SMOOTHING APPROXIMATIONS FOR PIECEWISE SMOOTH FUNCTIONS: A PROBABILISTIC APPROACH

ELMHEIDI AMHRAOUI* AND TAWFIK MASROUR

Laboratory of Mathematical Modeling, Simulation and Smart Systems (L2M3S)
ENSAM-Meknes, Moulay ISMAIL University, Meknes, Morocco

(Communicated by Chongyang Liu)

Abstract. In this article, we present a new approach to construct smoothing approximations for piecewise smooth functions. This approach proposes to formulate any piecewise smooth function as the expectation of a random variable. Based on this formulation, we show that smoothing all elements of a defined space of piecewise smooth functions is equivalent to smooth a single probability distribution. Furthermore, we propose to use the Boltzmann distribution as a smoothing approximation for this probability distribution. Moreover, we present the theoretical results, error estimates, and some numerical examples for this new smoothing method in both one-dimensional and multiple-dimensional cases.

1. Introduction. Piecewise smooth (P-S) functions are an important subclass of nonsmooth functions. They have been studied for a long time since they appear in many fields, including optimization, data modeling, and various equations [7, 13, 21–23]. The main deficiency of these functions is that they are not smooth at certain points in the domain. Therefore, their presence in any problem has many inconveniences. For example, in optimization, if the cost function is not at least differentiable, gradient-based methods are not applicable. A prominent approach for solving this problem is to construct a smoothing approximation for the P-S function. When a smoothing approximation is available, the P-S function in the original problem is replaced by its smoothing approximation, and the resulting problem can be solved by conventional methods.

One of the first studies on smoothing approximations is proposed to solve optimization problems where the cost function is not smooth due to presence of max or min operator [6]. For the same problem introduced in [6], a different smoothing approach is proposed in [29] and it is based on constructing a smooth transition in the kink points and leave the original cost function unchanged at almost every point. In [27], the authors proposed the hyperbolic smoothing technique to solve the min-sum-min problem which is a mathematical model for the clustering problem. The hyperbolic smoothing technique is also used in [3] to solve the finite variant

2010 Mathematics Subject Classification. Primary: 65D10, 26A27; Secondary: 49J52.
Key words and phrases. Piecewise smooth functions, Smoothing approximations, Smoothing techniques, Non-smooth analysis, Boltzmann distribution.

The first author’s work is supported by the national center for scientific and technical research, Morocco.

* Corresponding author: Elmehdi Amhraoui.
of the essential problem known as the minimax problem. Smoothing approximations are also studied in the context of penalty methods. In fact, most of the exact penalty functions are not smooth. Therefore, many papers studied the smoothing approximations for these exact penalty functions [11, 14, 17, 18, 25]. As a result, smoothing techniques of P-S functions play a fundamental role in optimization and penalty methods.

Smoothing approximations are also used to solve various equations, such as the so-called absolute value equation. The main difficulty of this equation is that it contains the component-wise absolute value, which is a P-S function, resulting in limits of analysis and resolution. Therefore, different smoothing approximations are used to solve this equation [10, 12, 19, 24].

Another application of smoothing approximations for P-S functions is data modeling and compression. For example, the model smoothness is an important property for capturing patterns and removing noise from data sets [28]. For these purposes, smoothing is regarded as a fundamental process in data modeling. There are many valuable contributions in this context; see for instance [1, 2, 8, 15, 16, 20].

Last but not least, the blending surface problem is another application of smoothing approximations for P-S functions. This problem aims to create intermediate surfaces that connect other surfaces smoothly. The study of this problem is fundamental in Computer Aided Geometric Design. Many important methods were developed especially for this problem; see for instance [4, 5, 9].

It is now clear that high-quality smoothing approximations of P-S functions are desired and have applications in a wide range of fields. In this paper, we present a new and comprehensive approach to construct smoothing approximations for P-S functions. In section 2, we first reformulate any P-S function as the expectation of a random variable and show how this new formulation reduces the problem of smoothing a defined space of P-S functions to smoothing a single probability distribution. In section 3, we propose to use the Boltzmann distribution to construct the smoothing approximations for P-S functions in the one-dimensional case. Theoretical results and numerical examples are given for this case. In section 4, we extend our approach to the n-dimensional case. Theoretical results and numerical examples are also given for this case. In section 5, we conclude this paper and discuss some of our future works.

2. A new formulation for piecewise smooth functions. Throughout this paper, $\Omega \neq \emptyset$ is an open subset of $\mathbb{R}^n$, $C(\Omega)$ denotes the set of all continuous real-valued functions over $\Omega$ and $C^k(\Omega)$ ($k \geq 1$) denotes the set of real-valued functions with continuous $k-th$ order partial derivatives on $\Omega$. A real-valued function $f : \Omega \rightarrow \mathbb{R}$ is considered to be smooth if it is at least of class $C^1(\Omega)$.

Now, we give the definition of a piecewise smooth function.

**Definition 2.1.** [21, 28] A real-valued function $f : \Omega \rightarrow \mathbb{R}$ is called piecewise smooth function on $\Omega$ if it is continuous on $\Omega$, there exists a finite collection of smooth functions $f_i : \Omega \rightarrow \mathbb{R}$, $i = 1, 2, \ldots, r$, and there exists $\{A_i\}_{i=1, \ldots, r}$ a partition of $\Omega$, such that

$$f(x) = f_i(x) \quad \forall x \in A_i, \ i = 1, \ldots, r.$$  

The collection $\{f_i\}_{i=1, \ldots, r}$ (not necessarily unique) is called a representation of $f$ on $\Omega$. 
Let \( \tau = \{A_1, \ldots, A_m\} \) be a partition of \( \Omega \) and let \( k \geq 1 \) be a positive integer. We define the set \( PS^k_\tau(\Omega) \) as follows:

\[
PS^k_\tau(\Omega) = \left\{ f \in C(\Omega) \mid \exists \{f_i\}_{i=1}^m \subset C^k(\Omega) \text{ such that } f(x) = f_i(x) \ \forall x \in A_i, i = 1, \ldots, m. \right\}
\]

**Remark 1.** \( PS^k_\tau(\Omega) \) is a vector subspace of \( C(\Omega) \).

Let \( f \in PS^k_\tau(\Omega) \) and let \( R_f = (f_1, \ldots, f_m) \in (C^k(\Omega))^m \) be a representation of \( f \).

Although the functions \( f_i : \Omega \rightarrow \mathbb{R}, i = 1, 2, \ldots, r \) are of class \( C^k(\Omega) \), the function \( f \) may not be differentiable over \( \Omega \); see example 1. In this paper, we are interested in constructing a function of class \( C^k(\Omega) \) that approximates the P-S function \( f \).

