Optimization Technical Report

Christina Oberlin · Stephen J. Wright

Active Set Identification in Nonlinear Programming

January 2005

Abstract Techniques that identify the active constraints at a solution of a nonlinear programming problem from a point near the solution can be a useful adjunct to nonlinear programming algorithms. They have the potential to improve the local convergence behavior of these algorithms, and in the best case can reduce an inequality constrained problem to an equality constrained problem with the same solution. This paper describes several techniques that do not require good Lagrange multiplier estimates for the constraints to be available a priori, but depend only on function and first derivative information. Computational tests comparing the effectiveness of these techniques on a variety of test problems are described. Many tests involve degenerate cases, in which the constraint gradients are not linearly independent and/or strict complementarity does not hold.

Keywords nonlinear programming · active constraint identification · degeneracy

Mathematics Subject Classification (2000) 90C30 · 90C46

Research supported by NSF Grants ATM-0296033, CNS-0127857, CCF-0113051, SCI-0330538, DMS-0427689, CCF-0430504, CTS-0456694, CNS-0540147, DOE grant DE-FG02-04ER25627, and an NSF Graduate Research Fellowship.

Christina Oberlin
Computer Sciences Department,
University of Wisconsin,
1210 W. Dayton Street,
Madison, WI 53706. E-mail: coberlin@cs.wisc.edu

Stephen J. Wright
Computer Sciences Department,
University of Wisconsin,
1210 W. Dayton Street,
Madison, WI 53706. E-mail: swright@cs.wisc.edu

1 Introduction

Consider the following nonlinear programming problem:

$$\min f(x) \text{ subject to } h(x) = 0, \ \ c(x) \leq 0, \tag{1}$$

where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$, $c : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^n \to \mathbb{R}^p$ are twice continuously differentiable functions.

In this paper, we examine identification of active inequality constraints—the components of c for which equality holds at a local solution x^*—using information available at a point x near x^*. We focus on identification schemes that do not require good estimates of the Lagrange multipliers to be available a priori. Rather, in some cases, such estimates are computed as an adjunct to the identification technique. In most of our results, we relax the "standard" nondegeneracy assumptions at x^* to allow linearly dependent active constraint gradients and weakly active constraints. We consider three schemes that require solution of linear programs and one that requires solution of a mixed integer program. We analyze the effectiveness of these schemes and discuss computational issues of solving the linear and mixed-integer programs. Finally, we present results obtained on randomly generated problems and on degenerate problems from the CUTEr test set [13].

One area in which identification schemes are useful is in "EQP" approaches to sequential quadratic programming (SQP) algorithms, in which each iteration consists of an estimation of the active set followed by solution of an equality constrained quadratic program that enforces the apparently active constraints and ignores the apparently inactive ones. The "IQP" variant of SQP, in which an inequality constrained subproblem is solved (thereby estimating the active set implicitly), has been more widely studied in the past two decades, but the EQP variant has been revived recently by Byrd et al. [5] [6].

1.1 Assumptions and Background

We describe here the notation and assumptions used in the remainder of the paper.

The Lagrangian for (1) is

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \mu^T h(x) + \lambda^T c(x), \tag{2}$$

where $\mu \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^m$ are Lagrange multipliers. First-order necessary conditions for x^* to be a solution of (1), assuming a constraint qualification, are that there exist multipliers (μ^*, λ^*) such that

$$\nabla_x \mathcal{L}(x^*, \mu^*, \lambda^*) = 0, \tag{3a}$$

$$h(x^*) = 0, \tag{3b}$$

$$0 \geq c(x^*) \perp \lambda^* \geq 0, \tag{3c}$$

where the symbol $\perp$ denotes vector complementarity; that is, $a \perp b$ means $a^T b = 0$. We define the "dual" solution set as follows:

$$\mathcal{S}_D \stackrel{\text{def}}{=} \{(\mu^*, \lambda^*) \text{ satisfying (3)}\}, \tag{4}$$

while the primal-dual solution set $\mathcal{S}$ is

$$\mathcal{S} \stackrel{\text{def}}{=} \{x^*\} \times \mathcal{S}_D.$$

The set of *active inequality constraints* at x^* is defined as follows:

$$\mathcal{A}^* = \{i = 1, 2, \ldots, m \,|\, c_i(x^*) = 0\}.$$

The *weakly active inequality constraints* $\mathcal{A}_0^*$ are those active constraints i for which $\lambda_i^* = 0$ for all optimal multipliers (μ^*, λ^*); that is,

$$\mathcal{A}_0^* = \{i \in \mathcal{A}^* \,|\, \lambda_i^* = 0 \text{ for all } (\mu^*, \lambda^*) \in \mathcal{S}_D\}. \tag{5}$$

The constraints $\mathcal{A}^* \setminus \mathcal{A}_0^*$ are said to be the *strongly active inequalities*.

In this paper, we make use of the following two constraint qualifications at x^*. The linear independence constraint qualification (LICQ) is that

$$\{\nabla h_i(x^*),\ i = 1, 2, \ldots, p\} \cup \{\nabla c_i(x^*),\ i \in \mathcal{A}^*\} \text{ is linearly independent.} \tag{6}$$

The Mangasarian-Fromovitz constraint qualification (MFCQ) is that there is a vector $v \in \mathbb{R}^n$ such that

$$\nabla c_i(x^*)^T v < 0,\ \ i \in \mathcal{A}^*; \quad \nabla h_i(x^*)^T v = 0,\ \ i = 1, 2, \ldots, p, \tag{7a}$$

$$\{\nabla h_i(x^*),\ i = 1, 2, \ldots, p\} \text{ is linearly independent.} \tag{7b}$$

In some places, we use the following second-order sufficient condition: Defining

$$\begin{aligned} \mathcal{C} \stackrel{\text{def}}{=} \{v \,|\, & \nabla c_i(x^*)^T v = 0,\ i \in \mathcal{A}^* \setminus \mathcal{A}_0^*,\ \ \nabla c_i(x^*)^T v \le 0,\ i \in \mathcal{A}_0^*, \\ & \nabla h_i(x^*)^T v = 0,\ i = 1, 2, \ldots, p\}, \end{aligned} \tag{8}$$

we require that

$$v^T \nabla^2_{xx} \mathcal{L}(x^*, \mu^*, \lambda^*) v > 0 \text{ for all } v \in \mathcal{C} \setminus \{0\} \text{ and all } (\mu^*, \lambda^*) \in \mathcal{S}_D. \tag{9}$$

The following notation is used for first derivatives of the objective and constraint functions at x:

$$g(x) = \nabla f(x), \quad J(x) = [\nabla h_i(x)^T]_{i=1,2,\ldots,p}, \quad A(x) = [\nabla c_i(x)^T]_{i=1,2,\ldots,m}.$$

We use $A_i(x) = \nabla c_i(x)^T$ to denote the ith row of $A(x)$, while for any index set $\mathcal{T} \subset \{1, 2, \ldots, m\}$, we use $A_{\mathcal{T}}(x)$ to denote the $|\mathcal{T}| \times n$ submatrix corresponding to $\mathcal{T}$. In some subsections, the argument x is omitted from the quantities $c(x)$, $A(x)$, $A_i(x)$, and $A_{\mathcal{T}}(x)$ if the dependence on x is clear from the context. In some instances, we also use ∇c_i^*, g^*, etc., to denote $\nabla c_i(x^*)$, $g(x^*)$, etc., respectively.

Given a matrix $B \in \mathbb{R}^{n\times q}$ we denote

$$\text{range}\,[B] = \{Bz \,|\, z \in \mathbb{R}^q\}, \quad \text{pos}[B] = \{Bz \,|\, z \in \mathbb{R}^q,\ z \geq 0\}.$$

The norms $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$ all appear in the paper. When the subscript is omitted, the Euclidean norm $\|\cdot\|_2$ is intended.

We use the usual definition of the distance function dist $(\cdot,\cdot)$ between sets, that is

$$\text{dist}\,(\mathcal{S}_1, \mathcal{S}_2) = \inf_{s_1\in\mathcal{S}_1, s_2\in\mathcal{S}_2} \|s_1 - s_2\|. \tag{10}$$

(Distance between a point and a set is defined similarly.)

For a vector z, function $\max(z, 0)$ (defined componentwise) is denoted by z_+, while z_- denotes $\max(-z, 0)$. We use the notation e throughout the paper to denote the vector $(1, 1, \ldots, 1)^T$. (The dimension of e is not specified but is clear from the context.)

In assessing the accuracy of an active set estimate, a *false positive* is an index i that is identified as active by our scheme but which actually does not belong to $\mathcal{A}^*$, while a *false negative* is an index $i \in \mathcal{A}^*$ which is wrongly identified as inactive.

1.2 Related Work

Some previous works have studied the behavior of nonlinear programming algorithms in identifying active constraint sets, more or less as a byproduct of their progress toward a solution. Other papers have described the use of these active-set estimates to speed the convergence of the algorithm in its final stages. We mention several works of both types here, in nonlinear programming and in the context of other optimization and complementarity problems.

Bertsekas [1] proposed a two-metric algorithm for minimizing a nonlinear function subject to bound constraints on the components of x. A key aspect of this method is estimation of the active bounds at the solution. (Different second-order scalings are applied to the apparently active components and the free components.) Strongly active constraints are identified for all feasible x in a neighborhood of x^*. The latter result is also proved by Lescrenier [16] for a trust-region algorithm.

Burke and Moré [3, 4] take a geometric approach, assuming the constraints to be expressed in the form $x \in \Omega$ for a convex set Ω. This set can be partitioned into faces, where a face F is defined to be a subset of Ω such that every line segment in Ω whose relative interior meets F is contained in F. In this context, active set identification corresponds to the identification of the face that contains the solution x^*. In [3], it is shown that "quasi-polyhedral" faces are identified for all x close to x^* provided that a geometric nondegeneracy condition akin to strict complementarity is satisfied. (Quasi-polyhedrality is defined in [3, Definition 2.5]; curved faces are not quasi-polyhedral.) Burke [2] takes a partly algebraic viewpoint and shows that the set of active indices of a linear approximation to the problem at x near x^*

are sufficient for the objective gradient to be contained in the cone of active constraint gradients—a result not unlike Theorem 3 below.

Wright [21] also uses a hybrid geometric-algebraic viewpoint and considers convex constraint sets Ω with (possibly curved) boundaries defined by (possibly nonlinear) inequalities. The concept of a "class-$\mathcal{C}^p$ identifiable surface" is defined and it is shown that this surface is identified at all x close to x^* provided that a nondegeneracy condition is satisfied. Hare and Lewis [15] extend these concepts to nonconvex sets, using concepts of prox-regularity and partly smooth functions developed elsewhere by Lewis [17] and others.

Facchinei, Fischer, Kanzow [10] describe a technique based on the algebraic representation of the constraint set, that uses estimates of the Lagrange multipliers (μ, λ) along with the current x to obtain a two-sided estimate of the distance of (x, μ, λ) to the primal-dual solution set. This estimate is used in a threshold test to obtain an estimate of $\mathcal{A}^*$. We discuss this technique further in Section 2.

Conn, Gould, and Toint [8, Chapter 12] discuss the case of convex constraints, solved with a trust-region algorithm in which a "generalized Cauchy point" is obtained via gradient projection. They prove that when assumptions akin to strict complementarity and LICQ hold at the solution x^*, their approach identifies the active set once the iterates enter a neighborhood of x^*; see for example [8, Theorem 12.3.8].

Active constraint identification has played an important role in finite termination strategies for linear programming. Ye [25] proposed such a strategy, which determined the active set estimate by a simple comparison of the primal variables with the dual slacks. (An equality constrained quadratic program, whose formulation depends crucially on the active set estimate, is solved in an attempt to "jump to" an optimal point.) El-Bakry, Tapia, and Zhang [9] discuss methods based on "indicators" for identifying the active constraints for linear programming.

Similar active identification and finite termination strategies are available for monotone linear complementarity problems; see for example, the paper of Monteiro and Wright [18]. For monotone nonlinear complementarity problems, Yamashita, Dan, and Fukusmima [24] describe a technique for classifying indices (including degenerate indices) at the limit point of a proximal point algorithm. This threshold is defined similarly to the one in [10], while the classification test is similar to that of [18].

1.3 Organization of the Paper

In Section 2, we review a technique for identifying the active set using an estimate (x, μ, λ) of the primal-dual optimum. This technique provides the basis for the identification techniques of Subsections 3.2 and 3.3. Section 3 describes the main techniques for identifying the active set without assuming that reliable estimates of the Lagrange multipliers (μ, λ) are available. Subsection 3.1 describes a technique used by Byrd et al. [5] [6] along with a dual variant; Subsection 3.2 describes a technique based on minimizing the primal-dual measure of Section 2, which can be formulated as a mixed integer program; Subsection 3.3 derives a linear programming approximation to

the latter technique. In all cases, we prove results about the effectiveness of these schemes and discuss their relationship to each other. In Section 4, we describe our implementation of the identification schemes and present results obtained on randomly generated problems (with controlled degeneracy) and on degenerate problems from the CUTEr test set. Some conclusions appear in Section 5.

2 Identification from a Primal-Dual Point

In this section, we suppose that along with an estimate x of the solution x^*, we have estimates of the Lagrange multipliers (μ, λ). We describe a threshold test based on the function ψ defined as follows:

$$\psi(x,\mu,\lambda) = \left\| \begin{bmatrix} \nabla_x \mathcal{L}(x,\mu,\lambda) \\ h(x) \\ \min(\lambda, -c(x)) \end{bmatrix} \right\|_1, \tag{11}$$

where the $\min(\cdot,\cdot)$ is taken componentwise. (Other norms could be used in this definition, including weighted norms, but the ℓ_1 norm is convenient for computation in later contexts.) The test based on ψ provides the starting point for the LPEC scheme of Subsection 3.2, where we fix x and choose (μ, λ) to minimize ψ, rather than assuming that (μ, λ) are given.

