Journal of Global Optimization **26:** 167–182, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

A Sequential Convexification Method (SCM) for Continuous Global Optimization

WENXING ZHU* and QINGXIANG FU
Department of Computer Science and Technology, Fuzhou University, Fuzhou 350002, China
**E-mail: wxzhu@fzu.edu.cn*

Abstract. A new method for continuous global minimization problems, acronymed SCM, is introduced. This method gives a simple transformation to convert the objective function to an auxiliary function with gradually 'fewer' local minimizers. All Local minimizers except a prefixed one of the auxiliary function are in the region where the function value of the objective function is lower than its current minimal value. Based on this method, an algorithm is designed which uses a local optimization method to minimize the auxiliary function to find a local minimizer at which the value of the objective function is lower than its current minimal value. The algorithm converges asymptotically with probability one to a global minimizer of the objective function. Numerical experiments on a set of standard test problems with several problems' dimensions up to 50 show that the algorithm is very efficient compared with other global optimization methods.

Key words: Global minimization, Sequential convexification method, Auxiliary function, Fewer local minimizers

1. Introduction

Many engineering applications and social problems, especially in recent years the molecular structure prediction problems, can be formulated as nonlinear function optimization problems in which the function to be optimized possesses many local minimizers in the solution space. In most cases, it is desired to find the local minimum at which the function takes its lowest value, i.e., the global minimum. Many methods have been developed for continuous global optimization problems. Some of them try to escape from local minima by modifying the objective function to a function with 'fewer' local minimizers, and then developing algorithms to minimize the modified objective function to find local minimizers of the original objective function with lower function values.

The diffusion equation method [7, 8], the effective energy method [2, 3, 12] and a continuation based integral transformation scheme [4, 13] approximate the coarse structure of the original objective function using a parameterized set of smoothed objective function with 'fewer' local minimizers. All these methods transform the original objective function into a family of smoothed functions via integration of the original objective function. Such integrations are too expensive to compute at run time.

Several methods do not integrate the original objective function. The Terminal Repeller Unconstrained Sub-energy Tunneling [1] modifies the objective function $f(x)$ as a Sub-energy Tunneling Function $E_{\text{sub}}(x, x_1^*)$, where x_1^* is a current minimal solution, which has a property that

$$E_{\text{sub}}(x, x_1^*) \simeq \begin{cases} 0, & f(x) \geqslant f(x_1^*), \\ f(x) - f(x_1^*), & f(x) < f(x_1^*). \end{cases}$$

However, to minimize the Sub-energy Tunneling Function this method has to integrate a dynamical system rather than use a local optimization method as a subroutine. The filled function methods [5, 6] modify the objective function as a filled function, and then use a local optimization method to minimize the filled function to find local minimizers of the original objective function with lower function values. The filled functions have an advantage that they have no local minimizers or stationary points in the region $\{x : f(x) \geqslant f(x_1^*)\}$ except prefixed, but they have a disadvantage that they have some parameters which are difficult to adjust.

In this paper, we give a simple transformation method to convert the original objective function $f(x)$ into an auxiliary function with gradually 'fewer' local minimizers. The auxiliary function contains one parameter which is easy to set. Let f_1^* be the best minimal value of the original objective function found up to now. The auxiliary function is 'convex' on the domain $\{x \in X : f(x) \geqslant f_1^*\}$, and also has the advantage that it has no local minimizers or stationary points in the region $\{x \in X : f(x) \geqslant f_1^*\}$ except prefixed, where X is the solution space. Moreover, if f_1^* is smaller, then the region is larger, and at last the auxiliary function is convex if f_1^* is the global minimal value of the original objective function. We use a local optimization method to minimize the auxiliary function to find a better local minimizer of $f(x)$. A stopping criterion of our method is developed based on the Bayesian stopping rules for multistart global optimization methods [10]. An algorithm is designed based on this method, and we prove that it can converge asymptotically with probability one to a global minimizer of the original problem. Numerical experiments on a set of standard test problems with several problems' dimensions up to 50 show that our algorithm is very efficient compared with other global optimization methods.

2. Sequential convexification method

Consider the following global minimization problem

$$(P) \quad \begin{cases} \min & f(x) \\ \text{s.t.} & x \in X, \end{cases}$$

where X is a bounded closed box in R^n, and $f(x)$ is continuously differentiable on X.

Suppose that f_1^* is the current minimal value of problem (P). Before solving problem (P), we can get f_1^* using any local optimization method to minimize $f(x)$

on X, or for simplicity, we just take any point x' in X, and let f_1^* be the value of $f(x')$.

