

A Novel Approach for Solving Convex Problems with Cardinality Constraints

Goran Banjac and Paul J. Goulart

Abstract

In this paper we consider the problem of minimizing a convex differentiable function subject to sparsity constraints. Such constraints are non-convex and the resulting optimization problem is known to be hard to solve. We propose a novel generalization of this problem and demonstrate that it is equivalent to the original sparsity-constrained problem if a certain weighting term is sufficiently large. We use the proximal gradient method to solve our generalized problem, and show that under certain regularity assumptions on the objective function the algorithm converges to a local minimum. We further propose an updating heuristic for the weighting parameter, ensuring that the solution produced is locally optimal for the original sparsity constrained problem. Numerical results show that our algorithm outperforms other algorithms proposed in the literature.

I. INTRODUCTION

We consider the following optimization problem

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x \in \mathcal{B} \\ & && \text{card}(x) \leq k, \end{aligned} \tag{P}$$

where $f : \mathbb{R}^n \rightarrow \tilde{\mathbb{R}}$ is a differentiable convex function, $\mathcal{B} = \{x \in \mathbb{R}^n \mid l \leq x_i \leq u\}$, with $-l, u \in \mathbb{R}_+ \cup \{\infty\}$, and $k \in \mathbb{N}$ is a given positive integer. Such problems arise in a number of engineering applications, such as compression [1], sparse regression [2], image processing [3], and sparse controller design [4].

The cardinality constraint in problem \mathcal{P} is used to impose sparsity of some degree on the optimal solution. However, such a constraint is non-convex and the resulting optimization problem is known to be \mathcal{NP} -hard in general. A widely used heuristic approach for enforcing sparsity of the solution is to add the ℓ_1 regularization term in the objective [5], *i.e.* to solve the following convex relaxation of problem \mathcal{P}

$$\begin{aligned} & \text{minimize} && f(x) + \gamma \|x\|_1 \\ & \text{subject to} && x \in \mathcal{B} \end{aligned} \tag{1}$$

The authors are with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK. Email: {goran.banjac, paul.goulart}@eng.ox.ac.uk

This work was supported by the European Commission research project FP7-PEOPLE-2013-ITN under grant agreement no. 607957 [*Training in Embedded Optimization and Predictive Control (TEMPO)*].

for different values of the weighting parameter γ , and to find a γ sufficiently large to obtain a solution in which the original cardinality constraint is satisfied. However, any such solution to (1) is not guaranteed to be optimal for the original problem \mathcal{P} .

This approach to obtaining a sparse solution has received significant attention, because the resulting convex problem can be solved very efficiently. Since the problem (1) has special structure, proximal methods can be used to obtain a solution when the dimension of the optimization variable is very large [6]. A related approach is to replace the ℓ_1 norm in (1) by a weighted ℓ_1 norm, and to then solve a sequence of weighted ℓ_1 optimization problems wherein the weights are updated based on previously obtained solutions [7].

However, the quality of a solution obtained by relaxing the non-convex constraint is not always satisfactory. One alternative approach is to tackle the non-convex problem \mathcal{P} directly. The exact minimization of such a non-convex problem is possible using techniques such as branch-and-bound, but the approach is only applicable to problems of relatively low dimension. The authors in [8], [9] proposed a projected gradient method for solving a variant of problem \mathcal{P} that does not consider the constraint $x \in \mathcal{B}$, and established convergence of the algorithm to a local minimum under some regularity assumptions on the objective function. The authors in [10] use the alternating direction method of multipliers (ADMM) to solve problem \mathcal{P} , with additional heuristic steps that may improve quality of the solution obtained. However, this method is not guaranteed to converge and the authors propose to terminate the algorithm after some predefined number of iterations.

In this paper we reformulate problem \mathcal{P} by replacing the ℓ_1 regularization term in (1) by a function that penalizes only the $(n - k)$ smallest elements in magnitude of x , rather than all the elements of x . We then apply the proximal gradient method (PGM) to the reformulated problem, and show that under some regularity assumptions on the objective function our algorithm converges to a local minimum. We show that, given a suitable selection of the weighting parameter, our algorithm is equivalent to the projected gradient method for solving problem \mathcal{P} . However, we propose an update scheme of the weighting parameter which often yields a better solution than when the parameter is held fixed.

The remainder of the paper is organized as follows. In Section II we show how to reformulate problem \mathcal{P} . In Section III we propose an algorithm for solving such a reformulated problem and establish convergence of the algorithm. Section IV provides a numerical example which shows that our algorithm outperforms other methods proposed in the literature. Finally, Section V concludes the paper.

