ON A NEW SMOOTHING TECHNIQUE FOR NON-SMOOTH, NON-CONVEX OPTIMIZATION

NURULLAH YILMAZ AND AHMET SAHINER∗
Department of Mathematics
Suleyman Demirel University, Isparta, Turkey

(Communicated by Gerhard-Wilhelm Weber)

Abstract. In many global optimization techniques, the local search methods are used for different issues such as to obtain a new initial point and to find the local solution rapidly. Most of these local search methods base on the smoothness of the problem. In this study, we propose a new smoothing approach in order to smooth out non-smooth and non-Lipschitz functions playing a very important role in global optimization problems. We illustrate our smoothing approach on well-known test problems in the literature. The numerical results show the efficiency of our method.

1. Introduction. We consider the following unconstrained optimization problem

\[(P) \quad \min_{x \in \mathbb{R}^n} f(x)\]

where the objective function \(f : \mathbb{R}^n \to \mathbb{R}\) is continuous and differentiable almost everywhere. The problem \((P)\) has drawn important attention in many branches of engineering, finance, and other sciences [31, 33]. However, the problem is very hard to solve because of the non-smooth and non-convex structure of the objective function [18, 47].

When the problem is non-smooth (perhaps non-Lipschitz), well developed important techniques for smooth optimization can be insufficient. In order to make these optimization techniques available the sub-differential and smoothing approaches are proposed. The sub-differential approach is related to generalization of the differentiability. In fact, with the help of the generalizations on differentiability, the conditions for differentiability are weakened and the classical techniques based on classical differentiability are upgraded [32, 7, 8, 9]. These upgraded techniques are used effectively for Lipschitz continuous problems but most of these techniques are unavailable for non-Lipschitz problems.

The smoothing studies deal with making the objective function continuously differentiable before the optimization procedure. The smoothing focuses on finding the appropriate modification which makes the objective function smooth and on designing approximation to the objective function by smooth functions. One of the

2010 Mathematics Subject Classification. Primary: 65D10, 90C26; Secondary: 49J52, 26A27.

Key words and phrases. Smoothing techniques, non-smooth analysis, non-Lipschitz problems, global optimization.

This study has been supported by the Teaching Staff Training Project Units of Suleyman Demirel University (OYP-05545-DR-13).

∗ Corresponding author: Ahmet Sahiner.
first studies on the smoothing approach was proposed in 1975 by Bertsekas [2], in 1980 by Zang [51] and in 1982 by Xavier [44, 46]. The smoothing approach has been used for many non-smooth but Lipschitz continuous problems such as min-max [1, 50], penalty expressions of constrained optimization problems [20], clustering problems [45], nonlinear complementarity problems [4]. Moreover, the smoothing approach is used for both Lipschitz and non-Lipschitz regularization problems in [5, 6] and for other problems [25, 26]. Therefore, the smoothing techniques are very important in terms of optimization. The smoothing function is described by the following definition:

**Definition 1.1.** [5] Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. The function $	ilde{f} : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is called a smoothing function of $f(x)$, if $\tilde{f}(\cdot, \beta)$ is continuously differentiable in $\mathbb{R}^n$ for any fixed $\beta$, and for any $x \in \mathbb{R}^n,$

$$\lim_{z \to x, \beta \to 0} \tilde{f}(z, \beta) = f(x).$$

If the function $f$ in problem (P) is non-convex then, it may have a lot of local minimizers. When the function $f$ has multiple local minimizers, determining the global one and escaping from the current local minimizer in order to get the lower one are important challenges. Although there are many valuable methods and algorithms which are dedicated to solve the global optimization problems, there is no exact method which efficiently solves all these different types of problems. For example; the global optimization methods based on gradient usage are not useful for non-smooth problems and the methods based on Lipschitz continuity are not useful for non-Lipschitz problems.

We assume the following assumptions hold for the problem (P):

**Assumption 1** The function $f$ is coercive, i.e. $f(x) \to \infty$ as $\|x\| \to \infty$.

As a result of Assumption 1, there exists a closed and bounded box $\Omega$ containing all the minimizers of $f$.

**Assumption 2** The function $f$ has a finite number of local minimizers.

Throughout the paper, we use $x^*_k$ to denote the $k$–th local minimizer of $f$ whereas by $x^*$ we mean the global minimizer.

Most of the global optimization methods concentrate on finding lower minimizer than the current one. These methods can be classified into three groups; heuristic, stochastic and deterministic. Each of these approaches can outmaneuver each other from different aspects. Heuristic approaches generally base on the simulation of physical, chemical or biological processes. The prominent heuristic approaches are Simulated Annealing, Genetic Algorithms, Particle Swarm Optimization and Fuzzy Optimization [15, 12, 37, 30]. Stochastic approaches are quite simple, very efficient in black box problems and robust with respect to the increment of dimension of the problem [53]. Although these methods have many important advantages, the convergence to solution is quite slow and they can give different results if they are run again [36, 16]. Deterministic approaches always find the same solution if they run again under the same conditions. They also have some advantages over the others in terms of easy implementations [38, 35]. There exist important methods such as Branch and Bound [54, 27], Cutting Plane [39] and space filling curves based methods [17] and other important methods [13, 28, 29].

One of the important deterministic approach is the auxiliary function approach which includes the Tunneling Method (Algorithm) [19, 3, 48], Filled Function Method [10, 11, 52, 49, 42, 22, 21], Global Descent Method [43, 23] and Cut-Peak Function Method [40]. These methods are established on finding a lower minimizer...
than the current one by making a suitable modification to the objective function. The modified function is generally called as auxiliary function (Filled Function, Tunneling Function and etc.). The auxiliary function contains parameters which control the modifications. In most auxiliary function methods, setting the parameters constitutes the important phase of their algorithms. Since the parameters may change from problem to problem, there is always a risk of undesirable situations in computations.