The following result shows that for (x, μ, λ) close to $\mathcal{S}$, this function provides a two-sided estimate of the distance to the solution. (See Facchinei, Fischer, and Kanzow [10, Theorem 3.6], Hager and Gowda [14], and Wright [22, Theorem A.1] for proofs of results similar or identical to this one.)

Theorem 1 *Suppose the KKT conditions (3), the MFCQ (7), and the second-order condition (9) are satisfied at x^*. There are constants $\epsilon \in (0, 1]$ and $C > 0$ such that, for all (x, μ, λ) with $\lambda \geq 0$ and $\operatorname{dist}((x,\mu,\lambda), \mathcal{S}) \leq \epsilon$, we have*

$$C^{-1}\psi(x,\mu,\lambda) \leq \operatorname{dist}((x,\mu,\lambda), \mathcal{S}) \leq C\psi(x,\mu,\lambda). \tag{12}$$

(The upper bound of 1 in the definition of ϵ is needed to simplify later arguments.)

For future reference, we define L to be a Lipschitz constant for the functions g, c, h, A, and J in the neighborhood $\|x - x^*\| \leq \epsilon$, for the ϵ given in Theorem 1. In particular, we have

$$\begin{aligned} &\|g(x) - g(x^*)\| \leq L\|x - x^*\|, \quad \|c(x) - c(x^*)\| \leq L\|x - x^*\|, \\ &\|A(x) - A(x^*)\| \leq L\|x - x^*\|, \quad \|h(x) - h(x^*)\| \leq L\|x - x^*\|, \\ &\|J(x) - J(x^*)\| \leq L\|x - x^*\|, \quad \text{for all } x \text{ with } \|x - x^*\| \leq \epsilon. \end{aligned} \tag{13}$$

We define a constant K_1 such that the following condition is satisfied:

$$K_1 = \max\left(\|c(x^*)\|_\infty, \max_{(\mu^*,\lambda^*)\in\mathcal{S}_D} \|(\mu^*, \lambda^*)\|_\infty \right) + 1, \tag{14}$$

(Note that finiteness of K_1 is assured under MFCQ.)

The active set estimate is a threshold test, defined as follows for a given parameter $\sigma \in (0,1)$:

$$\mathcal{A}(x,\mu,\lambda) = \{i \,|\, c_i(x) \geq -\psi(x,\mu,\lambda)^\sigma\}, \tag{15}$$

The following result is an immediate consequence of Theorem 1. It has been proved in earlier works (see, for example, [10]), but since the proof is short and illustrative, we repeat it here.

Theorem 2 *Suppose that the KKT conditions (3), the MFCQ (7), and the second-order condition (9) are satisfied at x^*. Then there is $\bar{\epsilon}_1 > 0$ such that for all (x,μ,λ) with $\lambda \geq 0$ and $\operatorname{dist}((x,\mu,\lambda),\mathcal{S}) \leq \bar{\epsilon}_1$, we have that $\mathcal{A}(x,\mu,\lambda) = \mathcal{A}^*$.*

Proof. First set $\bar{\epsilon}_1 = \epsilon$, where ϵ is small enough to satisfy the conditions in Theorem 1. Taking any $i \notin \mathcal{A}^*$, we can decrease $\bar{\epsilon}_1$ if necessary to ensure that the following inequalities hold for all (x,μ,λ) with $\operatorname{dist}((x,\mu,\lambda),\mathcal{S}) \leq \bar{\epsilon}_1$:

$$c_i(x) < (1/2)c_i(x^*) \leq -\psi(x,\mu,\lambda)^\sigma,$$

thus ensuring that $i \notin \mathcal{A}(x,\mu,\lambda)$.

We can reduce $\bar{\epsilon}_1$ again if necessary to ensure that the following relation holds for all $i \in \mathcal{A}^*$ and all (x,μ,λ) with $\operatorname{dist}((x,\mu,\lambda),\mathcal{S}) \leq \bar{\epsilon}_1$:

$$|c_i(x)| \leq L\|x - x^*\| \leq L \operatorname{dist}((x,\mu,\lambda),\mathcal{S}) \leq LC\psi(x,\mu,\lambda) \leq \psi(x,\mu,\lambda)^\sigma,$$

where L is the Lipschitz constant defined in (13). We conclude that $i \in \mathcal{A}(x,\mu,\lambda)$. □

High-quality estimates of the optimal Lagrange multipliers may be available in primal-dual interior-point algorithms and augmented Lagrangian algorithms. In SQP algorithms, an estimate (μ,λ) may be available from the QP subproblem solved at the previous iteration, or from an approximation procedure based on the current estimate of the active set (which usually also derives from the QP subproblem). However, the use of devices such as trust regions or ℓ_1 penalty terms in the subproblem may interfere with the accuracy of the Lagrange multiplier estimates. Moreover, in many algorithms, there is not a particularly strong motivation for obtaining accurate estimates of (μ,λ). For instance, in SQP algorithms that use exact second derivatives, rapid convergence of the primal iterates to x^* can be obtained even when (μ,λ) do not converge to $\mathcal{S}_D$; see Theorem 12.4.1 of Fletcher [11] and the comments that follow this result. The QP subproblem of the primal-dual algorithms in the `Knitro` software package may return only the primal variables, in which case the multipliers must be approximated using primal information [7].

Even in cases in which an estimate of (μ,λ) is available from the algorithm, it may be desirable to seek alternative values of (μ,λ) that decrease the value of $\psi(x,\mu,\lambda)$, thereby tightening the tolerance in the threshold test (15). This approach forms the basis of the techniques described in Subsections 3.2 and 3.3, which provide asymptotically accurate estimates of the Lagrange multipliers as well as of the active set $\mathcal{A}^*$.

3 Identification from a Primal Point

We describe a number of techniques, for estimating $\mathcal{A}^*$ for a given x near the solution x^*. We discuss the relationships between these techniques and conditions under which they provide asymptotically accurate estimates of $\mathcal{A}^*$.

3.1 Linear Programming Techniques

We describe here techniques based on a linearization of the ℓ_1-penalty formulation of (1). A linearized trust-region subproblem is solved and an estimate of $\mathcal{A}^*$ is extracted from the solution. One of these techniques is used by Byrd et al. [5] [6] as part of their SQP-EQP approach. (The idea of a linearized trust-region subproblem was proposed initially by Fletcher and Sainz de la Maza [12].)

The following subproblem forms the basis of the techniques in this section:

$$\min_d \; g^T d + \nu \|Jd + h\|_1 + \nu \|(c + Ad)_+\|_1 \quad \text{subject to } \|d\|_\infty \le \Delta, \tag{16}$$

where ν is a penalty parameter, Δ is the trust-region radius, and all functions are assumed to be evaluated at x. This problem can be formulated explicitly as a linear program by introducing auxiliary variables r, s, and t, and writing

$$\min_{(d,r,s,t)} \; g^T d + \nu e^T r + \nu e^T s + \nu e^T t, \quad \text{subject to} \tag{17a}$$

$$Ad + c \le r, \;\; Jd + h = t - s, \;\; -\Delta e \le d \le \Delta e, \;\; (r, s, t) \ge 0, \tag{17b}$$

where, as mentioned in the introduction, we have $e = (1, 1, \ldots, 1)^T$. The dual of this problem is as follows:

$$\min_{(\lambda,\mu,u,v)} \; -c^T \lambda - h^T \mu + \Delta e^T u + \Delta e^T v, \quad \text{subject to} \tag{18a}$$

$$A^T \lambda + J^T \mu + g = u - v, \;\; 0 \le \lambda \le \nu e, \;\; -\nu e \le \mu \le \nu e, \;\; (u, v) \ge 0. \tag{18b}$$

This formulation can be written more compactly as follows:

$$\min_{(\lambda,\mu)} \; -c^T \lambda - h^T \mu + \Delta \|A^T \lambda + J^T \mu + g\|_1, \quad \text{subject to} \tag{19a}$$

$$0 \le \lambda \le \nu e, \;\; -\nu e \le \mu \le \nu e. \tag{19b}$$

The formulations above are feasible and bounded. Moreover, they admit some invariance to scaling the constraints. Suppose for some constraint c_i, we have that the λ_i component of the dual solution is strictly less than its upper bound of ν. By duality, we then have $r_i = 0$ at the solution of (17). If we scale constraint c_i by some $\sigma_i > 0$ (that is, we set $c_i \leftarrow \sigma_i c_i$ and $A_i \leftarrow \sigma_i A_i$), constraints (17b) and (18b) continue to be satisfied, while the objectives (17a) and (18a) remain unchanged (and therefore optimal) if we set $\lambda_i \leftarrow \lambda_i / \sigma_i$, provided that $\lambda_i / \sigma_i \le \nu$. Similar comments apply regarding the components of h.

The active set estimate can be derived from the solution of these linear programs in different ways. We mention the following three possibilities:

$$\mathcal{A}_c(x) = \{i \mid A_i d + c_i \geq 0\}, \tag{20a}$$

$$\mathcal{A}_\lambda(x) = \{i \mid \lambda_i > 0\}, \tag{20b}$$

$$\mathcal{A}_B(x) = \{i \mid \lambda_i \text{ is in the optimal basis for (18)}\}. \tag{20c}$$

The first of these activity tests (20a) cannot be expected to identify weakly active constraints except when $x = x^*$. The second test (20b) will generally not identify weakly active constraints, and will also fail to identify a strongly active constraint i if the particular multiplier estimate used in the test happens to have $\lambda_i = 0$. The third test (20c) does not attempt to estimate the full active set, but rather a "sufficient" subset of it that can be used in subsequent calculations requiring a nonsingular basis matrix for the active constraint gradients.

For the remainder of this section, we focus on $\mathcal{A}_c(x)$. The following simple lemma shows that, for x sufficiently near x^* and Δ sufficiently small, this activity test does not contain false positives.

Lemma 1 *There are positive constants $\bar{\epsilon}_2$ and $\bar{\Delta}$ such that when $\|x - x^*\| \leq \bar{\epsilon}_2$ and $\Delta \leq \bar{\Delta}$, we have $\mathcal{A}_c(x) \subset \mathcal{A}^*$.*

Proof. We first choose $\bar{\epsilon}_2$ small enough such that for any x with $\|x - x^*\| \leq \bar{\epsilon}_2$ and any $i \notin \mathcal{A}^*$ we have $c_i(x) \leq \frac{1}{2} c_i(x^*) < 0$. By decreasing $\bar{\Delta}$ if necessary, we also have, for any $\|d\|_\infty \leq \Delta \leq \bar{\Delta}$ with $\|x - x^*\| \leq \bar{\epsilon}_2$, that $i \notin \mathcal{A}^* \Rightarrow A_i(x)d + c_i(x) < 0$. The result follows from the definition (20a) of $\mathcal{A}_c(x)$. □

When the trust-region radius Δ is bounded in terms of $\|x - x^*\|$ and a constraint qualification holds, we can show that the set identified by (20a) is at least extensive enough to "cover" the objective gradient g^*.

Theorem 3 *If MFCQ holds at x^*, for any $\zeta \in (0, 1)$, there are positive constants $\bar{\nu}$, $\bar{\epsilon}_2$, and $\bar{\Delta}$ such that whenever the conditions $\nu \geq \bar{\nu}$, $\|x - x^*\| \leq \bar{\epsilon}_2$, and $\Delta \in [\|x - x^*\|^\zeta, \bar{\Delta}]$ are satisfied, we have*

$$-g^* \in \text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i \in \mathcal{A}_c(x)}]. \tag{21}$$

Proof. We start by defining $\bar{\epsilon}_2$ and $\bar{\Delta}$ as in Lemma 1. For these values (and any smaller values) we have immediately that $\mathcal{A}_c(x) \subset \mathcal{A}^*$.

We require $\nu \geq \bar{\nu}$, where

$$\bar{\nu} \stackrel{\text{def}}{=} \max\ (\{\|(\mu^*, \lambda^*)\|_\infty \mid (\mu^*, \lambda^*) \in \mathcal{S}_D\}) + 1. \tag{22}$$

Note that $\bar{\nu}$ is well-defined because the KKT and MFCQ conditions guarantee the nonemptiness and boundedness of $\mathcal{S}_D$.

For any $(\mu^*, \lambda^*) \in \mathcal{S}_D$, the dual problem (18) at x^* with $(\mu, \lambda, u, v) = (\mu^*, \lambda^*, 0, 0)$ has objective value 0 because of the complementarity condition (3c). For the problem with $x \neq x^*$, we obtain a feasible point for (18) by setting

$$(\mu, \lambda, u, v) = \left(\mu^*, \lambda^*, (A^T\lambda^* + J^T\mu^* + g)_+, (A^T\lambda^* + J^T\mu^* + g)_-\right).$$

The objective at this point is

$$\begin{aligned}
&-c^T\lambda^* - h^T\mu^* + \Delta\|A^T\lambda^* + J^T\mu^* + g\|_1 \\
&= (c(x^*) - c(x))^T\lambda^* + (h(x^*) - h(x))^T\mu^* \\
&\qquad +\Delta\|(A(x^*)^T - A(x)^T)\lambda^* + (J(x^*)^T - J(x)^T)\mu^* + (g(x^*) - g(x))\|_1 \\
&(23) \; O(\|x - x^*\|).
\end{aligned}$$

The first equality is due to (3), while the second is due to the continuous differentiability of f, c, and h and the boundedness of $\mathcal{S}_D$. The optimal point for (18) must therefore have an objective value that is bounded above by a positive number of size $O(\|x - x^*\|)$.

