

4-1-2010

# Applying Metric Regularity to Compute Condition Measure of Smoothing Algorithm for Matrix Games

Boris S. Mordukhovich  
Wayne State University, boris@math.wayne.edu

Javier Peña  
Carnegie Mellon University, jfp@andrew.cmu.edu

Vera Roshchina  
Universidade de Évora, Portugal, veraroshchina@gmail.com

---

## Recommended Citation

Mordukhovich, Boris S.; Peña, Javier; and Roshchina, Vera, "Applying Metric Regularity to Compute Condition Measure of Smoothing Algorithm for Matrix Games" (2010). *Mathematics Research Reports*. Paper 75.  
[http://digitalcommons.wayne.edu/math\\_reports/75](http://digitalcommons.wayne.edu/math_reports/75)

This Technical Report is brought to you for free and open access by the Mathematics at DigitalCommons@WayneState. It has been accepted for inclusion in Mathematics Research Reports by an authorized administrator of DigitalCommons@WayneState.

**APPLYING METRIC REGULARITY TO COMPUTE  
CONDITION MEASURE OF SMOOTHING  
ALGORITHM FOR MATRIX GAMES**

**BORIS S. MORDUKHOVICH, JAVIER PEÑA and VERA ROSHCHINA**

**WAYNE STATE  
UNIVERSITY**

**Detroit, MI 48202**

**Department of Mathematics  
Research Report**

**2010 Series  
#4**

*This research was partly supported by the US National Science Foundation*

# APPLYING METRIC REGULARITY TO COMPUTE CONDITION MEASURE OF SMOOTHING ALGORITHM FOR MATRIX GAMES

BORIS S. MORDUKHOVICH<sup>1</sup>, JAVIER PEÑA<sup>2</sup> and VERA ROSHCINA<sup>3</sup>

**Abstract.** We develop an approach of variational analysis and generalized differentiation to conditioning issues for two-person zero-sum matrix games. Our major results establish precise relationships between a certain condition measure of the smoothing first-order algorithm proposed in [4] and the exact bound of metric regularity for an associated set-valued mapping. In this way we compute the aforementioned condition measure in terms of the initial matrix game data.

**Key words.** matrix games, smoothing algorithm, condition measure, variational analysis, metric regularity, generalized differentiation

**AMS subject classifications.** 90D10, 90C30, 49J53, 49J52, 49M45

**Abbreviated title.** Condition measure for matrix games

## 1 Introduction and formulation of main results

This paper is devoted to applications of advanced techniques in variational analysis and generalized differentiation to the study of conditioning in optimization. Our specific goal is to apply the key notions and generalized differential characterizations of Lipschitzian stability and metric regularity, fundamental in variational analysis, to computing a certain condition measure of the first-order smoothing algorithm proposed in [4] to find approximate Nash equilibria of two-person zero-sum matrix games.

To the best of our knowledge, applications of Lipschitzian stability and metric regularity to numerical aspects of optimization were initiated by Robinson in the 1970s; see, e.g., [16] and the references therein. In the complexity theory, Renegar [14, 15] established relationships between the rate of convergence of interior-point methods for linear and conic convex programs and his “distance to ill-posedness/infeasibility” and condition numbers. We refer the reader to [1, 2, 3, 6, 7, 13] and their bibliographies for more recent results in this direction for various algorithms in convex and nonconvex optimization problems.

In [4], a new condition measure was introduced to evaluate the complexity of a first-order algorithm for solving a two-person zero-sum game

$$(1.1) \quad \min_{x \in Q_1} \max_{y \in Q_2} x^T A y = \max_{y \in Q_2} \min_{x \in Q_1} x^T A y,$$

where  $A \in \mathbb{R}^{m \times n}$ , where the symbol  $^T$  stands for transposition, and where each of the sets  $Q_1$  and  $Q_2$  is either a simplex (in the matrix game formulation) or a more elaborate

---

<sup>1</sup>Wayne State University, Detroit, MI, USA (boris@math.wayne.edu). Research of this author was partly supported by the US National Science Foundation under grant DMS-0603846 and by the Australian Research Council under grant DP-12092508.

<sup>2</sup>Carnegie Mellon University, Pittsburgh, PA, USA (jfp@andrew.cmu.edu). Research of this author was partly supported by the US National Science Foundation under grant CCF-0830533.

<sup>3</sup>Universidade de Évora, Évora, Portugal (veraroshchina@gmail.com).

polytope (in the case of sequential games). Problems of this type arise in many interesting applications; see, e.g., [20] and the references therein.

It was shown in [4] that an iterative version of Nesterov's first-order smoothing algorithm [11, 12] computes an  $\varepsilon$ -equilibrium point (in the sense of Nash) for problem (1.1) in  $\mathcal{O}(\|A\|\kappa(A)\ln(1/\varepsilon))$  iterations, where  $\kappa(A)$  is a condition number of (1.1) depending only on  $A$ ; see (1.7) for the definition of the condition measure  $\kappa(A)$  in the case of matrix games. The dependence of this complexity bound on  $\varepsilon$  is exponentially better than the complexity bound  $\mathcal{O}(1/\varepsilon)$  in the original Nesterov's smoothing techniques. Furthermore, it was proved in [4] that the condition measure  $\kappa(A)$  is always finite while the proof therein was non-constructive. In particular, no explicit upper bound on  $\kappa(A)$  was given.

In this paper we focus on the *matrix game equilibrium problem*

$$(1.2) \quad \min_{x \in \Delta_m} \max_{y \in \Delta_n} x^T A y = \max_{y \in \Delta_n} \min_{x \in \Delta_m} x^T A y,$$

where the  $m$ -dimensional simplex

$$\Delta_m := \left\{ x \in \mathbb{R}^m \mid \sum_{i=1}^m x_i = 1, x \geq 0 \right\}$$

describes the set of mixed strategies for the  $x$ -player (Player 1) with  $m$  pure strategies; similarly for the  $y$ -player  $y \in \Delta_n$  (Player 2). This means that if Player 1 uses  $x \in \Delta_m$  and Player 2 uses  $y \in \Delta_n$ , then Player 1 gets payoff  $-x^T A y$  while Player 2 gets payoff  $x^T A y$ . Thus the equilibrium problem (1.2) can be reformulated as the following problem of *nonsmooth convex optimization*:

$$(1.3) \quad \text{minimize } F(x, y) \text{ subject to } (x, y) \in \Delta_m \times \Delta_n,$$

where the minimizing cost function  $F(x, y)$  is defined by the maximum

$$(1.4) \quad F(x, y) := \max \{ x^T A v - u^T A y \mid (u, v) \in \Delta_m \times \Delta_n \}.$$

It is easy to observe that

$$(1.5) \quad \min \{ F(x, y) \mid (x, y) \in \Delta_m \times \Delta_n \} = 0.$$

Taking (1.5) into account, we say [4, 20] that a feasible pair  $(\bar{x}, \bar{y}) \in \Delta_m \times \Delta_n$  is a *Nash equilibrium* to (1.2) if  $F(\bar{x}, \bar{y}) = 0$ , which corresponds to an optimal solution of the constrained optimization problem (1.3). Consider the optimal solution set

$$(1.6) \quad S := \{ (\bar{x}, \bar{y}) \in \Delta_m \times \Delta_n \mid F(\bar{x}, \bar{y}) = 0 \} = F^{-1}(0) \cap (\Delta_m \times \Delta_n)$$

and, following [4], define the *condition measure*  $\kappa(A)$  of the matrix game (1.2) depending on the underlying matrix  $A$  via the objective (1.4) and the optimal solution set (1.6) as

$$(1.7) \quad \kappa(A) := \inf \{ \kappa \geq 0 \mid \text{dist}((x, y); S) \leq \kappa F(x, y) \text{ for all } (x, y) \in \Delta_m \times \Delta_n \},$$

where  $\text{dist}(\cdot; S)$  stands for the standard Euclidean distance function.

In what follows we derive three major results concerning the computation of the condition measure  $\kappa(A)$  in (1.7). The first theorem shows that the condition measure  $\kappa(A)$

precisely relates to the exact bound of metric regularity for an associated set-valued mapping built upon the cost function (1.4). The second result expresses this exact regularity bound via the subdifferential of the convex function (1.4) and the normal cone to the simplex product  $\Delta_m \times \Delta_n$  and then computes the latter constructions in terms of the initial data of (1.2). Finally, we arrive at an exact formula for evaluating  $\kappa(A)$ , which is a key step towards performing further complexity analysis of the algorithm [4].

To formulate the first theorem, define a set-valued mapping  $\Phi: \mathbb{R}^{m+n} \rightrightarrows \mathbb{R}$  by

$$(1.8) \quad \Phi(x, y) := \begin{cases} [F(x, y), \infty) & \text{if } (x, y) \in \Delta_m \times \Delta_n, \\ \emptyset & \text{otherwise} \end{cases}$$

via the cost function  $F$  constructed in (1.4). Let  $\text{reg } \Phi((x, y), F(x, y))$  be the *exact bound of metric regularity* (or the *exact regularity bound*) of the mapping  $\Phi$  around the point  $((x, y), F(x, y)) \in \text{gph } \Phi$ ; see Section 2 for more details.

**Theorem 1.1 (condition measure via the exact regularity bound).** *Assume that  $(\Delta_m \times \Delta_n) \setminus S \neq \emptyset$  with  $S$  defined in (1.6). Then we have the precise relationship*

$$(1.9) \quad \kappa(A) = \sup_{(x, y) \in (\Delta_m \times \Delta_n) \setminus S} \text{reg } \Phi((x, y), F(x, y))$$

between the condition measure (1.7) and the exact regularity bound of (1.8).