Notation

Let \mathbb{R} denote the set of real numbers, \mathbb{R}_+ the set of nonnegative real numbers, $\tilde{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ the extended real line, and \mathbb{R}^n the n -dimensional real space equipped with inner product $\langle x, y \rangle$ and induced norm $\|x\|$. Given a vector $x \in \mathbb{R}^n$, we denote its cardinality by $\text{card}(x)$,

i.e. the number of nonzero elements in x , and its i -th element by x_i . We define a sequence of indices $\{i_s(x)\}_{s=1}^n$ such that

$$|x_{i_1(x)}| \geq \cdots \geq |x_{i_n(x)}|.$$

We denote the *largest- k norm* of a vector $x \in \mathbb{R}^n$ by $\|x\|_{[k]} : \mathbb{R}^n \rightarrow \mathbb{R}_+$, defined as the sum of the k elements of largest magnitude in x , *i.e.*

$$\|x\|_{[k]} := \sum_{s=1}^k |x_{i_s(x)}|.$$

Observe that $\|x\|_{[1]}$ corresponds to the ℓ_∞ norm and $\|x\|_{[n]}$ to the ℓ_1 norm, and that $\|x\|_1 \geq \|x\|_{[k]} \geq \|x\|_\infty$.

Let $\text{dom } f := \{x \in \mathbb{R}^n \mid f(x) < \infty\}$ denote the effective domain of a function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$. The proximal operator of f is given by $\text{prox}_f(x) \in \text{argmin}_y \{f(y) + \frac{1}{2}\|y - x\|^2\}$. The directional derivative of f at $x \in \text{dom } f$ in the direction $d \in \mathbb{R}^n$ is denoted by $f'(x, d)$. If f is differentiable, then $\nabla f(x)$ denotes the gradient of f evaluated at x , and $\nabla_i f(x)$ the i -th element of $\nabla f(x)$. The distance of x to a closed set $\mathcal{C} \subseteq \mathbb{R}^n$ is denoted by $\text{dist}(x, \mathcal{C}) := \min_{c \in \mathcal{C}} \|x - c\|$, and the projection of x onto \mathcal{C} by $\text{proj}_{\mathcal{C}}(x) \in \text{argmin}_{c \in \mathcal{C}} \|x - c\|$. Projection of x onto \mathbb{R}_+ is denoted by $x_+ := \max(x, 0)$. The *indicator function* of a set \mathcal{C} is given by

$$\mathcal{I}_{\mathcal{C}}(x) := \begin{cases} 0, & x \in \mathcal{C} \\ \infty, & \text{otherwise.} \end{cases}$$

The *soft thresholding operator* $S_\kappa : \mathbb{R} \rightarrow \mathbb{R}$, with $\kappa \geq 0$, is given by

$$S_\kappa(x) := \begin{cases} x + \kappa, & x < -\kappa \\ 0, & x \in [-\kappa, \kappa] \\ x - \kappa, & x > \kappa, \end{cases}$$

and the *saturation operator* $\text{sat}_{l,u} : \mathbb{R} \rightarrow \mathbb{R}$, with $-l, u \in \mathbb{R}_+ \cup \{\infty\}$, as

$$\text{sat}_{l,u}(x) := \begin{cases} l, & x < l \\ x, & x \in [l, u] \\ u, & x > u. \end{cases}$$

II. PROBLEM REFORMULATION

We make the following standing assumptions about the cardinality constrained optimization problem \mathcal{P} :

Assumption 1.

- a) The function f is convex and differentiable.
- b) The gradient ∇f is Lipschitz continuous on \mathcal{B} with Lipschitz constant L_f , *i.e.*

$$\|\nabla f(x) - \nabla f(y)\| \leq L_f \|x - y\|, \quad \forall x, y \in \mathcal{B}.$$

c) The function f is lower bounded on \mathcal{B} , i.e. there exists $\zeta \in \mathbb{R}$ such that $f(x) \geq \zeta$ for all $x \in \mathcal{B}$.

Notice that a sufficient condition for f to be lower bounded on \mathcal{B} is that \mathcal{B} is bounded.

In this section we will show how to reformulate problem \mathcal{P} . Observe that the cardinality constraint can be equivalently written as $\|x\|_1 = \|x\|_{[k]}$ [11]. Since $\|x\|_1 \geq \|x\|_{[k]}$ is always true, the cardinality constraint in problem \mathcal{P} can be replaced by $\|x\|_1 - \|x\|_{[k]} \leq 0$. By representing the resulting problem in the Lagrangian form, we obtain

$$\begin{aligned} & \text{minimize} && f(x) + \gamma \varphi_{[k]}(x) \\ & \text{subject to} && x \in \mathcal{B} \end{aligned} \quad (\mathcal{R})$$

where $\varphi_{[k]}(x) := \|x\|_1 - \|x\|_{[k]}$, and $\gamma \in \tilde{\mathbb{R}}_+$ is a weighting parameter. Observe that, compared to problem (1), in problem \mathcal{R} only the $(n - k)$ smallest elements in magnitude of x are penalized, rather than all the elements of x .