In this study, we propose a new smoothing approach for a wide class of non-smooth functions. In this new technique, first the smoothing approach is applied to a non-smooth (or non-Lipschitz) functions so that well-known classical optimization methods can be applied.

In the following section, we present the theoretical infrastructure of the approach analytically. In Section 3, we give the convergence results of the smoothing approach. In Section 4, the application of smoothing approach is presented and the numerical results on test problems are reported. In the last section, we give some concluding remarks.

2. A New Smoothing Technique. The non-smoothness of the problem \((P)\) mostly originates from the presence of “max”, “min”, “\(|·|\)” in the formulation of \(f(x)\). If \(p = 1\), then the problem \((P)\) becomes non-smooth but Lipschitz continuous. If \(0 < p < 1\), then the problem \((P)\) becomes non-Lipschitz. In order to obtain the smooth objective function in which the local minimizer finding algorithms can easily be applicable, we replace non-smooth functions including the above terms with their smoothed versions. For this purpose, if 
\[
\tilde{q}_p(t, \beta) = \begin{cases} 
q_p(t), & |t| > \beta, \\
\mu_p(t, \beta), & |t| \leq \beta,
\end{cases}
\]

where \(\mu_p(t, \beta)\) is continuously differentiable with respect to \(t\) and it is expressed by the following formula:
\[
\mu_p(t, \beta) = \beta^{p-3} \left( \frac{p-1}{4} \right) t^3 + \beta^{p-2} \left( \frac{p}{4} \right) t^2 + \beta^{p-1} \left( \frac{3-p}{4} \right) t + \beta^p \left( \frac{2-p}{4} \right),
\]

where \(\beta > 0\). If the second order differentiability of the function \(\tilde{q}_p(t, \beta)\) is needed then, the following function:
\[
\mu_p(t, \beta) = \frac{(p-3)(p-1)}{16} \beta^{p-5} t^5 + \frac{p(p-2)}{16} \beta^{p-4} t^4 - \frac{(p-5)(p-1)}{8} \beta^{p-3} t^3 - \frac{p(p-4)}{8} \beta^{p-2} t^2 + \frac{(p-5)(p-3)}{16} \beta^{p-1} t + \frac{(p-2)(p-4)}{16},
\]
is used instead of the function in (2). Therefore, not only Steepest Descent-type local search methods but also Newton-type local search methods can be considered in the optimization process of the problem \((P)\).

Lemma 2.1. Let \(\beta > 0\), then
\[
0 \leq \tilde{q}_p(t, \beta) - q_p(t) \leq \mu_p(0, \beta).
\]

Proof. Since the functions \(\tilde{q}_p(t, \beta)\) and \(q(t)\) have the same values outside the \(\beta\) neighborhood of 0, it is necessary to prove that the inequality (4) holds as \(-\beta \leq t \leq \beta\). Let us consider the case \(-\beta \leq t < 0\). Since, \(0 \leq \tilde{q}_p(t, \beta) - q_p(t) = \mu_p(t, \beta)\) and \(\tilde{q}_p(t, \beta)\) is increasing on the interval \([-\beta, 0]\), the maximum difference is obtained
at the right endpoint $t = 0$. For $0 \leq t \leq \beta$, $0 \leq \tilde{q}_p(t, \beta) - q_p(t) = \mu_p(t, \beta) - t^p$ and it is decreasing on the interval $[0, \beta]$. Thus, the maximum difference is obtained at the left end point $t = 0$. Therefore the inequality in (4) is obtained. \hfill \Box

The smoothing function is controlled by the parameter $\beta$. The effects of the parameter changes on $q_p(t, \beta)$ are illustrated in Fig. 1 (a) and (b).

**Corollary 1.** Let $\beta \to 0$, then the smoothing function $\tilde{q}_p(t, \beta)$ approaches $q_p(t)$.

**Proof.** From the Lemma 1 we have

$$0 \leq \tilde{q}_p(t, \beta) - q_p(t) \leq \mu_p(0, \beta).$$

Since $\mu_p(0, \beta) \to 0$ as $\beta \to 0$, we obtain $\tilde{q}_p(t, \beta) - q_p(t) \to 0$ as $\beta \to 0$. \hfill \Box

**Remark 1.** Let $g$ be a smooth function from $\mathbb{R}^n$ to $\mathbb{R}$. If we replace $t$ by $g(x)$ in the function $\tilde{q}_p(t)$ defined in (1), we obtain a smoothing function for $\max\{g(x), 0\}$. Moreover, it is sufficient to consider a smoothing function just for $\max\{g(x), 0\}$ since we have $\max\{g_1(x), g_2(x)\} = g_1(x) + \max\{g_2(x) - g_1(x), 0\}$ for smooth functions $g_1$ and $g_2$.

**Remark 2.** Since we have the identity such as $\min\{g(x), 0\} = -\max\{-g(x), 0\}$ and $|f(x)| = \max\{g(x), 0\} + \max\{-g(x), 0\}$, it is not needed to present the smoothing results for the operators $\min$ and $|\cdot|$.

3. **The Degree of Approximation of Smoothing Function.** Now, let us define the function $\theta$ from $\mathbb{R}^n$ to $\mathbb{R}$ as $\theta(x) = (\sum_{i=1}^m q_{p_i}(g_i(x)))$ and smoothed function $\tilde{\theta}(x, \beta) = (\sum_{i=1}^m \tilde{q}_{p_i}(g_i(x), \beta))$ where $g_i(x): \mathbb{R}^n \to \mathbb{R}$ are smooth functions and $0 < p_i \leq 1$ for $i = 1, 2, \ldots, m$. If $q_{p_i}(t) = \max\{0, t\}^{p_i}$, then the function $\theta$ can be considered as penalty term in constrained optimization [34]. If $q_{p_i}(t) = |t|^{p_i}$, the function $\theta$ can be used as a regularization term in solving inverse problems [6].