Suppose for contradiction that regardless of how small we choose $\bar{\epsilon}_2$, there is an x with $\|x-x^*\| \leq \bar{\epsilon}_2$ such that the active set $\mathcal{A}_c(x)$ has the property that $-g^* \notin \text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i\in\mathcal{A}_c(x)}]$. Since there are only a finite number of possible sets $\mathcal{A}_c(x)$, we pick one of them for which this property holds for x arbitrarily close to x^*, and call it $\mathcal{A}_1$. The set $\text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i\in\mathcal{A}_1}]$ is finitely generated and is therefore closed; see Rockafellar [19, Theorem 19.1].

Using the definition for $\text{dist}\,(\cdot,\cdot)$ (10), we have that τ defined by

$$\tau \stackrel{\text{def}}{=} (0.5)\text{dist}\;(-g^*, \text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i\in\mathcal{A}_1}]) \tag{24}$$

is strictly positive. After a possible reduction of $\bar{\epsilon}_2$, we have that

$$\text{dist}\;(-g(x), \text{range}\,[\nabla h(x)] + \text{pos}[(\nabla c_i(x))_{i\in\mathcal{A}_1}]) \geq \tau, \tag{25}$$

for the given $\mathcal{A}_1$ and *all* x with $\|x - x^*\| \leq \bar{\epsilon}_2$. (The proof of the latter claim makes use of standard arguments and appears in Appendix A.)

Given x with $\mathcal{A}_c(x) = \mathcal{A}_1$, let the solutions to the problems (17) and (18) at x be denoted by (d_x, r_x, s_x, t_x) and $(\mu_x, \lambda_x, u_x, v_x)$, respectively. For all $i \notin \mathcal{A}_1$, we have by (20a) and complementarity that $A_i d_x + c_i < 0$, $(r_x)_i = 0$, and

$$(\lambda_x)_i = 0, \;\; \text{for all } i \notin \mathcal{A}_1. \tag{26}$$

We now consider the objective of the dual problem (18) in two parts. We have by using the property (26) that

$$\begin{aligned}
\Delta e^T u_x + \Delta e^T v_x &\geq \Delta \min_{\lambda\geq 0,\mu} \|g + J^T\mu + \sum_{i\in\mathcal{A}_1} \lambda_i \nabla c_i\|_1 \\
&= \Delta\text{dist}\;(-g, \text{range}\,[\nabla h] + \text{pos}[(\nabla c_i)_{i\in\mathcal{A}_1}]) \\
&\geq \Delta\tau.
\end{aligned}$$

From $\nu \geq \bar{\nu}$ and (22), we also have

$$-c^T\lambda_x - h^T\mu_x \geq -\nu\|c_+\|_1 - \nu\|h\|_1.$$

By substituting these relations into the dual objective (18), we have

$$-c^T\lambda_x - h^T\mu_x + \Delta e^T u_x + \Delta e^T v_x \geq \Delta\tau - \nu\|c_+\|_1 - \nu\|h\|_1.$$

Finally, we decrease $\bar{\epsilon}_2$ further if necessary so that

$$\Delta\tau - \nu\|c(x)_+\|_1 - \nu\|h(x)\|_1 \geq (\tau/2)\|x - x^*\|^\zeta,$$

for $\|x - x^*\| \leq \bar{\epsilon}_2$. We note that such a choice is possible since $\Delta \geq \|x - x^*\|^\zeta$ and $h(x)$ and $(c(x))_+$ are both $O(\|x - x^*\|)$. Hence the optimal objective in (18) is bounded below by $(\tau/2)\|x - x^*\|^\zeta$. This bound contradicts our earlier observation in (23) that the optimal objective is bounded above by a multiple of $\|x - x^*\|$. We conclude that $\tau = 0$ in (24), so $-g^* \in \text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i \in \mathcal{A}_c(x)}]$, as claimed. □

When the assumptions are made stronger, we obtain the following result.

Corollary 1 *If LICQ holds at x^*, for any ζ, $\bar{\epsilon}_2$, Δ, ν, and x satisfying the conditions of Theorem 3, $\mathcal{A}^* \setminus \mathcal{A}_0^* \subset \mathcal{A}_c(x) \subset \mathcal{A}^*$. If strict complementarity also holds at x^*, then $\mathcal{A}_c(x) = \mathcal{A}^*$.*

Proof. When LICQ holds at x^*, the multiplier (μ^*, λ^*) which satisfies equations (3) is unique, and $\lambda_i^* > 0$ for all $i \in \mathcal{A}^* \setminus \mathcal{A}_0^*$. For ζ, $\bar{\epsilon}_2$, Δ, and ν defined in Theorem 3, we must have $i \in \mathcal{A}_c(x)$ whenever $\lambda_i^* > 0$, since otherwise (21) would not hold. Thus, $\mathcal{A}^* \setminus \mathcal{A}_0^* \subset \mathcal{A}_c(x)$. Lemma 1 supplies $\mathcal{A}_c(x) \subset \mathcal{A}^*$. The final statement follows trivially from the equivalence of strict complementarity with $\mathcal{A}_0^* = \emptyset$. □

The implementation of SQP-EQP known as `Active` [5] [6], which is contained in the `Knitro` package, solves the formulation (17) using variants of the simplex method. It is observed (Waltz [20]) that many simplex iterations are spent in resolving the trust-region bounds $-\Delta e \leq d \leq \Delta e$. This effort would seem to be wasted; we are much more interested in the question of which linearized inequality constraints from (1) are active at the solution of (17) (and, ultimately, of (1)) than in the trust-region bounds. The authors of `Active` have tried various techniques to terminate the solution of (17) prematurely at an inexact solution, but these appear to increase the number of "outer" iterations of the SQP algorithm.

Because there is no curvature, trust-region bounds in (16) may be active, regardless of the size of Δ, even when x is arbitrarily close to x^*. The theorems above highlight the importance of choosing Δ large enough to allow constraints in $\mathcal{A}^*$ to become active in (16) but small enough to prevent inactive constraints (those not in $\mathcal{A}^*$) becoming active in (16). Byrd et al. [5, Section 3] describe a heuristic for `Active` in which Δ is adjusted from its value at the previous iteration of the outer algorithm according to success of the QP step, the norms of the QP step and the solution d of (16), and whether or not the minimizer of the quadratic model in this direction d lies at the boundary of the trust region.

The performance of these schemes also depends strongly on the value of ν in (16) and (19). The bound

$$\nu \geq \max\left(\max_j \lambda_j^*, \max_k |\mu_k^*|\right) \tag{27}$$

ensures global convergence. However, excessively large estimates of ν can slow convergence. The heuristic used in `Active` [5, Section 9] re-solves the

LP for increasing values of ν whenever a substantial decrease in infeasibility is possible. In addition, ν is decreased whenever the bound (27) (using the current multiplier estimates) is inactive for several consecutive successful, feasible LP iterations.

Theorem 3 and Corollary 1 suggest that the approaches of this section may give false negatives for constraints that are weakly active, or which may have an optimal Lagrange multiplier of zero. However, it is not obvious that failure to identify such constraints would adversely affect the performance of nonlinear programming algorithms. To first order, they are not critical to satisfying the KKT conditions.

3.2 A Technique Based on the Primal-Dual Estimate

Here we describe a scheme based on explicit minimization of $\psi(x, \mu, \lambda)$ in (11) with respect to (μ, λ) for $\lambda \geq 0$. We show that this minimization problem can be formulated as a linear program with equilibrium constraints (LPEC), one that is related to the linear programs discussed in Subsection 3.1. However, in contrast to this earlier approach, we use a threshold test like (15) to estimate the active set, rather than active set or Lagrange multiplier information from the subproblem.

The LPEC subproblem is as follows:

$$\omega(x) \stackrel{\text{def}}{=} \min_{\lambda \geq 0, \mu} \sum_{i=1}^{m} |\min(\lambda_i, -c_i)| + \|h\|_1 + \|A^T\lambda + J^T\mu + g\|_1. \tag{28}$$

The activity test $\mathcal{A}_{\text{lpec}}$ is defined as

$$\mathcal{A}_{\text{lpec}}(x) = \{i \,|\, c_i(x) \geq -(\beta\omega(x))^\sigma\}, \tag{29}$$

where $\beta > 0$ and $\sigma \in (0, 1)$ are constants.

The problem (28) can be formulated as the following LPEC:

$$\omega(x) \stackrel{\text{def}}{=} \min_{(\lambda,\mu,s,u,v)} e^T s + \textstyle\sum_{c_i \geq 0} c_i + \|h\|_1 + e^T u + e^T v, \quad \text{subject to} \tag{30a}$$

$$0 \leq (-c)_+ - s \;\perp\; \lambda - s \geq 0, \tag{30b}$$

$$A^T\lambda + J^T\mu + g = u - v, \qquad (\lambda, s, u, v) \geq 0. \tag{30c}$$

By introducing a large constant M and binary variables y_i, $i = 1, 2, \ldots, m$ (which take on the value 0 if the minimum in $\min(-c_i, \lambda_i)$ is achieved by $-c_i$ and 1 if it is achieved by λ_i), we can write (30) as the following mixed integer (binary) program:

$$\omega(x) \stackrel{\text{def}}{=} \min_{(\lambda,\mu,s,y,u,v)} e^T s + \|h\|_1 + e^T u + e^T v, \quad \text{subject to} \tag{31a}$$

$$-c_i - s_i \leq -c_i y_i, \;\; i = 1, 2, \ldots, m, \tag{31b}$$

$$\lambda_i - s_i \leq M(1 - y_i), \;\; i = 1, 2, \ldots, m, \tag{31c}$$

$$A^T\lambda + J^T\mu + g = u - v, \tag{31d}$$

$$(\lambda, u, v) \geq 0, \qquad s \geq (c)_+, \qquad y_i \in \{0, 1\}, \;\; i = 1, 2, \ldots m. \tag{31e}$$

The validity of this formulation for (28) is based on the nonnegativity of λ and the minimization of the $e^T s$ term. The large parameter M is necessary for (31c) but not for (31b), because $-c$ is a parameter while λ is a variable in the program.

There are notable similarities between the formulation (28) of the LPEC subproblem and the dual formulation (19) of the previous subsection. First, the term $\|A^T\lambda + J^T\mu + g\|_1$ appears in both objectives, though in (19) it is weighted by the trust-region radius Δ. Second, the term $\|h\|_1$ in (28) (which is constant in (28) and (31a)) corresponds to the rather different term $-\mu^T h$ in (19). Third, the parameter ν which penalizes constraint violation does not appear in (28). Fourth, and perhaps most interestingly, the minimum $|\min(-c_i, \lambda_i)|$ in (28) is replaced by the product $(-c_i)\lambda_i$ in (19). While the use of the minimum may lead to stronger identification properties (see below), it is responsible for the presence of equilibrium constraints in (28) and therefore makes the subproblem much harder to solve. In addition, the attractive scale invariance property possessed by the $-c_i\lambda_i$ term is lost. If we multiply c_i and A_i by some $\sigma_i > 0$ and replace $\lambda_i \leftarrow \lambda_i/\sigma_i$ to maintain constancy of the product $A_i\lambda_i$, the minimum $|\min(-c_i, \lambda_i)|$ will be replaced by $|\min(-\sigma_i c_i, \lambda_i/\sigma_i)|$, which has a different value in general.

We now show that $\omega(x)$ defined in (28) provides a two-sided estimate of the distance to the solution and that the identification scheme (29) eventually is successful, under appropriate assumptions.

Theorem 4 *Suppose that the KKT conditions (3), the MFCQ (7), and the second-order condition (9) are satisfied at x^*, and let ϵ be as defined in Theorem 1. Then there are positive constants $\bar{\epsilon} \in (0, \epsilon/2]$ and $\bar{C}$ such that for all x with $\|x - x^*\| \le \bar{\epsilon}$, we have that*

(i) the minimum in (28) is achieved at some (μ, λ) with $\operatorname{dist}((\mu, \lambda), \mathcal{S}_D) \le \epsilon/2$;
(ii) $\bar{C}^{-1}\omega(x) \le \|x - x^*\| \le \bar{C}\omega(x)$; and
(iii) $\mathcal{A}_{\mathrm{lpec}}(x) = \mathcal{A}^*$.

Proof.
(i) Note first that for any $(\mu^*, \lambda^*) \in \mathcal{S}_D$ and any x with $\|x - x^*\| \le \epsilon$, we have that

$$\begin{aligned}
\omega(x) &\le \psi(x, \mu^*, \lambda^*) \\
&= \sum_{i=1}^{m} |\min(\lambda_i^*, -c_i(x))| + \|h(x)\|_1 + \|A(x)^T\lambda^* + J(x)^T\mu^* + g(x)\|_1 \\
&\le \sum_{i=1}^{m} |c_i(x) - c_i(x^*)| + \|h(x) - h(x^*)\|_1 \\
&\qquad + \|(A(x) - A(x^*))^T\lambda^* + (J(x) - J(x^*))^T\mu^* + (g(x) - g(x^*))\|_1 \\
&\le C_1\|x - x^*\|, \qquad (32)
\end{aligned}$$

for some constant C_1. (In the second-last inequality, we used $\min(\lambda_i^*, -c_i(x^*)) = 0$, which follows from (3c).) Hence, if the minimum in (28) occurs outside the set $\{(\mu, \lambda) | \lambda \ge 0, \operatorname{dist}((\mu, \lambda), \mathcal{S}_D) \le \epsilon/2\}$ for x arbitrarily close to x^*, we

must be able to choose a sequence (x^k, μ^k, λ^k) with $x^k \to x^*$, $\lambda^k \geq 0$, and $\text{dist}\,((\mu^k, \lambda^k), \mathcal{S}_D) > \epsilon/2$ such that

$$\psi(x^k, \mu^k, \lambda^k) \leq \psi(x^k, \mu^*, \lambda^*) \leq C_1 \|x^k - x^*\|, \text{ for all } k.$$

In particular we have $\psi(x^k, \mu^k, \lambda^k) \to 0$. Consider first the case in which (μ^k, λ^k) is unbounded. By taking a subsequence if necessary, we can assume that

$$\|(\mu^k, \lambda^k)\| \to \infty, \quad \frac{(\mu^k, \lambda^k)}{\|(\mu^k, \lambda^k)\|} \to (\mu^*, \lambda^*), \quad \|(\mu^*, \lambda^*)\| = 1, \quad \lambda^* \geq 0.$$

For any $i \notin \mathcal{A}^*$, we have by taking a further subsequence if necessary that $c_i(x^k) < (1/2)c_i(x^*) < 0$, for all k. Since $|\min(\lambda_i^k, -c_i(x^k))| \leq \psi(x^k, \mu^k, \lambda^k) \to 0$, we have that $\lambda_i^k \to 0$ and thus $\lambda_i^* = 0$ for all $i \notin \mathcal{A}^*$. We also have that $A(x^k)^T\lambda^k + J(x^k)^T\mu^k + g(x^k) \to 0$, so when we divide this expression by $\|(\mu^k, \lambda^k)\|$ and take limits, we obtain

$$A(x^*)^T\lambda^* + J(x^*)^T\mu^* = A_{\mathcal{A}^*}(x^*)^T\lambda^*_{\mathcal{A}^*} + J(x^*)^T\mu^* = 0.$$

We can now use the usual argument based on the MFCQ property (7) (see Appendix A) to deduce that $\lambda^*_{\mathcal{A}^*} = 0$ and then $\mu^* = 0$, contradicting $\|(\mu^*, \lambda^*)\| = 1$. Hence, the sequence (μ^k, λ^k) must be bounded.

