Doctoral Program in Statistics, Optimization and Applied Mathematics

# Semilocal Lipschitz Stability in Linear Optimization

**Jesús Camacho Moro**

DOCTORAL DISSERTATION

Director

**Dr. Juan Parra López**

Codirector

**Dr. María Josefa Cánovas Cánovas**

Universidad Miguel Hernández de Elche

2023

This Doctoral Thesis, entitled **"Semilocal Lipschitz Stability in Linear Optimization"**, is presented under the modality of **thesis by compendium** of the following **publications**:

A **From calmness to Hoffman constants for linear semi-infinite inequality systems.** J. Camacho, M.J. Cánovas & J. Parra (2022). SIAM Journal on Optimization, 32(4), 2859-2878. https://doi.org/10.1137/21M1418228

B **Lipschitz upper semicontinuity in linear optimization via local directional convexity.** J. Camacho, M.J. Cánovas & J. Parra (2022). Optimization. https://doi.org/10.1080/02331934.2022.2057851

C **Robust and continuous metric subregularity for linear inequality systems.** J. Camacho, M.J. Cánovas, M.A. López & J. Parra (2022). Computational Optimization and Applications. https://doi.org/10.1007/s10589-022-00437-0

D **Feasibility problems via paramonotone operators in a convex setting.** J. Camacho, M.J. Cánovas, J.E. Martínez-Legaz & J. Parra. https://doi.org/10.48550/arXiv.2307.00979

E **Hoffman constant of the argmin mapping in linear optimization.** J. Camacho, M.J. Cánovas, H. Gfrerer & J. Parra. https://doi.org/10.48550/arXiv.2307.01034

El Dr. D. Juan Parra López, director y la Dra. Dña. María Josefa Cánovas Cánovas, codirectora de la tesis doctoral titulada **“Semilocal Lipschitz Stability in Linear Optimization”**

**INFORMAN:**

Que D. Jesús Camacho Moro ha realizado bajo nuestra supervisión el trabajo titulado **“Semilocal Lipschitz Stability in Linear Optimization”** conforme a los términos y condiciones definidos en su Plan de Investigación y de acuerdo al Código de Buenas Prácticas de la Universidad Miguel Hernández de Elche, cumpliendo los objetivos previstos de forma satisfactoria para su defensa pública como tesis doctoral.

Lo que firmamos para los efectos oportunos, en Elche, julio de 2023

| | |
|---|---|
| Director de la tesis | Codirectora de la tesis |
| Dr. D. Juan Parra López | Dra. Dña. María Josefa Cánovas Cánovas |

El Dr. D. Domingo Morales González, Coordinador del Programa de Doctorado en **Estadística, Optimización y Matemática Aplicada,**

**INFORMA**:

Que D. Jesús Camacho Moro ha realizado bajo la supervisión de nuestro Programa de Doctorado el trabajo titulado **"Semilocal Lipschitz Stability in Linear Optimization"** conforme a los términos y condiciones definidos en su Plan de Investigación y de acuerdo al Código de Buenas Prácticas de la Universidad Miguel Hernández de Elche, cumpliendo los objetivos previstos de forma satisfactoria para su defensa pública como tesis doctoral.

Lo que firmo para los efectos oportunos, en Elche, julio de 2023

Prof. Dr. D. Domingo Morales González

Coordinador del Programa de Doctorado en Estadística, Optimización y Matemática Aplicada

This doctoral thesis has been financially supported by:

- Grant PGC2018-097960-B-C21 from MICINN, Spain, and ERDF, "A way to make Europe", European Union.
- Predoctoral Grant PRE2019-090676.
- Grant PROMETEO/2021/063 from Generalitat Valenciana, Spain.

# Agradecimientos

A mis directores por acompañarme en esta travesía por el desierto que es la tesis. Me han acogido como a un hijo adoptivo desde que llegué y eso es algo que no olvidaré nunca. Con ellos he aprendido la variedad de este oficio: investigador, profesor, presentador, y organizador de eventos, todo en uno.

A mis colaboradores por enseñarme otras formas de ver las matemáticas, además de recordarme el aspecto humano de este mundo. Me llevo grandes anécdotas y pasiones compartidas como el cine o las salidas en bici.

A mis compañeros del CIO por esas charlas infinitas en el comedor. Con vosotros se hacen más amenos esos días en los que no te sale nada.

A mi familia por haberme convertido en quien soy hoy día. Sin la constancia que me inculcaron esta tesis no sería posible. También por haberme animado desde la distancia durante estos años de trabajo. Se os ha echado de menos.

A mi novia, que ha recorrido casi todo este camino junto a mi, por hacerme más feliz en cada paso. Ella es mi recordatorio de que la vida no son sólo matemáticas.

A mi compañero de casa, despacho y planes, por tirar siempre de mí. Me llevo de aquí a mucho más que un amigo.

Por último, aunque sea poco ortodoxo, quiero darme las gracias a mi mismo por no abandonar el sueño que tuve desde pequeño. Y sobre todo, por no abandonarme cuando más lo necesitaba.

# Contents

# 1 Summary

## 1.1 English

This thesis has its core in variational analysis, a vast and modern mathematical field closely related to optimization. From its origin in the second half of the 20th century a fast growing literature has arisen; overviews on the topic can be found in the monographs [15, 20, 26, 29], among others. Over the years, hand by hand with the development of computers, the theory of variational analysis and optimization has become more and more important, especially because of its numerous applications into our society, finding examples of applications in logistics, location problems, project selection, resource managing, economy, etc. In particular, the current dissertation is mainly focused on the quantitative stability of feasible and optimal solutions to linear optimization problems, whose theoretical analysis comes from the early 1950s (see, e.g., [16, 18, 30]). Concerning theory, methods and applications of linear optimization problems with arbitrarily many constraints, the reader is addressed to the monograph [17].

As pointed out at the very beginning, the thesis is a compendium of three published papers, gathered in Appendices A, B, C, and two preprints, in Appendices D and E, presented in chronollogical order of production. These papers also appear at the begining of the bibliography chapter with the corresponding capital letters instead of numbers, so that the original contributions of the thesis can be easily identified. For instance, Theorem 4 of Appendix A will be cited as [A, Theorem 4]. Clearly, each appendix has its own reference list, and in the bibliography chapter of the thesis, before the appendices, we only include those references which are cited in Chapters 1 to 5 of the thesis.

In general terms, the main original contributions of the thesis (gathered in [A-E]) are concerned with linear optimization problems in different parametric settings and their associated feasible and optimal set (argmin) mappings. Some results have been established for general multifunctions between metric (or more structured) spaces. For pedagogical reasons, given such a multifunction $\mathcal{M} : Y \rightrightarrows X$, we shall refer to elements in $Y$ as parameters and to elements in $X$ as points or solutions, since this is the role played by these elements in our framework. In this way, $\mathcal{M}(y)$ can be thought as the (feasible or optimal) solution set associated with parameter $y \in Y$. Each Lipschitz-type property can be quantified by means of its associated sharp Lipschitz constant, which will be referred to as modulus. Roughly speaking, the thesis deals with variation rates of solutions with respect to parameters. Some of these rates (as the calmness modulus) are *local*, in the sense that we consider parameters in a neighborhood of a nominal (given) one, $\overline{y}$, and solutions around a nominal $\overline{x} \in X$. Some other rates (as the Lipschitz upper semicontinuity modulus) are *semilocal*, as they combine small perturbations of the parameter with the variation of the whole solution set. Finally, *global* rates (as the Hoffman constant) are concerned with all parameters and their whole solution sets.

The immediate antecedents which constitute the starting point of the thesis are papers [11], and [10] devoted to the calmness moduli of the feasible and the argmin mappings, respectively. Point-based (only involving nominal parameters and solutions) formulae for these moduli are provided there. One of the goals achieved in the thesis is to determine point-based formulae for semilocal and global moduli, specifically as the maximum of finitely many calmness moduli.

