A Smoothing Method for Sparse Optimization over Polyhedral Sets

Mounir Haddou, Tangi Migot

▶ **To cite this version:**

Mounir Haddou, Tangi Migot. A Smoothing Method for Sparse Optimization over Polyhedral Sets. Hoai An Le Thi; Tao Pham Dinh; Ngoc Thanh Nguyen. Modelling, computation and optimization in information systems and management sciences, 359, Springer, pp.369-379, 2015, Advances in intelligent systems and computing, 978-3-319-18160-8. 10.1007/978-3-319-18161-5_31 . hal-01269975v1

HAL Id: hal-01269975

https://hal.science/hal-01269975v1

Submitted on 5 Feb 2016 (v1), last revised 18 Feb 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Smoothing Method for Sparse Optimization over Polyhedral Sets

M. Haddou* T. Migot†

Abstract

In this paper, we investigate a class of heuristics schemes to solve the NP-hard problem of minimizing ℓ_0-norm over a polyhedral set. A well-known approximation is to consider the convex problem of minimizing ℓ_1-norm. We are interested in finding improved results in cases where the problem in ℓ_1-norm does not provide an optimal solution to the ℓ_0-norm problem. We consider a relaxation technique using a family of smooth concave functions depending on a parameter. Some other relaxations have already been tried in the literature and the aim of this paper is to provide a more general context. We use an homotopy algorithm, starting from a solution to the problem in ℓ_1-norm and ending in a solution of the problem in ℓ_0-norm. We show convergence results, a kind of monotonicity of the solutions as well as error estimates leading to an exact penalization theorem. We also provide keys for implementing the algorithm and numerical simulations.

Mathematics Subject Classification. 90-08 and 65K05
Keywords : smoothing functions ; sparse optimization ; concave minimization ; l0-norm

Introduction

Consider a compact polyhedron F defined by linear inequalities,
$F = \{x \in \mathbb{R}^n |\ Ax \leq b\} \cap \mathbb{R}^n_+$ for some $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$, which we suppose non-empty and not reduced to a singleton. One should note that the hypothesis of considering polyhedron in the non-negative orthant is not restrictive, it is only assumed to simplify the presentation and to avoid the absolute value in the definition of the problem. We are interested in finding the sparsest point over this polyhedron, which is equivalent to minimize the ℓ_0-norm

$$\forall x \in \mathbb{R}^n,\ \|x\|_0 = \sum_{i=1}^{n} s(|x_i|), \text{ where for } t \in \mathbb{R},\ s(t) = \{0 \text{ if } t = 0\ ;\ 1 \text{ otherwise}\}\ . \tag{1}$$

*IRMAR-INSA Rennes ; mounir.haddou@insa-rennes.fr
†IRMAR-INSA Rennes ; tangi.migot@insa-rennes.fr

Note that the ℓ_0-norm is not a norm as it does not have the homogeneity property. We will study in this document the following

$$(P_0) \ \min_{x \in F} \|x\|_0 \ , \tag{2}$$

which is an NP-hard problem as shown in [17]. This problem has several applications and received a considerable interest recently. Some of the possible applications are signal and image modelling [3], machine learning [14] and compressed sensing [5],[6] and [7].

The problem (P_0) being difficult to solve a more simple approach is often used, which consists in solving the convex problem in ℓ_1-norm. The ℓ_1-norm is denoted by

$$\forall x \in \mathbb{R}^n, \ ||x||_1 = \sum_{i=1}^{n} |x_i| \tag{3}$$

and then we defined the convex problem by

$$(P_1) \ \min_{x \in F} ||x||_1 \ . \tag{4}$$

It can be seen as a convexification of (P_0), because $|x|$ is the convex envelope of $s(x)$ for $x \in [-1, 1]$. This approach has been extensively studied, for instance in [5], and has the benefits that it can be reformulated as a linear program. Also several criteria have been found which guarantee that solving (P_1) will also solve (P_0) under various assumptions, see [6] for some examples. A more sophisticated version of this convex formulation could be to consider a reweighted-ℓ_1 problem, [22]. For more details about applications in image and signal modelling as well as a review on related questions see reviews in [3] or [20].

Also formulation (4) does not solve all the time the initial problem. Consider for instance the following example in two dimension.

Example 0.1. *Given a matrix $A \in \mathbb{R}^{n \times n}$ and a vector $b \in \mathbb{R}^n$ such that*

$$A = \begin{pmatrix} -0.1 & -1 \\ -10 & -1 \end{pmatrix} \text{ and } b = \begin{pmatrix} -1 \\ -10 \end{pmatrix} \ . \tag{5}$$

Geometrical observation allows to conclude that the solution of problem (4) is $(\frac{10}{11}, \frac{10}{11})^T$, while solution of problem (2) are of the form $(0, 10 + \epsilon)^T$ and $(10 + \epsilon, 0)^T$ with $\epsilon \geq 0$.

Nonconvex optimization has been one of the main approach to tackle this problem. For instance in a related formulations of the problem [8] proposed a reformulation as a mathematical program with complementarity constraint. One can find thresholding algorithm in [21]. Moreover we can also use DC decomposition of the ℓ_0-norm, see [12]. We are interested here in nonconvex methods to improve the solution we get by solving (P_1) in the general case where this approach does not solve the initial problem. In this aim several concave relaxation of $||.||_0$ have been tried in the literature. In [9, 4], they

consider a concave minimization problem approximating the ℓ_0-norm with the ℓ_p-norm for $0 < p \leq 1$. In [18], they study a selection of concave minimization problem using concave functions such that $(t+r)^p$ with $r > 0$ and $0 < p < 1$, $-(t+r)^{-p}$ with $r > 0$ and $1 < p$, $\log(t+r)$ with $0 < r << 1$ or $1 - e^{-rt}$ with $r > 0$ and $p \in \mathbb{N}$. Finally, in [16] is presented a more general family including the gaussian family and proposes an homotopy method starting from the ℓ_2-norm solutions.