![Figure 1](image.png)

**Figure 1.** (a) The green and solid graph is the graph of $q_1(t)$, the blue and dashed one is the graph of $\tilde{q}_1(t, 0.5)$, the red and dotted one is the graph of $\tilde{q}_1(t, 1)$. (b) The green and solid graph is the graph of $q_{1/2}(t)$, the blue and dashed one is the graph of $\tilde{q}_{1/2}(t, 0.5)$, the red and dotted one is the graph of $\tilde{q}_{1/2}(t, 1)$. 


Here, we present the error estimations of the difference between optimal solution and value of smoothed function and original function by the following results.

**Theorem 3.1.** Let \( x \in \mathbb{R}^n \) and \( \beta > 0 \)

\[
0 \leq \hat{\theta}(x, \beta) - \theta(x) \leq \sum_{i=1}^{m} \mu_{p_i}(0, \beta).
\]  

**Proof.** From Lemma 2.1 we obtain

\[
\hat{\theta}(x, \beta) - \theta(x) = \sum_{i=1}^{m} \tilde{q}_p_i(g_i(x), \beta) - \sum_{i=1}^{m} q_{p_i}(g_i(x))
\]

\[
= \sum_{i=1}^{m} (\tilde{q}_p_i(g_i(x), \beta) - q_{p_i}(g_i(x)))
\]

\[
\leq \sum_{i=1}^{m} \mu_{p_i}(0, \beta).
\]

Therefore the inequality (5) is obtained.

**Corollary 2.** Let \( \beta \to 0 \) then, the function \( \hat{\theta}(x, \beta) \) approaches \( \theta(x) \).

**Proof.** From the Lemma 2 we have

\[
0 \leq \hat{\theta}(x, \beta) - \theta(x) \leq \sum_{i=1}^{m} \mu_{p_i}(0, \beta).
\]

Since \( \mu_{p_i}(0, \beta) \to 0 \) and \( \sum_{i=1}^{m} \mu_{p_i}(0, \beta) \to 0 \) as \( \beta \to 0 \), we obtain \( \hat{\theta}(x, \beta) - \theta(x) \to 0 \) as \( \beta \to 0 \).

**Theorem 3.2.** Let \( \{\beta_j\} \to 0 \) and \( x^j \) be a minimizer of \( \hat{\theta}(x, \beta) \). Suppose that \( \tilde{x} \) is an accumulation point of \( \{x^j\} \). Then \( \tilde{x} \) is a minimizer of \( \hat{\theta}(x, \beta) \).

**Proof.** By considering the Theorem 3.1 and Corollary 2, the proof is obtained.

**Theorem 3.3.** Suppose that \( x^* \) is a local minimizer of \( \theta(x) \) and \( \tilde{x} \) is a local minimizer of \( \varphi(\hat{\theta}(x, \beta)) \). Then, we have

\[
|\hat{\theta}(\tilde{x}, \beta) - \theta(x^*)| \leq \sum_{i=1}^{m} \mu_{p_i}(0, \beta).
\]

Moreover, we have \( \varphi(\hat{\theta}(\tilde{x}, \beta)) \to \varphi(\theta(x^*)) \) for \( \beta \to 0 \).

**Proof.** From Lemma 2.1 and Theorem 3.1, we obtain

\[
|\hat{\theta}(\tilde{x}, \beta) - \theta(\tilde{x})| \leq |\hat{\theta}(\tilde{x}, \beta) - \theta(x^*)|
\]

\[
\leq |\hat{\theta}(x^*, \beta) - \theta(x^*)|
\]

\[
\leq \sum_{i=1}^{m} \mu_{p_i}(0, \beta).
\]

Therefore the inequality (6) is obtained and the proof is completed.

Let \( \varphi : \mathbb{R}_+ \to \mathbb{R}_+ \) be a second-order differentiable and increasing (or decreasing) function such that

\[
\varphi(t) \leq h(t),
\]

(7)
where \( h(t) \) is linear function on \( \mathbb{R} \). The well known examples of \( \varphi \) are \( t, \frac{1}{1+\alpha t}, \frac{\alpha t}{1+\alpha t} \) and \( \arctan(\alpha t) \) which are commonly used as potential function in solving inverse problems \cite{24} and considered in global optimization of mixed integer non-linear problems \cite{47}.

**Theorem 3.4.** Let the function \( \varphi \) is defined as in (7). Suppose that \( x^* \) is a local minimizer of \( \varphi(\theta(x)) \) and \( \bar{x} \) is a local minimizer of \( \varphi(\theta(x, \beta)) \). Then, we have

\[
|\varphi(\hat{\theta}(\bar{x}, \beta)) - \varphi(\theta(x))| \leq K \left( \sum_{i=1}^{m} \mu_p_i(0, \beta) \right),
\]

where \( K > 0 \) is a constant.

**Proof.** For any \( \beta > 0 \)

\[
\varphi(\hat{\theta}(x, \beta)) - \varphi(\theta(x)) \leq h\left(\hat{\theta}(x, \beta)\right) - h\left(\theta(x)\right) = h\left(\hat{\theta}(x, \beta) - \theta(x)\right) \]

\[
\leq K \sum_{i=1}^{m} \mu_p_i(0, \beta).
\]

Therefore the proof is completed. \( \square \)

**Corollary 3.** Let \( \beta \to 0 \) then, the function \( \varphi(\hat{\theta}(x, \beta)) \) approaches \( \varphi(\theta(x)) \).

