MATHEMATICS OF OPERATIONS RESEARCH
Vol. 36, No. 1, February 2011, pp. 55–70
ISSN 0364-765X | EISSN 1526-5471 | 11 | 3601 | 0055

informs®

DOI 10.1287/moor.1110.0481

Generic Optimality Conditions for Semialgebraic Convex Programs

Jérôme Bolte
TSE (Gremaq, Université Toulouse 1), 31000 Toulouse, France,
jerome.bolte@tse-fr.eu, www.ecp6.jussieu.fr/pageperso/bolte

Aris Daniilidis
Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, arisd@mat.uab.cat, mat.uab.es/~arisd

Adrian S. Lewis
ORIE, Cornell University, Ithaca, New York 14853, aslewis@orie.cornell.edu, people.orie.cornell.edu/~aslewis

We consider linear optimization over a nonempty convex semialgebraic feasible region F. Semidefinite programming is an example. If F is compact, then for almost every linear objective there is a unique optimal solution, lying on a unique "active" manifold, around which F is "partly smooth," and the second-order sufficient conditions hold. Perturbing the objective results in smooth variation of the optimal solution. The active manifold consists, locally, of these perturbed optimal solutions; it is independent of the representation of F and is eventually identified by a variety of iterative algorithms such as proximal and projected gradient schemes. These results extend to unbounded sets F.

Key words: convex optimization; sensitivity analysis; partial smoothness; identifiable surface; active sets; generic; second-order sufficient conditions; semialgebraic
MSC2000 subject classification: Primary: 90C31; secondary: 14P10, 32B20, 49K40, 90C05
OR/MS subject classification: Primary: programming; secondary: mathematics, convexity
History: Received January 19, 2009; revised March 30, 2009, October 9, 2009, September 13, 2010, and December 12, 2010.

1. Introduction. Optimizers approach problems from diverse perspectives, designing algorithms and analyzing their convergence behavior, studying sensitivity analysis with respect to data perturbation and, relatedly, investigating duality theory and shadow prices. In most cases, whether in theory or computational practice, we make assumptions about prevailing optimality conditions, a canonical example being the second-order sufficient conditions for classical nonlinear programming (Nocedal and Wright [28]).

That such optimality conditions typically hold can be justified rigorously via Sard's theorem. To take the most basic example, when considering smooth equality constraints $F(x) = b$, we usually assume that the Jacobian of the map F has full rank at some feasible solution of interest. Provided that F is sufficiently smooth, Sard's theorem guarantees that for a *generic* vector b, this assumption holds at *any* feasible solution. "Generic" in this context means that the assertion holds for almost all b, in the sense of Lebesgue measure, and hence, for example, almost surely for a random b having an everywhere strictly positive probability density function.

A classic paper of Spingarn and Rockafellar [40] explains how such arguments show the generic nature of the second-order sufficient conditions; see also Spingarn [39]. The idea of studying optimization from a generic perspective dates back further, at least to the 1973 study by Saigal and Simon [36] of the complementarity problem, and it has persisted; see, for example, the studies of generic strict complementarity and primal and dual nondegeneracy for semidefinite programming by Alizadeh et al. [1] and Shapiro [37] the study of general conic convex programs by Pataki and Tunçel [29].

An important consequence of the classical second-order conditions is the existence of an "active manifold," consisting of those feasible points satisfying all of the active constraints with equality. Perturbations to the objective function result in smooth perturbation of the optimal solution on this manifold. Classical active set algorithms attempt to find this manifold, thereby reducing the optimization problem to a much easier equality-constrained problem.

For convex programs in particular, a variety of algorithms, such as proximal and projected-gradient schemes, "identify" the active manifold automatically: after finite time, iterates generated by the algorithm must lie on the manifold. For example, Rockafellar [33] observed that the classical proximal point method converges finitely on a polyhedral function; the same holds for functions with the "weak sharp minimum" property introduced by Ferris [12]. Burke and Moré [5] present an early survey on identification in the polyhedral case.

Modern optimization considers a variety of models beyond the framework of classical nonlinear programming. Semidefinite programming is one example. This generality introduces more complexity into the active set methodology. For example, sensitivity analysis when a constraint involves a positive semidefinite matrix variable must consider not simply whether or not the variable is zero, but also its rank; see, for example, Bonnans and Shapiro [4]. Stability constraints on nonsymmetric matrices involve analogous subtleties (Burke et al. [6]).

In these more general frameworks, active set ideas can perhaps be better understood directly in terms of the geometry of the feasible region, rather than its defining constraints, adopting an *intrinsic* approach, rather than a *structural* one. This idea motivates the definition of a weak sharp minimum (Ferris [12]), and the generalization to "identifiable surfaces" in convex optimization attributable to Wright [43]. The equivalent idea of an active manifold with respect to which the feasible region is "partly smooth" was introduced by Lewis in Lewis [21]; like the weak sharp minimum idea, this approach is intrinsically geometric (and, furthermore, extends to the nonconvex case). A closely related idea, "$\mathcal{UV}$-decompositions" of convex functions, was developed by Lemaréchal and Sagastizábal [19], as well as by Lemaréchal et al. [20].

Just as in classical cases, partial smoothness guarantees that a wide variety of iterative algorithms necessarily arrive on the active manifold after finite time (Wright [43], Hare and Lewis [13]), raising the hope of accelerating the algorithms using second-order information. This hope motivated the $\mathcal{UV}$-decomposition idea and was pursued further by Mifflin and Sagastizábal [25, 26] and recently by Daniilidis et al. [9]. A general approach, based on a proximal algorithm for composite optimization, is sketched by Lewis and Wright [22], and broad techniques for estimating identifiable surfaces computationally are discussed in Lewis and Wright [23].

Partial smoothness is a relatively strong assumption: it does not, for example, subsume the subtleties of sensitivity analysis investigated extensively by Bonnans and Shapiro [4], and it evades the technical complexities of second-order nonsmooth analysis, as described in Rockafellar and Wets' monograph [34]. Nonetheless, Lewis [21] argues that partial smoothness is a broad, intuitive, and powerful conceptual tool, and that a corresponding analogue of the second-order sufficient conditions suffices for a thorough and substantial generalization of classical nonlinear programming sensitivity analysis. In further analogy to the classical case, we would naturally expect algorithmic convergence analysis to rely on the same assumption.

In this work, we follow the philosophy of Spingarn and Rockafellar [40] in asking whether an assumption of partial smoothness and second-order sufficiency is usually justified. We prove, for a large class of linear optimization problems over convex feasible regions, that this assumption indeed holds generically.

As remarked in Spingarn and Rockafellar [40], the key ingredients to any such argument about generic behavior is the class of problems under consideration and the precise notion of "generic" that we use. The results of Spingarn and Rockafellar [40] fixed an objective and constraint functions, allowed linear perturbations to the objective and constant perturbations to the constraints, and proved a measure-theoretic result about the second-order conditions via Sard's theorem. Both Alizadeh et al. [1] and Shapiro [37] use rather analogous arguments to prove that strict complementarity and primal and dual nondegeneracy are generic properties of semidefinite programs; using a very different technique based on the boundary behavior of convex sets, Pataki and Tunçel [29] generalized these results to general conic convex programs. Ioffe and Lucchetti [16] adopt a more abstract, topological approach, allowing very general perturbations to the optimization problem but proving a result instead about "well-posedness" (Dontchev and Zolezzi [10]). Our approach, while notable for the generality of the feasible regions considered, is more concrete, combining some of the spirit of Spingarn and Rockafellar [40] and Pataki and Tunçel [29]: we aim to understand the generic nature of second-order sufficient conditions, like Spingarn and Rockafellar [40], but we make no assumption about how the feasible region is presented; instead, we assume only that it is *semialgebraic*. In a recent survey (Ioffe [15]), Ioffe describes how a semialgebraic version of Sard's theorem applies to an analogue of Spingarn and Rockafellar's result on generic optimality conditions.[1]

A set defined by finitely many polynomial inequalities is called *basic semialgebraic*; any set that can be represented as a finite union of such sets is simply called *semialgebraic*. Semialgebraic sets comprise a rich class that is stable under many mathematical operations. They are often easy to recognize, even without an explicit representation as a union of basic sets, as a consequence of the Tarski-Seidenberg theorem, which states, loosely, that the projection of a semialgebraic set is semialgebraic. For example, the feasible region of any semidefinite program is semialgebraic. A good resource on semialgebraic geometry is Basu et al. [2].

As our main result, we prove that, given any fixed nonempty closed convex semialgebraic set, corresponding to a generic linear objective function is a unique optimal solution, lying on a unique active manifold, and for which the partly smooth second-order sufficient conditions hold. The active manifold is independent of any particular representation of the semialgebraic set as a union of basic sets, and the optimal solution varies smoothly on it (in fact, giving a complete local description of it) under local perturbations to the objective function. In particular, this result holds for any semidefinite program.

[1] Renegar, J. 1996. A semialgebraic version of generic strict complementarity and nondegeneracy result. Personal communication to G. Pataki and L. Tunçel.

The term "generic" for a large subset of Euclidean space has been used in a variety of mathematical senses. Spingarn and Rockafellar [40] mean a *full-measure* subset—its complement has Lebesgue measure zero—whereas *topologically generic* means that the subset contains a countable intersection of dense open subsets. These two notions are incompatible in general, but fortunately, as we explain in §3, the distinction collapses, and the idea dramatically simplifies for semialgebraic sets, because such sets "stratify" into finite unions of smooth manifolds. For semialgebraic sets, therefore, we have three equivalent properties: full-measure, topologically generic, and dense.