We show in the sequel that, for an appropriately selected parameter γ , local minimizers of problems \mathcal{P} and \mathcal{R} coincide. We first recall local optimality conditions for a nondifferentiable function h .

Lemma 1 (see e.g. [12]). *Let the function $h : \mathbb{R}^n \rightarrow \tilde{\mathbb{R}}$ be proper and let $x^* \in \text{dom } h$. Then x^* is a local minimizer of h if and only if $h'(x^*, d) \geq 0$ for all $d \in \mathbb{R}^n$.*

By introducing the following functions

$$\begin{aligned} h_{\mathcal{P}}(x) &:= f(x) + \mathcal{I}_{\mathcal{B}}(x) + \mathcal{I}_{\text{card}(x) \leq k}(x) \\ h_{\mathcal{R}}(x) &:= f(x) + \mathcal{I}_{\mathcal{B}}(x) + \gamma \varphi_{[k]}(x), \end{aligned}$$

we can characterize local minimizers of problems \mathcal{P} and \mathcal{R} in a way that they are equivalent to local minimizers of $h_{\mathcal{P}}$ and $h_{\mathcal{R}}$, respectively.

Theorem 1. *If $x^* \in \mathcal{B}$ is locally optimal for problem \mathcal{R} with $\gamma > \max_i |\nabla f_i(x^*)|$, then x^* is locally optimal for problem \mathcal{P} , and vice versa.*

Proof: 1) Suppose that x^* is locally optimal for problem \mathcal{P} , and thus $\varphi_{[k]}(x^*) = 0$. Let \mathcal{G} be the set of indices i for which $x_i^* \neq 0$, and $d \in \mathbb{R}^n$ any vector such that $\|d\|_{\infty} < \min_{i \in \mathcal{G}} |x_i^*|/2 =: \varepsilon$. Then the set of indices of the k largest elements in magnitude of $(x^* + d)$ is equivalent to \mathcal{G} , and thus $\varphi_{[k]}(x^* + d) = \sum_{j \notin \mathcal{G}} |d_j|$. Due to the convexity of f and $\mathcal{I}_{\mathcal{B}}$, and separability of $\mathcal{I}_{\mathcal{B}}$, we have

$$\begin{aligned} & h_{\mathcal{R}}(x^* + d) - h_{\mathcal{R}}(x^*) \\ & \geq \langle \nabla f(x^*), d \rangle + \mathcal{I}'_{\mathcal{B}}(x^*, d) + \gamma \sum_{j \notin \mathcal{G}} |d_j| \\ & = \sum_{i \in \mathcal{G}} (\nabla_i f(x^*) d_i + \mathcal{I}'_{\mathcal{B}}(x_i^*, d_i)) + \\ & \quad \sum_{j \notin \mathcal{G}} (\nabla_j f(x^*) d_j + \mathcal{I}'_{\mathcal{B}}(x_j^*, d_j)) + \gamma \sum_{j \notin \mathcal{G}} |d_j| \end{aligned}$$

Observe that $\mathcal{I}'_{\mathcal{B}}(x_j^*, d_j)$ is 0 if $(x_j^* + d_j) \in \mathcal{B}$, and ∞ otherwise. In particular, we always have $\mathcal{I}'_{\mathcal{B}}(x_j^*, d_j) \geq 0$. On the other hand, since x^* is a local minimizer of problem \mathcal{P} , this means that moving x from x^* along coordinates $i \in \mathcal{G}$ either leads to infeasibility or does not improve f , *i.e.* the first sum in the above equality is nonnegative. Therefore, the above inequality reduces to

$$\begin{aligned} h_{\mathcal{R}}(x^* + d) - h_{\mathcal{R}}(x^*) &\geq \sum_{j \notin \mathcal{G}} \nabla_j f(x^*) d_j + \gamma \sum_{j \notin \mathcal{G}} |d_j| \\ &\geq -\max_{j \notin \mathcal{G}} |\nabla_j f(x^*)| \sum_{j \notin \mathcal{G}} |d_j| + \gamma \sum_{j \notin \mathcal{G}} |d_j| \\ &= \left(\gamma - \max_{j \notin \mathcal{G}} |\nabla_j f(x^*)| \right) \sum_{j \notin \mathcal{G}} |d_j|. \end{aligned}$$

Provided that $\gamma > \max_i |\nabla f_i(x^*)| \geq \max_{j \notin \mathcal{G}} |\nabla f_j(x^*)|$, the above inequality implies

$$h_{\mathcal{R}}(x^* + d) \geq h_{\mathcal{R}}(x^*), \quad (2)$$

for all d such that $\|d\|_{\infty} < \varepsilon$, *i.e.* x^* is a local minimizer of \mathcal{R} [13]. This concludes the first direction of the proof.