More in detail, Appendix A is concerned with all these properties for a generic multifunction, showing a particularly good behavior when the graph of such a multifunction is closed and convex. This is the case of the feasible set mapping of linear inequality systems parameterized by its right-hand side (RHS in brief). This section introduces two new properties aside the mentioned above: the uniform calmness (semilocal) [A, Definition 1] and the Hoffman stability at $\overline{y}$ (between semilocal and global) [A, Formula (5)]. As main results, we draw the reader's attention to [A, Theorems 4,5,6], which will be discussed in Chapter 4.

The fact that the graph of the optimal set mapping is no longer convex, even when the objective function remains unperturbed, provokes notable differences in the methodological approach for computing moduli. Such a computation is carried out in Appendices B and E. Appendix B is mainly concerned with the Lipschitz upper semicontinuity of the optimal set mapping under canonical perturbations (RHS perturbations of the constrains and tilt perturbations of the objective). Moving from local to semilocal measures, [B] introduces a new concept named local directional convexity [B, Section 3]. Specifically, the graph of the optimal set mapping is convex when restricted to small perturbations of the RHS along a fixed direction [B, Theorem 3.1]. Thanks to this geometrical property, we are able to compute the Lipschitz upper semicontinuity modulus of the argmin mapping as a maximum of some specific calmness moduli [B, Theorem 4.2]. On the other hand, to jump from semilocal to global measures, Appendix E develops in [E, Section 3] a sort of finite piecewise linear procedure in order to compute the global Hoffman constant of the argmin mapping under RHS perturbations in [E, Theorem 5]. To this aim, the concept of well-connected polyhedral mapping is introduced in [E, Definition 2]; outstanding steps in this procedure are given in [E, Definition 3, Theorem 4 and Corollary 1]. For the sake of completeness let us mention that the global Hoffman constant under canonical perturbations is always infinite except in the trivial case when all left-hand side coefficients of the constraints are zero [E, Proposition 2].

Up to now we have dealt with different types of variation rates when the left-hand side of the constraints remains unperturbed. Appendix C presents an approach to the stability of the feasible set mapping under left-hand side perturbations from the broader paradigm of the so-called radius theorem, which has been widely studied in different frameworks (see, e.g. [8, 13, 14, 23]). Roughly speaking, the radius of a certain property for some multifunction is the smallest perturbation of this multifunction causing failure of such a property. In the light of the antecedents of the thesis, the natural property to consider would be the metric subregularity of the inverse feasible set mapping, since this property is known to be equivalent to the calmness of the feasible set mapping. However, this property holds for free for finite linear inequality systems, so that the radius turns out to be infinite, as pointed out in [13]. Therefore, it makes sense to consider desirable stronger properties whose fulfillment is not guaranteed. Appendix C introduces two of such properties, called robust and continuous subregularity in [C, Definition 1]. The main results to this respect are [C, Theorems 5 and 6 and Corollaries 2 and 3]. Previously, [C, Theorem 3] provides some technical results on the stability of the so-called end set of a convex polyhedron. As a consequence, the continuity of the subregularity modulus is characterized in [C, Theorem 4]. The radius of robust subregularity is computed through a point-based formula in [C, Theorem 6], whereas determining the radius of continuous subregularity remains as an open problem.

Finally, the thesis makes an incursion in monotone operator theory with the aim of incorporating this stuff to provide a new approach to feasibility problems. The close link between monotone operators and convex analysis is exhibited by the well-known fact that the subdifferential of a proper lower semicontinuous convex function is maximally monotone. In Apendix D we focus in operators that are simultaneously paramotone and bimonotone, which are shown to be constant on their domains [D, Corollary 9]. This fact is applied in two particular situations. The first one looks for the smallest perturbation (in the sense of translations) over a finite amount of convex sets in order to reach a nonempty intersection. This problem is directly related with simultaneous projections, solved in [D, Propositions 16 and 18 and Theorem 19] for the case of two closed convex sets in a Hilbert space, and it is extended to finitely many sets under some differentiability assumptions in [D, Theorem 21]. The second application deals with the distance to feasibility; more in detail, given an inconsistent convex inequality system, we derive lower and upper estimates for the smallest RHS perturbation that produces a feasible system [D, Proposition 24]. These estimates coincide in the linear case [D, Corollary 28], and an operative procedure to determine such a distance is provided in [D, Theorem 29].

## 1.2 Spanish

Esta tesis tiene su núcleo en el análisis variacional, un vasto y moderno campo matemático estrechamente relacionado con la optimización. Desde su origen en la segunda mitad del siglo XX ha surgido una literatura en rápido crecimiento; se pueden encontrar visiones generales sobre el tema en las monografías [15, 20, 26, 29], entre otras. A lo largo de los años, de la mano del desarrollo de los ordenadores, la teoría del análisis variacional y de la optimización se ha vuelto cada vez más importante, especialmente por sus numerosas aplicaciones en nuestra sociedad, encontrando ejemplos de aplicación en logística, problemas de localización, selección de proyectos, gestión de recursos, economía, etc. En particular, la presente disertación se centra principalmente en la estabilidad cuantitativa de soluciones factibles y óptimas de problemas de optimización lineal, cuyo análisis teórico se remonta a principios de los años cincuenta (véanse, por ejemplo, [16, 18, 30]). En cuanto a la teoría, los métodos y aplicaciones de problemas de optimización lineal con un número arbitrario de restricciones, el lector puede consultar la monografía [17].

Como se señala nada más comenzar, la tesis es un compendio de tres artículos publicados, recogidos en los Apéndices A, B y C, y de dos "preprints", en los Apéndices D y E, presentados por orden cronológico de producción. Estos trabajos también aparecen al principio del capítulo de bibliografía con las correspondientes mayúsculas en lugar de números para que las contribuciones originales de la tesis puedan identificarse fácilmente. Por ejemplo, el Teorema 4 del Apéndice A se citará como [A, Teorema 4]. Claramente, cada apéndice tiene su propia lista de referencias, y en el capítulo de bibliografía de la tesis, antes de los apéndices, sólo incluimos las referencias citadas en los capítulos 1 a 5 de la tesis.

En términos generales, las principales contribuciones originales de la tesis (recogidas en [A-E]) versan sobre problemas de optimización lineal en diferentes configuraciones paramétricas y sus multifunciones conjunto factible y conjunto óptimo (argmin) asociadas. Se han establecido algunos resultados para multifunciones generales entre espacios métricos (o más estructurados). Por razones pedagógicas, dada una multifunción $\mathcal{M} : Y \rightrightarrows X$, nos referiremos a los elementos de $Y$ como parámetros y a los elementos de $X$ como puntos o soluciones, ya que este es el papel que desempeñan estos elementos en nuestro marco de trabajo. De este modo, $\mathcal{M}(y)$ puede considerarse como el conjunto de soluciones (factibles u óptimas) asociado al parámetro $y \in Y$. Cada propiedad de tipo Lipschitz puede cuantificarse mediante su constante de Lipschitz ajustada asociada, que se denominará módulo. A grandes rasgos, la tesis trata de las tasas de variación de las soluciones con respecto a los parámetros. Algunas de estas tasas (como el módulo de calmness -no suele traducirse el término-) son *locales*, en el sentido de que que consideramos parámetros entorno a uno nominal (dado), $\overline{y}$, y soluciones alrededor de un $\overline{x} \in X$ nominal. Algunas otras ratios (como el módulo de Lipschitz upper semicontinuity) son *semilocales*, ya que combinan pequeñas perturbaciones del parámetro con la variación de todo el conjunto de soluciones. Por último, las tasas *globales* (como la constante de Hoffman) se refieren a todos los parámetros y a todo su conjunto de soluciones.