Previous work on generic optimality conditions has been mostly structural, focusing on some given functional presentation of the feasible region, rather than on its intrinsic geometry. Spingarn and Rockafellar's work [40] concerns classical smooth constraint systems, and Ioffe's semialgebraic version [15] is analogous. Alizadeh et al.[1] derive results for linear semidefinite programs, extended to general conic convex systems by Pataki and Tunçel [29]; Shapiro and Fan [38] and Shapiro [37] focus on nonlinear semidefinite programs. Such structural approaches reflect the presentation of optimization problems in practice and are very general: in particular, Ben-Tal and Nemirovski [3] is a powerful toolkit for *semidefinite representation* of convex sets (that is, as affine preimages of the semidefinite cone).

For comparison purposes, the approach to second-order conditions developed by Bonnans and Shapiro [4] is particularly general and instructive. They consider smooth preimages of general closed convex cones and, like Shapiro and Fan [38], Shapiro [37], Alizadeh et al. [1], present in Bonnans and Shapiro [4, §4.6.1] an appropriate generalization of the classical full rank condition. This *transversality condition*—called *constraint nondegeneracy* by Robinson [31] and also discussed in detail in the context of nonlinear semidefinite programming by Sun [41], is generic, as a consequence of Sard's theorem, provided that the problem parametrization is sufficiently rich (Bonnans and Shapiro [4, §5.3.1]). Another important ingredient of second-order analysis, the quadratic decay condition, is also generic in semidefinite programming, because it is equivalent (see Bonnans and Shapiro [4, Theorem 5.91]) to uniqueness of the optimal solution along with a suitable analogue of the classical "strict complementarity" condition, known to be generic according to Alizadeh et al. [1] and Pataki and Tunçel [29]. The active manifold also emerges naturally using this approach, via a standard application of the transversality condition, assuming a powerful property of the underlying cone called "cone reducibility" (Bonnans and Shapiro [4, Definition 3.135]). While cone-reducibility is in general nontrivial to verify, some careful calculations show that the semidefinite cone in particular is cone reducible (Bonnans and Shapiro [4], Example 3.140), and furthermore products of cone reducible cones are cone reducible.

Despite the great generality of these concrete structural approaches, an intrinsic theory, based on the feasible region itself rather than on functional descriptions thereof, has a certain appeal. We make no assumption whatsoever about the presentation of feasible region, assuming only that it is semialgebraic. In other respects, our assumptions are quite restrictive: we deal only with the convex case, and we only consider perturbations to the objective, taking what is possibly just a first step towards a more general theory. Even the theoretical gain in generality in considering semialgebraic sets is unclear because, despite considerable interest and effort, an example of a semialgebraic convex set that is not semidefinite representable remains undiscovered (Helton and Nie [14]). Nonetheless, this semialgebraic approach is interesting: the main result is independent of the choice of presentation of the feasible region (a choice that may influence the corresponding genericity result in a complex fashion), the proof technique is novel in this context, the generic conclusion is stronger and more concrete (holding on a set that is dense and open rather than just full measure), and the sole assumption of semialgebraicity is typically immediate to verify, according to the Tarski-Seidenberg theorem.

The stratification property on which our theoretical development fundamentally depends is not confined to the class of semialgebraic sets. It holds more generally for "subanalytic" sets and indeed for any "tame" class of sets: see, for example, Coste [8] for a short introduction to tame geometry. Consequently, while all results are stated for semialgebraic sets, the authors believe analogous results hold for tame sets, and furthermore, that tame-geometric techniques show great promise in optimization theory more generally (see, for example, Ioffe's survey [15]). However, to lighten the exposition, we do not pursue this extension here.

Our approach in this paper relies heavily on convexity. However, many of the basic ideas driving this development extend to nonconvex settings, a central example being the normal cone and its continuity properties. Nonconvex variational analysis has grown into a complete, powerful, and elegant theory over the past several decades: fine expositions may be found in Mordukhovich [27], Clarke et al. [7], and Rockafellar and Wets [34]. Extending the results described in this paper to nonconvex settings is the topic of ongoing research.

Our exposition blends three relatively unfamiliar techniques for an optimization audience: the notion of a generic problem instance, semialgebraic geometry, and partial smoothness. We strive, however, for a self-contained approach, introducing and discussing the key concepts as we need them, and assuming nothing beyond classical convex analysis.

2. Preliminary results. In this section, we begin with some routine convex analysis, following the notation of Rockafellar [32], unless otherwise stated. Throughout this work we deal with the Euclidean space $\mathbf{R}^n$ equipped with the usual scalar product $\langle \cdot, \cdot \rangle$ and the corresponding Euclidean norm $|\cdot|$. We denote by $B(x, r)$ the closed ball with center $x \in \mathbf{R}^n$ and radius $r > 0$. A subset of $\mathbf{R}^n$ is a *cone* if it contains zero and is closed under nonnegative scalar multiplication. We denote by S^{n-1} the unit sphere of $\mathbf{R}^n$. Given any set $E \subset \mathbf{R}^n$, we denote by $\bar{E}$ its closure.

Notation. Throughout this paper, unless otherwise stated, we consider a fixed nonempty compact convex set $F \subset \mathbf{R}^n$, and study the set of maximizers of the linear optimization problem

$$\max_F \langle c, \cdot \rangle$$

for vectors $c \in \mathbf{R}^n$. The linear case is in some sense not restrictive, because a nonlinear optimization problem $\max_F f$ could be rephrased as the linear problem $\max\{t\colon t \le f(x),\ x \in F,\ t \in \mathbf{R}\}$.

A point $\bar{x} \in F$ is a maximizer if and only if c lies in the normal cone

$$N_F(\bar{x}) = \{c \in \mathbf{R}^n\colon \langle c, x - \bar{x} \rangle \le 0 \text{ for all } x \in F\}.$$

We call a maximizer $\bar{x}$ *nondegenerate* if, in fact, c lies in the relative interior of the normal cone, $c \in \operatorname{ri} N_C(\bar{x})$.

For an arbitrary convex set $F \subset \mathbf{R}^n$ (possibly unbounded), we say that the objective function $\langle c, \cdot \rangle$ *decays quadratically* on F around a maximizer $\bar{x}$ if there exists a constant $\delta' > 0$ such that

$$\langle c, \bar{x} \rangle \ge \langle c, x \rangle + \delta' |x - \bar{x}|^2 \quad \text{for all } x \in F \cap B(\bar{x}, \delta'); \tag{1}$$

see, for example, [4]. Although this is a local condition, it implies uniqueness of the maximizer $\langle c, \bar{x} \rangle > \langle c, x \rangle$ whenever $\bar{x} \ne x \in B(\bar{x}, \delta')$ and, hence, whenever $\bar{x} \ne x$, by convexity. On the other hand, quadratic decay can easily fail, even around a unique nondegenerate maximizer. For example, the point zero is the unique maximizer for the problem $\max\{-x_2\colon x_2 \ge |x_1|^{3/2}\}$, and it is nondegenerate, but quadratic decay fails.

If the set F is compact, quadratic decay is in fact a global condition: it simplifies to the existence of a constant $\delta > 0$ such that

$$\langle c, \bar{x} \rangle \ge \langle c, x \rangle + \delta |x - \bar{x}|^2 \quad \text{for all } x \in F. \tag{2}$$

Indeed, if property (1) holds, then uniqueness of the maximizer $\bar{x}$ implies that the continuous function

$$x \mapsto \frac{\langle c, \bar{x} - x \rangle}{|x - \bar{x}|^2}$$

is strictly positive on the compact set $\{x \in F\colon |x - \bar{x}| \ge \delta'\}$; if we denote the minimum value of this function by $\delta'' > 0$, then property (2) holds with $\delta = \min\{\delta', \delta''\}$.

The set of maximizers $\operatorname{arg\,max}_F \langle c, \cdot \rangle$ is called the *exposed face* of the set F corresponding to the vector c. In particular, the set F is itself an exposed face (corresponding to $c = 0$): all other exposed faces we call *proper*. The optimal value $\max_F \langle c, \cdot \rangle$, as a function of c, is called the *support function*, denoted σ_F: under our standing assumption that F is compact, the support function is a continuous and positively homogeneous convex function. Via standard convex analysis (Rockafellar [32]), we know that $\operatorname{arg\,max}_F \langle c, \cdot \rangle$ is the nonempty compact convex set $\partial \sigma_F(c)$, where ∂ denotes the convex subdifferential. We denote by x_c the optimal solution of minimum norm as

$$x_c = \operatorname{arg\,min}\Big\{|x|\colon x \in \operatorname{arg\,max}_F \langle c, \cdot \rangle\Big\}.$$

Notice the homogeneity property,

$$x_{\lambda c} = x_c \quad \text{for all } c \in \mathbf{R}^n \text{ and } \lambda > 0.$$

Clearly $\langle c, \cdot \rangle$ decays quadratically around a maximizer if and only if there exists a constant $\delta_c > 0$ such that

$$\langle c, x_c \rangle \ge \langle c, x \rangle + \delta_c |x - x_c|^2 \quad \text{for all } x \in F. \tag{3}$$

We aim to show good behavior of the optimization problem $\max_F \langle c, \cdot \rangle$ for objective vectors c lying in some large subset of $\mathbf{R}^n$, or equivalently, by scaling c, the sphere S^{n-1}.

We begin our development with a well-known result (cf. Ewald et al. [11]) proved by an easy and standard argument.