The purpose of this paper is to provide a theoretical context of relaxation methods with concave minimization problems in a more general way, that is using a general family of concave reformulation functions. We propose an homotopy method starting from a solution of the convex problem (4) and converging to a solution of the problem (2). We provide a complete analysis of this algorithm with convergence results and a study on error estimate leading to an exact penalty theorem. We run our algorithm on random examples showing the interest of this approach.

This document is organised as follows. Sect. 1 presents a general formulation of the relaxation methods using concave functions. Sect. 2 presents convergence and kind of monotonicity results leading to an homotopy method. Sect. 3 contains error estimates and an exact penalization theorem. Finally in Sect. 4 we give the algorithm with several remarks concerning its implementation and in Sect. 5 we present numerical results.

1 A smoothing method

We consider a family of smooth function in order to reformulate the ℓ_0-norm. This family has already been used in the different context of complementarity [11] and image restoration [2]. These functions are non-decreasing continuous smooth concave functions such that

$$\theta : \mathbb{R} \to]-\infty, 1[\text{ with } \theta(t) < 0 \text{ if } t < 0,\ \theta(0) = 0 \text{ and } \lim_{t \to +\infty} \theta(t) = 1 \ . \tag{6}$$

One way to build θ functions is to consider non-increasing probability density functions $f : \mathbb{R}_+ \to \mathbb{R}_+$ and then take the corresponding cumulative distribution function

$$\forall t \geq 0,\ \theta(t) = \int_0^t f(x)dx \text{ and } \forall t < 0,\ \theta(t) < 0 \ . \tag{7}$$

By definition of f we can verify that

$$\lim_{t \to +\infty} \theta(t) = \int_0^{+\infty} f(x)dx = 1 \text{ and } \theta(0) = \int_0^0 f(x)dx = 0 \ . \tag{8}$$

The non-increasing hypothesis on f gives the concavity of θ.

Examples of this family are $\theta^1(t) = t/(t+1)$ if $t \geq 0$ and $\theta^1(t) = t$ if $t < 0$, $\theta^2(t) = 1 - e^{-t}$ with $t \in \mathbb{R}$.

Then using a scaling technique similar to the perspective functions in convex analysis we define $\theta(t,r) := \theta\left(\frac{t}{r}\right)$ for $r > 0$ and we get

$$\theta(0,r) = 0 \ \forall r > 0 \text{ and } \lim_{r\to 0} \theta(t,r) = 1 \ \forall t > 0 \ . \tag{9}$$

For the previous examples of this family and $t \geq 0$ we have $\theta^1(t,r) = t/(t+r)$, $\theta^2(t,r) = 1 - e^{-t/r}$. The function $\theta^1(t,r)$ will be extensively used in this paper.

Throughout this paper we will consider the concave optimization problem for $r > 0$

$$(P_r) \ \min_{x\in F} \sum_{i=1}^{n} \theta(x_i, r) \ . \tag{10}$$

2 Convergence

In this section, we will show the link between problems (P_0), (P_1) and (P_r). We denote $S^*_{||.||_0}$ the set of solutions of (P_0), $S^*_{||.||_1}$ the set of solutions of (P_1) and S^*_r the set of solutions of (P_r). Theorem 2.1 gives convergence of (P_r) to (P_0) for r decreasing to 0.

Theorem 2.1 (Convergence to ℓ_0-norm)**.** *Every limit point of any sequence $\{x_r\}_r$, such that $x_r \in S^*_r$ and $r \downarrow 0$, is an optimal solution of (P_0).*

Proof. Given $\bar{x}$ the limit of the sequence $\{x_r\}_r$, up to a subsequence, and $x^* \in S^*_{||.||_0}$. Since F is a closed set one has $\bar{x} \in F$. Furthermore we have for any r in the corresponding subsequence

$$\sum_{i=1}^{n} \theta(\bar{x}_i, r) \leq \sum_{i=1}^{n} \theta(x^*_i, r) \ . \tag{11}$$

Moreover the definition of $\theta(.,r)$ functions, for $r > 0$ and $t \in \mathbb{R}^n$ give

$$\sum_{i=1}^{n} \lim_{r\downarrow 0} \theta(t_i, r) = ||t||_0 \ . \tag{12}$$

Replacing into (11) we get

$$||\bar{x}||_0 \leq ||x^*||_0 \ , \tag{13}$$

and thanks to the definition of $\bar{x}$

$$||\bar{x}||_0 = ||x^*||_0 \ . \tag{14}$$

□

We now give another convergence result from [9], which adds that the convergence appears in a finite number of iteration.

Proposition 2.1. *Given a non-empty polyhedron $F \subset \mathbb{R}^n_+$. Then there exists a $\bar{r}$ such that for all $r \leq \bar{r}$ a vertex of F is an optimal solution of (P_0) and $(P_{r\leq\bar{r}})$.*

Proof. (P_r) is a problem of minimizing a concave function over a polyhedron F. We can use Corollary 32.3.4 of [19], since there is no half-line in F such that $\theta(.,r)$ is unbounded below, so the infimum over F is attained and it is attained at one of the extremal point of F.
Given that there is a finite number of extremal point, for a given extremal point x' we can find an infinite non-increasing sequence $R = (r_0, r_1, r_2, ...)$ with $\forall j \;\; r_j \geq 0$ where x' is solution of (P_R). Moreover the objective function of (P_r) is non-increasing and bounded below by the infimum of ℓ_0-norm, so

$$\sum_{i=1}^{n} \theta(x'_i, r_j) = \min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r_j) \leq \inf_{x \in F} ||x||_0. \tag{15}$$

Going through the limit in R for $j \to \infty$ and as the concave function is continuous and $x' \in F$, we have the results. □

The next theorem shows for r sufficiently large that solutions of (P_r) are the same than solutions of (P_1).

