ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
"Antonio Ruberti"
CONSIGLIO NAZIONALE DELLE RICERCHE

F. Lampariello, G. Liuzzi

A FILLING FUNCTION METHOD FOR UNCONSTRAINED GLOBAL OPTIMIZATION

R. 15, 2012

Francesco Lampariello – Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti", Consiglio Nazionale delle Ricerche (IASI-CNR), Viale manzoni 30, 00185 Rome, Italy.

Giampaolo Liuzzi – Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti", Consiglio Nazionale delle Ricerche (IASI-CNR), Viale manzoni 30, 00185 Rome, Italy.
giampaolo.liuzzi@iasi.cnr.it.

ISSN: 1128–3378

Collana dei Rapporti dell'Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", CNR

viale Manzoni 30, 00185 ROMA, Italy

tel. ++39-06-77161
fax ++39-06-7716461
email: iasi@iasi.cnr.it
URL: http://www.iasi.cnr.it

Abstract

The problem is considered of finding a global minimum point of a given continuously differentiable function. The strategy is adopted of a sequential nonmonotone improvement of local optima. In particular, to escape the basin of attraction of a local minimum, a suitable gaussian-based filling function is constructed using the quadratic model of the objective function, and added to the objective to fill the basin. Then, a procedure is defined where some new minima are determined, and that of them with the lowest function value is selected as the subsequent restarting point, even if its basin is higher than the starting one. The algorithm is applied to a set of test functions from the literature and the numerical results are reported.

Key words: Global optimization, unconstrained minimization, gradient methods.

A filling function method for unconstrained global optimization

F. Lampariello*, G.Liuzzi*

January 17, 2013

Abstract. The problem is considered of finding a global minimum point of a given continuously differentiable function. The strategy is adopted of a sequential nonmonotone improvement of local optima. In particular, to escape the basin of attraction of a local minimum, a suitable gaussian-based filling function is constructed using the quadratic model of the objective function, and added to the objective to fill the basin. Then, a procedure is defined where some new minima are determined, and that of them with the lowest function value is selected as the subsequent restarting point, even if its basin is higher than the starting one. The algorithm is applied to a set of test functions from the literature and the numerical results are reported.

Keywords: global optimization, unconstrained minimization, gradient methods.

1 Introduction

We consider the problem of finding a global minimum point of a given function, i.e., the unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1}$$

where $f : \mathbb{R}^n \to R$ is a continuously differentiable function. We assume, as usual, that $f(x)$ is radially unbounded, i.e., $f(x) \to +\infty$, as $\|x\| \to +\infty$, so that there exists a closed bounded domain $\Omega \subset \mathbb{R}^n$, containing all the minimum points of $f(x)$. Moreover we assume that $f(x)$ has a finite number of minimum points, and therefore, any of them is isolated.
As it is known, there are, in general, no global optimality conditions, and many approaches have been proposed to tackle this problem, ranging from a stochastic framework to a deterministic one (see, e.g. [17, 16, 8, 4, 15]).
Among the various approaches, we refer in particular to the strategy of global descent, i.e., of a sequential improvement of local optima. In this context, two main methods have been devised to determine, starting from a local minimum point x_o, another one with a lower objective function value.
The *Tunneling* method (introduced in [10]) constructs a new function of the kind $T(x; x_o) = (f(x) - f(x_o))/\|x - x_o\|^\lambda$, in order to find a point $\hat{x} \neq x_o$ where $f(\hat{x}) = f(x_o)$, by solving the eq. $T(x; x_o) = 0$. This last problem, however, has the same difficulty as the original one.
The *Filled function* method [6, 14, 13] constructs an auxiliary function from the objective $f(x)$ in such a way that it has a maximum in x_o, it does not admit minimum points or saddle points in higher basins than that in x_o, and it has a minimum point in a lower basin, if the latter exists. The

*Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti", Consiglio nazionale delle Ricerche. Viale Manzoni 30, 00185 Rome, Italy. Emails: francesco.lampariello@iasi.cnr.it, giampaolo.liuzzi@iasi.cnr.it

fulfilment of all these conditions through a suitable choice of the parameters in the filled function, may result a critical task.
We observe that both methods modify the objective function in the whole space.
The method described here consists in a sequential nonmonotone improvement of local optima, in the sense that even minimum points of $f(x)$ in higher basins than that of the current one are accepted, in order to allow the space exploration in a greater extent. Moreover, the objective function is modified only locally, by suitably filling it in larger and larger neighborhoods of the current minimum.

2 A gaussian-based filling function

We assume that a gradient-based local optimization routine, such as a steepest descent, Newton, or quasi-Newton algorithm, is available. Therefore, starting from a point randomly chosen in $\mathbb{R}^n$, let x_o be the local minimum reached by the local optimization routine. Moreover, let H_o be the Hessian matrix $\nabla^2 f(x_o)$ (or an appropriate approximation of it, e.g. computed by finite differences of the gradient or by quasi-Newton formulas), which is positive definite, since by assumption x_o is an isolated local minimum point.
In order to search over the various minima for finding that corresponding to the lowest f value, we have to escape the basin of attraction of the current minimum x_o. To do this, we consider the quadratic model of f at x_o

$$q(x; x_o) = \frac{1}{2}(x - x_o)^T H_o (x - x_o), \tag{2}$$

and we construct the following gaussian-based function

$$\varphi(x; x_o) = \frac{\beta}{1 - \exp(-\alpha q(x; x_o))} - \beta, \tag{3}$$

where α and β are positive scalars.
Then, we consider the new function obtained by adding $\varphi(x; x_o)$ to the original function $f(x)$

$$\hat{f}(x; x_o) = f(x) + \varphi(x; x_o). \tag{4}$$

Since $q(x_o; x_o) = 0$, we have $\varphi(x_o; x_o) = +\infty$, so that $\hat{f}(x_o; x_o) = +\infty$. Moreover, since $\varphi(x; x_o) > 0$, $\forall x$, $\hat{f}(x; x_o) > f(x)$, $\forall x$, $\varphi(x; x_o) \to 0$ as $\|x - x_o\| \to +\infty$, so that $\hat{f}(x; x_o) \to f(x)$, and far from x_o in any direction, where $\alpha q(x; x_o)$ is sufficiently high, we have $\varphi(x; x_o) \approx 0$ so that $\hat{f}(x; x_o) \approx f(x)$.
Therefore, function (4) is substantially the original function modified only locally, i.e., obtained by filling f within an ellipsoidal neighborhood of x_o, whose shape is determined by the matrix H_o, and hence according to the basin shape in its lowest side, and whose amplitude can be varied by taking different values of the parameter α. For large values of α, the original function is filled within a relatively small neighborhood of x_o, and viceversa. Thus, by applying the local search routine to $\hat{f}(x; x_o)$ starting from a point near x_o, we reach a minimum point $\hat{x}$ which can not be x_o, and which (in general) is not a minimum of f. Now, by reapplying the local search to the original function $f(x)$ starting from $\hat{x}$, either the same point x_o is reached, or a new minimum point is found. In the first case, the value of α is not sufficiently low to escape the basin of attraction of x_o, and a lower α value is needed.
The same purpose could be achieved by taking more simply $H_o = I$, the identity matrix, which corresponds to fill f within a spherical neighborhood of x_o. However, in this case, a much lower α value could be necessary, and the basin of a possible minimum point near x_o could also be filled, and hence that minimum lost.
Note that the local search routine can be easily applied to the modified function $\hat{f}(x; x_o)$, since

$$\nabla \hat{f}(x; x_o) = \nabla f(x) - \psi(x; x_o) H_o (x - x_o), \tag{5}$$

where $\psi(x;x_o) = \dfrac{\alpha\beta\exp(-\alpha q(x;x_o))}{[1-\exp(-\alpha q(x;x_o))]^2}$, and when a Newton-type method is used,

$$\nabla^2\hat{f}(x;x_o) = \nabla^2 f(x) + \psi(x;x_o)\left(\alpha\chi(x;x_o)H_o(x-x_o)(x-x_o)^T H_o - H_o\right),$$

where $\chi(x;x_o) = \dfrac{1+\exp(-\alpha q(x;x_o))}{1-\exp(-\alpha q(x;x_o))}$.