Los antecedentes inmediatos que constituyen el punto de partida de la tesis son los trabajos [11] y [10], dedicados a los módulos de calmness de las multifunciones conjunto factible y argmin, respectivamente. En estos trabajos se proporcionan fórmulas de las que llamamos "point-based" (que sólo implican parámetros y soluciones nominales) para estos módulos. Uno de los objetivos alcanzados en la tesis es determinar fórmulas de este tipo para los módulos semilocales y globales, concretamente como el máximo de una cantidad finita de módulos de calmness.

Más en detalle, el Apéndice A se ocupa de todas estas propiedades para una multifunción genérica, mostrando un comportamiento particularmente bueno cuando el grafo de dicha multifunción es cerrado y convexo. Este es el caso de la multifunción conjunto factible para sistemas de desigualdades lineales parametrizados por su lado derecho (RHS en sus siglas en inglés). En este apartado se introducen dos nuevas propiedades aparte de las mencionadas anteriormente: uniform calmness (semilocal) [A, Definición 1] y la estabilidad de Hoffman en $\overline{y}$ (entre semilocal y global) [A, Fórmula (5)]. Como principales resultados, llamamos la atención del lector sobre [A, Teoremas 4, 5 y 6], que se discutirán en el Capítulo 4.

El hecho de que el grafo de la multifunción conjunto óptimo deje de ser convexo, incluso cuando la función objetivo permanece inalterada, provoca notables diferencias en el enfoque metodológico para calcular los módulos. Dicho cálculo se lleva a cabo en los Apéndices B y E. El Apéndice B se ocupa principalmente de la propiedad de Lipschitz upper semicontinuity de la multifunción conjunto óptimo bajo perturbaciones canónicas (perturbaciones RHS de las restricciones y perturbaciones de inclinación del objetivo). Pasando de medidas locales a semilocales, [B] introduce un nuevo concepto denominado convexidad direccional local [B, Sección 3]. Específicamente, el grafo de la multifunción conjunto óptimo es convexo cuando se restringe a pequeñas perturbaciones RHS a lo largo de una dirección fija [B, Teorema 3.1]. Gracias a esta propiedad geométrica podemos calcular el módulo de Lipschitz upper semicontinuity de la multifunción argmin como un máximo de algunos módulos de calmness específicos [B, Teorema 4.2]. Por otra parte, para saltar de medidas semilocales a globales, el Apéndice E desarrolla en [E, Sección 3] una especie de procedimiento finito lineal a trozos para para calcular la constante global de Hoffman de la multifunción argmin bajo perturbaciones RHS en [E, Teorema 5]. Con este objetivo, el concepto de well-connected polyhedral mapping se introduce en [E, Definición 2]; pasos destacados en este procedimiento se dan en [E, Definición 3, Teorema 4 y Corolario 1]. En aras de la exhaustividad mencionemos que la constante global de Hoffman bajo perturbaciones canónicas es siempre infinita excepto en el caso trivial en el que todos los coeficientes del lado izquierdo de las restricciones son cero [E, Proposición 2].

Hasta ahora hemos tratado diferentes tipos de tasas de variación cuando el lado izquierdo de las restricciones permanece inalterado. En el Apéndice C se presenta una aproximación a la estabilidad del conjunto factible bajo perturbaciones del lado izquierdo desde el paradigma más amplio del llamado teorema del radio, que ha sido ampliamente estudiado en diferentes marcos (véase, por ejemplo. [8, 13, 14, 23]). En términos generales, el radio de una cierta propiedad para alguna multifunción es la perturbación más pequeña de esta multifunción que causa el fallo de dicha propiedad. A la luz de los antecedentes de la tesis, la propiedad natural a considerar sería la subregularidad métrica de la inversa de la multifunción conjunto factible, ya que esta propiedad se sabe que es equivalente a la calmness de la multifunción conjunto factible. Sin embargo, esta propiedad se cumple siempre para sistemas de desigualdades lineales finitos, por lo que el radio resulta ser infinito, como se señala en [13]. Por lo tanto, tiene sentido considerar propiedades deseables más fuertes cuyo cumplimiento no esté garantizado. En el Apéndice C se presentan dos de estas propiedades, denominadas subregularidad robusta y subregularidad continua en [C, Definición 1]. Los principales resultados a este respecto son [C, Teoremas 5 y 6 y Corolarios 2 y 3]. Anteriormente, [C, Teorema 3] proporciona algunos resultados técnicos sobre la estabilidad del llamado end set de un poliedro convexo. Como consecuencia, la continuidad del módulo de subregularidad se caracteriza en [C, Teorema 4]. El radio de subre-gularidad robusta se calcula a través de una fórmula "point-based" en [C, Teorema 6], mientras que determinar el radio de subregularidad continua sigue siendo un problema abierto.

Por último, la tesis hace una incursión en la teoría de operadores monótonos con el objetivo de incorporar estas herramientas para proporcionar un nuevo enfoque a los problemas de factibilidad. El estrecho vínculo entre los operadores monótonos y el análisis convexo queda de manifiesto en el hecho conocido de que el subdiferencial de una función convexa, semicontinua inferiormente y propia es maximalmente monótono. En el apéndice D nos centramos en operadores que son simultáneamente paramonotonos y bimonotonos, los cuales resultan ser constantes en sus dominios [D, Corolario 9]. Este hecho se aplica en dos situaciones particulares. En la primera se busca la menor perturbación (en el sentido de traslaciones) sobre una cantidad finita de conjuntos convexos con el fin de alcanzar una intersección no vacía. Este problema está directamente relacionado con las proyecciones simultáneas, resuelto en [D, Proposiciones 16 y 18 y Teorema 19] para el caso de dos conjuntos convexos cerrados en un espacio de Hilbert, y se extiende a una cantidad finita de conjuntos bajo algunos supuestos de diferenciabilidad en [D, Teorema 21]. La segunda aplicación trata de la distancia a la factibilidad; más en detalle, dado un sistema de desigualdades convexo inconsistente, obtenemos acotaciones inferiores y superiores para la menor perturbación RHS que produce un sistema consistente [D, Proposición 24]. Estas estimaciones coinciden en el caso lineal [D, Corolario 28], y se proporciona un procedimiento operativo para determinar tal distancia en [D, Teorema 29].

# 2 | Methodology

The usual methodology for a doctoral thesis in mathematics has been followed. The very first step is to introduce the doctoral candidate in the field through some selected literature. Once we have an overview of the topic, we search for open problems, either well-known ones or interesting new ones, and set our objectives. To this end, an in-depth study of the state of the art is needed, including the possibility of discussing directly with relevant researchers in the specific area under consideration. Next, we conjecture some results, generally based on patterns observed in academic examples. Finally, we provide formal proofs for those statements that turn out to be right and counterexamples otherwise. Once the contents are considered of enough quality, they are submitted to publication and shared with the scientific community in different conferences and workshops.

# 3 Introduction

This chapter is devoted to introduce our model and the necessary notation and definitions to enable the reader to obtain, in the next chapter, a more detailed perspective of the thesis and its main original contributions than that sketched in the summary. In addition, we include just a basic selection on background results and references. In order to avoid unnecessary redundancies, for more detailed information the reader is addressed to the publications gathered in Appendices A-E.