To formulate the second major result, let  $a_i$  as  $i = 1, \dots, n$  and  $-b_k^T$  as  $k = 1, \dots, m$  stand for the columns and the rows of the matrix  $A$ , respectively. By  $e_j$ ,  $j = 1, \dots, m+n$ , we denote the unit vectors in  $\mathbb{R}^{m+n}$ , i.e.,

$$(e_j)_l = 0 \text{ for all } l \neq j \text{ and } (e_j)_j = 1 \text{ as } j = 1, \dots, m+n.$$

For a positive integer  $p$ , let  $\mathbf{1}_p := [1 \ \dots \ 1] \in \mathbb{R}^p$ . Given finally a feasible point  $(x, y) \in \Delta_m \times \Delta_n$ , form the corresponding index sets by

$$(1.10) \quad \begin{cases} I(x) := \{\bar{i} \in \{1, \dots, n\} \mid a_{\bar{i}}^T x = \max_{i \in \{1, \dots, n\}} a_i^T x\}, \\ K(y) := \{\bar{k} \in \{1, \dots, m\} \mid b_{\bar{k}}^T y = \max_{k \in \{1, \dots, m\}} b_k^T y\}, \\ J(x, y) := \{j \in \{1, \dots, m\} \mid x_j = 0\} \cup \{j = m+p \mid y_p = 0\}. \end{cases}$$

**Theorem 1.2 (computing the exact bound of metric regularity).** *For any point  $(x, y) \in (\Delta_m \times \Delta_n) \setminus S$  the exact regularity bound of the mapping  $\Phi$  from (1.8) around the point  $((x, y), F(x, y))$  admits the representation*

$$(1.11) \quad \text{reg } \Phi((x, y), F(x, y)) = \frac{1}{\text{dist}(0; \partial F(x, y) + N_{\Delta_m \times \Delta_n}(x, y))}$$

via the subdifferential of the convex function (1.4) and the normal cone to the simplex product  $\Delta_m \times \Delta_n$  at  $(x, y)$ . Furthermore, the latter constructions are computed by

$$(1.12) \quad \partial F(x, y) = \text{co} \{ (a_i, b_k) \in \mathbb{R}^m \times \mathbb{R}^n \mid i \in I(x), k \in K(y) \},$$

$$(1.13) \quad N_{\Delta_m \times \Delta_n}(x, y) = \text{span}\{\mathbf{1}_m\} \times \text{span}\{\mathbf{1}_n\} - \text{cone}[\text{co}\{e_j \mid j \in J(x, y)\}]$$

in the notation above, where the symbols  $\text{span}$ ,  $\text{cone}$ , and  $\text{co}$  stand respectively for the linear, conic, and convex hulls of the sets in question.

Unifying the results of Theorem 1.1 and Theorem 1.2, we get the following precise formula for computing the condition measure of the smoothing algorithm for matrix games.

**Theorem 1.3 (computing the condition measure).** *Let  $(\Delta_m \times \Delta_n) \setminus S \neq \emptyset$ . Then, in the notation above, the condition measure  $\kappa(A)$  defined in (1.7) is computed by*

$$\kappa(A) = \sup_{(x,y) \in (\Delta_m \times \Delta_n) \setminus S} \left[ \text{dist}\left(0; \text{co}\{(a_i, b_k) \mid i \in I(x), k \in K(y)\} + \text{span}\{\mathbf{1}_m\} \times \text{span}\{\mathbf{1}_n\} - \text{cone}[\text{co}\{e_j \mid j \in J(x, y)\}]\right) \right]^{-1}.$$

The proofs of Theorem 1.1 and Theorem 1.2 given below are based on applying advanced techniques of variational analysis and generalized differentiation. This approach leads us therefore to deriving the precise formula for the condition measure in Theorem 1.3. For additional insight, we also present a direct, independent proof of the latter theorem relying on more conventional while somewhat more laborious techniques of convex optimization employing particularly Lagrangian duality.

**Remark 1.4 (numerical implementation and further research).** Numerical implementation of the formula for the condition measure in Theorem 1.3 is not a purpose of this paper and in fact is not an easy job. It has been well recognized in complexity theory that evaluating condition numbers may be in general as difficult as to solve the original problem. This is true, e.g., in the cases of such fundamental complexity measures as the condition number of a matrix [5] used in estimating complexity of numerical linear algebra algorithms, Renegar's condition number [13, 14, 15] that characterizes difficulty of solving conic feasibility problems, the "measure of condition" for finding zeros of complex polynomials introduced by Shub and Smale [19], etc.

The main purpose of this paper is not obtaining an easily computable expression for the condition measure  $\kappa(A)$ , but rather gaining a better understanding on how exactly the problem data influence the condition measure. Observe that the formula for  $\kappa(A)$  obtained in Theorem 1.3 is much easier to evaluate and analyze than the original construction (1.7). This is valuable for the average-case and smoothed analysis of the algorithm, singling out classes of well-conditioned problems, preconditioning issues, and making further improvements to the algorithm. We will pursue these goals in our subsequent research.

The rest of the paper is organized as follows. In Section 2 we recall some basic definitions and facts of variational analysis and generalized differentiation crucial for deriving the main results of the paper. Section 3 is devoted to variational proofs of the main results formulated above. Finally, in Section 4 we present an alternative direct proof of Theorem 1.3 by employing tools of convex optimization.

Throughout the paper we use standard notation and terminology of variational analysis; see, e.g., the basic texts [10, 18].

## 2 Preliminaries from variational analysis and generalized differentiation

Here we confine ourselves to finite-dimensional Euclidean spaces sufficient for the subsequent applications. The reader is referred to [10, 18] for more details and related material.

Given a set-valued mapping  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , consider its inverse  $G^{-1}: \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  with  $x \in G^{-1}(z) \iff z \in G(x)$  as well as its graphs, domain, and range defined respectively by

$$\text{gph } G := \{(x, z) \mid z \in G(x)\}, \quad \text{dom } G := \{x \mid G(x) \neq \emptyset\}, \quad \text{rge } G := \text{dom } G^{-1}.$$

The notion of metric regularity is of primary interest in our development.

**Definition 2.1 (metric regularity).** *A set-valued mapping  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is METRICALLY REGULAR around  $(\bar{x}, \bar{z}) \in \text{gph } G$  with modulus  $\mu \geq 0$  if there exist neighborhoods  $U$  of  $\bar{x}$  and  $V$  of  $\bar{z}$  such that*

$$(2.1) \quad \text{dist}(x; G^{-1}(z)) \leq \mu \text{dist}(z; G(x)) \text{ whenever } x \in U \text{ and } z \in V.$$

The infimum of  $\mu \geq 0$  over all  $(\mu, U, V)$  for which (2.1) holds is called the EXACT REGULARITY BOUND of  $G$  around  $(\bar{x}, \bar{z})$  and is denoted by  $\text{reg } G(\bar{x}, \bar{z})$ .

It is well known in variational analysis that the fundamental property of metric regularity is closely related to Lipschitzian behavior of inverse mappings. Recall that a mapping  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is *Lipschitz-like* (or has the Aubin property) around  $(\bar{x}, \bar{z}) \in \text{gph } G$  with modulus  $\ell \geq 0$  if there are neighborhoods  $U$  of  $\bar{x}$  and  $V$  of  $\bar{z}$  such that

$$(2.2) \quad G(x) \cap V \subset G(u) + \ell \|x - u\| B \text{ for all } x, u \in U,$$

where  $B$  stands for closed unit ball of the space in question. The infimum of  $\ell \geq 0$  over all the combinations  $(\ell, U, V)$  for which (2.2) holds is called the *exact Lipschitzian bound* of  $G$  around  $(\bar{x}, \bar{z})$  and is denoted by  $\text{lip } G(\bar{x}, \bar{z})$ .

The following result can be found, e.g., in [10, Theorem 1.49].

**Proposition 2.2 (relationships between metric regularity and Lipschitz-like properties).** *Let  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , and let  $(\bar{x}, \bar{z}) \in \text{gph } G$ . Then the mapping  $G$  is metrically regular around  $(\bar{x}, \bar{z})$  if and only if its inverse  $G^{-1}: \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  is Lipschitz-like around  $(\bar{z}, \bar{x})$ . Furthermore, we have the equality*

$$\text{reg } G(\bar{x}, \bar{z}) = \text{lip } G^{-1}(\bar{z}, \bar{x}).$$

One of the key advantages of modern variational analysis is the possibility to completely characterize Lipschitzian and metric regularity properties of set-valued mappings in terms of appropriate generalized differential constructions enjoying full calculus. Let us recall such constructions used in this paper.

