J Glob Optim (2008) 42:619–623
DOI 10.1007/s10898-008-9281-2

A note on global optimization via the heat diffusion equation

Vladimir Shikhman

Received: 10 July 2007 / Accepted: 19 January 2008 / Published online: 19 February 2008

Abstract We consider the interesting smoothing method of global optimization recently proposed in Lau and Kwong (J Glob Optim 34:369–398, 2006) . In this method smoothed functions are solutions of an initial-value problem for a heat diffusion equation with external heat source. As shown in Lau and Kwong (J Glob Optim 34:369–398, 2006), the source helps to control global minima of the smoothed functions—they are not shifted during the smoothing. In this note we point out that for certain (families of) objective functions the proposed method unfortunately does not affect the functions, in the sense, that the smoothed functions coincide with the respective objective function. The key point here is that the Laplacian might be too weak in order to smooth out critical points.

Keywords Global optimization · Heat diffusion equation · Smoothing method

1 Introduction

We consider the following unconstrained global optimization problem:

$$\text{minimize } f : \mathbb{R}^n \longrightarrow \mathbb{R} \tag{1.1}$$

where f is a twice continuously differentiable function.

Among various approaches trying to solve (1.1) there is the interesting one of smoothing. The idea of a smoothing method is to evolve the objective function f into a new (smoothed) function possessing much fewer local minima, in the best case only one. Afterwards one searches for a global minimum of the smoothed function using local minimization technique. However, a global minimum of the smoothed function determined in a such way can happen to differ from a global minimum of the objective function. It can be caused by shifting effects while smoothing.

V. Shikhman (✉)
Department of Mathematics – C, RWTH Aachen University, 52056 Aachen, Germany
e-mail: shikhman@mathc.rwth-aachen.de

So it is necessary to employ a kind of reversing procedure to get back to a global minimum of the original objective function.

As one can imagine, two properties of a such smoothing method are crucial:

($\star$) the number of local minima of the smoothed function is significantly smaller than that of the objective function,

($\star\star$) the information about a global minimum of the smoothed function must lead to a global minimum of the objective function.

Among different smoothing methods proposed in the literature we mention: diffusion equation methods [1,2,6], the effective energy method [4] and the effective energy transformation scheme [7]. Most of them stem from molecular chemistry applications and deal in particular with the Lennard-Jones problem (see [5] for further details and more references). Here we want to discuss a smoothing method recently suggested by Lau and Kwong [3]. The smoothed function for the problem (1.1) is defined in their method as $u(t,x) : [0,T)\times\mathbb{R}^n \to \mathbb{R}$, $T > 0$, satisfying the following initial-value problem:

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \min\{0, \Delta u(t,x)\} & (t,x) \in (0,T)\times\mathbb{R}^n, \\ u(0,x) = f(x) & x \in \mathbb{R}^n, \end{cases} \tag{1.2}$$

where Δ denotes the Laplacian in x-coordinates.

Due to the equality $\min\{0, \Delta u(t,x)\} = \Delta u(t,x) - \max\{0, \Delta u(t,x)\}$, the governing equation is a heat diffusion equation with the external heat source $-\max\{0, \Delta u(t,x)\}$. Two main results for this initial-value problem were obtained in [3], assuming some growth conditions on f:

(i) the initial-value problem (1.2) has the unique viscosity solution $u(t,x)$,
(ii) a global minimum of $u(t,\cdot)$ coincides with that of f for all $t \in (0,T)$.

Obviously the above result (ii) corresponds to the property ($\star\star$) of a smoothing method: a reversing procedure is not to be carried out, because global minima of f are not shifted during smoothing at all.

Our aim here is to discuss the property ($\star$) of the underlying smoothing method from [3]. The Laplacian on the right-hand side of (1.2) gives rise to suppose, that the local structure of a critical point $\bar{x}$ of f (i.e. with $Df(\bar{x}) = 0$) will be changed during the smoothing process. However, the Laplacian only describes the sum of the eigenvalues of the Hessian $D^2 f(\bar{x})$, but not their signs. In this sense the Laplacian might be too weak in order to smooth.