Most of the thesis is concerned with the parameterized linear optimization problem

$$\begin{aligned}\pi : \text{minimize } & c'x \\ \text{subject to } & a_t'x \leq b_t, \quad t \in T := \{1, 2, \ldots, m\},\end{aligned}$$

where $x \in \mathbb{R}^n$ is the decision variable, regarded as a column vector, and the prime denotes transposition, so that $c'x$ is the usual inner product of $c$ and $x$ in $\mathbb{R}^n$. In all the thesis except Appendix C, the left-hand side coefficient function $a : t \mapsto a_t \in \mathbb{R}^n$ is fixed. The most common parametric setting is that of the so-called *canonical perturbations,* in which $c \in \mathbb{R}^n$ and $b : t \mapsto b_t \in \mathbb{R}$ are considered as parameters. Accordingly, we identify problem $\pi$ with the pair $(c, b) \in \mathbb{R}^n \times \mathbb{R}^m$. Regarding the topology, the space of variables, $\mathbb{R}^n$, is endowed with an arbitrary norm $\|\cdot\|$ unless otherwise specified (for instance, the end of Appendix D deals with the Euclidean norm). The corresponding dual norm is given by $\|u\|_* := \max_{\|x\|\leq 1} |u'x|$. The parameter space, $\mathbb{R}^m$, of RHS perturbations of the constraint system is endowed with the supremum norm $\|b\|_\infty := \max_{t\in\{1,2,\ldots,m\}} |b_t|$ and in the case of canonical perturbations we set

$$\|(c,b)\| := \max\left\{\|c\|_*, \|b\|_\infty\right\},$$

since $c$ is viewed as a linear functional.

In this framework, we consider the *feasible set mapping*, $\mathcal{F} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, given by

$$\mathcal{F}(b) := \{x \in \mathbb{R}^n \mid a_t' x \leq b_t,\ t \in T\}.$$

Part of Appendix A deals with linear semi-infinite inequality systems (possibly infinitely many constraints), where, specifically, $T$ is a compact Hausdorff space. The reader is addressed to [17, Chapter 6] for the stability theory of linear semi-infinite inequality systems in the more general case when $T$ is arbitrary, with no topological structure. In the case when function $a \equiv (a_t)_{t\in T}$ is also perturbed (in Appendix C) we represent by $\sigma \equiv (a,b)$ the constraint system $\{a_t' x \leq b_t,\ t \in T\}$.

The *optimal set mapping*, also called *argmin mapping*, $\mathcal{F}^{op} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, is given by

$$\mathcal{F}^{op}(c,b) := \arg\min\{c'x : x \in \mathcal{F}(b)\},\ \ (c,b) \in \mathbb{R}^n \times \mathbb{R}^m.$$

In the case when $c$ remains fixed at its nominal value $\overline{c}$, we are dealing with $\mathcal{F}^{op}_{\overline{c}} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, given by

$$\mathcal{F}^{op}_{\overline{c}}(b) := \mathcal{F}^{op}(\overline{c}, b),\ \ b \in \mathbb{R}^m.$$

A significant part of the results affecting $\mathcal{F}$, $\mathcal{F}^{op}$ or $\mathcal{F}^{op}_{\overline{c}}$ and most of definitions are given for a generic multifunction $\mathcal{M} : Y \rightrightarrows X$ between metric spaces, with both distances being denoted by $d$. In this case, as pointed out in the summary, elements $y \in Y$ are called parameters and elements $x \in X$ will be referred to as points or solutions.

Next, we introduce the main (upper) Lipschitz type properties dealt with in the thesis. Mapping $\mathcal{M} : Y \rightrightarrows X$ is said to be:

- *calm* at $(\overline{y},\overline{x}) \in \operatorname{gph}\mathcal{M}$ (the graph of $\mathcal{M}$) if there exist a constant $\kappa \geq 0$ along with a neighborhood of $(\overline{y},\overline{x})$, $V \times U$, such that

  $$d(x,\mathcal{M}(\overline{y})) \leq \kappa d\left(y,\overline{y}\right) \text{ for all } x \in \mathcal{M}(y) \cap U \text{ and all } y \in V.$$

  This property is known to be equivalent to the *metric subregularity* of $\mathcal{M}^{-1}$ (the inverse of $\mathcal{M}$) at $(\overline{x},\overline{y})$ which reads as the existence of $\kappa \geq 0$ and a (possibly smaller) neighborhood $U$ of $\overline{x}$ such that

  $$d(x,\mathcal{M}(\overline{y})) \leq \kappa d\left(\overline{y},\mathcal{M}^{-1}(x)\right) \text{ for all } x \in U\ .$$

- *Lipschitz upper semicontinuous* at $\overline{y} \in \operatorname{dom}\mathcal{M}$ (the domain of $\mathcal{M}$) if there exists a neighborhood $V$ of $\overline{y}$ and a constant $\kappa \geq 0$ such that

  $$d(x,\mathcal{M}(\overline{y})) \leq \kappa d\left(y,\overline{y}\right) \text{ for all } y \in V \text{ and all } x \in \mathcal{M}(y). \tag{3.1}$$

- *Hoffman stable* at $\overline{y} \in \operatorname{dom}\mathcal{M}$ if (3.1) holds for all $(y,x) \in \operatorname{gph}\mathcal{M}$.

- *uniformly calm* at $\overline{y} \in \operatorname{dom}\mathcal{M}$ (introduced in [A, Definition 1]) if there exist a neighborhood $V$ of $\overline{y}$ along with $\varepsilon > 0$ and $\kappa \geq 0$ such that

  $$d(x,\mathcal{M}(\overline{y})) \leq \kappa d\left(y,\overline{y}\right) \text{ for all } y \in V \text{ and all } x \in \mathcal{M}_{\varepsilon}(y),$$

  where

  $$\mathcal{M}_{\varepsilon}(y) := \mathcal{M}(y) \cap B\left(\mathcal{M}\left(\overline{y}\right),\varepsilon\right) \text{ for } y \in Y,$$

  or, equivalently, if $\mathcal{M}_{\varepsilon}$ is Lipschitz upper semicontinuous at $\overline{y}$ for some $\varepsilon > 0$.

For each of the previously defined properties we can associate the corresponding modulus at $(\overline{y},\overline{x}) \in \operatorname{gph}\mathcal{M}$ or $\overline{y} \in \operatorname{dom}\mathcal{M}$ as the infimum of constants $\kappa$ such that the corresponding variational inequality holds for some associated neighborhoods, except the modulus of Hoffman stability, which does not involve any neighborhood. These moduli may be written as follows (the first and the third come directly from the definitions, while the second and the fourth are established in [A, Proposition 2]):

- *calmness modulus* of $\mathcal{M}$ at $(\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}$ :

$$\operatorname{clm} \mathcal{M} \left(\overline{y}, \overline{x}\right) = \limsup_{x \to \overline{x}} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)},$$

- *Lipschitz upper semicontinuity modulus* of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$ :

$$\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \limsup_{y \to \overline{y}} \left( \sup_{x \in \mathcal{M}(y)} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(y, \overline{y}\right)} \right),$$

- *Hoffman modulus* of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$ :

$$\operatorname{Hof} \mathcal{M} \left(\overline{y}\right) = \sup_{x \in X} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)},$$

- uniform *calmness modulus* of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$ :

$$\operatorname{uclm} \mathcal{M}(\overline{y}) = \limsup_{d(x, \mathcal{M}(\overline{y})) \to 0} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)}.$$

Finally, the global Hoffman constant of $\mathcal{M}$, inspired by the pioneer work [18], is defined as

$$\operatorname{Hof} \mathcal{M} = \sup_{(y,x) \in (\operatorname{dom} \mathcal{M}) \times X} \frac{d(x, \mathcal{M}(y))}{d\left(y, \mathcal{M}^{-1}(x)\right)}.$$

Here we use the convention $\frac{0}{0} := 0$ and $\limsup_{z \to \overline{z}}$ is understood as the supremum (maximum, indeed) of all possible sequential upper limits for all possible sequences $\{z_r\}_{r \in \mathbb{N}}$ converging to $\overline{z}$ as $r \to \infty$.