**Example 1.** Let \( \tau_0 = \{(-\infty, 0], (0, +\infty)\} \). We have \( f(x) = |x| \) and \( g(x) = \max(0, x) \) are two elements of \( PS^\infty_{\tau_0}(\mathbb{R}) \).

\[
|x| = \begin{cases} -x, & x \in (-\infty, 0], \\ x, & x \in (0, +\infty), \end{cases}, \quad \max(0, x) = \begin{cases} 0, & x \in (-\infty, 0], \\ x, & x \in (0, +\infty). \end{cases}
\]

Furthermore, we have \( f_1(x) = x \) and \( f_2(x) = -x \) are of class \( C^\infty(\mathbb{R}) \), but \( f(x) = |x| \) is not differentiable at \( 0 \).

In the following, we will formulate any piecewise smooth functions as an expectation of a discrete random variable. Let \((\mathcal{E}, \mathcal{F}, \mathbb{P})\) be a space of probability and let \( X : \mathcal{E} \rightarrow \Omega \) be a random variable with values on \( \Omega \). For each possible outcome \( x \in \Omega \) of the random variable \( X \), we consider the conditional random variable \( C\mid (X = x) : \mathcal{E} \rightarrow E \) which assigns a label \( i \in E = \{1, 2, \ldots, m\} \) to \( x \) with probability \( p_i(x) \). The label \( i \) denotes the index of the set \( A_i \) and \( p_i(x) \) can be seen as the probability that \( x \) belongs to the set \( A_i \).

**Remark 2.** \( p = (p_1, \ldots, p_m) \) can be any function from \( \Omega \) to \([0, 1]^m\) such that

\[
\sum_{i=1}^m p_i(x) = 1, \ \forall x \in \Omega. \tag{1}
\]

Let \( R = (f_1, \ldots, f_m) \) be an element of \((C^k(\Omega))^m\). We consider the conditional random variable \( Y_R|X = x : \mathcal{E} \rightarrow F(x) \), where

\[
F(x) = \{f_1(x), f_2(x), \ldots, f_m(x)\},
\]

and

\[
(Y_R|X = x) = f_i(x) \iff (C|X = x) = i. \tag{2}
\]

By considering (2), we have:

\[
\mathbb{P}[Y_R = f_i(x)|X = x] = \mathbb{P}[C = i|X = x] = p_i(x). \tag{3}
\]

From (3), we notice that the probability \( \mathbb{P}[Y_R = f_i(x)|X = x] = p_i(x) \) is independent of the vector \( R \), that is, for every \( R, R' \in (C^k(\Omega))^m \), such that \( R = (f_1, \ldots, f_m) \) and \( R' = (f'_1, \ldots, f'_m) \), we have:

\[
\mathbb{P}[Y_R = f_i(x)|X = x] = \mathbb{P}[Y_{R'} = f'_i(x)|X = x] = p_i(x), \ \forall i \in E.
\]

We define \( \hat{Y}_R(x) \) as the mathematical expectation of \( Y_R|X = x \) with respect to the probability distribution \( p(x) = (p_1(x), \ldots, p_m(x)) \).

\[
\hat{Y}_R(x) := \mathbb{E}_p(Y_R|X = x) = \sum_{i=1}^m \mathbb{P}[Y_R = f_i(x)|X = x] \times f_i(x) = \sum_{i=1}^m p_i(x)f_i(x).
\]
In addition, we define $S_p$ as the operator that associates to each $R \in (C^k(\Omega))^m$ the function $Y_R : \Omega \rightarrow \mathbb{R}$. We call $S_p$ the smoothing operator of the vector space $PS^k_{r}(\Omega)$.

If we define $q_i$ as the indicator function of the set $A_i$, that is

$$q_i(x) = \begin{cases} 1, & x \in A_i, \\ 0, & x \notin A_i. \end{cases}$$

then, we have $\sum_{i=1}^{m} q_i(x) = 1$ for each $x \in \Omega$ (because $\{A_1, \cdots, A_m\}$ is a partition of $\Omega$). Thus, $q(x) = (q_1(x), \cdots, q_m(x))$ can be seen as a probability distribution for the conditional random variable $C|(X = x)$. We call $q$ the characteristic probability of the space $PS^k_{r}(\Omega)$. In addition, For each $f \in PS^k_{r}(\Omega)$ with representation $R_f = (f_1, \cdots, f_m)$, we have:

$$S_q(R_f)(x) = \mathbb{E}_q(Y_{R_f}|X = x) = \sum_{i=1}^{m} q_i(x)f_i(x)$$

$$= \begin{cases} f_1(x), & x \in A_1, \\ f_2(x), & x \in A_2, \\ \vdots \\ f_m(x), & x \in A_m. \end{cases} = f(x).$$

Therefore, for each $f \in PS^k_{r}(\Omega)$ with representation $R_f = (f_1, \cdots, f_m)$, we have:

$$f(x) = \mathbb{E}_q(Y_{R_f}|X = x), \forall x \in \Omega.$$

Thus, we obtain a new formulation for $f \in PS^k_{r}(\Omega)$.

For each $f \in PS^k_{r}(\Omega)$, we have $S_q(R_f) = f$, where $R_f = (f_1, \cdots, f_m)$ is any possible representation of the function $f$. However, $q = (q_1, \cdots, q_m)$ is a discontinuous function over $\Omega$, which, on the one hand, explains the non-smoothness of $f$ in most cases and, on the other hand, does not guarantee the smoothness of $S_q(R_f)$ despite the fact that $f_i, i = 1, \cdots, m$ are all smooth. As a result, if we construct a smooth probability distribution $p$ that approximates the characteristic probability $q$, we will obtain a smoothing approximation $S_p(R_f)$ for each $f \in PS^k_{r}(\Omega)$. So smoothing all elements of the vector space $PS^k_{r}(\Omega)$ is equivalent to smooth the characteristic probability $q$ of the space $PS^k_{r}(\Omega)$.

**Theorem 2.2.** Assuming that the function $p$ satisfies the conditions of Remark 2 and that $p_i, i = 1, \cdots, m$ are of class $C^k(\Omega)$. For every $f \in PS^k_{r}(\Omega)$ with representation $R_f$, we have:

(a) $S_p(R_f)$ is of class $C^k(\Omega)$,
(b) $S_p : (C^k(\Omega))^m \rightarrow C^k(\Omega)$ is a linear operator,
(c) $S_p(R_f) \rightarrow f$ pointwise as $p \rightarrow q$ pointwise.

**Proof.** (a) Let $R = (f_1, \cdots, f_m) \in (C^k(\Omega))^m$. Since $p_i, i = 1, \cdots, m$ and $f_i, i = 1, \cdots, m$ are of class $C^k(\Omega)$, then $S_p(R)(x) = \sum_{i=1}^{m} p_i(x) \times f_i(x)$ is of class $C^k(\Omega)$. Particularly, for each $f \in PS^k_{r}(\Omega)$, we have $S_p(R_f)$ is of class $C^k(\Omega)$, where $R_f \in (C^k(\Omega))^m$ is a representation of $f$ on $\Omega$. 