Given a nonempty subset  $\Omega \subset \mathbb{R}^n$  and a point  $\bar{x} \in \Omega$ , define the Fréchet/regular normal cone to  $\Omega$  at  $\bar{x}$  by

$$\widehat{N}_\Omega(\bar{x}) := \left\{ v \in \mathbb{R}^n \mid \limsup_{x \rightarrow \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\},$$

where the symbol  $x \xrightarrow{\Omega} \bar{x}$  means that  $x \rightarrow \bar{x}$  with  $x \in \Omega$ . Then the Mordukhovich (basic/limiting) *normal cone* to  $\Omega$  at  $\bar{x} \in \Omega$  is defined by

$$(2.3) \quad N_{\Omega}(\bar{x}) := \operatorname{Lim\,sup}_{x \xrightarrow{\Omega} \bar{x}} \widehat{N}_{\Omega}(x),$$

where ‘Lim sup’ stands for the Painlevé-Kuratowski outer/upper limit of a set-valued mapping  $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  given by

$$\operatorname{Lim\,sup}_{x \rightarrow \bar{x}} M(x) := \left\{ v \in \mathbb{R}^m \mid \exists x_k \rightarrow \bar{x}, v_k \rightarrow v \text{ as } k \rightarrow \infty \text{ such that} \right. \\ \left. v_k \in M(x_k) \text{ for all } k = \mathbb{N} := \{1, 2, \dots\} \right\}.$$

If the set  $\Omega$  is locally closed around  $\bar{x}$ , the normal cone (2.3) admits the equivalent description (which was in fact the original definition in [8])

$$N_{\Omega}(\bar{x}) = \operatorname{Lim\,sup}_{x \rightarrow \bar{x}} [\operatorname{cone}(x - \Pi_{\Omega}(x))]$$

in terms of the projection operator  $\Pi_{\Omega}(x) := \{y \in \Omega \mid \|y - x\| = \operatorname{dist}(x, \Omega)\}$ .

Note that the normal cone (2.3) may be nonconvex even for simple sets  $\Omega \subset \mathbb{R}^n$ , e.g., for the graph of  $|x|$  and the epigraph of  $-|x|$  at  $(0, 0)$ . Due to its nonconvexity, the normal cone (2.3) cannot be polar to any tangent cone. Nevertheless, this nonconvex normal cone and the corresponding subdifferential/coderivative constructions for extended-real-valued (i.e., with values in  $(-\infty, \infty]$ ) functions and set-valued mappings satisfy comprehensive calculus rules, which are derived by using variational arguments, particularly the extremal principle of variational analysis; see [10, 18] and the references therein.

A set  $\Omega$  is called *normally regular* at  $\bar{x} \in \Omega$  if  $N_{\Omega}(\bar{x}) = \widehat{N}_{\Omega}(\bar{x})$ . The class of normally regular sets covers “nice” sets having a local convex-like structure. A major example is provided by convex sets; see, e.g., [10, Proposition 1.5].

**Proposition 2.3 (normal regularity of convex sets).** *Let  $\Omega \subset \mathbb{R}^n$  be convex. Then it is normally regular at every point  $\bar{x} \in \Omega$ , and its normal cone (2.3) reduces to the normal cone in the sense of convex analysis:*

$$N_{\Omega}(\bar{x}) = \{v \in \mathbb{R}^n \mid \langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega\}.$$

Given next a set-valued mapping  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  and a point  $(\bar{x}, \bar{z}) \in \operatorname{gph} G$ , define a generalized derivative of  $G$  at  $(\bar{x}, \bar{z})$  induced by the normal cone (2.3) to the graph  $\Omega = \operatorname{gph} G$ . Namely, the *coderivative* of  $G$  at  $(\bar{x}, \bar{z})$  is a set-valued mapping  $D^*G(\bar{x}, \bar{z}): \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  with the values

$$(2.4) \quad D^*G(\bar{x}, \bar{z})(v) := \{u \in \mathbb{R}^n \mid (u, -v) \in N_{\operatorname{gph} G}(\bar{x}, \bar{z})\}.$$

Observe that  $0 \in D^*G(\bar{x}, \bar{x})(0)$  and  $D^*G(\bar{x}, \bar{x})(\lambda v) = \lambda D^*G(\bar{x}, \bar{x})(v)$ , i.e., the coderivative (2.4) is a positively homogeneous mapping. If  $G: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is single-valued and smooth around  $\bar{x}$  with the derivative  $\nabla G(\bar{x})$ , we have

$$D^*G(\bar{x})(v) = \{\nabla G(\bar{x})^T v\} \text{ for all } v \in \mathbb{R}^m.$$

The latter signifies that the coderivative (2.4) is an appropriate extension of the adjoint/transpose derivative operator to the case of nonsmooth and set-valued mappings. Note also that, by the nonconvexity of the normal cone (2.3), the coderivative (2.4) is not dual to any tangentially generated graphical derivative, except of the case when  $G$  is *graphical regular* at  $(\bar{x}, \bar{z})$  meaning that

$$N_{\text{gph}G}(\bar{x}, \bar{z}) = \widehat{N}_{\text{gph}G}(\bar{x}, \bar{z}).$$

As mentioned above, the coderivative (2.4) satisfies comprehensive calculus rules for general set-valued mappings. In this paper we only need the following one, which is a consequence of [10, Proposition 3.12]. To formulate it, recall that the indicator mapping  $\delta_\Omega: \mathbb{R}^n \rightarrow \mathbb{R}$  of a set  $\Omega \subset \mathbb{R}^n$  is defined by

$$\delta_\Omega(x) := \begin{cases} 0 & \text{if } x \in \Omega, \\ \emptyset & \text{otherwise,} \end{cases}$$

(a bit different from the indicator functions) and that we easily have the relationship

$$D^*(\delta_\Omega)(\bar{x})(v) = N_\Omega(\bar{x}) \text{ for any } \bar{x} \in \Omega \text{ and } v \in \mathbb{R}.$$

**Proposition 2.4 (coderivative sum rule).** *Let  $\Omega \subset \mathbb{R}^n$  be locally closed around  $\bar{x} \in \Omega$ , and let  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}$  be closed-graph and Lipschitz-like around  $(\bar{x}, \bar{z}) \in \text{gph}G$ . Then for all  $v \in \mathbb{R}$  we have the inclusion*

$$D^*(G + \delta_\Omega)(\bar{x}, \bar{z})(v) \subset D^*G(\bar{x}, \bar{z})(v) + N_\Omega(\bar{x}),$$

which holds as equality if  $\Omega$  is normally regular at  $\bar{x}$  and  $F$  is graphically regular at  $(\bar{x}, \bar{z})$ .

In what follows we employ the norm of the coderivative as a positively homogeneous mapping. The norm of a positively homogeneous mapping  $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is defined by

$$\|M\| := \sup \{ \|u\| \mid u \in M(v) \text{ with } \|v\| \leq 1 \}$$

and admits (by passing to the inverse) the useful distance function representation below established in [3, Proposition 2.5].

**Proposition 2.5 (norm of positively homogeneous mappings).** *Let  $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  be positively homogeneous. Then the norm of its inverse is computed by*

$$\|M^{-1}\| = \sup_{\|v\|=1} \frac{1}{\text{dist}(0; M(v))}.$$

The final and most important result presented in this section provides a complete coderivative characterization of the Lipschitz-like property (known as the Mordukhovich criterion [18]) with computing the exact bound of Lipschitzian moduli; see [9, Theorem 5.7], [10, Theorem 4.10], and [18, Theorem 9.40] for different proofs.

**Theorem 2.6 (coderivative characterization of the Lipschitz-like property for set-valued mappings).** *Let  $G: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  be closed-graph around  $(\bar{x}, \bar{z}) \in \text{gph}G$ . Then  $G$  is Lipschitz-like around this point if and only if  $D^*G(\bar{x}, \bar{z})(0) = \{0\}$ . In this case*

$$\text{lip}G(\bar{x}, \bar{z}) = \|D^*G(\bar{x}, \bar{z})\|.$$

### 3 Proofs of main results

We give here complete proofs of Theorem 1.1 and Theorem 1.2, and thus derive the condition measure formula of Theorem 1.3 by variational arguments.

Let us start with the proof of Theorem 1.1. To proceed, we first establish a more convenient representation of the condition measure (1.7) for our further analysis, which in turn is preceded by a technical claim.

Observe that the function  $F(x, y)$  defined by (1.4) can be written as follows:

$$F(x, y) = \max_{\substack{i=1, \dots, n \\ j=1, \dots, m}} (a_i^T x + b_j^T y).$$

In addition we represent the simplex product  $\Delta_m \times \Delta_n$  by:

$$\Delta_m \times \Delta_n = \{w = (x, y) \mid w \geq 0, Ew = f\} \quad \text{with} \quad E := \begin{bmatrix} \mathbf{1}_m^T & 0 \\ 0 & \mathbf{1}_n^T \end{bmatrix} \quad \text{and} \quad f := \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

To simplify notation, rewrite the function  $F$  as

$$F(w) = \max_{\ell \in L} c_\ell^T w,$$

where  $L := \{1, \dots, n\} \times \{1, \dots, m\}$  and  $c_\ell^T := [a_i^T \ b_j^T]$  for each  $\ell = (i, j) \in L$ . Denote further  $\Omega := \Delta_n \times \Delta_m = \{w \mid w \geq 0, Ew = f\}$  and rewrite (1.7) as

$$(3.1) \quad k_F := \inf \{k \geq 0 \mid \text{dist}(w; S) \leq kF(w) \text{ for all } w \in \Omega\}$$

with  $S$  given by (1.6). Observe that  $\min_{\Omega} F(x) = 0$  by (1.5). It is also convenient for us to define the moving sets

$$(3.2) \quad S(z) := \{w \in \Omega \mid F(w) = z\} = F^{-1}(z) \cap \Omega \quad \text{with thus} \quad S = S(0)$$

and to represent the mapping  $\Phi$  in (1.8) and its inverse by

$$(3.3) \quad \Phi(w) = [F(w), \infty) + \delta_{\Omega}(w) \quad \text{and} \quad \Phi^{-1}(z) = \{w \in \Omega \mid F(w) \leq z\}.$$

Let us finally denote  $\mathcal{J} := \{1, \dots, n, n+1, \dots, m+n\}$  and define the corresponding counterparts of the index sets  $I(\cdot)$  and  $J(\cdot)$  from (1.10) given by

$$\mathcal{I}(w) := \{\ell \in L \mid c_\ell^T w = F(w)\}, \quad \mathcal{J}(w) := \{j \in \mathcal{J} \mid w_j = 0\}.$$

It is not difficult to verify the following technical claim, where  $\mathbb{B}_\gamma$  stands for the closed ball of radius  $\gamma > 0$  centered at the origin.