Theorem 2.2 (Convergence to ℓ_1-norm)**.** *Every limit point of any sequence* $\{x_r\}_r$*, such that* $x_r \in S_r^*$ *for* $r \uparrow \infty$*, is an optimal solution of* (P_1)*.*

Proof. As $r > 0$, we can use a scaling technique for $S_r^{*(2)} = \arg\min_{x \in F} \sum_{i=1}^{n} r\theta(x_i, r)$

$$\min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r) \quad \Longleftrightarrow \quad \min_{x \in F} \sum_{i=1}^{n} r\theta(x_i, r) \tag{16}$$

$$S_r^* \quad = \quad S_r^{*(2)}. \tag{17}$$

Given $x^r \in S_r^{*(2)}$ and $\bar{x} \in S^*_{||.||_1}$. We use the first order Taylor's theorem for $\theta(t)$ in 0,

$$\theta(t) = t\theta'(0) + g(t), \text{ where } \lim_{t \to 0} \frac{g(t)}{t} = 0. \tag{18}$$

Functions θ are concave, so we have $\theta'(0) > 0$. By definition of $\bar{x}$ and using (18)

we have

$$\sum_{i=1}^{n} r\theta(x_i^r, r) \leq \sum_{i=1}^{n} r\theta(\bar{x}_i, r) \tag{19}$$

$$\sum_{i=1}^{n} x_i^r\theta'(0) + rg(\frac{x_i^r}{r}) \leq \sum_{i=1}^{n} x_i^r\theta'(0) + rg(\frac{\bar{x}_i}{r}) \tag{20}$$

$$\sum_{i=1}^{n} x_i^r - \sum_{i=1}^{n} \bar{x}_i \leq \frac{r}{\theta'(0)} \sum_{i=1}^{n} g(\frac{\bar{x}_i}{r}) - \frac{r}{\theta'(0)} \sum_{i=1}^{n} g(\frac{x_i^r}{r}) \tag{21}$$

$$\leq \frac{1}{\theta'(0)} \left| \sum_{i=1}^{n} \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \bar{x}_i \right| + \frac{1}{\theta'(0)} \left| \sum_{i=1}^{n} \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} x_i^r \right| \tag{22}$$

$$\leq \frac{1}{\theta'(0)} \left(\sum_{i=1}^{n} \left| \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \right| \right) \left(\sum_{i=1}^{n} \bar{x}_i \right) \tag{23}$$

$$+ \frac{1}{\theta'(0)} \left(\sum_{i=1}^{n} \left| \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} \right| \right) \left(\sum_{i=1}^{n} x_i^r \right) \tag{24}$$

$$\sum_{i=1}^{n} x_i^r \leq \left(\sum_{i=1}^{n} \bar{x}_i \right) \frac{1 + \frac{1}{\theta'(0)} \left(\sum_{i=1}^{n} \left| \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} \right| \right)}{1 - \frac{1}{\theta'(0)} \left(\sum_{i=1}^{n} \left| \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} \right| \right)}. \tag{25}$$

We have

$$\lim_{r \to +\infty} \frac{\bar{x}}{r} = 0, \tag{26}$$

and now

$$\sum_{i=1}^{n} \theta(x_i^r, r) \leq \sum_{i=1}^{n} \theta(\bar{x}_i, r) \tag{27}$$

$$\lim_{r \to +\infty} \sum_{i=1}^{n} \theta(x_i^r, r) \leq \lim_{r \to +\infty} \sum_{i=1}^{n} \theta(\bar{x}_i, r) \tag{28}$$

$$\leq 0 \tag{29}$$

$$\lim_{r \to +\infty} \sum_{i=1}^{n} \theta(x_i^r, r) = 0. \tag{30}$$

As for $r > 0 :\ \theta(x_i, r) \in [0, 1[$ and $\theta^{-1}(0, r) = 0$, we have

$$\lim_{r \to +\infty} \frac{x_i^r}{r} = 0 \ \forall i\ . \tag{31}$$

Using (26) and (31) it becomes

$$\lim_{r \to +\infty} \frac{\bar{x}_i}{r} = 0 \implies \lim_{r \to +\infty} \frac{g(\frac{\bar{x}_i}{r})}{\frac{\bar{x}_i}{r}} = 0 \tag{32}$$

$$\lim_{r \to +\infty} \frac{x_i^r}{r} = 0 \implies \lim_{r \to +\infty} \frac{g(\frac{x_i^r}{r})}{\frac{x_i^r}{r}} = 0, \tag{33}$$

and then going to the limit

$$\lim_{r\to+\infty} \sum_{i=1}^{n} x_i^r \leq \sum_{i=1}^{n} \bar{x}_i. \tag{34}$$

So, with (34) and by definition of $\bar{x}$ we have the equality and then the result. □

The next theorem gives a monotonicity result, which enlightens the relation between the three problems (P_0), (P_1) and (P_r). By monotonicity, we mean that for a given feasible point we want a relation of monotony in r for the objective function of (P_1), (P_r) and (P_0). As the components of the ℓ_0-norm and the $\theta_r(t)$ are in $[0,1[$ it is necessary to put the components of ℓ_1-norm in a similar box, which explains the change of variable in the theorem.