It is possible to show that there exist values of the parameters α and β such that the function $\hat{f}(x;x_o)$ does not have minimum points inside the basin of attraction of $f(x)$ at x_o, provided that the simple basin is bounded. We recall the following definitions (see, e.g., [6]).

Definition 1 *The basin of $f(x)$ at an isolated minimum point x_o is a connected domain $B(x_o)$ containing x_o, consisting of all the points such that, starting from any of them, the steepest descent trajectory of $f(x)$, i.e., the solution of the eq. $dx(t)/dt = -\nabla f(x(t))$, converges to x_o.*

Definition 2 *The simple basin of $f(x)$ at an isolated minimum point x_o is a connected domain $S(x_o) \subseteq B(x_o)$ such that, for any $x \neq x_o$, $(x-x_o)^T\nabla f(x) > 0$.*

Proposition 1 *Let x_o be an isolated minimum point of $f(x)$, and assume that the simple basin $S(x_o)$ is bounded. Then, for any given value of $\beta > 0$, there exist values of $\alpha > 0$ such that the modified function $\hat{f}(x;x_o)$ does not admit stationary points in $B(x_o)$.*

Proof. Let x be any point such that $x \in B(x_o)$ and $x \notin S(x_o)$, so that

$$(x-x_o)^T\nabla f(x) \leq 0.$$

Then, from (5), since $\psi(x;x_o) > 0$, and $(x-x_o)^T H_o(x-x_o) > 0$, $\forall x$, $x \neq x_o$, we have

$$(x-x_o)^T\nabla\hat{f}(x;x_o) = (x-x_o)^T\nabla f(x) - \psi(x;x_o)(x-x_o)^T H_o(x-x_o) < 0,$$

regardless the values of α and β.
Let now x be any point such that $x \in S(x_o)$. Then, in order to have

$$(x-x_o)^T\nabla\hat{f}(x;x_o) < 0,$$

it must be

$$\psi(x;x_o)(x-x_o)^T H_o(x-x_o) > (x-x_o)^T\nabla f(x),$$

and to ensure this it is sufficient that

$$\psi(x;x_o)(x-x_o)^T H_o(x-x_o) = 2\beta\frac{\alpha q(x;x_o)\exp(-\alpha q(x;x_o))}{[1-\exp(-\alpha q(x;x_o))]^2} > DL,$$

where $D = \max\limits_{x\in S(x_o)} \|x-x_o\|$, and $L = \max\limits_{x\in S(x_o)} \|\nabla f(x)\|$, and which are both finite, since $S(x_o)$ is bounded.
Denoting by $y = \alpha q(x;x_o)$, the condition can be rewritten as

$$b(y) = \frac{y\exp(-y)}{[1-\exp(-y)]^2} > \frac{DL}{2\beta} \tag{6}$$

and, since $\lim\limits_{y\to 0^+} b(y) = +\infty$, $\lim\limits_{y\to+\infty} b(y) = 0$, and $b(y)$ is continuous, for a given value of β, there exists a value $\tilde{y} > 0$ such that inequality (6) is verified for all $0 < y < \tilde{y}$, and hence a value $\tilde{\alpha} = \tilde{y}/\max\limits_{x\in S(x_o)} q(x;x_o)$ such that the same holds for all $0 < \alpha < \tilde{\alpha}$. Therefore, there exist values of the parameters such that, in any point x of $B(x_o)$, the direction $x - x_o$ is of descent for $\hat{f}(x;x_o)$.□

We remark that the function $\hat{f}(x;x_o)$ is not a *filled function* of $f(x)$ at x_o, since it may have minimum points or saddle points in a basin of $f(x)$ higher than $B(x_o)$.