**Claim 3.1 (relationships between index sets).** *For every  $\bar{w} \in \Omega$  there exists  $\gamma > 0$  such that  $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$  and  $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$  whenever  $w \in \bar{w} + \mathbb{B}_\gamma$ .*

**Proof.** Fix an arbitrary element  $\bar{w} \in \Omega \setminus S$  and let

$$0 < \gamma < \min \left\{ \min_{\substack{\ell \in L, \\ c_\ell \neq 0}} \frac{1}{2\|c_\ell\|} \min_{i \in \mathcal{I} \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^\top \bar{w}), \min_{j \in 1:m+n \setminus \mathcal{J}(\bar{w})} \bar{w}_j \right\},$$

where  $\min \emptyset = \infty$  by convention. It is easy to observe that such a number  $\gamma$  always exists. When  $\mathcal{I}(\bar{w}) = L$ , the inclusion  $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$  is obvious. Assume thus that  $L \setminus \mathcal{I}(\bar{w}) \neq \emptyset$ . For every  $\ell_0 \in L \setminus \mathcal{I}(\bar{w})$  and every  $w \in \bar{w} + B_\gamma$  we have

$$c_{\ell_0}^\top w = c_{\ell_0}^\top \bar{w} + c_{\ell_0}^\top (w - \bar{w}) \leq c_{\ell_0}^\top \bar{w} + \|c_{\ell_0}\| \gamma < c_{\ell_0}^\top \bar{w} + \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^\top \bar{w}),$$

which implies the relationships

$$(3.4) \quad \begin{aligned} \max_{i \in \mathcal{I} \setminus \mathcal{I}(\bar{w})} (a_i^\top w) &< \max_{i \in \mathcal{I} \setminus \mathcal{I}(\bar{w})} (c_i^\top \bar{w}) + \frac{1}{2} \min_{\ell \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_\ell^\top \bar{w}) \\ &= \frac{1}{2} \left( F(\bar{w}) + \max_{i \in L \setminus \mathcal{I}(\bar{w})} a_i^\top \bar{w} \right). \end{aligned}$$

Similarly, for every  $\ell_0 \in \mathcal{I}(\bar{w})$  we have

$$c_{\ell_0}^\top w = c_{\ell_0}^\top \bar{w} + c_{\ell_0}^\top (w - \bar{w}) \geq c_{\ell_0}^\top \bar{w} - \|c_{\ell_0}\| \gamma > c_{\ell_0}^\top \bar{w} - \frac{1}{2} \min_{i \in \mathcal{I} \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_i^\top \bar{w}),$$

which in turn implies that

$$(3.5) \quad \begin{aligned} \max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^\top w) &> \max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^\top \bar{w}) - \frac{1}{2} \min_{i \in L \setminus \mathcal{I}(\bar{w})} (F(\bar{w}) - c_i^\top \bar{w}) \\ &= \frac{1}{2} \left( F(\bar{w}) + \max_{i \in L \setminus \mathcal{I}(\bar{w})} c_i^\top \bar{w} \right). \end{aligned}$$

Combining (3.4) and (3.5) gives us the strict inequality

$$\max_{\ell \in \mathcal{I}(\bar{w})} (c_\ell^\top w) - \max_{\ell \in \mathcal{I} \setminus \mathcal{I}(\bar{w})} (c_\ell^\top w) > 0,$$

and hence justifies the first claimed inclusion  $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$ .

It remains to show that  $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$  for all  $w \in \bar{w} + B_\gamma$ . The latter inclusion is obvious when  $\mathcal{J}(\bar{w}) = \mathcal{J}$ . Assume now that  $\mathcal{J} \setminus \mathcal{J}(\bar{w}) \neq \emptyset$  and then get for every  $w \in \bar{w} + B_\gamma$  and  $j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})$  the relationships

$$w_j = \bar{w}_j + w_j - \bar{w}_j > \bar{w}_j - \gamma \geq \bar{w}_j - \min_{j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})} \bar{w}_j \geq 0.$$

Thus  $w_j > 0$  whenever  $j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})$ , and thus we arrive at  $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$ .  $\triangle$

The next result provides a useful representation of the condition measure (3.1) convenient for our subsequent analysis.

**Lemma 3.2 (representation of condition measure).** *Assume that  $\Omega \setminus S \neq \emptyset$ . Then there exists  $\bar{w} \in \Omega \setminus S$  such that*

$$(3.6) \quad k_F = \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)} = \frac{\text{dist}(\bar{w}; S)}{F(\bar{w})}.$$

**Proof.** For every  $w \in \Omega \setminus S$  define

$$k(w) := \frac{\text{dist}(w; S)}{F(w)}$$

and find a sequence  $\{w_p\} \subset \Omega \setminus S$  such that  $k_F = \lim_{p \rightarrow \infty} k(w_p)$ . Fixing a number  $p \in \mathbb{N}$ , let  $\bar{w}_p$  be a point of  $S$  closest to  $w_p$  and denote

$$w_t^p := \bar{w}_p + t(w^p - \bar{w}_p).$$

By Claim 3.1 there is  $t_p \in (0, 1]$  such that  $\mathcal{I}(w_t^p) \subset \mathcal{I}(\bar{w}_p)$  and  $\mathcal{J}(w_t^p) \subset \mathcal{J}(\bar{w}_p)$  for all  $t \in (0, t_p)$ . Denoting  $v^p := w_{t_p}^p$ , observe that  $\text{dist}(v^p; S) = t_p \text{dist}(w^p; S)$  and thus  $F(v^p) \leq t_p F(w^p)$  by the convexity of the cost function  $F$ . This yields  $k(v^p) \geq k(w^p)$ , and hence  $k_F = \lim_{p \rightarrow \infty} k(v^p)$ . Since the set  $L$  is finite, assume without loss of generality that the corresponding active indices are the same for all  $v^p$  and also the same for all  $\bar{w}_p$ , i.e.,

$$(3.7) \quad \tilde{\mathcal{I}} := \mathcal{I}(v^p) \subset \mathcal{I}(\bar{w}_p) =: \hat{\mathcal{I}}.$$

Considering further the closure

$$\tilde{\Omega} := \text{cl} \{w \in \Omega \setminus S \mid \mathcal{I}(w) = \tilde{\mathcal{I}}, \mathcal{I}(\bar{x}) = \hat{\mathcal{I}}\},$$

conclude that both sets  $\tilde{\Omega}$  and  $\tilde{S} := \tilde{\Omega} \cap S$  are polytopes, that  $\{v^p\} \subset \tilde{\Omega}$ , and that for every  $w \in \tilde{\Omega}$  we have  $\bar{w} \in \tilde{S}$ . Thus

$$\tilde{S} = \text{co}\{s_1, \dots, s_N\} \quad \text{and} \quad \tilde{\Omega} = \text{co}\{w_1, w_2, \dots, w_M, s_1, \dots, s_N\}$$

with some  $N, M \in \mathbb{N}$ , where  $w_i \notin \tilde{S}$  for all  $i = 1, \dots, M$ . Picking any  $w \in \tilde{\Omega}$ , we get

$$w = \sum_{i=1}^M \lambda_i w_i + \sum_{k=1}^N t_k s_k \quad \text{with} \quad \sum_{i=1}^M \lambda_i + \sum_{k=1}^N t_k = 1 \quad \text{and} \quad \lambda_i, t_k \geq 0$$

whenever  $i = 1, \dots, M$  and  $k = 1, \dots, N$ . Fix now  $i = 1, \dots, M$  and let  $\bar{w}_i$  be a point in  $\tilde{S}$  closest to  $w_i$ . Then there are coefficients  $\mu_j^i$  as  $j = 1, \dots, N$  such that

$$\bar{w}_i = \sum_{j=1}^N \mu_j^i s_j, \quad \sum_{j=1}^N \mu_j^i = 1, \quad \text{and} \quad \mu_j^i \geq 0 \quad \text{for all } j = 1, \dots, N.$$

Combining the above, we arrive at the representations

$$\begin{aligned} \min_{w' \in \tilde{S}} \|w - w'\| &= \min_{\substack{\sum_{k=1}^N \tau_k = 1 \\ \tau_k \geq 0}} \left\| \sum_{i=1}^M \lambda_i w_i + \sum_{k=1}^N t_k s_k - \sum_{k=1}^N \tau_k s_k \right\| \\ &= \min_{\substack{\sum_{k=1}^N \tau_k = 1 \\ \tau_k \geq 0}} \left\| \sum_{i=1}^M \lambda_i \left( w_i - \frac{\sum_{k=1}^N (\tau_k - t_k) s_k}{\sum_{i=1}^M \lambda_i} \right) \right\|. \end{aligned}$$

Replacing  $\tau_k$  by  $\gamma_k := (\tau_k - t_k) / \sum_{i=1}^M \lambda_i$ , the latter can be written as

$$\min_{w' \in S} \|w - w'\| = \min_{\substack{\sum_{k=1}^N \gamma_k = 1 \\ \gamma_k \geq -\frac{t_k}{\sum_{i=1}^M \lambda_i}}} \left\| \sum_{i=1}^M \lambda_i \left( w_i - \sum_{k=1}^N \gamma_k s_k \right) \right\|.$$

