Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 14, No. 2, pp. 273–305, 2018

COMPLEMENTARITY FORMULATIONS OF ℓ_0-NORM OPTIMIZATION PROBLEMS

Mingbin Feng*, John E. Mitchell†, Jong-Shi Pang‡
Xin Shen§ and Andreas Wächter¶

Abstract: In a number of application areas, it is desirable to obtain sparse solutions. Minimizing the number of nonzeroes of the solution (its ℓ_0-norm) is a difficult nonconvex optimization problem, and is often approximated by the convex problem of minimizing the ℓ_1-norm. In contrast, we consider exact formulations as mathematical programs with complementarity constraints and their reformulations as smooth nonlinear programs. We discuss properties of the various formulations and their connections to the original ℓ_0-minimization problem in terms of stationarity conditions, as well as local and global optimality. We prove that any limit point of a sequence of local minimizers to a relaxation of the formulation satisfies certain desirable properties. Numerical experiments using randomly generated problems show that standard nonlinear programming solvers, applied to the smooth but nonconvex equivalent reformulations, are often able to find sparser solutions than those obtained by the convex ℓ_1-approximation.

Key words: *ℓ_0-norm minimization, complementarity constraints, nonlinear programming*

Mathematics Subject Classification: *90C33, 90C46, 90C30*

1 Introduction

Denoted by $\| \bullet \|_0$, the so-called ℓ_0-norm of a vector is the number of nonzero components of the vector. In recent years, there has been an increased interest in solving optimization problems that minimize or restrict the number of nonzero elements of the solution vector [2, 3, 8, 10, 11, 12, 34, 37]. A simple example of such a problem is that of finding a solution to a system of linear inequalities with the least ℓ_0-norm:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \| x \|_0 \\ \text{subject to} & Ax \geq b \quad \text{and} \quad Cx = d, \end{array} \tag{1.1}$$

*This author is supported by National Science Foundation grant DMS-1216920.

†This author was supported by the National Science Foundation under Grant Number CMMI-1334327 and by the Air Force Office of Scientific Research under Grant Number FA9550-11-1-0260.

‡This author was supported by the National Science Foundation under Grant Number CMMI-1333902 and by the Air Force Office of Scientific Research under Grant Number FA9550-11-1-0151.

§This author was supported by the National Science Foundation under Grant Number CMMI-1334327 and by the Air Force Office of Scientific Research under Grant Number FA9550-11-1-0260.

¶This author was supported by National Science Foundation grant DMS-1216920 and grant CMMI-1334639.

where $A \in \mathbb{R}^{m\times n}$, $C \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^m$ and $d \in \mathbb{R}^k$ are given matrices and vectors, respectively. Since this problem is NP-hard, one popular solution approach replaces the nonconvex discontinuous ℓ_0-norm in (1.1) by the convex continuous ℓ_1-norm, leading to a linear program:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & \|\, x \,\|_1 \\ \text{subject to} & Ax \geq b \quad \text{and} \quad Cx = d. \end{array} \tag{1.2}$$

Theoretical results are known that provide sufficient conditions under which an optimal solution to (1.2) is also optimal to (1.1) [7, 14, 21, 38]. Yet these results are of limited practical value as the conditions can not easily be verified or guaranteed for specific realizations of (1.1); thus in general, optimal solutions to (1.2) provide suboptimal solutions to (1.1).

It is our contention that, from a practical perspective, improved solutions to (1.1) can be obtained by reformulating the ℓ_0-norm in terms of complementarity constraints [29]. This leads to a linear program with linear complementarity constraints (LPCC) which can be solved with specialized algorithms that do not depend on the feasibility and/or boundedness of the constraints [23, 24]. In the event that bounds are known on the solutions of the problem, the LPCC can be further reformulated as a mixed-integer linear program (MILP). However, the solution of this MILP is usually too time-consuming for large instances.

As an alternative to the MILP approach, the LPCC can be expressed directly as a smooth continuous nonlinear program (NLP). It is the main purpose of this research to examine the quality of solutions computed by standard NLP solvers applied to these smooth reformulations of the ℓ_0-norm. There are two features of the NLP reformulations that make them difficult to solve. First, the NLPs are highly nonconvex, and, consequently, the solutions returned by the NLP solvers depend strongly on the starting point, because the NLP methods are typically only able to find local minimizers or Karush-Kuhn-Tucker (KKT) points, instead of global minimizers. Secondly, the NLPs are not well-posed in the sense that they do not satisfy the assumptions that are made usually for the convergence analysis of standard NLP algorithms, such as the Mangasarian-Fromovitz constraint qualification. Nevertheless, our numerical results show that these methods often generate high-quality solutions for the ℓ_0-norm minimization problem (1.1), thus providing a testament of the effectiveness of the NLP solvers applied to a very challenging class of nonconvex problems.

The remainder of this paper is organized as follows. In Section 1.1 we present two basic complementarity formulations for the ℓ_0-norm. One of them leads to an LPCC formulation of the problem (1.1) which is reformulated as a smooth NLP using different approaches, including a new construction based on squared complementarities. The other complementarity formulation results in a nonlinear program with bilinear, disjunctive constraints. These formulations are generalized to the nonlinear case in Section 2 where we introduce an NLP model whose objective comprises a weighted combination of a smooth term and a discontinuous ℓ_0-term. This model is sufficiently broad to encompass many optimization problems that include applications arising from compressive sensing [8, 12], basis pursuit [3, 11], LASSO regression [34, 37], image deblurring [2], the least misclassification (as opposed to the well-known least-residual) support-vector machine problem with a soft margin; the latter problem was first introduced by Mangasarian [10, 30], and a cardinality minimization problem [27].

To give some theoretical background for the expected convergence behavior for (local) NLP solvers, connections between the KKT points of the smooth formulations of the complementarity problems and the original ℓ_0-norm minimization problem are established in Section 3. Further insights are obtained in Section 4 by considering ε-relaxations of the smooth NLP formulations. In particular, convergence of points satisfying second order conditions for the relaxations are discussed in Section 4.3 and this convergence is related to

solutions to the ℓ_1-norm approximation. The practical performance of standard NLP codes for the solution of ℓ_0-norm minimization problems is assessed in Section 5. We present numerical results for an extensive set of computational experiments that show that the solutions obtained by some NLP formulations of the ℓ_0-minimization are significantly better than those obtained from the convex ℓ_1-formulation, often close to the globally optimal objective value. Conclusions and an outlook for future research are given in the final section.

1.1 Equivalent formulations

We start by introducing two basic ways to formulate the ℓ_0-norm using complementarity constraints.

Full complementarity. A straightforward way of formulating the ℓ_0-norm using complementarity constraints is to first express $x = x^+ - x^-$ with $x^\pm$ being the non-negative and non-positive parts of x, respectively; this is followed by the introduction of a vector $\xi \in [0,1]^n$ that is complementary to $|x|$, the absolute-value vector of x. This maneuver leads to the following formulation:

$$\begin{array}{lll} \underset{x,\, x^\pm,\, \xi}{\text{minimize}} & \mathbf{1}_n^T(\,\mathbf{1}_n - \xi\,) = \displaystyle\sum_{j=1}^{n}(\,1-\xi_j\,) & \\ \text{subject to} & Ax \geq b, \quad Cx = d, \quad \text{and} \quad x = x^+ - x^- & (1.3) \\ & 0 \leq \xi \perp x^+ + x^- \geq 0, \quad \xi \leq \mathbf{1}_n & \\ \text{and} & 0 \leq x^+ \perp x^- \geq 0 & \end{array}$$

where $\mathbf{1}_n$ is the n-vector of all ones. It is not difficult to deduce that if x is an optimal solution of (1.1), then by letting $x^\pm \triangleq \max(0, \pm x)$ and

$$\xi_j \triangleq \left\{ \begin{array}{ll} 0 & \text{if } x_j \neq 0 \\ 1 & \text{if } x_j = 0 \end{array} \right\} \quad j = 1, \cdots, n, \tag{1.4}$$

the resulting triple $(x^\pm, \xi)$ is an optimal solution of (1.3) with objective value equal to $\|x\|_0$. Conversely, if $(x^\pm, \xi)$ is an optimal solution of (1.3), then $x \triangleq x^+ - x^-$ is an optimal solution of (1.1) with the same objective value as the optimal objective value of (1.3). The definition (1.4) provides a central connection between (1.1) and its "pieces" to be made precise in Section 3. Such pieces are smooth programs in which some of the x-variables are fixed at zero and correspond in some way to the enumeration of the zero versus nonzero components of x. The scalar $1 - \xi_i$ is the indicator of the support of x_i; we call $\mathbf{1}_n - \xi$ the *support vector* of x.

It is easy to see that the complementarity between the variables $x^\pm$ is not needed in (1.3); this results in the following equivalent formulation of this problem, and thus of (1.1):

$$\begin{array}{lll} \underset{x,\, x^\pm,\, \xi}{\text{minimize}} & \mathbf{1}_n^T(\,\mathbf{1}_n - \xi\,) = \displaystyle\sum_{j=1}^{n}(\,1-\xi_j\,) & \\ \text{subject to} & Ax \geq b, \quad Cx = d, \quad \text{and} \quad x = x^+ - x^- & (1.5) \\ & 0 \leq \xi \perp x^+ + x^- \geq 0 & \\ & x^\pm \geq 0 \quad \text{and} \quad \xi \leq \mathbf{1}_n, & \end{array}$$

In terms of the global resolution of (1.3), maintaining the complementarity between $x^\pm$ could potentially allow sharper cutting planes to be derived in a branch-and-cut scheme for solving

this disjunctive program. This led to better numerical results in our experiments reported in Section 5. We give below several equivalent formulations of the complementarity condition $0 \leq y \perp z \geq 0$ in (1.3) and (1.5) that lead to a smooth continuous NLP formulation:

- $(y, z) \geq 0$ and $y^T z \leq 0$ (inner product complementarity);
- $(y, z) \geq 0$ and $y \circ z \leq 0$, where $u \circ v$ denotes the Hadamard, i.e., componentwise, product of two vectors u and v (componentwise or Hadamard complementarity);
- Adding the penalty term $My^T z$ in the objective for some large scalar $M > 0$ (penalized complementarity);
- $(y, z) \geq 0$ and $(y^T z)^2 \leq 0$ (squared complementarity).

Interestingly, the last formulation, which has never been used in the study of complementarity constraints, turns out to be quite effective for solving some instances of the ℓ_0-norm minimization problem. It can be shown that the only KKT point of this formulation, if it exists, is $(x, x^+, x^-, \xi) = (0, 0, 0, \mathbf{1}_n)$. From a theoretical perspective, this result suggests that it is not a good idea to use an NLP algorithm to solve the ℓ_0-norm minimization problems (1.1) or (2.1) transformed by the squared reformulation. Nevertheless, our numerical experiments reported in Section 5 suggest otherwise. Indeed, the encouraging computational results are the primary reason for us to introduce this squared formulation.

We point out that, with the exception of the penalizing complementarity approach, none of these reformulations of the complementarity problem give rise to a well-posed NLP model in the sense that the Mangasarian-Fromovitz constraint qualification (MFCQ) fails to hold at any feasible point, and the existence of KKT points is not guaranteed. Nevertheless, some NLP solvers have been found to be able to produce good numerical solutions for these reformulations [17].

Half complementarity. There is a simpler formulation, which we call the *half complementarity formulation*, that requires only the auxiliary ξ-variable:

$$\begin{array}{ll} \underset{x,\,\xi}{\text{minimize}} & \mathbf{1}_n^T(\,\mathbf{1}_n - \xi\,) \\ \text{subject to} & Ax \geq b; \quad Cx = d \\ & 0 \leq \xi \leq \mathbf{1}_n; \quad \text{and} \quad \xi \circ x = 0, \end{array} \tag{1.6}$$

The equivalence of (1.1) and (1.6) follows from the same definition (1.4) of ξ. Strictly speaking, the constraints in (1.6) are not of the complementarity type because there is no non-negativity requirement on the variable x; yet the Hadamard constraint $\xi \circ x = 0$ contains the disjunctions: either $\xi_i = 0$ or $x_i = 0$ for all i.

Finally, if a scalar $M > 0$ is known such that $M \geq \|x^*\|_\infty$ for an optimal solution x^*, then the ℓ_0-norm minimization problem (1.1) can be formulated as a mixed-integer linear program with the introduction of a binary variable $\zeta \in \{0, 1\}^n$:

$$\begin{array}{ll} \underset{x,\,\zeta}{\text{minimize}} & \mathbf{1}_n^T \zeta = \displaystyle\sum_{j=1}^{n} \zeta_j \\ \text{subject to} & Ax \geq b \quad \text{and} \quad Cx = d, \\ & -M\zeta \leq x \leq M\zeta, \quad \zeta \in \{0, 1\}^n. \end{array} \tag{1.7}$$

2 A General ℓ_0-norm Minimization Problem

Together, the ℓ_0-norm and its complementarity formulation allow a host of minimization problems involving the count of variables to be cast as disjunctive programs with complementarity constraints. A general NLP model of this kind is as follows: for two finite index sets $\mathcal{E}$ and $\mathcal{I}$,

$$\begin{array}{lll} \underset{x}{\text{minimize}} & f(x) + \gamma \| \, x \, \|_0 & \\ \text{subject to} & c_i(x) \, = \, 0, & i \, \in \, \mathcal{E} \\ \text{and} & c_i(x) \, \leq \, 0, & i \, \in \, \mathcal{I}, \end{array} \tag{2.1}$$

where $\gamma > 0$ is a prescribed scalar and the objective function f and the constraint functions c_i are all continuously differentiable. Let S denote the feasible set of (2.1). A distinguished feature of the problem (2.1) is that its objective function is discontinuous, in fact only lower semicontinuous; as such, it attains its minimum over any compact set. More generally, we have the attainment result as stated in Proposition 2.1. We recall that a function θ is coercive on a set X if $\lim \theta(x) = \infty$ as $\|x\| \to \infty$ for x feasible to (2.1).

Proposition 2.1. *Let the functions f and $\{c_i\}_{i \in \mathcal{I} \cup \mathcal{E}}$ be continuous. If (2.1) is feasible and f is coercive on S, then (2.1) has an optimal solution.*

Proof. Let x^0 be a feasible vector. Since f is continuous and the ℓ_0-norm is lower semicontinuous, the level set $\{ \, x \, \in \, S \, \mid \, f(x) + \gamma \, \| \, x \, \|_0 \, \leq \, f(x^0) + \gamma \, \| \, x^0 \, \|_0 \, \}$ is nonempty and compact, by the coercivity of f. The desired conclusion now follows readily. □

Similar to (1.3), we can derive an equivalent complementarity constrained formulation for (2.1) as follows:

$$\begin{array}{lll} \underset{x, \, x^{\pm}, \, \xi}{\text{minimize}} & f(x) + \boldsymbol{\gamma}^T \, (\, \mathbf{1}_n - \xi \,) & \\ \text{subject to} & c_i(x) \, = \, 0, & i \, \in \, \mathcal{E} \\ & c_i(x) \, \leq \, 0, & i \, \in \, \mathcal{I} \\ & x = x^+ - x^- & \\ & 0 \, \leq \, \xi \, \perp \, x^+ + x^- \, \geq \, 0 & \\ & 0 \, \leq \, x^+ \, \perp \, x^- \, \geq \, 0, \quad \text{and} \quad \xi \, \leq \, \mathbf{1}_n, & \end{array} \tag{2.2}$$

where we have used an arbitrary positive vector $\boldsymbol{\gamma}$ instead of a scalar γ-multiple of the vector of ones. Since both the objective and constraint functions are nonlinear, (2.2) is an instance of a Mathematical Program with Complementarity Constraints (MPCC).