3 A global optimization procedure

Using the simple device described above, our procedure for solving Problem (1) can be summarized as follows. Once a minimum x_o has been found, we perform a prefixed number p of local searches applied to the modified function $\hat{f}(x; x_o)$ by setting decreasing values of the parameter α, and the corresponding local searches applied to $f(x)$. The initial value of α is chosen in such a way that the neighborhood of x_o where $f(x)$ is filled, is sufficiently small in order to locate possible minima near x_o, while the last value is taken low to discover minima far from x_o. Then, a certain number $m \leq p$ of new minimum points of f are determined. Among these, that for which the function value is the lowest, say x_ℓ, is selected as new starting point of the procedure even if $f(x_\ell) > f(x_o)$. Obviously, if $f(x_\ell) < f(x_o)$, the point x_o is replaced by x_ℓ as the current best minimum point. At the subsequent restarts, the minimum point x_ℓ corresponding to the lowest function value is discarded if it coincides with one of the minima from which the process was already restarted, and the new restart is performed from the best minimum point among the remaining. When no new minimum point is found, i.e., $m = 0$ (to avoid as much as possible this last occurrence, the number p should be taken at least of the order of some tens), or when all the new minima were already used as restarting points, the procedure ends. Obviously, it is possible to define an algorithm where a further point is chosen at random and the above procedure is repeated. In this case, to avoid repetition of the same path, this further point is discarded if the new local minimum, determined by the local search routine, is near any one of the minima from which the procedure was already restarted.

As regards the parameters α and β in function (3), in order to establish the amplitude of the neighborhood of a minimum x_o where the function f is filled, let s be a distance from x_o, and, denoting by h_m the mean of the diagonal elements of H_o, $h_m = (1/n)\sum_{j=1}^{n} h_o(j,j)$, we take

$$\alpha = \frac{9}{h_m s^2}\,, \quad \beta = \frac{\sqrt{2\pi}}{3} s\,. \tag{7}$$

Thus, in the direction along which the curvature of f is h_m, the prefixed distance s corresponds to 3σ, where σ is the standard deviation of the Gaussian function in (3):

$$\varphi(x; x_o) = \frac{\beta}{1 - \beta f_G(x; x_o, \sigma)} - \beta,$$

where

$$f_G(x; x_o, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{\|x - x_o\|^2}{2\sigma^2}\right\}.$$

The procedure employs p increasing values of the stepsize s, and correspondingly the function f is filled within neighborhoods of x_o larger and larger.
The procedure is formalized in the following scheme.

Global Optimization ALgorithm (GOAL)

Data. Scalars s_o, $\varepsilon > 0$, an integer $p > 1$, $\Delta x \in \mathbb{R}^n$, and an increasing scalar function $\delta(t)$, with $\delta(0) = 0$.

Step 1. Generate at random a point $\tilde{x} \in \mathbb{R}^n$, compute $f(\tilde{x})$, and apply the local search routine for minimizing $f(x)$. Let x_o be the minimum reached.
Set $r = 1$, $x^{(r)} = x_o$, $f_{\min} = f(x_o)$ and $x_{\min} = x_o$.

Step 2. Compute $H_o = \nabla^2 f(x_o)$, and $h_m = (1/n)\sum_{j=1}^{n} h_o(j,j)$. Set $i = 1$ and $f_\ell = f(\tilde{x})$.

Step 3. Compute the parameters α and β in (7), where $s = s_o + \delta(i-1)$, and starting from $x_o + \Delta x$, minimize $\hat{f}(x; x_o)$. Let $\hat{x}_i$ be the minimum reached, minimize $f(x)$ starting from $\hat{x}_i$, and let x_i be the minimum obtained.