2) Suppose that x^* is locally optimal for problem \mathcal{R} with $\gamma > \max_i |\nabla f_i(x^*)|$. We first show that $\varphi_{[k]}(x^*) = 0$. Assume that $\varphi_{[k]}(x^*) \neq 0$ which means that the $(k+1)$ -th largest element in magnitude of x^* , whose index is denoted by i , is not equal zero. Then the directional derivative of $h_{\mathcal{R}}$ at x^* in the direction $d = (0, \dots, 0, -x_i^*, 0, \dots, 0)$ is

$$\begin{aligned} h'_{\mathcal{R}}(x^*, d) &= \langle \nabla f(x^*), d \rangle + \mathcal{I}'_{\mathcal{B}}(x^*, d) + \gamma \varphi'_{[k]}(x^*, d) \\ &= -\nabla_i f(x^*) x_i^* - \gamma |x_i^*| \\ &\leq (|\nabla f_i(x^*)| - \gamma) |x_i^*| \\ &< 0, \end{aligned}$$

where we used the fact that $\mathcal{I}'_{\mathcal{B}}(x_i^*, -x_i^*) = 0$ coming from $0 \in \mathcal{B}$, $\gamma > |\nabla f_i(x^*)|$, and $|x_i^*| > 0$. In the sense of Lemma 1, the above inequality means that x^* is not a local minimizer of \mathcal{R} , which is a contradiction.

Now observe that if $d \in \mathbb{R}^n$ points in a feasible direction from x^* with respect to the cardinality constraints, then $\mathcal{I}'_{\text{card}(x) \leq k}(x^*, d) = \gamma \varphi'_{[k]}(x^*, d) = 0$, and otherwise $\mathcal{I}'_{\text{card}(x) \leq k}(x^*, d) = \infty$ and $\gamma \varphi'_{[k]}(x^*, d) \leq \gamma \|d\|_{[k]} < \infty$. Thus we have $\mathcal{I}'_{\text{card}(x) \leq k}(x^*, d) \geq \gamma \varphi'_{[k]}(x^*, d)$ for all d , and therefore

$$\begin{aligned} h'_{\mathcal{P}}(x^*, d) &= \langle \nabla f(x^*), d \rangle + \mathcal{I}'_{\mathcal{B}}(x^*, d) + \mathcal{I}'_{\text{card}(x) \leq k}(x^*, d) \\ &\geq \langle \nabla f(x^*), d \rangle + \mathcal{I}'_{\mathcal{B}}(x^*, d) + \gamma \varphi'_{[k]}(x^*, d) \\ &= h'_{\mathcal{R}}(x^*, d) \\ &\geq 0, \end{aligned}$$

The above inequality means that x^* is a local minimizer of problem \mathcal{P} . This concludes the proof. \blacksquare

III. SOLUTION METHOD

Operator splitting methods were originally designed for solving optimization problems in the form

$$\text{minimize } f(x) + g(x) \quad (3)$$

where both f and g are convex. However, they are sometimes used as heuristics in non-convex optimization [4], [6], [9], [10]. The advantage of these methods is that the functions f and g can be tackled separately. For instance, in the PGM the function f is tackled through its gradient, and g through its proximal operator. In the case when both f and g are convex, and ∇f is Lipschitz continuous, the method converges to a global minimum [14]. Another method that can be used for solving problem (3) is ADMM, which uses the proximal operators of both f and g . This method does not require any differentiability assumptions on the functions and, provided that both f and g are convex and that problem (3) is solvable, the method always converges to a global minimum [14].

If we set

$$g(x; \gamma, l, u) := \mathcal{I}_{\mathcal{B}}(x) + \gamma \varphi_{[k]}(x), \quad (4)$$

where (γ, l, u) are parameters of the function, then problem \mathcal{R} can be represented in the form (3). In order to use proximal methods for solving problem \mathcal{R} efficiently, the proximal operator of the function g should be easy to evaluate. We will show in the sequel how to evaluate the proximal operator of g , and that the PGM can be used for computing a local minimum of problem \mathcal{R} .

Note that the authors in [15] use the same reformulation to tackle problem \mathcal{P} , but have a different approach for solving it. In particular, the authors represent the objective in problem \mathcal{R} as a difference of convex functions (*i.e.* a *DC function*) and apply so-called DC algorithms to the problem.