(b) Let \( a, b \in \mathbb{R} \) and let \( R = (f_1, \ldots, f_m) \) and \( R' = (f'_1, \ldots, f'_m) \) be two elements of \( (C^k(\Omega))^m \).

Since \( aR + bR' = (af_1 + bf'_1, \ldots, af_m + bf_m) \in (C^k(\Omega))^m \), then we have:

\[
S_p(aR + bR')(x) = \sum_{i=1}^{m} p_i(x)(af_i(x) + bf'_i(x)) \\
= a \sum_{i=1}^{m} p_i(x)f_i(x) + b \sum_{i=1}^{m} p_i(x)f'_i(x) \\
= aS_p(R)(x) + bS_p(R')(x)
\]

Therefore, \( S_p \) is linear.

(c) Let \( f \in P_{S^k}^l(\Omega) \) and let \( R_f = (f_1, \ldots, f_m) \) be a representation of \( f \). We have:

\[
\left| S_p(R_f)(x) - f(x) \right| = \left| S_p(R_f)(x) - S_q(R_f)(x) \right| \\
= \left| \sum_{i=1}^{m} p_i(x)f_i(x) - \sum_{i=1}^{m} q_i(x)f_i(x) \right| \\
= \left| \sum_{i=1}^{m} (p_i(x) - q_i(x))f_i(x) \right| \\
\leq \sum_{i=1}^{m} |p_i(x) - q_i(x)| \times |f_i(x)| \\
\leq \max_{i=1}^{m} |f_i(x)| \times \sum_{i=1}^{m} |p_i(x) - q_i(x)| \\
= \max_{i=1}^{m} |f_i(x)| \times \|p(x) - q(x)\|_1
\]

thus we obtain:

\[
\left| S_p(R_f)(x) - f(x) \right| \leq c(x) \times \|p(x) - q(x)\|_1, \forall x \in \Omega. \quad (4)
\]

where \( c(x) = \max_{i=1}^{m} |f_i(x)| \). As a result, \( S_p(R_f) \rightarrow f \) pointwise as \( p \rightarrow q \) pointwise.

Considering the formulation above, in what follows, we will try to construct a smooth function \( p \) that satisfies the conditions of Remark 2 and approximates the characteristic probability of the vector space \( P_{S^k}^l(\Omega) \). Once the smooth function \( p \) is available, the smoothing operator \( S_p \) will generate a smoothing approximation for any element of the space \( P_{S^k}^l(\Omega) \).

3. Smoothing approximations for one-dimensional case. Let us consider the partition \( \tau = \{ A_0, A_1, \ldots A_m \} \) of \( \mathbb{R} \), where

\[
A_i = \begin{cases} (-\infty, a_i], & i = 0, \\ (a_i, a_{i+1}], & i = 1, \ldots, m - 1, \\ (a_m, +\infty), & i = m. \end{cases} \quad (5)
\]

and \( a_1 < a_2 < \cdots < a_m \).
We define the functions $h_i, i = 0, \cdots, m$ as follows

$$h_i(x) = \begin{cases} -(x - a_i), & i = 0, \\ -(x - a_i)(x - a_{i+1}), & i = 1, \cdots, m - 1, \\ x - a_m, & i = m. \end{cases} \quad (6)$$

**Proposition 1.** Take $A_i$ and $h_i$ as in (5) and (6), respectively.

For $i = 0, \cdots, m$, we have:

$$(x \in \hat{A}_i) \iff \left( h_i(x) > 0 \text{ and } h_k(x) < 0, \forall k \neq i \right),$$

where $\hat{A}_i$ is the interior of $A_i$.

Furthermore, for $i = 1, \cdots, m$, we have:

$$x = a_i \iff \left( h_i(x) = h_{i-1}(x) = 0 \text{ and } h_k(x) < 0, \forall k \notin \{i, i-1\} \right).$$

To transform the functions $h_i, i = 0, \cdots, m$ into a probability distribution for the random variable $C(X = x)$, we propose the use of the Boltzmann distribution which is expressed in (7)

$$p_i^\alpha(x) = \frac{e^{\alpha h_i(x)}}{\sum_{k=0}^{m} e^{\alpha h_k(x)}}, \quad i = 0, \cdots, m \quad (7)$$

where $\alpha > 0$.

For every $x \in \mathbb{R}$, we have $p_i(x) \in [0, 1], i = 1, \cdots, m$ and $\sum_{i=1}^{m} p_i^\alpha(x) = 1$. Thus, $p^\alpha(x) = (p_0^\alpha(x), \cdots, p_m^\alpha(x))$ is a probability distribution.

The main advantage of the Boltzmann distribution is that it exaggerates the difference between the functions $h_i, i = 0, \cdots, m$. Therefore, $p_i^\alpha(x)$ becomes close to 1 if $x \in A_i$ and close to 0 otherwise. This property makes the Boltzmann distribution widely used. For example, it is used to normalize the output of a neural network to a probability distribution over all possible classes.

For the sake of simplicity, the smoothing operator of the space $PS_\tau^k(\mathbb{R})$ is denoted $S_\alpha$ instead of $S_{\alpha, \tau}$.

**Theorem 3.1.** Take $p_i^\alpha, i = 0, \cdots, m$ as in (7), $\alpha > 0$, and $k \geq 1$, then we have:

(a) $p_i^\alpha : \mathbb{R} \rightarrow [0, 1]$ is of class $C^\infty(\mathbb{R}), i = 0, \cdots, m$,

(b) $\forall f \in PS_\tau^k(\mathbb{R})$, we have $S_\alpha(R_f) \in C^k(\mathbb{R})$, where $R_f$ is a representation of $f$,

(c) $p^\alpha \rightarrow q$ pointwise almost everywhere as $\alpha \rightarrow +\infty$, where $p^\alpha = (p_0^\alpha, \cdots, p_m^\alpha)$ and $q = (q_0, \cdots, q_m)$ is the characteristic probability of $PS_\tau^k(\mathbb{R})$,

(d) $\forall f \in PS_\tau^k(\mathbb{R})$, we have $S_\alpha(R_f) \rightarrow f$ pointwise as $\alpha \rightarrow +\infty$, where $R_f$ is any representation of $f$.

**Proof.** (a) Easy to prove.

(b) Obtained immediately by applying the part (a) of Theorem 2.2.