Remark 2.1. *In the following theorem we use the hypothesis that θ functions are convex in r. This is not so restrictive as we think several functions verify it. If we take the three examples of θ functions given in the introduction, θ^1 and $\theta^{log} := \log(1+x)/\log(1+x+r)$ are convex in r but not θ^2.*

Theorem 2.3 (Monotonicity of solutions)**.** *Given $x \in F$, let $y = x/(||x||_\infty + \epsilon)$ where $\epsilon > 0$, so $y \in [0,1[^n$. Set a function $\Psi(t,r) : [0,1[\to [0,1[$ as*

$$\Psi(t,r) = \frac{\theta(t,r)}{\theta(1,r)}, \tag{35}$$

where $\theta(t,r)$ is the smooth function described in the introduction, which we will consider here as convex in r. For r and $\bar{r}$ such that $0 < \bar{r} < r < +\infty$, then one has

$$||y||_1 \leq \sum_{i=1}^{n} \Psi(y_i, r) \leq \sum_{i=1}^{n} \Psi_(y_i, \bar{r}) \leq ||y||_0 \ . \tag{36}$$

Proof. Since functions θ are concave and $\theta(0) = 0$ one has subadditivity of this functions, then

$$\theta(y_i, r) \geq y_i \ \theta(1, r) \ . \tag{37}$$

Therefore

$$\sum_{i=1}^{n} \Psi(y_i, r) - \ ||y||_1 \quad = \quad \sum_{i=1}^{n} (\frac{\theta(y_i, r)}{\theta(1, r)} - \ y_i) \tag{38}$$

$$\geq \quad 0, \tag{39}$$

which leads to the first inequality

$$||y||_1 \leq \sum_{i=1}^{n} \Psi(y_i, r) \ . \tag{40}$$

We continue with the second inequality showing that $\Psi(y,r)$ functions are non-increasing in r, i.e

$$\sum_{i=1}^{n} \Psi(y_i, r) \leq \sum_{i=1}^{n} \Psi_(y_i, \bar{r}). \tag{41}$$

The functions $\Psi(y,r)$ is non-increasing in r if its derivative with respect to r

$$\frac{\partial}{\partial r}\Psi(y,r) = \frac{(\frac{\partial}{\partial r}\theta(y,r))\theta(1,r) - (\frac{\partial}{\partial r}\theta(1,r))\theta(y,r)}{\theta(1,r)^2} \tag{42}$$

is negative. Since $\theta(y,r)$ is an non-decreasing function in y we have

$$\frac{\theta(y,r)}{\theta(1,r)} < 1 \tag{43}$$

and

$$\frac{\partial}{\partial r}\theta(y,r) = -\frac{1}{r^2}\frac{\partial}{\partial y}\theta(y,r) < 0 \ . \tag{44}$$

So $\theta(y,r)$ is non-increasing function in r. Using convexity of $\theta(y,r)$ in r it follows

$$\frac{\frac{\partial}{\partial r}\theta(y,r)}{\frac{\partial}{\partial r}\theta(1,r)} = \frac{\frac{\partial}{\partial r}\theta(1,r/y)}{\frac{\partial}{\partial r}\theta(1,r)} > 1 \ . \tag{45}$$

Then in (42) the derivative with respect to r is negative and we have (41). Finally, since $\theta(y,r)$ is non-decreasing in y and $y \in [0,1[^n$ one has

$$||y||_0 - \sum_{i=1}^{n}\Psi(y_i,\bar{r}) = \sum_{i=1;y_i\neq 0}^{n} 1 - \frac{\theta(y_i,\bar{r})}{\theta(1,\bar{r})} \ \geq \ 0 \ , \tag{46}$$

which gives the last inequality and complete the theorem. □

Remark 2.2. *Both choice of scaling parameter in Theorem 2.2 and Theorem 2.3 are linked. In the former, we set that* $\lim_{r\to+\infty} r\sum_{i=1}^{n}\theta(x_i,r) = \sum_{i=1}^{n} x_i$, *so evaluating in one dimension and* $x = 1$ *we have* $\lim_{r\to+\infty} r\theta(1,r) = 1$ *and then we see that* r *and* $1/\theta(1,r)$ *have the same behaviour for* r *sufficiently large.*

All this results lead us to the general behaviour of the method. First, we start from one solution of (P_1) then by decreasing parameter r the solution of (P_r) becomes closer to a solution of (P_0).

Another approach would be to define a new problem which selects one solution of the possibly many optimal solution of (P_0). We consider the following problem which is a selective version of (P_r)

$$(P_{r-sel}) \ \min_{x\in F}\sum_{i=1}^{n}\theta(x_i,r) + \sum_{i=1}^{n} r^{\frac{i}{2n+1}}x_i \ . \tag{47}$$

We will use a lexicographic norm and we note

$$||y||_{lex} < ||x||_{lex} \Longleftrightarrow \exists i \in \{1,...,n\}, \ y_i < x_i \text{ and } \forall \ 1 \leq j < i, \ y_j = x_j \ . \tag{48}$$

In the next theorem we want to choose the solution of (P_0) which has the smallest lexicographic norm. From the previous equation it is clear that this optimal solution is unique.

Theorem 2.4 (Convergence of the selective concave problem)**.** *We use functions θ such that $\theta \geq \theta^1$. Given $\{x^r\}_r$ the sequence of solutions of (P_{r-sel}) for $r \in R = \{r_0,\ r_1,\ ...\}$ with R a non-increasing sequence such that $r_j > 0$ and $\bar{x}$ the limit point of this sequence. Then, $\bar{x}$ is the unique solution of $S^*_{||.||_0}$ such that $\forall y \in S^*_{||.||_0},\ ||\bar{x}||_{lex} \leq ||y||_{lex}$.*

Proof. Given x^* an optimal solution of (P_0) such that $\forall y \in S^*_{||.||_0},\ ||x^*||_{lex} \leq ||y||_{lex}$ and $\bar{x}$ the limit of a sequence of $\{x^r\}_r$ solution of (P_{r-sel}). So, there exists a $\bar{r}$ such that for every $r < \bar{r}$ we have $\bar{x}$ solution of (P_{r-sel}) and

$$\sum_{i=1}^{n} \theta(\bar{x}_i, r) + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^{n} \theta(x^*_i, r) + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} x^*_i \ . \tag{49}$$