A. Evaluating the proximal operator

The authors in [9] show that projection of a vector in \mathbb{R}^n onto subset of \mathbb{R}^n with cardinality k can be obtained by setting its $(n - k)$ smallest elements in magnitude to zero. Since the choice of these elements does not have to be unique, neither is the projection on the cardinality constraint. This corresponds to a projection onto the constraint set in \mathcal{P} when setting $-l = u = \infty$. In the case that $-l = u = M > 0$, the projection can be obtained by additional clipping of the elements with magnitude larger than M , as shown in [10]. It is easy to show that in the case when $l = 0$ and $u = M > 0$, the projection can be obtained by first setting all the nonnegative elements to zero, and then performing projection as in the case when $-l = u = M > 0$.

We can similarly evaluate the proximal operator of functions $g_M(x; \gamma) := g(x; \gamma, -M, M)$

$$(\text{prox}_{g_M}(x; \gamma))_{i_s(x)} = \begin{cases} \text{sat}_{-M, M}(S_\gamma(x_{i_s(x)})), & s > k \\ \text{sat}_{-M, M}(x_{i_s(x)}), & s \leq k \end{cases}$$

Algorithm 1 Proximal gradient method

- 1: **Initialization**
 - 2: $t = 0$
 - 3: Set $x_0 \in \mathcal{B}$
 - 4: **Repeat until convergence**
 - 5: $y_{t+1} = x_t - (1/L)\nabla f(x_t)$
 - 6: $x_{t+1} \in \text{prox}_{g/L}(y_{t+1})$
 - 7: $t \leftarrow t + 1$
-

and $g_{M_+}(x; \gamma) := g(x; \gamma, 0, M)$

$$(\text{prox}_{g_{M_+}}(x; \gamma))_{i_s(x_+)} = \begin{cases} \text{sat}_{0,M}(S_\gamma(x_{i_s(x_+)})), & s > k \\ \text{sat}_{0,M}(x_{i_s(x_+)}), & s \leq k \end{cases}$$

Note that in order to evaluate the proximal operators of g_M and g_{M_+} , we must first sort the vectors x and x_+ , respectively. Notice that when $\gamma = \infty$, the proximal operator of g is equivalent to projection onto the constraint set in \mathcal{P} . This implies that the PGM for solving problem \mathcal{R} generalizes the projected gradient method for solving problem \mathcal{P} for which the convergence to a local minimum was established in [9]. We will show in the next subsection that we can extend this result to any $\gamma > 0$ in the PGM.

B. Convergence of the proximal gradient method

The proximal gradient method, described in Algorithm 1, is known to converge when f and g are convex, ∇f has Lipschitz continuous gradient with constant L_f , and $L > L_f/2$ [16]. The authors in [9] establish convergence of the PGM when g is an indicator function of a general closed set, provided that $L > L_f$. We will use similar arguments to establish convergence of the PGM for an arbitrary non-convex function g . We first require a supporting lemma:

Lemma 2. Any fixed-point x^* of Algorithm 1, i.e. a point satisfying

$$x^* \in \text{prox}_{g/L}(x^* - \frac{1}{L}\nabla f(x^*))$$

is a local minimizer of problem (3).

Proof: From the definition of the proximal operator, x^* must satisfy

$$x^* \in \underset{z}{\text{argmin}} \left\{ g(z) + \frac{L}{2} \|z - (x^* - \frac{1}{L}\nabla f(x^*))\|^2 \right\}. \quad (5)$$

According to Lemma 1, the above inclusion is equivalent to

$$g'(x^*, d) + \langle \nabla f(x^*), d \rangle \geq 0, \quad \forall d \in \mathbb{R}^n,$$

which again by Lemma 1 implies that x^* is a local minimizer of problem (3). This concludes the proof. ■

Theorem 2. Let $\{x_t\}_{t \geq 0}$ be a sequence of iterates generated by Algorithm 1 for solving problem (3). Suppose that $(f + g)$ is lower bounded, f is convex, ∇f is Lipschitz continuous with constant L_f , $\inf L > L_f$ and $\sup L < \infty$. Then the sequence $\{f(x_t) + g(x_t)\}_{t \geq 0}$ converges to a local minimum of problem (3).

Proof: Let $h(x) := f(x) + g(x)$ and

$$h_L(z; x) := f(x) + \langle \nabla f(x), z - x \rangle + \frac{L}{2} \|z - x\|^2 + g(z).$$

We first show that Algorithm 1 generates a non-increasing sequence $\{h(x_t)\}_{t \geq 0}$. Similar to (5), x_{t+1} can be characterized as

$$x_{t+1} \in \operatorname{argmin}_z \left\{ g(z) + \frac{L}{2} \left\| z - \left(x_t - \frac{1}{L} \nabla f(x_t) \right) \right\|^2 \right\}. \quad (6)$$

Lipschitz continuity of ∇f implies that $h_{L_f}(z; x)$ is an upper bound on $h(z)$ [16], i.e.