**Proof.** The proof is similar to Corollary 2. \( \square \)

**Theorem 3.5.** Let the function \( \varphi \) is defined as in (7), \( \{\beta_j\} \to 0 \) and \( x^j \) be a minimizer of \( \varphi(\hat{\theta}(x, \beta)) \). Suppose that \( \bar{x} \) is an accumulation point of \( \{x^j\} \). Then \( \bar{x} \) is a minimizer of \( \varphi(\hat{\theta}(x, \beta)) \).

**Proof.** By considering the Theorem 3.4 and Corollary 3, the proof is obtained. \( \square \)

**Theorem 3.6.** Let the function \( \varphi \) is defined as in (7). Suppose that \( x^* \) is a local minimizer of \( \varphi(\theta(x)) \) and \( \bar{x} \) is a local minimizer of \( \varphi(\theta(x, \beta)) \). Then, we have

\[
|\varphi(\hat{\theta}(\bar{x}, \beta)) - \varphi(\theta(x^*))| \leq K \left( \sum_{i=1}^{m} \mu_p_i(0, \beta) \right),
\]

where \( K > 0 \) is a constant. Moreover, we have \( \varphi(\hat{\theta}(\bar{x}, \beta)) \to \varphi(\theta(x^*)) \) for \( \beta \to 0 \).

**Proof.** Since the function \( \varphi \) is increasing,

\[
\varphi\left(\hat{\theta}(\bar{x}, \beta)\right) - \varphi\left(\theta(\bar{x})\right) \leq \varphi\left(\hat{\theta}(\bar{x}, \beta)\right) - \varphi\left(\theta(x^*)\right)
\]

\[
\leq \varphi\left(\hat{\theta}(x^*, \beta)\right) - \varphi\left(\theta(x^*)\right).
\]

By considering the properties of the function \( \varphi \) we have

\[
\varphi\left(\hat{\theta}(x^*, \beta)\right) - \varphi\left(\theta(x^*)\right) \leq h\left(\hat{\theta}(x^*, \beta)\right) - h\left(\theta(x^*)\right)
\]

\[
= K \left(\hat{\theta}(x^*, \beta) - \theta(x^*)\right)
\]

\[
\leq K \left( \sum_{i=1}^{m} \mu_p_i(0, \beta) \right)
\]

for \( K > 0 \). The proof is obtained similarly when \( \varphi \) is a decreasing function. \( \square \)
4. Applications of the Smoothing Technique in Global Optimization. Our main contribution with the smoothing approach is to smooth out the non-smooth problem \((P)\) and make it suitable for applying global optimization techniques which are used for smooth problems. According to the above smoothing approach, the smoothing function of \(f(x)\) is shown by \(\tilde{f}(x)\) (see Definition 1) and the problem \((P)\) is restated as

\[
\min_{x \in \mathbb{R}^n} \tilde{f}(x)
\]

where the objective function \(\tilde{f} : \mathbb{R}^n \to \mathbb{R}\) is continuously differentiable. Since the problem \((P)\) transforms into a smooth global optimization problem, a new global optimization method based on smooth local searches can, thus, be used to solve the problem \((P)\).

The smoothing approach can be applied to many different branches of optimization such as exact penalty functions, inverse problems etc. The smoothing approach can also be used for the global optimization of non-smooth problem by combining with gradient based global optimization algorithm. In this section, we use our smoothing approach and consider the auxiliary function based global optimization algorithm AFA proposed in [41]. The algorithm is programmed in MATLAB and has been executed on Intel Core i5-3337U 1.8GHz with 7.8 Gb Ram. In the following example we show the usage of the new smoothing approach in order to smooth out non-convex, non-smooth and non-Lipschitz functions.

Problem. Let us first define the following \(n\)-dimensional unconstrained global optimization problem:

\[
\min_{x \in \mathbb{R}} f(x) = -\sum_{j=1}^{m} \frac{a_j}{1 + \|x - b_j\|_{p_j}^p}, \quad (10)
\]

where \(m\) is the number of local minimizer, \(n\) is the dimension of the problem, \(b_j = (b_{j,1}, b_{j,2}, \ldots, b_{j,n}) \in \mathbb{R}^n\), \(a_j\) is the real number and \(\|x\|_{p_j} = (\sum_{i=1}^{n} |x_i|^{p_j})^{1/p_j}\), \(p_j > 0\) for \(j = 1, 2, \ldots, m\).

It can be observed that the number of local minimizers, the dimension, the location of the local minimizers of the of the objective function \(f(x)\) can be chosen. Moreover, the biggest value of \(a_j\) for \(j = 1, 2, \ldots, m\) gives the location of the global minimizer of \(f(x)\). Depending the parameter \(p\), the minimization problem at \((10)\) may be smooth, non-smooth and non-Lipschitz.