In the same way as in Theorem 2.1, going to the limit for $r \downarrow 0$ we have

$$||\bar{x}||_0 \leq ||x^*||_0 \ , \tag{50}$$

which is an equality by definition of x^* and prove the first part of the theorem. Now we need to verify the selection of the solution. Using that $\theta(x, r) \leq 1$ in (49) one has

$$\sum_{i=1}^{n} (\theta(\bar{x}_i, r) - 1) + k + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq k + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} x^*_i \tag{51}$$

$$\sum_{i=1}^{n} (\theta(\bar{x}_i, r) - 1) + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^{n} r^{\frac{i}{2n+1}} x^*_i \ . \tag{52}$$

Consider $\theta \geq \theta^1$, r sufficiently small ($\min_{\{i|x_i \neq 0\}} \bar{x}_i \geq \sqrt{r}$) and using that $||\bar{x}||_0 = k$ we have

$$-k \frac{r}{r + \sqrt{r}} + \sum_{i=1}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=1}^{n} r^{\frac{i}{2n+1}} x^*_i \ . \tag{53}$$

Diving by $r^{\frac{1}{2n+1}}$, we get

$$-k \frac{r}{r^{\frac{1}{2n+1}}(r + \sqrt{r})} + x_1 + \sum_{i=2}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq x^*_1 + \sum_{i=2}^{n} r^{\frac{i}{2n+1}} x^*_i \ . \tag{54}$$

Therefore, going to the limit for $r \downarrow 0$ one has

$$\bar{x_1} \leq x^*_1 \tag{55}$$

which is an equality by hypothesis on x^* being the smallest $||.||_{lex}$ solution of (P_0). So, as $\bar{x_1} = x^*_1$ in (53) one has

$$-k \frac{r}{r + \sqrt{r}} + \sum_{i=2}^{n} r^{\frac{i}{2n+1}} \bar{x}_i \leq \sum_{i=2}^{n} r^{\frac{i}{2n+1}} x^*_i \ . \tag{56}$$

By induction we get $\bar{x}_i = x_i^*,\ \forall i \in \{1, ..., n\}$ and so $\bar{x} = x^*$, because we have

$$\forall j \in \{1, ..., n\},\ \lim_{r \to 0} \frac{r}{r^{\frac{j}{2n+1}}(r + \sqrt{r})} = 0\ . \tag{57}$$

Finally we have the results as $\bar{x}$ is the optimal solution which has the smallest lexicographic norm. □

Remark 2.3. *If we try to get an equivalent result as in Theorem 2.2 for this selection problem, it is clear that for r sufficiently large we will solve the ℓ_1-norm problem but with a reversed lexicographical order than the one we are looking for, i.e for a non-decreasing sequence of r_j*

$$\bar{x} = \lim_{j \to \infty} \{x^{r_j}\}_{r_j} \text{ with } x^{r_j} \in S^*_{r_j - sel} \Longrightarrow \bar{x} \in S^*_{||.||_1} \text{ and } \bar{x} = arg \max_{y \in S^*_{||.||_1}} ||y||_{lex}. \tag{58}$$

This will definitely prevent us of any kind of monotonicity result such as Theorem 2.3. So, unless $S^*_{||.||_0}$ admits only one solution, the initial point as a solution of (P_1) has no chance of being a good initial point. This argument and the fact that this problem looks numerically not advisable lead us not to follow the study of this selective problem.

3 Error estimate

In this section we focus on what happen when r becomes small. We denote $\text{card}(I)$ the number of elements in a set I. Note that the following results are given for functions $\theta \geq \theta^1$ with $\theta^1(t, r) = t/(t + r)$ for $t, r \in \mathbb{R}_+$.

Lemma 3.1. *Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = ||x^*||_0 < n$ be the optimal value of problem (P_0) and $I(x, r) = \{i | x_i \geq\ kr\}$. Then one has*

$$x^r \in \arg\min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r) \Rightarrow card(I(x^r, r)) \leq k\ . \tag{59}$$

Proof. We use a proof by contradiction. Consider that $\text{card}(I(x^r, r)) \geq k + 1$ and we have $x^r \in \arg\min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r)$, then

$$\sum_{i=1}^{n} \theta(x_i^r, r) \geq (k + 1)\theta(kr, r) \geq (k + 1)\theta^1(kr, r) = (k + 1)\frac{kr}{kr + r} = k\ , \tag{60}$$

which is a contradiction with the definition of x^r. □

This lemma gives us a theoretical stopping criterion for the decrease of r, as for $r < \bar{r} = \min_{x_i^r \neq 0} x_i^r/k$, x^r becomes an optimal solution. In the following lemma we look at the consequences in the evaluation of θ.

Lemma 3.2. *Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = ||x^*||_0 < n$ be the optimal value of problem (P_0) and*

$$\bar{r} = \min_{x_i^r \neq 0} x_i^r / k \ .$$

Then one has

$$r \leq \bar{r} \Longleftrightarrow \theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \ . \tag{61}$$

Proof. By using the expression of θ^1 we have

$$\theta^1(\min_{x_i^r \neq 0} x_i^r, r) = \frac{\min_{x_i^r \neq 0} x_i^r}{\min_{x_i^r \neq 0} x_i^r + r} \geq \frac{k}{k+1} \tag{62}$$

$$\Longleftrightarrow \quad \min_{x_i^r \neq 0} x_i^r (k+1) \geq k(\min_{x_i^r \neq 0} x_i^r + r) \tag{63}$$

$$\Longleftrightarrow \quad \min_{x_i^r \neq 0} x_i^r \geq kr \tag{64}$$

$$\Longleftrightarrow \quad \bar{r} = \frac{\min_{x_i^r \neq 0} x_i^r}{k} \geq r \ , \tag{65}$$

and so the results. □

Both previous lemma lead us to the following theorem, which is an exact penalization result for our method.