By taking a subsequence if necessary, we can define a vector $(\hat{\mu}, \hat{\lambda})$ such that

$$(\mu^k, \lambda^k) \to (\hat{\mu}, \hat{\lambda}), \quad \hat{\lambda} \geq 0.$$

The limit $\psi(x^k, \mu^k, \lambda^k) \to 0$ thus implies that $\psi(x^*, \hat{\mu}, \hat{\lambda}) = 0$, which in turn implies that $(\hat{\mu}, \hat{\lambda}) \in \mathcal{S}_D$, contradicting $\text{dist}\,((\mu^k, \lambda^k), \mathcal{S}_D) > \epsilon/2$. Thus, there is some $\bar{\epsilon}$ such that for all x with $\|x - x^*\| \leq \bar{\epsilon}$ the minimum occurs in the set $\{(\mu, \lambda) | \lambda \geq 0, \text{dist}\,((\mu, \lambda), \mathcal{S}_D) \leq \epsilon/2\}$. Since this set is compact (boundedness of $\mathcal{S}_D$ follows from the MFCQ (7)), we conclude that the minimum in (28) is attained by some (μ, λ) in this set.

(ii) The left-hand inequality is already proved by (32). We now show that, for the $\bar{\epsilon} \in (0, \epsilon/2]$ determined in part (i), we have

$$\|x - x^*\| \leq C\omega(x) \text{ for all } x \text{ with } \|x - x^*\| \leq \bar{\epsilon}, \tag{33}$$

for C defined in Theorem 1. First note that for any (μ, λ) with $\text{dist}\,((\mu, \lambda), \mathcal{S}_D) \leq \epsilon/2$, we have

$$\text{dist}\,((x, \mu, \lambda), \mathcal{S}) \leq \|x - x^*\| + \text{dist}\,((\mu, \lambda), \mathcal{S}_D) \leq \bar{\epsilon} + \epsilon/2 \leq \epsilon,$$

so that from Theorem 1 we have

$$\|x - x^*\| \leq \text{dist}\,((x, \mu, \lambda), \mathcal{S}) \leq C\psi(x, \mu, \lambda), \tag{34}$$

for all (μ, λ) with $\text{dist}\,((\mu, \lambda), \mathcal{S}_D) \leq \epsilon/2$ and $\lambda \geq 0$. We showed in part (i) that the minimum of $\psi(x, \mu, \lambda)$ is attained in this set, for sufficiently small choice of $\bar{\epsilon}$. Hence, we have

$$\|x - x^*\| \leq C \min_{\lambda \geq 0, \text{dist}\,((\mu,\lambda),\mathcal{S}_D) \leq \epsilon/2} \psi(x, \mu, \lambda) = C\omega(x),$$

as required. The result follows by taking $\bar{C} = \max(C, C_1)$, where C_1 is from (32).

(iii) The proof of this final claim follows from an argument like that of Theorem 2. □

We note that an exact solution of (28) (or (31)) is not needed to estimate the active set. In fact, any approximate solution whose objective value is within a chosen fixed factor of the optimal objective value will suffice to produce an asymptotically accurate estimate. Computationally speaking, we can terminate the branch-and-bound procedure at the current incumbent once the lower bound is within a fixed factor of the incumbent objective value. Moreover, we can derive an excellent starting point for (31) from the solution of the dual subproblem (18) of the previous subsection or from the linear programming subproblem of the next section. (As our experiments of Section 4 show, the branch and bound procedure often terminates at the root node, without doing any expansion of the branch-and-bound tree at all. When this occurs, the main computational cost is the cost of solving a single linear programming relaxation of (31)).

The main differences between the schemes of this subsection and the previous one can be summarized as follows:

- When a second-order sufficient condition holds, the scheme of this subsection accurately estimates $\mathcal{A}^*$ (including the weakly active constraints), whereas the schemes of the previous subsection may only identify those active constraints that are instrumental in satisfying the first KKT condition (3a).
- Effectiveness of the techniques of the previous subsection depends critically on the choice of trust-region radius Δ, whereas no such parameter is present in this subsection. However, the practical performance of the latter approach may depend on the scaling of the constraints c_i and their multipliers λ_i. Performance may be improved for some problems by changing the relative weightings of the terms $\|h\|_1$ and $\|A^T\lambda + J^T\mu + g\|_1$ in $\omega(x)$. However, it is difficult to determine a choice of weights that works reliably for a range of problems.

3.3 A Linear Programming Approximation to the LPEC

In this section, we describe a technique that has the same identification properties as the scheme of the previous subsection, as described in Theorem 4, but requires only the solution of a linear program, rather than an LPEC. The key to the scheme is to obtain a two-sided bound on $\omega(x)$, defined in (28), that can be obtained by solving a single linear program.

We start by defining the following functions:

$$\rho(x,\mu,\lambda) \stackrel{\text{def}}{=} \sum_{c_i<0} -c_i\lambda_i + \sum_{c_i\geq 0} c_i + \|h\|_1 + \|A^T\lambda + J^T\mu + g\|_1, \tag{35}$$

$$\bar{\rho}(x,\mu,\lambda) \stackrel{\text{def}}{=} \sum_{c_i<0} (-c_i\lambda_i)^{1/2} + \sum_{c_i\geq 0} c_i + \|h\|_1 + \|A^T\lambda + J^T\mu + g\|_1. \tag{36}$$

These functions are related in the following elementary fashion.

Lemma 2 *For any (μ, λ) with $\lambda \geq 0$, we have*

$$\bar{\rho}(x, \mu, \lambda) \leq \rho(x, \mu, \lambda) + \sqrt{m}\rho(x, \mu, \lambda)^{1/2}.$$

Proof.

$$\begin{aligned}
\bar{\rho}(x, \mu, \lambda) &= \left\| \left[(-c_i \lambda_i)^{1/2} \right]_{c_i < 0} \right\|_1 + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|A^T \lambda + J^T \mu + g\|_1 \\
&\leq \sqrt{m} \left\| \left[(-c_i \lambda_i)^{1/2} \right]_{c_i < 0} \right\|_2 + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|A^T \lambda + J^T \mu + g\|_1 \\
&= \sqrt{m} \left[\sum_{c_i < 0} (-c_i \lambda_i) \right]^{1/2} + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|A^T \lambda + J^T \mu + g\|_1 \\
&\leq \sqrt{m}\rho(x, \mu, \lambda)^{1/2} + \rho(x, \mu, \lambda).
\end{aligned}$$

□

The next result defines the relationship between ρ, $\bar{\rho}$, and the proximality measure ψ defined in (11).

Lemma 3 *Let $K_2 \geq 1$ be given. Then for all (x, μ, λ) with $\lambda \geq 0$ and*

$$\|c\|_\infty \leq K_2, \quad \|\lambda\|_\infty \leq K_2, \tag{37}$$

we have that

$$K_2^{-1} \rho(x, \mu, \lambda) \leq \psi(x, \mu, \lambda) \leq \bar{\rho}(x, \mu, \lambda). \tag{38}$$

Proof. For $c_i < 0$ and $\lambda_i \geq 0$, we have

$$-c_i \lambda_i = \min(-c_i, \lambda_i) \max(-c_i, \lambda_i) \geq \min(-c_i, \lambda_i)^2, \tag{39}$$

and also

$$-c_i \lambda_i \leq K_2 \min(-c_i, \lambda_i). \tag{40}$$

From (39) we have

$$\begin{aligned}
\psi(x, \mu, \lambda) &= \sum_{c_i < 0} |\min(-c_i, \lambda_i)| + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|g + A^T \lambda + J^T \mu\|_1 \\
&\leq \sum_{c_i < 0} (-c_i \lambda_i)^{1/2} + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|g + A^T \lambda + J^T \mu\|_1 \\
&= \bar{\rho}(x, \mu, \lambda),
\end{aligned}$$

thereby proving the right-hand inequality in (38).

For the left-hand inequality, we have from (40) and $K_2 \geq 1$ that

$$\psi(x, \mu, \lambda) \geq K_2^{-1} \sum_{c_i < 0} (-c_i \lambda_i) + \sum_{c_i \geq 0} c_i + \|h\|_1 + \|A^T \lambda + J^T \mu + g\|_1 \geq K_2^{-1} \rho(x, \mu, \lambda),$$

as required. □

We are particularly interested in the solution (μ_x, λ_x) to the program

$$\min_{\mu, 0 \le \lambda \le K_1 e} \rho(x, \mu, \lambda), \tag{41}$$

where K_1 is the constant defined in (14). The problem of determining (μ_x, λ_x) can also be expressed as the following linear program:

$$\min_{(\lambda,\mu,u,v)} \sum_{c_i<0}(-c_i\lambda_i) + \sum_{c_i \ge 0} c_i + \|h\|_1 + e^T u + e^T v, \quad \text{subject to} \tag{42a}$$

$$A^T\lambda + J^T\mu + g = u - v, \;\; 0 \le \lambda \le K_1 e, \;\; (u, v) \ge 0. \tag{42b}$$

We define the activity test associated with $\bar{\rho}$ as follows:

$$\mathcal{A}_{\bar{\rho}}(x) = \{i = 1, 2, \ldots, m \,|\, c_i(x) \ge -(\beta\bar{\rho}(x, \mu_x, \lambda_x))^{\bar{\sigma}}\}, \tag{43}$$

for given constants $\beta > 0$ and $\bar{\sigma} \in (0, 1)$.

We now prove a result similar to Theorem 4, showing in particular that under the same assumptions as the earlier result, the identification scheme above is asymptotically successful.

Theorem 5 *Suppose that the KKT conditions (3), the MFCQ (7), and the second-order condition (9) are satisfied at x^*, and let ϵ be as defined in Theorem 1. Then there exists a positive constant $\hat{\epsilon} \in (0, \epsilon/2]$ such that for all x with $\|x - x^*\| \le \hat{\epsilon}$, we have*

(i) the minimum in (42) is attained at some (μ, λ) with $\mathrm{dist}\,((\mu, \lambda), \mathcal{S}_D) \le \epsilon/2$;

(ii) $K_1^{-1}\rho(x, \mu_x, \lambda_x) \le \omega(x) \le \bar{\rho}(x, \mu_x, \lambda_x)$, where K_1 is the constant defined in (14); and

(iii) $\mathcal{A}_{\bar{\rho}}(x) = \mathcal{A}^*$.

Proof.

(i) Note that for any $(\mu^*, \lambda^*) \in \mathcal{S}_D$ and any x with $\|x - x^*\| < \epsilon$, we have

$$\begin{aligned}
&\rho(x, \mu^*, \lambda^*) \\
&= \sum_{c_i<0} -c_i(x)\lambda_i^* + \sum_{c_i \ge 0} c_i(x) + \|h(x)\|_1 + \|A(x)^T\lambda^* + J(x)^T\mu^* + g(x)\|_1 \\
&\le \sum_{c_i<0} (c_i(x^*) - c_i(x))\lambda_i^* + \sum_{c_i \ge 0} (c_i(x) - c_i(x^*)) + \|h(x) - h(x^*)\|_1 \\
&\qquad + \|(A(x) - A(x^*))^T\lambda^* + (J(x) - J(x^*))^T\mu^* + (g(x) - g(x^*))\|_1 \\
&\le C_2\|x - x^*\|,
\end{aligned}$$

for some constant C_2. (In the first inequality above, we used the fact $\lambda_i^* c_i^* = 0$ for all i to bound the first summation, and the fact that $c_i^* \le 0$ for all i to bound the second summation.) Note that since $\|(\mu^*, \lambda^*)\|_\infty \le K_1$, we have $0 \le \lambda^* \le K_1 e$, so that (μ^*, λ^*) together with an obvious choice of (u, v), is feasible for (42). We note also that any $(\hat{\mu}, \hat{\lambda})$ for which $\rho(x^*, \hat{\mu}, \hat{\lambda}) = 0$ and $\hat{\lambda} \ge 0$ satisfies $(\hat{\mu}, \hat{\lambda}) \in \mathcal{S}_D$. Using these observations, the remainder of the proof closely parallels that of Theorem 4 (i), so we omit the details.

(ii) Reduce $\hat{\epsilon}$ if necessary to ensure that $\hat{\epsilon} \leq \bar{\epsilon} \leq \epsilon/2$, where $\bar{\epsilon}$ is defined in Theorem 4. Reduce $\hat{\epsilon}$ further if necessary to ensure that $\|c(x)\|_\infty \leq K_1$ for all x with $\|x - x^*\| \leq \hat{\epsilon}$. Note that by Theorem 4(i), the minimizer of (28) has dist $((\mu, \lambda), \mathcal{S}_D) \leq \epsilon/2$, and therefore $\|\lambda\|_\infty \leq \|\lambda^*\|_\infty + 1 \leq K_1$ for any $(\mu^*, \lambda^*) \in \mathcal{S}_D$.