It is clear that

$$\sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M} \left(\overline{y}, x\right) \leq \operatorname{uclm} \mathcal{M} \left(\overline{y}\right) \leq \operatorname{Lipusc} \mathcal{M} \left(\overline{y}\right) \leq \operatorname{Hof} \mathcal{M} \left(\overline{y}\right) \leq \operatorname{Hof} \mathcal{M}. \quad (3.2)$$

Concerning background results, [22] expresses the calmness modulus of the feasible set mapping (in the equivalent terminology of local error bounds) in terms of limits of subdifferentials. A point-based expression for this modulus is given in [11, Theorem 4] under RHS perturbations and an extension to semi-infinite systems under a certain regularity condition is established in [25]. The case of full perturbations is solved in [11, Theorem 5]. Concerning the calmness modulus of $\mathcal{F}^{op}$ at a nominal $\left(\left(\overline{c},\overline{b}\right),\overline{x}\right)$, which turns out to coincide with that of $\mathcal{F}^{op}_{\overline{c}}$ at $\left(\overline{b},\overline{x}\right)$, a point-based expression is given in [10, Theorem 4.1] in terms of a certain type of feasible set mappings. For additional details, the reader is addressed to the preliminary sections of Apendices A and B. In relation to the global Hoffman constant for $\mathcal{F}$, its finiteness and some first estimates were given in the seminal work of Hoffman [18]. Some exact formulae from different points of view can be traced out from [7], [21] and [27], among others.

In order to introduce the framework of Appendix C, let us denote by $\mathcal{G}_a$ the inverse multifunction of $\mathcal{F}_a$, by specifying the left-hand side $a$ in our feasible set mapping $\mathcal{F}$, since it is also subject to perturbations in this appendix. Given any property $\mathcal{P}$ of $\mathcal{G}_{\overline{a}}$ fulfilled at the nominal $\left(\overline{x},\overline{b}\right) \in \text{gph}\mathcal{G}_{\overline{a}}$, the *radius* of $\mathcal{P}$-stability at that point is defined as

$$\text{rad}_{\mathcal{P}}\, \mathcal{G}_{\overline{a}}(\overline{x},\overline{b}) := \inf_{g\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)} \left\{ \|g\| \;\middle|\; \mathcal{G}_{\overline{a}} + g \text{ does not have } \mathcal{P} \text{ at } \left(\overline{x}, \overline{b} + g\left(\overline{x}\right)\right)\right\}, \tag{3.3}$$

where $\mathcal{L}\left(\mathbb{R}^n,\mathbb{R}^m\right)$ stands for the space of linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ endowed with the norm subordinated to the norms under consideration in these spaces. This definition of radius is inspired by the one given in [14, Definition 1.4] for the metric regularity property in more general contexts; see also [13] for the property of metric subregularity.

**Definition 1 (adaptation of [C, Definition 1])** *Given system* $\overline{\sigma} \equiv (\overline{a},\overline{b})$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, *we say that*

*(i)* $\mathcal{G}_{\overline{a}}$ *is* robustly subregular *at* $\left(\overline{x},\overline{b}\right)$ *if there exist constants* $\kappa \geq 0$ *and* $\varepsilon > 0$ *along with a neighborhood* $U$ *of* $\overline{x}$ *such that*

$$d\left(x, \mathcal{F}_a\left(\overline{b} + (a-\overline{a})'\overline{x}\right)\right) \leq \kappa d\left(\overline{b} + (a-\overline{a})'\overline{x}, \mathcal{G}_a(x)\right) \tag{3.4}$$

*for all* $x \in U$ *and all* $a \in \mathbb{R}^n$ *such that* $\|a-\overline{a}\| < \varepsilon$. *The infimum of constants* $\kappa$ *over the triplets* $(\kappa,\varepsilon,U)$ *satisfying* (3.4) *is called the* robust subregularity modulus *of* $\mathcal{G}_{\overline{a}}$ *at* $\left(\overline{x},\overline{b}\right)$ *and will be denoted by* $\operatorname{rob}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)$. *As stated in [C, Theorem 5 (iii)]*

$$\operatorname{rob}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)=\limsup_{a\to\overline{a}}\mathcal{S}\left(a\right), \tag{3.5}$$

*where* $\mathcal{S}\left(a\right):=\operatorname{clm}\mathcal{F}_{a}\left(\overline{b}+(a-\overline{a})'\overline{x},\overline{b}\right)$ *(see the beginning of [C, Section 3]).*

$(ii)$ $\mathcal{G}_{\overline{a}}$ *is* continuously subregular *at* $\left(\overline{x},\overline{b}\right)$ *if* $\mathcal{S}$ *is continuous at* $\overline{a}$.

For the reader's convenience, we finish this chapter by extracting some relevant information from the introduction of Appendix D. Note that each appendix has its own notation and now $T$ is used to represent an operator. Specifically, let $X$ be a real Banach space, with topological dual $X^*$, and denote by $\langle\cdot,\cdot\rangle$ the corresponding canonical pairing. A set-valued operator $T : X \rightrightarrows X^*$ is said to be *monotone* if

$$\langle x-y, x^*-y^*\rangle \geq 0 \text{ whenever } (x,x^*),(y,y^*)\in \operatorname{gph}T.$$

In the case when both $T$ and $-T$ are monotone, then $T$ is called *bimonotone.* If $T$ is monotone and, in addition, $\operatorname{gph}T$ is maximal in the sense of inclusion order, it is said to be *maximally monotone.* A well-known example of maximally monotone operator is the subdifferential operator of a proper lower semicontinuous (lsc, for short) convex function $f : X \to \left]-\infty,+\infty\right]$, denoted by $\partial f$ (see [D, Section 2] for details). Monotone operators are fundamental tools of nonlinear analysis and optimization; see, e.g., the books [2, 4, 6, 28, 29, 31, 32]. A monotone operator $T$ is called *paramonotone* if the following implication holds:

$$\left.\begin{array}{c}(x,x^*),(y,y^*)\in \operatorname{gph}T\\ \langle x-y,x^*-y^*\rangle=0\end{array}\right\}\Rightarrow (x,y^*),(y,x^*)\in\operatorname{gph}T.$$

The term paramonotonicity was introduced in [12] (although the condition was previously presented in [5] without a name). The initial motivation for the introduction of paramonotone operators comes from its crucial role regarding interior point methods for variational inequalities (see again [5] and [12], and also [19]). Some important examples of paramonotone operators are gathered in [D, Section

2]. At this moment, let us mention that subdifferentials of proper lsc convex functions enjoy this property (see [19, Proposition 2.2] in the Euclidean space and [3, Fact 3.1] for its extension to Banach spaces).

# 4 Discussion of main original results

This chapter is intended to gather the main original contributions of the thesis, so that the reader has a panoramic overview. For further details, see Appendices A-E. The chapter is divided into three sections. The first one, corresponding to Appendices A, B and E, is focused on the semilocal and global Lipschitz-type moduli appearing in (3.2) and their particularization to $\mathcal{F}$, $\mathcal{F}^{op}$ and $\mathcal{F}^{op}_{\underline{c}}$; the second one, corresponding to Appendix C, deals with robust and continuous subregularity and the radius of the first one; the third section is focused on paramonotone and bimonotone operators, as well as their applications to feasibility problems, and corresponds to Appendix D.

## 4.1 Semilocal and global moduli

First of all, we show that for appropriate convex graph multifunctions all inequalities in (3.2) except the last one hold as equalities. This is the case of $\mathcal{F}$ (under RHS perturbations).

**Theorem 1 ([A, Theorem 4])** *Let $\mathcal{M} : Y \rightrightarrows X$, with $Y$ being a normed space and $X$ being a reflexive Banach space, and assume that* $\operatorname{gph}\mathcal{M}$ *is a nonempty convex set. Let* $\overline{y} \in \operatorname{dom}\mathcal{M}$ *with* $\mathcal{M}\left(\overline{y}\right)$ *closed. Then one has*

$$\sup_{x\in\mathcal{M}(\overline{y})} \operatorname{clm}\mathcal{M}\left(\overline{y},x\right) = \operatorname{uclm}\mathcal{M}\left(\overline{y}\right) = \operatorname{Lipusc}\mathcal{M}\left(\overline{y}\right) = \operatorname{Hof}\mathcal{M}\left(\overline{y}\right).$$

In the next theorem, $T$ is assumed to be a compact metric space and we consider the parameter space $C(T,\mathbb{R})$ of continuous functions $b : t \mapsto b_t \in \mathbb{R}$ endowed with the supremum norm. In this context we provide a formula for $\operatorname{Hof} \mathcal{F}$ which, combined with Theorem 3, shows that the last inequality of (3.2) may be strict for $\mathcal{F}$. This theorem generalizes the previous known formulae commented in the previous chapter; specifically [27, (3) and (4)], [21, Theorem 2.7] and [7, Theorem 8]. Hereafter 'conv' means convex hull.

