m(S_2)} = \frac{1}{2}$. The variance given by $\sigma^2(\delta) =< \delta^2 > - < \delta >^2$

tends to zero as $k \to \infty$. The uncovered fraction of S is given by:

$$\frac{|m(S) - m(S_2) < \delta > |}{m(S)} = |1 - 2 < \delta > | \tag{6}$$

with a standard deviation $2\sigma(\delta)$, that yields an estimate for the measure of the uncovered fraction of the search space. We permit iterating without finding new minima until $\sigma^2(\delta) < \frac{1}{2}\sigma^2_{last}(\delta)$, where $\sigma_{last}(\delta)$ is the standard deviation at the iteration during which the most recent minimum was found. With the above definitions the algorithm can be stated as:

1. Initially set $\alpha = 0$
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: $\alpha = \sigma^2(\delta)/2$ 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 $L^{(w)}_J$ 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:

$$L^{(w)}_J = L^{(w-1)}_J + (n_w - 1)\frac{m(A_J)}{m(S)} \tag{7}$$

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 (observed frequency) probability is correspondingly given by $L_j/\sum_{i=1}^{w} L_i$. At 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. (7) L_i in place of $m(A_i)$ leading to:

$$\begin{aligned} L_J^{(w)} &= L_J^{(w-1)} + (n_w - 1)\frac{L_J}{\sum_{i=1}^{w-1} L_i} \\ &= L_J^{(w-1)} + (n_w - 1)\frac{L_J}{\sum_{i=1}^{w-1} n_i} \end{aligned} \tag{8}$$

with $n_1 = 1, \quad J \le w - 1$ 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:

$$\mathcal{L}_J^{(w)}(K) = \mathcal{L}_J^{(w)}(K-1) + \frac{L_J}{K + \sum_{i=1}^{w} n_i} \tag{9}$$

with $\mathcal{L}_J^{(w)}(0) = L_J^{(w)}$.
A criterion based on the variance $\sigma^2(E_2)$ of the quantity

$$E_2(w, K) \equiv \frac{1}{w}\sum_{J=1}^{w}\left(\frac{\mathcal{L}_J^{(w)}(K) - L_J}{\sum_{l=1}^{w} L_l}\right)^2 \tag{10}$$

may be stated as:
Stop if $\sigma^2(E_2) < \frac{1}{2}\sigma_{last}^2(E_2)$
where $\sigma_{last}^2(E_2)$ is the variance of E_2 calculated at the time when the last minimum was retrieved.

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:

$$P_m^l = (1 - \sum_{i=1}^{l-1} \pi_i) P_{m-1}^{l-1} + (\sum_{i=1}^{l} \pi_i) P_{m-1}^l \tag{11}$$

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. (11) is as follows. The probability that at the m^{th} draw l minima are recovered, is connected with the probabilites 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:

$$\sum_{l=1}^{m} P_m^l = 1 \tag{12}$$

To prove the above let us define the quantity $s_l = \sum_{i=1}^{l} \pi_i$. Perform a sumation over l on both sides of eq. (11) and obtain:

$$\sum_{l=1}^{m} P_m^l = \sum_{l=1}^{m} P_{m-1}^{l-1} - \sum_{l=1}^{m} s_{l-1} P_{m-1}^{l-1} + \sum_{l=1}^{m} s_l P_{m-1}^l \tag{13}$$

Note that since $P_{m-1}^0 = 0$ and $P_{m-1}^m = 0$ the last two sums in eq. (13) cancel, and hence we get: $\sum_{l=1}^{m} P_m^l = \sum_{l=1}^{m-1} P_{m-1}^l$. This can be repeated to show that the sum $\sum_{l=1}^{m} P_m^l$ is constant independent of m, and so it suffices to calculate it for $m = 1$, that can be trivially seen to equal 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:

$$\sigma^2(L)_m = \sum_{l=1}^{m} l^2 P_m^l - (\sum_{l=1}^{m} l P_m^l)^2 \tag{14}$$

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:

$$\pi_i^{(m)} \equiv \frac{L_i^{(m)}}{m} \rightarrow \frac{m(A_i)}{m(S)} = \pi_i \tag{15}$$

where $L_i^{(m)}$ is the number of times the minimizer x_i^* is found after m draws. Hence eq. (11) is modified and reads:

$$P_m^l = (1 - \sum_{i=1}^{l-1} \pi_i^{(m-1)}) P_{m-1}^{l-1} + (\sum_{i=1}^{l} \pi_i^{(m-1)}) P_{m-1}^l \tag{16}$$

The termination rule can be stated as:
Stop if $\sigma^2(L)_m < \frac{1}{2}\sigma^2(L)_{last}$, where again $\sigma^2(L)_{last}$ is the variance at the time when the last minimum was found.