Theorem 3.1 (Exact Penalization Theorem)**.** *Consider θ functions where $\theta \geq \theta^1$. Let $\mathbb{N} \ni k = ||x^*||_0 < n$ be the optimal value of problem (P_0) and $x^r \in S_r^*$. Then one has*

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \Longrightarrow x^r \in S^*_{||.||_0} \ . \tag{66}$$

Proof. By Lemma 3.2 and with $\bar{r} = \min_{x_i^r \neq 0} x_i^r / k$ one has

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{k}{k+1} \Longleftrightarrow r \leq \bar{r} \ . \tag{67}$$

Then by Lemma 3.1 and using $x_r \in S_r^*$ we have

$$x^r \in \arg\min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r) \Rightarrow \mathrm{card}(I(x_r, r)) \leq k \ . \tag{68}$$

Finally, using $r \leq \bar{r}$ and that k is the optimal value of problem in ℓ_0-norm we have the result. □

We use in the previous result the minimum non-zero component of x^r, which is logical as we expect that for r sufficiently small the sequence of $\{\min_{x_i^r \neq 0} x_i^r\}_r$ should be increasing. The following lemma gives us a clue on this behaviour.

Lemma 3.3. *Consider θ functions where $\theta \geq \theta^1$. Let $x^* \in S^*_{||\cdot||_0}$, $||x^*||_0 = k$ and*

$$r^* = \frac{1}{k} \min_{x_i^* \neq 0} x_i^* \ .$$

Then one has

$$\forall r \leq r^*, x^r \in S_r^* \Longrightarrow \min_{x_i^r \neq 0} \ x_i^r \leq \min_{x_i^* \neq 0} \ x_i^* \ . \tag{69}$$

Proof. Suppose that $\min_{x_i \neq 0} \ x_i > \min_{x_i^* \neq 0} \ x_i^*$. Since $x^r \in S_r^*$ we have

$$\begin{aligned} \sum_{i=1}^{n} \theta(x^r, r) &\geq \sum_{i=1}^{n} \theta(x^r, r^*) && (70)\\ &> (k+1)\theta(\min_{x_i^* \neq 0} \ x_i^*, r^*) && (71)\\ &> (k+1)\frac{kr^*}{kr^* + r^*} && (72)\\ &= k \ , && (73) \end{aligned}$$

which is in contradiction with the definition of x^r. □

4 Algorithm

The previous results allow us to build a generic algorithm

$$[\text{Thetal0}] \quad \begin{cases} \{r^k\}_{k \in \mathbb{N}}, \ r^0 > 0 \text{ and } \lim_{k \to +\infty} r^k = 0 \\ \text{find } x^k : \ x^k \in \arg\min_{x \in F} \sum_{i=1}^{n} \theta(x_i, r^k) \end{cases} . \tag{74}$$

Now, several questions remain to be answered such as initialization, choice of the sequence $\{r^k\}$ and the method used to solve the concave minimization problems. In Sect. 3 we have shown an exact penalization result, which will help us building a stopping criterion. We make a few remarks about these questions. Note that interesting related remarks can be found in [16].

Remark 4.1 (On the behaviour of θ functions)**.** *These concave functions are acting as step function for r sufficiently small. That is one has the following behaviour*

$$\theta(t, r) \simeq \begin{cases} 1 \ \textit{if } t >> r \\ 0 \ \textit{if } t << r \end{cases} \ , \tag{75}$$

which gives us a strategy to update r. Let x^k be our current iterate and r^k the corresponding parameter. We divide our iterate into two sets, those with indices in $I = \{i \mid x_i^k \geq r^k\}$ and the others with indices in $\bar{I} = \{i \mid x_i^k < r^k\}$. We can see I as the set of indices of the "non-zero" components and $\bar{I}$ as the set of indices of the "zero" components of x^k. So we will choose r^{k+1} around $\max_{i \in \bar{I}} x_i^k$ to ask whether or not it belongs to zeros and we repeat this operation until r is

sufficiently small to consider $\bar{I}$ the set of effective zeros. Also this is a general behaviour, to be sure to have decrease of r one should add a fixed parameter of minimum decrease.

Remark 4.2 (Initialization)**.** *It is the main purpose of our method to start with the solution x^0 of the problem (P_1), which is a convex problem. So, we need to find the r^0 related to x^0. A natural, but non-trivial, way of doing this would be to find the parameter which minimizes the following problem*

$$\min_{r>0} || \sum_{i=1}^{N} \theta(x_i^0, r) - ||x^0||_1 \ ||_2^2 \ . \tag{76}$$

A simpler idea is to be inspired from last remark and put r^0 as a value which is just beyond the top value of x_i^0.

Remark 4.3 (Stopping criterion)**.** *It has been shown, in Sect. 3, an exact penalization theorem using the quantity $k/(k+1)$, which depends on the solution we are looking for. Numerically, we can make more iterations but being sure to satisfy this criterion using the fact that $||x^0||_0 \geq k$, which gives us the following criterion*

$$\theta(\min_{x_i^r \neq 0} x_i^r, r) \geq \frac{||x^0||_0}{||x^0||_0 + 1} \geq \frac{k}{k+1} \ . \tag{77}$$

Remark 4.4 (Algorithm for concave minimization)**.** *In the same way as in [9] and [18] we will use a successive linearization algorithm (SLA) algorithm to solve the concave minimization problem at each iteration in r. This algorithm is a finitely timestep Franck & Wolf algorithm, [13].*

Proposition 4.1 (SLA algorithm for concave minimization)**.** *Given ϵ sufficiently small and r^k. We know x^k and we find x^{k+1} as a solution of the linear problem*

$$\min_{x \in F} \ x^t \nabla_x \theta(x^k, r^k) \ , \tag{78}$$

with x^0 a solution of the problem (P_1). We stop when

$$x^{k+1} \in F \ \text{and} \ (x^{k+1} - x^k)^t \nabla_x \theta(x^k, r^k) \leq \epsilon \ . \tag{79}$$

This algorithm generates a finite sequence with strictly decreasing objective function values.