Let $u(t)$ be a continuously differentiable univariate function, $u(0) = 0$, $u'(0) = 0$, and let $u'(t) > 0$, $\forall t > 0$ so that $u(t)$ is a strictly monotonically increasing function of t for $t \geqslant 0$. Let $v(x) \geqslant 0$ be a continuously differentiable convex function on X which has only one minimizer x_v and has no stationary points except x_v on X. Then we construct the following auxiliary function

$$F(x, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})],$$

where $h \geqslant 0$ is a parameter. Obviously, $F(x, f_1^*)$ is a smooth function, and by choosing $v(x)$ and $u(t)$ appropriately, we can make $F(x, f_1^*)$ as smooth as $f(x)$.

Construct the following auxiliary global minimization problem

$$(AP) \quad \begin{cases} \min & F(x, f_1^*) \\ \text{s.t.} & x \in X. \end{cases}$$

Problem (AP) has the following properties.

THEOREM 1. *Suppose that f_1^* is already the global minimal value of problem (P). Then problem (AP) has a unique minimizer x_v.*

Proof. If f_1^* is already the global minimal value of problem (P), then $\forall x \in X$, $f(x) \geqslant f_1^*$, and $\max\{0, f_1^* - f(x)\} = 0$, $F(x, f_1^*) = v(x)$. By the assumption that $v(x)$ has only one minimizer x_v on X, we conclude that problem (AP) has a unique minimizer x_v. □

By the proof of Theorem 1, we know that if f_1^* is the global minimal value of problem (P), then $F(x, f_1^*)$ is a convex function, and problem (AP) is a convex programming problem. Moreover we have the following theorem.

THEOREM 2. *If $f(x_v) > f_1^*$, or x_v is a local minimizer of problem (P) with $f(x_v) \geqslant f_1^*$, then x_v is a local minimizer of problem (AP).*

Proof. If $f(x_v) > f_1^*$, or x_v is a local minimizer of problem (P) with $f(x_v) \geqslant f_1^*$, then there exists a neighbourhood $B(x_v)$ of x_v such that $\forall x \in B(x_v) \cap X$, $f(x) \geqslant f_1^*$. Thus $\forall x \in B(x_v) \cap X$,

$$F(x, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})] = v(x).$$

Since $v(x)$ has only one local minimizer x_v, we have

$$F(x, f_1^*) = v(x) \geqslant v(x_v), \forall x \in B(x_v) \cap X,$$

i.e., x_v is a local minimizer of problem (AP). □

Next we discuss the locations of other local minimizers of problem (AP).

THEOREM 3. *Any one of local minimizers or stationary points of problem* (AP) *except* x_v *cannot be in the region* $S_1 = \{x \in X : f(x) \geqslant f_1^*\}$.

Proof. The derivative of $F(x, f_1^*)$ is

$$\nabla F(x, f_1^*) = \nabla v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})] - v(x) \cdot h \cdot$$
$$u'(\max\{0, f_1^* - f(x)\}) \cdot (- \nabla f(x)).$$

$\forall x \in S_1$, we have $f(x) \geqslant f_1^*$, $\max\{0, f_1^* - f(x)\} = 0$ and $u(\max\{0, f_1^* - f(x)\}) = 0$, $u'(\max\{0, f_1^* - f(x)\}) = 0$. Thus

$$\nabla F(x, f_1^*) = \nabla v(x), \quad \forall x \in S_1.$$

Hence if $x \in S_1$ and $x \neq x_v$, then $\nabla F(x, f_1^*) \neq 0$. Moreover for such x, let $d = x_v - x$, we have $d^T \nabla v(x) < 0$ and $d^T \nabla F(x, f_1^*) = d^T \nabla v(x) < 0$, since $v(x)$ is a convex function with a unique minimizer x_v. So d is a descent direction of $F(x, f_1^*)$ at the point x with $x \in S_1$ and $x \neq x_v$. Hence any one of local minimizers or stationary points except x_v of problem (AP) cannot be in the region $S_1 = \{x \in X : f(x) \geqslant f_1^*\}$. □

Since $S_1 \cup \{x \in X : f(x) < f_1^*\} = X$, Theorem 3 implies the following corollary.

COROLLARY 1. *Any one of local minimizers or stationary points of problem* (AP) *except* x_v *must be in the region* $S_2 = \{x \in X : f(x) < f_1^*\}$.

However, if $h = 0$, then $F(x, f_1^*) = v(x)$, and $F(x, f_1^*)$ has a unique minimizer x_v. Thus if $x_v \in \{x \in X : f(x) \geqslant f_1^*\}$, then by Theorem 3, $F(x, f_1^*)$ has no local minimizers or stationary points in the region S_2. So we have one question that for what h the function $F(x, f_1^*)$ has local minimizers or stationary points in the region $S_2 = \{x \in X : f(x) < f_1^*\}$. In fact, suppose that f^* is the global minimal value of problem (P), we have the following theorem.