Proposition 2.1 (Generic Uniqueness). *Consider a nonempty compact convex set $F \subset \mathbf{R}^n$. For all nonzero vectors c lying in a topologically generic and full-measure cone in $\mathbf{R}^n$, the linear functional $\langle c, \cdot\rangle$ has a unique maximizer over F.*

Proof. The set of optimal solutions is a singleton (namely, $\{x_c\}$) if and only if the support function σ_F is differentiable at c. Being a finite convex function, the set of points of differentiability is both topologically generic and full-measure in $\mathbf{R}^n$ (Phelps [30]), and by positive homogeneity, it is also closed under strictly positive scalar multiplication. □

In fact a stronger result holds almost surely.

Proposition 2.2 (Generic Quadratic Decay). *Consider a nonempty compact convex set $F \subset \mathbf{R}^n$. Denote by K the set of vectors $c \in \mathbf{R}^n$ such that the linear functional $\langle c, \cdot\rangle$ decays quadratically around a maximizer over F. Then the cone $K \cup \{0\}$ is full-measure in $\mathbf{R}^n$.*

Proof. It is easy to check that the set K is closed under strictly positive scalar multiplication, so the set $K \cup \{0\}$ is certainly a cone. Alexandrov's theorem (Rockafellar and Wets [34, Theorem 13.51]) applied to the finite convex function σ_F shows that there exists a full-measure subset A of $\mathbf{R}^n$ at each point of which σ_F has a quadratic expansion. In particular, σ_F has gradient $\nabla\sigma_F(c) = x_c$ for all vectors $c \in A$, and in view of Rockafellar and Wets [34, Definition 13.1(c)], we have, for any fixed vector $\bar{c} \in A$, that there exists a positive semidefinite matrix S such that vectors $c \in \mathbf{R}^n$ near $\bar{c}$ satisfy

$$\sigma_F(c) = \sigma_F(\bar{c}) + \langle \nabla\sigma_F(\bar{c}), c - \bar{c}\rangle + \frac{1}{2}\langle S(c-\bar{c}), c-\bar{c}\rangle + o(|c-\bar{c}|^2).$$

Hence, there exist constants $\varepsilon > 0$ and $\rho > 0$ such that for all $c \in B(\bar{c}, \varepsilon)$ we have

$$\sigma_F(c) \le \sigma_F(\bar{c}) + \langle x_{\bar{c}}, c-\bar{c}\rangle + \frac{\rho}{2}|c-\bar{c}|^2.$$

Furthermore, we can clearly assume that

$$\varepsilon^{-1}\operatorname{diam}(F) < \rho. \tag{4}$$

Now consider any point $x \in F$. Because the Fenchel conjugate of the function σ_F is just the indicator function of F, we deduce successively

$$\begin{aligned} 0 = \sigma_F^*(x) &= \sup_{c\in\mathbf{R}^n}\{\langle x, c\rangle - \sigma_F(c)\} \\ &\ge \sup_{c\in B(\bar{c},\varepsilon)}\{\langle x, c\rangle - \sigma_F(c)\} \\ &\ge \sup_{c\in B(\bar{c},\varepsilon)}\left\{\langle x, c\rangle - \sigma_F(\bar{c}) - \langle x_{\bar{c}}, c-\bar{c}\rangle - \frac{\rho}{2}|c-\bar{c}|^2\right\} \\ &= \sup_{c\in B(\bar{c},\varepsilon)}\left\{\langle x - x_{\bar{c}}, c\rangle - \frac{\rho}{2}|c-\bar{c}|^2\right\} \\ &= \langle x - x_{\bar{c}}, \bar{c}\rangle + \sup_{u\in B(0,\varepsilon)}\left\{\langle x - x_{\bar{c}}, u\rangle - \frac{\rho}{2}|u|^2\right\}. \end{aligned}$$

In view of inequality (4), it is easy to see that the above supremum is attained at the point $u = \rho^{-1}(x - x_{\bar{c}}) \in B(0, \varepsilon)$. Replacing this value in the above inequality, we deduce that

$$0 \ge \langle x - x_{\bar{c}}, \bar{c}\rangle + \frac{1}{2\rho}|x - x_{\bar{c}}|^2, \quad \text{for all } x \in F,$$

which yields the asserted equation with $\delta_c = (2\rho)^{-1}$. Thus we have shown $A \subset K$, and the result follows. □

We next follow an argument analogous to that of Pataki and Tunçel [29] to show that nondegeneracy is also a generic property. To prove this result, it suffices to consider the special case when the interior of the set F contains the point zero. We then relate nondegeneracy to the facial structure of the *polar* set

$$F^\circ = \{c \in \mathbf{R}^n \colon \langle c, x\rangle \le 1 \text{ for all } x \in F\},$$

another compact convex set whose interior contains zero.

Proposition 2.3 (Normal Representation of Polar Exposed Faces). *Suppose zero lies in the interior of the compact convex set $F\subset \mathbf{R}^n$. Then the proper exposed faces of the polar set F° are those sets of the form*

$$G_{\bar x}=\{c\in N_F(\bar x)\colon \langle c,\bar x\rangle =1\}$$

for points $\bar x$ on the boundary of F. Furthermore, any such exposed face has relative interior given by

$$\operatorname{ri} G_{\bar x}=\{c\in \operatorname{ri} N_F(\bar x)\colon \langle c,\bar x\rangle =1\}.$$

Proof. For any point $\bar x\in F$, it is easy to see that

$$\{c\in N_F(\bar x)\colon \langle c,\bar x\rangle =1\}=\{c\in F^\circ\colon \langle c,\bar x\rangle =1\}=\arg\max_{F^\circ}\langle \cdot,\bar x\rangle.$$

Thus, any such set is certainly an exposed face, and if $\bar x$ is a boundary point of F (and hence nonzero), then this exposed face must be proper, because it does not contain zero.

Conversely, by definition, any exposed face of F° has the form

$$G=G_z=\{c\in F^\circ\colon \sigma_{F^\circ}(z)=\langle c,z\rangle\}$$

for some vector $z\in\mathbf{R}^n$, and assuming G is proper implies $z\neq 0$. By standard convex analysis, the support function σ_{F° is identical to the *gauge function* $\gamma_F\colon \mathbf{R}^n\to\mathbf{R}_+$ defined by

$$\gamma_F(z)=\inf\{\lambda\in\mathbf{R}_+\colon z\in\lambda F\}.$$

Because $z\neq 0$, we know that $\gamma_F(z)>0$, so we can define a point $\bar x=\gamma_F(z)^{-1}z$. By positive homogeneity, $\gamma_F(\bar x)=1$, so $\bar x$ lies on the boundary of F, and

$$G=\{c\in F^\circ\colon \gamma_F(z)=\langle c,z\rangle\}.$$

The first part of the result follows.

To show the last equation, it suffices to prove that the sets $\{c\in\mathbf{R}^n\colon \langle c,\bar x\rangle=1\}$ and $\operatorname{ri} N_F(\bar x)$ have nonempty intersection (see Rockafellar [32, Theorem 6.5]). If not, there exists a separating hyperplane, and hence a nonzero vector $y\in\mathbf{R}^n$ and a number $\alpha\in\mathbf{R}$ satisfying, for all vectors $c\in\mathbf{R}^n$,

$$c\in N_F(\bar x)\Rightarrow \langle c,y\rangle\le\alpha,$$
$$\langle c,\bar x\rangle=1\Rightarrow \langle c,y\rangle\ge\alpha.$$

The second implication easily shows $y=\lambda\bar x$ for some number $\lambda>0$. Because $0\in N_F(\bar x)$, the first implication shows $\alpha\ge 0$, and consequently, by positive homogeneity,

$$c\in N_F(\bar x)\ \Rightarrow\ \langle c,y\rangle\le 0,$$

and consequently

$$c\in N_F(\bar x)\ \Rightarrow\ \langle c,\bar x\rangle\le 0.$$

Because $\bar x$ lies on the boundary of the set F, there exists a nonzero vector $\bar c\in N_F(\bar x)$, and because zero lies in the interior of F, there exists a number $\delta>0$ such that $\delta\bar c\in F$. Hence

$$0\ge\langle \bar c,\delta\bar c-\bar x\rangle>-\langle\bar c,\bar x\rangle,$$

which is a contradiction. □

Corollary 2.1. *Suppose the compact convex set $F\subset\mathbf{R}^n$ contains zero in its interior. Then a vector c lies in the relative interior of a proper exposed face of the polar set F° if and only if the problem $\max_F\langle c,\cdot\rangle$ has a nondegenerate maximizer with optimal value $\sigma_F(c)=1$.*

Proof. By the preceding proposition, there exists a point $\bar x$ on the boundary of F such that $c\in\operatorname{ri} N_F(\bar x)$ and $\langle c,\bar x\rangle=1$. This point $\bar x$ is the desired nondegenerate maximizer, and clearly $\sigma_F(c)=\langle c,\bar x\rangle=1$.

Conversely, if $\bar x$ is a nondegenerate maximizer and $\sigma_F(c)=1$, then by definition $c\in\operatorname{ri} N_F(\bar x)$ and $\langle c,\bar x\rangle=1$. Clearly $c\neq 0$, so $\bar x$ must lie on the boundary of F, and the result now follows by the preceding proposition. □

Theorem 2.1 (Generic Nondegeneracy). *For any nonempty compact convex set $F \subset \mathbf{R}^n$, the set of vectors $c \in \mathbf{R}^n$ with the property that the problem $\max_F \langle c, \cdot\rangle$ has no nondegenerate maximizers has measure zero.*

Proof. To prove this result, we use the idea of Hausdorff measure, for which a good basic reference is Rogers [35]. We consider the subset H of the unit sphere consisting of vectors c such that the linear function $\langle c, \cdot\rangle$ has no nondegenerate maximizers over the set F. The unit sphere has dimension $n-1$; we show that H has $(n-1)$-dimensional Hausdorff measure zero. It follows easily that the cone $\mathbf{R}_+ H$ has measure zero, which is the result we desire.