5 Computational Experiments

We compare four stopping rules in the framework of three global optimization methods. Namely the rules presented in this article, along with Kan's stopping rule. This rule (Boender and Kan (1987)), uses for the expected number of minima the quantity $w_{est} = w(t-1)/(t-w-2)$, where t stands for the total number of points used in the reduced sample and w for the number of minima recovered so far. The employed methods are the simple *Multistart*, TMLSL (Ali and Storey (1994)), and the recently developed GTC method by Tsoulos and Lagaris (2004). In the tables we present the averages of ten runs for each experiment, each run using a different seed for the random number generator. Under the column headings FUNCTION, NOEM, NORM, NOFC and NOGC, we list correspondingly the problem names, the number of the existing minima, the number of recovered minima, the number of function calls and the number of gradient calls.

5.1 Simple Multistart

As can be deduced from tables I-IV,there are five problems that are solved completely i.e. all existing minima are found by Multistart with all the four stopping rules. For these cases direct comparison is possible (Shubert, Rastrigin, Guilin, Hansen, Camel). It is clear that for the first four of these, that correspond to a larger number of minima, the order of performance favors our “Expected Minimizers” stopping rule (EMSR), followed by the “Observables” (OSR), the “Double Box” (DBSR) and finally Kan’s rule. However for the small problem (Camel) Kan’s rule moves to the top of the list. From the incompletely solved problems the rules that gathered the most minima are DBSR and OSR followed by EMSR and then by Kan’s rule.

TABLE I. Multistart, Kan’s Stopping rule

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	7713921	7713921
Rastrigin	49	49	116122	116122
GKLS(3,30)	30	15	5085	5085
GKLS(3,100)	100	56	82622	82622
GKLS(4,100)	100	25	26083	26083
Guilin(20,100)	100	100	1741346	1741346
Griewank-2	529	529	18941873	18941873
Hansen	527	527	13661461	13661461
Camel	6	6	2830	2830
Shekel-10	10	9	10701	10701

TABLE II. Multistart, DOUBLE BOX

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	615904	615904
Rastrigin	49	49	18661	18661
GKLS(3,30)	30	28	195546	195546
GKLS(3,100)	100	97	7467472	7467472
GKLS(4,100)	100	97	9890042	9890042
Guilin(20,100)	100	100	1617099	1617099
Griewank-2	529	529	1832773	1832773
Hansen	527	527	543990	543990
Camel	6	6	6527	6527
Shekel-10	10	10	36198	36198

TABLE III. Multistart, OBSERVABLES

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	406010	406010
Rastrigin	49	49	13559	13559
GKLS(3,30)	30	27	51896	51896
GKLS(3,100)	100	97	5735713	5735713
GKLS(4,100)	100	98	10837290	10837290
Guilin(20,100)	100	100	805965	805965
Griewank-2	529	528	856865	856865
Hansen	527	527	352066	352066
Camel	6	6	5422	5422
Shekel-10	10	10	12645	12645

TABLE IV. Multistart, EXPECTED MINIMIZERS

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	213896	213896
Rastrigin	49	49	9369	9369
GKLS(3,30)	30	27	57159	57159
GKLS(3,100)	100	91	2061695	2061695
GKLS(4,100)	100	95	6494378	6494378
Guilin(20,100)	100	100	659681	659681
Griewank-2	529	528	439584	439584
Hansen	527	527	217501	217501
Camel	6	6	5179	5179
Shekel-10	10	9	17442	17442

5.2 Topographical Multilevel Single Linkage

From tables V-VIII we see that Kan's rule does not perform well with TMLSL. The rest three stopping rules solve completely the same five problems as in *Multistart*. The order of performance is again from best to worse EMSR, OSR, DBSR. For the incompletely solved problems, the three new rules recover almost the same number of minimizers.

TABLE V. TML, Kan's Stopping rule

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	159	154926	26306
Rastrigin	49	20	2597	969
GKLS(3,30)	30	9	821	521
GKLS(3,100)	100	12	1490	1004
GKLS(4,100)	100	3	535	135
Guilin(20,100)	100	99	400226	380386
Griewank-2	529	215	1470660	1246240
Hansen	527	423	528331	82855
Camel	6	4	415	215
Shekel-10	10	6	646	276

TABLE VI. TML, DOUBLE BOX

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	360416	61132
Rastrigin	49	49	27850	15056
GKLS(3,30)	30	29	125141	78494
GKLS(3,100)	100	94	16905021	9324132
GKLS(4,100)	100	94	9719338	5165518
Guilin(20,100)	100	100	1676572	1616692
Griewank-2	529	528	2006690	1693136
Hansen	527	527	808274	127356
Camel	6	6	3641	2881
Shekel-10	10	10	44162	39044