Step 4. If $\|x_i - x^{(k)}\| > \varepsilon$, for all $k = 1, \ldots, r$, and $f(x_i) < f_\ell$, set $f_\ell = f(x_i)$, and $x_\ell = x_i$.
Set $i = i + 1$, and if $i \leq p$ go to Step 3.

Step 5. If $f_\ell < f_{\min}$, set $f_{\min} = f_\ell$, and $x_{\min} = x_\ell$.
Set $r = r + 1$. If $f_\ell < f(\tilde{x})$, then set $x^{(r)} = x_\ell$, $x_o = x_\ell$, and go to Step 2; otherwise, stop.

The points $x^{(r)}$ are the points from which the filling procedure restarts. The increment Δx is chosen in such a way that $\|\Delta x\| \ll s_o$. The scalar function $\delta(t)$ defines the rule for increasing the amplitude of the neighborhood of $x^{(r)}$ where the function f is filled. The value of ε, which is the lowest distance between two minimum points over which they are considered distinct, should be chosen taking into account the stopping criterion of the local search routine.
Note that there is no stopping rule; GOAL stops automatically, since, by assumption, the number of local minima is finite.

4 A convergence property of GOAL

As already observed, it is possible to include GOAL into an algorithm where the procedure is repeated starting from a suitably generated sequence of initial points. In particular, we refer to the generator of starting points DIRGEN [3], which is a DIRECT-type strategy [9].
Differently from the simple uniform random sampling, this choice allows to densely cover the search domain (see, e.g. [9, 12]), so that the overall algorithm will enjoy the so-called everywhere dense convergence property [12].

More precisely, we introduce lower and upper bounds on the problem variables by difining the compact domain (which is the space exploration region):

$$\mathcal{D} = \{x \in \Re^n : l \leq x \leq u\},$$

where $l, u \in \Re^n$ with $l_i < u_i$, $i = 1, \ldots, n$, and such that $\Omega \subset \mathcal{D}$.
Since GOAL uses a gradient-based local search routine, we have that, for every local minimum $x^* \in \Omega$, an open neighborhood $\mathcal{L}$ of x^* exists, such that if the starting point x belongs to $\mathcal{L}$, the local routine converges to x^* ([1], Prop. 1.12). Hence, for a given global minimum x_g^*, let $\mathcal{L}_g$ be the corresponding convergence neighborhood. Then, we define the following algorithm.

GOALr (repeated)

Data. $l, u \in \Re^n$. Set $x_0^* = (u - l)/2$, $\mathcal{P}_0 = \emptyset$, and $j = 0$.

Step 1. Set $j = j + 1$. If $\mathcal{P}_{j-1} = \emptyset$, then Apply DIRGEN to generate a set of initial points $\mathcal{P}_j \subset \mathcal{D}$, otherwise set $\mathcal{P}_j = \mathcal{P}_{j-1}$.
Choose a point $\tilde{x}_j \in \mathcal{P}_j$, set $\mathcal{P}_j = \mathcal{P}_j \setminus \{\tilde{x}_j\}$ and let $x_{\min,j}$ be the final minimum point obtained by GOAL sarting from $\tilde{x}_j$ (instead of an initial random point in Step 1).

Step 2. If $f(x_{\min,j}) < f(x_{j-1}^*)$, set $x_j^* = x_{\min,j}$, otherwise set $x_j^* = x_{j-1}^*$, and go to Step 1.

By using the same arguments as in the proof of Prop. 3.1 in [11], it is possible to show the following convergence property.