We first restrict attention to the case when F contains zero in its interior. The general case follows easily, first by a translation to ensure zero lies in $\operatorname{ri} F$ and then by considering $\mathbf{R}^n$ as the direct sum of the span of F and its orthogonal complement.

Assume henceforth, therefore, that the compact convex set F contains zero in its interior. Its polar F° is then another compact convex set containing zero in its interior. Consider the following map from the boundary of F° to the unit sphere S^{n-1}. We define $\Phi\colon \operatorname{bd}(F^\circ) \to S^{n-1}$ by $\Phi(x) = |x|^{-1}x$. Because the function $x \mapsto |x|^{-1}$ is locally Lipschitz on $\mathbf{R}^n \setminus \{0\}$, it is globally Lipschitz on the compact subset $\operatorname{bd}(F^\circ)$, and hence so is Φ. Furthermore, Φ is invertible, with inverse $\Phi^{-1}\colon S^{n-1} \to \operatorname{bd}(F^\circ)$ given by $\Phi^{-1}(x) = (\gamma_{F^\circ}(x))^{-1}x$. The sublinear function $\gamma_{F^\circ} = \sigma_F$ is everywhere finite, and hence Lipschitz, so the function $x \mapsto (\gamma_{F^\circ}(x))^{-1}$ is locally Lipschitz on $\mathbf{R}^n\setminus\{0\}$, and hence globally Lipschitz on S^{n-1}, and therefore so is Φ^{-1}. Thus Φ is a Lipschitz homeomorphism between the compact sets $\operatorname{bd}(F^\circ)$ and S^{n-1}, with Lipschitz inverse. Consequently, it is easy to see that Φ and Φ^{-1} preserve sets of Hausdorff measure zero. It therefore suffices to prove the result with $\operatorname{bd}(F^\circ)$ in place of S^{n-1}.

Consider, therefore, any vector $c \in \operatorname{bd}(F^\circ)$ such that the optimization problem $\max_F \langle c, \cdot\rangle$ has no nondegenerate maximizers. The intersection of F° with a supporting hyperplane at c gives a proper exposed face G of F containing c. Furthermore, Corollary 2.1 implies $c \notin \operatorname{ri} G$, so c must lie in the relative boundary of G. The result now follows, because by a result of Larman [17], the union of the relative boundaries of the proper faces of an n-dimensional compact convex set has $(n-1)$-dimensional Hausdorff measure zero. □

Corollary 2.2 (Generic Nondegenerate Maximization and Quadratic Decay). *Denote by L the set of vectors $c \in \mathbf{R}^n$ such that the linear functional $\langle c, \cdot\rangle$ decays quadratically around a nondegenerate maximizer over the set F. Then the cone $L \cup \{0\}$ is full measure in $\mathbf{R}^n$.*

Proof. This result follows by combining Theorem 2.1 with Proposition 2.2. □

It is interesting to compare this approach to nondegeneracy with the development of Pataki and Tunçel [29]. Their framework consists of a primal conic convex program,

$$\inf_{(L+b)\cap K} \langle d, \cdot\rangle,$$

where L is a linear subspace and K is a pointed closed convex cone with nonempty interior, and a corresponding dual problem

$$\inf_{(L^\perp+d)\cap K^+} \langle b, \cdot\rangle,$$

where K^+ is the dual cone $\{s\colon \langle z, s\rangle \geq 0\ \forall z \in K\}$. For simplicity, suppose $0 \in \operatorname{int} F$. Then by choosing

$$d = (-c, 0) \in \mathbf{R}^n \times \mathbf{R}, \qquad L = \mathbf{R}^n \times \{0\}, \qquad b = (0, 1), \qquad K = \mathbf{R}_+(F \times \{1\}),$$

we arrive at a primal conic convex program equivalent to $\max_F \langle c, \cdot\rangle$. The dual problem is easy to construct (because $K^+ = \mathbf{R}_+(-F^\circ \times \{1\})$) and reduces to $\inf\{r\colon (c, r) \in \operatorname{epi} \sigma_F\}$, so the dual optimal solution is $-(c, \sigma_F(c))$.

Now suppose the problem $\max_F \langle c, \cdot\rangle$ has unique solution x_c. Then the unique optimal solution of the corresponding primal conic convex program is $(x_c, 1)$, which lies in the relative interior of the exposed face $G = \mathbf{R}_+(x_c, 1)$. In the language of Pataki and Tunçel [29], *strict complementarity* in this case amounts to the optimal dual solution lying in the relative interior of the *conjugate face*, which reduces to

$$G^\triangle = \{(-y, \langle x_c, y\rangle)\colon y \in N_F(x_c)\}$$

(via calculations similar to those above). Strict complementarity therefore holds exactly when $c \in \operatorname{ri} N_F(x_c)$, the property we refer to as “nondegeneracy.” Although Pataki and Tunçel [29] shows that strict complementarity is a generic property, that result concerns variations to the whole data triple (d, b, L), and so it is independent of ours.

3. Semialgebraic functions and stratification. When our underlying feasible region F is semialgebraic, the rather classical arguments we presented in the previous section have much stronger implications. We therefore next introduce the semialgebraic tools we use and discuss their immediate implications.

As we remarked in the Introduction, a subset of $\mathbf{R}^n$ is *semialgebraic* if it is a finite union of sets, each defined by finitely many polynomial inequalities. A function (or set-valued mapping) is *semialgebraic* if its graph is semialgebraic.

Semialgebraic sets and functions enjoy many structural properties. In particular, every semialgebraic set can be written as a finite disjoint union of manifolds (or "strata") that fit together in a regular "stratification": see Basu et al. [2], for example, or the exposition in van den Dries and Miller [42, §4.2]. In particular, the *dimension* of a semialgebraic set is the maximum of the dimensions of the strata, a number independent of the stratification; see Coste [8, Definition 9.14] for more details.

This paper concerns "generic" properties. It is therefore worth emphasizing that, as an immediate consequence of stratification, the following four properties of semialgebraic sets $S \subset \mathbf{R}^n$ are equivalent:

- S is dense;
- the dimension of the complement of S is strictly less than n;
- S is full-measure;
- S is topologically generic.

To see this, note that the complement S^c is semialgebraic and hence is a finite union of (relatively open) manifolds M_j, and by definition, $\dim S^c < n$ if and only if $\dim M_j < n$ for all j. If S^c is dense, then none of the manifolds M_j can be open, and hence $\dim M_j < n$ for all j. Conversely, if $\dim M_j < n$ for all j, then each complement M_j^c is both full-measure and topologically generic, whence so is their intersection, namely, S. Because, in general, full-measure or topologically generic sets are always dense, the equivalence follows.

In this paper we make fundamental use of a stratification result. We present a particular case—adapted to our needs—of a more general result: see van den Dries and Miller [42, §1.19 (2), p. 502] or Loi [24] for the statement in its full generality. The result describes a decomposition of the domain of a semialgebraic function into subdomains on which the function has "constant rank": a smooth function has *constant rank* if its derivative has constant rank throughout its domain. Such functions have a simple canonical form: they are locally equivalent to projections, as described by the following result from basic differential geometry Lee [18, Theorem 7.8].

Proposition 3.1 (Constant Rank Theorem). *Let M_1 and M_2 be two differentiable manifolds of dimensions m_1 and m_2, respectively, and let $g\colon M_1 \to M_2$ be a differentiable mapping of constant rank $r \le \min\{m_1, m_2\}$. Then for every point $x \in M_1$, there exist neighborhoods O_i of zero in $\mathbf{R}^{m_i}$ and local diffeomorphisms $\psi_i\colon O_i \to M_i$ (for $i = 1, 2$) with $\psi_1(0) = x$ and $\psi_2(0) = g(x)$, such that mapping $\psi_2^{-1} \circ g \circ \psi_1$ is just the projection $\pi\colon O_1 \to O_2$ defined by*

$$\pi(y_1, y_2, \ldots, y_{m_1}) = (y_1, y_2, \ldots, y_r, 0, \ldots, 0) \in \mathbf{R}^{m_2}, \quad (y \in O_1). \tag{5}$$

The key stratification result we use follows. For our purposes in this work, readers unfamiliar with differential geometry could simply replace the term "stratification" with "decomposition into a disjoint union of subsets."

Proposition 3.2 (Map Stratification). *Let $f\colon M \to \mathbf{R}^m$ be a semialgebraic function, where M is a semialgebraic subset of $\mathbf{R}^n$. Then, for any integer $k = 1, 2, \ldots,$ there exist a stratification $\mathcal{S} = \{S_i\}_i$ of M into C^k semialgebraic submanifolds and a stratification $\mathcal{T}$ of $\mathbf{R}^m$ into C^k semialgebraic submanifolds, such that the restriction f_i of f onto each stratum $S_i \in \mathcal{S}$ is a C^k semialgebraic function, $f_i(S_i) \in \mathcal{T}$, and f_i is of constant rank in S_i.*

Note. The term "C^k" can in fact be replaced by "real-analytic."