Similar to the half complementarity formulation (1.6) of (1.1), we may associate with (2.1) the following smooth NLP with an auxiliary variable ξ:

$$\begin{array}{lll} \underset{x, \, \xi}{\text{minimize}} & f(x) + \boldsymbol{\gamma}^T \, (\, \mathbf{1}_n - \xi \,) & \\ \text{subject to} & c_i(x) \, = \, 0, & i \, \in \, \mathcal{E} \\ & c_i(x) \, \leq \, 0, & i \, \in \, \mathcal{I} \\ & 0 \, \leq \, \xi \, \leq \, \mathbf{1}_n \quad \text{and} \quad \xi \, \circ \, x \, = \, 0. & \end{array} \tag{2.3}$$

Subsequently, we will relate various properties of the two programs (2.1) and (2.3). Similar results for the full complementarity formulation can be proved if the variable x is nonnegatively constrained. To avoid repetition, we focus on the above half complementarity formulation with no (explicit) sign restriction on x.

The misclassification minimization problem that arises from the literature in support-vector machines [10, 30] provides an example of problem (2.1). Cardinality constrained optimization problems provide a large class of problems where the ℓ_0-norm appears in the constraint that restricts the cardinality of the nonzero elements of the decision variables. These problems are of growing importance; two recent papers study them from different perspectives. Referencing many applications, the paper [39] proposes a piecewise linear approximation of the ℓ_0 constraint as a dc (for difference of convex) constraint. Extending a previous work [5], the report [4] follows a related approach to ours by reformulating the cardinality constraint using complementarity conditions.

3 A Touch of Piecewise Theory

In practice, the problem (2.3) provides a computational platform for solving the problem (2.1). Thus it is important to understand the basic connections between these two problems. Due to the presence of the bilinear constraints: $\xi \circ x = 0$, (2.3) is a nonconvex program even if the original NLP (2.1) with $\boldsymbol{\gamma} = 0$ is convex. The discussion in this section focuses on the half-complementarity formulation (2.3) and omits the results for the full-complementarity formulation of the problem (2.1).

The discussion in the next several subsections proceeds as follows. We begin with a reformulation of (2.3) as a nonlinear program with "piecewise structures" [32], which offers a global perspective of this nonconvex program. Next, we turn to the local properties of the problem, establishing the constant rank constraint qualification (CRCQ) [16, page 262] [25] of the problem under a constant rank condition on the constraint functions, which naturally holds when the latter functions are affine. With the CRCQ in place, we then present the Karush-Kuhn-Tucker (KKT) conditions of (2.3) and relate them to the KKT conditions of the "pieces" of the problem. We also briefly consider second-order optimality results for these problems. In Section 4, we undertake a similar analysis of a relaxed formulation of (2.3). Incidentally, beginning with Scholtes [33], there has been an extensive literature on regularization methods for general MPCCs; a recent reference is [27] which contains many related references on this topic.

3.1 Piecewise formulation

While the computation of a globally optimal solution to the ℓ_0-norm minimization problem (2.1) is practically very difficult, we describe a piecewise property of this problem and identify its pieces. For any index set $\mathcal{J} \subseteq \{1, \cdots, n\}$ with complement $\mathcal{J}^{\mathrm{c}}$, consider the nonlinear program

$$\begin{array}{lll} \underset{x}{\text{minimize}} & f(x) & \\ \text{subject to} & c_i(x) = 0, & i \in \mathcal{E} \\ & c_i(x) \leq 0, & i \in \mathcal{I} \\ \text{and} & x_i = 0, & i \in \mathcal{J}, \end{array} \tag{3.1}$$

which may be thought of as a "piece" of (2.1) in the sense of piecewise programming. Indeed, provided that (2.1) is feasible, we have

$$\boxed{\infty > \text{ minimum of (2.1)} = \underset{\mathcal{J}}{\text{minimum}} \{ \text{ minimum of (3.1)} + \gamma \, | \, \mathcal{J}^{\mathrm{c}} \, | \} \geq -\infty,} \tag{3.2}$$

where the value of $-\infty$ is allowed in both the left- and right-hand sides. [We adopt the convention that the minimum value of an infeasible optimization problem is taken to be ∞.] To prove (3.2), we note that the left-hand minimum is always an upper bound of the right-hand minimum. Let $\mathcal{N}(x)$ be the support of the vector x, with complement $\mathcal{N}(x)^{\rm c}$ in $\{1, \cdots, n\}$. It follows that, for any feasible x of (2.1), we have, with $\mathcal{J} = \mathcal{N}(x)^{\rm c}$,

$$f(x) + \gamma \| x \|_0 \;=\; f(x) + \gamma \, | \, \mathcal{J}^{\rm c} \, | \;\geq\; \{ \text{ minimum of (3.1) with } \mathcal{J} = \mathcal{N}(x)^{\rm c} \} + \gamma \, | \, \mathcal{J}^{\rm c} \, |.$$

This bound establishes the equality of the two minima in (3.2) when the left-hand minimum is equal to $-\infty$. A moment's thought shows that these two minima are also equal when the right-hand minimum is equal to $-\infty$. Thus it remains to consider the case where both the left and right minima are finite. Let $x^{\mathcal{J}}$ be an optimal solution of (3.1) that attains the right-hand minimum in (3.2). We have

$$\text{minimum of (2.1)} \;\leq\; f(x^{\mathcal{J}}) + \gamma \| x^{\mathcal{J}} \|_0 \;\leq\; f(x^{\mathcal{J}}) + \gamma \, | \, \mathcal{J}^{\rm c} \, |,$$

establishing the equality (3.2).

3.2 Constraint qualifications

As the constraints of (2.3) are nonlinear, it is important that they satisfy some constraint qualification in order to gain some understanding of the optimality properties of the problem. Let x be a feasible solution to (2.1). We wish to postulate a constraint qualification at x with respect to (2.1) under which the CRCQ will hold for the constraints of (2.3) at (x, ξ), where ξ is defined by (1.4). For this purpose, we introduce the index set

$$\mathcal{A}(x) \;\triangleq\; \{ \, i \in \mathcal{I} \;\mid\; c_i(x) \;=\; 0 \, \}.$$

The gradients of the active constraints in (2.3) at the pair (x, ξ) are of several kinds:

$$\left\{ \begin{pmatrix} \nabla c_i(x) \\ 0 \end{pmatrix} : i \in \mathcal{E} \cup \mathcal{A}(x) \right\}, \left\{ - \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \mathcal{N}(x) \right\}, \left\{ \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \mathcal{N}(x)^{\rm c} \right\}$$
$$\text{and} \underbrace{\left\{ x_i \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \mathcal{N}(x) \right\}, \left\{ \xi_i \begin{pmatrix} e_i \\ 0 \end{pmatrix} : i \in \mathcal{N}(x)^{\rm c} \right\}}_{\text{gradients of the equality constraint } \xi \circ x = 0}$$

where e_i is the n-vector of zeros except for a 1 in the ith position. We assume that for every index set $\alpha \subseteq \mathcal{A}(x)$ the family of vectors

$$\left\{ \left(\frac{\partial c_i(x')}{\partial x_j} \right)_{j \in \mathcal{N}(x)} : i \in \mathcal{E} \cup \alpha \right\} \tag{3.3}$$

has the same rank for all vectors x' sufficiently close to x that are also feasible to (2.1). Each vector (3.3) is a subvector of the gradient vector $\nabla c_i(x')$ with the partial derivatives $\partial c_i(x')/\partial x_j$ for $j \in \mathcal{N}(x)^{\rm c}$ removed. If this assumption holds at x, the CRCQ is valid for the constraints of the problem (2.3) at the pair (x, ξ). To show this, it suffices to verify that

for any index sets $\alpha \subseteq \mathcal{A}(x)$, $\beta_1, \gamma_1 \subseteq \mathcal{N}(x)$ and $\beta_2, \gamma_2 \subseteq \mathcal{N}(x)^{\mathrm{c}}$, the family of vectors:

$$\left\{ \begin{pmatrix} \nabla c_i(x') \\ 0 \end{pmatrix} : i \in \mathcal{E} \cup \alpha \right\}, \quad \left\{ - \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \beta_1 \right\}, \quad \left\{ \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \beta_2 \right\}$$
$$\text{and} \quad \underbrace{\left\{ x_i' \begin{pmatrix} 0 \\ e_i \end{pmatrix} : i \in \gamma_1 \right\}, \quad \left\{ \xi_i' \begin{pmatrix} e_i \\ 0 \end{pmatrix} : i \in \gamma_2 \right\}}_{\substack{\text{gradients of the equality constraint } \xi \circ x = 0 \\ \text{evaluated at } (x', \xi')}} \tag{3.4}$$

has the same rank for all pairs (x', ξ') sufficiently close to the given pair (x, ξ) that are also feasible to (2.3). Clearly, this assumption is satisfied when the constraint functions $c_i(x)$ are affine. Consider such a pair (x', ξ'). We must have $\mathcal{N}(x) \subseteq \mathcal{N}(x')$; moreover, if $i \in \mathcal{N}(x)^{\mathrm{c}}$, then $\xi_i = 1$; hence $\xi' > 0$. By complementarity, it follows that $i \in \mathcal{N}(x')^{\mathrm{c}}$. Consequently, $\mathcal{N}(x) = \mathcal{N}(x')$. This is sufficient to establish the rank invariance of the vectors (3.4) when the pair (x', ξ') varies near (x, ξ).

An immediate corollary of the above derivation is that if the constraints of (2.1) are all affine, then the CRCQ holds for the constraints of (2.3) at (x, ξ) for any x feasible to (2.1) and ξ defined by (1.4).

3.3 KKT conditions and local optimality

With the CRCQ in place, it follows that the KKT conditions are necessary for a pair (x, ξ) to be optimal to (2.3), provided that ξ is defined by (1.4). Letting λ, η, and μ be the multipliers to the constraints of (2.3), the KKT conditions of this problem are:

$$\begin{aligned} 0 &= \nabla f(x) + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(x) + \mu \circ \xi \\ 0 &\le \xi \perp -\boldsymbol{\gamma} + \mu \circ x + \eta \ge 0, \quad 0 = \xi \circ x \\ 0 &\le \eta \perp \mathbf{1}_n - \xi \ge 0 \\ 0 &= c_i(x), \qquad\qquad i \in \mathcal{E} \\ 0 &\le \lambda_i \perp c_i(x) \le 0, \qquad i \in \mathcal{I} \end{aligned} \tag{3.5}$$

We wish to compare the above KKT conditions with those of the pieces of (2.1) exemplified by (3.1) for an index subset $\mathcal{J}$ of $\{1, \cdots, n\}$. Letting λ denote the multipliers of the functional constraints in (3.1), we can write the KKT conditions of (3.1) as

$$\begin{aligned} 0 &= \frac{\partial f(x)}{\partial x_j} + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j}, \quad j \in \mathcal{J}^{\mathrm{c}} \\ 0 &= c_i(x), \qquad\qquad i \in \mathcal{E} \\ 0 &\le \lambda_i \perp c_i(x) \le 0, \qquad i \in \mathcal{I} \\ 0 &= x_j, \qquad\qquad j \in \mathcal{J}. \end{aligned} \tag{3.6}$$

We have the following result connecting the KKT systems (3.5) and (3.6), which can be contrasted with the equality (3.2) that deals with the global minima of these two problems.

Proposition 3.1. *Let x be a feasible solution of (2.1) and let ξ be defined by (1.4). The following three statements hold for any positive vector $\boldsymbol{\gamma}$:*

(a) *If (x, ξ) is a KKT point of (2.3) with multipliers λ, μ, and η, then x is a KKT point of (3.1) for any $\mathcal{J}$ satisfying $\mathcal{N}(x) \subseteq \mathcal{J}^{\rm c} \subseteq \mathcal{N}(x) \cup \{j \mid \mu_j = 0\}$.*

(b) *Conversely, if x is a KKT point of (3.1) for some $\mathcal{J}$, then (x, ξ) is a KKT point of (2.3).*

(c) *If x is a local minimum of (3.1) for some $\mathcal{J}$, then (x, ξ) is a local minimum of (2.3).*

Proof. To prove (a), it suffices to note that for such an index set $\mathcal{J}$, we must have $\mathcal{J} \subseteq \mathcal{N}(x)^{\rm c}$; moreover, if $j \in \mathcal{J}^{\rm c}$, then either $\mu_j = 0$ or $\xi_j = 0$. To prove part (b), it suffices to define the multipliers μ and η. An index set $\mathcal{J}$ for which (3.6) holds must be a subset of $\mathcal{N}(x)^{\rm c}$ so that $\mathcal{N}(x) \subseteq \mathcal{J}^{\rm c}$. Let

$$\mu_j \triangleq \begin{cases} \dfrac{\gamma_j}{x_j} & \text{if } j \in \mathcal{N}(x) \\ -\left[\dfrac{\partial f(x)}{\partial x_j} + \displaystyle\sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \dfrac{\partial c_i(x)}{\partial x_j} \right] & \text{if } j \in \mathcal{J} \\ 0 & \text{if } j \in \mathcal{J}^{\rm c} \cap \mathcal{N}(x)^{\rm c} \end{cases}$$

$$\eta_j \triangleq \begin{cases} 0 & \text{if } j \in \mathcal{N}(x) \\ \gamma_j & \text{if } j \in \mathcal{J} \\ \gamma_j & \text{if } j \in \mathcal{J}^{\rm c} \cap \mathcal{N}(x)^{\rm c}. \end{cases}$$

It is not difficult to verify that the KKT conditions (3.5) hold at the triple $(x, \xi, \lambda, \mu, \eta)$.