**Theorem 2 ([A, Theorem 5])** *Consider $\mathcal{F} : C(T,\mathbb{R}) \rightrightarrows \mathbb{R}^n$. We have*

$$\operatorname{Hof} \mathcal{F} = \sup_{\substack{J \subset T \text{ compact} \\ 0_n \notin \operatorname{conv}\{a_t,\, t\in J\}}} d_*(0_n, \operatorname{conv}\{a_t,\ t \in J\})^{-1}.$$

The next theorem essentially refines the expression $\sup_{x\in\mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}\left(\overline{y}, x\right)$ of Theorem 1 when applied to $\mathcal{F}$ by reducing the supremum to a maximum of a finite amount of calmness moduli. To do this we need some extra notation. From now on we consider the set

$$\mathcal{E}(b) := \operatorname{extr}\left(\mathcal{F}(b) \cap \operatorname{span}\left\{a_t,\ t \in T\right\}\right), \text{ with } b \in \operatorname{dom} \mathcal{F}, \tag{4.1}$$

where 'extr' and 'span' stand for the set of extreme points and the linear subspace generated by the corresponding sets, respectively. It is known that $\mathcal{E}(b)$ is always a nonempty and finite set when $T$ is finite (see Appendix A for details). This construction is inspired by the one of [24, p. 142]. The theorem also appeals to the family $\mathcal{D}(x)$ of subsets $D \subset T(x)$ (set of active indices at $x$ for the linear system under consideration) such that system

$$\left\{\begin{array}{ll} a_t' d = 1, & t \in D, \\ a_t' d < 1, & t \in T(x) \setminus D \end{array}\right\} \tag{4.2}$$

is consistent (in the variable $d \in \mathbb{R}^n$).

**Theorem 3 ([A, Theorem 6])** *Let $\overline{b} \in \operatorname{dom} \mathcal{F}$ and assume that* $\{a_t' x \le \overline{b}_t,\ t \in$

$T\}$ *is a locally polyhedral system (see [A, Section 4]). Then*

$$\operatorname{Hof}\mathcal{F}\left(\overline{b}\right)=\sup_{x\in\mathcal{E}\left(\overline{b}\right)}\operatorname{clm}\mathcal{F}\left(\overline{b},x\right)=\sup_{x\in\mathcal{E}\left(\overline{b}\right)}\sup_{D\in\mathcal{D}(x)}d_{*}\left(0_n,\operatorname{conv}\left\{a_t,\ t\in D\right\}\right)^{-1}.$$

The previous theorem provides a point-based expression for $\operatorname{Hof}\mathcal{F}\left(\overline{b}\right)$ in contrast to that of [1, Theorem 2.6].

Our next focus is on the optimal set mapping. Despite the striking resemblance between Theorems 3 and 5, the methodology for deriving them is completely different, due to the fact that $\operatorname{gph}\mathcal{F}^{op}_{\overline{c}}$ is no longer convex (and hence neither is $\operatorname{gph}\mathcal{F}^{op}$). Nevertheless, we extract geometrical patterns from the graph that serve as a tool to overcome this drawback. Specifically, we construct the *local directional optimal set mapping* for a nominal problem $\overline{\pi}\equiv\left(\overline{c},\overline{b}\right)$, some $\varepsilon>0$ and direction $d\in\mathbb{R}^m$ with $\|d\|_\infty=1$ as the multifunction $\mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}:=[0,\varepsilon]\rightrightarrows\mathbb{R}^n$ given by

$$\mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}(\mu)=\mathcal{F}^{op}\left(\overline{c},\overline{b}+\mu d\right),\quad \mu\in[0,\varepsilon].$$

**Theorem 4 ([B, Theorem 3.1])** *Let* $\overline{\pi}=\left(\overline{c},\overline{b}\right)\in\operatorname{dom}\mathcal{F}^{op}$ *and* $\varepsilon>0$ *be as in [B, Lemma 3.2]. Then* $\operatorname{gph}\mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}$ *is convex for all* $d\in\mathbb{R}^m$ *with* $\|d\|_\infty=1$.

In the following theorem we appeal to the set of "generalized extreme points" $\mathcal{E}^{op}(\overline{\pi}):=\operatorname{extr}(\mathcal{F}^{op}(\overline{\pi})\cap\operatorname{span}\{a_t,\ t\in\{1,\ldots,m\}\})$.

**Theorem 5 ([B, Theorem 4.2])** *Let* $\overline{\pi}\in\operatorname{dom}\mathcal{F}^{op}$*, then*

$$\operatorname{Lipusc}\mathcal{F}^{op}(\overline{\pi})=\max_{x\in\mathcal{E}^{op}(\overline{\pi})}\operatorname{clm}\mathcal{F}^{op}(\overline{\pi},x).$$

In the remaining part of this section we are concerned with the global Hoffman constant of the argmin mapping. First we show that the case of canonical perturbations is essentially trivial.

**Proposition 6 ([E, Proposition 2])** *We have that*

$$\operatorname{Hof} \mathcal{F}^{op} = \begin{cases} 0 \textit{ if } \{a_t,\ t \in T\} = \{0_n\}, \\ +\infty \textit{ otherwise.} \end{cases}$$

This reduces our study to mapping $\mathcal{F}^{op}_{\overline{c}}$, whose graph presents a rich structure. We isolate one of the relevant features of this graph in the following definition. Here we use symbol $\mathcal{S}$ with a different meaning to that of Appendix C (see Definition 1 in Chapter 3), with the aim of keeping the same notation as in the appendices, i.e., of the original publications (preprint [E] in this case).

**Definition 7 ([E, Definition 2])** *Let $I$ be a finite index set and, for each $i \in I$, consider a multifunction $\mathcal{S}_i : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with a polyhedral graph. We say that $\mathcal{S} := \cup_{i\in I}\mathcal{S}_i$ is a* well-connected polyhedral mapping *if the following properties hold:*

*(i)* $\operatorname{dom}\mathcal{S}$ *is a convex set in* $\mathbb{R}^m$*;*

*(ii)* $\mathcal{S}|_{\operatorname{dom}\mathcal{S}_i} = \mathcal{S}_i$*, for all* $i \in I$*.*

The previous structure is key to obtain the desired formula of the global Hoffman constant of $\mathcal{F}^{op}_{\overline{c}}$, as well as the following construction:

**Definition 8 ([E, Definition 3])** *Let $\mathcal{S} = \cup_{i\in I}\mathcal{S}_i$ be a well-connected polyhedral mapping. Let $b, \overline{b} \in \operatorname{dom}\mathcal{S}$. We call a subdivision $0 =: \mu_0 < \mu_1 < \ldots < \mu_N := 1$ together with a family of indices $i_1, \ldots, i_N \in I$ connecting $b$ with $\overline{b}$ if for all $k \in \{1, \ldots, N\}$ and all $\mu \in [\mu_{k-1}, \mu_k]$ there holds $\overline{b} + \mu\left(b - \overline{b}\right) \in \operatorname{dom}\mathcal{S}_{i_k}$. In other words, $\mathcal{S}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right) = \mathcal{S}_{i_k}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right)$ whenever $\mu \in [\mu_{k-1}, \mu_k]$.*

The following technical result constitutes the key tool to derive the aimed formula for $\operatorname{Hof}\mathcal{S}$ in terms of the supremum of calmness moduli, which is established in the subsequent corollary.