Proof. see [[13], Theorem 4.2]. □

We note that this algorithm didn't provide necessarily a global optimum as it ends in a local solution, so we don't expect global solutions in our algorithm. Also when considering the objective function of this linear program the gradient of functions θ tends to be very large as $\theta'_r(t) \approx O(1/r)$, so it can be numerically efficient to add a scaling parameter of order r.

5 Numerical Simulations

Thanks to the previous sections we have keys for an algorithm. We will show now some numerical results. These simulations have been done using MATLAB language, [15], with the linear programming solver GUROBI, [10].

The precision in our simulations is $\epsilon = 10^{-8}$. We generate various polyhedron $F = \{x \in \mathbb{R}^n |\ b \in \mathbb{R}^m,\ Ax \leq b\} \cap \mathbb{R}^n_+$ with $m < n$. In the same way as in [9] we choose $n = (500, 750, 1000)$ and in each case $m = (40\%, 60\%, 80\%)$. For each pair (n,m) we choose randomly one hundred problems. We take a random matrix A of size $m \times n$ and a default sparse solution x_{init} with 10% of non-zero components. We get b by calculating the matrix-vector product $b = Ax_{init}$. In the end, we will compare the sparsity of the solution from Thetal0-algorithm using θ^1 ($\#\theta^1$), the default sparse solution ($\#\ell_0$) and the initial iterate ($\#\ell_1$). We get the initial iterate as a solution of problem (P_1). The item # indicates the number of non-zero components in a vector.

Table 1: Numerical results with random $F = \left\{x \in \mathbb{R}^n_+ |\ b \in \mathbb{R}^m,\ Ax \leq b\right\}$, dimensions of the problem are first 2 columns. Compare a default sparse solution with 10% of non-zero components, $\#l_0$, the initial iterate solution of (P_1), $\#l_1$, and the solution by θ-algorithm with function θ^1, $\#\theta^1$. The item # indicates the number of non-zeros.

n	m	$\#\ell_0 \geq \#\theta^1$	$\#\ell_1 \leq \#\ell_0$	$\#\theta^1 < \#\ell_1$
1000	800	100	100	0
1000	600	100	98	2
1000	400	50	1	99
750	600	100	100	0
750	450	100	98	2
750	300	54	0	100
500	400	100	100	0
500	300	100	94	6
500	200	63	0	100

Results are sum up in Table 1. The first two columns give the dimensions of the problems, Column 3 gives the number of problems where the solution of θ algorithm has at least the same sparsity as the default sparse solution. Column 4 in the same vein compare the sparsity of the solution in ℓ_1-norm with the default sparse solution. Column 5 gives the number of problems where the solution by θ-algorithm improves strictly the solution by ℓ_1-norm.

This results validate our algorithm, as in the majority of the cases it manages to find at least an equivalent solution to the default sparse solution. One may notice that in many cases the ℓ_1-norm minimization solution solves the problem in ℓ_0-norm, which is not surprising according to [6].

In Figure 1, we show the behaviour of the minimum non-zero component of the current iterate along the iterations in r for one example. We can see the increasing behaviour that is the general behaviour expected in the Remark 4.3.

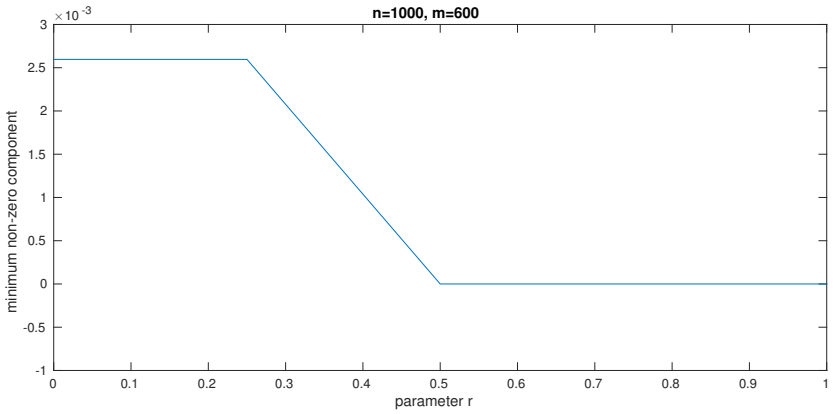

Figure 1: Evolution of the minimum non-zero component of x_r in function of the parameter $r \downarrow 0$.

6 Conclusion and Outlook

We proposed a class of heuristics schemes to solve the NP-hard problem of minimizing the ℓ_0-norm. Our methods require to find a sequence of solution from concave minimization problem, which we solved with a successive linearization algorithm. This methods have the benefit that they can only improve the solution we get by solving the ℓ_1-norm problem. We gave convergence results, an exact penalization theorem and keys to implement the methods. To confirm the validity of this algorithm we gave numerical results from randomly generated problems.