From this intuitive observation we get an idea to construct an objective function f with the non-negative Laplacian on $\mathbb{R}^n$ ($\Delta f(x) \geq 0$). From the fact $\min\{0, \Delta f(x)\} = 0$ for such a function we obtain: $u(t,x) := f(x)$ is a solution for the initial-value problem (1.2), so the smoothing method based on it does not change the objective function and generates the same smoothed function as f in this case. If the objective function f has an additional property of possessing several local minima, we can see that the underlying method would not make any progress in smoothing out these local minima.

Moreover, as the main result we construct a quite wide family of such objective functions, each of them having a non-trivial set of local minima and saddle points and being the smoothed function with respect to (1.2) itself. It shows, that the smoothing method from [3] can not be applied successfully at least for such families of objective functions.

2 A special family of objective functions

We start with a two-dimensional example communicated by Wagner (pers. commun.).

Example 2.1 Consider $f(x, y) = (x^2 - 1)^2 + ay^2$, where $a \in \mathbb{R}$. It holds:

$$Df(x, y) = (4(x^2 - 1)x, 2ay); \quad D^2 f(x, y) = \begin{pmatrix} 12x^2 - 4 & 0 \\ 0 & 2a \end{pmatrix}.$$

For $a > 2$ we obtain $\Delta f(x, y) = 12x^2 - 4 + 2a > 0$. The set of local minima for f is $\{(1, 0), (-1, 0)\}$ and the set of critical points with quadratic index 1 is the singleton $\{(0, 0)\}$.

Next we generalize the idea of Example 2.1.

Theorem 2.2 *Let $n \geq 2$ and $g_i : \mathbb{R} \to \mathbb{R}$ $i = 1, \ldots, n-1$ be twice continuously differentiable functions with second derivatives bounded from below on $\mathbb{R}$. Let $A_i^{\min}$ and $A_i^{\max}$ denote the set of non-degenerate local minima and local maxima of g_i respectively.*

Then there exists a real number $a > 0$, such that the function $f : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x_1, \ldots, x_n) := \sum_{i=1}^{n-1} g_i(x_i) + ax_n^2 \tag{2.1}$$

has following properties:

(i) $\Delta f(x_1, \ldots, x_n) \geq 0$ *for all* $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

(ii) *the number of non-degenerate critical points of f with quadratic index* $0 \leq k \leq n-1$ *is equal to*

$$\sum_{\substack{P \subset \{1, \ldots, n-1\} \\ |P| = k}} \prod_{i \in P} |A_i^{\max}| \prod_{j \in \{1, \ldots, n-1\} \setminus P} |A_j^{\min}|. \tag{2.2}$$

Here $|\cdot|$ denotes the cardinality of a set (∞ is allowed). To avoid misleadings by multiplication in formula (2.2) we assume, as usual, for $c \geq 0$:

$$c \cdot \infty = \begin{cases} 0, & c = 0, \\ \infty, & else. \end{cases}$$

Proof We obtain for f:

$$Df(x_1, \ldots, x_n) = (g_1'(x_1), \ldots, g_{n-1}'(x_{n-1}), 2ax_n),$$

$$D^2 f(x_1, \ldots, x_n) = diag(g_1''(x_1), \ldots, g_{n-1}''(x_{n-1}), 2a),$$

$$\Delta f(x_1, \ldots, x_n) = \sum_{i=1}^{n-1} g_i''(x_i) + 2a.$$

The last equality and the boundness from below of g_i'' allow us to find $a > 0$ sufficiently large, so that (i) holds.