THEOREM 4. *Suppose that* f_1^* *is not the global minimal value of problem* (P), *i.e.*, $f_1^* > f^*$, *and suppose that parameter* h *satisfies that* $h > 1/u(f_1^* - f^*)$. *Then all global minimizers of problem* (AP) *are in the region* $S_2 = \{x \in X : f(x) < f_1^*\}$.

Proof. Since $F(x, f_1^*)$ is a continuous function in the closed box X, it has a global minimizer in X. For any $x \in S_1$, we have $f(x) \geqslant f_1^*$, $\max\{0, f_1^* - f(x)\} = 0$ and

$$F(x, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})] = v(x) \geqslant 0. \tag{2.1}$$

Since $f_1^* > f^*$, it is true that $S_2 \neq \emptyset$. By the supposition of Theorem 4, for a global minimizer x^* of problem (P), we have $f(x^*) = f^*$, and

$$1 - h \cdot u(\max\{0, f_1^* - f(x^*)\}) = 1 - h \cdot u(f_1^* - f^*\}) < 0,$$

$$F(x^*, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x^*)\})] < 0.$$

Thus for any $x \in S_2$, we have $f(x) < f_1^*$ and $\max\{0, f_1^* - f(x)\} = f_1^* - f(x) > 0$, and if such x satisfies that $u(\max\{0, f_1^* - f(x)\}) \leqslant \frac{1}{h}$, then

$$F(x, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})] \geqslant 0, \tag{2.2}$$

and if such x satisfies that $u(\max\{0, f_1^* - f(x)\}) > \frac{1}{h}$, then

$$F(x, f_1^*) = v(x)[1 - h \cdot u(\max\{0, f_1^* - f(x)\})] < 0. \tag{2.3}$$

So by (2.1)-(2.3), we know that all global minimizers of problem (AP) are in the region $S_2 = \{x \in X : f(x) < f_1^*\}$. □

Theorem 4 shows that all global minimizers of problem (AP) are in the region $S_2 = \{x \in X : f(x) < f_1^*\}$ if parameter h is large enough. Hence $F(x, f_1^*)$ has local minimizers or stationary points in the region S_2 if h is large enough.

According to the above analysis, if f_1^* is smaller, then region S_1 is larger, and S_2 is smaller. Especially, if f_1^* is already the global minimal value of problem (P), then $S_1 = X$, and $S_2 = \emptyset$. Thus with the decreasing of f_1^*, the number of local minimizers or stationary points of problem (AP) will decrease, and at last if f_1^* is the global minimal value of problem (P), then problem (AP) is a convex programming problem and has only one minimizer x_v.

Moreover if we use any local optimization method to minimize the auxiliary function $F(x, f_1^*)$ on X from any initial point, then by Theorems 3 and 4, it is obvious that the minimization sequence converges either to x_v or to a point $x' \in X$ such that $f(x') < f_1^*$. If we find such x', then using any local optimization method to minimize function $f(x)$ on X from initial point x', we can find a point $x'' \in X$ such that $f(x'') \leqslant f(x')$. This is also the main idea of our algorithm presented in the next section to find a global minimizer of problem (P).

However, generally we do not know the global minimal value of problem (P), so we do not know how large of h would have to be to satisfy the condition of Theorem 4. But for practical consideration, given a desired optimality tolerance $\epsilon > 0$, problem (P) might be considered solved if we can find an $x \in X$ such that $f(x) < f^* + \epsilon$. So we develop a lower bound of h in the following theorem which depends only on the parameter ϵ.

THEOREM 5. *Let ϵ be a sufficiently small positive parameter, and let $h > \frac{1}{u(\epsilon)}$. Then all global minimizers of problem (AP) are in the region $\{x \in X : f(x) < f^* + \epsilon\}$, and if $f_1^* \geqslant f^* + \epsilon$, then $\{x \in X : f(x) < f^* + \epsilon\}$ is contained in $S_2 = \{x \in X : f(x) < f_1^*\}$.*

Proof. Since $h > \frac{1}{u(\epsilon)} = \frac{1}{u((f^*+\epsilon)-f^*)}$, it is direct by Theorem 4 that all global minimizers of problem (AP) are in the region $\{x \in X : f(x) < f^* + \epsilon\}$. Moreover by the assumption that $f_1^* \geqslant f^* + \epsilon$, it is obvious that $\{x \in X : f(x) < f^* + \epsilon\} \subseteq S_2 = \{x \in X : f(x) < f_1^*\}$. □

Thus given a desired optimality tolerance $\epsilon > 0$, for $h > \frac{1}{u(\epsilon)}$, if we find one global minimizer x of problem (AP), then by Theorem 5, $x \in \{x \in X : f(x) < f^* + \epsilon\}$ and it is an approximate global minimal solution of problem (P).

By Theorem 5, the following corollary holds obviously.