(c) $x \in \mathbb{R} \setminus \{a_1, a_2, \cdots, a_m\}$

$\implies \exists i \in \{0, 1, \cdots, m\}$ such that $x \in \hat{A}_i$

$\implies h_i(x) > 0$ and $h_k(x) < 0, \forall k \neq i$ (Proposition 1)

$\implies h_k(x) - h_i(x) < 0, \forall k \neq i$
Since we obtain:

\[ e^\alpha(h_u(x) - h_i(x)) \rightarrow 0, \text{ as } \alpha \rightarrow +\infty \]

\[ \frac{1}{1 + \sum_{k \neq i} e^\alpha(h_u(x) - h_i(x))} \rightarrow 1, \text{ as } \alpha \rightarrow +\infty \]

\[ p^\alpha_i(x) \rightarrow 1, \text{ as } \alpha \rightarrow +\infty \]

\[ p^\alpha_i(x) \rightarrow 1 \text{ and } p^\alpha_k(x) \rightarrow 0 \text{ } \forall k \neq i, \text{ as } \alpha \rightarrow +\infty. \]

( because \( \sum_{k=0}^m p^\alpha_k(x) = 1 \))

Therefore, \( \forall x \in \mathbb{R} \setminus \{a_1, a_2, \cdots, a_m\} \) we have, \( p^\alpha \rightarrow q \) as \( \alpha \rightarrow +\infty \).

(d) Let \( f \in PS^\alpha_k(\mathbb{R}) \) and let \( R_f = (f_0, f_1, \cdots, f_m) \) be a representation of \( f \). If \( x \in \mathbb{R} \setminus \{a_1, a_2, \cdots, a_m\} \), we have \( p^\alpha(x) \rightarrow q(x) \) as \( \alpha \rightarrow +\infty \). By applying Theorem 2.2, we obtain:

\[ \forall x \in \mathbb{R} \setminus \{a_1, a_2, \cdots, a_m\}, \ S_\alpha(R_f)(x) \rightarrow f(x) \text{ as } \alpha \rightarrow +\infty. \tag{8} \]

If \( x = a_i \), then we have:

\[
\left| S_\alpha(R_f)(a_i) - f(a_i) \right| = \left| \sum_{k=0}^m p^\alpha_k(a_i)f_k(a_i) - f(a_i) \right| \\
= \left| \sum_{k=0}^m p^\alpha_k(a_i)f_k(a_i) - f_{i-1}(a_i) \right| \left( \text{because } f(a_i) = f_{i-1}(a_i) \right) \\
= \left| \sum_{k=0}^m p^\alpha_k(a_i)(f_k(a_i) - f_{i-1}(a_i)) \right| \\
\leq \sum_{k=0}^m p^\alpha_k(a_i)|f_k(a_i) - f_{i-1}(a_i)| \\
= p^\alpha_{i-1}(a_i)|f_{i-1}(a_i) - f_{i-1}(a_i)| + p^\alpha_i(a_i)|f_i(a_i) - f_{i-1}(a_i)| \\
+ \sum_{\substack{k=0 \atop k \neq i, k \neq i-1}}^m p^\alpha_k(a_i)|f_k(a_i) - f_{i-1}(a_i)| \\
= p^\alpha_i(a_i)|f_i(a_i) - f_{i-1}(a_i)| + \sum_{\substack{k=0 \atop k \neq i, k \neq i-1}}^m p^\alpha_k(a_i)|f_k(a_i) - f_{i-1}(a_i)|
\]

Since \( f \) is assumed to be continuous, then we have \( |f_i(a_i) - f_{i-1}(a_i)| = 0 \). Therefore, we obtain:

\[
\forall i \in \{0, 1, \cdots, m\}, \left| S_\alpha(R_f)(a_i) - f(a_i) \right| \leq \sum_{\substack{k=0 \atop k \neq i, k \neq i-1}}^m p^\alpha_k(a_i)|f_k(a_i) - f_{i-1}(a_i)|. \tag{9}
\]
Furthermore, we have:

\[
p_{\alpha k}(a_i) = \frac{e^{\alpha h_k(a_i)}}{\sum_{j=0}^{m} e^{\alpha h_j(a_i)}}
\]

\[
= \frac{1}{\sum_{j=0}^{m} e^{\alpha (h_j(a_i) - h_k(a_i))}}
\]

\[
\leq \frac{1}{1 + e^{\alpha (h_i(x) - h_k(x))}} , \text{ (because } \sum_{j=0}^{m} e^{\alpha (h_j(x) - h_k(x))} \geq 0)\]

\[
= \frac{1}{1 + e^{-\alpha h_k(a_i)}}, \text{ (because } h_i(a_i) = 0)\]

Since \( h_k(a_i) < 0 \forall k \notin \{i, i-1\} \), then we have \( \frac{1}{1 + e^{-\alpha h_k(a_i)}} \to 0 \text{ as } \alpha \to +\infty \) and \( k \notin \{i, i-1\} \). Therefore, we obtain:

\[
\forall i \in \{0, \cdots, m\}, \forall k \notin \{i, i-1\}, \ p_{\alpha k}(a_i) \to 0 \text{ as } \alpha \to +\infty. \quad (10)
\]

From (9) and (10), we obtain

\[
\forall i \in \{0, \cdots, m\}, \ |S_{\alpha}(R_f)(a_i) - f(a_i)| \to 0 \text{ as } \alpha \to +\infty. \quad (11)
\]

From (8) and (11), we obtain

\[
\forall x \in \mathbb{R}, \ S_{\alpha}(R_f)(x) \to f(x) \text{ as } \alpha \to +\infty.
\]

Remark 3. In practice, \( \alpha = 1 \) is sufficient to obtain better approximations, that is the main advantage of Boltzmann distribution.

To simplify and illustrate our approach, we consider the following partition of \( \mathbb{R} \)

\[
\tau_a = \{(-\infty, a], (a, +\infty)\}. \quad (12)
\]

The partition \( \tau_a \) is a special case of (5), and it is studied extensively; see for instance [26, 28].

For the partition \( \tau_a \), the probability \( p^a \) will take the following form:

\[
p^a(x) = (p_0^a(x), p_1^a(x)) = \left( \frac{1}{1 + e^{\alpha(x-a)}}, \frac{1}{1 + e^{-\alpha(x-a)}} \right).
\]

Let \( f \in PS_{\tau_a}^k(\mathbb{R}) \) and let \( R_f = (f_0, f_1) \) be a representation of \( f \), that is

\[
f(x) = \begin{cases} f_0(x), & x \leq a, \\ f_1(x), & x > a. \end{cases}
\]

the smoothing approximation of \( f \) is

\[
S_{\alpha}(R_f)(x) = \frac{1}{1 + e^{\alpha(x-a)}}f_0(x) + \frac{1}{1 + e^{-\alpha(x-a)}} f_1(x). \quad (15)
\]
According to Theorem 3.1, $S_\alpha(R_f)$ is of class $C^k(\mathbb{R})$ and converges to $f$ as $\alpha$ goes to infinity. Now, let us analyze the error estimates.