Finally, to prove (c), let (x', ξ') be a feasible pair of (2.3) sufficiently close to (x, ξ). Since $x_j = 0$ for all $j \in \mathcal{J}$, it follows that $\xi_j = 1$ for all such j. Since ξ'_j is sufficiently close to ξ_j, we deduce $\xi'_j > 0$; hence $x'_j = 0$ by complementarity. Thus x' is feasible to (3.1). Moreover, if $x_j \neq 0$, then $x'_j \neq 0$; hence $\xi'_j = 0$.

$$\begin{aligned} & f(x') + \boldsymbol{\gamma}^T (\, \mathbf{1}_n - \xi' \,) \\ & \quad \geq f(x) + \sum_{j \,:\, x_j = 0} \gamma_j \, \left(1 - \xi'_j \right) + \sum_{j \,:\, x_j \neq 0} \gamma_j \, \left(1 - \xi'_j \right) \\ & \quad \geq f(x) + \sum_{j \,:\, x_j = 0} \gamma_j \, (\, 1 - \xi_j \,) + \sum_{j \,:\, x_j \neq 0} \gamma_j \, (\, 1 - \xi_j \,) \, = \, f(x) + \boldsymbol{\gamma}^T (\, \mathbf{1}_n - \xi \,), \end{aligned}$$

establishing the desired conclusion. □

The next proposition states a result for the inner-product reformulations of the full-complementarity formulation (2.2). Similar statements hold for the componentwise formulation, which are omitted.

Proposition 3.2. *Let x be a feasible solution of (2.1), let ξ be defined by (1.4), and $x^{\pm} \triangleq \max\{0, \pm x\}$. Then the following three statements hold for any positive vector $\boldsymbol{\gamma}$:*

(a) *If* $(x, x^{\pm}, \xi)$ *is a KKT point of*

$$\begin{array}{ll} \underset{x,\, x^{\pm},\, \xi}{\textit{minimize}} & f(x) + \boldsymbol{\gamma}^T (\, \mathbf{1}_n - \xi \,) \\ \textit{subject to} & c_i(x) \,=\, 0, \quad i \,\in\, \mathcal{E} \\ & c_i(x) \,\leq\, 0, \quad i \,\in\, \mathcal{I} \\ & x = x^+ - x^-, \quad \xi^T(x^+ + x^-) \leq 0, \quad (x^+)^T(x^-) \leq 0 \\ & 0 \,\leq\, \xi \,\leq\, \mathbf{1}_n \quad \textit{and} \quad x^{\pm} \geq 0, \end{array} \tag{3.7}$$

then x *is a KKT point of (3.1) for any* $\mathcal{J}$ *satisfying* $\mathcal{N}(x) \subseteq \mathcal{J}^{\rm c} \subseteq \mathcal{N}(x) \cup \{j \mid \mu_j = 0\}$.

(b) *Conversely, if* x *is a KKT point of (3.1) for some* $\mathcal{J}$*, then* $(x, x^{\pm}, \xi)$ *is a KKT point of (3.7).*

(c) *If* x *is a local minimum of (3.1) for some* $\mathcal{J}$*, then* $(x, x^{\pm}, \xi)$ *is a local minimum of (3.7).*

We now look at the second-order optimality conditions. In the proposition below, we examine the sufficient conditions; an analogous result can be derived for the second-order necessary conditions in a similar manner. Results similar to the following Proposition 3.3 and Corollary 3.4 hold for the full-complementarity and half-complementarity formulations.

Proposition 3.3. *Let* (x, ξ) *be a point so that* (x, ξ) *is a KKT point of (2.3) with multipliers* λ*,* μ*, and* η*, and so that* x *is a KKT point of (3.1) for any* $\mathcal{J}$ *satisfying* $\mathcal{N}(x) \subseteq \mathcal{J}^{\rm c} \subseteq \mathcal{N}(x) \cup \{j \mid \mu_j = 0\}$*. If the second-order sufficient condition holds at* x *of (3.1), then it holds at* (x, ξ) *of (2.3). In addition, if* $\mathcal{J}^{\rm c} = \mathcal{N}(x)$*, the converse holds.*

Proof. The second-order conditions examine directions d in the critical cone, i.e., those directions that satisfy the linearization of each equality constraint and active inequality constraints, and that keep active the linearization of any inequality constraint with a positive multiplier. From the KKT conditions (3.5) of (2.3), if $x_j = 0$ for some $j \in \{1, \ldots, n\}$ then $\xi_j = 1$ and the corresponding multiplier $\eta_j > 0$. If $x_j \neq 0$ then the linearization of the complementarity constraint restricts the direction. Thus, all directions $d = (d_x, d_\xi)$ in the critical cone satisfy

$$d_\xi = 0.$$

Therefore, we only need to consider the x part of the Hessian,

$$H = \nabla^2 f(x) + \sum_{i \,\in\, \mathcal{E} \,\cup\, \mathcal{I}} \lambda_i \nabla^2 c_i(x),$$

for both (2.3) and (3.1). Let $D_1 \subseteq \mathbb{R}^n$ be the set of directions d_x that satisfy

$$\begin{array}{ll} \nabla c_i(x)^T d_x = 0, & i \,\in\, \mathcal{E} \\ \nabla c_i(x)^T d_x = 0, & i \,\in\, \mathcal{I} \text{ with } \lambda_i > 0 \\ \nabla c_i(x)^T d_x \leq 0, & i \,\in\, \mathcal{I} \text{ with } \lambda_i = 0 \text{ and } c_i(x) = 0 \end{array} \tag{3.8}$$

together with $d_x \circ \xi = 0$. The second-order sufficient condition for (2.3) holds at x if and only if $d^T H d > 0$ for all $d \in D_1$ with $d \neq 0$. Similarly, let $D_2 \subseteq \mathbb{R}^n$ be the set of directions that satisfy (3.8) together with $(d_x)_j = 0$ for all $j \in \mathcal{J}$. Then the second-order sufficient condition for (3.1) holds at x if and only if $d^T H d > 0$ for all $d \in D_2$ with $d \neq 0$.

To prove the first part of the claim, we need to show that $D_1 \subseteq D_2$. Let $d^1 \in D_1$. Since $\mathcal{J} \subseteq \mathcal{N}(x)^c$, we have $\xi_j = 1$ for all $j \in \mathcal{J}$. Because the direction satisfies the linearization of the half-complementarity constraint, it follows that $d^1_j = 0$, which implies $d^1 \in D_2$. To prove the second part, let $d^2 \in D_2$. Since $\mathcal{N}(x)^c = \mathcal{J}$, we have $x_j = 0$ and hence $\xi_j = 1$ for all $j \in \mathcal{J}$. Further, $x_j \neq 0$ and $\xi_j = 0$ for all $j \in \mathcal{J}^c$. Thus $d^2 \in D_1$. □

Let us consider what these results imply for the simple model problem (1.1). It is easy to see that *any feasible point* x with $Ax \geq b$ and $Cx = d$ is a KKT point for (3.1) with $\mathcal{J} = \{j : x_j = 0\}$ and $f(x) = 0$. Propositions 3.1(b) and 3.3 then imply that x corresponds to a KKT point of (2.3) that satisfies the second-order necessary optimality conditions. In other words, finding a second-order necessary KKT point for (2.3) merely implies that we found a *feasible* point for (1.1), but this says nothing about its ℓ_0-norm. We summarize this observation in the following corollary.

Corollary 3.4. *A vector $\widehat{x}$ is feasible to the system $Ax \geq b$ and $Cx = d$ if and only if $(\widehat{x}, \xi)$, where $\mathbf{1}_n - \xi$ is the support vector of $\widehat{x}$, is a KKT pair of the nonlinear program* (2.3) *with $\mathcal{J} = \{j : \widehat{x}_j = 0\}$ that satisfies the second-order necessary optimality conditions.*

A principal goal in this study is assessing the adequacy of (local) NLP solvers for solving ℓ_0-norm minimization problems, such as (1.1) and (2.1), using the equivalent full- or half-complementarity reformulations. The results in this section cast a negative shadow on this approach. NLP solvers typically aim to find KKT points, ideally those that satisfy second-order optimality conditions. Propositions 3.1–3.3 establish that the reformulations for (2.1) may have an exponential number of points (one for each set $\mathcal{J} \subseteq \{1, \ldots, n\}$ in (3.1)), to which the NLP solvers might be attracted to. [This conclusion is not too surprising in light of the piecewise structure of these problems as seen from Subsection 3.1.] In the particular case of the model problem (1.1), Corollary 3.4 paints an even more gloomy picture because *any* feasible point for (1.1) has the characteristics that an NLP solver looks for, and most of those points have sub-optimal objective function values. Interestingly, these theoretical reservations do not seem to materialize in practice. Our computational results attest that usually points of rather good objective values are returned by the NLP solvers. The discussions related to the relaxed formulations in Section 4 shed some light on this observation. In particular, convergent sequences of points satisfying the second order necessary conditions for a relaxed problem are shown to converge to locally optimal solutions on a piece, under certain assumptions; in the linear case (1.1) they converge to nondominated solutions, and the set of limit points is a subset of the set of possible solutions to a weighted ℓ_1-norm approach.

4 Relaxed Formulations

As mentioned at the end of the Introduction, in general, the exact reformulations of an MPCC result in NLPs that are not well-posed in the sense that the MFCQ does not hold at any feasible point. To overcome this shortcoming, relaxation schemes for MPCCs have been proposed [33, 13, 26], where the complementarity constraints are relaxed. The resulting NLPs have better properties, and the solution of the original MPCC is obtained by solving a sequence of relaxed problems, for which the relaxation parameter is driven to zero. In this section, we investigate the stationarity properties of relaxed reformulations for the ℓ_0-norm minimization problem.

We introduce the following relaxation of the new half-complementarity formulation (2.3), which we denote by NLP(ε), for a given relaxation scalar $\varepsilon > 0$:

$$\begin{array}{lll} \underset{x,\xi}{\text{minimize}} & f(x) + \boldsymbol{\gamma}^T (\, \mathbf{1}_n - \xi \,) & \\ \text{subject to} & c_i(x) \,=\, 0, \quad i \,\in\, \mathcal{E} & (\lambda_{\mathcal{E}}) \\ & c_i(x) \,\leq\, 0, \quad i \,\in\, \mathcal{I} & (\lambda_{\mathcal{I}}) \\ & \xi \,\leq\, \mathbf{1}_n & (\eta) \\ & \xi \circ x \,\leq\, \varepsilon\, \mathbf{1}_n & (\mu^+) \\ & -\xi \circ x \,\leq\, \varepsilon\, \mathbf{1}_n & (\mu^-) \\ \text{and} & \xi \geq 0, & \end{array} \tag{4.1}$$

where λ, η, and $\mu^{\pm}$ are the associated multipliers for the respective constraints. The problem NLP(0) is the original half-complementary formulation (2.3). In essence, we wish to examine the limiting properties of the NLP(ε) as $\varepsilon \downarrow 0$. Observations analogous to those in the following subsections are valid for relaxations of the full complementarity formulations (3.7).

Similarly to Subsection 3.2, we give a sufficient condition for the Abadie constraint qualification (ACQ) to hold for (4.1) at a feasible pair $(\bar{x}, \bar{\xi})$, i.e., for the linearization cone of the constraints of (4.1) at this pair to equal the tangent cone. Explicitly stated, this CQ stipulates that if $(dx, d\xi)$ is a pair of vectors such that

$$\begin{array}{ll} \nabla c_i(x)^T dx \,=\, 0 & \forall\, i \,\in \mathcal{E} \\ \nabla c_i(x)^T dx \,\leq\, 0 & \forall\, i \,\in \mathcal{I} \text{ such that } c_i(\bar{x}) = 0 \\ d\xi_i \,\leq\, 0 & \forall\, i \text{ such that } \bar{\xi}_i = 1 \\ d\xi_i \,\geq\, 0 & \forall\, i \text{ such that } \bar{\xi}_i = 0 \\ \bar{x}_i\, d\xi_i + \bar{\xi}_i\, dx_i & \left\{ \begin{array}{ll} \leq 0 & \text{if } \bar{x}_i \bar{\xi}_i \,=\, \varepsilon \\ \geq 0 & \text{if } \bar{x}_i \bar{\xi}_i \,=\, -\varepsilon, \end{array} \right. \end{array} \tag{4.2}$$

then there exist a sequence of pairs $\{(x^k, \xi^k)\}$ such that each (x^k, ξ^k) is feasible to (4.1) and a sequence of positive scalars $\{\tau_k\}$ tending to zero such that

$$dx \,=\, \lim_{k\to\infty} \frac{x^k - \bar{x}}{\tau_k} \quad \text{and} \quad d\xi \,=\, \lim_{k\to\infty} \frac{\xi^k - \bar{\xi}}{\tau_k}\,.$$

The sufficient condition that we postulate is on the functional constraints

$$\begin{array}{ll} c_i(x) \,=\, 0 & i \,\in\, \mathcal{E} \\ c_i(x) \,\leq\, 0 & i \,\in\, \mathcal{I} \end{array} \tag{4.3}$$

at the given $\bar{x}$. Namely, the linearization cone of the constraints (4.3) at $\bar{x}$ is equal to the tangent cone; i.e., for every dx satisfying the first two conditions in (4.2), there exists a sequence $\{x^k\}$ of vectors converging to $\bar{x}$ and a sequence of positive scalars $\{\tau_k\}$ converging to zero such that

$$dx \,=\, \lim_{k\to\infty} \frac{x^k - \bar{x}}{\tau_k}\,;$$

equivalently, for some sequence $\{e^k\}$ of vectors satisfying $\lim_{k\to\infty} \dfrac{e^k}{\tau_k} = 0$, we have $x^k \triangleq \bar{x} + \tau_k\, dx + e^k$ satisfies (4.3) for all k. This by itself is a CQ on these functional constraints that is naturally satisfied if such constraints are affine.

Theorem 4.1. *The Abadie constraint qualification holds for NLP(ε) at the feasible pair $(\bar{x}, \bar{\xi})$ if the linearization cone of the constraints (4.3) at $\bar{x}$ is equal to the tangent cone.*

Proof. For any feasible solution to a nonlinear program, the tangent cone is a subset of the linearization cone [29, page 117]; we show the reverse also holds under the assumptions of the theorem. Let $d\xi$ satisfy the remaining three conditions in (4.2). We claim that the pair $(dx, d\xi)$ is in the tangent cone of the constraints in the relaxed formulation (4.1) at $(\bar{x}, \bar{\xi})$. It suffices to show that there exists a sequence $\{\eta^k\}$ such that $\lim_{k\to\infty} \frac{\eta^k}{\tau_k} = 0$ and $\widehat{\xi}^k \triangleq \bar{\xi} + \tau_k\, d\xi + \eta^k$ satisfies

$$0 \le \widehat{\xi}_i^k \le 1 \text{ and } |\,(\,\bar{x}_i + \tau_k\, dx_i \,+\, e_i^k\,)\,(\,\bar{\xi}_i + \tau_k\, d\xi_i \,+\, \eta_i^k\,)\,| \,\le\, \varepsilon, \quad \forall\, i.$$

For a component i such that $|\bar{x}_i\bar{\xi}_i| < \varepsilon$, it suffices to choose $\eta_i^k = 0$. Consider a component i for which $\varepsilon = |\bar{x}_i\bar{\xi}_i|$. We consider only the case where $\bar{x}_i\bar{\xi}_i = \varepsilon$ and leave the other case to the reader. Thus, both $\bar{x}_i$ and $\bar{\xi}_i$ must be positive; hence so are both $x_i^k \triangleq \bar{x}_i + \tau_k\, dx_i + e_i^k$ and $\widehat{\xi}_i^k \triangleq \bar{\xi}_i + \tau_k\, d\xi_i + \eta_i^k$ for all k sufficiently large. It remains to show that we can choose η_i^k so that $\widehat{\xi}_i^k \le 1$ and $x_i^k\xi_i^k \le \varepsilon$ for all k sufficiently large.