COROLLARY 2. *Let f_L^* be the least local minimal value of problem (P) which is larger than the global minimal value f^* of problem (P). If $h > \frac{1}{u(f_L^* - f^*)}$, then all global minimizers of problem (AP) are in the region $\{x \in X : f(x) < f_L^*\}$.*

If a local search used to minimize $F(x, f_1^*)$ and $f(x)$ is strictly descent and can converge to a local minimizer, then during the minimization of $F(x, f_1^*)$, if we find an $x' \in \{x \in X : f(x) < f_L^*\}$, then by Lemma 1 in the next section, starting from x' to minimize $f(x)$ on X will converge to a global minimizer of problem (P). So the parameter ϵ need not be too small. Theoretically, by Corollary of Theorem 5, ϵ should not be greater than $f_L^* - f^*$.

3. The algorithm and its asymptotic convergence

Basing on the theory developed in Section 2, we use multistart local search method to solve problem (AP) to find a better minimal value of problem (P) than the current one f_1^*. The algorithm is described as follows.

ALGORITHM 1.

Step 1. Select randomly a point $x_0 \in X$, set $f_1^* = f(x_0)$, $x_1^* = x_0$. Let h be a sufficiently large positive number, and let N_L be a sufficiently large integer.

Step 2. Choose a convex function $v(x)$ with a unique minimizer x_v, and construct an auxiliary function $F(x, f_1^*)$. Set $N = 0$.

Step 3. If $N \geqslant N_L$, then go to Step 6.

Step 4. Set $N = N + 1$. Draw an initial point from a uniform distribution over X, and start from which to minimize $F(x, f_1^*)$ on X using any local optimization method. Suppose that x' is an obtained local minimizer. If $x' = x_v$, then go to Step 3, otherwise go to Step 5.

Step 5. Minimize $f(x)$ on X with initial point x', and obtain a local minimizer x_2^* of $f(x)$. Let $x_1^* = x_2^*$, $f_1^* = f(x_2^*)$ and go to Step 2.

Step 6. Stop the algorithm, and output x_1^* and f_1^* as an approximate global minimal solution and an approximate global minimal value of problem (P) respectively.

In Step 1 of the above algorithm, parameter h is set large enough. In fact, if one is satisfied with a solution x such that $f(x) < f^* + \epsilon$, where ϵ is a sufficiently small positive number, then by Theorem 5 in Section 2, the value of h should be greater than $\frac{1}{u(\epsilon)}$.

Now we discuss the stopping rule of Algorithm 1. Note that after getting a current minimal value f_1^* of problem (P), Algorithm 1 uses multistart local search method to find a better local minimal value of problem (P) than f_1^*, where parameter N is used to count how many local searches have been performed, and N_L is the maximum number of local searches to be performed to minimize the function $F(x, f_1^*)$ on X. Then how large would N_L have to be to terminate the minimization of $F(x, f_1^*)$ on X? Boender and Rinnooy Kan [10] developed some Bayesian stopping rules for multistart local search method. We use one of them for our algorithm.

Assume that w is the number of different local minimizers of $F(x, f_1^*)$ on X having been discovered, and N is the number of minimizations of $F(x, f_1^*)$ for finding these w local minimizers. Boender and Rinnooy Kan [10] discovered that the Bayesian estimate of the total number of local minimizers of $F(x, f_1^*)$ on X is $\frac{w(N-1)}{N-w-2}$. Hence if the value $\frac{w(N-1)}{N-w-2}$ is very close to w, then one can probabilistically say that $F(x, f_1^*)$ has only w local minimizers on X, which have already been found, and we can terminate the algorithm.

Generally speaking, if the dimension of a global minimization problem is higher, then more local searches must be performed to find a global minimizer. So we think it would be better if a stopping criterion can be related to the dimension of the problem to be solved. Therefore an appealing simple stopping criterion is to terminate the algorithm if

$$\frac{w(N-1)}{N-w-2} \leqslant w + \frac{1}{n}, \tag{3.1}$$

where n is the dimension of the problem to be solved.

Furthermore, for the function $F(x, f_1^*)$ with parameter h large enough, if we can conclude that it has only one local minimizer, then we can conclude that a global minimizer of problem (P) has been found. So by the stopping rule (3.1), for $w = 1$, we have $N \geqslant 2n + 3$. This means that if we can only find the prefixed local minimizer x_v of $F(x, f_1^*)$ after $2n + 3$ minimizations of $F(x, f_1^*)$ with initial points drawn from a uniform distribution over X, then we can conclude probabilistically that $F(x, f_1^*)$ has only one local minimizer on X, and a global minimizer of problem (P) has been found. So we set $N_L = 2n + 3$, and terminate the computation if $N \geqslant N_L = 2n + 3$.

Next we prove the asymptotic convergence of Algorithm 1.