Proposition 2 *Under the above assumptions, an iteration index k exists such that $\tilde{x}_k$ belongs to $\mathcal{L}_g$, and hence the final point $x_{\min,k}$ determined by GOAL starting from $\tilde{x}_k$ is a global minimum point of f.*

5 Numerical results

The problems used in our experimentation are those described in [11] (listed in Table 1 along with the number of variables and the known global minimum function value) and those in [7] (listed in Table 2). Moreover, we consider the class of test problems two times continuously differentiable introduced in [5]. This class is obtained by the GKLS generator which modifies a paraboloid with vertex in $x_p \in [-1,1]^n$ by means of a function constructed by using cubic polynomials. In this way the resulting function has a prefixed number v of local minima, and in the global minimizer x^* the function value is -1. Then, by specifying the number of variables n, the number v, the distance $r = \|x^* - x_p\|$, and the radius ρ of the spherical basin of attraction of the global minimum, the class consists of a set of 100 randomly generated smooth functions. Note that, by choosing ρ relatively small and v relatively large, the problems become very difficult. Moreover, the values of r and ρ are taken in such a way that the global minimizer is inside the admissible region $[-1,1]^n$ (see [5]).
In order to distinguish between simple and difficult problems, we have applied to them a multistart algorithm starting from $1000n$ initial points generated by DIRGEN and, from these, using the limited memory ℓ-BFGSB quasi-Newton method [2, 18] as a routine to reach a minimum point. We consider as "difficult" the problems where that algorithm fails to find the global minimizer. For the remaining problems, GOALr also finds the global solution.

The data in GOAL are taken as follows:

$$
\begin{aligned}
&s_o = 0.1,\\
&\varepsilon = 10^{-2},\\
&p = 25,\\
&\delta(t) = s_o(\gamma^t - 1), \quad \text{with } \gamma = 1.15.
\end{aligned}
$$

The choice of $\delta(t)$ is of an exponential-type function to allow space exploration in a great extent, and the minimization routine employed is again the ℓ-BFGSB quasi-Newton method.

In Table 3 we list the problems (in Tables 1 and 2) which turn out to be difficult according to the preceding criterion (i.e. 9 problems). GOALr fails only on one problem (Perm with $\beta = 0.5$) over them.
In Table 4, for the GKLS problems, we report the choice made of the parameter values together with the number of failures occurred by applying both algorithms.

In conclusions, from these results, it appears that our algorithm GOALr may represent a valuable alternative to other methods for the solution of continuous global optimization problems.

References

[1] D.P. Bertsekas. *Constrained Optimization and Lagrange Multipliers Methods.* Academic Press, NewYork, USA, 1982.

[2] R. H. Byrd, P. Lu, and J. Nocedal. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 16(5):1190–1208, 1995.

[3] E.F. Campana, G. Liuzzi, S. Lucidi, D. Peri, V. Piccialli, and A. Pinto. New global optimization methods for ship design problems. *Optimization and Engineering*, 10:533–555, 2009.

[4] Gintautas Dzemyda, Vydūnas Šaltenis, and Antanas Žilinskas, editors. *Stochastic and Global Optimization*. Series on Nonconvex Optimization and its Applications. Kluwer Academic Publisher, Dordrecht, The Netherlands, 2002.

[5] M. Gaviano, D.E. Kvasov, D. Lera, and Ya.D. Sergeyev. Algorithm 829: Software for generation of classes of test functions with known local and global minima for global optimization. *ACM Transactions on Mathematical Software*, 29(4):469–480, 2003.

[6] R. Ge. A filled function method for finding a global minimizer of a function of several variables. *Mathematical Programming*, 46:191–204, 1990.

[7] A. Hedar. http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO.htm.

[8] R. Horst, Panos M. Pardalos, and Nguyen Van Thoai. *Introduction to Global Optimization*. Series on Nonconvex Optimization and its Applications. Kluwer Academic Publisher, Dordrecht, The Netherlands, 2^{nd} edition, 2000.

[9] D.R. Jones, C.D. Perttunen, and B.E. Stuckman. Lipschitzian optimization without the lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993.

[10] A.V. Levy and A. Montalvo. The tunneling algorithm for the global minimization of functions. *SIAM Journal on Scientific and Statistical Computing*, 6:15–29, 1985.