We first derive an inequality on some of the terms in the product $x_i^k\xi_i^k$. We show

$$(\,\bar{x}_i + \tau_k\, dx_i\,)\,(\,\bar{\xi}_i + \tau_k\, d\xi_i\,) \,=\, \varepsilon + \tau_k\, \left[\,(\,\bar{x}_i\, d\xi_i + \bar{\xi}_i\, dx_i\,) + \tau_k\, dx_i\, d\xi_i\,\right] \,\le\, \varepsilon.$$

In particular, since $\bar{x}_i d\xi_i + \bar{\xi}_i dx_i \le 0$, there are two cases to consider. If $\bar{x}_i d\xi_i + \bar{\xi}_i dx_i = 0$, then $dx_i d\xi_i \le 0$ and the claim holds for all $\tau_k \ge 0$. If $\bar{x}_i d\xi_i + \bar{\xi}_i dx_i < 0$, then $\bar{x}_i d\xi_i + \bar{\xi}_i dx_i + \tau_k dx_i d\xi_i < 0$ for all $\tau_k > 0$ sufficiently small.

We can now choose $\eta_i^k \triangleq -\alpha|e_i^k|$, where $\alpha > 0$ will be determined from the following derivation. With this choice, we easily have $\widehat{\xi}_i^k \le 1$ for all k sufficiently large. Furthermore,

$$\begin{aligned} x_i^k\,\xi_i^k \;&=\; (\,\bar{x}_i + \tau_k\, dx_i + e_i^k\,)\,(\,\bar{\xi}_i + \tau_k\, d\xi_i \,+ \eta_i^k\,) \\ &=\; \varepsilon + \tau_k\, \underbrace{\left[\,(\,\bar{x}_i\, d\xi_i + \bar{\xi}_i\, dx_i\,) + \tau_k\, dx_i\, d\xi_i\,\right]}_{\le\, 0 \text{ as above}} \\ &\quad + e_i^k \left(\underbrace{\bar{\xi}_i + \tau_k\, d\xi_i}_{\text{positive}} \right) - \alpha\,|\,e_i^k\,| \left(\underbrace{\bar{x}_i + \tau_k\, dx_i}_{\text{positive}} \right) - \alpha\, e_i^k\,|\,e_i^k\,|. \end{aligned}$$

It is now clear that we may choose $\alpha > 0$ so that $x_i^k\xi_i^k \le \varepsilon$ for all k sufficiently large. □

4.1 Convergence of KKT points for NLP(ε)

In this subsection we examine the limiting behavior of KKT points $x(\varepsilon)$ for NLP(ε) as ε converges to zero. This is of interest because algorithms based on the sequential solution of the relaxation NLP(ε) aim to compute limit points of $x(\varepsilon)$ [13, 26, 33]. However, our analysis also gives insight into the behavior of standard NLP solvers that are applied directly to one of the unrelaxed NLP-reformulations of the MPCC, such as (2.3) and (3.7). For instance, some implementations of NLP solvers, such as the IPOPT solver [35] used for the numerical experiments in Section 5, relax all inequality and bound constraints by a small amount that

is related to the convergence tolerance before solving the problem at hand. This modification is done in order to make the problem somewhat "nicer"; for example, a feasible problem is then guaranteed to have a nonempty relative interior of the feasible region. However, because this alteration is on the order of the user-specified convergence tolerance, it usually does not lead to solutions that are far away from solutions of the original unperturbed problem. In the current context this means that such an NLP code solves the relaxation NLP(ε) even if the relaxation is not explicitly introduced by the user.

With a CQ in place, we may write down the KKT conditions for NLP(ε):

$$0 = \nabla f(x) + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(x) + (\mu^+ - \mu^-) \circ \xi \tag{4.4a}$$

$$0 \le \xi \perp -\boldsymbol{\gamma} + (\mu^+ - \mu^-) \circ x + \eta \ge 0 \tag{4.4b}$$

$$0 \le \eta \perp \mathbf{1}_n - \xi \ge 0 \tag{4.4c}$$

$$0 = c_i(x), \qquad i \in \mathcal{E} \tag{4.4d}$$

$$0 \le \lambda_i \perp c_i(x) \le 0, \qquad i \in \mathcal{I} \tag{4.4e}$$

$$0 \le \mu^+ \perp \varepsilon \mathbf{1}_n - \xi \circ x \ge 0 \tag{4.4f}$$

$$0 \le \mu^- \perp \varepsilon \mathbf{1}_n + \xi \circ x \ge 0. \tag{4.4g}$$

We may draw the following observations from the above conditions:

• For $j \in \{1, \ldots, n\}$ with $x_j > \varepsilon > 0$, by (4.4g) we have $\mu_j^- = 0$, and by (4.4f) we have $\xi_j < 1$. It follows from (4.4c) that $\eta_j = 0$ and then (4.4b) and (4.4a) give the relationships:

$$\xi_j = \frac{\varepsilon}{x_j} < 1, \quad \eta_j = 0, \quad \mu_j^+ = \frac{\gamma_j}{x_j}, \quad \mu_j^- = 0, \quad \frac{\partial f(x)}{\partial x_j} + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j} = -\frac{\varepsilon \gamma_j}{x_j^2} < 0. \tag{4.5}$$

• For $j \in \{1, \ldots, n\}$ with $x_j < -\varepsilon < 0$, by (4.4f) we have $\mu_j^+ = 0$, and by (4.4g) we have $\xi_j < 1$. It follows from (4.4c) that $\eta_j = 0$ and then (4.4b) and (4.4a) give the relationships:

$$\xi_j = \frac{\varepsilon}{-x_j} < 1, \quad \eta_j = 0, \quad \mu_j^+ = 0, \quad \mu_j^- = \frac{\gamma_j}{-x_j}, \quad \frac{\partial f(x)}{\partial x_j} + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j} = \frac{\varepsilon \gamma_j}{x_j^2} > 0. \tag{4.6}$$

• For $j \in \{1, \ldots, n\}$ with $-\varepsilon < x_j < \varepsilon$, (4.4f) and (4.4g) give $\mu_j^+ = \mu_j^- = 0$. Then (4.4b) implies $\eta_j > 0$, giving $\xi_j = 1$ by (4.4c) and so $\eta_j = \gamma_j$ by (4.4b). Together with (4.4a), this overall implies

$$\xi_j = 1, \quad \eta_j = \gamma_j, \quad \mu_j^+ = 0, \quad \mu_j^- = 0, \quad \frac{\partial f(x)}{\partial x_j} + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j} = 0. \tag{4.7}$$

• For $j \in \{1, \ldots, n\}$ with $x_j = \varepsilon$ we have $\mu_j^- = 0$ from (4.4g). Also, we must have $\xi_j = 1$ since otherwise $\eta_j = 0$ from (4.4c) and $\mu_j^+ = 0$ from (4.4f), which would then violate (4.4b). Thus from (4.4b) and (4.4a) we have

$$\xi_j = 1, \quad \eta_j = \gamma_j - \varepsilon \mu_j^+, \quad 0 \le \mu_j^+ \le \frac{\gamma_j}{\varepsilon}, \quad \mu_j^- = 0, \quad \frac{\partial f(x)}{\partial x_j} + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j} + \mu_j^+ = 0. \tag{4.8}$$

• For $j \in \{1, \ldots, n\}$ with $x_j = -\varepsilon$ we have $\mu_j^+ = 0$ from (4.4f). Also, we must have $\xi_j = 1$ since otherwise $\eta_j = 0$ from (4.4c) and $\mu_j^- = 0$ from (4.4g), which would then violate (4.4b). Thus from (4.4b) and (4.4a) we have

$$\begin{aligned} \xi_j \;=\; 1, \quad \eta_j \;=\; \gamma_j - \varepsilon\mu_j^-, \quad \mu_j^+ \;=\; 0, \quad 0 \;\leq\; \mu_j^- \;\leq\; \frac{\gamma_j}{\varepsilon}, \\ \frac{\partial f(x)}{\partial x_j} + \sum_{i \,\in\, \mathcal{E} \,\cup\, \mathcal{I}} \lambda_i \, \frac{\partial c_i(x)}{\partial x_j} - \mu_j^- \;=\; 0. \end{aligned} \tag{4.9}$$

We show that any limit of a subsequence of KKT points to (4.1) is a KKT point to the problem (3.1) for a particular $\mathcal{J}$, under certain assumptions.

Theorem 4.2. *Let $(x(\varepsilon), \xi(\varepsilon))$ be a KKT point for for NLP(ε) with multipliers $(\lambda(\varepsilon), \eta(\varepsilon), \mu(\varepsilon))$. Let (x^*, λ^*) be a limit of a subsequence of $(x(\varepsilon), \lambda(\varepsilon))$ as $\varepsilon \downarrow 0$. Assume $f(x)$ and each $c_i(x)$, $i \in \mathcal{E} \cup \mathcal{I}$, is Lipschitz continuous. Let $\mathcal{J} = \mathcal{N}(x^*)^c$. Then (x^*, λ^*) is a KKT point for (3.1).*

Proof. From observations (4.5) and (4.6), we have for components j with $x_j^* \neq 0$ that

$$\frac{\partial f(x(\varepsilon))}{\partial x_j} + \sum_{i \,\in\, \mathcal{E} \,\cup\, \mathcal{I}} \lambda_i \, \frac{\partial c_i(x(\varepsilon))}{\partial x_j} \;=\; -\frac{\varepsilon\gamma_j}{x(\varepsilon)_j^2} \;\to\; 0.$$

For the remaining components, we can choose the multiplier for the constraint $x_j = 0$ to be equal to $\lim_{\varepsilon \to 0} \mu_j^+(\varepsilon) - \mu_j^-(\varepsilon)$, which exists from the Lipschitz continuity assumption and observations (4.8) and (4.9). Thus, the KKT gradient condition holds for (3.1). The remaining KKT conditions hold from continuity of the functions. □

4.2 Second-order conditions

We analyze the second-order necessary (sufficient) conditions of the relaxed NLP(ε). Such a necessary (sufficient) condition stipulates that the Hessian of the Lagrangian function with respect to the primary variables (i.e., (x, ξ)) is (strictly) copositive on the critical cone of the feasible set; that is to say, at a feasible pair $(\bar{x}, \bar{\xi})$, if $(dx, d\xi)$ is a pair satisfying (4.2) and $\nabla f(\bar{x})^T dx - \boldsymbol{\gamma}^T d\xi = 0$, then for all (some) multipliers $(\lambda, \eta, \mu^{\pm})$ satisfying the KKT conditions (4.4),

$$\begin{pmatrix} dx \\ d\xi \end{pmatrix}^T \begin{bmatrix} \nabla^2 f(\bar{x}) + \sum\limits_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \, \nabla c_i(\bar{x})^2 & \mathrm{Diag}(\mu^+ - \mu^-) \\ \mathrm{Diag}(\mu^+ - \mu^-) & 0 \end{bmatrix} \begin{pmatrix} dx \\ d\xi \end{pmatrix} \;\geq\; (>)\, 0,$$

or equivalently, (focusing only on the necessary conditions),

$$dx^T \left[\nabla^2 f(\bar{x}) + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \, \nabla^2 c_i(\bar{x}) \right] dx + \sum_{i=1}^{n} \left(\mu_i^+ - \mu_i^- \right) dx_i \, d\xi_i \;\geq\; 0. \tag{4.10}$$

Taking into account the KKT conditions (4.4), the pair $(dx, d\xi)$ satisfies the following conditions (see [16, Lemma 3.3.2]):

$$\begin{array}{ll}
\nabla c_i(\bar{x})^T dx \;=\; 0 & \forall\, i \in \mathcal{E} \\
\nabla c_i(\bar{x})^T dx \;\le\; 0 & \forall\, i \in \mathcal{I} \text{ such that } c_i(\bar{x}) = 0 \\
\nabla c_i(\bar{x})^T dx \;=\; 0 & \forall\, i \in \mathcal{I} \text{ such that } c_i(\bar{x}) = 0 < \lambda_i \\
d\xi_i \;\le\; 0 & \forall\, i \text{ such that } \bar{\xi}_i = 1 \\
d\xi_i \;=\; 0 & \forall\, i \text{ such that } 1 - \bar{\xi}_i = 0 < \eta_i \\
d\xi_i \;\ge\; 0 & \forall\, i \text{ such that } \bar{\xi}_i = 0 \\
d\xi_i \;=\; 0 & \forall\, i \text{ such that } \bar{\xi}_i = 0 < -\gamma_i + (\mu_i^+ - \mu_i^-)\bar{x}_i \\
\bar{x}_i\, d\xi_i + \bar{\xi}_i\, dx_i & \begin{cases} \le 0 & \text{if } \bar{x}_i\, \bar{\xi}_i \;=\; \varepsilon \\ = 0 & \text{if } \varepsilon - \bar{x}_i\, \bar{\xi}_i \;=\; 0 \;<\; \mu_i^+ \\ = 0 & \text{if } \varepsilon + \bar{x}_i\, \bar{\xi}_i \;=\; 0 \;<\; \mu_i^- \\ = 0 & \text{if } \bar{x}_i\, \bar{\xi}_i \;=\; -\varepsilon \end{cases}
\end{array} \tag{4.11}$$

Note from (4.4f) that $\mu_i^+ > 0$ only if $\bar{x}_i \ge \varepsilon$ and $\bar{x}_i\bar{\xi}_i = \varepsilon$, in which case dx_i and $d\xi_i$ cannot have the same sign. Similarly from (4.4g), $\mu_i^- > 0$ only if $\bar{x}_i \le -\varepsilon$ and $\bar{x}_i\bar{\xi}_i = -\varepsilon$, in which case dx_i and $d\xi_i$ cannot have opposite signs. Then (4.10) becomes:

$$dx^T \left[\nabla^2 f(\bar{x}) + \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i\, \nabla^2 c_i(\bar{x}) \right] dx + \sum_{i : \mu_i^+ > 0} \mu_i^+ \underbrace{dx_i\, d\xi_i}_{\le\, 0} - \sum_{i : \mu_i^- > 0} \mu_i^- \underbrace{dx_i\, d\xi_i}_{\ge\, 0} \ge 0. \tag{4.12}$$

It follows that if $dx_i d\xi_i \ne 0$ for some i such that $\mu_i^+ > 0$ or $\mu_i^- > 0$, then the above inequality, and thus the second-order necessary condition (SONC) for the NLP(ε) (which is a minimization problem) cannot hold, by simply scaling up $d\xi_i$ and fixing dx_i. Therefore, if this SONC holds, then we must have $dx_i d\xi_i = 0$ for all i for which $(\mu_i^+ + \mu_i^-) > 0$.