The above proposition yields that each restriction $f_i\colon S_i \to f_i(S_i)$ is surjective, C^k, and of constant rank r_i. Thus r_i is also equal to the dimension of the manifold $f_i(S_i)$:

$$r_i = \operatorname{rank} f_i = \dim \operatorname{Im}(df_i(x)) = \dim T_{f_i(S_i)}(f_i(x)), \quad \text{for all } x \in S_i.$$

Semialgebraic assumptions strengthen conclusions about sensitivity analysis, as the following refinement of Corollary 2.2 shows.

Corollary 3.1 (Generic Stability). *For any nonempty semialgebraic compact convex set $F \subset \mathbf{R}^n$, and any integer $k = 1, 2, \ldots,$ there exists a dense open semialgebraic set $G \subset \mathbf{R}^n$ with the following property. The semialgebraic map $c \mapsto x_c$, taking vectors $c \in \mathbf{R}^n$ to the minimum-norm solution of the optimization problem $\max_F \langle c, \cdot\rangle$, is C^k-smooth throughout G. Furthermore, for all vectors $c \in G$, the objective $\langle c, \cdot\rangle$ decays quadratically over F around the optimal solution x_c, and $c \in \operatorname{ri} N_F(x_c)$.*

Proof. That the map $c \mapsto x_c$ is semialgebraic follows by a standard argument using the Tarski-Seidenberg theorem (Basu [2], Coste [8]). The existence of a dense open semialgebraic set $G' \subset \mathbf{R}^n$ on which this map is C^k follows by applying the stratification result. Consider the full-measure set L guaranteed by Corollary 2.2. Because the set F is semialgebraic, another application of the Tarski-Seidenberg theorem shows that so is the set L, and hence it must contain a dense, open semialgebraic subset, again using the stratification result. The result follows by defining $G = G' \cap L$. □

We can replace the term "C^k" with "real-analytic."

4. Partial Smoothness. The generic stability result, Corollary 3.1, does not yet capture the full force of classical sensitivity analysis under second-order sufficient conditions, because it lacks the active set philosophy. We turn next, therefore, to the idea of partial smoothness introduced in Lewis [21], specialized to the case of a convex set F. This idea involves a continuity property of the normal cone mapping N_F, so we first recall the definition of continuity for a set-valued mapping.

In general, for two metric spaces X and Y and a set-valued mapping $T\colon X \rightrightarrows Y$, we say (cf. Phelps [30], Rockafellar and Wets [34]) that T is *outer semicontinuous* at a point $\bar{x} \in X$ if, for any sequence of points $x^r \in X$ converging to $\bar{x}$ and any sequence of points $y^r \in T(x^r)$ converging to $\bar{y}$, we must have $\bar{y} \in T(\bar{x})$. On the other hand, we say that T is *inner semicontinuous* at $\bar{x}$ if, for any sequence of points $x^r \in X$ converging to $\bar{x}$ and any point $\bar{y} \in T(\bar{x})$, there exists a sequence $y^r \in Y$ converging to $\bar{y}$ such that $y^r \in T(x^r)$ for all large r. If both properties hold, we call T *continuous* at $\bar{x}$.

Definition 4.1. A closed convex set $F \subset \mathbf{R}^n$ is called *partly smooth* at a point $\bar{x} \in F$ relative to a set $\mathcal{M} \subset F$ if the following properties hold:

(i) $\mathcal{M}$ is a C^2 manifold around $\bar{x}$ (called the *active manifold*).

(ii) The normal cone mapping $x \mapsto N_F(x)$, restricted to the domain $\mathcal{M}$, is continuous at $\bar{x}$.

(iii) $N_{\mathcal{M}}(\bar{x}) = N_F(\bar{x}) - N_F(\bar{x})$.

Some comments are in order about this definition which, in this convex case, was shown in Lewis [21] to be exactly equivalent to the definition of an *identifiable surface* introduced by Wright [43].

The idea of a manifold that we use here is rudimentary. Following Rockafellar and Wets [34], we say that a set $\mathcal{M} \subset \mathbf{R}^n$ is a C^2 *manifold (of codimension m) around* a point $\bar{x} \in \mathcal{M}$ if there exists a C^2 map $G\colon \mathbf{R}^n \to \mathbf{R}^m$ with the properties that $G(\bar{x}) = 0$, the derivative map $dG(\bar{x})\colon \mathbf{R}^n \to \mathbf{R}^m$ is surjective, and the inverse image $G^{-1}(0)$ coincides with $\mathcal{M}$ on a neighborhood of $\bar{x}$: we refer to $G(x) = 0$ as a *local equation* for $\mathcal{M}$. We remark that, in the standard language of differential geometry, C^2 submanifolds of $\mathbf{R}^n$ have this property around every point.

The definition of partial smoothness involves several interpretations of the normal cone. Given a point $x \in F$, the set $N_F(x)$ is the usual normal cone, in the sense of convex analysis (Rockafellar [32]). We can regard N_F as a set-valued mapping $x \mapsto N_F(x)$. On the other hand, $N_{\mathcal{M}}(\bar{x})$ is the normal space in the usual sense of differential geometry: the orthogonal complement of the tangent space $T_{\mathcal{M}}(\bar{x})$. Given a representation as above of the set $\mathcal{M}$ as $G^{-1}(0)$ (locally), $T_{\mathcal{M}}(\bar{x})$ is just the null space of the derivative $dG(\bar{x})$ (independent of the choice of the map G). The definition of both $N_F(\bar{x})$ and $N_{\mathcal{M}}(\bar{x})$ is subsumed by the more general variational-analytic idea of the normal cone developed in Rockafellar and Wets [34] and Mordukhovich [27].

Geometrically, condition (iii) guarantees that the set F has a "sharpness" property around the active manifold $\mathcal{M}$, as illustrated in the following simple examples:

Example 4.1 (Sharpness). In the space $\mathbf{R}^3$, define sets

$$F_1 = \{(u, v, w)\colon w \geq u^2 + |v|\},$$
$$F_2 = \{(u, v, w)\colon w \geq (|u| + |v|)^2\},$$
$$\mathcal{M} = \{(t, 0, t^2)\colon t \in (-1, 1)\}.$$

Then the closed convex set F_1 is partly smooth at the point $\bar{x} = (0, 0, 0)$ relative to the active manifold $\mathcal{M}$. On the other hand, the closed convex set F_2 is not: an easy calculation shows that

$$N_{F_2}(0, 0, 0) = \{(0, 0, w)\colon w \leq 0\},$$
$$N_{\mathcal{M}}(0, 0, 0) = \{(0, v, w)\colon v, w \in \mathbf{R}\},$$

so condition (iii) fails. □

It is standard and easy to check that the set-valued mapping $x \mapsto N_F(x)$ is always outer semicontinuous on the set F, and hence on $\mathcal{M}$. Part (ii) of the definition of partial smoothness therefore reduces to the inner semicontinuity property. The following example illustrates how this continuity property can fail.

Example 4.2 (Failure of Normal Continuity). In the space $\mathbf{R}^3$, define sets

$$F = \{(u, v, w): v \geq 0,\ w \geq 0,\ v + w \geq u^2\},$$
$$\mathcal{M} = \{(t, t^2, 0): t \in (-1, 1)\}.$$

The set F is closed and convex, and conditions (i) and (iii) are satisfied at the point $\bar{x} = (0, 0, 0)$. However, condition (ii) fails, because the normal cone mapping is discontinuous at zero, relative to $\mathcal{M}$. □

While not obvious from the above definition, the active manifold for a partly smooth convex set is locally unique around the point of interest: see Hare and Lewis [13, Cor. 4.2].

For the purposes of sensitivity analysis, partial smoothness is most useful when combined with a second-order sufficiency condition, captured by the following definition.

Definition 4.2. Consider a vector $\bar{c} \in \mathbf{R}^n$ and a closed convex set $F \subset \mathbf{R}^n$ that is partly smooth at a point $\bar{x} \in \arg\max_F \langle \bar{c}, \cdot \rangle$ relative to a set $\mathcal{M} \subset F$. We say that $\bar{x}$ is *strongly critical* if the following properties hold:

(i) *Nondegeneracy*: $\bar{c} \in \operatorname{ri} N_F(\bar{x})$.

(ii) *Quadratic decay*: There exists a constant $\delta > 0$ such that

$$\langle \bar{c}, \bar{x} \rangle \geq \langle \bar{c}, x \rangle + \delta |x - \bar{x}|^2 \quad \text{for all } x \in \mathcal{M} \text{ sufficiently near } \bar{x}.$$

The nondegeneracy condition is analogous to the "strict complementarity" condition in classical nonlinear programming. The classical tangent cone $T_F(\bar{x}) = \operatorname{cl} \mathbf{R}_+(F - \bar{x})$ and the normal cone $N_F(\bar{x})$ are mutually dual, so

$$\bar{c} \in N_F(\bar{x}) \quad \Leftrightarrow \quad \langle \bar{c}, d \rangle \leq 0 \quad \text{for all } d \in T_F(\bar{x}),$$

whereas Rockafellar [32, Thm 13.1] shows

$$\bar{c} \in \operatorname{ri} N_F(\bar{x}) \quad \Leftrightarrow \quad \langle \bar{c}, d \rangle < 0 \text{ for all } d \in T_F(\bar{x}) \backslash (-T_F(\bar{x})).$$

Just as in the nonlinear programming case, assumptions (i) and (ii) yield uniqueness of the maximizer $\bar{x}$ for the optimization problem $\max_F \langle \bar{c}, \cdot \rangle$. The following analogue of classical sensitivity results shows that strong criticality also implies that good sensitivity properties hold, and that the active manifold, locally, is simply the set of optimal solutions of perturbed problems.