[11] G. Liuzzi, S. Lucidi, and V. Piccialli. A direct-based approach exploiting local minimizations for the solution of large-scale global optimization problems. *Computational Optimization and Applications*, 45(2):353–375, 2010.

[12] G. Liuzzi, S. Lucidi, and V. Piccialli. A partition-based global optimization algorithm. *Journal of Global Optimization*, 48:113–128, 2010.

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

[14] S. Lucidi and V. Piccialli. New classes of globally convexized filled functions for global optimization. *Journal of Global Optimization*, 24(2):219–236, 2002.

[15] Panos M. Pardalos and H. Edwin Romeijn, editors. *Handbook of Global Optimization Volume 2*. Series on Nonconvex Optimization and its Applications. Kluwer Academic Publisher, Dordrecht, The Netherlands, 2002.

[16] János D. Pintér. *Global Optimization in Action. Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications*. Series on Nonconvex Optimization and its Applications. Kluwer Academic Publisher, Dordrecht, The Netherlands, 1996.

[17] A. Törn and A. Žilinskas. *Global Optimization*. Springer-Verlag, Berlin, 1989.

[18] C. Zhu, R. H. Byrd, and J. Nocedal. L-bfgs-b: Algorithm 778: L-bfgs-b, fortran routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4):550–560, 1997.

Problem	**n**	$\mathbf{f}^*$
Shubert	2	-186.7309
Shub. pen. 1	2	-186.7309
Shub. pen. 2	2	-186.7309
S-H. Camel B.	2	-1.0316
Goldstein-Price	2	3.0
Treccani mod.	2	0.0
Quartic	2	-0.352
Shekel ($m = 5, 7, 10$)	4	-10.1532
	4	-10.4029
	4	-10.5364
Espon. mod.	2,4	-1.0
Cos-mix mod.	2	-0.2
	4	-0.4
Hartman	3	-3.8627
Hartman	6	-3.3223
$5n$ loc-min	2,5,10,20	0.0
$10n$ loc-min	2,5,10,20	0.0
$15n$ loc-min	2,5,10,20	0.0
Griewank	2,5,10,20	0.0
Pinter	2,5,10,20	0.0
Rotated Griewank	2,10,30,50	-180.0
Ackley	2,10,30,50	0.0
Dixon Price	2,10,25,50	0.0
Easom	2	-1
Michalewics	2	-1.8013
Michalewics	5	-4.687658
Michalewics	10	-9.66015
Rastrigin	2,10,30,50	0.0

Table 1: Test Problems from [11]

Problem	**n**	$\mathbf{f}^*$
Beale	2	0.0
Bohachevsky 1,2,3	2	0.0
Booth	2	0.0
Colville	4	0.0
Perm $P^0_{n,\beta}$ ($\beta = 10$)	2,10,30,50	0.0
Perm $P_{n,\beta}$ ($\beta = 0.5$)	2,10,30,50	0.0
Powell	4,12,20,40	0.0
Powersum	4	0.0
Schwefel	2,5,10,20	0.0

Table 2: Test Problems from [7]

Problem	**n**
Easom	2
Griewank	2,5
Ackley	10,30,50
Dixon Price	25,50
Michalewics	10
Rastrigin	10,30,50
Schwefel	10,30,50
Perm $P^0_{n,\beta}$ ($\beta = 10$)	10,30,50
Perm $P_{n,\beta}$ ($\beta = 0.5$)	**10,30,50**

Table 3: Problems considered difficult as defined above. GOALr fails only on Perm $P_{n,\beta}$ with $\beta = 0.5$

				nr. of failures	
n	v	r	ρ	Multistart	GOALr
2	10^4	0.95	$0.03r$	32	0
3	10^3	0.95	$0.1r$	34	0
4	10^2	0.95	$0.2r$	37	16
5	10	0.95	$0.31r$	57	47

Table 4: Results for the GKLS problems