Using the result of Lemma 3 (with K_1 replacing K_2), we have that

$$K_1^{-1}\rho(x, \mu_x, \lambda_x)$$
$$= \min_{\mu, 0 \leq \lambda \leq K_1 e} K_1^{-1}\rho(x, \mu, \lambda) \leq \min_{\mu, 0 \leq \lambda \leq K_1 e} \psi(x, \mu, \lambda) \leq \psi(x, \mu_x, \lambda_x) \leq \bar{\rho}(x, \mu_x, \lambda_x).$$

However, as we showed in Theorem 4(i), the minimizer of $\psi(x, \mu, \lambda)$ over the set of (μ, λ) with $\lambda \geq 0$ is attained at values of (μ, λ) that satisfy the restriction $\|\lambda\|_\infty \leq K_1$, so we can write

$$K_1^{-1}\rho(x, \mu_x, \lambda_x) \leq \min_{\mu, 0 \leq \lambda} \psi(x, \mu, \lambda) \leq \bar{\rho}(x, \mu_x, \lambda_x),$$

which yields the result, by (28).

(iii) We have from Lemma 2, Theorem 4(ii), and part (ii) of this theorem that $\bar{\rho}(x, \mu_x, \lambda_x) \to 0$ as $x \to x^*$. Therefore, using continuity of c_i, $i = 1, 2, \ldots, m$, we can decrease $\hat{\epsilon}$ if necessary to ensure that for $\|x - x^*\| \leq \hat{\epsilon}$, we have

$$c_i(x) < (1/2)c_i(x^*) \leq -(\beta\bar{\rho}(x, \mu_x, \lambda_x))^{\bar{\sigma}}, \quad \text{for all } i \notin \mathcal{A}^*.$$

Hence, $i \notin \mathcal{A}_{\bar{\rho}}(x)$ for all such x.

For $i \in \mathcal{A}^*$, we have for the Lipschitz constant L defined in (13), and using Theorem 4(ii) and part (ii) of this theorem that

$$\begin{aligned}
|c_i(x)| &\leq L\|x - x^*\| \\
&= L\|x - x^*\|^{1-\bar{\sigma}}\|x - x^*\|^{\bar{\sigma}} \\
&\leq L\|x - x^*\|^{1-\bar{\sigma}}\bar{C}^{\bar{\sigma}}\omega(x)^{\bar{\sigma}} \\
&\leq \left[L\|x - x^*\|^{1-\bar{\sigma}}\bar{C}^{\bar{\sigma}}/\beta^{\bar{\sigma}}\right](\beta\bar{\rho}(x, \mu_x, \lambda_x))^{\bar{\sigma}} \\
&\leq (\beta\bar{\rho}(x, \mu_x, \lambda_x))^{\bar{\sigma}},
\end{aligned}$$

for $\hat{\epsilon}$ sufficiently small. Hence, we have $i \in \mathcal{A}_{\bar{\rho}}(x)$ for all x with $\|x - x^*\| \leq \hat{\epsilon}$. □

Near the solution, $\omega(x)$ may be (and often is) much smaller than $\bar{\rho}(x, \mu_x, \lambda_x)$, because of the looseness of the estimate (39). To compensate for this difference, we set $\bar{\sigma}$ in the definition of $\mathcal{A}_{\bar{\rho}}$ (43) to be larger than σ in the definition of $\mathcal{A}_{\text{lpec}}$ (29) in the tests described in the next section.

A referee has pointed out that some interesting insights are available from examination of the dual of the subproblem (42). Ignoring the upper bound $\lambda \leq K_1 e$, we can write the dual as

$$\begin{aligned}
\min g^T d, &\quad \text{subject to} \\
A_i d + c_i \leq 0 &\quad \text{for } i \text{ with } c_i < 0, \\
A_i d \leq 0 &\quad \text{for } i \text{ with } c_i \geq 0, \\
Jd = 0, &\quad -e \leq d \leq e.
\end{aligned}$$

	LP-D	LP-P	LPEC-A	LPEC
Rows	n	$m+p$	n	$5m+n$
Columns	$m+2n+p$	$m+n+2p$	$m+2n+p$	$4m+2n+p$

Table 1 Problem dimensions as a function of the number of inequalities (m), variables (n), and equalities (p).

It is not difficult to construct examples for which $A_i d + c_i = 0$ for an inactive constraint i, even when x is arbitrarily close to x^*. (The referee gave the example of minimizing a scalar x^2 subject to $-x - 0.5 \leq 0$, for x slightly greater than the optimum $x^* = 0$.) Thus, if the active set estimate were obtained from formulae such as (20), it may not be asymptotically accurate. Hence, the use of the threshold test (43) in place of activity tests (20) are key to the effectiveness of the approach of this section. In this vein, it can be shown that if the (μ, λ) components of the solution of the earlier linear programming problem (18) are inserted into the threshold test (43) in place of (μ_x, λ_x), an asymptotically accurate estimate is obtained, under certain reasonable assumptions. We omit a formal statement and proof of this claim, as we believe (μ_x, λ_x) to be a better choice of the Lagrange multipliers, because their calculation does not depend on the parameters ν and Δ that appear in (18).

4 Computational Results

In this section, we apply the techniques of Section 3 to a variety of problems in which x is slightly perturbed from its (approximately) known solution x^*. The resulting active-set estimate is compared with our best guess of the active set at the solution. We report the false positives and false negatives associated with each technique, along with the runtimes required to execute the tests.

The linear programming techniques of Subsection 3.1 are referred to as LP-P for the primal formulation (17) and LP-D for the dual formulation (18). For these formulations, we use both activity tests $\mathcal{A}_c$ and $\mathcal{A}_\lambda$ of (20), modified slightly with activity thresholds. We also implement the LPEC scheme of Subsection 3.2 and the linear programming approximation scheme of Subsection 3.3, which we refer to below as LPEC-A. We implemented all tests in C, using the CPLEX callable library (version 9.0) to solve the linear and mixed integer programs.

The times required to implement the tests are related to the size and density of the constraint matrix for each formulation. The matrix dimensions for each formulation are given in Table 1. Except for problems with many equality constraints, the LPEC formulation has the largest constraint matrix. Further, it is the only formulation with binary variables.

Subsection 4.1 discusses some specifics of the formulations, such as the choice of parameters and tolerances in the identification procedures. In Subsection 4.2, we apply the identification techniques to a set of random problems, for which we have control over the dimensions and amount of degeneracy. In Subsection 4.3, we consider a subset of constrained problems from

the CUTEr test set, a conglomeration of problems arising from theory, modeling, and real applications [13]. While the random problems are well scaled with dense constraint Jacobians, the CUTEr problems may be poorly scaled and typically have sparse constraint Jacobians. Subsection 4.4 contains some remarks about additional testing.

4.1 Implementation Specifics

4.1.1 Choosing Parameters and Tolerances

The following implementation details are common to both random and CUTEr test sets. We bound the ℓ_∞ norm of the perturbation x from the (approximately) optimal point x^* by a noise parameter `noise`. Denoting by ϕ a random variable drawn from the uniform distribution on $[-1, 1]$, we define the perturbed point x as follows:

$$x_i = x_i^* + \frac{\texttt{noise}}{n}\phi, \quad i = 1, 2, \ldots, n. \tag{44}$$

A second parameter `DeltaFac` controls the bound on the trust-region radius for the LP-P and LP-D programs. We set

$$\Delta = \texttt{DeltaFac}\frac{\texttt{noise}}{n},$$

so that when $\texttt{DeltaFac} \geq 1$, the trust region is large enough to contain the true solution x^*. For the results tabulated below, we use $\texttt{DeltaFac} = 4$. This value is particularly felicitous for the LP-P and LP-D schemes, as it yields a Δ large enough to encompass the solution yet small enough to exclude many inactive constraints. The number of false positives therefore tends to be small for LP-P and LP-D in our tables. The relatively small trust region also allows the CPLEX presolver to streamline the linear programming formulations before calling the simplex code, thus reducing the solve times for the linear programs in LP-P and LP-D. (Specifically, for each inequality constraint that is inactive over the entire trust region, the LP-P subproblem is reduced by one row and column, while the LP-D subproblem is reduced by one column.) It is unlikely that a nonlinear programming algorithm that uses LP-P or LP-D as its identification technique could choose a value of Δ as nice as the one used in these tests in practice.

The activity tests $\mathcal{A}_c$ (20a) and $\mathcal{A}_\lambda$ (20b) were modified to include a tolerance, as follows:

$$\mathcal{A}_c(x) = \{i \,|\, A_i d + c_i \geq -\epsilon_0\} \tag{45}$$

and

$$\mathcal{A}_\lambda(x) = \{i \,|\, \lambda_i \geq \epsilon_0\}, \tag{46}$$

with $\epsilon_0 = 10^{-4}$.

In the tests $\mathcal{A}_{\rm lpec}(x)$ (29) for LPEC and $\mathcal{A}_{\bar{\rho}}$ (43) for LPEC-A, we set $\beta = 1/(m+n+p)$. The value 0.75 is used for σ in $\mathcal{A}_{\rm lpec}$ (29), while the larger value 0.90 is used for $\bar{\sigma}$ in $\mathcal{A}_{\bar{\rho}}$ (43).

By default, the mixed-integer solver in CPLEX makes use of various cut generation schemes, including flow covers, MIR cuts, implied bound cuts, and Gomory fractional cuts. We disabled these schemes because, given our usually excellent starting point for the LPEC test, the cost of cut generation is excessive compared to the cost of solving the root relaxation. However, for all tests, we allowed both linear and mixed-integer solvers to perform their standard presolving procedures, as they generally improved the performance. For the mixed-integer solver in LPEC, we accept the solution if it is within a factor of 2 of the lower bound.

For both linear and integer programming solvers, we tightened the general feasibility tolerance `eprhs` from 10^{-6} to 10^{-9} because problems in the CUTEr test set (notably BRAINPC0 and BRAINPC3) report infeasibilities after scaling for LPEC-A under the default tolerance. In addition, we observed LP-P objective values that were too negative when using default `eprhs` values. Specifically, for some of the constraints $A_i d + c_i - r_i \leq 0$ in (17), the solver would find a d with $A_i d + c_i$ slightly positive, while setting r_i to zero. Thus, the constraint would be satisfied to the specified tolerance, while avoiding the larger value of $g^T d$ that would be incurred if it were satisfied exactly.

4.1.2 Formulation Details

For all test problems, the parameter ν of the LP-P and LP-D programs is assigned a value large enough to ensure that the known optimal multipliers (μ^*, λ^*) are feasible for (19). The results of this paper, theoretical and computational, are otherwise insensitive to the choice of ν. (However, the choice of ν appears to be important for global convergence of the nonlinear programming algorithm, as discussed in Byrd et al. [6].)

The computational efficiency of the LPEC mixed integer program (31) is sensitive to the magnitude of M. Recall that the formulation (31) is identical to the LPEC (30) provided that M is sufficiently large, in particular, larger than $\|c+\lambda^*\|_\infty$, where λ^* is an optimal multiplier. However, excessively large M values may result in long runtimes. We observed runtime reductions of as much as 50% when we replaced heuristically-chosen values of M with near-minimal values.

We describe some heuristics for setting M and ν in the following subsections.

In solving LPEC, we use a starting point based on the solution for LPEC-A. Specifically, we set λ, μ, u, and v to their optimal values from (42); set $y_i = 0$ if $-c_i < \lambda_i$ and $y_i = 1$ otherwise; and set $s_i = |\min(-c_i, \lambda_i)|$. In most cases, this starting point is close to an acceptable solution for LPEC and little extra work is needed beyond solving an LP relaxation of the LPEC at the root node and verifying that the starting point is not far from the lower bound obtained from this relaxation. The solution to LP-D also provides a useful starting point for LPEC in most cases.

For LPEC and LPEC-A, no attempt is made to scale the constraints c_i or the components of the threshold functions $\omega(x)$ and $\bar{\rho}(x, \mu_x, \lambda_x)$. Heuristics to adjust such weightings may improve the performance of LPEC and LPEC-A techniques.

4.2 Random Problems

We generate random problems involving dense Jacobians J and A to mimic the behavior of a nonlinear program near a local solution x^*. Besides choosing the dimensions n, m, and p, we influence the amount of degeneracy in the problem by specifying the row rank of J and A and the proportion of weakly active constraints.

4.2.1 Problem Setup

Parameters specific to the random problem setup are `fStrong` and `fWeak` (approximate proportion of strongly and weakly active inequality constraints, respectively) and `degenA` and `degenJ` (proportional to the ranks of the null spaces of A and J, respectively). We first fill out the first $(1 - \texttt{degenA})m$ rows of the optimal inequality constraint Jacobian A^* with components 5ϕ (where, as above, ϕ represents a random variable uniformly distributed in $[-1, 1]$). We then set the last $(\texttt{degenA})m$ rows of A^* to be random linear combinations of the first $(1 - \texttt{degenA})m$ rows, where the coefficients of the linear combinations are chosen from ϕ. A similar process is used to choose the optimal equality constraint Jacobian J^* using the parameter `degenJ`.

We set the solution to be $x^* = 0$. Recall that x is a perturbation of x^* (44). First, we set each component of μ^* to $\frac{1}{2}\phi(\phi + 1)$. Next, we randomly classify each index $i \in \{1, 2, \ldots, m\}$ as "strongly active," "weakly active," or "inactive," such that the proportion in each category is approximately `fStrong`, `fWeak`, and $(1 - \texttt{fStrong} - \texttt{fWeak})$, respectively. For the inactive components, we set $c_i^* = -\frac{5}{2}(\phi + 1)^2$, while for the strongly active components, we set $\lambda_i^* = \frac{5}{2}(\phi + 1)^2$. Other components of c^* and λ^* are set to zero. To make the optimality condition (3a) consistent, we now set $g^* = -(A^*)^T\lambda^* - (J^*)^T\mu^*$. Naturally, $h^* = 0$.