Theorem 4.1 (Second-Order Sufficiency). *Consider a vector $\bar{c} \in \mathbf{R}^n$ and a compact convex set $F \subset \mathbf{R}^n$ that is partly smooth at a point $\bar{x} \in \arg\max_F \langle \bar{c}, \cdot \rangle$ relative to a set $\mathcal{M} \subset F$. Assume $\bar{x}$ is strongly critical. Then $\langle \bar{c}, \cdot \rangle$ is maximized over F by $\bar{x}$ and decays quadratically around it. Furthermore, there exists a neighborhood U of $\bar{c}$ and a C^1 map from U to $\mathcal{M}$, denoted as $c \mapsto x_c$, mapping a vector $c \in U$ to the unique optimal solution x_c of the perturbed problem* $\max_F \langle c, \cdot \rangle$. *The set $\{x_c: c \in U\}$ is locally identical to $\mathcal{M}$ near $\bar{x}$.*

With the exception of the last statement, this result can be found in Lewis [21] and Hare and Lewis [13, Theorem 6.2]. The last statement amounts to the observation that every point in $\mathcal{M}$ near $\bar{x}$ can be written as x_c for some vector $c \in U$. Consider any sequence $x^r \in \mathcal{M}$ approaching $\bar{x}$ for which this representation fails. Because the normal cone mapping N_F is continuous when restricted to $\mathcal{M}$, there exist vectors $c^r \in N_F(x^r)$ approaching $\bar{c}$. Because $c^r \in U$ for all large indices r, we deduce $x_r = x_{c_r}$, which is a contradiction.

Just as in classical sensitivity analysis, the assumptions and conclusions of this theorem can fail, even on straightforward examples. A good illustration is the following example.

Example 4.3. Consider the convex optimization problem over $\mathbf{R}^3$,

$$\inf\{w: w \geq (|u| + |v|)^2\}. \tag{6}$$

As we perturb the linear objective function slightly, the corresponding optimal solution does not vary smoothly: it may lie on not one but on two distinct manifolds containing zero. More precisely, consider the perturbed problem

$$\inf\{-2au - 2bv + w: w \geq (|u| + |v|)^2\},$$

for parameters $a, b \in \mathbf{R}$. When $|a| = |b| \neq 0$, the optimal solution is not unique. Furthermore, whenever $|a| \neq |b|$, the optimal solution is unique but is given by

$$\begin{cases} (a, 0, a^2) & (|a| > |b|), \\ (0, b, b^2) & (|a| < |b|). \end{cases}$$

Clearly, this optimal solution is not a smooth function of the parameters. □

Nonetheless, following Spingarn and Rockafellar [40], we argue that, for a broad class of concrete optimization problems, such breakdowns are rare, and therefore that partial smoothness and strong criticality are reasonable assumptions. Specifically, we prove in the next section that, for nonempty semialgebraic compact convex sets F, the second-order sufficiency assumptions of Theorem 4.1 hold generically.

5. Main result. Henceforth we assume that the nonempty compact convex set $F \subset \mathbf{R}^n$ is semialgebraic. Our main result asserts that a generic linear optimization problem over F has a unique optimal solution, that F is partly smooth there, and that strong criticality holds. The proof we develop shows how the corresponding active manifold arises naturally by means of Proposition 3.2 (constant rank stratification) applied to an appropriate function.

Theorem 5.1 (Generic Partial Smoothness). *Given any nonempty compact convex semialgebraic set $F \subset \mathbf{R}^n$, for almost all linear objective functions $\langle c, \cdot\rangle$ (indeed, for all nonzero vectors c in a dense open semialgebraic cone in $\mathbf{R}^n$), the optimization problem $\max_F \langle c, \cdot\rangle$ has the following properties.*

(i) *Existence of a* nondegenerate maximizer with quadratic decay, $x_c \in F$.

(ii) Partial smoothness *of F at x_c relative to a semialgebraic C^2 manifold $\mathcal{M}_c \subset F$.*

(iii) Local uniqueness *of the* active manifold *$\mathcal{M}_c$ near x_c, and "local constancy": there exists a semialgebraic C^2 manifold $\mathcal{M}$ such that, for all vectors c' sufficiently near c, we can choose $\mathcal{M}_{c'} = \mathcal{M}$.*

(iv) C^1-smooth dependence *of the unique optimal solution of the perturbed problem $\max_F \langle c', \cdot\rangle$ as the vector c' varies near c: furthermore, this optimal solution lies on the active manifold $\mathcal{M}$.*

Notes: Before we begin the proof, we make some comments about this result. As we have observed, any dense semialgebraic subset of $\mathbf{R}^n$ must be full-measure and topologically generic, with a complement whose dimension is strictly less than n, and it must contain a dense open semialgebraic subset. By an obvious positive homogeneity argument, it suffices to find a dense semialgebraic subset of the unit sphere S^{n-1} on which the conclusions of the theorem hold.

It is interesting to revisit the simple convex optimization problem (6). Truncating the feasible region (by intersecting with the unit ball, for example), we obtain a convex compact semialgebraic set over which the functional $\bar{c} = (0, 0, -1)$ has a unique maximizer (the origin) satisfying the first generic condition asserted in the theorem, while failing the last three. Thus the vector $\bar{c}$ lies outside the asserted dense semialgebraic set.

Proof. Let us consider the set-valued mapping $\tilde{\Phi}: S^{n-1} \rightrightarrows F$ defined by

$$\tilde{\Phi}(c) = \arg\max_F \langle c, \cdot\rangle, \tag{7}$$

so in fact $\tilde{\Phi} = (\partial\sigma_F)|_{S^{n-1}}$ and

$$\tilde{\Phi}^{-1}(x) = N_F(x) \cap S^{n-1}. \tag{8}$$

The Tarski-Seidenberg theorem (Basu [2], Coste [8]) shows that $\tilde{\Phi}$ is semialgebraic. Let L denote the subset of $\mathbf{R}^n$ asserted in Corollary 2.2. Because $L \cup \{0\}$ is a dense cone, the set $D = L \cap S^{n-1}$ is dense in the sphere S^{n-1}. Because the set F is semialgebraic, so is D, again by the Tarski-Seidenberg theorem. Consequently, the set

$$N_* = S^{n-1} \backslash D \tag{9}$$

has dimension strictly less than $n - 1$.

Let Φ: $D \to F$ denote the restriction of the mapping $\tilde{\Phi}$ to D, so in other words $\Phi = (\partial\sigma_F)|_D$. Observe that Φ is single-valued: in our previous notation, for all vectors $c \in D$, we have $\Phi(c) = x_c$, and the strict complementarity and quadratic decay conditions hold:

(i) $c \in \operatorname{ri} N_F(\Phi(c))$;

(ii) for some $\delta > 0$,

$$\langle c, \Phi(c)\rangle \geq \langle c, x\rangle + \delta|x - \Phi(c)|^2 \quad \text{for all } x \in F.$$

Applying Proposition 3.2 (map stratification) to the semialgebraic function Φ, we arrive at a stratification $\mathcal{S} = \{S_j\}_{j\in J}$ of D such that for every index $j \in J$,
- $\Phi_j := \Phi|_{S_j}$ is a C^2 function of constant rank;
- $\Phi_j(S_j)$ is a manifold of dimension equal to the rank of Φ_j;
- the image strata $\{\Phi(S_j)\}_j$ belong to a stratification of $\mathbf{R}^n$.

In particular,

$$D = \bigcup_{j\in J} S_j \tag{10}$$

and

$$j_1 \neq j_2 \;\Rightarrow\; \Phi(S_{j_1}) = \Phi(S_{j_2}) \quad \text{or} \quad \Phi(S_{j_1}) \cap \Phi(S_{j_2}) = \varnothing. \tag{11}$$

Denote the set of strata of full dimension by $\{S_{j_1}, \ldots, S_{j_l}\}$. Observe that the set

$$U = \bigcup_{i=1}^{\ell} S_{j_i}$$

is dense in D and, hence, in S^{n-1}.

Our immediate objective is to show that for every vector $c \in U$, the set F is partly smooth at $\Phi(c)$ with respect to some set $\mathcal{M} \subset F$. To this end, fix any point $\bar{x} \in \Phi(U)$. For any point $x \in \Phi(U)$, define the set of "active" indices as

$$I(x) := \{j \in J\colon\, x \in \Phi(S_j)\}.$$

We aim to show that the set F is partly smooth at $\bar{x}$ relative to the set

$$\mathcal{M} = \Phi_j(S_j)$$

for any index $j \in I(\bar{x})$. Note that, in view of property (11), the definition of $\mathcal{M}$ is in fact independent of the choice of index j, and for the same reason the set of active indices $I(x)$ is independent of the point $x \in \mathcal{M}$. In what follows, we simply write I for the set $I(\bar{x})$.

Clearly, property (i) of the definition of partial smoothness (Definition 4.1) holds. If we can prove properties (ii) and (iii), then our result will follow from Corollary 2.2, because $U \subset D$.

Step 1. **Normal cone continuity.** We establish inner continuity (and hence continuity) at $\bar{x}$ of the normal cone mapping $x \mapsto N_F(x)$ as x moves along the manifold $\mathcal{M}$. It suffices to prove that the truncated normal cone mapping $\tilde{\Phi}^{-1}$ defined by Equation (8) is inner semicontinuous. We proceed by decomposition with respect to the active strata.