**Theorem 3.2.** Take $p^n$, $f$, and $S_\alpha(R_f)$ as in (13), (14), and (15), respectively. We have:

$$\forall x \in \mathbb{R}, \forall \alpha > 0, \ |S_\alpha(R_f)(x) - f(x)| \leq \frac{1}{2} |f_1(x) - f_0(x)|.$$  \hspace{1cm} (16)

and for every compact $C$ of $\mathbb{R}$, for every $\epsilon > 0$, there exists $\alpha > 0$ such that

$$x \in C \Rightarrow |S_\alpha(R_f)(x) - f(x)| \leq \epsilon.$$  \hspace{1cm} (17)

**Proof.** Let $x \in \mathbb{R}$, we have:

$$|S_\alpha(R_f)(x) - f(x)| = |p_0^n(x)f_0(x) + p_1^n(x)f_1(x) - f(x)|$$

$$= |p_0^n(x)f_0(x) + (1 - p_0^n(x))f_1(x) - f(x)|$$

$$= |p_0^n(x)(f_0(x) - f_1(x)) + f_1(x) - f(x)|$$

$$= \begin{cases} (1 - p_0^n(x))|f_1(x) - f_0(x)|, & x \leq a, \\ p_0^n(x)|f_1(x) - f_0(x)|, & x > a. \end{cases}$$

$$= \begin{cases} \frac{1}{1 + e^{\alpha|x-a|}}|f_1(x) - f_0(x)|, & x \leq a, \\ \frac{1}{1 + e^{\alpha|x-a|}}|f_1(x) - f_0(x)|, & x > a. \end{cases}$$

Therefore, we obtain (16).

Let $C$ be a compact of $\mathbb{R}$. Since $g(x) := f_1(x) - f_0(x)$ is continuous in $a$ and $g(a) = 0$, then for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|x - a| \leq \delta \Rightarrow |g(x) - g(a)| \leq \epsilon$$

$$\Rightarrow |f_1(x) - f_0(x)| \leq \epsilon$$

$$\Rightarrow \frac{1}{1 + e^{\alpha|x-a|}}|f_1(x) - f_0(x)| \leq \frac{1}{1 + e^{\alpha|x-a|}}\epsilon, \forall \alpha > 0$$

$$\Rightarrow \frac{1}{1 + e^{\alpha|x-a|}}|f_1(x) - f_0(x)| \leq \epsilon, \forall \alpha > 0$$

$$\Rightarrow |S_\alpha(R_f)(x) - f(x)| \leq \epsilon, \forall \alpha > 0.$$

Therefore, we have:

$$\forall \epsilon > 0, \exists \delta > 0 \text{ such that } (|x - a| \leq \delta \Rightarrow |S_\alpha(R_f)(x) - f(x)| \leq \epsilon, \forall \alpha > 0.) \hspace{1cm} (18)$$

Furthermore, for every $\beta > 0$ we have:

$$|x - a| > \beta \Rightarrow |S_\alpha(R_f)(x) - f(x)| < \frac{1}{1 + e^{\alpha\beta}}|f_1(x) - f_0(x)|, \forall \alpha > 0$$

$$\Rightarrow |S_\alpha(R_f)(x) - f(x)| < \frac{M}{1 + e^{\alpha\beta}}, \forall \alpha > 0$$

( where $M = \max_{x \in C}|f_1(x) - f_0(x)|$)

$$\Rightarrow \forall \epsilon > 0, \exists \hat{\alpha} > 0 \text{ such that } (|S_{\hat{\alpha}}(R_f)(x) - f(x)| < \epsilon, \forall x \in C.)$"
(because \( \lim_{\alpha \to \infty} \frac{MC}{1 + e^{\alpha\delta}} = 0 \))

Therefore, for every \( \beta > 0 \), for every \( \epsilon > 0 \), there exists \( \hat{\alpha} > 0 \) such that
\[
|x - a| > \beta \Rightarrow |S_{\hat{\alpha}}(R_f)(x) - f(x)| \leq \epsilon, \forall x \in C.
\] (19)

For \( \beta = \delta \), we have:
\[
\forall \epsilon > 0, \exists \hat{\alpha} > 0 \text{ such that } (|x - a| > \delta \Rightarrow |S_{\hat{\alpha}}(R_f)(x) - f(x)| \leq \epsilon, \forall x \in C.)
\] (20)

Since (18) is true for every \( \alpha > 0 \) and every \( x \in \mathbb{R} \). Therefore, it is true for \( \hat{\alpha} \) and for every \( x \in C \), that is
\[
\forall \epsilon > 0, \exists \delta > 0 \text{ such that } (|x - a| \leq \delta \Rightarrow |S_{\hat{\alpha}}(R_f)(x) - f(x)| \leq \epsilon, \forall x \in C.)
\] (21)

From (20) and (21), we obtain
\[
\forall \epsilon > 0, \exists \hat{\alpha} > 0 \text{ such that } (|S_{\hat{\alpha}}(R_f)(x) - f(x)| \leq \epsilon, \forall x \in C.)
\]

This completes the proof of (17).

**Example 2.** Consider the two functions \( f(x) = |x| \) and \( g(x) = \max(0, x) \). These functions are important and appear in many problems. However, these functions are not smooth and their presence in any problem leads to limits in analysis. In the following, we will construct a smoothing approximation for \( f \) and \( g \) using our smoothing method. We have:
\[
|x| = \begin{cases} 
-x, & x \in (-\infty, 0], \\
x, & x \in (0, +\infty). 
\end{cases}
\]

and
\[
\max(0, x) = \begin{cases} 
0, & x \in (-\infty, 0], \\
x, & x \in (0, +\infty). 
\end{cases}
\]

So \( f \) and \( g \) are two elements of \( PS_{\infty}^{\infty}(\mathbb{R}) \). As a result, the smoothing approximation of \( f \) and \( g \), denoted \( \hat{f}_\alpha \) and \( \hat{g}_\alpha \), respectively, are obtained by applying the smoothing operator of the space \( PS_{\infty}^{\infty}(\mathbb{R}) \), that is
\[
\hat{f}_\alpha(x) = \frac{1}{1 + e^{\alpha x}} \times (-x) + \frac{1}{1 + e^{-\alpha x}} \times x,
\]

and
\[
\hat{g}_\alpha(x) = \frac{1}{1 + e^{-\alpha x}} \times x.
\]

According to Theorem 3.1, \( \hat{f}_\alpha \) and \( \hat{g}_\alpha \) are of class \( C^\infty(\mathbb{R}) \) and converge to \( f \) and \( g \), respectively, as \( \alpha \) goes to infinity. The graphs of the function \( f \), the smoothing approximation \( \hat{f}_\alpha \), and the error of approximation are shown in Figure 1. In Figure 2, we show the graphs of the function \( g \), the smoothing approximation \( \hat{g}_\alpha \), and the error of the smoothing approximation.

Since \( S_\alpha \) is a linear operator, the smoothing approximation of \( h(x) = |x| + \max(0, x) \) is \( \hat{h}_\alpha(x) = \hat{f}_\alpha(x) + \hat{g}_\alpha(x) \). The graphs of the function \( h \), the smoothing approximation \( \hat{h}_\alpha \), and the error between \( h \) and \( \hat{h}_\alpha \) are shown in Figure 3.