$$h_{L_f}(z; x) \geq h(z), \quad \forall z \in \mathbb{R}^n. \quad (7)$$

Since

$$\begin{aligned} h_L(z; x_t) &= g(z) + \frac{L}{2} \left\| z - \left(x_t - \frac{1}{L} \nabla f(x_t) \right) \right\|^2 \\ &\quad + f(x_t) - \frac{1}{2L} \|\nabla f(x_t)\|^2, \end{aligned}$$

where the last two summands do not depend on z , inclusion (6) is equivalent to

$$x_{t+1} \in \operatorname{argmin}_z h_L(z; x_t),$$

which implies

$$h_L(x_{t+1}; x_t) \leq h_L(x_t; x_t) = h(x_t). \quad (8)$$

Inequalities (7) and (8) imply

$$\begin{aligned} h(x_t) - h(x_{t+1}) &\geq h_L(x_{t+1}; x_t) - h_{L_f}(x_{t+1}; x_t) \\ &= \frac{L - L_f}{2} \|x_{t+1} - x_t\|^2. \end{aligned} \quad (9)$$

The last inequality shows that $\{h(x_t)\}_{t \geq 0}$ is a strictly decreasing sequence but for the case where $x_{t+1} = x_t$, for which x_t is a fixed-point of Algorithm 1 and, according to Lemma 2, a local minimizer of (3). Lower boundedness of h together with monotonicity of $\{h(x_t)\}_{t \geq 0}$ implies convergence by the monotone convergence theorem. \blacksquare

Corollary 1. Suppose that Assumption 1 holds, $\inf L > L_f$, $\sup L < \infty$, and let $\{x_t\}_{t \geq 0}$ be a sequence of iterates generated by Algorithm 1 for solving problem \mathcal{R} . Then $\{h_{\mathcal{R}}(x_t)\}_{t \geq 0}$ converges to a local minimum of \mathcal{R} .

C. Termination criterion

As shown in the previous subsection, the Algorithm 1 for solving problem \mathcal{R} always produces a monotonically decreasing sequence $\{h_{\mathcal{R}}(x_t)\}_{t \geq 0}$ that eventually converges to some local minimum $h_{\mathcal{R}}^*$. Since we do not know $h_{\mathcal{R}}^*$, a reasonable termination criterion is that the difference in objective value in subsequent iterations is small relative to the objective value, *i.e.*

$$h_{\mathcal{R}}(x_t) - h_{\mathcal{R}}(x_{t+1}) \leq \varepsilon h_{\mathcal{R}}(x_{t+1}),$$

where $\varepsilon \in \mathbb{R}_+$ is the optimality tolerance. Notice from (9) that this condition implies that $\|x_{t+1} - x_t\|$ is also small.

D. Heuristic for selection of the weighting parameter

In Section II we showed that local minimizers of problems \mathcal{P} and \mathcal{R} coincide provided that $\gamma > \max_i |\nabla f_i(x^*)|$. Although $\nabla f(x^*)$ is not known prior to the algorithm runtime, in some cases we can find an upper bound on $\max_i |\nabla f_i(x^*)|$ over \mathcal{B} , and use it in order to select an appropriate γ . However, such a selection rule usually results in a relatively large value of γ , and consequently in equivalence between our proposed method and the projected gradient method for solving problem \mathcal{P} .

We thus propose a heuristic for updating the weighting parameter γ at each iteration. After y_{t+1} is computed, we update the weighting parameter in each iteration according to the following rule

$$\gamma_{t+1} = \max_i |\nabla_i f(y_{t+1})|. \quad (10)$$

According to Theorem 1, this selection rule does not guarantee that a limit point of the algorithm will satisfy the original cardinality constraint. In order to obtain a vector that satisfies the cardinality constraint, it is sufficient to project the obtained vector onto the constraint set in problem \mathcal{P} . We will show in the next section that this strategy usually results in higher quality solutions.

IV. NUMERICAL RESULTS

We consider the following sparse least-squares problem

$$\begin{aligned} & \text{minimize} && \|Ax - b\|_2^2 \\ & \text{subject to} && \|x\|_\infty \leq M \\ & && \text{card}(x) \leq k, \end{aligned} \quad (11)$$

with decision variable $x \in \mathbb{R}^n$ and problem data $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $M \in \mathbb{R}_+$, and $k \in \mathbb{N}$.

This problem arises, for instance, in compressed sensing where a sparse vector x must be recovered from linear measurements $Ax = b$ [9]. Observe that problem (11) satisfies Assumption 1 with $\zeta = 0$ and $L_f = 2 \|A^T A\|_2$ where $\|\cdot\|_2$ denotes the matrix spectral norm.