For any point $x \in \mathcal{M}$, define

$$N_j(x) = N_F(x) \cap S_j \quad (j \in J). \tag{12}$$

Note that

$$N_j(x) \neq \emptyset \;\Leftrightarrow\; j \in I \;\Leftrightarrow\; \mathcal{M} = \Phi(S_j). \tag{13}$$

We therefore have

$$N_F(x) \cap S^{n-1} = N_*(x) \cup \bigcup_{j\in I} N_j(x),$$

where $N_*(x) = N_F(\bar{x}) \cap N_*$ for the set N_* is defined by Equation (9).

Claim 1 A. *For every point $x \in \mathcal{M}$ the set $\cup_{j\in I} N_j(x)$ is dense in $N_F(x) \cap S^{n-1}$.*

Proof of Claim A. Because we are assuming $\bar{x} \in \Phi(U)$, there exists an active index j_p with $p \in \{1, \ldots, \ell\}$ corresponding to a full-dimensional stratum S_{j_p} such that $\mathcal{M} = \Phi_{j_p}(S_{j_p})$ (see property (13)). This yields that for every point $x \in \mathcal{M}$, there exists a vector $c \in S_{j_p}$ with $x = \Phi(c)$. Hence

$$c \in N_F(x) \cap S_{j_p} = N_{j_p}(x) \subset \bigcup_{j\in I} N_j(x).$$

Now fix any vector $c_* \in N_F(x) \cap S^{n-1}$, and consider the spherical path

$$c_t := \frac{c + t(c_* - c)}{|c + t(c_* - c)|} \quad (t \in [0, 1]).$$

Because $x = \Phi(c)$, we in fact know $c \in \operatorname{ri} N_F(x)$. It follows that $c_t \in \operatorname{ri} N_F(x)$, for all $t \in [0, 1)$. Because $c \in S_{j_p} \subset D$, there exists a constant $\delta_c > 0$ such that

$$\langle c, x\rangle \geq \langle c, x'\rangle + \delta_c |x - x'|^2 \quad \text{for all } x' \in F.$$

By the definition of the normal cone, we also have

$$\langle c_*, x\rangle \geq \langle c_*, x'\rangle \quad \text{for all } x' \in F.$$

Multiplying these inequalities by $(1-t)$ and t, respectively, and adding, we infer that point x is a maximizer of the function $\langle c_t, \cdot\rangle$ over the set F, with quadratic decay, for all $0 \leq t < 1$. In other words, $c_t \in N_F(x) \cap D$, which in view of Equation (10) yields $c_t \in \bigcup_{j\in I} N_j(x)$, for $t \in [0, 1)$. Since $c_t \to c_*$ as $t \uparrow 1$, Claim A follows. □

In view of Claim A, it is sufficient to establish the inner continuity of the mapping

$$x \mapsto \bigcup_{j\in I} N_j(x) \quad x \in \mathcal{M}. \tag{14}$$

To see this, we use the following simple and routine result: for completeness, we provide a proof.

Lemma 5.1. *Let X and Y be metric spaces, and consider two set-valued mappings $G, T\colon X \rightrightarrows Y$ such that $\operatorname{cl}(G(x)) = T(x)$ for all points $x \in X$. If G is inner semicontinuous at a point $\bar{x} \in X$, then so is T.*

Proof of Lemma 5.1. Assume (towards a contradiction) that there exists a constant $\rho > 0$, a sequence $\{x^k\} \subset X$ with $x^k \to \bar{x}$ and a point $\bar{y} \in T(\bar{x})$, such that

$$\operatorname{dist}(\bar{y}, T(x^k)) > \rho > 0.$$

Then pick any point $\hat{y} \in B(\bar{y}, \rho/2) \cap G(\bar{x})$ and use the inner semicontinuity of G to get a sequence of points $y^k \in G(x^k) \subset T(x^k)$ for $k \in \mathbf{N}$ such that $y^k \to \hat{y}$. This gives a contradiction, proving the lemma. □

Applying this lemma to the set-valued mappings

$$G(x) = \bigcup_{j\in I} N_j(x) \qquad \text{and} \qquad T(x) = N_F(x) \cap S^{n-1}$$

accomplishes the reduction we seek.

Let us therefore prove the inner semicontinuity of the mapping defined in (14) at the point $\bar{x}$. To this end, fix any vector $\bar{c} \in \bigcup_{j\in I} N_j(\bar{x})$ and consider any sequence $\{x^k\} \subset \mathcal{M}$ approaching $\bar{x}$. For some index $j \in I$ we have $\bar{c} \in S_j$. Let us restrict our attention to the constant-rank surjective mapping $\Phi_j\colon S_j \to \mathcal{M}$ and let us recall that

$$\Phi_j(S_j) = \mathcal{M} \qquad \text{and} \qquad \Phi_j(\bar{c}) = \bar{x}.$$

Let d be the dimension of the stratum S_j, so

$$\operatorname{rank}(d\Phi_j) = \dim \mathcal{M} := r \leq d \leq n-1.$$

Denote by 0_d (respectively 0_r) the zero vector of the space $\mathbf{R}^d$ (respectively $\mathbf{R}^r$). Then applying the constant rank theorem (Proposition 3.1), we infer that for some constants $\delta, \varepsilon > 0$ there exist diffeomorphisms

$$\psi_1\colon B(0_d, \delta) \to S_{j_0} \cap B(\bar{c}, \varepsilon) \qquad \text{and} \qquad \psi_2\colon B(0_r, \delta) \to \mathcal{M} \cap B(\bar{x}, \varepsilon) \tag{15}$$

such that

$$\psi_1(0_d) = \bar{c} \qquad \text{and} \qquad \psi_2(0_r) = \bar{x} \tag{16}$$

and such that all vectors $y \in B(0_d, \delta)$ satisfy

$$(\psi_2^{-1} \circ \Phi_j \circ \psi_1)(y) = \pi(y), \tag{17}$$

where for $y = (y_1, \ldots, y_d) \in \mathbf{R}^d$ we have

$$\pi(y_1, \ldots, y_r, y_{r+1} \ldots, y_d) = (y_1, \ldots, y_r) \in B(0_r, \delta) \subset \mathbf{R}^r.$$

We may assume that the sequence $\{x^k\}$ lies in $\mathcal{M} \cap B(\bar{x}, \varepsilon)$. Thus, in view of definition (15), for every integer $k \in \mathbf{N}$ there exists a vector $z^k = (z_1^k, \ldots, z_r^k) \in B(0_r, \delta)$ with

$$\psi_2(z^k) = x^k. \tag{18}$$

Note $z^k \to 0_r = (\psi_2)^{-1}(\bar{x})$. Define vectors

$$y^k := (z_1^k, \ldots, z_r^k, 0, \ldots, 0) \in \mathbf{R}^d$$

for every $k \in \mathbf{N}$. Because $z^k \in B(0_r, \delta)$, we know $y^k \in B(0_d, \delta)$, and clearly

$$y^k \to 0_d. \tag{19}$$

We now define vectors $c^k := \psi_1(y^k)$ for each k. In view of definition (15) we see that $c^k \in S_j \cap B(\bar{c}, \varepsilon)$, and in view of properties (19) and (16),

$$c^k \to \psi_1(0_d) = \bar{c} \quad \text{as } k \to \infty.$$

To complete the proof of inner semicontinuity, it remains to show that $c^k \in N_F(x^k)$. Because $\Phi_j(c^k) = \Phi_j(\psi_1(y^k))$ we infer by properties (17) and (19) that

$$\psi_2^{-1}(\Phi_j(c^k)) = (\psi_2^{-1} \circ \Phi_j \circ \psi_1)(y^k) = \pi(y^k) = z^k.$$

Using now the fact that ψ_2 is a diffeomorphism, we deduce from Equation (18) that $\Phi_j(c^k) = \psi_2(z^k) = x^k$. Thus $c^k \in \Phi_j^{-1}(x^k) \subset N_F(x^k)$, which completes the proof of inner semicontinuity and hence of Step 1.

Step 2. **Sharpness.** It remains to verify that condition (iii) of Definition 4.1, namely

$$N_{\mathcal{M}}(\bar{x}) = N_F(\bar{x}) - N_F(\bar{x}), \tag{20}$$

is also fulfilled.

To this end, as in the proof of Claim A, we can choose an index $j \in I$ corresponding to a stratum S_j of full dimension $(n-1)$ such that $\mathcal{M} = \Phi_j(S_j)$. Recall that the semialgebraic C^2-mapping $\Phi_j\colon S_j \to \mathcal{M}$ is surjective and has constant rank $r = \dim \mathcal{M}$, so $\dim N_{\mathcal{M}}(\bar{x}) = n - r$. It follows directly from the inclusion $\mathcal{M} \subset F$ that $N_F(\bar{x}) \subset N_{\mathcal{M}}(\bar{x})$. Because the right-hand side is a subspace, we deduce

$$N_F(\bar{x}) - N_F(\bar{x}) \subset N_{\mathcal{M}}(\bar{x}). \tag{21}$$

Because Φ_j is surjective and of maximal rank, we deduce easily that $\Phi_j^{-1}(\bar{x})$ is a semialgebraic submanifold of S^{n-1} of dimension $(n-1) - r$, which, in view of definition (8) and Equation (12), yields

$$\dim(N_F(\bar{x}) \cap S^{n-1}) \geq \dim N_j(\bar{x}) \geq (n-1) - r.$$

Thus $\dim N_F(\bar{x}) \geq n - r$, which, along with inclusion (21), yields Equation (20), as required. □

A simple argument extends the main result to unbounded feasible regions.