Summarizing the above discussion, we have proved the following result.

Proposition 4.3. *Let $(\bar{x}, \bar{\xi})$ be a feasible pair to the NLP(ε). Suppose that for all KKT multipliers $(\lambda, \eta, \mu^{\pm})$ satisfying (4.4),*

$$dx^T \left[\nabla^2 f(\bar{x}) + \sum_{i \in \mathcal{E}} \lambda_i\, \nabla^2 c_i(\bar{x}) + \sum_{i \in \mathcal{I} \text{ with } \lambda_i > 0} \lambda_i\, \nabla^2 c_i(\bar{x}) \right] dx \;\ge\; 0$$

for all vectors dx satisfying the first three conditions of (4.11). The SONC holds for the NLP(ε) if and only if for all critical pairs $(dx, d\xi)$ satisfying (4.11), $dx_i d\xi_i = 0$ for all i for which $(\mu_i^+ + \mu_i^-) > 0$.

4.3 Convergence of points satisfying SONC for NLP(ε)

For most of this subsection, we restrict our attention to problems of the form (1.1), which we write in the following form in order to simplify the notation:

$$\begin{array}{ll} \underset{x \,\in\, \mathbb{R}^n}{\text{minimize}} & \| x \|_0 \\ \text{subject to} & x \in X := \{x : Ax \ge b\}. \end{array} \tag{4.13}$$

The corresponding version of NLP(ε) can be written

$$\begin{array}{lll} \underset{x,\,\xi}{\text{minimize}} & \gamma^T(\,\mathbf{1}_n - \xi\,) & \\ \text{subject to} & b - Ax \le 0 & (\lambda) \\ & \xi \le \mathbf{1}_n & (\eta) \\ & \xi \circ x \le \varepsilon\,\mathbf{1}_n & (\mu^+) \\ & -\xi \circ x \le \varepsilon\,\mathbf{1}_n & (\mu^-) \\ \text{and} & \xi \ge 0, & \end{array} \tag{4.14}$$

It follows from Theorem 4.1 that the Abadie CQ holds for (4.14). Further, the same argument as in Section 3.2 can be used to show that the constant rank constraint qualification holds for (4.14). It is proved in [1] that if the CRCQ holds then any local minimizer must satisfy the second order necessary conditions. Hence, any local minimizer to (4.14) must satisfy the second-order necessary conditions. In this subsection, we investigate limits of local minimizers to (4.14) as $\varepsilon \downarrow 0$.

Theorem 4.4. *Any limit x^* of a subsequence of the x part of the locally optimal solutions $(x(\varepsilon), \xi(\varepsilon))$ to (4.14) must be an extreme point of the piece defined by x^*, namely*

$$\mathcal{P}(x^*) := \{x : Ax \ge b\} \cap \{x : x_i = 0 \text{ if } x_i^* = 0\}.$$

Proof. Assume x^* is not an extreme point of $\mathcal{P}(x^*)$, so $x^* \neq 0$ and there exists a feasible direction $dx \neq 0$ with $x^* \pm \alpha dx \in \mathcal{P}(x^*)$ for sufficiently small positive α. For sufficiently small ε, we have $(b - Ax(\varepsilon))_i = 0$ only if $(b - Ax^*)_i = 0$, so dx satisfies the first three conditions of (4.11). We now construct $d\xi(\varepsilon)$ to satisfy the remaining conditions of (4.11).

Let $\xi(\varepsilon)$ be the ξ part of the solution to (4.14) and let $(\lambda(\varepsilon), \eta(\varepsilon), \mu^{\pm}(\varepsilon)))$ be the corresponding KKT multipliers. For sufficiently small ε, $x_j(\varepsilon) > \varepsilon$ if $x_j^* > 0$ and $x_j(\varepsilon) < -\varepsilon$ if $x_j^* < 0$. It follows from the observations (4.5) and (4.6) that $\xi_j(\varepsilon) = \dfrac{\varepsilon}{x_j(\varepsilon)} < 1$ and $\mu_j^+(\varepsilon) > 0$ if $x_j^* > 0$, and $\xi_j(\varepsilon) = -\dfrac{\varepsilon}{x_j(\varepsilon)} < 1$ and $\mu_j^-(\varepsilon) > 0$ if $x_j^* < 0$ for sufficiently small ε.

The direction $d\xi(\varepsilon)$ is defined as

$$d\xi_j(\varepsilon) = \begin{cases} -\dfrac{\xi_j(\varepsilon)}{x_j(\varepsilon)}\,dx_j & \text{if } x_j^* \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

With this choice, (4.11) is satisfied while the left hand side of (4.10) evaluates to be negative. Hence, the point $(x(\varepsilon), \xi(\varepsilon))$ is not a local optimum to (4.14). □

The following corollary is immediate.

Corollary 4.5. *If the feasible region is contained within the nonnegative orthant then any limit x^* of a subsequence of locally optimal solutions to (4.14) must be an extreme point of $\{x : Ax \ge b\}$.*

In order to relate the solutions obtained as limits of solutions to NLP(ε) to solutions of weighted ℓ_1-norm minimization problems, we make the following definition. First, we introduce some notation: if $u \in \mathbb{R}^n$ then $|u|$ is the vector in $\mathbb{R}^n_+$ with $|u|_i = |u_i|$ for $i = 1, \ldots, n$.

Definition 4.6. A point $y \in S \subseteq \mathbb{R}^n$ is *dominated* in S if there exists a vector $\tilde{y} \in S$ with $|\tilde{y}| \neq |y|$ and $|\tilde{y}| \leq |y|$. Otherwise it is *nondominated* in S.

A point $\bar{x}$ is nondominated in X if and only if the optimal value of the following linear program is zero:

$$\begin{array}{ll} \underset{x,\, t \in \mathbb{R}^n}{\text{maximize}} & \mathbf{1}_n^T t \\ \text{subject to} & Ax \geq b, \quad t \geq 0 \\ & x + t \leq |\bar{x}| \quad \text{and} \quad x - t \geq -|\bar{x}| \end{array} \tag{4.16}$$

The following proposition relates the limiting points of subsequences of solutions to (4.14) to nondominated points.

Theorem 4.7. *Any limit x^* of a subsequence of the x-part of locally optimal solutions $(x(\varepsilon), \xi(\varepsilon))$ to (4.14) must be nondominated.*

Proof. We prove that if x^* is dominated then $(x(\varepsilon), \xi(\varepsilon))$ is not a local optimum for sufficiently small ε. Suppose there exists a feasible point $(\tilde{x}, \tilde{\xi})$ such that $|\tilde{x}| \leq |x^*|$, $\tilde{x} \neq x^*$. Set

$$dx = x^* - \tilde{x} \neq 0,$$

a feasible direction from $x(\varepsilon)$ in (4.14) since $\mathcal{A}(x(\varepsilon)) \subseteq \mathcal{A}(x^*)$ for sufficiently small ε. Note that $dx_i = 0$ if $x_i^* = 0$. For small positive α, a lower objective value can be achieved by setting $x = x(\varepsilon) + \alpha dx$ and updating the value of $\xi(\varepsilon)_j$ to be $\min\left\{1, \dfrac{\varepsilon}{|\, x(\varepsilon)_j + \alpha dx_j \,|}\right\}$, so $(x(\varepsilon), \xi(\varepsilon))$ is not a locally optimal point to (4.14). □

This theorem allows us to relate the set of solutions obtained as limits of sequences of locally optimal solutions to (4.14) to the solutions obtained using an ℓ_1-norm approach. A weighted ℓ_1-norm approach for finding a solution to (1.1) is to solve a problem of the form

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \displaystyle\sum_{j=1}^{n} w_j \,|\, x_j \,| \\ \text{subject to} & Ax \geq b \end{array} \tag{4.17}$$

for a positive weight vector $w \in \mathbb{R}^n$. Iterative reweighted ℓ_1 schemes [9] seek choices of weights w that lead to sparse solutions.

Proposition 4.8. *Any nondominated point satisfying $Ax \geq b$ is a solution to a weighted ℓ_1-norm minimization problem.*

Proof. Let $\bar{x}$ be a nondominated point in X. The dual of the linear program (4.16) is

$$\begin{array}{ll} \underset{y\, \lambda\, \pi}{\text{minimize}} & |\bar{x}|^T(\lambda + \pi) - b^T y \\ \text{subject to} & A^T y - \lambda + \pi = 0_n, \\ & \lambda + \pi \geq \mathbf{1}_n, \\ & y, \lambda, \pi \geq 0 \end{array} \tag{4.18}$$

The nondominance of $\bar{x}$ indicates that the optimal objective of (4.16) is 0. Let (λ^*, π^*, y^*) be an optimal solution to (4.18). Set $w = \lambda^* + \pi^* > 0$. Expressing (4.17) as a linear program and examining its dual shows that $\bar{x}$ is then optimal for (4.17), because the optimal value of (4.18) is zero.

□

Note that minimizing a weighted ℓ_1-norm minimization problem may give a solution that cannot be obtained as a limit of a subsequence of solutions to NLP(ε), so the latter set of solutions can be a proper subset of the set of solutions that can be obtained by solving (4.17). This is illustrated by the following example:

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & \| x \|_0 \\ \text{subject to} & l\,x_1 + x_2 \geq l\,p + 1 \\ & x_1 + p\,x_2 \geq 2\,p \end{array} \tag{4.19}$$

where l and p are positive parameters with $l\,p > 1$ and $p > 1$. This problem is solved by points of the form $(r, 0)$ and $(0, s)$ with $r \geq 2p$ and $s \geq lp+1$. The point $(p, 1)$ is a nondominated extreme point which does not solve (4.19). For any feasible $\hat{x} = (\hat{x}_1, \hat{x}_2)$ sufficiently close to $(p, 1)$ with $\hat{\xi} = \left(\dfrac{\varepsilon}{\hat{x}_1}, \dfrac{\varepsilon}{\hat{x}_2} \right)$, the direction $dx = (p, -1)$ and $d\xi = \varepsilon \left(-\dfrac{p}{\hat{x}_1^2}, \dfrac{1}{\hat{x}_2^2} \right)$ is a feasible improving direction for (4.14), provided $\gamma_1 < \gamma_2(p - \delta)$ for some positive parameter δ which determines the size of the neighborhood. Thus, such an $\hat{x}$ and $\hat{\xi}$ cannot be optimal to (4.14), and $x = (p, 1)$ cannot be a limit of solutions to (4.14). Nonetheless, $x = (p, 1)$ is optimal to a weighted ℓ_1-norm formulation

$$\begin{array}{ll} \underset{x \in \mathbb{R}^2}{\text{minimize}} & w_1|x_1| + w_2|x_2| \\ \text{subject to} & l\,x_1 + x_2 \geq l\,p + 1 \\ & x_1 + p\,x_2 \geq 2\,p \end{array} \tag{4.20}$$

provided

$$l > \frac{w_1}{w_2} > \frac{1}{p}, \text{ with } w_1, w_2 \geq 0.$$

As p and l increase, the point $(p, 1)$ becomes the optimal solution to (4.20) for more choices of w, and it is optimal in (4.14) for fewer choices of γ.

Returning to the general problem (2.1) and its relaxation NLP(ε), we prove a result regarding the limit of a sequence of points satisfying the second order necessary KKT conditions. This result is analogous to Theorem 4.7, although the earlier theorem is not implied by the result for the general problem.

Theorem 4.9. *Assume that $f(x)$ and each $c_i(x)$, $i \in \mathcal{E} \cup \mathcal{I}$, have continuous second derivatives. Let $(x(\varepsilon), \xi(\varepsilon))$ be a local minimizer for NLP(ε) that satisfies the second order necessary conditions with multipliers $(\lambda(\varepsilon), \eta(\varepsilon), \mu(\varepsilon))$. Let (x^*, λ^*) be a limit of a subsequence of $(x(\varepsilon), \lambda(\varepsilon))$ as $\varepsilon \downarrow 0$. Let $\mathcal{J} = \mathcal{N}(x^*)^c$. If the gradients of the constraints in $\mathcal{E} \cup \mathcal{A}(x^*)$ are linearly independent and if $\lambda_i^* \neq 0$ for all $i \in \mathcal{E} \cup \mathcal{A}(x^*)$ then x^* satisfies the second order necessary conditions for the problem (3.1).*

Proof. Assume the conclusion is false, so there is a direction $\widehat{dx}$ in the critical cone of (3.1) at x^* satisfying $\widehat{dx}^T \nabla^2_{xx} L(x^*, \lambda^*) \widehat{dx} < 0$. We construct a direction that shows that the second order necessary conditions are violated at $x(\varepsilon)$ for sufficiently small positive ε.