Suppose that the local optimization method used in Algorithm 1 is strictly descent and can converge to a local minimizer of the problem being solved. Let f_L^* be the least local minimal value of problem (P) which is larger than the global minimal value f^* of problem (P), and suppose that the Lebesgue measure of the set $S_L^* = \{x \in X : f(x) < f_L^*\}$ is $m(S_L^*) > 0$.

LEMMA 1. *With an initial point $x' \in S_L^* = \{x \in X : f(x) < f_L^*\}$, the minimization sequence generated by the minimization of $f(x)$ on X will converge to a global minimizer of $f(x)$ on X.*

Proof. Starting from an initial point $x' \in S_L^* = \{x \in X : f(x) < f_L^*\}$, the minimization sequence generated by the minimization of $f(x)$ on X will converge to a local minimizer x_1^* of $f(x)$ on X. Since $f(x') < f_L^*$, and the minimization sequence is strictly descent, it follows that $f(x_1^*) < f_L^*$. By the assumption that f_L^* is the least local minimal value of problem (P) larger than f^*, we have $f(x_1^*) = f^*$, i.e., the minimization sequence converges to a global minimizer of $f(x)$ on X. □

During the k-th iteration of Algorithm 1, let x_k be the k-th random point drawn from a uniform distribution over X at Step 4, and let y_k be the best local minimal value found so far, i.e., $y_0 = f(x_0)$, and $y_k = y_{k-1}$ if starting from the initial point x_k the minimization sequence generated by the minimization of $F(x, f_1^*)$ on X converges to the prefixed point x_v, otherwise y_k is a new value less than y_{k-1}.

THEOREM 6. *If $h > \frac{1}{u(f_L^* - f^*)}$ and $N_L = +\infty$, then y_k converges to the global minimal value f^* of problem (P) with probability 1, i.e., $P\{\lim_{k\to\infty} y_k = f^*\} = 1$.*

Proof. To prove Theorem 6 is equivalent to prove that

$$P\{\bigcap_{k=1}^{\infty}\bigcup_{l=k}^{\infty}(|y_l - f^*| \geqslant \epsilon)\} = 0, \forall\epsilon > 0. \tag{3.2}$$

If $h > \frac{1}{u(f_L^* - f^*)}$, then by Theorems 3 and 5 of section 2, it is obvious that $f^* \leqslant y_k \leqslant y_{k-1}$, i.e., $\{y_k\}_{k=0}^{\infty}$ is monotonically decreasing.

Let $q = 1 - \frac{m(S_L^*)}{m(X)}$. By Lemma 1 and the monotonicity of $\{y_k\}_{k=0}^{\infty}$, $\forall\epsilon > 0$, we have

$$\begin{aligned} P\{y_l - f^* \geqslant \epsilon\} &= P\{\bigcap_{i=1}^{l}(y_i - f^* \geqslant \epsilon)\} \\ &\leqslant P\{\bigcap_{i=1}^{l}(x_i \notin S_L^*)\} = \prod_{i=1}^{l} P\{x_i \notin S_L^*\} \\ &= \prod_{i=1}^{l}(1 - \frac{m(S_L^*)}{m(X)}) = q^l. \end{aligned}$$

Thus

$$\begin{aligned} P\{\bigcap_{k=1}^{\infty}\bigcup_{l=k}^{\infty}(y_l - f^* \geqslant \epsilon)\} &\leqslant \lim_{k\to\infty} P\{\bigcup_{l=k}^{\infty}(y_l - f^* \geqslant \epsilon)\} \\ &\leqslant \lim_{k\to\infty}\sum_{l=k}^{\infty} P\{y_l - f^* \geqslant \epsilon\} \leqslant \lim_{k\to\infty}\sum_{l=k}^{\infty} q^l = \lim_{k\to\infty}\frac{q^k}{1-q}. \end{aligned}$$

Since $m(S_L^*) > 0$, we have $0 \leqslant q < 1$, and $\frac{q^k}{1-q} \to 0$ if $k \to \infty$. Hence (3.2) holds. □

Hence by Theorem 6, the current local minimizer x_1^* of problem (P) generated by Algorithm 1 converges with probability one to a global minimizer of problem (P).

4. Test results and comparison with other methods

In this section, we test our algorithm on a set of standard global minimization problems and compare the performance of our algorithm with some well known methods for global minimization problems.

We choose the unique minimizer x_v of $v(x)$ as $x_v = x_1^*$, and choose $v(x) = (x - x_v)^T(x - x_v)$, $u(t) = t^2$. We set $N_L = 2n + 3$, where n is the dimension of the global minimization problem to be solved. We take h to be a sufficiently large positive number, say $h = 10^6$.

In practical implementation of Algorithm 1, the initial point in Step 4 of Algorithm 1 is drawn from a uniform distribution over the boundary of X instead of over X. Since by Theorems 3 and 4 in Section 2, from any initial point the minimization sequence generated by the minimization of $F(x, f_1^*)$ on X converges either to x_v or to a point $x' \in X$ such that $f(x') < f_1^*$, to avoid converging to x_v too frequently, we should start the minimization of $F(x, f_1^*)$ from points at the boundary of X.