In accordance with the assumed Lipschitz properties, we set

$$\begin{aligned} g_i &= g_i^* + (\texttt{noise}/n)\phi, \quad i = 1, 2, \ldots, m, \\ A_{ij} &= A_{ij}^* + (\texttt{noise}/n)\phi, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n, \\ J_{ij} &= J_{ij}^* + (\texttt{noise}/n)\phi, \quad i = 1, 2, \ldots, p; \quad j = 1, 2, \ldots, n. \end{aligned}$$

Since $c(x) = c^* + A^*(x - x^*) + O(\|x - x^*\|^2) = c^* + A^*x + O(\|x\|^2)$, we set

$$c_i = c_i^* + A_i^*x + (\texttt{noise}/n)^2\phi, \quad i = 1, 2, \ldots, m.$$

A similar scheme is used to set h.

The data thus generated is consistent to first order, but there is no explicit assurance that the second-order condition holds. (This condition is required

m	n	fStrong	LP-D		LP-P		LPEC-A	LPEC
			$\mathcal{A}_c$	$\mathcal{A}_\lambda$	$\mathcal{A}_c$	$\mathcal{A}_\lambda$		
50	200	0.10	0/0	0/0	0/0	0/0	1/0	1/0
50	200	0.50	0/0	0/0	0/0	0/0	0/0	0/0
50	1000	0.10	0/0	0/0	0/0	0/0	0/0	0/0
50	1000	0.50	0/0	0/0	0/0	0/0	0/0	0/0
100	200	0.10	0/0	0/0	0/0	0/0	0/0	0/0
100	200	0.50	0/0	0/0	0/0	0/0	1/0	0/0
100	1000	0.10	0/0	0/0	0/0	0/0	0/0	0/0
100	1000	0.50	0/0	0/1	0/0	0/1	0/0	0/0
400	200	0.10	0/0	0/0	0/0	0/0	2/0	1/1
400	1000	0.10	1/0	0/0	1/0	0/0	3/0	1/0
400	1000	0.50	1/0	0/0	1/0	0/0	4/0	3/0

Table 2 Nondegenerate Random Problems: False Positives/False Negatives. $p = n/5$, `noise`$= 10^{-3}$, `fWeak`$= 0.00$, `degenA`$= 0.00$, `DeltaFac`$= 4.00$.

m	n	fStrong	LP-D	LP-P	LPEC-A	LPEC
50	200	0.10	0.07	0.03	0.11	0.16
50	200	0.50	0.09	0.05	0.11	0.13
50	1000	0.10	7.08	4.61	7.34	7.94
50	1000	0.50	7.73	4.79	6.98	7.85
100	200	0.10	0.08	0.03	0.19	0.26
100	200	0.50	0.13	0.10	0.18	0.26
100	1000	0.10	7.05	4.46	9.35	9.99
100	1000	0.50	8.84	6.77	9.41	10.40
400	200	0.10	0.15	0.11	0.47	8.05
400	1000	0.10	9.80	5.59	20.70	171.24
400	1000	0.50	17.87	18.17	21.57	26.35

Table 3 Nondegenerate Random Problems: Time (secs). $p = n/5$, `noise`$= 10^{-3}$, `fWeak`$= 0.00$, `degenA`$= 0.00$, `DeltaFac`$= 4.00$.

for Theorems 4 and 5 concerning the exact identification properties of the LPEC and LPEC-A schemes.)

By setting ν to the large value 100, we ensure that solutions of LP-P and LP-D have $r = 0$ and $s = t = 0$. For the LPEC problem (31), we define $M = 5 \max_j(|c_j|)$. This value is large enough to secure local optimality of the LPEC programs of our test problems.

4.2.2 Nondegenerate Problems

Results for a set of random nondegenerate problems are shown in Table 2, with runtimes in Table 3. Nondegeneracy is assured by setting `fWeak` $= 0$, `degenA` $= 0$, and `degenJ` $= 0$. The number of equality constraints p is $n/5$ and we set `noise` $= 10^{-3}$. Each entry in Table 2 shows the numbers of false positives and false negatives for the problem and test in question. For LP-P and LP-D, we report both active sets $\mathcal{A}_c$ (45) and $\mathcal{A}_\lambda$ (46). The case $m = 400$, $n = 200$, `fStrong=0.50` does not appear because the expected number of degrees of freedom $(n - p - (\texttt{fStrong} + \texttt{fWeak})m)$ is nonpositive.

The identification techniques are accurate on these problems. Because the LICQ conditions hold (to high probability), even the LP-P and LP-D

m	n	fWeak	degenA	LP-D		LP-P		LPEC-A	LPEC
				$\mathcal{A}_c$	$\mathcal{A}_\lambda$	$\mathcal{A}_c$	$\mathcal{A}_\lambda$		
50	200	0.05	0.0	1/1	0/2	1/1	0/2	1/0	1/0
50	200	0.05	0.1	0/2	0/2	0/2	0/2	0/0	0/1
50	200	0.05	0.3	0/0	0/2	0/0	0/2	0/0	0/0
50	200	0.20	0.0	1/3	0/6	1/3	0/6	1/0	1/0
50	200	0.20	0.1	0/2	0/6	0/2	0/6	0/0	0/0
50	200	0.20	0.3	0/4	0/6	0/4	0/6	0/0	0/0
50	1000	0.05	0.0	0/0	0/2	0/0	0/2	0/0	0/0
50	1000	0.05	0.1	0/2	0/2	0/2	0/2	0/0	0/0
50	1000	0.05	0.3	0/2	0/2	0/2	0/2	0/1	0/1
50	1000	0.20	0.0	0/3	0/6	0/3	0/6	0/0	0/0
50	1000	0.20	0.1	0/3	0/6	0/3	0/6	0/0	0/0
50	1000	0.20	0.3	0/4	0/6	0/4	0/6	0/2	0/1
400	200	0.05	0.0	0/10	0/19	0/10	0/19	2/0	1/0
400	200	0.05	0.1	1/5	0/19	1/5	0/19	7/0	3/5
400	200	0.05	0.3	1/11	0/19	1/11	0/19	6/1	2/6
400	1000	0.05	0.0	2/8	0/19	2/9	0/19	3/0	1/0
400	1000	0.05	0.1	1/8	0/19	1/8	0/19	1/0	0/1
400	1000	0.05	0.3	4/7	0/19	4/7	0/19	6/0	5/7
400	1000	0.20	0.0	1/25	0/77	1/25	0/77	4/0	1/0
400	1000	0.20	0.1	0/22	0/77	0/22	0/77	1/0	0/6
400	1000	0.20	0.3	1/28	0/77	1/28	0/77	4/2	2/10

Table 4 Degenerate Random Problems: False Positives/False Negatives: $p = n/5$, `noise`$= 10^{-3}$, `fStrong`$= 0.20$, `DeltaFac`$= 4.00$

procedures are guaranteed to be asymptotically correct. Indeed, the LP-P and LP-D schemes generally perform best; the LPEC-A and LPEC schemes show a few false positives on the larger examples. For these problems, it is not necessary for the LPEC to search beyond the root node in the branch-and-bound tree, except in the case $m = 400$, $n = 200$, `fStrong` $= 0.10$, for which one additional node is considered.

In agreement with the theory of Section 3, the false positives reported in LPEC-A and LPEC results disappear for smaller noise values. In particular, for `noise` $= 10^{-7}$ the identification results are perfect for the LPEC-A and LPEC methods, while the LP-D and LP-P methods still give some errors.

Runtimes are shown in Table 3. The differences between the approaches are not significant, except for two of the $n = 400$ cases, for which LPEC is substantially slower than LPEC-A.

4.2.3 Degenerate Problems

Results for a set of random degenerate problems are shown in Table 4, with runtimes in Table 5. In these tables, we fixed `fStrong` $= 0.2$, `noise` $= 10^{-3}$, `degenJ` $= 0$, and $p = n/5$. The values of `fWeak` and `degenA` were varied, along with the dimensions m and n.

All methods perform well when $m = 50$. The LPEC and LPEC-A approaches rarely make an identification error on these problems, whereas LP-P and LP-D record a few false negatives. For the problems with 400 inequality constraints, the numbers of errors made by LPEC-A and LPEC are lower

m	n	fWeak	degenA	LP-D	LP-P	LPEC-A	LPEC
50	200	0.05	0.0	0.08	0.04	0.11	0.15
50	200	0.05	0.1	0.08	0.04	0.10	0.15
50	200	0.05	0.3	0.08	0.04	0.12	0.16
50	200	0.20	0.0	0.06	0.04	0.10	0.13
50	200	0.20	0.1	0.09	0.05	0.11	0.16
50	200	0.20	0.3	0.09	0.04	0.11	0.38
50	1000	0.05	0.0	7.51	4.05	6.99	7.81
50	1000	0.05	0.1	7.26	4.70	6.71	7.70
50	1000	0.05	0.3	7.08	4.47	7.22	17.15
50	1000	0.20	0.0	7.54	4.43	7.47	8.04
50	1000	0.20	0.1	8.11	4.28	6.80	7.75
50	1000	0.20	0.3	7.52	4.80	7.28	17.04
400	200	0.05	0.0	0.23	0.28	0.41	2.71
400	200	0.05	0.1	0.27	0.28	0.41	9.69
400	200	0.05	0.3	0.22	0.30	0.39	3.81
400	1000	0.05	0.0	12.06	7.99	21.35	71.89
400	1000	0.05	0.1	10.99	9.73	21.99	29.76
400	1000	0.05	0.3	11.36	9.71	22.78	138.38
400	1000	0.20	0.0	15.59	12.33	20.98	141.82
400	1000	0.20	0.1	13.97	10.94	22.39	29.67
400	1000	0.20	0.3	13.47	11.24	22.56	131.26

Table 5 Degenerate Random Problems: Time (secs). $p = n/5$, `noise`$= 10^{-3}$, `fStrong`$= 0.20$, `DeltaFac`$= 4.00$.

than those made by LP-P and LP-D. The misidentifications for LP-D and LP-P tend to be false negatives, and their numbers increase with the number of weakly active constraints. This experience is in accordance with the theory of Subsection 3.1, which gives no guarantee that the weakly active constraints will be identified. The numbers of false negatives are larger for test $\mathcal{A}_\lambda$ than for $\mathcal{A}_c$—nearly as large as the number of degenerate constraints. (For $m = 400$, `fWeak` $= 0.05$ there are 20 such constraints while for $m = 400$ and `fWeak` $= 0.20$ there are 80.) This observation indicates that the multiplier (μ, λ) determined by the LP-P and LP-D solution is similar to the optimal multiplier (μ^*, λ^*) to the original problem, for which $\lambda_i^* = 0$ when constraint i is weakly active. The errors for LPEC-A and LPEC contain both false positives and false negatives, indicating that the values of σ and $\bar{\sigma}$ and the factor β that we use in the activity test are appropriate. (For larger values of σ and $\bar{\sigma}$, the numbers of false negatives increase dramatically.)

The methods can usually be ranked in order of speed as LP-P, LP-D, LPEC-A, and LPEC. The differences between LP-P and LP-D are likely due to problem size reductions by the presolver, which are greater for LP-P, and which are significant because the matrix is dense. As expected (see our discussion in Subsection 4.1.1), we observed size reductions corresponding to the number of inactive constraints. In contrast, no presolver reductions were observed for LPEC-A.

For the mixed-integer program arising in the LPEC test, an additional node beyond the root node of the branch-and-bound tree is considered only for the case $m = 400$, $n = 200$, `fWeak` $= 0.05$, and `degenA` $= 0.1$. We observed large initial scaled dual infeasibilities and runtimes that are sensitive to the

LPEC parameter M. For the case $m = 400$, the relative slowness of the LPEC method may be due to the relatively large size of the matrix generated by the LPEC (see Table 1).

4.3 CUTEr Problems

We now consider a subset of constrained minimization problems from the CUTEr test set [13]. The subset contains degenerate problems of small or medium size for which the `Interior/Direct` algorithm of `Knitro 4.x` terminates successfully within 3000 iterations (with default parameter values). From the output of this code, we obtain approximations x^* to a solution and (μ^*, λ^*) to the optimal Lagrange multipliers.

The format of the CUTEr test problems differs from that of (1) in that bound constraints are treated separately from general constraints and all constraints are two-sided, that is, they have both lower and upper bounds. We implemented alternative formulations for our four tests which treated the bounds explicitly and combined them with the trust-region constraints, thereby reducing the total number of constraints and/or variables. We found that these formulations gave little or no improvement in performance, so we do not report on them further. For the results below, we rewrite the CUTEr test problems in the format (1), treating bound constraints in the same way as general inequality constraints.

4.3.1 Determining the "True" Active Set

In contrast to the random problems of Section 4.2, the true active set $\mathcal{A}^*$ is not known, but must be estimated from the solution determined by `Interior/Direct`. Inevitably, this solution is approximate; the code terminates when constraint-multiplier product for each inequality constraint falls below a given tolerance, set by default to 10^{-6} (see Byrd et al. [7]). If one of λ_i or $-c_i$ is much smaller than the other, classification of the constraint is easy, but in many cases these two quantities are of comparable magnitude. For example, the `Interior/Direct` solutions of problems such as CAR2, BRAINPC*, and READING1 (when formulated as (1)) display patterns in which $-c_i$ increases steadily with i while λ_i decreases steadily, or vice versa. It is difficult to tell at which index i the line should be drawn between activity and inactivity.