**Theorem 9 ([E, Theorem 4])** *Let* $\mathcal{S} = \bigcup_{i \in I} \mathcal{S}_i$ *be a well-connected polyhedral mapping. Let* $b, \overline{b} \in \operatorname{dom} \mathcal{S}$ *with* $b \neq \overline{b}$ *and consider a subdivision* $0 =: \mu_0 < \mu_1 < \ldots < \mu_N := 1$ *together with a family of indices* $i_1, \ldots, i_N \in I$ *connecting* $b$ *with* $\overline{b}$. *Then, for every* $x \in \mathcal{S}(b)$, *there exist points* $x^k \in \mathcal{S}\left(\overline{b} + \mu_k d\right)$ *with* $k = 0, \ldots, N-1$ *such that*

$$\frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} \leq \max\{\operatorname{clm} \mathcal{S}\left(\overline{b} + \mu_k\left(b - \overline{b}\right), x^k\right) \mid k = 0, \ldots, N-1\}$$

$$\leq \max\{\operatorname{Lipusc} \mathcal{S}\left(\overline{b} + \mu_k\left(b - \overline{b}\right)\right) \mid k = 0, \ldots, N-1\}.$$

**Corollary 10 ([E, Corollary 1])** *Let* $\mathcal{S}$ *be a well-connected polyhedral mapping. Then*

$$\operatorname{Hof} \mathcal{S} = \sup\{\operatorname{Lipusc} \mathcal{S}(b) \mid b \in \operatorname{dom} \mathcal{S}\} = \sup\{\operatorname{clm} \mathcal{S}(b, x) \mid (b, x) \in \operatorname{gph} \mathcal{S}\}.$$

Once the well-connected polyhedral structure has proven to be useful in order to get a formula for the global Hoffman constant, we look for a computable expression for $\operatorname{Hof} \mathcal{F}_{\overline{c}}^{op}$. The role of $I$ will be played by the family of all possible *minimal Karush-Kuhn-Tucker subsets of indices* which is defined by

$$\mathcal{M}_{\overline{c}} := \left\{ D \subset T \;\middle|\; \begin{array}{l} -\overline{c} \in \operatorname{cone}\left\{a_t,\, t \in D\right\} \text{ and } D \text{ is} \\ \text{minimal w.r.t. the inclusion order} \end{array} \right\},$$

where '$\operatorname{cone}$' means 'convex cone generated by'.

After checking that $\mathcal{F}_{\overline{c}}^{op}$ is a well-connected polyhedral mapping (see [E, Proposition 4]), and combining this with some mentioned background results on calmness moduli (in terms of the so-called end set, see e.g. [E, Theorem 2] for details) we obtain the following theorem, providing a point-based expression for $\operatorname{Hof} \mathcal{F}_{\overline{c}}^{op}$.

**Theorem 11 ([E, Theorem 5])** *Let* $-\overline{c} \in \operatorname{cone}\left\{a_t,\ t \in T\right\}$. *One has*

$$\begin{aligned}\operatorname{Hof} \mathcal{F}^{op}_{\overline{c}} &= \max_{b \in \operatorname{dom} \mathcal{F}} \operatorname{Lipusc} \mathcal{F}^{op}\left(\overline{c}, b\right) \\ &= \max_{b \in \operatorname{dom} \mathcal{F}} \max_{x \in \mathcal{E}^{op}\left(\overline{c}, b\right)} \operatorname{clm} \mathcal{F}^{op}\left(\left(\overline{c}, b\right), x\right) \\ &= \max_{\substack{D \subset S \subset T \\ D \in \mathcal{M}_{\overline{c}}}} \left\{d_*\left(0_n, \operatorname{end} \operatorname{conv}\left\{a_t,\ t \in S;\ \ -a_t,\ t \in D\right\}\right)\right\}^{-1}.\end{aligned}$$

## 4.2 Robust and continuous subregularity

Another incursion into the world of stability analysis is been made in this thesis from the study of the radius of a given property. The topic started with the seminal work [14], where the idea of perturbing the nominal problem conserving a desired property is embodied and developed in the context of metric regularity for different types of mappings: linear, sublinear, differentiable, etc. In contrast to metric regularity, the particular behavior of the metric subregularity property, which does always hold in the finite linear inequality system setup under data perturbation, puts us in a dichotomy: tackling some of the already existing open questions from the last 20 years on specific types of structured perturbations or studying the linear structure from a different point of view. We advance that the second path was the chosen one in [C], leading to new stuff centered on the variation of the classical moduli. We focus on studying the continuity of the metric subregularity modulus, but there is still room for further research in this new direction.

More in depth, we consider left-hand side (LHS in brief) perturbations over a nominal system and record the continuity of the metric subregularity modulus. First, we make use of the characterization recalled in [C, Theorem 2] for the former modulus in terms of the end set or, equivalently, in terms of the family $\mathcal{D}_{\overline{a}}$, which is the same introduced in (3.2) above but making explicit the dependence on the LHS coefficients. Therefore, describing the behavior of such end sets will translate into immediate consequences on the behavior of the moduli themselves. With this aim, recall the definition of set $S(a)$ in Definition 1$(i)$, which can be also written as (see [C, Formula (17) and Theorem 1])

$$\mathcal{S}(a) = d_*(0_n, E(a))^{-1},$$

with

$$E(a) = \bigcup_{D \in \mathcal{D}_a} \operatorname{conv}\{a_t,\, t \in D\},$$

for $a \in \mathbb{R}^{n \times m}$.

The next two technical results concern the continuity behavior of the previous mappings $E$ and $\mathcal{S}$.

**Theorem 12 ([C, Theorem 3])** *Let $\overline{a} \in \mathbb{R}^{n \times m}$. We have:*

(i) $\operatorname*{Liminf}_{a \to \overline{a}} E(a) = \bigcup_{D \in \mathcal{D}_{\overline{a}}} \operatorname{conv}\left\{\overline{a}_t,\, t \in D\right\} = E(\overline{a});$

(ii) $\operatorname*{Limsup}_{a \to \overline{a}} E(a) = \bigcup_{D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0} \operatorname{conv}\left\{\overline{a}_t,\, t \in D\right\} \supset E(\overline{a}).$

Therefore, $E$ is lower semicontinuous in the sense of Berge. Moreover, the inclusion in (ii) involving $\mathcal{D}_{\overline{a}}^0$ defined by replacing 1 with 0 in $\mathcal{D}_{\overline{a}}$ may be strict.

**Theorem 13 ([C, Theorem 4])** *Let $\overline{\sigma} \equiv (\overline{a}, \overline{b})$ and $\overline{x} \in \mathcal{F}_{\overline{a}}(\overline{b})$. Then:*

(i) $\operatorname*{Liminf}_{a \to \overline{a}} \mathcal{S}(a) = \left[d_*\left(0_n, \bigcup_{D \in \mathcal{D}_{\overline{a}}} \operatorname{conv}\left\{\overline{a}_t,\, t \in D\right\}\right)\right]^{-1} = \mathcal{S}(\overline{a});$

(ii) $\operatorname*{Limsup}_{a \to \overline{a}} \mathcal{S}(a) = \left[d_*\left(0_n, \bigcup_{D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0} \operatorname{conv}\left\{\overline{a}_t,\, t \in D\right\}\right)\right]^{-1} \geq \mathcal{S}(\overline{a}).$

Again, $\mathcal{S}$ may fail to be upper semicontinuous, finding situations of finite continuity gaps and others of infinite nature.