Example 1 (2-dimensional non-Lipschitz problem). Based on the Problem 4, we design a two dimensional non-smooth and non-Lipschitz problem as

\[
\min_{x \in \mathbb{R}} f(x) = -\sum_{j=1}^{5} \frac{a_j}{1 + \|x - b_j\|_{p_j}^p}, \quad (11)
\]

where the values of \(b_j = (b_{j,1}, b_{j,2}) \in \mathbb{R}^2\), \(a_j \in \mathbb{R}\) and \(p_j > 0\) for \(j = 1, 2, \ldots, 5\) are presented in the Table 1. It can be observed from the formulation of the problem, the objective function \(f(x)\) has local minimizers approximately at the points \((4, 3), (-2, -3)\) and \((-3, 3), (4, -4)\) and \((0, 0)\) and, \(f\) attains the global minimum at the point \(x^* \approx (0, 0)\). At the points \(x = (-4, 4)\) and \(x = (0, 0)\), the function \(f\) is not differentiable. At the point \(x = (-4, 4)\), the directional derivatives of \(f\) are exist but not equal and at \(x = (0, 0)\), the directional derivatives of \(f\) are not exist. Clearly, \(f\) is continuous but it is non-smooth and non-Lipschitz. Therefore, it is not possible to find the local and global minimizers of this function by using classical gradient-based...
Table 1. The values of $a_j$, $b_j$ and $p_j$ in Example 1

| $j$ | $1$  | $2$  | $3$  | $4$  | $5$  |
|-----|------|------|------|------|------|
| $a_j$ | $3$  | $2$  | $3$  | $3.5$  | $5$  |
| $b_j$ | $(4, 3)$  | $(-2, -3)$  | $(-3, 3)$  | $(4, -4)$  | $(0, 0)$  |
| $p_j$ | $2$  | $2$  | $2$  | $1$  | $\frac{2}{3}$  |

local searches. The graph of that function is shown in Fig. 2 (a). As it is studied

in this paper, the objective function $f$ is smoothed out by applying the smoothing approach. We construct the following smoothing function of $f$

$$
\tilde{f}(x, \beta) = -\left( -\sum_{j=1}^{3} \frac{a_j}{1 + \|x - b_j\|^{p_j}} + \frac{a_4}{1 + \tilde{Q}(x, \beta)} + \frac{a_5}{1 + \tilde{R}(x, \beta)} \right),
$$

where $\tilde{Q}(x, \beta) = \tilde{q}_1(x_1, \beta) + \tilde{q}_2(x_2, \beta)$ and $\tilde{R}(x, \beta) = \tilde{r}_1(x_1, \beta) + \tilde{r}_2(x_2, \beta)$. The function $\tilde{q}_i$ is defined as

$$
\tilde{q}_i(x_i, \beta) = \begin{cases} 
|x_i - b_i|, & |x_i - b_i| \geq \beta, \\
\frac{1}{2\beta}(x_i - b_i)^2 + \frac{1}{2}\beta, & |x_i - b_i| < \beta,
\end{cases}
$$

where $b_1 = -4$, $b_2 = 4$ and the function $\tilde{r}_i$ is defined as

$$
\tilde{r}_i(x_i, \beta) = \begin{cases} 
|x_i|^2, & |x_i| \geq \beta, \\
\frac{1}{3\beta^2}(x_i)^2 + \frac{2}{3}\beta^2, & |x_i| < \beta,
\end{cases}
$$

for $i = 1, 2$ with $\beta > 0$. We consider the following problem

$$
\min_{x \in \mathbb{R}} \tilde{f}(x, \beta), \tag{12}
$$

instead of (11). The new problem can be defined as the global minimization of the continuously differentiable function. The graph of the function $\tilde{f}$ is shown in Fig 2 (b) with $\beta = 0.1$. At this stage, any of gradient based global optimization algorithms can be used to solve the problem (12). For example, we apply AFA to problem (12) and present the results of the minimization process in Table 2. Table 2 is constituted by applying
AFA to problem (12) and reporting the value of smoothing parameter $\beta$, the iteration number by $k$, the $k-$th local minimizer of $f(x, \beta)$ by $\bar{x}_k$, the corresponding value of $k-$th local minimizer in smoothing function by $f(\bar{x}_k, \beta)$, the corresponding value of $k-$th local minimizer in original function by $f(\bar{x}_k)$, the $k-$th local minimizer of $f(x)$ by $\bar{x}_k^*$ and the corresponding value of $k-$th local minimizer in original function by $f(\bar{x}_k^*)$. It can be observed from the Table 2 that the local minimizers of original and smoothing functions are located at the points with very small differences. If the smoothing parameter $\beta$ approaches to 0, the differences between the locations and values of local minimizers of original and smoothing functions becomes smaller.

**Example 2.** Let us define the following constrained optimization problem:

$$
\min f(x) = -\sum_{j=1}^{6} \frac{a_j}{1 + \|x - b_j\|^p_j},
$$

s.t. $g_1(x) = x + 9 \leq 0,$

$$
g_2(x) = 8 - x \leq 0,
$$

where the values of $b_j, a_j \in \mathbb{R}$ and $p_j > 0$ for $j = 1, 2, \ldots, 6$ are presented in the Table 3. We consider the penalty approach in [34] to solve the problem given in (13)

**Table 3. The values of $a_j, b_j$ and $p_j$ in Example 2**

| $j$ | 1   | 2   | 3   | 4   | 5   | 6   |
|-----|-----|-----|-----|-----|-----|-----|
| $a_j$ | 1.8 | 1.2 | 4.8 | 3.5 | 2.8 | 2.1 |
| $b_j$ | -7  | 2   | -3.5| 6   | -1  | 4   |
| $p_j$ | 2   | 2   | 1   | 1   | $\frac{2}{3}$ | $\frac{1}{2}$ |

defining the surrogate unconstrained problem as

$$
\min_{x \in \mathbb{R}} F_\lambda(x) = -\sum_{j=1}^{6} \frac{a_j}{1 + \|x - b_j\|^p_j} + \lambda \sum_{i=1}^{2} G_i(x),
$$

where $\lambda > 0$ and $G_i(x) = \max\{g_i(x), 0\}$. It can be seen that the function $F_\lambda(x)$ in (14) is non-smooth and non-Lipschitz (see Fig. 3 (a)). By employing smoothing approach, the smooth function that mimics the original function is obtained. We construct the following smoothing function of $F_\lambda$ as