From Figures 1, 2 and 3, we see that our smoothing approximation has a small error. In addition, the error becomes close to zero if the value of the parameter \( \alpha \) is increased.
4. Smoothing approximations for Multiple-dimensional case. In this section, we aim to extend the above techniques to n-dimensional case. First, we consider the partition \( \tau_G = \{ A, A^c \} \) of \( \mathbb{R}^n \), where

\[
A = \{ x \in \mathbb{R}^n \mid G(x) \leq 0 \},
\]

and \( A^c \) is the complement of \( A \). In this case, the probability \( p^\alpha \) will take the following form

\[
p^\alpha(x) = (p_0^\alpha(x), p_1^\alpha(x)) = \left( \frac{1}{1 + e^{\alpha G(x)}}, \frac{1}{1 + e^{-\alpha G(x)}} \right).
\]

Let \( f \in PS^k_{\tau_G}(\mathbb{R}^n) \) and let \( R_f = (f_0, f_1) \) be a representation of \( f \), that is

\[
f(x) = \begin{cases} 
    f_0(x), & G(x) \leq 0, \\
    f_1(x), & G(x) > 0.
\end{cases}
\]

the smoothing approximation of \( f \) is

\[
S_\alpha(R_f)(x) = \frac{1}{1 + e^{\alpha G(x)}} f_0(x) + \frac{1}{1 + e^{-\alpha G(x)}} f_1(x).
\]

**Theorem 4.1.** Take \( S_\alpha \) as in (24). If \( G : \mathbb{R}^n \rightarrow \mathbb{R} \) is of class \( C^k(\mathbb{R}^n) \), then for every \( f \in PS^k_\tau(\mathbb{R}^n) \) with representation \( R_f \), we have:
Figure 2. (a), (b) and (c) contain the graphs of the function \( g(x) = \max(0, x) \) and its smoothing approximation \( \hat{g}_\alpha(x) \) for \( \alpha = 1, \alpha = 2 \) and \( \alpha = 3 \), respectively. (d) The error of the smoothing approximation for different values of \( \alpha \).

(a) \( S_\alpha(R_f) \) is of class \( C^k(\mathbb{R}^n) \),
(b) \( S_\alpha(R_f) \to f \) as \( \alpha \to +\infty \).

Proof. (a) Let \( f \in PS^k_F(\mathbb{R}^n) \) and let \( R_f = (f_0, f_1) \) be a representation of \( f \). Since \( G : \mathbb{R}^n \to \mathbb{R} \) is of class \( C^k(\mathbb{R}^n) \), then \( p_0^\alpha \) and \( p_1^\alpha \) are of class \( C^k(\mathbb{R}^n) \). Furthermore, we have \( f_0 \) and \( f_1 \) are of class \( C^k(\mathbb{R}^n) \). According to equation (24), \( S_\alpha(R_f) \) is of class \( C^k(\mathbb{R}^n) \).

(b) Let \( f \in PS^k_F(\mathbb{R}^n) \) and let \( R_f = (f_0, f_1) \) be a representation of \( f \). For each \( x \in \mathbb{R}^n \), we have:

\[
|S_\alpha(R_f)(x) - f(x)| = |p_0^\alpha(x)f_0(x) + p_1^\alpha(x)f_1(x) - f(x)|
= |p_0^\alpha(x)f_0(x) + (1 - p_0^\alpha(x))f_1(x) - f(x)|
= |p_0^\alpha(x)(f_0(x) - f_1(x)) + f_1(x) - f(x)|
= \begin{cases} 
(1 - p_0^\alpha(x))|f_1(x) - f_0(x)|, & G(x) \leq 0 \\
\frac{1}{1 + e^{-\alpha G(x)}}|f_1(x) - f_0(x)|, & G(x) > 0 
\end{cases}
\]
Figure 3. (a), (b) and (c) contain the graphs of the function $h(x) = |x| + \max(0, x)$ and its smoothing approximation $\hat{h}_\alpha(x)$ for $\alpha = 1$, $\alpha = 2$ and $\alpha = 3$, respectively. (d) The error of the smoothing approximation for different values of $\alpha$.

If $G(x) \neq 0$, we have $\frac{1}{1 + e^{\alpha|G(x)|}} \rightarrow 0$ as $\alpha \rightarrow +\infty$. Therefore, we have:

$$\forall x \in \mathbb{R}^n \setminus \{x \in \mathbb{R}^n | G(x) = 0\}, S_\alpha \big(R_f\big)(x) \rightarrow f(x) \text{ as } \alpha \rightarrow +\infty. \quad (25)$$

If $G(x) = 0$, then $|f_1(x) - f_0(x)| = 0$ (because $f$ is assumed to be continuous). Therefore, we have:

$$\forall x \in \{x \in \mathbb{R}^n | G(x) = 0\}, \ S_\alpha \big(R_f\big)(x) = f(x). \quad (26)$$

From (25) and (26), we obtain $S_\alpha \big(R_f\big) \rightarrow f$ pointwise, as $\alpha \rightarrow +\infty$. \hfill \square

Now, let us analyze the error estimates.

**Theorem 4.2.** Take $p^\alpha$, $f$, and $S_\alpha \big(R_f\big)$ as in (22), (23), and (24), respectively. We have:

$$\forall x \in \mathbb{R}^n, \forall \alpha > 0, \ |S_\alpha \big(R_f\big)(x) - f(x)| \leq \frac{1}{2} |f_1(x) - f_0(x)|, \quad (27)$$

and for every compact $C$ of $\mathbb{R}^n$, for every $\epsilon > 0$, there exists $\alpha > 0$ such that

$$x \in C \Rightarrow |S_\alpha \big(R_f\big)(x) - f(x)| \leq \epsilon. \quad (28)$$
Proof. The proof is similar to the proof of Theorem 3.2. □

Example 3. Let us consider the following function

\[ f(x, y) = \begin{cases} 
-4 \log(1 + (xy)^2), & xy \leq 0, \\
(e^{-(xy)^2} - 1), & xy > 0.
\end{cases} \]

It is clear that \( f \) is not continuously differentiable if \( x = 0 \) or \( y = 0 \). Since \( f \) is an element of the space \( PS_{\tau_G}^\infty(\mathbb{R}^2) \), where \( G(x, y) = xy \), its smoothing approximation, denoted \( \hat{f}_\alpha \), is obtained by applying the smoothing operator of the space \( PS_{\tau_G}^\infty(\mathbb{R}^2) \), that is

\[ \hat{f}_\alpha(x, y) = \frac{1}{1 + e^{\alpha xy}}(-4 \log(1 + (xy)^2)) + \frac{1}{1 + e^{-\alpha xy}}(e^{-(xy)^2} - 1). \]

According to Theorem 4.1, \( \hat{f}_\alpha \) is of class \( C^\infty(\mathbb{R}^n) \) and converges to \( f \) as \( \alpha \) goes to infinity. The graphs of the function \( f(x, y) \), the smoothing approximation \( \hat{f}_\alpha(x, y) \), and the error of approximation are presented in Figure 4. It is clear from Figure 4 that our smoothing approximation has a small error.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{(a) The graph of \( f(x, y) \). (b) The graph of the smoothing approximation \( \hat{f}_\alpha(x, y) \) for \( \alpha = 4 \). (c) The error of the smoothing approximation for \( \alpha = 3 \). (d) The error of the smoothing approximation for \( \alpha = 4 \).}
\end{figure}

In the following, we will look at a more general partition of \( \mathbb{R}^n \).