For ε sufficiently small, any constraint with $\lambda_i^* \neq 0$ must have $\lambda_i(\varepsilon) \neq 0$ so it must be active at $x(\varepsilon)$. In addition, no constraint in $\mathcal{I} \setminus \mathcal{A}(x^*)$ is active at $x(\varepsilon)$, for sufficiently small ε. For such ε, the gradients of the constraints in $\mathcal{E} \cup \mathcal{A}(x^*)$ are linearly independent at $x(\varepsilon)$ and close to their values at x^*, and the same set of constraints is active. Let the rows of the matrix $\hat{B}$ comprise the gradients of the active constraints at x^* and let $B(\varepsilon)$ be the analogous matrix at $x(\varepsilon)$, and let $M(\varepsilon) = B(\varepsilon) - \hat{B}$. Let $dx(\varepsilon)$ denote the projection of $\widehat{dx}$ onto the nullspace of $B(\varepsilon)$, so

$$\begin{aligned} dx(\varepsilon) &= \widehat{dx} - (\hat{B} + M(\varepsilon))^T \left((\hat{B} + M(\varepsilon))(\hat{B} + M(\varepsilon))^T \right)^{-1} (\hat{B} + M(\varepsilon))\widehat{dx} \\ &= \widehat{dx} - (\hat{B} + M(\varepsilon))^T \left((\hat{B} + M(\varepsilon))(\hat{B} + M(\varepsilon))^T \right)^{-1} M(\varepsilon)\widehat{dx}, \end{aligned}$$

which is well defined for sufficiently small ε since BB^T is positive definite. By continuity of the gradients, $||M(\varepsilon)|| \to 0$ as $\varepsilon \downarrow 0$, so $dx(\varepsilon) \to \widehat{dx}$ as $\varepsilon \downarrow 0$. Further, by continuity of the Hessians, we then have for sufficiently small ε that

$$dx(\varepsilon)^T \nabla^2_{xx} L(x(\varepsilon), \lambda(\varepsilon))\, dx(\varepsilon) \; < \; 0.$$

For sufficiently small ε, we have $|x_i(\varepsilon)| > \varepsilon$ if $x_i^* \neq 0$, so $0 < \xi_i(\varepsilon) < 1$ for these components. A direction $d\xi(\varepsilon)$ can be constructed from (4.15) (taking $dx = dx(\varepsilon)$) so that the direction $(dx(\varepsilon), d\xi(\varepsilon))$ is in the critical cone at $(x(\varepsilon), \xi(\varepsilon))$ for the problem NLP(ε) and satisfies (4.11). Therefore, from (4.12), this point violates the second order necessary conditions. □

We now use these observations to characterize the limit points of KKT points $x(\varepsilon)$ for NLP(ε) in a particular case. Consider the problem

$$\text{minimize} \quad ||\, x \,||_0 \quad \text{subject to} \quad x_1 + x_2 + x_3 \geq 1 \quad \text{and} \quad x \geq 0.$$

The relaxation of the corresponding half-complementarity formulation is

$$\begin{array}{lllll} \underset{x,\,\xi}{\text{minimize}} & \mathbf{1}_3^T\,(\,\mathbf{1}_3 - \xi\,) & & & \\ \text{subject to} & 1 - x_1 - x_2 - x_3 \leq 0, & (\lambda_0) & & \\ & -x_i \leq 0, & (\lambda_i) & \text{for } i = 1,2,3 & \\ & \xi \leq \mathbf{1}_3 & (\eta) & & \\ & \xi \circ x \leq \varepsilon \mathbf{1}_3 & (\mu^+) & & \\ \text{and} & \xi \geq 0, & & & \end{array} \tag{4.21}$$

which is a special case of (4.1) where $f(x) = 0$, $\boldsymbol{\gamma} = \mathbf{1}_3$, $c_0(x) = 1 - x_1 - x_2 - x_3$, and $c_i(x) = -x_i$ for $i = 1, 2, 3$. Because here $\xi, x \geq 0$ for any feasible point, the constraint corresponding to μ^- in (4.1) can never be active, and we may ignore this constraint without loss of generality. The following proposition shows that there are exactly seven limit points of $x(\varepsilon)$ as ε converges to zero, and that the KKT points converging to non-global solutions of the unrelaxed half–complementarity formulation (2.3) are not local minimizers of (4.21).

Proposition 4.10. *The following statements are true.*

(a) *The set of x parts of the limit points of KKT points for* (4.21) *as $\varepsilon \downarrow 0$ is exactly*

$$\left\{ \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right),\ \left(0, \tfrac{1}{2}, \tfrac{1}{2}\right),\ \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right),\ \left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right),\ \left(0, 0, 1\right),\ \left(0, 1, 0\right),\ \left(1, 0, 0\right) \right\}.$$

(b) *The x parts of the KKT points $(x(\varepsilon), \xi(\varepsilon))$ of (4.21) that converge to*

$$\widehat{x} \in \left\{ \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right),\ \left(0, \tfrac{1}{2}, \tfrac{1}{2}\right),\ \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right),\ \left(\tfrac{1}{2}, \tfrac{1}{2}, 0\right) \right\}$$

as $\varepsilon \downarrow 0$ are not local minimizers.

Proof. Part (b) follows easily from Theorem 4.4. Part (a) can be proven by specializing the conclusions (4.5)–(4.9) to (4.21). Let $x(\varepsilon)$ be a sequence of KKT points for (4.21) that converges to some limit point $\widehat{x}$ as $\varepsilon \downarrow 0$, and let $\xi(\varepsilon)$, $\eta(\varepsilon)$, $\mu^+(\varepsilon)$, $\lambda(\varepsilon)$ satisfy the KKT conditions (4.4). Due to the first constraint in (4.21), $x(\varepsilon)$ has at least one component $x_j(\varepsilon) \geq \frac{1}{3} > \varepsilon$ if ε is sufficiently small. For such j we know that $\lambda_j(\varepsilon) = 0$ due to $x_j(\varepsilon) > 0$. Moreover, based on (4.5) we have

$$\frac{\partial f(x)}{\partial x_j} + \sum_{i \,\in\, \mathcal{E} \cup \mathcal{I}} \lambda_i \frac{\partial c_i(x)}{\partial x_j} \;=\; -\lambda_0(\varepsilon) - \lambda_i(\varepsilon) \;=\; -\frac{\varepsilon}{x_j(\varepsilon)^2}$$

so that $\lambda_0(\varepsilon) = \frac{\varepsilon}{x_j(\varepsilon)^2} > 0$ for such j. Hence the constraint corresponding to $\lambda_0(\varepsilon)$ is active and

$$x_1(\varepsilon) + x_2(\varepsilon) + x_3(\varepsilon) = 1. \tag{4.22}$$

The problem can be broken into three cases depending on the nonzero structure of $\widehat{x}$. We prove one case and leave the other cases to the reader.

Case 1: Suppose $\widehat{x}_j > 0$ for all $j = 1, 2, 3$. For sufficiently small ε we have $x_j(\varepsilon) > \varepsilon$ and therefore $\lambda_j(\varepsilon) = 0$ for all $j = 1, 2, 3$. Then, based on (4.5), we know that $\lambda_0(\varepsilon) = \frac{\varepsilon}{x_j(\varepsilon)^2}$ holds for all $j = 1, 2, 3$, so that $x_1(\varepsilon) = x_2(\varepsilon) = x_3(\varepsilon)$. Using (4.22) and (4.5) we see that the KKT quantities must satisfy

$$\begin{aligned} x(\varepsilon) \;&=\; \left(\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}\right), \qquad \xi(\varepsilon) \;=\; \big(3\varepsilon, 3\varepsilon, 3\varepsilon\big), \qquad \lambda(\varepsilon) \;=\; \big(9\varepsilon, 0, 0, 0\big) \\ \mu^+(\varepsilon) \;&=\; \big(3, 3, 3\big), \qquad \eta(\varepsilon) \;=\; \big(0, 0, 0\big). \end{aligned} \tag{4.23}$$

This shows that $\lim\limits_{\varepsilon \downarrow 0} x(\varepsilon) = \left(\dfrac{1}{3}, \dfrac{1}{3}, \dfrac{1}{3}\right)$ is the only potential limit point with the nonzero structure considered in this case. Conversely, it is easy to verify that the quantities in (4.23) satisfy the KKT conditions (4.4), so that $\widehat{x} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is indeed a limit point. □

For this particular example, this result is as strong as we could hope for. There are only a few KKT points for each sufficiently small ε with an equal number of limit points as $\varepsilon \downarrow 0$. This is in stark contrast to Corollary 3.4, which does not differentiate between *any feasible* point. In addition, if the relaxation NLP(ε) is solved by an NLP solver that is able to escape local maximizers, the point returned by the solver for small ε is close to a global minimizer of the original ℓ_0-norm minimization problem (1.1).

5 Computational Results

After having discussed theoretical properties of the NLP reformulations, we now examine the practical performance of NLP solvers as solution methods of ℓ_0-norm minimization problems. One premise for our experiments is that black-box NLP codes are used with

default settings. Those are applied directly to the NLP reformulations described in the previous sections, without modifications, despite the fact that some of these optimization models are not well-posed (i.e., the MFCQ does not hold at any feasible point for (3.7)).

The goal of these brute-force experiments is to assess the potential of NLP algorithms as solution approaches for hard ℓ_0-norm optimization problems. If these initial experiments give encouraging results, it motivates further research that aims at a deeper understanding of the underlying mechanisms and the development of specialized methods.

The experiments were conducted using the NLP solvers CONOPT 3.15C [15], IPOPT 3.10.4 [35], KNITRO 8.0.0 [6], MINOS 5.51 [31], and SNOPT 7.2-8 [20]. We did not alter the solvers' default options, except that KNITRO was run with the option "`hessopt=5`", which avoids the (potentially time-consuming) computation of the full Hessian matrix. In addition, any arising linear program (LP), mixed-integer linear programs (MILP), quadratic program (QP), and mixed-integer quadratic programs (MIQP) was solved with CPLEX 12.5.1.0. Matlab R2012b and the AMPL modeling software [18] were used as scripting languages and to generate the random problem instances. All numerical experiments reported in this paper were obtained on a 8-core 3.4GHz Intel Core i7 computer with 32GB of memory, running Ubuntu Linux.

5.1 Sparse solutions of linear inequalities

We first consider random instances of the model problem (1.1) of the form

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & \| x \|_0 \\ \text{subject to} & Ax \geq b \quad \text{and} \quad -M\mathbf{1}_n \leq x \leq M\mathbf{1}_n, \end{array} \tag{5.1}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $M > 0$. The test instances were generated using AMPL's internal random number generator, where the elements of A and b are independent uniform random variables between -1 and 1.

Our numerical experiments compare the performance of different NLP optimization codes when they are applied to the different NLP reformulations. Because these problems are nonconvex, we also explore the effect of different starting points.

The NLP reformulations considered are (the first five were discussed in previous sections):

- "Half": The half-complementarity formulation (1.6);
- "Aggregate": The full complementarity formulation where the complementarity constraints are reformulated by the inner product (3.7);
- "Individual": The full complementarity formulation where the complementarity constraints are reformulated by the Hadamard product;
- "Agg-Relaxed" and "Ind-Relaxed" are variants of "Aggregate" and "Individual" formulations, respectively, where the "≤ 0" complementarity constraints are relaxed to "$\leq 10^{-8}$";
- "Squared": The full complementarity formulation where the complementarity constraints are reformulated using the square of the inner product.
- "AMPL": This formulation uses the keyword `complements` in order to pose (1.3) directly as an LPCC in AMPL. It is then up to the particular chosen optimization code to handle the complementarity constraints appropriately. Among the solvers considered here, only KNITRO is able to handle the `complements` keyword. KNITRO then internally refor-

mulates the complementarity constraints using a penalty term that is added to the objective function; see [28] for details.

• "MILP": The MILP formulation (1.7). Our test problems (5.1) explicitly include a bound on the optimal value x^*, so that the same M can be used in (5.1) and (1.7). The solution for this formulation is the global solution for (1.1).

• "LP": The linear programming formulation (1.2).

Because the nonlinear optimization methods aim at finding only local (and not global) optimal solutions of the nonconvex NLP reformulations, the choice of the starting point is crucial. In the experiments, we considered the following options:

• Start1: Set $x^+ = x^- = 0$ and $\xi = 0$;

• Start2: Set $x^+ = x^- = 0$ and $\xi = \mathbf{1}_n$;

• Start3: Let x^{LP} be the optimal solution of the LP formulation (1.2). Then set $x^+ \triangleq \max\{0, x^{\mathrm{LP}}\}$, $x^- \triangleq \max\{0, -x^{\mathrm{LP}}\}$, and $\xi = 0$;

• Start4: Let x^{LP} be the optimal solution of the LP formulation (1.2). Then set $x^+ \triangleq \max\{0, x^{\mathrm{LP}}\}$, $x^- \triangleq \max\{0, -x^{\mathrm{LP}}\}$, and $\xi = \mathbf{1}_n$;

• Start5: Let x^{LP} be the optimal solution of the LP formulation (1.2). Then set $x^+ \triangleq \max\{0, x^{\mathrm{LP}}\}$, $x^- \triangleq \max\{0, -x^{\mathrm{LP}}\}$, and ξ according to (1.4).

5.1.1 Pilot study on small problems

As a pilot study, we considered small problems with 30 constraints and 50 variables (i.e., $m = 30$, $n = 50$), and M was chosen to be 100. To make statements with statistical significance, we generated 50 different random instances. Each of these instances is solved by 140 combinations of NLP solver, problem reformulation, and starting point, in addition to the LP and MILP formulations.

For each individual run, the point x^* returned by the solver is accepted as solution if it satisfies $Ax^* \geq b$, independent of the solvers' exit status. In particular, we accept a solution as feasible if $\|Ax^* - b\|_1 \geq 1e-8$. The number of nonzeros (i.e., $\|x^*\|_0$) is computed by counting the number of elements with $|x_j^*| > 10^{-6}$. Table 1 lists the mean and standard deviation of $\|x^*\|_0$ for the different combinations. We note that all 50 problems were solved (i.e., the returned point satisfied the linear inequalities) for each combination, except for one CONOPT combination. As a reference, for the LP option, the mean was 14.32 with standard deviation 2.97, and for the MILP option, the mean was 4.88 with standard deviation 0.82.

To present the results in more detail, Figures 1(a)–1(e) depict, for each of the 50 instances, the best ℓ_0-norm obtained by the different NLP solvers, in comparison to the LP approximation and the (globally optimal) MILP solution. For each solver, the "Best-" line is the smallest $\|x^*\|_0$ value obtained over all configurations for the same solver. In addition, the figures show the results obtained with the formulation/starting-point combination giving the smallest mean by the respective NLP solver. For example, in Figure 1(a), the line "`CONOPT-Relaxed-3`" shows the outcome for the relaxed aggregate formulation and the 3-th starting point (Start3) with the CONOPT solver. As we can see, all of the NLP solvers are able to find solutions that are sparser than those obtained by the common ℓ_1-approximation. Indeed, the optimal solutions of some NLP solvers, particularly IPOPT and KNITRO, are able to find points with sparsity very close to the sparsest solution possible, as computed by the MILP formulation. Finally, Figure 1(f) shows the sparsest solution obtained by *any* of