Moreover, the set of non-degenerate critical points of f is $A_1 \times \cdots \times A_{n-1} \times \{0\}$, where $A_i := A_i^{\min} \cup A_i^{\max}$. This observation and the formula for $D^2 f$ prove (ii). □

Example 2.3 For particular g_i from the Theorem 2.2 we can take

(1) each polynomial of an even degree with a positive leading coefficient (finite number of local minima),

(2) sine and cosine functions (infinite number of local minima),
(3) a smooth function defined for $x_0 \in \mathbb{R}$ and $r > 0$ as following:

$$\nu_{x_0,r}(x) = \begin{cases} \exp\left(\frac{1}{\left(\frac{x-x_0}{r}\right)^2 - 1}\right), & |x - x_0| < r, \\ 0, & |x - x_0| \geq r. \end{cases}$$

Adding such functions $\nu_{x_0,r}$ to the already constructed g_i in the small enough neighborhoods of its critical points x_0 we can additionally achieve different critical values for the objective function f.

3 Final remarks

Due to the referee's observation, the method in [3] may still be able to handle f from (2.1) indirectly. Each global minimum of f over $\mathbb{R}^n$ has the form $(x_1^*, x_2^*, \ldots, x_{n-1}^*, 0)$, where $(x_1^*, x_2^*, \ldots, x_{n-1}^*)$ is a global minimum of $\sum_{i=1}^{n-1} g_i$ over $\mathbb{R}^{n-1}$. So, a global minimum of f can be obtained by minimizing $\sum_{i=1}^{n-1} g_i$.

However, one may slightly modify the proposed function f from Theorem 2.2:

$$\widetilde{f}(x_1, \ldots, x_n) := \sum_{i=1}^{n-2l} g_i(x_i) + a \sum_{j=1}^{l} h_j(x_{2j-1+n-2l}, x_{2j+n-2l}). \quad (3.1)$$

Here, $0 \leq l \leq \left\lfloor \frac{n}{2} \right\rfloor$, $a > 0$ and g_i are given as in Theorem 2.2. Moreover, $h_j(x, y) := u_j(x, y) + v_j(x) + w_j(y)$, where $u_j : \mathbb{R}^2 \to \mathbb{R}$ are harmonic functions and $v_j, w_j : \mathbb{R} \to \mathbb{R}$ are twice continuously differentiable functions with second derivatives bounded from below on $\mathbb{R}$ by positive constants.

For $a > 0$ sufficiently large, we obtain $\Delta \widetilde{f} \geq 0$. It means: the smoothed function with respect to (1.2) is $\widetilde{f}$ itself. In this case, if possible, an indirect treating would be more sophisticated.

For further research it would be very interesting to find out, for what (classes of) functions arising in applications, the method in [3] does make progress in smoothing out critical points.

Acknowledgements The author would like to thank Prof. H. Th. Jongen for drawing his attention to Ref. [3]. The author would also like to thank the anonymous referee for his substantial remarks.

References

1. Kostrowicki, J., Scheraga, H.: Some approaches to the multi-minima problem in protein folding. In: Pardalos, P.M., Shalloway, D., Xue, G. (eds.) Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding, vol. 95, pp. 123–132. AMS (1996)
2. Kostrowicki, J., Piela, L., Cherayil, B.J., Scheraga, H.: Performance of the diffusion equation method in searches for optimum structures of clusters of Lennard-Jones atoms. J. Phys. Chem. **95**, 4113–4119 (1991)
3. Lau, M.S.K., Kwong, C.P.: A smoothing method of global optimization that preserves global minima. J. Glob. Optim. **34**, 369–398 (2006)
4. Shalloway, D.: Packet annealing: a deterministic method for global minimization, with applications, to molecular conformation. In: Pardalos, P.M., Floudas, C. (eds.) Global Optimization. Princeton University Press (1991)
5. Shao, C.S., Byrd, R., Eskow, E., Schnabel, R.B.: Global optimization for molecular clusters using a new smoothing approach. J. Glob. Optim. **16**, 167–196 (2000)

6. Vese, L.: A method to convexify functions via curve evolution. Commun. Part. Diff. Eq. **24**, 1573–1591 (1999)
7. Wu, Z.: The effective energy transformation scheme as a special continuation approach to global optimization with application to molecular conformation. SIAM J. Optim. **6**, 748–768 (1996)