$$
\tilde{F}_\lambda(x, \beta) = -\sum_{j=1}^{2} \frac{a_j}{1 + \|x - b_j\|^2} - \sum_{j=3}^{5} \frac{a_j}{1 + Q_j(x, \beta)} - \sum_{j=5}^{6} \frac{a_j}{1 + R_j(x, \beta)} + \lambda \sum_{i=1}^{2} \tilde{G}_i(x, \beta),
$$
where

\[ \tilde{Q}_j(x, \beta) = \begin{cases} |x - b_j|, & |x - b_j| \geq \beta, \\ \frac{1}{2\beta}(x - b_j)^2 + \frac{1}{2}\beta, & |x - b_j| < \beta, \end{cases} \]

for \( j = 3, 4 \) and

\[ \tilde{R}_j(x, \beta) = \begin{cases} |x - b_j|^{p_j}, & |x - b_j| \geq \beta, \\ \frac{p_j}{2\beta^{p_j-1}}(x - b_j)^2 + \frac{\beta^{p_j}(2-p_j)}{2}, & |x - b_j| < \beta, \end{cases} \]

for \( j = 5, 6 \) with \( \beta > 0 \). The smoothing penalty term is defined as

\[ \tilde{G}_i(x, \beta) = \begin{cases} G_i(x), & |g_i(x)| \geq \beta, \\ \frac{1}{4\beta}(g_i(x))^2 + \frac{1}{2}(g_i(x)) + \frac{1}{4}\beta, & |g_i(x)| < \beta, \end{cases} \]

for \( i = 1, 2 \) with \( \beta > 0 \). Therefore, we consider the following problem

\[
\min_{x \in \mathbb{R}} \tilde{F}_{\lambda}(x, \beta),
\]

instead of the problem (14). The new problem can be defined as the global minimization of the continuously differentiable function. The graph of the function \( \tilde{F}_{\lambda} \) is shown in Fig. 3 (b). At this stage, any of Lipschitz constant based or gradient

based global optimization algorithms can use to solve the problem (15). As it is used in Example 1, we again consider the AFA and present the results of the minimization process in Table 4. Table 4 is constituted by applying AFA to problem (15) and reporting the value of smoothing parameter \( \beta \), the iteration number by \( k \), the \( k \)-th local minimizer of smoothing penalty function \( \tilde{F}_{\lambda}(x, \beta) \) by \( \tilde{x}_k \), the corresponding value of \( k \)-th local minimizer in smoothing penalty function by \( \tilde{F}_{\lambda}(\tilde{x}_k, \beta) \), the corresponding value of \( k \)-th local minimizer in penalty function by \( F_{\lambda}(\tilde{x}_k) \), the corresponding value of \( k \)-th local minimizer in original function by \( f(\tilde{x}_k) \), the \( k \)-th local minimizer of the original function \( f(x) \) by \( x^*_k \) and the corresponding value of \( k \)-th local minimizer in original function by \( f(x^*_k) \).

The numerical results in Table 4 show that our smoothing approach is viable for constrained global optimization, too. It can be observed again that, if the smoothing parameter \( \beta \) approaches to 0, the differences between the locations and values of local minimizers of original and smoothing functions are getting smaller.
Table 4. Table of minimization process of the Example 2

| $\beta$ | $k$ | $\tilde{x}_k$ | $F_\lambda(\tilde{x}_k, \beta)$ | $F_\lambda(\tilde{x}_k)$ | $f(\tilde{x}_k)$ | $x_k^*$ | $f(x_k^*)$ |
|---------|-----|----------------|-------------------------------|-------------------|----------------|--------|----------|
| 10^{-4} | 1   | 4.0001         | -5.0017                      | -5.0320          | -5.0320        | 4.0000 | -5.0504 |
|         | 2   | 6.0000         | -5.7891                      | -5.7893          | -5.7893        | 6.0000 | -5.7893 |
|         | 3   | -3.5000        | -7.0153                      | -7.0156          | -7.0156        | -3.5000| -7.0156 |
| 10^{-6} | 1   | -0.9998        | -5.2906                      | -5.3361          | -5.3361        | -1.0000| -5.4260 |
|         | 2   | 3.9999         | -2.5950                      | -2.5966          | -2.5966        | 4.0000 | -2.5967 |
|         | 3   | -3.5000        | -7.0132                      | -7.0156          | -7.0156        | -3.5000| -7.0156 |

We have given two different one-dimensional numerical examples in order to illustrate how the presented smoothing approach works with global optimization algorithms. We consider some well-known non-smooth test problems and apply AFA to find the global minimizer of non-smooth test problems. Furthermore, we compare our results with the results obtained from Global Descent Algorithm (GDA) presented in [14] in terms of total iteration numbers, function evaluations, total computation time. The list of test problems is shown in Table 5. The results of

Table 5. The list of non-smooth test problems

| Problem No. | Function Name | Dimension $n$ | Region | Global minimum |
|-------------|---------------|---------------|--------|----------------|
| 2           | Problem 1 in [14] | 1             | [−10, 10] | 1              |
| 3           | Problem 2 in [14] | 1             | [−10, 10] | 1              |
| 4           | Problem 3 in [14] | 2             | [−10, 10] | 2              |
| 5, 6, 7, 8, 9, 10, 11 | Problem 7 in [14] | 2, 3, 4, 5, 10, 15, 20 | [−10, 10] | 0              |