In Steps 4 and 5 of Algorithm 1, we prefer using an inexact line search method in the local optimization method, since in this way we can explore the solution space with more trial points. Hence in practical implementation of Algorithm 1, we use BFGS local optimizer with an inexact line search method to minimize $F(x, f_1^*)$ and $f(x)$ in Steps 4 and 5 of Algorithm 1.

The above implementation details can be summarized in an algorithm as follows.

ALGORITHM 2.
Step 1. Select randomly a point $x_0 \in X$, set $f_1^* = f(x_0)$, $x_1^* = x_0$, $h = 10^6$. Let $N_L = 2n + 3$, where n is the dimension of the global minimization problem to be solved.

Step 2. Let $x_v = x_1^*$, $v(x) = (x - x_v)^T(x - x_v)$, $u(t) = t^2$, and construct an auxiliary function $F(x, f_1^*)$. Set $N = 0$.

Step 3. If $N \geqslant N_L$, then go to Step 6.

Step 4. Set $N = N + 1$. Draw an initial point from a uniform distribution over the boundary of X, and start from which to minimize $F(x, f_1^*)$ on X using BFGS local optimizer with an inexact line search method. During the minimization of $F(x, f_1^*)$, if there exists a point x_k with $f(x_k) < f_1^*$, then stop minimizing $F(x, f_1^*)$ and go to Step 5, otherwise go to Step 3.

Step 5. Minimize $f(x)$ on X with initial point x_k, and obtain a local minimizer x_2^* of $f(x)$. Let $x_1^* = x_2^*$, $f_1^* = f(x_2^*)$ and go to Step 2.

Step 6. Stop the algorithm, and output x_1^* and f_1^* as an approximate global minimal solution and an approximate global minimal value of problem (P), respectively.

We test the above algorithm on a set of standard global minimization problems on a personal computer with Intel CPU 586/200 and internal memory 128M.

The algorithm finds global minimal solutions of these problems. We compare the performance of our algorithm with TRUST [1], the Diffusion Equation method [7], the Filled Function method [6], the Multilevel Single Linkage method [11] and Levy's Tunneling Algorithm [9]. The comparison criterion is the number of function evaluations. The computational results of these methods are taken from the papers cited.

For our algorithm, the number of function evaluations includes the number of function evaluations of the objective function $f(x)$ and that of the auxiliary function $F(x, f_1^*)$. Also, the number of function evaluations of the Filled Function method recorded in the tables of this section includes the number of function evaluations of the objective function $f(x)$ and that of the filled function.

In the tables of this section, the symbol '–' means the number of function evaluations is too large to record.

PROBLEM 1. *Branin's function*

$$f(x) = (x_2 - \frac{5.1x_1^2}{4\pi^2} + \frac{5x_1}{\pi} - 6)^2 + 10(1 - \frac{1}{8\pi}) \cos x_1 + 10$$

has many local minimizers in the domain $-5 \leqslant x_1 \leqslant 10$, $0 \leqslant x_2 \leqslant 15$, *but the global minima are* $x^* = (3.141593, 2.275000)^T$, $x^* = (-3.141593, 12.275000)^T$. *The global minimal value is* $f^* = 0.397667$.

PROBLEM 2. *The three hump camel function*

$$f(x) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} - x_1x_2 + x_2^2$$

has three local minimizers in the domain $-3 \leqslant x_i \leqslant 3$, $i = 1, 2$, *and the global minimizer is* $x^* = (0, 0)^T$. *The global minimal value is* $f^* = 0$.

PROBLEM 3. *The Treccani function*

$$f(x) = x_1^4 + 4x_1^3 + 4x_1^2 + x_2^2$$

has two local minimizers $x^* = (-2, 0)^T$ *and* $x^* = (0, 0)^T$ *in the domain* $-3 \leqslant x_1 \leqslant 3$, $i = 1, 2$. *The global minimal value is* $f^* = 0$.

PROBLEM 4. *The six hump camel function*

$$f(x) = 4x_1^2 - 2.1x_1^4 + \frac{x_1^6}{3} + x_1x_2 - 4x_2^2 + 4x_2^4$$

has six local minimizers in the domain $-3 \leqslant x_1 \leqslant 3$, $-1.5 \leqslant x_2 \leqslant 1.5$, *and two of them are global minimizers:* $x^* = (-0.089842, 0.712656)^T$, $x^* = (0.089842, -0.712656)^T$. *The global minimal value is* $f^* = -1.031628$.