We further let

$$\gamma_k^* := \sum_{i=1}^M \lambda_i \mu_k^i$$

and observe that

$$\gamma_k^* \geq 0 \geq -\frac{t_k}{\sum_{i=1}^M \lambda_i} \quad \text{and} \quad \sum_{k=1}^N \gamma_k^* = \sum_{k=1}^N \sum_{i=1}^M \lambda_i \mu_k^i = \sum_{i=1}^M \sum_{k=1}^N \lambda_i \mu_k^i = \sum_{i=1}^M \lambda_i \sum_{k=1}^N \mu_k^i = 1.$$

This gives therefore that

$$\begin{aligned} \min_{w' \in S} \|w - w'\| &\leq \left\| \sum_{i=1}^M \lambda_i \left( w_i - \sum_{k=1}^N \gamma_k^* s_k \right) \right\| = \left\| \sum_{i=1}^M \lambda_i w_i - \sum_{k=1}^N \gamma_k^* s_k \right\| = \left\| \sum_{i=1}^M \lambda_i w_i - \sum_{k=1}^N \sum_{i=1}^M \lambda_i \mu_k^i s_k \right\| \\ &= \left\| \sum_{i=1}^M \lambda_i \left( w_i - \sum_{k=1}^N \mu_k^i s_k \right) \right\| = \left\| \sum_{i=1}^M \lambda_i (w_i - \bar{w}_i) \right\| \leq \sum_{i=1}^M \lambda_i \text{dist}(w_i; S), \end{aligned}$$

which ensures the estimate

$$(3.8) \quad \text{dist}(w; S) \leq \sum_{i=1}^M \lambda_i \text{dist}(w_i; S).$$

On the other hand,

$$(3.9) \quad F(w) = a_{i_0}^T \left( \sum_{i=1}^M \lambda_i w_i + \sum_{k=1}^N t_k s_k \right) = \sum_{i=1}^M \lambda_i F(w_i)$$

for  $i_0 \in \hat{I}$ , where  $\hat{I}$  is defined in (3.7). Letting

$$K := \max_{i=1, \dots, M} \frac{\text{dist}(w_i; S)}{F(w_i)},$$

it follows from (3.8) and (3.9) that

$$\frac{\text{dist}(w; S)}{F(w)} \leq \frac{\sum_{i=1}^M \lambda_i \text{dist}(w_i; S)}{\sum_{i=1}^M \lambda_i F(w_i)} \leq K,$$

which implies that

$$k(v^p) \leq \max_{i=1, \dots, M} \frac{\text{dist}(w_i; S)}{F(w_i)} \quad \text{for all } p \in \mathbb{N}.$$

Passing finally to the limit as  $p \rightarrow \infty$ , we get

$$k_F \leq \max_{i=1, \dots, M} \frac{\text{dist}(w_i; S)}{F(w_i)},$$

which means that the supremum in (3.6) is attained and thus completes the proof.  $\triangle$

Now we are ready to proof our first main result given in Theorem 1.1.

### 3.1 Proof of Theorem 1.1

We split the proof of the theorem into three major steps.

**Step 1: metric regularity via condition measure.** For every  $\bar{w} \notin S$  and every  $\bar{z} > 0$  we have the distance estimate

$$(3.10) \quad \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) \leq k_F \text{dist}(\Phi(\bar{w}); \bar{z}).$$

**Proof.** When  $\bar{w} \notin \Omega$ , the right-hand side of (3.10) becomes infinity (by the construction of  $\Phi$  in (3.3) and the standard convention on  $\inf \emptyset = \infty$ ) while the left-hand side is finite, i.e., there is nothing to prove. Considering the case of  $\bar{w} \in \Omega$ , observe that the left-hand side of (3.10) becomes zero when  $F(\bar{w}) \leq \bar{z}$ , and thus the inequality holds automatically. It remains to examine the case when  $0 < \bar{z} < F(\bar{w})$  with  $\bar{w} \in \Omega$ .

To proceed, take a point  $w^*$  closest to  $\bar{w}$  in  $\Phi^{-1}(\bar{z})$  and observe that  $F(w^*) = \bar{z}$ , since otherwise the continuity of  $F$  allows us to find a closer point in  $[w^*, \bar{w}] \cap \Phi^{-1}(\bar{z})$ . Thus

$$(3.11) \quad \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) = \text{dist}(\bar{w}; S(\bar{z})) = \|\bar{w} - w^*\|,$$

where  $S(\cdot)$  is defined in (3.2). Picking a point  $w_0$  closest to  $\bar{w}$  in  $S$  and employing again the continuity of  $F$ , we find  $\lambda \in (0, 1)$  such that

$$(3.12) \quad F(w_0 + \lambda(\bar{w} - w_0)) = \bar{z}.$$

In addition the convexity of  $F$  yields that

$$(3.13) \quad \bar{z} = F(w_0 + \lambda(\bar{w} - w_0)) \leq F(w_0) + \lambda(F(\bar{w}) - F(w_0)) = \lambda F(\bar{w}).$$

Combining the above, we have the relationships

$$(3.14) \quad \begin{aligned} \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) &= \text{dist}(\bar{w}; S(\bar{z})) && \text{(by (3.11))} \\ &\leq \|w_0 + \lambda(\bar{w} - w_0) - \bar{w}\| && \text{(by (3.12))} \\ &= (1 - \lambda)\|w_0 - \bar{w}\| = (1 - \lambda) \text{dist}(\bar{w}; S); \end{aligned}$$

$$(3.15) \quad \begin{aligned} \text{dist}(\Phi(\bar{w}); \bar{z}) &= F(\bar{w}) - \bar{z} && \text{(as } \bar{z} < F(\bar{w})) \\ &\geq F(\bar{w}) - \lambda F(\bar{w}) && \text{(by (3.13))} \\ &= (1 - \lambda)F(\bar{w}), \end{aligned}$$

which finally give

$$\begin{aligned} \text{dist}(\bar{w}; \Phi^{-1}(\bar{z})) &\leq (1 - \lambda) \text{dist}(\bar{w}; S) && \text{(by (3.14))} \\ &\leq (1 - \lambda)k_F F(\bar{w}) && \text{(as } \text{dist}(\bar{w}; S) \leq k_F F(\bar{w})) \\ &\leq k_F \text{dist}(\Phi(\bar{w}); \bar{z}) && \text{(by (3.15))} \end{aligned}$$

and thus allow us to arrive at (3.10). △

**Step 2: distance properties.** Let  $\bar{w} \in \Omega \setminus S$  be such that

$$(3.16) \quad k_F = \frac{\text{dist}(\bar{w}; S)}{F(\bar{w})},$$

let  $w_0$  be a point in  $S$  closest to  $\bar{w}$ , and let  $\lambda \in (0, 1)$

$$w_\lambda := \bar{w} + \lambda(w_0 - \bar{w}), \quad 0 < \lambda < 1.$$

Then for any  $\lambda \in (0, 1)$  we have the properties:

$$(i) \quad F(w_\lambda) = (1 - \lambda)F(\bar{w}).$$

$$(ii) \quad \text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \lambda \text{dist}(\bar{w}; S).$$

**Proof.** To justify (i), observe that by the convexity of  $F$  we have

$$(3.17) \quad F(w_\lambda) \leq F(\bar{w}) + \lambda(F(w_0) - F(\bar{w})) \leq (1 - \lambda)F(\bar{w})$$

in the notation above. On the other hand, the definition of  $k_F$  and the choice of  $\bar{w}$  yield

$$(3.18) \quad F(w_\lambda) \geq \frac{\text{dist}(w_\lambda; S)}{k_F} = \frac{(1 - \lambda)\|\bar{w} - w_0\|}{k_F} = \frac{(1 - \lambda) \text{dist}(\bar{w}; S)}{k_F} = (1 - \lambda)F(\bar{w}).$$

Thus assertion (i) follows from (3.17) and (3.18).

To justify (ii), it suffices to show that

$$\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \|\bar{w} - w_\lambda\|.$$

Proceeding by contradiction, assume that  $\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) < \lambda \text{dist}(\bar{w}; S) = \lambda k_F F(\bar{w})$  and take a point  $w^*$  closest to  $\bar{w}$  in  $S(F(w_\lambda))$ . By the continuity of  $F$  we have

$$\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) = \text{dist}(\bar{w}; S(F(w_\lambda))) = \|\bar{w} - w^*\|,$$

which yields therefore that

$$(3.19) \quad \|\bar{w} - w^*\| = \text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda))) < \lambda k_F F(\bar{w}).$$

Taking further a point  $\tilde{w}$  closest to  $w^*$  in  $S$ , we get by (3.1) and by part (i) above that

$$(3.20) \quad \|\tilde{w} - w^*\| \leq k_F F(w^*) = k_F F(w_\lambda) = k_F(1 - \lambda)F(\bar{w}).$$

Since  $\tilde{w} \in S$ , the latter implies that

$$\begin{aligned} \text{dist}(\bar{w}; S) &\leq \|\bar{w} - \tilde{w}\| \\ &\leq \|\bar{w} - w^*\| + \|w^* - \tilde{w}\| \quad (\text{by the triangle inequality}) \\ &< \lambda k_F F(\bar{w}) + (1 - \lambda)k_F F(\bar{w}) \quad (\text{by (3.19) and (3.20)}) \\ &= k_F F(\bar{w}), \end{aligned}$$

which contradicts (3.16) and thus completes the proof of Step 2.  $\triangle$

**Step 3: condition measure via metric regularity.** *We have the equality*

$$(3.21) \quad k_F = \sup_{w \in \Omega \setminus S} \text{reg} \Phi(w, F(w)).$$

**Proof.** Let us first show that

$$(3.22) \quad k_F \geq \sup_{w \in \Omega \setminus S} \text{reg} \Phi(w, F(w)).$$

Assuming the contrary, find  $(w', z') \in \text{gph } F$ ,  $w' \in \Omega \setminus S$  satisfying

$$\text{reg} \Phi(w', z') > k_F.$$

Observe that there exists a neighborhood of  $(w', z')$  such that for all points  $(w, z)$  in that neighborhood  $w \notin S$  and  $z > 0$ . By the definition of metric regularity we can find  $\bar{w}, \bar{z}$  in such a neighborhood of  $(w', z')$  for which

$$\text{dist}(\bar{w}, \Phi^{-1}(\bar{z})) > k_F \text{dist}(\Phi(\bar{w}); \bar{z}).$$

The latter contradicts Step 1 and thus ensures (3.22).