Algorithm 2 ADMM

- 1: **Initialization**
 - 2: $t = 0$
 - 3: Set $z_0 \in \mathcal{B}$, $u_0 = 0$
 - 4: **Repeat until** $t = N$
 - 5: $x_{t+1} = \text{prox}_{f/\rho}(z_t - u_t)$
 - 6: $z_{t+1} = \text{prox}_{g/\rho}(x_{t+1} + u_t)$
 - 7: $u_{t+1} = u_t + x_{t+1} - z_{t+1}$
 - 8: $t \leftarrow t + 1$
-

We will show in the sequel that for problem (11) the algorithm proposed in the previous section, denoted here by *prox-PGM*, produces solutions of a higher quality than other algorithms reported in the literature. We briefly describe these algorithms.

- **Lasso:** The least absolute shrinkage and selection operator (Lasso) is a well-known heuristic for solving problem (11) and is based on a convex relaxation as in problem (1). We solve the problem for different values of the weighting parameter γ , and keep the solution obtained with the smallest value of the parameter for which the cardinality constraint is satisfied.
- **proj-PGM:** The projected gradient method was used in [9] for solving a variant of problem (11), and is equivalent to Algorithm 1 when $\gamma = \infty$. The method converges to a local minimum, and the same termination criterion can be used as proposed for the prox-PGM.
- **proj-ADMM:** A method proposed in [10] for solving problem (11) is equivalent to Algorithm 2, when $\gamma = \infty$. We set the algorithm parameter $\rho = L$. The method does not necessarily converge, and we terminate the algorithm after some fixed number of iterations is reached, as proposed in [10]. Here, we stop the algorithm after 100 iterations. Notice that evaluating the proximal operator of f requires solving a linear system, which is computationally more expensive than evaluating the gradient of f . However, since the left-hand side of the linear system does not change, we can factor the matrix once and use the cached factorization in the subsequent iterations.
- **prox-ADMM:** As Algorithm 1 can be used for solving both problems \mathcal{P} and \mathcal{R} , we can use Algorithm 2 also for solving problem \mathcal{R} . We denote this method by prox-ADMM and it can be seen as a generalization of proj-ADMM. We use the same strategy as in (10) to update the weighting parameter of the problem, with y_{t+1} replaced by x_{t+1} .

The data are generated as described in [10, Section 6.1], *i.e.* $A \in \mathbb{R}^{m \times 2m}$ with i.i.d. $\mathcal{N}(0, 1)$ entries, $b = A\hat{x} + v$ with \hat{x} drawn uniformly at random from the set of vectors satisfying $\text{card}(x) \leq \lfloor m/5 \rfloor$ and $\|x\|_\infty \leq M = 1$, and $v \in \mathbb{R}^m$ being a noise vector drawn from $\mathcal{N}(0, \sigma^2 I)$. We set $\sigma = \|A\hat{x}\|/(20\sqrt{m})$ so that the signal-to-noise ratio is near 20.

For each value of m we generate 100 instances of the problem. Since the quality of the solutions obtained depends on the initial point (except for the Lasso method), for each problem

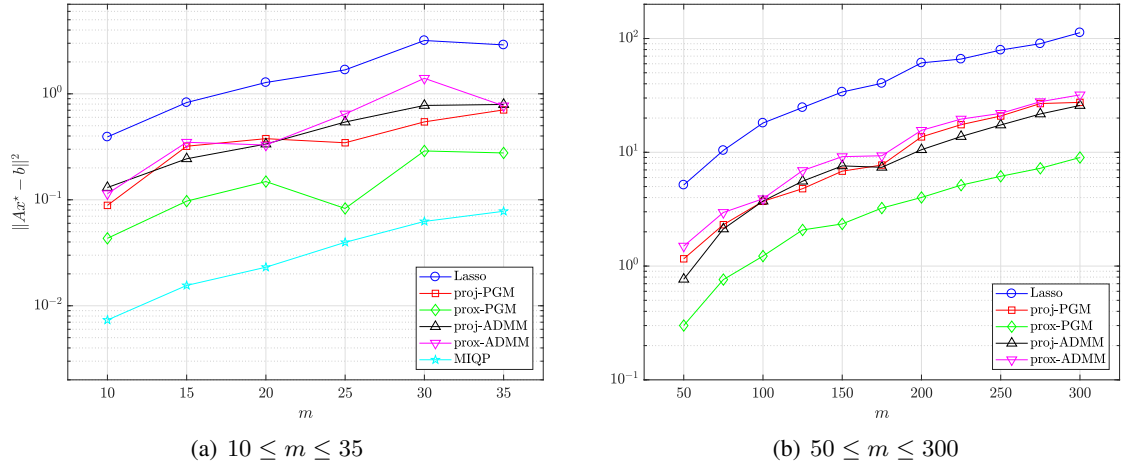


Fig. 1. The average objective function of solutions found by various algorithms for 100 instances of problem (11).

instance we run the algorithms from 10 initial points drawn from $\mathcal{N}(0, I)$, and keep the solution with the smallest objective value. The resulting solution is then *polished* in a sense that we replace the cardinality constraint in (11) with the sparsity pattern of the solution, and solve the resulting convex optimization problem to obtain a final solution. Although the authors in [10] propose additional heuristic steps such as the *neighbor search* that may improve quality of the obtained solutions, we do not include these heuristics in our implementations. Note that these heuristics are not restricted to a specific algorithm, but can be used in any of the proposed methods that solve non-convex problems.