Let \( \{\tau_i\}_{i=1,\ldots,n} \) be a finite collection of partitions of \( \mathbb{R} \). It is clear that

\[ \tau = \{ A_1 \times A_2, \cdots, A_n | A_i \in \tau_i, i = 1, \cdots, n \}. \tag{29} \]
is a partition of \( \mathbb{R}^n \).

Let \( x = (x_1, \cdots, x_n) \in \mathbb{R}^n \). We define \( p(x) \) as follows:

\[
p(x) = \left( \prod_{i=1}^{n} p^i_j(x_i) \right)_{j_i=1,2,\cdots,|\tau_i|, \forall i}
\]  

(30)

where \( p^i = (p^i_1, \cdots, p^i_{|\tau_i|}) \) denotes the smoothing approximation of the characteristic distribution of \( PS^k_{\tau} (\mathbb{R}^n) \) and \( |\tau_i| \) denotes the cardinal of \( \tau_i \).

**Theorem 4.3.** Take \( \tau \) and \( p \) as in (29) and (30), respectively. We have:

\[
p \rightarrow q \text{ pointwise, as } p^i \rightarrow q^i \text{ pointwise } \forall i = 1, \cdots, n,
\]

(31)

where \( q \) and \( q^i \), \( i = 1, \cdots, n \) are the characteristic probabilities of \( PS^k_{\tau} (\mathbb{R}^n) \) and \( PS^k_{\tau} (\mathbb{R}) \), \( i = 1, \cdots, n \), respectively.

**Proof.** The proof is obtained by noticing that

\[
q(x) = \left( \prod_{i=1}^{n} q^i_j(x_i) \right)_{j_i=1,2,\cdots,|\tau_i|, \forall i}.
\]

\[ \square \]

**Theorem 4.4.** Take \( \tau \) and \( p \) as in (29) and (30), respectively. If \( p^i_j, i = 1 \cdots, n; j_i = 1, \cdots, |\tau_i| \) are of class \( C^k(\mathbb{R}) \), then for every \( f \in PS^k_{\tau} (\mathbb{R}^n) \) we have:

(a) \( S_p(R_f) \in C^k(\mathbb{R}^n) \),

(b) \( S_p(R_f) \rightarrow f \text{ pointwise, as } p^i \rightarrow q^i \text{ pointwise } \forall i = 1, \cdots, n \),

where \( q^i \), \( i = 1, \cdots, n \) are the characteristic probabilities of \( PS^k_{\tau} (\mathbb{R}) \), \( i = 1, \cdots, n \), respectively.

**Proof.** The proof is obtained by applying Theorem 2.2 and Theorem 4.3.\[ \square \]

To simplify, we consider the following two partitions of \( \mathbb{R} \)

\[
\tau_a = \{ (-\infty, a], (a, +\infty) \}
\]

\[
\tau_b = \{ (-\infty, b], (b, +\infty) \}
\]

It is clear that

\[
\tau_{a,b} = \{ (-\infty, a] \times (-\infty, b], (\infty, a] \times (b, +\infty), (a, +\infty) \times (-\infty, b], (a, +\infty) \times (b, +\infty) \}
\]

is a partition of \( \mathbb{R}^2 \). Furthermore, the smoothing approximation of the characteristic distribution of \( PS^k_{\tau_{a,b}} (\mathbb{R}^2) \) has the following form

\[
p^a(x, y) = \left( \begin{array}{c}
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}}
\end{array} \right) \left( \begin{array}{c}
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}}
\end{array} \right) \times \left( \begin{array}{c}
f_1(x, y) \\
f_2(x, y) \\
f_3(x, y) \\
f_4(x, y) \\
\end{array} \right)
\]

(32)

and for every \( f \in PS^k_{\tau_{a,b}} (\mathbb{R}^2) \) with representation \( R_f = (f_1, f_2, f_3, f_4) \), the smoothing approximation is given by

\[
S_a(R_f)(x, y) = \left( \begin{array}{c}
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(x-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}} \\
\frac{1}{1+e^{a(y-a)}}
\end{array} \right) \times \left( \begin{array}{c}
f_1(x, y) \\
f_2(x, y) \\
f_3(x, y) \\
f_4(x, y) \\
\end{array} \right)
\]

(33)
Theorem 4.5. Take $S_\alpha$ as in (33), then for every $f \in PS_k^\lambda(\mathbb{R}^2)$ with representation $R_f$, we have:

(a) $S_\alpha(R_f)$ is of class $C^k(\mathbb{R}^2)$,
(b) $S_\alpha(R_f) \to f$ as $\alpha \to +\infty$.

Proof. The proof is similar to the proof of Theorem 4.1.

Example 4. Let us consider the function $f(x,y) = |x| + |y|$ which is not differentiable if $x = 0$ or $y = 0$. It is clear that $f$ is an element of $PS_{70,0}^\infty(\mathbb{R}^2)$ and that $R_f = (x+y, x-y, -x+y, -x-y)$ is a representation of $f$. Therefore, the smoothing approximation of $f$, denoted $\hat{f}_\alpha$, is obtained by applying the smoothing operator of $PS_{70,0}^\infty(\mathbb{R})$, that is

$$\hat{f}_\alpha(x,y) = \begin{pmatrix}
\frac{1}{1+e^{-\alpha x}} & \frac{1}{1+e^{-\alpha y}} & t \\
\frac{1}{1+e^{-\alpha x}} & \frac{1}{1+e^{-\alpha y}} & -x+y \\
\frac{1}{1+e^{-\alpha x}} & \frac{1}{1+e^{-\alpha y}} & x+y \\
\frac{1}{1+e^{-\alpha x}} & \frac{1}{1+e^{-\alpha y}} & x-y
\end{pmatrix}$$

According to Theorem 4.5, $\hat{f}_\alpha(x,y)$ is of class $C^\infty(\mathbb{R}^2)$ and converge to $f(x,y)$ as $\alpha$ goes to infinity. The graphs of the function $f(x,y)$, the smoothing approximation $\hat{f}_\alpha(x,y)$, and the error of the smoothing approximation are shown in Figure 5.