To prove the opposite inequality in (3.21), by Lemma 3.2 find  $\bar{w} \in \Omega \setminus S$  such that

$$\text{dist}(\bar{w}; S) = k_F F(\bar{w}).$$

Let  $w_0$  be a point in  $S$  closest to  $\bar{w}$  and define

$$w_\lambda := \bar{w} + \lambda(w_0 - \bar{w}), \quad 0 < \lambda < 1.$$

It follows from Step 2 that for every  $\lambda \in (0, 1)$  and the above choice of  $\bar{w}$  we have

$$\frac{\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda)))}{\text{dist}(F(w_\lambda); \Phi(\bar{w}))} = \frac{\lambda \text{dist}(\bar{w}; S)}{\lambda F(\bar{w})} = k_F.$$

The latter implies, since  $w_\lambda \rightarrow \bar{w}$  and  $F(w_\lambda) \rightarrow F(\bar{w})$  as  $\lambda \downarrow 0$ , that

$$\text{reg} \Phi(\bar{w}, F(\bar{w})) \geq \limsup_{\lambda \downarrow 0} \frac{\text{dist}(\bar{w}; \Phi^{-1}(F(w_\lambda)))}{\text{dist}(F(w_\lambda); \Phi(\bar{w}))} = k_F,$$

which therefore yields

$$k_F \leq \sup_{w \in \Omega \setminus S} \text{reg} \Phi(w, F(w))$$

and completes the the proof of the theorem.  $\triangle$

### 3.2 Proof of Theorem 1.2

First of all, observe by Step 1 in the proof of Theorem 1.1 that the multifunction  $\Phi$  is metrically regular around  $(w, z) \in \text{gph} \Phi$  for every  $w \in \Omega \setminus S$ . Employing the corresponding results of Section 2, for  $(w, z) \in \text{gph} \Phi$  with  $w \in \Omega \setminus S$  we get

$$\begin{aligned} \text{reg} \Phi(w, z) &= \text{lip} \Phi^{-1}(z, w) && \text{( by Theorem 2.2)} \\ &= \|D^* \Phi^{-1}(z, w)\| && \text{( by Theorem 2.6)} \\ &= \|(D^* \Phi(w, z))^{-1}\| && \text{( by the definition of } D^* \Phi(w, z)) \\ &= \sup_{|v|=1} \frac{1}{\text{dist}(0; D^* \Phi(w, z)(v))} && \text{( by Proposition 2.5 with } m = 1). \end{aligned}$$

This gives therefore the regularity exact bound formula

$$(3.23) \quad \text{reg} \Phi(w, z) = \frac{1}{\min \{ \text{dist}(0; D^* \Phi(w, z)(-1)), \text{dist}(0; D^* \Phi(w, z)(1)) \}}.$$

Defined next a set-valued mapping  $\tilde{F}: \mathbb{R}^{m+n} \rightrightarrows \mathbb{R}$  by

$$(3.24) \quad \tilde{F}(w) := [F(w), \infty) \text{ with } \text{gph} \tilde{F} = \text{epi} F$$

and observe that it is Lipschitz-like at every point of its graph, with is the epigraph of a Lipschitz continuous function. Furthermore, the graph of  $\tilde{F}$  is convex, and hence  $\tilde{F}$  is graphically regular at any point of its graph by Proposition 2.3, which also ensures the normal regularity of the convex set  $\Omega$ . Applying Proposition 2.4 to the sum  $\Phi = \tilde{F} + \delta_\Omega$ , we get the equality

$$(3.25) \quad D^*\Phi(w, z)(\lambda) = D^*\tilde{F}(w, z) + N_\Omega(w) \text{ for all } \lambda \in \mathbb{R}.$$

It follows from the structure of  $\tilde{F}$  in (3.24), the coderivative definition (2.4), and the well-known subdifferential representation

$$\partial\varphi(\bar{x}) = \{v \in \mathbb{R}^n \mid (v, -1) \in N((\bar{x}, \varphi(\bar{x})); \text{epi } \varphi)\}, \quad \bar{x} \in \text{dom } \varphi,$$

for any convex function  $\varphi: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  that

$$(3.26) \quad D^*\tilde{F}(w, z)(1) = \begin{cases} \partial F(w), & z = F(w), \\ \emptyset, & z > F(w), \end{cases} \quad D^*\tilde{F}(w, z)(-1) = \emptyset.$$

Combining (3.23), (3.25), and (3.26) gives us the formula

$$(3.27) \quad \text{reg } \Phi(w, z) = \frac{1}{\text{dist}(0; \partial F(w) + N_\Omega(w))},$$

which is (1.11). It remains to justify the subdifferential representation (1.12).

To proceed, we recall first the calculus formula

$$N_{A \cap B}(w) = N_A(w) + N_B(w)$$

held at every  $w \in A \cap B$  for arbitrary convex polyhedra in finite dimensions; see, e.g., [17, Corollary 23.8.1]). Thus we have in our case that

$$N_\Omega(w) = N_{\{Eu=f\}}(w) + N_{\{u \geq 0\}}(w).$$

Moreover, it is easy to see that

$$N_{\{Eu=f\}}(w) = (\ker E)^\top = \text{span}\{\mathbf{1}_m\} \times \text{span}\{\mathbf{1}_n\},$$

$$N_{\{u \geq 0\}}(w) = -\text{cone}[\text{co}\{e_j \mid j \in J(w)\}].$$

Thus for any  $w = (x, y) \in \Omega = \Delta_n \times \Delta_m$  we have

$$N_{\Delta_m \times \Delta_n}(x, y) = \text{span}\{\mathbf{1}_m\} \times \text{span}\{\mathbf{1}_n\} - \text{cone}[\text{co}\{e_j \mid j \in J(x, y)\}],$$

which is (1.13) in the notation of Section 1. Further, the classical subdifferential formula for max-functions (see, e.g., [18, Exercise 8.31]) gives us

$$\partial(\max_{\ell \in L} c_\ell^\top w) = \text{co}\{c_\ell \mid \ell \in L, c_\ell^\top w = \max_{\ell \in L} c_\ell^\top w\}.$$

This implies by the max-structure of the function  $F$  in (1.4) that

$$\partial F(x, y) = \text{co}\{(a_i, b_k) \mid i \in I(x), k \in K(y)\},$$

which is (1.12). Substituting finally the above calculations resulting in (1.12) and (1.13) into formula (3.27) with  $w = (x, y)$  and  $z = F(x, y)$ , we arrive at the precise computing the exact regularity bound in (1.11) and thus completes the proof of the theorem.  $\triangle$

## 4 Condition measure formula via alternative proof

In this section we give another proof of Theorem 1.3 based on convex optimization. This proof is split into three lemmas and the preceding technical claim.

Given a point  $\bar{w} \in \Omega \setminus S$  and keeping the notation above, consider the following two problems of parametric optimization (with the parameter  $z \in \mathbb{R}$ ) defined by

$$\begin{aligned} V_z(\bar{w}) &:= \min_w \|w - \bar{w}\| \\ \text{s.t. } c_\ell^\top w &\leq z \quad \forall \ell \in \mathcal{I}(\bar{w}) \\ Ew &= f \\ w_j &\geq 0 \quad \forall j \in \mathcal{J}(\bar{w}) \end{aligned} \quad (P_z) \quad \text{and} \quad \begin{aligned} \tilde{V}_z(\bar{w}) &:= \min_w \|w - \bar{w}\| \\ \text{s.t. } \max_{\ell \in L} \{c_\ell^\top w\} &= z \\ Ew &= f \\ w &\geq 0 \end{aligned} \quad (\tilde{P}_z)$$

and name  $(P_z)$  and  $(\tilde{P}_z)$  the *first* and *second parametric problem*, respectively. Observe that for every  $\bar{w} \in \Omega \setminus S$  and  $z \in \mathbb{R}_+$  the optimal value  $\tilde{V}_z(\bar{w})$  in problem  $(\tilde{P}_z)$  is equal to  $\text{dist}(\bar{w}; S(z))$ . Although the proof of the following claim is straightforward, we provide it for completeness and the reader's convenience.