In our tables below, we define $\mathcal{A}^*$ by applying the LPEC-A test (43) with $\bar{\sigma} = .75$ and $\beta = 1/(m + n + p)$ to the solution returned by `Knitro` . (LPEC could be used in place of LPEC-A to estimate $\mathcal{A}^*$ because both schemes are theoretically guaranteed to return the true active set for x close enough to x^*.) We also wish to determine the weakly active inequality set $\mathcal{A}_0^*$, defined by (5). Procedure ID0 from [23, Section 3], which involves repeated solution of linear programs, could be used to determine this set. However, for purposes of Table 6, we populated $\mathcal{A}_0^*$ with those indices in the estimated $\mathcal{A}^*$ that fail the test (46) when applied to the multipliers returned by LPEC-A at x^*. Note that this technique produces a superset of $\mathcal{A}_0^*$ in general.

4.3.2 Implementation Details

The penalty parameter in the LP-P and LP-D formulations is defined by

$$\nu = 1.5 \max(\max_j(\lambda_j^*), \max_k(|\mu_k^*|), 1),$$

where (μ^*, λ^*) are the approximately optimal multipliers that were reported by
`Interior/Direct`. This heuristic guarantees that these particular multipliers (μ^*, λ^*) are feasible for the LP-D formulation (18) at the `Interior/Direct` approximate solution x^*. For the parameter M in (31) we use

$$M = 3 \max(\max_j(\lambda_j^*), \max_j(|c_j(x)|)).$$

Function and gradient evaluations are obtained through the Fortran and C tools contained in the CUTEr distribution, and through a driver modeled after the routine `loqoma.c` (an interface for the code LOQO), which is also contained in CUTEr.

4.3.3 Test Results and Runtimes

Results for `noise` $= 10^{-3}$ are shown in Table 6. The numbers of elements in our estimate of the optimal active set $\mathcal{A}^*$ and weakly active set $\mathcal{A}_0^*$ are listed as $|\mathcal{A}^*|$ and $|\mathcal{A}_0^*|$. Each entry in the main part of the table contains the false positive/false negative count for each combination of test problem and identification technique. Table 7 shows the dimensions of each problem in the format $m/n/p$, with the main part of the table displaying runtimes in seconds. The LPEC column additionally reports the number of nodes beyond the root needed to solve the LPEC to the required (loose) tolerance. For many of the problems, the root node is within a factor of two of the optimal solution, and the reported number is therefore zero. If the LPEC test exceeds our time limit of 180 seconds, we qualify the approximate solution with the symbol "†".

Trends In Table 6, the LP-D and LP-P results are nearly identical, indicating that the two methods usually find the same solution. The errors for both these tests are mostly false negatives, which is expected, because the theory of Section 3.1 gives no guarantee that weakly active constraints will be identified. Further, false positives are unlikely because the nice value of Δ (set up by the choice of parameter `DeltaFac` $= 4$) excludes most inactive constraints from the trust region. The number of false negatives for the $\mathcal{A}_\lambda$ test is usually higher than for $\mathcal{A}_c$, because weakly active constraints will generally fail the $\mathcal{A}_\lambda$ test (46), while they may pass the $\mathcal{A}_c$ test (45). This behavior is highlighted in the results for the problems GMNCASE4 and OET7. Their numbers of false negatives for the $\mathcal{A}_\lambda$ test correspond exactly to $|\mathcal{A}_0^*|$, while the corresponding numbers of false negatives for the $\mathcal{A}_c$ test are much lower.

In contrast to the results for LP-D and LP-P, the results for LPEC-A and LPEC show a mixture of false positives and false negatives. Further, the

results for LPEC-A and LPEC are similar for most problems. For several problems, for example, TWIRIMD1 and ZAMB2, the results for LPEC-A agree with those for LPEC but not with those for LP-D and LP-P.

The runtimes given in Table 7 are typically much shorter than for the random problems in Tables 3 and 5 because the constraint Jacobians in the CUTEr problems are usually sparse (OET7 is an exception). The LP-D times are similar to the LP-P times, and LPEC-A times are generally comparable. With few exceptions, LPEC requires more execution time than LPEC-A. In cases for which LPEC requires significantly more time than LPEC-A, the LPEC identification performance is not better in general.

The LPEC method is usually the slowest, despite initialization from a good starting point. (The use of this starting point reduced significantly the solve time and the number of searched nodes for several problems, including HANGING, NGONE, SMMPSF, TRIMLOSS, and TWIRISM1.)

Anomalies For several problems, the numbers of false negatives for LP-D and LP-P with test $\mathcal{A}_\lambda$ are larger than $|\mathcal{A}_0^*|$; see for example SREADIN3. This may happen because the LP-P and LP-D programs find a sparse λ, one that has many more zeros than the λ^* that we used to form our estimate of $\mathcal{A}_0^*$ as described above.

For certain problems, *all* methods return large numbers of false negatives. These problems often contain many bound constraints; for example C-RELOAD, READING1, SREADIN3, TRIMLOSS, TWIRIMD1, and ZAMB2. We note that these errors still occurred when we reformulated the tests to treat the bound constraints explicitly.

For LPEC-A and LPEC, the BRAINPC* and OET7 problems have many false positives, as a result of many inactive constraints having values of $c_i(x)$ close to zero, below the threshold for determining activity.

For the problem SOSQP1, a quadratic program, only the LPEC method detects any active constraints; in fact, it makes no identification errors. A smaller choice for the parameter $\bar{\sigma}$ in the LPEC-A identification test would produce perfect identification for the LPEC-A technique also.

We remark on a few more of the anomalies in Table 7. Runtimes for HANGING are especially large, given its size. The LP solvers performed many perturbations and the MIP solver for the LPEC test reports trouble identifying an integer solution. For LPEC, an extremely large number of iterations and nodes are reported for C-RELOAD, again due to difficulty finding feasible integer solutions. Allowing the use of heuristics by the MIP solver yielded a large reduction in the number of considered nodes for this problem, but the runtime did not change significantly.

4.4 Additional Remarks

We conclude this section with some general comments on the numerical results and on additional testing not reported above.

In general, the LP-P and LP-D tests give similar identification results, with a tendency to underestimate the active set (that is, false negatives).

Problem	$\lvert\mathcal{A}^*\rvert/\lvert\mathcal{A}_0{}^*\rvert$	LP-D		LP-P		LPEC-A	LPEC
		$\mathcal{A}_c$	$\mathcal{A}_\lambda$	$\mathcal{A}_c$	$\mathcal{A}_\lambda$		
A4X12	191/88	0/21	0/122	0/21	0/122	6/0	0/32
AVION2	21/5	0/6	0/9	0/6	0/10	7/0	4/0
BIGBANK	0/0	0/0	0/0	0/0	0/0	0/0	0/0
BRAINPC0	3/3	0/0	0/1	0/0	0/3	65/0	67/0
BRAINPC1	3/3	4/0	0/2	4/0	0/3	33/0	38/0
BRAINPC3	3/3	0/0	0/1	0/0	0/3	67/0	69/0
BRAINPC4	9/9	6/0	0/9	6/0	0/9	22/0	56/0
CAR2	883/1	0/312	0/883	0/321	0/883	0/112	53/0†
CORE1	21/3	0/0	0/7	0/0	0/7	0/0	0/0
CORKSCRW	505/6	0/3	0/190	0/3	0/189	0/3	0/6
C-RELOAD	136/7	0/38	0/124	0/38	0/124	0/19	0/18†
DALLASM	3/1	0/0	0/1	0/0	0/1	0/0	0/0
DALLASS	1/0	0/0	0/1	0/0	0/1	0/0	0/0
DEMBO7	21/8	0/1	0/7	0/1	0/11	0/0	0/1
FEEDLOC	20/19	0/0	0/19	0/0	0/19	0/7	0/0
GMNCASE4	350/175	0/0	0/175	0/0	0/175	0/0	0/0
GROUPING	100/100	0/0	0/44	0/0	0/44	0/8	0/0
HANGING	2310/40	0/48	0/72	0/48	0/72	0/12	0/68
HELSBY	8/2	0/0	0/5	0/0	0/1	0/0	0/0
HIMMELBK	20/10	0/0	0/10	0/0	0/9	0/0	1/0
HUES-MOD	277/0	0/1	0/78	0/1	0/78	0/1	0/277
KISSING2	181/87	0/0	0/88	0/0	0/88	0/0	0/2
LISWET10	1999/0	0/2	0/237	0/2	0/254	1/0	0/6
LSNNODOC	3/1	0/0	0/1	0/0	0/1	0/0	0/0
MAKELA3	20/19	0/0	0/19	0/0	0/19	0/0	0/20
MINPERM	0/0	0/0	0/0	0/0	0/0	0/0	0/0
NET1	7/2	0/0	0/2	0/0	0/2	0/0	0/0
NGONE	102/0	0/0	0/86	0/0	0/86	0/0	0/5
OET7	110/105	0/15	0/105	0/15	0/105	38/21	86/20
POLYGON	105/4	0/0	0/4	0/0	0/4	0/0	0/17
PRIMALC8	505/2	0/4	0/505	0/4	0/505	0/0	0/0
PRODPL0	39/0	0/0	0/0	0/0	0/0	0/0	0/0
QPCBLEND	80/42	0/24	0/45	0/24	0/45	0/12	0/24
QPCBOEI1	309/49	0/0	0/47	0/0	0/49	4/0	0/18
QPCSTAIR	163/20	0/50	0/59	0/50	0/56	20/0	0/51
READING1	174/147	0/173	0/174	0/173	0/174	0/141	0/86†
SARO	675/2	0/43	0/675	0/43	0/675	0/44	0/58†
SAROMM	343/0	0/0	0/343	0/0	0/343	0/0	0/10†
SMBANK	0/0	0/0	0/0	0/0	0/0	0/0	0/0
SMMPSF	481/1	0/5	0/66	0/5	0/66	0/1	0/10
SOSQP1	2500/2500	0/2500	0/2500	0/2500	0/2500	0/2500	0/0
SREADIN3	180/154	0/180	0/180	0/180	0/180	0/146	0/104†
SSEBNLN	133/25	0/2	0/35	0/2	0/25	0/0	0/2
STEENBRA	381/95	0/0	0/55	0/0	0/51	0/0	0/0
TRIMLOSS	94/51	1/69	0/93	0/67	0/93	0/7	0/33
TRUSPYR2	8/1	0/0	0/1	0/0	0/0	0/0	4/0
TWIRIMD1	660/80	0/257	0/659	0/258	0/659	0/56	0/56†
TWIRISM1	140/29	0/15	0/83	0/15	0/84	0/15	0/18
ZAMB2	1259/0	0/673	0/1259	0/673	0/1259	0/102	0/102†

Table 6 CUTEr problems: False Positives/False Negatives. `noise`$= 10^{-3}$, $\sigma = 0.75$, $\bar{\sigma} = 0.90$, `DeltaFac`$= 4.00$

.

Problem	m/n/p	$\lvert\mathcal{A}^*\rvert/\lvert\mathcal{A}_0{}^*\rvert$	LP-D	LP-P	LPEC-A	LPEC/nodes
A4X12	384/ 130/ 16	191/ 88	0.02	0.03	0.01	5.62/66
AVION2	98/ 49/ 15	21/ 5	0.00	0.00	0.00	0.03/5
BIGBANK	3844/2230/1420	0/ 0	0.42	0.17	0.12	1.16/0
BRAINPC0	6905/6907/6902	3/ 3	3.98	6.82	22.52	31.47/0
BRAINPC1	6905/6907/6902	3/ 3	4.84	18.81	0.44	1.44/0
BRAINPC3	6905/6907/6902	3/ 3	2.87	6.64	25.42	58.82/0
BRAINPC4	6905/6907/6902	9/ 9	4.06	5.98	8.75	2.95/0
CAR2	4997/5999/4004	883/ 1	5.13	0.54	0.65	$189^{\dagger}$/7065
CORE1	139/ 65/ 41	21/ 3	0.00	0.00	0.00	0.01/0
CORKSCRW	4500/4506/3009	505/ 6	0.92	0.80	0.19	93.25/1
C-RELOAD	684/ 342/ 200	136/ 7	0.10	0.07	0.04	$181^{\dagger}$/40420
DALLASM	392/ 196/ 151	3/ 1	0.01	0.01	0.00	0.13/0
DALLASS	92/ 46/ 31	1/ 0	0.00	0.00	0.00	0.03/0
DEMBO7	53/ 16/ 0	21/ 8	0.00	0.00	0.00	0.03/1
FEEDLOC	462/ 90/ 22	20/ 19	0.00	0.00	0.00	0.18/0
GMNCASE4	350/ 175/ 0	350/ 175	0.05	0.08	0.04	0.12/0
GROUPING	200/ 100/ 125	100/ 100	0.00	0.00	0.00	0.01/0
HANGING	2330/3600/ 12	2310/ 40	27.40	44.70	6.46	10.04/0
HELSBY	685/1408/1399	8/ 2	0.22	0.20	0.03	0.50/0
HIMMELBK	24/ 24/ 14	20/ 10	0.00	0.00	0.00	0.00/0
HUES-MOD	5000/5000/ 2	277/ 0	1.34	0.15	0.24	1.08/0
KISSING2	625/ 100/ 6	181/ 87	0.01	0.02	0.01	14.89/492
LISWET10	2000/2002/ 0	1999/ 0	0.31	0.12	0.31	20.04/0
LSNNODOC	6/ 5/ 4	3/ 1	0.00	0.00	0.00	0.01/0
MAKELA3	20/ 21/ 0	20/ 19	0.00	0.00	0.00	0.00/0
MINPERM	1213/1113/1033	0/ 0	0.20	0.06	0.11	15.55/0
NET1	65/ 48/ 43	7/ 2	0.00	0.00	0.00	0.01/0
NGONE	5246/ 200/ 3	102/ 0	0.02	0.03	0.02	21.18/1
OET7	1002/ 7/ 0	110/ 105	0.01	0.02	0.00	0.29/0
POLYGON	5445/ 200/ 2	105/ 4	0.02	0.03	0.02	28.98/1
PRIMALC8	511/ 520/ 0	505/ 2	0.04	0.03	0.01	0.11/0
PRODPL0	69/ 60/ 20	39/ 0	0.00	0.00	0.00	0.02/0
QPCBLEND	114/ 83/ 43	80/ 42	0.01	0.01	0.00	0.50/141
QPCBOEI1	971/ 384/ 9	309/ 49	0.03	0.02	0.02	17.07/26
QPCSTAIR	532/ 467/ 291	163/ 20	0.05	0.03	0.02	2.65/38
READING1	8002/4002/2001	174/ 147	0.61	0.36	0.22	$192^{\dagger}$/2600
SARO	2920/4754/4025	675/ 2	3.17	4.57	2.67	$182^{\dagger}$/347
SAROMM	2920/5120/4390	343/ 0	4.68	7.13	1.71	$182^{\dagger}$/19
SMBANK	234/ 117/ 64	0/ 0	0.01	0.00	0.00	0.02/0
SMMPSF	743/ 720/ 240	481/ 1	0.11	0.04	0.03	4.82/255
SOSQP1	10000/5000/2501	2500/2500	0.08	0.08	0.30	7.94/0
SREADIN3	8004/4002/2001	180/ 154	0.80	0.32	0.23	$187^{\dagger}$/2385
SSEBNLN	384/ 194/ 74	133/ 25	0.01	0.01	0.01	0.03/0
STEENBRA	432/ 432/ 108	381/ 95	0.02	0.01	0.01	0.44/1
TRIMLOSS	319/ 142/ 20	94/ 51	0.00	0.00	0.01	0.45/31
TRUSPYR2	16/ 11/ 3	8/ 1	0.00	0.00	0.00	0.01/0
TWIRIMD1	2685/1247/ 521	660/ 80	3.86	0.97	1.02	$182^{\dagger}$/110
TWIRISM1	775/ 343/ 224	140/ 29	0.09	0.06	0.04	18.87/501
ZAMB2	7920/3966/1446	1259/ 0	0.50	0.14	0.21	$190^{\dagger}$/2025