PROBLEM 5. *The two dimensional Shubert function*

$$f(x) = \{\sum_{i=1}^{5} i\cos[(i+1)x_1 + i]\}\{\sum_{i=1}^{5} i\cos[(i+1)x_2 + i]\}$$

has 760 local minimizers in the domain $-10 \leqslant x_i \leqslant 10$, $i = 1, 2$, *and 18 of them are global minimizers. The global minimal value is* $f^* = -186.730909$.

PROBLEM 6. *The two-dimensional Shubert function*

$$f(x) = \{\sum_{i=1}^{5} i\cos[(i+1)x_1 + i]\}\{\sum_{i=1}^{5} i\cos[(i+1)x_2 + i]\} + \frac{1}{2}[(x_1 + 0.80032)^2 + (x_2 + 1.42513)^2]$$

has roughly the same behavior as the function presented in Problem 5 in the domain $-10 \leqslant x_i \leqslant 10$, $i = 1, 2$, *but has a unique global minimizer* $x^* = (-0.80032, -1.42513)^T$. *The global minimal value is* $f^* = -186.730909$.

PROBLEM 7. *The two-dimensional Shubert function*

$$f(x) = \{\sum_{i=1}^{5} i\cos[(i+1)x_1 + i]\}\{\sum_{i=1}^{5} i\cos[(i+1)x_2 + i]\} + [(x_1 + 0.80032)^2 + (x_2 + 1.42513)^2]$$

in the domain $-10 \leqslant x_i \leqslant 10$, $i = 1, 2$ *has roughly the same behavior and the same global minimizer and global minimal value as the function presented in Problem 6, but with steeper slope around global minimizer.*

PROBLEM 8. *Shekel's function*

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

has m *local minimizers in the domain* $0 \leqslant x_i \leqslant 10$, $i = 1, 2, 3, 4$, *but only one global minimizer, where* m=5, *or* 7, *or* 10. *The parameters are presented in the following table.*

Table 1. Comparison of sequentianl convexification method(SCM) and other algorithms on number of function evaluations

No.	n	SCM	TRUST	Diffusion equation	Filled function	MSL	Tunneling
1	2	57	60		189	206	
2	2	46			429		
3	2	61			313		
4	2	55	168	120	184		1496
5	2	103	256		290		12160
6	2	166			234		2912
7	2	232			439		2180
8(m=5)	4	156		12000	390	404	
8(m=7)	4	159		12000	410	not found	
8(m=10)	4	not found		12000	not found	564	
9	2	98		120	148	148	

i	a_i				c_i
1	*4*	*4*	*4*	*4*	*0.1*
2	*1*	*1*	*1*	*1*	*0.2*
3	*8*	*8*	*8*	*8*	*0.2*
4	*6*	*6*	*6*	*6*	*0.4*
5	*3*	*7*	*3*	*7*	*0.4*
6	*2*	*9*	*2*	*9*	*0.6*
7	*5*	*5*	*3*	*3*	*0.3*
8	*8*	*1*	*8*	*1*	*0.7*
9	*6*	*2*	*6*	*2*	*0.5*
10	*7*	*3.6*	*7*	*3.6*	*0.5*

PROBLEM 9. *The Goldstein–Price function*

$$f(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$$
$$\times[30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$

has four local minimizers in the domain $-2 \leqslant x_i \leqslant 2$, $i = 1, 2, 3, 4$, *but only one global minimizer* $x^* = (0, -1)^T$ *with the global minimal value* $f^* = 3$.

PROBLEM 10. *The function*

$$f(x) = \frac{1}{10}\{\sin^2(3\pi x_1) + \sum_{i=1}^{n-1}(x_i - 1)^2[1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2[1 + \sin^2(2\pi x_n)]\}$$

Table 2. Comparison of sequentianl convexification method(SCM) and other algorithms on number of function evaluations

No.	n	SCM	Filled function	Tunneling
10	2	143	252	2653
	3	331	339	6955
	4	674	1012	3865
	5	1312	938	10715
	6	824	2262	12786
	7	675	2951	16063
	8	1058	3634	
	9	2736	3623	
	10	1567	2969	
	11	4396	–	
	12	1821	–	
	13	1666	–	
	14	21830	–	
	15	2127	1555	
	16	2987	–	
	17	38603	–	
	18	3752	–	
	19	4509	–	
	20	3593	–	
	21	3485	4668	
	22	6703	–	
	23	23058	–	
	24	2558	–	
	25	23141	2361	
	30	13732		
	40	10975		
	50	11736		

has roughly 30^n *local minimizers in the domain* $-10 \leqslant x_i \leqslant 10$, $i = 1, \cdots, n$, *but only one global minimizer* $x^* = (1, 1, \cdots, 1)^T$ *with the global minimal value* $f^* = 0$.