Figure 5. (a) The graph of the function $f(x,y) = |x| + |y|$. (b) The graph of the smoothing approximation $\hat{f}_\alpha(x,y)$ for $\alpha = 1$. (c) The graph of the smoothing approximation $\hat{f}_\alpha(x,y)$ for $\alpha = 5$. (d) The error of the smoothing approximation for $\alpha = 5$. 
5. **Conclusion and perspectives.** Throughout this paper, we have studied the problem of constructing the smoothing approximations for piecewise smooth functions. Therefore, we first defined the space $PS_k^{	au}(\Omega)$ and formulated elements of that space as the expectation of a random variable. This formulation has two advantages, the first one is that smoothing all elements of $PS_k^{	au}(\Omega)$ is equivalent to smooth its characteristic distribution; the second advantage is that smoothing the characteristic distribution can be accomplished using the Boltzmann distribution (also known as softmax function).

We have studied the properties of our smoothing approximation and applied it to important examples. This study shows that our approach has important properties and is easy to apply; in fact, it suffices to apply the smoothing operator to the target P-S function to obtain a good smoothing approximation. The error between the original function and our smoothing approximation is very small and is controlled by a single parameter. Although only real-valued piecewise smooth functions are considered in this paper, it must be emphasized that this approach can be easily extended to vector-valued piecewise smooth functions.

For future works, we intend to use our smoothing approach to solve non-smooth optimization problems and penalty methods.

**REFERENCES**

[1] S. Amat, S. Busquier, A. Escudero and J. C. Trillo, Lagrange interpolation for continuous piecewise smooth functions, *Journal of computational and applied mathematics*, 221 (2008), 47–51.

[2] M. Aslam, S. Riemenschneider and L. Shen, Smoothing transforms for wavelet approximation of piecewise smooth functions, *IET Image Processing*, 2 (2008), 239–248.

[3] A. Bagirov, A. Al Nuaimat and N. Sultanova, Hyperbolic smoothing function method for minimax problems, *Optimization*, 62 (2013), 759–782.

[4] G. Z. Bai, S. Y. Liu, H. Wang, W. Zhi and Z. J. Yin, A novel method for smooth blending cylindrical surfaces whose axes are non-coplanar with intersecting line of two tapered surfaces, in *Applied Mechanics and Materials*, Trans. Tech. Publ., 687 (2014), 1470–1473.

[5] B. Belkhatir and A. Zidna, Construction of flexible blending parametric surfaces via curves, *Mathematics and Computers in Simulation*, 79 (2009), 3599–3608.

[6] D. P. Bertsekas, Non-differentiable optimization via approximation, in *Non-differentiable Optimization*, Springer, (1975), 1–25.

[7] M. Fukushima and L. Qi, Reformulation: Nonsmooth, Piecewise Smooth, Semismooth And Smoothing Methods, Vol. 22, Springer Science & Business Media, 2013.

[8] X. Han, A degree by degree recursive construction of hermite spline interpolants, *Journal of computational and Applied Mathematics*, 225 (2009), 113–123.

[9] E. Hartmann, Blending an implicit with a parametric surface, *Computer Aided Geometric Design*, 12 (1995), 825–835.

[10] F. Hashemi and S. Ketabchi, Numerical comparisons of smoothing functions for optimal correction of an infeasible system of absolute value equations, *Numerical Algebra, Control & Optimization*, 10 (2020), 13.

[11] M. Jiang, R. Shen, X. Xu and Z. Meng, Second-order smoothing objective penalty function for constrained optimization problems, *Numerical Functional Analysis and Optimization*, 35 (2014), 294–309.

[12] X. Jiang and Y. Zhang, A smoothing-type algorithm for absolute value equations, *Journal of Industrial & Management Optimization*, 9 (2013), 789.

[13] L. Kuntz and S. Scholtes, Qualitative aspects of the local approximation of a piecewise differentiable function, *Nonlinear Analysis: Theory, Methods & Applications*, 25 (1995), 107–215.

[14] S. Lian and Y. Duan, Smoothing of the lower-order exact penalty function for inequality constrained optimization, *Journal of Inequalities and Applications*, 2016 (2016), 1–12.

[15] Y. Lipman and D. Levin, Approximating piecewise-smooth functions, *IMA Journal of Numerical Analysis*, 30 (2010), 1159–1183.
[16] A. Mazroui, D. Sbibih and A. Tijini, A simple method for smoothing functions and compressing hermite data, Advances in Computational Mathematics, 23 (2005), 279–297.

[17] Z. Meng, C. Dang and X. Yang, On the smoothing of the square-root exact penalty function for inequality constrained optimization, Computational Optimization and Applications, 35 (2006), 375–398.

[18] J. Min, Z. Meng, G. Zhou and R. Shen, On the smoothing of the norm objective penalty function for two-cardinality sparse constrained optimization problems, Neurocomputing.

[19] C. T. Nguyen, B. Saheya, Y.-L. Chang and J.-S. Chen, Unified smoothing functions for absolute value equation associated with second-order cone, Applied Numerical Mathematics, 135 (2019), 206–227.

[20] P. Prandoni and M. Vetterli, Approximation and compression of piecewise smooth functions, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 357 (1999), 2573–2591.

[21] L. Qi and P. Tseng, On almost smooth functions and piecewise smooth functions, Nonlinear Analysis: Theory, Methods & Applications, 67 (2007), 773–794.

[22] D. Ralph and S. Scholtes, Sensitivity analysis of composite piecewise smooth equations, Mathematical Programming, 76 (1997), 593–612.

[23] R. T. Rockafellar, A property of piecewise smooth functions, Computational Optimization and Applications, 25 (2003), 247–250.

[24] B. Saheya, C.-H. Yu and J.-S. Chen, Numerical comparisons based on four smoothing functions for absolute value equation, Journal of Applied Mathematics and Computing, 56 (2018), 131–149.

[25] A. Sahiner, G. Kapusuz and N. Yilmaz, A new smoothing approach to exact penalty functions for inequality constrained optimization problems, Numerical Algebra, Control & Optimization, 6 (2016), 161.

[26] H. Wu, P. Zhang and G.-H. Lin, Smoothing approximations for some piecewise smooth functions, Journal of the Operations Research Society of China, 3 (2015), 317–329.

[27] A. E. Xavier, The hyperbolic smoothing clustering method, Pattern Recognition, 43 (2010), 731–737.

[28] N. Yilmaz and A. Sahiner, New smoothing approximations to piecewise smooth functions and applications, Numerical Functional Analysis and Optimization, 40 (2019), 513–534.

[29] I. Zang, A smoothing-out technique for minmax optimization, Mathematical Programming, 19 (1980), 61–77.

Received May 2021; Final revision July 2021; Early access August 2021.

E-mail address: e.amhraoui@edu.umi.ac.ma
E-mail address: t.masrour@ensam.umi.ac.ma