Further studies can investigate the special case where the ℓ_1-norm solves the ℓ_0-norm problem, to find an improved stopping condition. Thanks to several studies, for instance [6], we have criteria which can help us identifying the cases where the solution we get by solving (P_1) is an optimal solution of (P_0). We can wonder if there exists a better sufficient condition than the one presented here in the case where $x^r \in S^*_{||.||_1} \cap S^*_r$

We can also study a very similar problem which is the one of minimizing ℓ_0-norm with noise, see for instance [7] or [1], that is

$$(P_{0,\delta}) \ \min ||x||_0 \text{ s.t. } Ax \leq b + \delta. \tag{80}$$

As a first step in this direction we run our heuristic schemes on some perturbed problems. We generate polyhedron in a similar way as in the previous section with noise in $b = \ Ax_{init} + \vartheta$, where ϑ follows $\mathcal{N}(0, \sigma^2 I_{n\times n})$. We build a signal to noise ratio (SNR) for several values of σ^2 from 0.5 to 0,

$$SNR = 20 \log(\frac{||x^*||_2}{||x^* - x_b||_2}) \ , \tag{81}$$

where x^* and x_b are generated by our algorithm, the former comes from the

problem without noise and the later from the perturbed problem. We choose dimensions $n = 500$ and $m = 200$. Then for one hundred randomly selected problems we compute the mean of the SNR. Results in Figure 2 show very

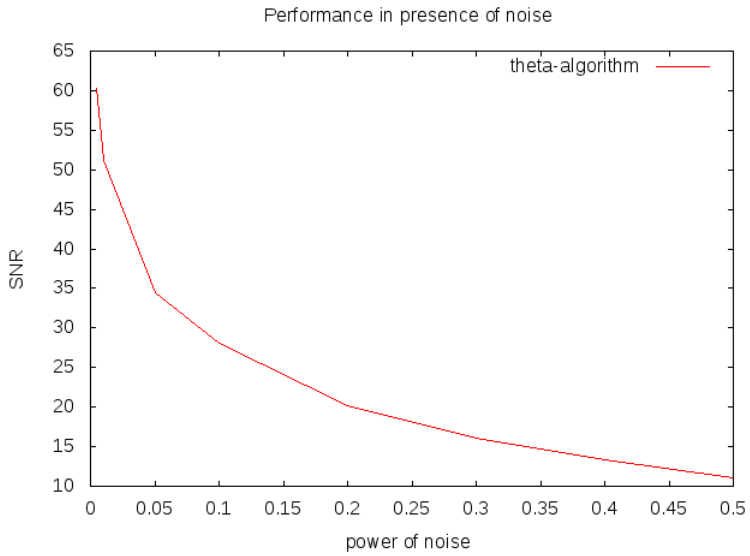

Figure 2: Performance of θ-algorithm in presence of noise, using function $\theta^1(t,r) = t/(t+r)$. $n = 500$, $m = 200$. Mean of SNR for 100 random problems in function of σ^2.

logical behaviour as more noise is present more informations are lost. Further work could compare these results with existing methods and shows theoretical study, which could help building an improved algorithm.

References

[1] Massoud Babaie-Zadeh and Christian Jutten. On the stable recovery of the sparsest overcomplete representations in presence of noise. *Signal Processing, IEEE Transactions on*, 58(10):5396–5400, 2010.

[2] Maïtine Bergounioux and Mounir Haddou. A new relaxation method for a discrete image restoration problem. *Journal of Convex Analysis*, 17(3):421–435, 2010.

[3] Alfred M Bruckstein, David L Donoho, and Michael Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM review*, 51(1):34–81, 2009.

[4] Rick Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *Signal Processing Letters, IEEE*, 14(10):707–710, 2007.

[5] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[6] David L Donoho. Neighborly polytopes and sparse solutions of underdetermined linear equations. Technical report, Department of Statistics, Stanford University, 2004.

[7] David L Donoho, Michael Elad, and Vladimir N Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *Information Theory, IEEE Transactions on*, 52(1):6–18, 2006.

[8] Mingbin Feng, John E Mitchell, Jong-Shi Pang, Xin Shen, and Andreas Wächter. Complementarity formulations of l0-norm optimization problems. Technical report, Technical report, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY, 2013.

[9] G.M. Fung and O.L. Mangasarian. Equivalence of minimal ℓ_0 and ℓ_p norm solutions of linear equalities, inequalities and linear programs for sufficiently small p. *Journal of Optimization Theory and Applications*, 151(1):1–10, 2011.

[10] Inc. Gurobi Optimization. Gurobi optimizer reference manual, 2015.

[11] M. Haddou and P. Maheux. Smoothing methods for nonlinear complementarity problems. *Journal of Optimization Theory and Applications*, 160(3):711–729, 2014.

[12] Hoai An Le Thi, T Pham Dinh, Hoai Minh Le, and Xuan Thanh Vo. Dc approximation approaches for sparse optimization. *European Journal of Operational Research*, 244(1):26–46, 2015.

[13] O. L. Mangasarian. Machine learning via polyhedral concave minimization. In *Applied Mathematics and Parallel Computing*, pages 175–188. Springer, 1996.

[14] OL Mangasarian. Machine learning via polyhedral concave minimization. In *Applied Mathematics and Parallel Computing*, pages 175–188. Springer, 1996.

[15] MATLAB. *version 8.4.0.150421 (R2014b)*. The MathWorks Inc., Natick, Massachusetts, 2014.

[16] Hosein Mohimani, Massoud Babaie-Zadeh, and Christian Jutten. A fast approach for overcomplete sparse decomposition based on smoothed norm. *Signal Processing, IEEE Transactions on*, 57(1):289–301, 2009.

[17] Balas Kausik Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.

[18] Francesco Rinaldi. *Mathematical programming methods for minimizing the zero-norm over polyhedral sets*. PhD thesis, PhD thesis, Sapienza, University of Rome, 2009.

[19] R Tyrrell Rockafellar. *Convex analysis*. Princeton university press, 1970.

[20] Joel Tropp, Stephen J Wright, et al. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6):948–958, 2010.

[21] Sergey Voronin and Rick Chartrand. A new generalized thresholding algorithm for inverse problems with sparsity constraints. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 1636–1640. IEEE, 2013.

[22] Yun-Bin Zhao and Michal Kočvara. A new computational method for the sparsest solutions to systems of linear equations. *SIAM Journal on Optimization*, 25(2):1110–1134, 2015.