**Claim 4.1 (stability of optimal solutions to first parametric problem).** *For any  $\bar{w} \in \Omega \setminus S$  and any  $\gamma > 0$  there is  $\varepsilon > 0$  (depending on  $\bar{w}$  and  $\gamma$ ) such that whenever  $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$  a unique solution  $w_z$  to problem  $(P_z)$  exists and satisfies the continuity property  $\|w_z - \bar{w}\| \leq \gamma$  with respect to the parameter  $z$ .*

**Proof.** Fix  $\gamma > 0$  and pick a  $w^S$  closest to  $\bar{w}$  in  $S$ . Let

$$\tilde{w} := \bar{w} + \tau(w^S - \bar{w}) \quad \text{with} \quad \tau := \min \left\{ \frac{\gamma}{\|w^S - \bar{w}\|}, 1 \right\} \in (0, 1].$$

Setting  $\varepsilon := F(\bar{w}) - F(\tilde{w})$ , observe by the convexity of  $F$  that

$$\varepsilon = F(\bar{w}) - F(\tilde{w}) \geq \tau F(\bar{w}) > 0.$$

We have furthermore that

$$E\tilde{w} = E\bar{w} + \min \left\{ \frac{\gamma}{\|w^S - \bar{w}\|}, 1 \right\} (Ew^S - E\bar{w}) = f,$$

$$\tilde{w}_j = (1 - \tau)\bar{w}_j + \tau w_j^S \geq 0 \quad \text{for all } j \in \mathcal{J}(\bar{w}), \quad \text{and}$$

$$\|\tilde{w} - \bar{w}\| = \tau \|w^S - \bar{w}\| \leq \gamma,$$

which imply the inclusion

$$\tilde{w} \in \Delta := \{w \mid \|w - \bar{w}\| \leq \gamma, Ew = f, w_j \geq 0 \text{ for all } j \in \mathcal{J}(\bar{w})\}.$$

Since the set  $\Delta$  is obviously convex with  $\bar{w} \in \Delta$ , we get

$$w_t := \bar{w} + t(\tilde{w} - \bar{w}) \in \Delta \quad \text{whenever } t \in [0, 1].$$

It follows from  $F(\tilde{w}) = F(\bar{w}) - \varepsilon$  and the continuity of  $F$  that for every  $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$  there is  $t_z \in [0, 1]$  such that  $w_{t_z}$  satisfies the equation

$$F(w_{t_z}) = z.$$

For any  $z$  from the above we easily get that  $w_{t_z}$  is feasible to problem  $(P_z)$ , that the set of feasible solutions to this problem is surely closed and bounded, and that the cost function is continuous with respect to  $w$ . Thus  $(P_z)$  admits an optimal solution, which is unique as a unique projection of  $\bar{w}$  on the convex feasible set. Finally,

$$V_z(\bar{w}) = \|w_z - \bar{w}\| \leq \|w_{t_z} - \bar{w}\| \leq t_z \|\tilde{w} - \bar{w}\| \leq \gamma,$$

and hence the optimal solution  $w_z$  belongs to the ball  $\bar{w} + \mathcal{B}_\gamma$ . △

The next result, whose proof is based on Claims 3.1 and 4.1, indicates the parameter region on which the optimal values in the first and second parametric problems agree.

**Lemma 4.2 (optimal values agree for both parametric problems).** *Let  $\bar{w} \in \Omega \setminus S$ . Then there exists  $\varepsilon_{\bar{w}} \in (0, F(\bar{w}))$  such that for every parameter  $z \in [F(\bar{w}) - \varepsilon_{\bar{w}}, F(\bar{w})]$  the optimal values of problems  $(P_z)$  and  $(\tilde{P}_z)$  coincide.*

**Proof.** Fix  $\bar{w} \in \Omega \setminus S$  and observe that the set of feasible solutions for  $(\tilde{P}_z)$  obviously belongs to the set of feasible solutions for  $(P_z)$ . Thus we have  $\tilde{V}_z(\bar{w}) \geq V_z(\bar{w})$  for all  $z \in \mathbb{R}$ . It remains to show that there exists  $\varepsilon_{\bar{w}} > 0$  such that  $\tilde{V}_z(\bar{w}) \leq V_z(\bar{w})$  whenever  $z \in [F(\bar{w}) - \varepsilon_{\bar{w}}, F(\bar{w})]$ .

Employing Claim 3.1, find  $\gamma > 0$  for which  $\mathcal{I}(w) \subset \mathcal{I}(\bar{w})$  and  $\mathcal{J}(w) \subset \mathcal{J}(\bar{w})$  when  $w \in \bar{w} + \mathcal{B}_\gamma$ . Further, it follows from Claim 4.1 that for such  $\gamma$  there is  $\varepsilon > 0$  with the property: whenever  $z \in [F(\bar{w}) - \varepsilon, F(\bar{w})]$  there exists a unique solution  $w_z$  to problem  $(P_z)$  satisfying  $w_z \in \bar{w} + \mathcal{B}_\gamma$ . Our choice of  $\gamma$  ensures the feasibility of  $w_z$  in problem  $(\tilde{P}_z)$ , and therefore we have the relationships

$$\tilde{V}_z(\bar{w}) \leq \|w_z - \bar{w}\| = V_z(\bar{w}),$$

which thus complete the proof of the lemma. △

**Lemma 4.3 (distances to solution sets).** *For every  $w \in \Omega \setminus S$  denote  $z_w := F(w) - \varepsilon_w$  with  $\varepsilon_w$  taken from Lemma 4.2. Then we have*

$$(4.1) \quad \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S(z_w))}{F(w) - z_w} = \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)}.$$

**Proof.** Fix  $w \in \Omega \setminus S$  and  $z \in (0, F(w))$ , and then let

$$k(w, z) := \frac{\text{dist}(w; S(z))}{F(w) - z}.$$

We first justify the inequality

$$(4.2) \quad \sup_{w \in \Omega \setminus S} k(w, z_w) \leq \sup_{w \in \Omega \setminus S} k(w, 0),$$

which gives the one in (4.1). Pick any  $\bar{w} \in \Omega \setminus S$ , let  $w^S := \arg \min_{w \in S} \|w - \bar{w}\|$  and

$$w_t := (1 - t)w^S + t\bar{w}, \quad t \in [0, 1].$$

Since  $z_{\bar{w}} \in (0, F(\bar{w}))$ , there is  $\tau \in (0, 1)$  with  $F(w_\tau) = z_{\bar{w}}$ . By the convexity of  $F$  we have

$$(4.3) \quad z_{\bar{w}} = F(w_\tau) \leq (1 - \tau)F(w^S) + \tau F(\bar{w}) = \tau F(\bar{w}).$$

Further, it follows from  $w_\tau \in S(z_{\bar{w}})$  that

$$(4.4) \quad \text{dist}(\bar{w}; S(z_{\bar{w}})) \leq \|\bar{w} - w_\tau\| = (1 - \tau)\|w^S - \bar{w}\| = (1 - \tau) \text{dist}(\bar{w}; S).$$

Combining (4.3) and (4.4) gives us

$$k(\bar{w}, z_{\bar{w}}) = \frac{\text{dist}(\bar{w}; S(z_{\bar{w}}))}{\text{dist}(F(\bar{w}); z_{\bar{w}})} \leq \frac{(1 - \tau) \text{dist}(\bar{w}; S)}{(1 - \tau)F(\bar{w})} = k(\bar{w}, 0),$$

which yields (4.2) and thus the corresponding inequality in (4.1).

It remains to show that

$$(4.5) \quad \sup_{w \in \Omega \setminus S} k(w, z_w) \geq \sup_{w \in \Omega \setminus S} k(w, 0),$$

which ensures the equality in (4.1). By Lemma 3.2 we have that  $k_F = \sup_{w \in \Omega \setminus S} k(w, 0)$  and that the maximum is attained at some  $\bar{w} \in \Omega \setminus S$ . Given  $z \in (0, F(\bar{w}))$ , let

$$w_z := \arg \min_{w \in S(z_w)} \|w - \bar{w}\|$$

and observe the estimate

$$\text{dist}(\bar{w}; S) \leq \|\bar{w} - w_z\| + \text{dist}(w_z; S),$$

implying in turn that

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \geq k(\bar{w}, z_{\bar{w}}) = \frac{\|\bar{w} - w_{z_{\bar{w}}}\|}{F(\bar{w}) - z_{\bar{w}}} \geq \frac{\text{dist}(\bar{w}; S) - \text{dist}(w_{z_{\bar{w}}}; S)}{F(\bar{w}) - z_{\bar{w}}}.$$

On the other hand, we have the equality  $\text{dist}(\bar{w}; S) = F(\bar{w})k(\bar{w}, 0)$  by the definition of  $k(w, z)$  and also the relationships

$$\text{dist}(w_{z_{\bar{w}}}; S) = F(w_{z_{\bar{w}}})k(w_{z_{\bar{w}}}, 0) \leq z_{\bar{w}}k(\bar{w}, 0)$$

due to  $F(w_{z_{\bar{w}}}) = z_{\bar{w}}$  and  $k(\bar{w}, 0) = \sup_{w \in \Omega \setminus S} k(w, 0)$ . Thus

$$\sup_{w \in \Omega \setminus S} k(w, z_w) \geq \frac{F(\bar{w})k(\bar{w}, 0) - z_{\bar{w}}k(\bar{w}, 0)}{F(\bar{w}) - z_{\bar{w}}} = k(\bar{w}, 0) = \sup_{w \in \Omega \setminus S} k(w, 0),$$

which justifies (4.5) and completes the proof of the lemma.  $\triangle$

The last lemma establishes, by employing Lagrangian duality, a precise formula for computing the optimal value of the cost function in the parametric problem  $(P_z)$ —and hence in  $(\tilde{P}_z)$ —via the initial data.