Table 6. The numerical results for the non-smooth test problems

| Problem No. | $x_0$ | AFA | GDA |
|-------------|-------|-----|-----|
|             | $f^*$ | f.eval | Time(sec) | $f^*$ | f.eval | Time(sec) |
| 2           | 10    | 1.0000 | 106 | 0.0155 | 1.0000 | 362 | 0.1552 |
| 3           | 10    | 1.0000 | 219 | 0.0178 | 1.0000 | 524 | 0.3779 |
| 4           | (10, 10) | 2.0000 | 539 | 0.2676 | 2.0000 | 1136 | 2.4738 |
| 5           | (5, 5) | 1.0000 | 367 | 0.2432 | 0.0000 | 1212 | 2.9559 |
| 6           | (5, 5, 5) | 2.9806e-005 | 471 | 0.2458 | 0.0000 | 1678 | 5.1616 |
| 7           | (5, . . ., 5) | 2.9986e-005 | 614 | 0.2771 | 0.0000 | 2614 | 7.5685 |
| 8           | (5, . . ., 5) | 4.9992e-005 | 733 | 0.3065 | 2e-004 | 3484 | 10.0016 |
| 9           | (5, . . ., 5) | 9.9991e-005 | 2270 | 0.4037 | 6e-004 | 10344 | 37.5289 |
| 10          | (5, . . ., 5) | 1.4982e-006 | 3278 | 0.4986 | 8e-004 | 14390 | 55.3796 |
| 11          | (5, . . ., 5) | 1.6520e-013 | 8262 | 0.7262 | 5e-004 | 42446 | 118.4403 |

AFA and their comparison with the results of GDA on the considered non-smooth test problems are presented in Table 6. We consider the smoothing function of the objective function instead of the original objective function. The table is constituted by applying AFA to the smoothed version of test problems from the starting point $x_0$ and reporting the global minimum value of original objective function by $f^*$, the number of function evaluations including local searches by f.eval and CPU time by
Time (sec). We find the global minimizer of the smoothing function and compute the corresponding global minimum value by putting it in the original function. The results for GDA on test problems are taken from [14]. It can be observed that AFA presents satisfactory results with less function evaluations and computation time than GDA.

5. Conclusion. In this study, we have introduced a new smoothing technique for non-smooth functions. The proposed smoothing technique can be used in solving penalty, minimax, inverse etc. type problems and also global optimization of continuous, multi-modal but non-smooth, non-convex and also non-Lipschitz problems as well as many optimization problems which contain any of “min, min^p, max, max^p, |·| and |·|^p” for 0 < p < 1 operators. Our smoothing technique can be easily controlled by one parameter. This parameter gives a possibility to obtain sensitive approximation (or representation) to the original non-smooth function.

We visualize the process of our smoothing approach on two different examples. According to numerical results in Table 6, the smoothing approach notably decrease the computational cost in solving the test problems.

For future studies, the cooperation of smoothing approach with DIRECT method proposed in [13] can be interesting for this purpose. Finally, the application of this approach on real-world problems could be promising.

REFERENCES

[1] A. M. Bagirov, A. Al Nuaimat and N. Sultanova, Hyperbolic smoothing function method for minimax problems, Optimization, 62 (2013), 759–782.
[2] D. Bertsekas, Nondifferentiable optimization via approximation, Mathematical Programming Study, 3 (1975), 1–25.
[3] B. C. Cetin, J. Barhen and J. W. Burdick, Terminal repeller unconstrained subenergy tunneling (TRUST) for fast global optimization, J. Optimiz. Theory Appl., 77 (1993), 97–126.
[4] C. Chen and O. L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problem, Comput. Optim. Appl., 5 (1996), 91–138.
[5] X. Chen, Smoothing methods for nonsmooth, nonconvex minimization, Math. Program. Ser. B, 134 (2012), 71–99.
[6] X. Chen, M. K. Ng and C. Zhang, Non-lipschitz l^p-regularization and box constrained model for image processing, IEEE Trans. Image. Process., 21 (2012), 4709–4721.
[7] F. H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
[8] V. F. Demyanov and V. N. Malozemov, Introduction to Minimax, 2nd edition, Dover, New York 1990.
[9] V. F. Demyanov and A. M. Rubinov, Constructive Nonsmooth Analysis, Verlag Peter Lang, Frankfurt, 1995.
[10] R. P. Ge, A filled function method for finding global minimizer of a function of several variables, Math. Program., 46 (1990), 191–204.
[11] R. P. Ge, The theory of filled function method for finding global minimizer of a nonlinearily constrained minimization problem, J. Comput. Math., 5 (1987), 1–9.
[12] J. H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, 1975.
[13] D. R. Jones, C. D. Perttunen and B. E. Stuckman, Lipschitzian optimization without the Lipschitz constant, J. Optim. Theory Appl., 79 (1993), 157–181.
[14] A. Ketfi-Cherif and A. Ziadi, Global descent method for constrained continuous global optimization, Appl. Math. Comput., 244 (2014), 209–221.
[15] S. Kirkpatrick, Jr. C. D. Gelatt and M. P. Vecchi, Optimization by simulated annealing, Science, 220 (1983), 671–680.
[16] D. E. Kvasov and M. S. Mukhametzhanov, Metaheuristic vs. deterministic global optimization algorithms: The univariate case, Appl. Math. Comput., 318 (2018), 245–259.
[17] D. E. Kvasov and Y. D. Sergeyev, Deterministic approaches for solving practical black-box global optimization problems, Adv. Eng. Softw., 80 (2015), 58–66.
[18] J. Lee and D. Skipper, *Virtuous smoothing for global optimization*, *J. Glob. Optim.*, 69 (2017), 677–697.

[19] A. V. Levy and A. Montalvo, *The tunneling algorithm for the global minimization of functions*, *SIAM J. Sci. Stat. Comput.*, 6 (1985), 15–29.