TABLE VII. TML, OBSERVABLES

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	328692	55712
Rastrigin	49	49	13405	5981
GKLS(3,30)	30	28	71765	45005
GKLS(3,100)	100	93	8788908	4837151
GKLS(4,100)	100	94	9348926	4935214
Guilin(20,100)	100	100	1293021	1244421
Griewank-2	529	528	1872271	1577327
Hansen	527	527	791335	124777
Camel	6	6	7286	6040
Shekel-10	10	10	41629	36975

TABLE VIII. TML, EXPECTED MINIMIZERS

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	277160	46806
Rastrigin	49	49	9564	3738
GKLS(3,30)	30	29	152654	95658
GKLS(3,100)	100	98	74265337	41033355
GKLS(4,100)	100	96	32536099	17259619
Guilin(20,100)	100	100	682524	653080
Griewank-2	529	526	602591	501127
Hansen	527	527	791177	124753
Camel	6	6	1803	1303
Shekel-10	10	9	9124	7481

5.3 Gradient–controlled, Typical–distance Clustering (GTC)

In tables VII-X the results with the GTC method are displayed. For the GKLS problems Kan's rule terminates prematurely. At the same problems the new rules perform quite well since they recover the existing minima to a large extend ($\approx$ 95%). For the completely solved problems the three new rules behave similarly. DBSR performs better in Rastrigin and Camel, OSR in Shubert, Hansen and Shekel and EMSR in Guilin. For the Guilin function, Kan's rule ranks second, but for all the other problems ranks last. For the GKLS problems, DBSR and EMSR have an edge over OSR. Over all the new rules in the GTC framework perform similarly. Again Kan's rule seems to be expensive and in some cases terminates prematurely.

TABLE VII. GTC, Kan's Stopping rule

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	187846	501505
Rastrigin	49	49	20657	24591
GKLS(3,30)	30	9	537	817
GKLS(3,100)	100	13	1103	1555
GKLS(4,100)	100	8	466	842
Guilin(20,100)	100	100	366280	382841
Griewank-2	529	529	4265130	4711675
Hansen	527	527	1223824	1731469
Camel	6	5	248	443
Shekel-10	10	9	4748	5046

TABLE VIII. GTC, DOUBLE BOX

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	31674	59044
Rastrigin	49	49	4449	5090
GKLS(3,30)	30	30	80275	132247
GKLS(3,100)	100	97	4160897	5663373
GKLS(4,100)	100	96	2564480	3878680
Guilin(20,100)	100	100	792052	827892
Griewank-2	529	529	1032445	1140113
Hansen	527	527	82572	109020
Camel	6	6	844	1705
Shekel-10	10	10	20226	21597

TABLE IX. GTC, OBSERVABLES

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	27021	45768
Rastrigin	49	49	4832	5542
GKLS(3,30)	30	29	44969	75270
GKLS(3,100)	100	96	2540167	3460014
GKLS(4,100)	100	96	2127872	3232187
Guilin(20,100)	100	100	646329	675530
Griewank-2	529	529	850197	939255
Hansen	527	527	62578	80650
Camel	6	6	993	2048
Shekel-10	10	10	18685	19946

TABLE X. GTC, EXPECTED MINIMIZERS

FUNCTION	NOEM	NORM	NOFC	NOGC
Shubert	400	400	49728	110785
Rastrigin	49	49	4527	5200
GKLS(3,30)	30	30	88805	148888
GKLS(3,100)	100	95	1192827	1626290
GKLS(4,100)	100	98	20955312	31663319
Guilin(20,100)	100	100	362688	379044
Griewank-2	529	528	407043	449271
Hansen	527	527	75446	98883
Camel	6	6	1484	3076
Shekel-10	10	10	20185	21539

6 Conclusions

We presented three new stopping rules for use in conjuction with *Multistart*–based methods for global optimization. These rules, although quite different in nature, perform similarly and significantly better than Kan's rule, which is the rule most widely used. 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 the stable conformations of large molecules, that correspond to local minimizers of the steric energy function. 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).
   $x \in [-1, 1]^n, n \in [2, 100]$

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 $n = 20$ and arranged k_i so that the number of minima is 100.

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. **Shekel-10**.

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

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

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

References

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

Tsoulos I. G., Lagaris I. E. (2004), *Gradient–Controlled, Typical–Distance Clustering for Global Optimization*, University of Ioannina, Dept. of Computer Science, Technical Report 4-2004.

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