		Start1		Start2		Start3		Start4		Start5	
		Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
CONOPT	Aggregate	22.94	5.14	N/A	all failed	14.32	2.97	9.16	2.24	14.32	2.97
	Agg-Relaxed	8.86	2.36	N/A	all failed	7.60	2.03	7.64	2.08	7.88	2.26
	Individual	22.12	4.80	N/A	all failed	14.32	2.97	8.96	1.99	14.32	2.97
	Ind-Relaxed	9.02	2.39	N/A	all failed	8.10	2.22	7.68	1.71	7.90	1.91
	Squared	10.54	3.86	8.22	1.78	9.10	2.73	7.98	2.07	8.96	2.82
IPOPT	Half	6.36	1.45	7.10	1.82	6.36	1.52	6.62	1.60	6.58	1.65
	Aggregate	6.30	1.40	6.98	1.81	6.36	1.51	6.62	1.60	6.54	1.64
	Agg-Relaxed	19.74	4.53	8.40	2.19	10.76	3.03	7.62	1.78	10.78	2.92
	Individual	6.62	1.40	7.02	1.76	6.52	1.34	6.88	1.72	6.86	1.69
	Ind-Relaxed	6.50	1.40	7.08	1.66	6.46	1.43	6.76	1.74	6.82	1.56
	Squared	5.94	1.20	6.76	1.68	6.08	1.40	6.30	1.56	6.56	1.54
KNITRO	AMPL	8.48	1.34	8.42	1.40	10.28	1.81	10.34	1.67	10.14	1.67
	Half	8.44	1.36	8.40	1.26	10.12	1.67	10.34	1.67	10.04	1.62
	Aggregate	7.54	1.99	7.60	1.95	7.52	1.84	7.60	2.01	7.64	1.91
	Agg-Relaxed	8.40	1.96	8.06	1.78	7.46	1.80	7.48	1.72	7.46	1.90
	Individual	7.98	1.36	7.46	1.51	8.42	1.47	8.30	1.47	8.26	1.59
	Ind-Relaxed	7.98	1.36	7.42	1.47	8.50	1.52	8.30	1.49	8.26	1.52
	Squared	8.10	1.84	8.30	1.59	8.34	1.65	8.12	1.77	8.08	1.76
MINOS	Aggregate	17.76	3.42	16.24	3.98	14.32	2.97	14.32	2.97	14.32	2.97
	Agg-Relaxed	17.76	3.42	16.04	4.09	14.32	2.97	14.16	2.84	14.32	2.97
	Individual	17.76	3.42	17.74	3.42	14.32	2.97	14.36	3.02	14.32	2.97
	Ind-Relaxed	17.76	3.42	17.72	3.40	14.28	2.92	14.36	2.91	14.32	2.97
	Squared	11.18	2.78	11.90	3.26	8.28	2.62	8.20	2.33	7.76	1.94
SNOPT	Aggregate	12.80	3.04	9.86	2.47	14.50	2.89	14.08	3.04	14.26	3.10
	Agg-Relaxed	12.80	3.04	9.76	2.47	14.50	2.89	14.08	3.04	14.28	3.05
	Individual	13.40	3.11	13.40	3.11	15.04	3.46	14.26	3.32	14.30	2.96
	Ind-Relaxed	13.40	3.11	13.40	3.11	15.08	3.66	14.32	3.41	14.22	2.99
	Squared	8.04	2.03	7.92	1.87	7.88	2.02	7.70	1.79	7.84	2.01

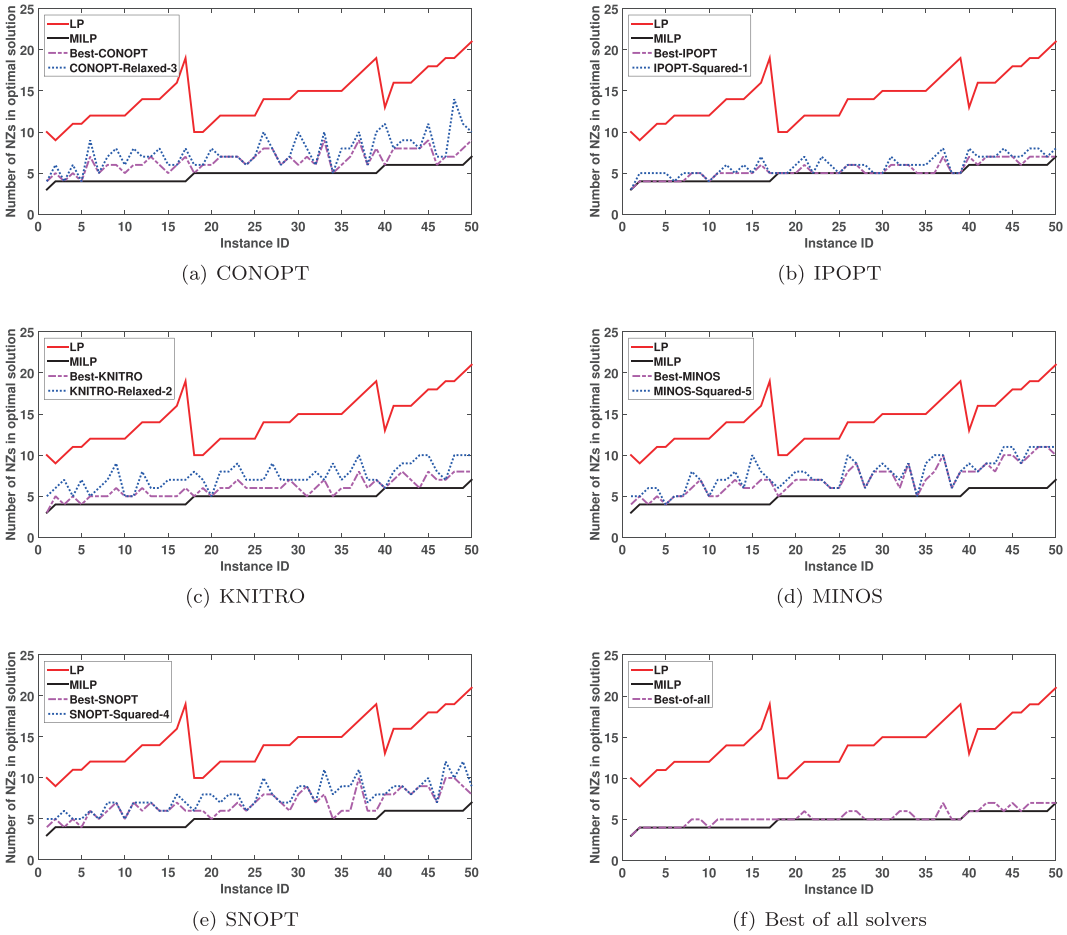
Table 1: Solution quality statistics for pilot study, grouped by solvers.

the solver, formulation, and starting point combinations. We see that, for each instance, at least one combination resulted in a solution that is equal or at most two nonzero elements worse than the global solution.

These results indicate that the application of (standard) NLP solvers to complementarity formulations of the ℓ_0-norm minimization problem results in high-quality solutions, considerably better than what is obtained by the common ℓ_1-approximation. This promising observation is noteworthy, given that the problems are highly nonconvex. From a theoretical standpoint, the NLP solvers are only guarateed to converge to a KKT point (at least when a constraint qualification holds), and as shown in Corollary 3.4, there are exponentially many (undesirable) KKT points to which the NLP solver might potentially converge. In practice, however, the line-search or trust-region globalization mechanisms usually guide the NLP solvers to local minimizers. As discussed in Section 4.3, for the relaxed formulations these correspond to non-dominated points from which there is no obvious direction to improve the objective.

It is also somewhat surprising that, in this preliminary experiment, the squared formulation resulted in the lowest sparsity in some cases, even though no KKT point exists for any instance. The fact that the solvers terminate nevertheless can be explained by looking at the optimality conditions for the squared formulation. These conditions can only be satisfied in exact arithmetic if the product of some Lagrangian multipliers with the (partial) Jacobian matrix $J_\xi\phi(x^+, x^-, \xi)$ of the squared reformulation $\phi(x^+, x^-, \xi) = (\xi^T(x^+ + x^-) + (x^+)^T(x^-))^2$ is nonzero. Note that $J_\xi\phi(x^+, x^-, \xi)$ is zero at every complementary point. However, as the iterations of the NLP solver converge to such a point, the product of $J_\xi\phi(x^+, x^-, \xi)$ with the multipliers can converge to a nonzero value when the multipliers become arbitrarily large. So, even though there is no finite KKT point, the NLP solvers' termination tests can be satisfied by diverging multipliers.

We also note that the CONOPT, MINOS, and SNOPT solvers usually do not converge to good solutions for the Aggregate and Individual formulations when started from a point obtained by the LP formulation (Start3, Start4, and Start5). Indeed, in many cases the solvers terminate immediately at such a starting point. This can be explained by the fact that any feasible point is a KKT point for these formulations, and the active set solvers



(a) CONOPT (b) IPOPT

(c) KNITRO (d) MINOS

(e) SNOPT (f) Best of all solvers

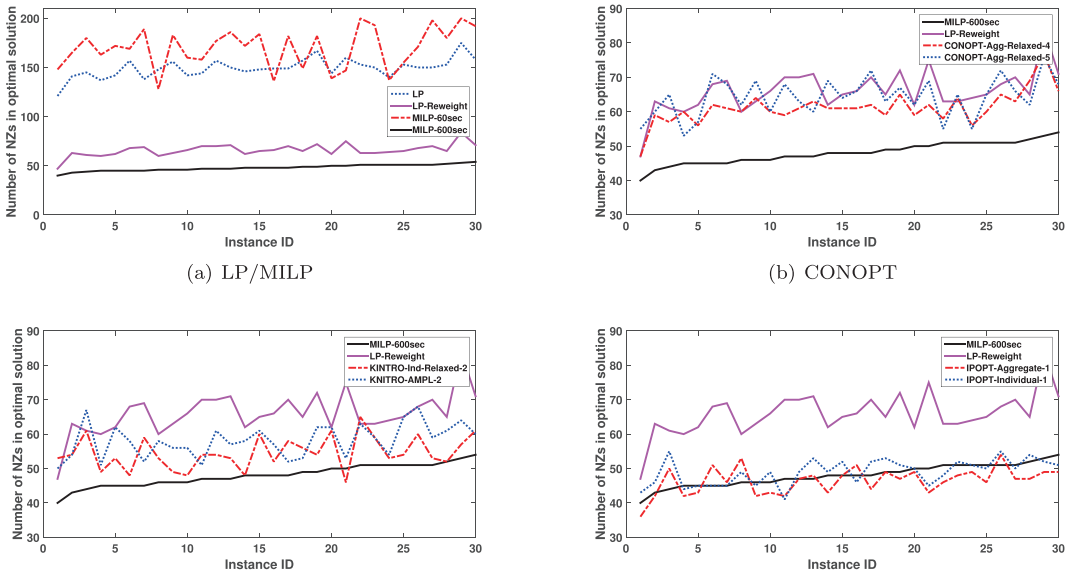
Figure 1: Sparsest solutions for 50 random problems using different NLP solvers, ordered by MILP solutions.

simply compute the corresponding multipliers at the starting point, so that the termination test is immediately satisfied. This is in contrast to the interior point solvers IPOPT and KNITRO, which are required to move the starting point away from bound constraints. This modification results in violations of the respective reformulation of the complementarity constraints and forces the algorithm to take steps.

5.1.2 Large-scale problems

Based on the results in the pilot study, we pursued further numerical studies on larger problems, using the NLP solvers CONOPT, IPOPT and KNITRO. For this set of experiments we generated 30 random instances of (5.1) with 300 constraints and 1,000 variables.

With this problem size, obtaining the true global optimum with the MILP formulation (1.7) is not possible with reasonable computational effort, even though we chose a reasonable



(a) LP/MILP (b) CONOPT

(c) KNITRO (d) IPOPT

Figure 2: Solution sparsities for 30 large-scale random problems using different NLP solvers, ordered by MILP solutions.

big-M constant ($M = 100$) in (5.1). In order to get an idea of what the sparsest solution for an instance might be, we ran CPLEX in multi-threaded mode for 10 minutes, which is roughly equivalent to more than an hour of CPU time (this option is labeled MILP-600sec), and we report the sparsity of the best incumbent. Similarly, we also explored the quality of a heuristic solution that an MILP solver is able to find in a time that is comparable to that taken by the NLP solvers. For this purpose, the MILP60 option reports the best incumbent obtained in one minute, equivalent to about 2.5 minutes of CPU time. In these experiments, CPLEX was run with the `mipemphasis=1` option, to focus on finding good heuristic solutions quickly.

As we observed in the small-case study, the ℓ_1-approximation (1.2) did not lead to good solutions. However, it is common to enhance the LP solution by some improvement heuristic. One such approach is the iterative re-weighted ℓ_1-minimization scheme proposed in [9]. Starting with the optimal solution $x^{*,0}$ of (1.2), this procedure optimizes a sequence of LPs for $k = 0, 1, 2, \ldots$ to generate iterates from

$$x^{*,k+1} = \operatorname*{argmin}_{x\in\mathbb{R}^n} \left\{ \sum_{i=1} w_{k,i}|x_i|, \quad \text{subject to } Ax \geq b \text{ and } -M\mathbf{1}_n \leq x \leq M\mathbf{1}_n \right\},$$

with weights $w_{k,i} = 1/|x_i^{*,k}|$. Here, we understand $w_{k,i} = 1/0 = \infty$ as x_i being fixed to zero. We ran this procedure for 30 iterations (after which the iterates had settled), and report the outcome of this procedure under the label "LPReweight".

The results of the experiments are depicted in Figure 2. First, we see in Figure 2(a) that the iterative re-weighting procedure indeed improves the standard ℓ_1-approximation considerably; it more than halves the objective function. However, there is still a significant gap (17% – 66%) between LP-Reweight and the best solution found by the MILP solver

	$\|x^*\|_0$		CPU time	
	Mean	StdDev	Mean	StdDev
LP	149.33	9.93	4.87	0.59
LP-Reweight	66.00	6.22	5.26	0.60
MILP-60sec	169.83	20.42	373.65	12.89
MILP-600sec	48.03	3.20	4116.37	142.83
CONOPT-Agg-Relaxed-4	61.23	4.90	14.33	2.49
CONOPT-Agg-Relaxed-5	64.23	5.69	32.30	3.84
IPOPT-Aggregate-1	46.47	3.88	544.27	354.91
IPOPT-Individual-1	49.00	3.73	44.99	45.50
KNITRO-AMPL-2	58.13	4.90	46.46	8.95
KNITRO-Ind-Relaxed-2	54.57	4.72	121.52	93.42

Table 2: Summary statistics for large-scale study.

within an hour of CPU time. We note that the MILP solver is not able to find any good solution within about 2.5 minutes of CPU time.

Figure 2(b) shows the solution quality obtained with different reformulations of the ℓ_0-norm minimization problem when solved with CONOPT. To limit the amount of data in the graphs, we plot only selected representative combinations, including the best ones. In this experiment, the relaxed aggregate formulation obtains solutions of similar quality to those obtained by the LP-Reweight option.

The KNITRO results are reported in Figure 2(d) for the relaxed individual formulation and AMPL keyword option which both give better solutions than the LP-Reweight option. IPOPT results are reported in Figure **??** for the aggregate and individual formulations. We see that the solution quality is comparable with that obtained by the MILP solver, and for the Aggregate formulation often even better.

For practical purposes it is also important to consider the computation time required to solve the NLP formulations. Table 2 lists the average solution quality and required CPU time for representative combinations of formulations and starting points. We note that, on average, solutions within 4% of the MILP-600sec objective can be computed in less than one minute ("IPOPT-Individual-1"). Solutions comparable and better than the MILP-600sec objective can be computed in 10 minutes on average ("IPOPT-Aggregate-1"), but with significant variation in the computation time.