[20] S. J. Lian, *Smoothing approximation to \( l_1 \) exact penalty for inequality constrained optimization*, *Appl. Math. Comput.*, 219 (2012), 3113–3121.

[21] H. Lin, Y. Wang, Y. Gao and X. Wang, *A filled function method for global optimization with inequality constraints*, *Comput. Appl. Math.*, 37 (2018), 1524–1536.

[22] S. Ma, Y. Yang and H. Liu, *A parameter free filled function for unconstrained global optimization*, *Appl. Math. Comput.*, 215 (2010), 3610–3619.

[23] C. K. Ng, D. Li and L. S. Zhang, *Global descent method for global optimization*, *SIAM J. Optim.*, 20 (2010), 3161–3184.

[24] M. Nikolova, C. K. Ng and C. P. Tam, *Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction*, *IEEE T. Image Process.*, 19 (2010), 3073–3088.

[25] A. Ozmen, E. Kropat and G.-W. Weber, *Spline regression models for complex multi-modal regulatory networks*, *Optim. Method Softw.*, 29 (2014), 515–534.

[26] A. Ozmen and G. W. Weber, *RMARS: Robustification of multivariate adaptive regression spline under polyhedral uncertainty*, *J. Comput. Appl. Math.*, 259 (2014), 914–924.

[27] R. Paulavicius, J. Žilinskas and A. Grothey, *Parallel branch and bound for global optimization with combination of Lipschitz bounds*, *Optim. Method Softw.*, 26 (2011), 487–498.

[28] R. Paulavicius, Y. D. Sergeyev, D. E. Kvasov and J. Žilinskas, *Globally-biased DISIMPL algorithm for expensive global optimization*, *J. Glob. Optim.*, 59 (2014), 545–567.

[29] R. Paulavicius, C. Lakhdar and J. Žilinskas, *Global optimization based on bisection of rectangles, function values at diagonals, and a set of Lipschitz constants*, *J. Glob. Optim.*, 79 (2018), 5–20.

[30] S. S. Rao, K. Sundararaju, B. G. Prakash and C. Balakrishna, *A fuzzy goal programming approach for structural optimization*, *AIAA Journal*, 30 (1992), 1425–1432.

[31] M. Resener, S. Haffner, L. A. Pereira and P. M. Pardalos, *Optimization techniques applied for inequality constrained optimization problems*, *Numer. Algebra Cont. Optim.*, 6 (2016), 232–237.

[32] Y. D. Sergeyev, R. G. Strongin and D. Lera, *Introduction to Global Optimization Exploiting Space-Filling Curves*, Springer, New York, 2013.

[33] Y. D. Sergeyev, D. E. Kvasov and M. S. Mukhametzhanov, *On the efficiency of nature-inspired metaheuristics in expensive global optimization with limited budget*, *Sci. Rep.*, 8 (2018), 1–9.

[34] Y. Shi and R. Eberhart, *A modified particle swarm optimizer*, In *Proceedings of IEEE International Conference on Evolutionary Computation*, (1998), 69–73.

[35] R. G. Strongin and Y. D. Sergeyev, *Global Optimization with Nonconvex Constraints: Sequential and Parallel Algorithms*, Kluwer Academic Publishers, Dordrecht, 2000.

[36] H. Tuy, *Cutting plane methods for global optimization*, In *Encyclopedia of Optimization* (eds. C. Floudas and P. Pardalos), 2nd Edition, Springer, New York, (2009), 590–594.

[37] Y. Wang, W. Fang and T. Wu, *A cut-peak function method for global optimization*, *J. Comput. Appl. Math.*, 230 (2009), 135–142.

[38] F. Wei, Y. Wang and H. Lin, *A new filled function method with two parameters for global optimization*, *J. Optim. Theory Appl.*, 163 (2014), 510–527.

[39] A. E. Xavier, *Penalizacao Hiperbolica: Um Novo Metodo para Resolututo de Problemas de Otimizao*, M. Sc. Thesis, COPPE–Federal University of Rio de Janeiro, 1982.

[40] A. E. Xavier, *The hyperbolic smoothing clustering method*, *Patt. Recog.*, 43 (2010), 731–737.

[41] A. E. Xavier and A. A. F. Oliveira, *Optimal covering of plane domains by circles via hyperbolic smoothing*, *J. Glob. Optim.*, 31 (2005), 493–504.
[47] L. Xu, J. Lee and D. Skipper, More virtuous smoothing, *SIAM J. Optim.*, 29 (2019) 1240–1259.
[48] Y. T. Xu, Y. Zhang and S. G. Wang, A modified tunneling function method for non-smooth global optimization and its application in artificial neural network, *Appl. Math. Model.*, 39 (2015), 6438–6450.
[49] Y. Yang and Y. Shang, A new filled function method for unconstrained global optimization, *Appl. Math. Comput.*, 173 (2006), 501–512.
[50] H. Yin, An adaptive smoothing method for continuous minimax problems, *Asia-Pac. J. Oper. Res.*, 32 (2015), 1–19.
[51] I. Zang, A smoothing out technique for min-max optimization, *Math. Program.*, 19 (1980), 61–77.
[52] L. S. Zhang, C. K. Ng, D. Li and W. W. Tian, A new filled function method for global optimization, *J. Glob. Optim.*, 28 (2004), 17–43.
[53] A. Zhigljavsky and A. Zilinskas, *Stochastic Global Optimization*, Springer, New York, 2008.
[54] J. Zilinskas, Branch and bound with simplicial partitions for global optimization, *Math. Model. Anal.*, 13 (2008), 145–159.

Received May 2019; 1st revision September 2019; Final revision October 2019.

E-mail address: nurullahyilmaz@sdu.edu.tr
E-mail address: ahmetsahiner@sdu.edu.tr