PROBLEM 11. *The function*

$$f(x) = \frac{1}{2}\sum_{i=1}^{n}(x_i^4 - 16x_i^2 + 5x_i)$$

Table 3. Comparison of sequentianl convexification method(SCM) and TRUST on number of function evaluations

No.	n	SCM	TRUST
11	1	23	38
	2	26	22
	3	27	21
	4	29	21
	5	27	
	6	32	
	7	117	
	8	105	
	9	74	
	10	126	
	11	83	
	12	132	
	13	303	
	14	581	
	15	469	
	16	795	
	17	250	
	18	2234	
	19	1297	
	20	1306	
	21	566	
	22	1370	
	23	1619	
	24	883	
	25	1540	
	30	1928	
	40	2073	
	50	1164	

has 2^n *local minimizers and only one global minimizer* $x^* = (-2.90354, -2.90354, \cdots, -2.90354)^T$.

The test results of the above problems are presented in Tables 1–3. In Table 1, our algorithm did not find a global minimal solution of Problem 8 with $m = 10$, but after more iterations with 4617 function evaluations, the method found a global minimal solution of the problem.

5. Conclusions

This paper has introduced SCM, a new effective and simple method for continuous global minimization problems. The method is based on a transformation of the objective function $f(x)$ into an auxiliary function $F(x, f_1^*)$, where f_1^* is the current minimal value of the objective function. All local minimizers or stationary points except a prefixed one of the auxiliary function are in the region $\{x \in X : f(x) < f(x_1^*)\}$. Minimizing $F(x, f_1^*)$ with any local optimizer, the minimization sequence will converge either to the prefixed point or to a point x' with $f(x') < f_1^*$. Benchmark comparisons with other global optimization procedures have demonstrated that our method is very efficient, as measured by the number of function evaluations. Further research should be to construct new auxiliary functions with the same properties of $F(x, f_1^*)$, and find the best one among them.

6. Acknowledgement

The authors are indebted to anonymous referees whose suggestions and comments are very helpful for improving the quality of the manuscript.

Note

Research supported by NKBRSF of China, Grant G1998030600, the Natural Science Foundation of Fujian Province, Grant A0010010 and the Foundation of Education Committee of Fujian Province, Grant JA00143.

References

1. Cetin, B.C., Barhne, J. and Burdick, J.W. (1993), Terminal Repeller unconstrained subenergy tunneling (TRUST) for fast global optimization, *Journal of Optimization Theory and Applications* 77(1), 97–126.
2. Coleman, T., Shalloway, D. and Wu, Z. (1992), Isotropic effective energy simulated annealing searches for low energy molecular cluster states, Technical Report CTC-92-TR113, Center for Theory and Simulation in Science and Engineering, Cornell University.
3. Coleman, T., Shalloway, D. and Wu, Z. (1993), A parallel build-up algorithm for global energy minimizations of molecular clusters using effective energy simulated annealing, Technical Report CTC-93-TR130, Center for Theory and Simulation in Science and Engineering, Cornell University.
4. Coleman, T. and Wu, Z. (1994), Parallel continuation-based global optimization for molecular conformation and protein folding, Technical Report CTC-94-TR175, Center for Theory and Simulation in Science and Engineering, Cornell University.
5. Ge, R.P. (1990), A filled function method for finding a global minimizer of a function of several variables, *Mathematical Programming* 46, 191–204.
6. Ge, R.P. and Qin Y.F. (1990), The globally convexized filled functions for global optimization, *Applied Mathematics and Computation* 33, 131–158.
7. Kostrowicki, J. and Piela, L. (1991), Diffusion equation method of global minimization: performance for standard test functions, *Journal of Optimization Theory and Applications* 69(2), 97–126.

8. Kostrowicki, J., Piela, L., Cherayil B.J. and Scheraga A. (1991), A performance of the diffusion equation method in searches for optimum structures of clusters of Lennard-Jones atoms, *Journal of Physical Chemistry* 95, 4113–4119.
9. Levy, A.V. and Montalvo, A. (1985), The tunneling algorithm for the global minimization of functions, *SIAM Journal on Scientific and Statistical Computing* 6, 15–29.
10. Rinnooy Kan, A.H.G. and Timmer, G.T. (1987), Stochastic global optimization methods, Part I: Clustering methods, *Mathematical Programming* 39, 27–56.
11. Rinnooy Kan, A.H.G. and Timmer, G.T. (1987), Stochastic global optimization methods, Part II: Multi-level methods, *Mathematical Programming* 39, 57–78.
12. Shalloway, D. (1991), Packet annealing: a deterministic method for global minimization with applications to molecular conformation, Preprint, Section of Biochemistry, Molecular and Cell Biology, *Global Optimization*, C. Floudas and P. Pardalos (eds.), Princeton University Press, Princeton, NJ.
13. Wu, Z. (1993), The effective energy transformation scheme as a special continuation approach to global optimization with application to molecular conformation, Technical Report CTC-93-TR143, Center for Theory and Simulation in Science and Engineering, Cornell University, 1993.