5.2 Traffic network problems with fixed costs

The final set of numerical experiments considers traffic planning problems, formulated as the following network flow problem with nodes ($i \in \mathcal{N}$) and arcs ($(i,j) \in \mathcal{A} \subseteq \mathcal{N} \times \mathcal{N}$):

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \displaystyle\sum_{(i,j)\in\mathcal{A}} f_{ij}\|x_{ij}\|_0 + \sum_{i\in\mathcal{N}} v_i\Big(\sum_{(i,j)\in\mathcal{A}} x_{ij}\Big) + \sum_{i\in\mathcal{N}} q_i\Big(\sum_{(i,j)\in\mathcal{A}} x_{ij}\Big)^2 \\ \text{subject to} & Ax = b,\ x \le c,\ \text{and}\ x \ge 0. \end{array} \tag{5.2}$$

Here, $b \in \mathbb{R}^{|\mathcal{N}|}$ are the external traffic volumes (inflows/outflows) at each node, and $A \in \mathbb{R}^{|\mathcal{N}|\times|\mathcal{N}|}$ represents the connectivity (or adjacency) matrix of the network. The vector $c \in \mathbb{R}^{|\mathcal{A}|}$ denotes the capacity of each arc. With each arc (i,j), we associate some fixed

costs f_{ij} that has to be paid if the arc is used (has non-zero flow). In addition, we consider node-based congestion costs that grow quadratically in the total inflow into a node.

To investigate the performance of the NLP formulations for a problem of realistic size, we obtained the incidence matrices A of the Austin network (7,388 nodes and 18,691 arcs) from the transportation network problem collection provided at `http://www.bgu.ac.il/~bargera/tntp/`. The other parameters in (5.2) were chosen as $f_{ij} = 5$ and $c_{ij} = 10$ for each arc, and the variable and quadratic costs v_i and q_i were both independently chosen from uniform random variables between 1 and 5. Based on the rationale that the nodes in the instances are ordered according to their geographic positions, we chose the first five nodes as sources with inflow uniformly chosen at random between $(2, 10)$, and the last five as sinks with equal amounts of outflow. In this way, long paths are generated in the optimal solution.

We compare the "Aggregate" and "Individual" NLP formulations of the complementarity reformulation (2.2) of (5.2) with the standard MIQP formulation of the fixed-cost term in (5.2) (similar to the MILP reformulation (1.7) of the complemenarity reformulation (2.2)).

We point out that it has been demonstrated that the solution time for the MIQP formulation of problems of a similar kind can be dramatically reduced using perspective formulations [19, 22]. That approach, however, can only be applied when the objective function is separable, which is not the case here. Therefore, we are comparing the NLP reformulation proposed in this paper with the time required to solve the straight-forward MIQP formulation.

We performed experiments for random 20 instances, where the NLP formulations are solved with IPOPT. As starting points we chose

- all-zero: $x_{ij} = 0$ and $\xi_{ij} = 0$ for all $(i, j) \in \mathcal{A}$.
- all-off: $x_{ij} = 0$ and $\xi_{ij} = 1$ for all $(i, j) \in \mathcal{A}$.
- all-on: $x_{ij} = c_{ij}$ and $\xi_{ij} = 0$ for all $(i, j) \in \mathcal{A}$.
- all-max: $x_{ij} = c_{ij}$ and $\xi_{ij} = 1$ for all $(i, j) \in \mathcal{A}$.

In Table 3 and Figure 3 we present some of the NLP runs that consistently achieved better objective values than the incumbent found by CPLEX after 10 minutes wall clock time for the MIQP formulation. In analogy to the "LP" formulation in Section 5.1, we also include results in which the ℓ_0-norm in (5.2) was replaced by the ℓ_1-norm, leading to a QP. However, the objective values reported for this option are those with the original ℓ_0-norm.

Clearly, the NLP solvers are able to achieve significantly better objective values than the global MILP solver, in a very small fraction of the time (around 13 CPU secs vs. 1 CPU hour). In particular, we observe that the solutions obtained by the NLP solver are much sparser than those found for the discrete formulation. This indicates that NLP solvers applied to complementarity formulations of ℓ_0-norm structures such as startup costs might be a promising alternative to mixed-integer formulations and deserve further investigation.

6 Conclusions and Outlook

We presented several nonlinear programming reformulations of the ℓ_0-norm minimization problem. Our goal was to study the practical performance of standard NLP codes on these NP-hard problems. We found that the solvers are often able to generate solutions that have objective values close to the global solution. This is somewhat remarkable because the NLP

	Opt. Obj.		CPU time		Sparsity $\|x^*\|_0$	
	Mean	StdDev	Mean	StdDev	Mean	StdDev
L1-QP	24885.58	4804.71	0.99	0.07	2724.90	418.42
MIQP-600sec	27324.94	8796.15	3643.73	108.77	2253.55	2065.14
IPOPT-Ind-all-max	16636.52	4093.09	12.79	1.97	893.75	264.41
IPOPT-Ind-all-off	16983.78	4219.95	44.62	70.10	466.30	105.41
IPOPT-Agg-all-max	17061.89	4444.53	40.82	19.65	1010.15	379.71

Table 3: Summary statistics for Austin traffic network.

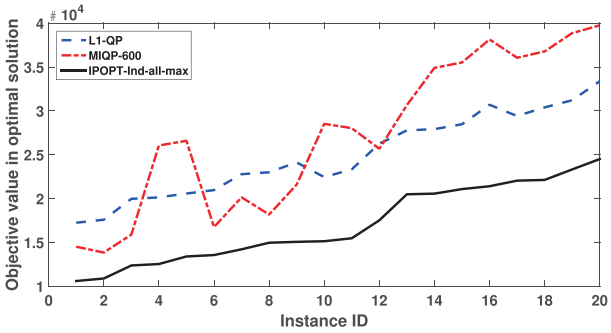


Figure 3: Optimal objective values for different formulations of the traffic problem, ordered by NLP objectives.

formulations are highly nonconvex and the usual constraint qualifications, such as MFCQ, do not hold.

Typically, NLP algorithms are designed to find a KKT point, ideally one that satisfies the second-order necessary optimality conditions. Our analysis pertaining to the optimality conditions of the NLP formulations finds that, for the simple problem with linear constraints in the introduction, any feasible point for the ℓ_0-norm minimization problem is such a KKT point. Consequently, from this perspective, any feasible point seems equally attractive to the NLP solver, and therefore these considerations do not explain the observed high quality of the solutions.

We also discussed the properties of solutions for relaxations of the NLP formulations as the relaxation parameter is driven to zero. For a small example problem with linear constraints we showed that there are only a few KKT points for the relaxed problem, and that those converge to a small number of limit points as the relaxation parameter goes to zero. This is in contrast to the earlier result that does not distinguish between any two feasible points. In addition, we established that a KKT point for the relaxed NLP that is not close to a local minimizer of the original problem is a local *maximizer* for the relaxation. As a consequence, an NLP solver, when applied to the relaxation with a small relaxation parameter, will most likely converge to a point that is close to a local minimizer of the original ℓ_0-norm minimization problem.

These observations might help to explain why the NLP solvers compute points with objective values close to the globally optimal value in our experiments. Some solvers relax any given NLP by a small amount by default, and therefore explicitly solve a relaxation of the complementarity reformulation. For other solvers, numerical inaccuracies or the lineariza-

tion of the nonlinear reformulations of the complementarity constraints at infeasible points might have an effect that is similar to that of a relaxation. The details of such an analogy, as well as the generalization of the results beyond the particular small example, are subject to future research.

Our numerical experiments did not identify a clear winner among the different reformulations of the ℓ_0-norm minimization problems. Similarly, while some NLP codes tended to produce better results than others, it is not clear which specific features of the algorithms or their implementations are responsible for finding good solutions. We point out that each software implementation includes enhancements, such as tricks to handle numerical problems due to round-off error or heuristics that are often not included in the mathematical description in scientific papers. Because the NLP reformulations of the ℓ_0-norm minimization problems are somewhat ill-posed, these enhancement are likely to be crucial for the solver's performance.

Acknowledgements

We wish to thank two referees for a careful reading of the manuscript and constructive comments.

References

[1] R. Andreani, CE. Echagüe and M.L. Schuverdt, Constant-rank condition and second-order constraint qualification, *J. Optim. Theory Appl.* 146(2) (2010) 255–266.

[2] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.* 2 (2009) 183–202.

[3] E. van den Berg and M.P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.* 31 (2008) 890–912.

[4] O.P. Burdakov, C. Kanzow, and A. Schwartz, Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method, *SIAM J. Optim.* 26 (2016) 397–425.

[5] O.P. Burdakov, C. Kanzow, and A. Schwartz, On a reformulation of mathematical programs with cardinality constraints, in *Advances in Global Optimization*, Proceedings of the 3rd World Congress of Global Optimization (July 8-12, 2013, Huangshan, China), D.Y. Gao, N. Ruan, and W.X. Xing (eds), Springer, 2015, pp. 3–14.

[6] R.H. Byrd, J. Nocedal, and R.A. Waltz, KNITRO: An Integrated Package for Nonlinear Optimization, in *Large-Scale Nonlinear Optimization*, G. di Pillo and M. Roma (eds.), Springer-Verlag, 2006, pp. 35-59.

[7] E. Candès, M. Rudelson, T. Tao, and R. Vershynin, Error correction via linear programming, in *46th Annual IEEE Symposium on Foundations of Computer Science, 2005. FOCS 2005*, 2005, pp. 668–681.

[8] E.J. Candès and M. Wakin, An introduction to compressive sampling, *IEEE Signal Process. Mag.* 25 (2008) 21–30.

[9] E.J. Candès, M. Wakin, and S. Boyd, Enhancing sparsity by reweighted l_1 minimization, *J. Fourier Anal. Appl.* 14 (2007) 877–905.

[10] C. Chen and O.L. Mangasarian, Hybrid misclassification minimization. *Adv. Comput. Math.* 5 (1996) 127–136.

[11] S.S. Chen, D.L. Donoho and M.A. Saunders, Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.* 20 (1998) 33–61.

[12] M. Davenport, M. Duarte, Y. Eldar, and G. Kutyniok, Introduction to compressed sensing, Chapter 1 of *Compressed Sensing: Theory and Applications*, Y. Eldar, and G. Kutyniok (eds.), Cambridge University Press, 2012.

[13] A.V. de Miguel, M. Friedlander, F.J. Nogales and S. Scholtes, A two-sided relaxation scheme for mathematical programs with equilibrium constraints, *SIAM J. Optim.* 16 (2006) 587–609.

[14] D.L. Donoho and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization, *Proc. Natl. Acad. Sci. USA* 100 (2003) 2197–2202.

[15] A. Drud, CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems, *Math. Program.* 31 (1985) 153–191.

[16] F. Facchinei and J.S. Pang, *Finite-dimensional variational inequalities and complementarity problems: Volumes I and II*, Springer-Verlag, New York, 2003.

[17] R. Fletcher and S. Leyffer, Solving mathematical programs with complementarity constraints as nonlinear programs, *Optim. Methods Softw.* 19 (2004) 15–40.

[18] R. Fourer, D.M. Gay and B. W. Kernighan, *AMPL: A modeling language for mathematical programming*, Second edition, Thomson, Toronto, Canada, 2003.

[19] A. Frangioni, C. Gentile, E. Grande, and A. Pacifici, Projected perspective reformulations with applications in design problems, *Oper. Res.* 59 (2011) 1225–1232.

[20] P.E. Gill, W. Murray and M.A. Saunders, SNOPT: An SQP algorithm for large-scale constrained optimization, *SIAM Rev.* 47 (2005) 99–131.

[21] R. Gribonval and M. Nielsen, Sparse representations in unions of bases, *IEEE Trans. Inform. Theory* 49 (2003) 3320–3325.

[22] O. Günlük and J. Linderoth, Perspective reformulations of mixed integer nonlinear programs with indicator variables, *Math. Program.* 124 (2010) 183–205.

[23] J. Hu, J.E. Mitchell, J.S. Pang, K.P. Bennett and G. Kunapuli, On the global solution of linear programs with linear complementarity constraints, *SIAM J. Optim.* 19 (2008) 445–471.

[24] J. Hu, J.E. Mitchell, J.S. Pang and B. Yu, On linear programs with linear complementarity constraints, *J. Global Optim.* 53 (2012) 29–51.

[25] R. Janin, Directional derivative of the marginal function in nonlinear programming, *Math. Program. Study* 21 (1984) 110–126.

[26] C. Kanzow and A. Schwartz, The price of inexactness: Convergence properties of relaxation methods for mathematical programs with equilibrium constraints revisited, *Math. Oper. Res.* 40 (2015) 253–275.

[27] C. Kanzow and A. Schwartz, A new regularization method for mathematical programs with complementarity constraints with strong convergence properties, *SIAM J. Optim.* 23 (2013) 770–798.

[28] S. Leyffer, G. López-Calva and J. Nocedal, Interior methods for mathematical programs with complementarity constraints, *SIAM J. Optim.* 17 (2006) 52–77.

[29] Z.-Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[30] O. Mangasarian, Misclassification minimization, *J. Global Optim.* 5 (1994) 309–323.

[31] B.A. Murtagh and M.A. Saunders, MINOS 5.5 User's Guide, Report SOL 83–20R, Systems Optimization Laboratory, Stanford University, December 1983 (revised February 1995).

[32] S. Scholtes, Nonconvex structures in nonlinear programming, *Oper. Res.*, 52 (2004) 368–383.

[33] S. Scholtes, Convergence properties of a regularisation scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.* 11 (2001) 918–936.

[34] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1996) 267–288.

[35] A. Wächter and L. T. Biegler, On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming, *Math. Program.* 106 (2006) 25–57.

[36] S.J. Wright, R.D. Nowak and M A. T. Figueiredo, Sparse reconstruction by separable approximation, *IEEE Trans. Signal Process.* 57 (2009) 2479–2493.

[37] T.T. Wu, Y.F. Chen, T. Hastie, E. Sobel and K. Lange, Genome-wide association analysis by lasso penalized logistic regression, *Bioinform.* 25 (2009) 714–721.

[38] Y. Zhang, Theory of compressive sensing via ℓ_1-minimization: a non-RIP analysis and extensions, *J. Oper. Res. Soc. China* 1 (2013) 79–105.

[39] X. Zheng, X. Sun, D. Li and J. Sun, Successive convex approximations to cardinality-constrained convex programs: a piecewise-linear DC approach, *Comput. Optim. Appl.* 59 (2014) 379–397.

Manuscript received 17 May 2016
revised 25 August 2016
accepted for publication 3 September 2016

Mingbin Feng
Department of Industrial Engineering and Management Sciences
Northwestern University, Evanston, IL 60208, USA
E-mail address: mingbinfeng2011@u.northwestern.edu

John E. Mitchell
Department of Mathematical Sciences
Rensselaer Polytechnic Institute, Troy, NY 12180, USA
E-mail address: mitchj@rpi.edu

Jong-Shi Pang
Department of Industrial and Systems Engineering
University of Southern California, Los Angeles, CA 90089, USA
E-mail address: jongship@usc.edu

Xin Shen
Department of Mathematical Sciences
Rensselaer Polytechnic Institute, Troy, NY 12180, USA
E-mail address: shenx5@rpi.edu

Andreas Wächter
Department of Industrial Engineering and Management Sciences
Northwestern University, Evanston, IL 60208, USA
E-mail address: andreas.waechter@northwestern.edu