To make the comparison fair, we run all the algorithms for solving non-convex problems for not more than 100 iterations. We use Gurobi [17] to solve convex quadratic programs (QPs) arising in Lasso and solution polishing, and mixed-integer QPs.

The numerical results obtained are shown in Figure 1. For each value of m we show the average value of the objective function over the 100 generated instances. The exact solutions are obtained for small values of m by solving mixed-integer QPs. Solutions obtained with the Lasso approach have values of the objective function around one order of magnitude larger than solutions obtained by prox-PGM. Also, prox-PGM consistently outperforms all the other methods by at least a factor of 2 (relative to the exact solution for $m \leq 35$ and relative to zero for $m \geq 50$) for all values of m .

Figure 2 shows the average times for solving one instance of the problem. The time required to solve mixed-integer QPs grows rapidly with m and the approach is applicable only for solving small problems. On the other hand, runtimes of the operator splitting methods scale much better with the problem dimensions making them suitable for solving large-scale problems.

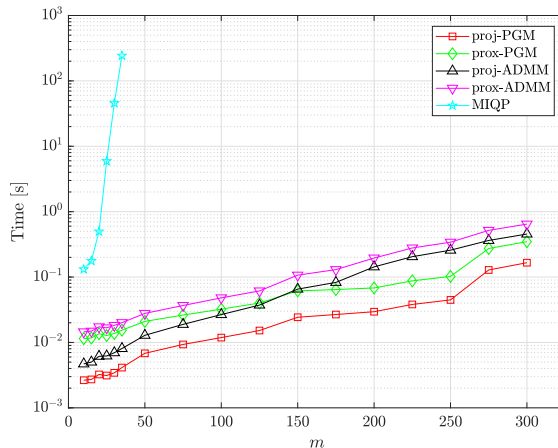


Fig. 2. The average runtimes of various algorithms for solving 100 instances of problem (11).

V. CONCLUSION

In this paper we propose a method for minimizing a convex differentiable function subject to sparsity constraints. We show that under suitably selected weighting parameter of a reformulated problem, the proximal gradient method converges to a local minimum of the original problem. We also propose a heuristic that updates the weighting parameter in each iteration of the algorithm. The performance of our method is compared to other methods proposed in the literature for solving such problems. Our method consistently outperforms all the other solution methods considered in this paper by more than a factor of 2.

REFERENCES

- [1] D. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] A. Miller, *Subset Selection in Regression*, ser. Chapman & Hall, 2002.
- [3] J. Mairal, F. Bach, and J. Ponce, “Sparse modeling for image and vision processing,” *Foundations and Trends in Computer Graphics and Vision*, vol. 8, no. 2-3, pp. 85–283, 2014.
- [4] F. Lin, M. Fardad, and M. Jovanović, “Design of optimal sparse feedback gains via the alternating direction method of multipliers,” *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2426–2431, 2013.
- [5] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.
- [6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2010.
- [7] E. J. Candés, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.
- [8] T. Blumensath and M. E. Davies, “Iterative thresholding for sparse approximations,” *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 629–654, 2008.

- [9] A. Beck and Y. C. Eldar, "Sparsity constrained nonlinear optimization: Optimality conditions and algorithms," *SIAM Journal on Optimization*, vol. 23, no. 3, pp. 1480–1509, 2013.
- [10] S. Diamond, R. Takapoui, and S. Boyd, "A general system for heuristic solution of convex problems over nonconvex sets," *Optimization Methods and Software*, 2017.
- [11] A. B. Hempel and P. J. Goulart, "A novel method for modelling cardinality and rank constraints," in *IEEE Conference on Decision and Control (CDC)*, 2014, pp. 4322–4327.
- [12] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of Optimization Theory and Applications*, vol. 109, no. 3, pp. 475–494, 2001.
- [13] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.
- [14] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer-Verlag, 2011.
- [15] J. Gotoh, A. Takeda, and K. Tono, "DC formulations and algorithms for sparse optimization problems," *Technical report, The University of Tokyo*, pp. 1–36, 2015.
- [16] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*, ser. Applied optimization. Kluwer Academic Publishers, 2004.
- [17] Gurobi Optimization Inc. (2016) Gurobi optimizer reference manual. [Online]. Available: <http://www.gurobi.com>