**Lemma 4.4 (computing optimal values of parametric problems).** *Let  $\bar{w} \in \Omega \setminus S$  and  $z \in (0, F(\bar{w}))$ . Then the optimal value  $V_z(\bar{w})$  of problem  $(P_z)$  is computed by*

$$V_z(\bar{w}) = \frac{F(\bar{w}) - z}{\text{dist}(0; \text{co}\{c_\ell, i \in \mathcal{I}(w)\} + (\ker E)^\perp - \text{cone}[\text{co}\{e_j, j \in \mathcal{J}(w)\}])}.$$

**Proof.** Observe that problem  $(P_z)$  can be reformulated as

$$V_z(\bar{w}) = \inf_w \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[ u^T(w - \bar{w}) + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (a_i^T w - z) + v^T(Ew - f) - \sum_{j \in \mathcal{J}(\bar{w})} \mu_j w_j \right].$$

Since the convex optimization problem  $(P_z)$  satisfies the Slater condition, we can interchange the supremum and the infimum above by Lagrangian duality. This gives

$$V_z(\bar{w}) = \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \inf_w \left[ u^T(w - \bar{w}) + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (a_i^T w - z) + v^T(Ew - f) - \sum_{j \in \mathcal{J}(\bar{w})} \mu_j w_j \right].$$

Regrouping the terms inside the square brackets, we obtain

$$V_z(\bar{w}) = \sup_{\substack{\|u\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ v \in \mathbb{R}^m, \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \inf_w \left[ \left( u + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu \right)^T w - u^T \bar{w} - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i z - v^T f \right].$$

Observe further that, whenever the term  $(u + \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu)$  is not zero, the inner infimum in  $w$  necessarily becomes  $-\infty$ . This allows us to put

$$u = - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i - E^T v + \mu$$

and consequently rewrite the expression for  $V_z(\bar{w})$  as follows:

$$V_z(\bar{w}) = \sup_{\substack{\| \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu \| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[ \left( \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu \right)^T \bar{w} - \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i z - v^T f \right].$$

Regrouping again gives us the formula

$$V_z(\bar{w}) = \sup_{\substack{\|\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \left[ \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i (a_i^T \bar{w} - z) + v^T (E\bar{w} - f) - \mu^T \bar{w} \right].$$

Noting that  $E\bar{w} = f$ ,  $\mu^T \bar{w} = 0$  and  $a_i^T \bar{w} = F(\bar{w})$  for  $i \in \mathcal{I}(\bar{w})$ , we have

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\|\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i + E^T v - \mu\| \leq 1, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \sum_{i \in \mathcal{I}(\bar{w})} \lambda_i.$$

Since  $V_z(\bar{w}) \neq 0$ , the latter yields

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\left\| \frac{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} + E^T \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} - \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} \right\| \leq \frac{1}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i}, \\ \lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{1/\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i},$$

which can be written as

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\lambda_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \mu_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \mu_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{\left\| \frac{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i a_i}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} + E^T \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} - \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})} \lambda_i} \right\|}.$$

Changing further the variables by

$$\tilde{\lambda} := \frac{\lambda}{\sum_{i \in \mathcal{I}(\bar{w})}}, \quad \tilde{\mu} := \frac{\mu}{\sum_{i \in \mathcal{I}(\bar{w})}}, \quad \tilde{v} := \frac{v}{\sum_{i \in \mathcal{I}(\bar{w})}},$$

we arrive at the expression

$$V_z(\bar{w}) = (F(\bar{w}) - z) \sup_{\substack{\tilde{\lambda}_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i = 1, \\ \tilde{\mu}_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \tilde{\mu}_j = 0, j \in \mathcal{J} \setminus \mathcal{J}(\bar{w})}} \frac{1}{\left\| \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i a_i + E^T \tilde{v} - \tilde{\mu} \right\|},$$

which is equivalently written as

$$V_z(\bar{w}) = \frac{F(\bar{w}) - z}{\inf \left\{ \begin{array}{l} \tilde{\lambda}_i \geq 0, i \in \mathcal{I}(\bar{w}), \\ \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i = 1, \\ \tilde{\mu}_j \geq 0, j \in \mathcal{J}(\bar{w}), \\ \tilde{\mu}_j = 0, j \in J \setminus \mathcal{J}(\bar{w}) \end{array} \right\} \left\| \sum_{i \in \mathcal{I}(\bar{w})} \tilde{\lambda}_i a_i + E^T \tilde{v} - \tilde{\mu} \right\|}.$$

Recalling the notation of Section 1 allows us to reduce the latter expression to the one in the lemma formulation and thus finish the proof.  $\triangle$

Combining the obtained lemmas with the definitions above, we can now complete the alternative proof of the condition measure formula in Theorem 1.3.

**Proof of Theorem 1.3.** For every  $w \in \Omega \setminus S$  choose the parameter  $z_w$  as in Lemma 4.3, i.e., put  $z_w = F(w) - \varepsilon_w$ , where  $\varepsilon_w$  is taken from Lemma 4.2. Then we have

$$\begin{aligned} \kappa_F &= \inf \{ k \geq 0 \mid \text{dist}(w; S) \leq kF(w) \text{ for all } w \in \Omega \setminus S \} \quad (\text{by definition (3.1)}) \\ &= \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S)}{F(w)} \quad (\text{as } \Omega \setminus S \neq \emptyset) \\ &= \sup_{w \in \Omega \setminus S} \frac{\text{dist}(w; S(z_w))}{F(w) - z_w} \quad (\text{by Lemma 4.3}) \\ &= \sup_{w \in \Omega \setminus S} \frac{\tilde{V}_{z_w}(w)}{F(w) - z_w} \quad (\text{by the definition of } (P'_z)) \\ &= \sup_{w \in \Omega \setminus S} \frac{V_{z_w}(w)}{F(w) - z_w} \quad (\text{by Lemma 4.2 and the choice of } z_w) \\ &= \sup_{w \in \Omega \setminus S} \frac{1}{\text{dist}(0; \text{co}_{i \in \mathcal{I}(w)} \{c_\ell\} + (\ker E)^\perp - \text{cone}[\text{co}_{j \in \mathcal{J}(w)} \{e_j\}])} \quad (\text{by Lemma 4.4}). \end{aligned}$$

Letting  $w = (x, y) \in \Omega$ , observe finally that

$$\{c_\ell \mid \ell \in \mathcal{I}(w)\} = \{(a_i, b_k) \mid i \in \mathcal{I}(x), k \in K(y)\}, \quad \ker E = \text{span}\{1_n\} \times \text{span}\{1_m\},$$

and  $\kappa(A) = \kappa_F$ , which complete the proof of the theorem.  $\triangle$

**Acknowledgments.** The authors gratefully acknowledge helpful discussions with Arkadi Nemirovski and Yurii Nesterov on the topics and results of this paper.

## References

- [1] F. J. ARAGÓN ARTACHO, A. L. DONTCHEV AND M. H. GEOFFROY, *Convergence of the proximal point method for metrically regular mappings*, ESAIM Proc., 17 (2007), pp. 1–8.
- [2] H. ATTOUCH, J. BOLTE, P. REDONT AND A. SOUBEYRAN, *Alternating minimization and projection methods for nonconvex problems, An approach based on the Kurdyka-Lojasiewicz inequality*, Math. Oper. Res., to appear.

- [3] A. L. DONTCHEV, A. D. LEWIS AND R. T. ROCKAFELLAR, *The radius of metric regularity*. Trans. Amer. Math. Soc., 355 (2003), pp. 493–517.
- [4] A. GILPIN, J. PEÑÁ AND T. SANDHOLM, *First-order algorithm with  $O(\ln(1/\varepsilon))$  convergence for  $\varepsilon$ -equilibrium in two-person zero-sum games*, in Proc. 23rd Nat. Conf. Art. Intel. (AAAI), 2008, pp. 75–82.
- [5] N. HIGHAM *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, PA, 1996.
- [6] D. KLATTE AND B. KUMMER, *Optimization methods and stability of inclusions in Banach spaces*, Math. Program., 117 (2009), pp. 305–330.
- [7] A. S. LEWIS, D. R. LUKE AND J. MALICK, *Local linear convergence for alternating and averaged nonconvex projections*, Found. Comp. Math., 9 (2009), pp. 485–513.
- [8] B. S. MORDUKHOVICH, *Maximum principle in problems of time optimal control with nonsmooth constraints*, Appl. Math. Mech., 40 (1976), 960–969.
- [9] B. S. MORDUKHOVICH, *Complete characterizations of openness, metric regularity, and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 340 (1993), pp. 1–35.
- [10] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 330, Springer, Berlin, 2006.
- [11] YU. NESTEROV, *Excessive gap technique in nonsmooth convex minimization*, SIAM J. Optim., 16 (2005), pp. 235–249.
- [12] YU. NESTEROV, *Smooth minimization of nonsmooth functions*, Math. Program., 103 (2005), pp. 127–152.
- [13] J. PEÑA AND J. RENEGAR, *Computing approximate solutions for conic systems of constraints*, Math. Program., 87 (2000), pp. 351–383.
- [14] J. RENEGAR, *Incorporating condition measures into the complexity theory of linear programming*, SIAM J. Optim., 5 (1995), pp. 506–524.
- [15] J. RENEGAR, *Linear programming, complexity theory and elementary functional analysis*, Math. Program., 70 (1995), pp. 279–351.
- [16] S. M. ROBINSON, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.
- [17] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [18] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational Analysis*, Grundlehren Series (Fundamental Principles of Mathematical Sciences), Vol. 317, Springer, Berlin, 1998.

- [19] M. SHUB AND S. SMALE, *Complexity of Bézout's theorem, I: Geometric aspects*, J. Amer. Math. Soc., 6 (1993), pp. 459–501.
- [20] B. VON STENDEL, *Equilibrium computation for two-player games in strategic and extensive form*, in *Algorithmic Game Theory*, N. Nisan et al., eds, Cambridge University Press, Cambridge, UK, 2007, pp. 53–78.