Corollary 5.1 (Unbounded Feasible Regions). *Given any nonempty closed convex semialgebraic set $F \subset \mathbf{R}^n$, for almost all vectors c in the domain of the support function σ_F, the optimization problem $\max_F \langle c, \cdot \rangle$ has all the properties described in Theorem 5.1.*

Proof. It suffices to prove the result for all compact subsets G of the interior of the domain of σ_F. Because σ_F is locally Lipschitz throughout the interior of its domain, it is globally Lipschitz on any such set G. Denoting the Lipschitz constant by L, we note for all vectors $c \in G$ the property

$$\operatorname*{arg\,max}_F \langle c, \cdot \rangle = \partial \sigma_F(c) \subset B(0, L).$$

Thus the original problem $\max_F \langle c, \cdot \rangle$ is equivalent to a problem with a compact feasible region: $\max_{F \cap B(0,L)} \langle c, \cdot \rangle$. Applying the main result to this latter problem completes the proof. □

An assumption like semialgebraicity (or, more generally, tameness) is crucial for results like those above. To see this, consider first any closed proper convex function f on $\mathbf{R}$. For any point x in the interior of the domain of f, the epigraph of f is partly smooth at the point $(x, f(x))$ if and only if f is either nondifferentiable at x or C^2 around x.

Now consider any strictly increasing function $h\colon [0,1] \to \mathbf{R}_+$ that is discontinuous on a dense set. Define functions $g, f\colon [0,1] \to \mathbf{R}$ by

$$g(y) = \int_0^y h(t)\,dt \qquad f(x) = \int_0^x g(y)\,dy.$$

Then f is C^1 and convex on the interval $(0,1)$, but nowhere C^2. Finally, let f take the value $+\infty$ outside of the interval $[0,1]$. Then, as we have observed, the closed convex set $F = \operatorname{epi} f$ is not partly smooth at any point $(x, f(x))$ with $x \in (0,1)$.

However, for this set F and any vector $c \in \mathbf{R}^2$ satisfying $-c_1/c_2 = g(x)$ and $c_2 < 0$, it is easy to check that the unique maximizer for the problem $\max_F \langle c, \cdot\rangle$ is exactly the point $\{(x, f(x))\}$. Thus the conclusion of Corollary 5.1 fails whenever the ratio $-c_1/c_2$ lies in the interval $(g(0), g(1))$ and $c_2 < 0$. A straightforward extension of this idea gives an example of a compact convex set $F \subset \mathbf{R}^2$ over which the function $\langle c, \cdot\rangle$ has a unique maximizer for every nonzero vector $c \in \mathbf{R}^2$, but such that F is never partly smooth around that maximizer.

Acknowledgments. A major part of this work was accomplished during a research visit of the second and third authors at the University Pierre et Marie Curie (Paris 6) in May 2008. The authors thank Dmitriy Drusviyatskiy for the observation following Theorem 4.1 and Victor Vinnikov for the suggestion to include the final example. The two first authors acknowledge the support of Grants ANR-08-BLAN-0294-01 (France) and MTM2008-06695-C03-03 (Spain). Research of the third author was supported in part by NSF Grant DMS-0806057.

References

[1] Alizadeh, F., J.-P. A. Haeberly, M. L. Overton. 1997. Complementarity and nondegeneracy in semidefinite programming. *Math. Programming* **77** 111–128.
[2] Basu, S., R. Pollack, M.-F. Roy. 2003. *Algorithms in Real Algebraic Geometry*. Springer, Berlin.
[3] Ben-Tal, A., A. Nemirovski. 2001. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM, Philadelphia.
[4] Bonnans, J. F., A. Shapiro. 2000. *Perturbation Analysis of Optimization Problems*. Springer, New York.
[5] Burke, J. V., J. J. Moré. 1988. On the identification of active constraints. *SIAM J. Numer. Anal.* **25** 1197–1211.
[6] Burke, J. V., A.S. Lewis, M. L. Overton. 2001. Optimal stability and eigenvalue multiplicity. *Foundations Comput. Math.* **1** 205–225.
[7] Clarke, F. H., Y. S. Ledyaev, R. J. Stern, P. R. Wolenski. 1998. *Nonsmooth Analysis and Control Theory*. Springer-Verlag, New York.
[8] Coste, M. 1999. An Introduction to o-minimal geometry. RAAG Notes, Institut de Recherche Mathématique de Rennes, Rennes, France.
[9] Daniilidis, A., C. Sagastizábal, M. Solodov. 2009. Identifying structure of nonsmooth convex functions by the bundle technique. *SIAM J. Optim.* **20** 820–840.
[10] Dontchev, A. L., T. Zolezzi. 1993. *Well-Posed Optimization Problems*. Springer-Verlag, Berlin.
[11] Ewald, G., D. G. Larman, C. A. Rogers. 1970. The directions of the line segments and of the r-dimensional balls on the boundary of a convex body in Euclidean space. *Mathematika* **17** 1–20.
[12] Ferris, M. C. 1991. Finite termination of the proximal point algorithm. *Math. Programming* **50** 359–366.
[13] Hare, W. L., A. S. Lewis. 2004. Identifying active constraints via partial smoothness and prox-regularity. *J. Convex Anal.* **11** 251–266.
[14] Helton, J. W., J. Nie. 2010. Semidefinite representation of convex sets. *Math. Programming* **122** 21–64.
[15] Ioffe, A. D. 2009. An invitation to tame optimization. *SIAM J. Optim.* **19** 1894–1917.
[16] Ioffe, A. D., R. Lucchetti. 2005. Typical convex program is very well posed. *Math. Programming* **104** 483–499.
[17] Larman, D. G. 1971. On a conjecture of Klee and Martin for convex bodies. *Proc. London Math. Soc.* **23** 668–682. Corrigendum: Vol. **36**, p. 86.
[18] Lee, J. M. 2003. *Introduction to Smooth Manifolds*. Springer, New York.
[19] Lemaréchal, C., C. Sagastizábal. 1997. Practical aspects of the Moreau-Yosida regularization: Theoretical preliminaries. *SIAM J. Optim.* **7** 367–385.
[20] Lemaréchal, C., F. Oustry, C. Sagastizábal. 2000. The $\mathcal{U}$-Lagrangian of a convex function. *Trans. Amer. Math. Soc.* **352** 711–729.
[21] Lewis, A. S. 2003. Active sets, nonsmoothness and sensitivity. *SIAM J. Optim.* **13** 702–725.
[22] Lewis, A. S., S. J. Wright. 2008. A proximal method for composite minimization. arXiv:0812.0423v1.
[23] Lewis, A. S., S. J. Wright. 2010. Identifying activity. *SIAM J. Optim.* Forthcoming. arXiv:0901.2668v1.
[24] Loi, T. L. 1997. Thom stratifications for functions defined in o-minimal structures on $(\mathbb{R}, +, \cdot)$. *Comptes Rendus de l'Académie des Sciences de Paris, Série 1* **324** 1391–1394.
[25] Mifflin, R., C. Sagastizábal. 2002. Proximal points are on the fast track. *J. Convex Anal.* **9** 563–579.
[26] Mifflin, R., C. Sagastizábal. 2005. A VU algorithm for convex minimization. *Math. Programming* **104** 583–608.
[27] Mordukhovich, B. S. 2006. *Variational Analysis and Generalized Differentiation, Vol. I, Basic Theory, and Vol. II, Applications*. Springer, New York.
[28] Nocedal, J., S. J. Wright. 1997. *Nonlinear Programming*. Springer-Verlag, New York.
[29] Pataki, G., L. Tunçel. 2001. On the generic properties of convex optimization problems in conic form. *Math. Programming* **89** 449–457.

[30] Phelps, R. R. 1993. *Convex Functions, Monotone Operators and Differentiability*, 2nd ed. *Lecture Notes in Mathematics*, Vol. 1364. Springer-Verlag, New York.
[31] Robinson, S. M. 2003. Constraint nondegeneracy in variational analysis. *Math. Oper. Res.* **28** 201–232.
[32] Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.
[33] Rockafellar, R. T. 1976. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14** 877–898.
[34] Rockafellar, R. T., R. J.-B. Wets. 1998. *Variational Analysis*. Springer, Berlin.
[35] Rogers, C. A. 1970. *Hausdorff Measures*. Cambridge University Press, Cambridge, UK.
[36] Saigal, R., C. Simon. 1973. Generic properties of the complementarity problem. *Math. Programming* **4** 324–335.
[37] Shapiro, A. 1997. First and second order analysis of nonlinear semidefinite programs. *Math. Programming* **77** 301–320.
[38] Shapiro, A., M. K. H. Fan. 1995. On eigenvalue optimization. *SIAM J. Optim.* **5** 552–568.
[39] Spingarn, J. E. 1982. On optimality conditions for structured families of nonlinear programming problems. *Math. Programming* **22** 82–92.
[40] Spingarn, J. E., R. T. Rockafellar. 1979. The generic nature of optimality conditions in nonlinear programming. *Math. Oper. Res.* **4** 425–430.
[41] Sun, D. 2006. The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their consequences. *Math. Oper. Res.* **31** 761–776.
[42] van den Dries, L., C. Miller. 1996. Geometric categories and o-minimal structures. *Duke Math. J.* **84** 497–540.
[43] Wright, S. J. 1993. Identifiable surfaces in constrained optimization. *SIAM J. Control Optim.* **31** 1063–1079.