Table 7 CUTEr problems: Time (secs). `noise`= 10^{-3}, $\sigma = 0.75$, $\bar{\sigma} = 0.90$, `DeltaFac`= 4.00..

The primal activity test $\mathcal{A}_c$ is superior to the dual activity test $\mathcal{A}_\lambda$ for these methods. LP-P tended to take less time to solve, probably because of the greater reductions due to presolving.

We tested the effect of using a much larger Δ in the LP-P and LP-D formulations for the random test problems. Runtimes were slightly longer on the largest problems, and the time advantage that LP-P has for smaller Δ disappears. The $\mathcal{A}_\lambda$ activity test returned the same poor underestimate of the active set as for the smaller Δ, while the $\mathcal{A}_c$ activity test made many more identification errors.

The LPEC-A test obviously should be used in preference to LPEC, as the results are similar (with the anomalies easily explained) and the runtimes are sometimes much shorter. We note that it might be possible to improve the performance of these methods by better scaling of the constraints.

We used a `noise` value of 10^{-3} in all reported results, but performed additional experiments with other values of this parameter. For smaller values of `noise`, LP-P and LP-D tend to have similar false positive counts, but show higher false negative counts in some cases. LPEC and LPEC-A show an overall improvement; for example, at `noise` $= 10^{-7}$ the BRAINPC* problems' results for LPEC and LPEC-A are nearly perfect. However, more false negatives are reported on some CUTEr problems. These difficult problems are the ones for which our estimate of the true active set $\mathcal{A}^*$ is sensitive to the parameters β, σ, and $\bar{\sigma}$ used in the threshold test (see Subsection 4.3.1), and for which the estimate of the true active set changes significantly if we use LPEC in place of LPEC-A. Specifically, on problems LISWET10, OET7, and READING1, the additional false negatives that were reported when `noise` was decreased from 10^{-3} to 10^{-7} disappeared when $\bar{\sigma}$ was changed or when LPEC was used in place of LPEC-A in the determination of $\mathcal{A}^*$.

As expected, the results of the random problems in Tables 2 and 3 for the LPEC and LPEC-A techniques are nearly perfect for `noise` $= 10^{-7}$. (`noise` must be decreased to an even smaller value to remove a single false positive in some cases; this identification error is caused by a constraint that is only very slightly inactive.)

For values of `noise` larger than 10^{-3}, LP-P and LP-D report more false positives on the random problems and fewer false negatives on the CUTEr problems. The LPEC and LPEC-A tests tend to give more false positives, while the false negative count decreases on the CUTEr problems and increases on the random problems.

Following a suggestion of a referee, and in line with the discussion at the end of Section 3, we obtained a new identification technique by inserting the solution of (18) in place of (μ_x, λ_x) in the threshold test (43). We found that, indeed, this "threshold LP-D" estimate of the active set was more accurate than those obtained from (20a) and (20b), as is done in the standard LP-D technique. On the random problem set, the results for threshold LP-D for `noise` $= 10^{-7}$ are identical to those for LPEC-A, in accordance with our claim that both techniques are asymptotically exact.

5 Conclusions

We have described several schemes for predicting the active set for a nonlinear program with inequality constraints, given an estimate x of a solution x^*. The effectiveness of some of these schemes in identifying the correct active set for x sufficiently close to x^* is proved, under certain assumptions. In particular, the scheme of Subsection 3.3 has reasonable computational requirements and strong identification properties and appears to be novel. Computational tests are reported which show the properties of the various schemes on random problems and on degenerate problems from the CUTEr test set.

Knowledge of the correct active set considerably simplifies algorithms for inequality constrained nonlinear programming, as it removes the "combinatorial" aspect from the problem. However, it remains to determine how the schemes above can be used effectively as an element of a practical algorithm for solving nonlinear programs. It may be that reliable convergence can be obtained in general without complete knowledge of the active set; some "sufficient subset" may suffice. What are the required properties of such a subset, and can we devise inexpensive identification schemes, based on the ones described in this paper, that identify it? We leave these and other issues to future research.

Acknowledgments

We thank Richard Waltz for many discussions during the early part of this project, for supplying us with the `Knitro` results, and for advice about the implementations. We also thank Dominique Orban for providing helpful advice about using CUTEr. Finally, we are most grateful to two anonymous referees for extremely thorough and helpful comments on the first version of this paper.

A Proof of (25)

We prove this statement by contradiction. Suppose that there is a sequence $\{x^k\}$ with $x^k \to x^*$ such that

$$\text{dist}\left(-g(x^k), \text{range}\,[\nabla h(x^k)] + \text{pos}[(\nabla c_i(x^k))_{i\in\mathcal{A}_1}]\right) < \tau, \tag{47}$$

for all k. By closedness, there must be vectors z^k and $y^k \geq 0$ such that the

$$\begin{aligned}&\text{dist}\left(-g(x^k), \text{range}\,[\nabla h(x^k)] + \text{pos}[(\nabla c_i(x^k))_{i\in\mathcal{A}_1}]\right)\\ &= \left\|\nabla h(x^k)z^k + \sum_{i\in\mathcal{A}_1} \nabla c_i(x^k)y_i^k + g(x^k)\right\| \leq \tau,\end{aligned}$$

for all k. If $\{(z^k, y^k)\}$ is unbounded, we have by compactness of the unit ball, and by taking a subsequence if necessary, that $\|(z^k, y^k)\| \uparrow \infty$ and $(z^k, y^k)/\|(z^k, y^k)\| \to$

(z^*, y^*) with $\|(z^*, y^*)\| = 1$ and $y^* \geq 0$. Hence, by dividing both sides in the expression above by $\|(z^k, y^k)\|$ and taking limits, we have

$$(\nabla h^*) z^* + \sum_{i \in \mathcal{A}_1} (\nabla c_i^*) y_i^* = 0. \tag{48}$$

From Lemma 1, we have $\mathcal{A}_1 \subset \mathcal{A}^*$, so that the MFCQ condition (7) holds at x^* for $\mathcal{A}_1$ replacing $\mathcal{A}^*$. Hence, for the vector v in this condition, we have that $\nabla h(x^*)$ has full column rank, and that $(\nabla h^*)^T v = 0$ and $\nabla (c_i^*)^T v < 0$ for all $i \in \mathcal{A}_1$. By taking inner products of (48) with v, we can deduce first that $y^* = 0$ and subsequently that $z^* = 0$, by a standard argument, contradicting $\|(z^*, y^*)\| = 1$. Therefore, the sequence $\{(z^k, y^k)\}$ must be bounded. Since the sequence remains in a ball about the origin (that is, a compact set), it has an accumulation point.

By taking a subsequence again if necessary, suppose that $(z^k, y^k) \to (\hat{z}, \hat{y})$. We then have that

$$\begin{aligned}
&\left\| \nabla h(x^k) \hat{z} + \sum_{i \in \mathcal{A}_1} \nabla c_i(x^k) \hat{y}_i + g(x^k) \right\| \\
&\leq \left\| \nabla h(x^k) z^k + \sum_{i \in \mathcal{A}_1} \nabla c_i(x^k) y_i^k + g(x^k) \right\| + \|\nabla h(x^k)\| \|z^k - \hat{z}\| + \sum_{i \in \mathcal{A}_1} \|\nabla c_i(x^k)\| \, |y_i^k - \hat{y}_i| \\
&\leq \tau + o(1),
\end{aligned}$$

for all k sufficiently large. By taking limits in this expression, we deduce that

$$\text{dist } (-g^*, \text{range}\,[\nabla h^*] + \text{pos}[(\nabla c_i^*)_{i \in \mathcal{A}_1}]) \leq \tau,$$

which contradicts the definition of τ, for $\tau > 0$. Hence, a sequence $\{x^k\}$ with the property (47) cannot exist, so (25) holds for all $\bar{\epsilon}_2$ sufficiently small.

References

1. Bertsekas, D.P.: Projected Newton methods for optimization problems with simple constraints. SIAM Journal on Control and Optimization **20**, 221–246 (1982)
2. Burke, J.V.: On the identification of active constraints II: The nonconvex case. SIAM Journal on Numerical Analysis **27**(4), 1081–1102 (1990)
3. Burke, J.V., Moré, J.J.: On the identification of active constraints. SIAM Journal on Numerical Analysis **25**, 1197–1211 (1988)
4. Burke, J.V., Moré, J.J.: Exposing constraints. SIAM Journal on Optimization **4**(3), 573–595 (1994)
5. Byrd, R., Gould, N.I.M., Nocedal, J., Waltz, R.A.: An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. Mathematical Programming, Series B **100**, 27–48 (2004)
6. Byrd, R., Gould, N.I.M., Nocedal, J., Waltz, R.A.: On the convergence of successive linear-quadratic programming algorithms. SIAM Journal on Optimization **16**(2), 471–489 (2005)
7. Byrd, R.H., Hribar, M.E., Nocedal, J.: An interior point algorithm for large scale nonlinear programming. SIAM Journal on Optimization **9**(4), 877–900 (1999)
8. Conn, A.R., Gould, N.I.M., Toint, P.: Trust-Region Methods. MPS-SIAM Series on Optimization. SIAM (2000)
9. El-Bakry, A.S., Tapia, R.A., Zhang, Y.: A study of indicators for identifying zero variables in interior-point methods. SIAM Review **36**, 45–72 (1994)

10. Facchinei, F., Fischer, A., Kanzow, C.: On the accurate identification of active constraints. SIAM Journal on Optimization **9**(1), 14–32 (1998)
11. Fletcher, R.: Practical Methods of Optimization, second edn. John Wiley and Sons, New York (1987)
12. Fletcher, R., Sainz de la Maza, E.: Nonlinear programming and nonsmooth optimization by successive linear programming. Mathematical Programming **43**, 235–256 (1989)
13. Gould, N.I.M., Orban, D., Toint, P.L.: CUTEr (and SifDec), a constrained and unconstrained testing environment, revisited*. Technical Report TR/PA/01/04, CERFACS (2001)
14. Hager, W.W., Gowda, M.S.: Stability in the presence of degeneracy and error estimation. Mathematical Programming, Series A **85**, 181–192 (1999)
15. Hare, W., Lewis, A.: Identifying active constraints via partial smoothness and prox-regularity. Journal of Convex Analysis **11**(2), 251–266 (2004)
16. Lescrenier, M.: Convergence of trust region algorithms for optimization with bounds when strict complementarity does not hold. SIAM Journal on Numerical Analysis **28**(2), 476–495 (1991)
17. Lewis, A.: Active sets, nonsmoothness, and sensitivity. SIAM Journal on Optimization **13**, 702–725 (2003)
18. Monteiro, R.D.C., Wright, S.J.: Local convergence of interior-point algorithms for degenerate monotone LCP. Computational Optimization and Applications **3**, 131–155 (1994)
19. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton, N.J. (1970)
20. Waltz, R.: An active-set trust-region algorithm for nonlinear optimization. Presentation at ISMP Copenhagen (2003)
21. Wright, S.J.: Identifiable surfaces in constrained optimization. SIAM J. Control Optim. **31**, 1063–1079 (1993)
22. Wright, S.J.: Modifying SQP for degenerate problems. SIAM Journal on Optimization **13**, 470–497 (2002)
23. Wright, S.J.: Constraint identification and algorithm stabilization for degenerate nonlinear programs. Mathematical Programming, Series B **95**, 137–160 (2003)
24. Yamashita, N., Dan, H., Fukushima, M.: On the identification of degenerate indices in the nonlinear complementarity problem with the proximal point algorithm. Mathematical Programming, Series A **99**, 377–397 (2004)
25. Ye, Y.: On the finite convergence of interior-point algorithms for linear programming. Mathematical Programming **57**, 325–336 (1992)