Since we encounter two types of continuity gaps, we proceed to characterize the finiteness in a first stage. Unexpectedly, the characterization provided in the next Theorem 14 is nothing else but another kind of uniform regularity property, which we name as *robust subregularity*. Moreover, its associated modulus, denoted by $\operatorname{rob}\mathcal{G}_{\overline{a}}(\overline{x},\overline{b})$, is precisely the upper limit of $\mathcal{S}$ at the nominal $\overline{a}$. Secondly, the continuity of $\mathcal{S}$ at $\overline{a}$ is captured in the *continuous subregularity* property, which is characterized in Corollary 15.

**Theorem 14 ([C, Theorem 5])** *Given* $\overline{\sigma} \equiv \left(\overline{a},\overline{b}\right)$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, *the following statements are equivalent:*

(i) $\limsup_{a\to\overline{a}} \mathcal{S}(a)$ is finite;

(ii) $0_n \notin \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right)\right\}$;

(iii) There exist constants $\kappa \geq 0$ and $\varepsilon > 0$ along with a neighborhood $U$ of $\overline{x}$ such that

$$d\left(x,\mathcal{F}_a\left(\overline{b}+(a-\overline{a})'\overline{x}\right)\right) \leq \kappa d\left(\overline{b}+(a-\overline{a})'\overline{x},\mathcal{G}_a(x)\right)$$

for all $x \in U$ and all $a \in \mathbb{R}^n$ such that $\|a-\overline{a}\| < \varepsilon$.

Moreover, $\limsup_{a\to\overline{a}} \mathcal{S}(a)$ coincides with the infimum of constants $\kappa$ over the triplets $(\kappa,\varepsilon,U)$ satisfying (3.4).

**Corollary 15 ([C, Corollary 2])** *For the nominal data* $\overline{\sigma} \equiv (\overline{a},\overline{b})$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, *the following statements are equivalent:*

(i) $\mathcal{G}_{\overline{a}}$ is continuously subregular at $\left(\overline{x},\overline{b}\right)$;

(ii) $\operatorname{rob}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right) = \mathcal{S}\left(\overline{a}\right)$;

(iii) It holds

$$0 \neq d_*\left(0_n, \bigcup D \in \mathcal{D}_{\overline{a}} \operatorname{conv}\{\overline{a}_t,\ t \in D\}\right) = d_*\left(0_n, \bigcup D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0 \operatorname{conv}\{\overline{a}_t,\ t \in D\}\right).$$

Finally, their radii are studied with moderate success. A computable point-based formula is given:

**Theorem 16 ([C, Theorem 6])** *Assume that $\mathcal{G}_{\overline{a}}$ is robustly subregular at $(\overline{x}, \overline{b})$. Then*

$$\mathrm{rad}_{\mathrm{rob}}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right) = d_*(0_n, \mathrm{bd}\,\mathrm{conv}\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\}).$$

As for the radius of continuous subregularity, apart from the direct bound

$$\mathrm{rad}_{\mathrm{rob}}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right) \geq \mathrm{rad}_{\mathrm{cont}}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right),$$

some hints for future development can be extracted from the last example in Appendix C.

**Corollary 17 ([C, Corollary 3])** *One has*

$$\mathrm{rad}_{\mathrm{rob}}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right) \leq \frac{1}{\mathrm{rob}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right)},$$

*and the inequality may be strict. Moreover,* $\mathrm{rob}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right) < +\infty$ *implies* $\mathrm{rad}_{\mathrm{rob}}\, \mathcal{G}_{\overline{a}} \left(\overline{x}, \overline{b}\right) > 0$.

## 4.3 Feasibility problems via paramonotone operators

To conclude the current chapter, let us comment one last area of interest in which we have taken part in the thesis. Up to this point, all systems considered are feasible since the properties under consideration so far involve elements of the domain or the graph of the corresponding multifunctions. When feasibility is not guaranteed, the question of reaching it arises. The distance to feasibility problem serves also as a motivation to review the literature related to monotone operators, which has a wide range of uses in applied mathematics. As pointed out in Chapter 3, we focus on those operators which are simultaneously paramonotone and bimonotone, whose characterization is given the following corollary of [D, Proposition 8], which we omit here for brevity.

**Corollary 18 ([D, Corollary 9])** *Let $T : X \rightrightarrows X^*$, the following conditions are equivalent:*

(i) $T$ is paramonotone and bimonotone;

(ii) $T$ is monotone and constant on its domain;

(iii) $(\operatorname{dom} T - \operatorname{dom} T) \perp (\operatorname{range} T - \operatorname{range} T)$ and $\operatorname{gph} T = \operatorname{dom} T \times \operatorname{range} T$.

The first application verses about simultaneous projections. More in depth, we study the minimal weighted distance to two disjoint non-empty closed convex sets in a Hilbert space, denoted by $S_1$ and $S_2$ (see also [D, Theorem 21] for an extension to finitely sets under some differentiability assumptions). If we define

$$\mathcal{A}(\alpha_1, \alpha_2, p) := \arg\min \alpha_1 d(\cdot, S_1)^p + \alpha_2 d(\cdot, S_2)^p,$$

then there are three differentiated classes of problems depending on the chosen alphas and $p$:

1. $\underline{p = 1,\ \alpha_1 \neq \alpha_2}$. Then [D, Proposition 16] states, assuming without loss of generality that $\alpha_1 > \alpha_2$, that

$$\mathcal{A}(\alpha_1, \alpha_2, 1) = \arg\min_{S_1} d(\cdot, S_2).$$

2. $\underline{p = 1,\ \alpha_1 = \alpha_2 = 1/2}$. Appealing to [D, Proposition 18 (ii)] we have

$$\mathcal{A}\left(\frac{1}{2}, \frac{1}{2}, 1\right) = \{x \in X : x \in ]P_1(x), P_2(x)[\} \cup \arg\min_{S_1} d(\cdot, S_2) \cup \arg\min_{S_2} d(\cdot, S_1),$$

where $P_i$ stands for the projection over the set $S_i$, $i = 1, 2$.

3. $\underline{p > 1}$. We can describe $\mathcal{A}(\alpha_1, \alpha_2, p)$ as the set of fixed points of

$$\frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_1 + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_2$$

according to [D, Theorem 19 (iii)].

The second application deals with the distance to feasibility in the framework of convex inequality systems in $\mathbb{R}^n$ under RHS perturbations. Let us consider the parameterized system,

$$\sigma(b) := \left\{ g_i(x) \leq b_i,\ i = 1, \ldots, m \right\},$$

where $x \in \mathbb{R}^n$, $(b_i)_{i=1,\ldots,m} \equiv b \in \mathbb{R}^m$, and $g_i : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $i = 1, 2, ..., m$, where the space of variables, $\mathbb{R}^n$, is endowed with an arbitrary norm, $\|\cdot\|$, with dual norm $\|\cdot\|_*$ and the associated distances denoted by $d$ and $d_*$, respectively. The space of parameters, $\mathbb{R}^m$, is endowed with any $p$-norm, $\|\cdot\|_p$, provided that $p \geq 2$, and the associated distance is denoted by $d_p$. We denote by $\Theta_c$ the set of consistent parameters; i.e.,

$$\Theta_c := \{ b \in \mathbb{R}^m \mid \sigma(b) \text{ is consistent} \}.$$

The distance from $\overline{b} \in \mathbb{R}^m \setminus \Theta_c$ to feasibility is

$$d_p\left(\overline{b}, \Theta_c\right) = \inf\left\{ \left\|\overline{b} - b\right\|_p : b \in \Theta_c \right\},$$

and verifies (see [D, Proposition 23])

$$d_p\left(\overline{b}, \Theta_c\right)^p = \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p.$$

The next key result extends the well-known Ascoli formula, for the distance from a point to a half space, to the convex case. Here '$\partial$' stands for the classical subdifferential of convex analysis.

**Proposition 19 ([D, Proposition 24])** *Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $b \in \mathbb{R}$ be such that the corresponding sublevel set, $S$, is nonempty. Then we have:*

(i) For any $x \in \mathbb{R}^n$,

$$d_S(x) \geq \frac{[g(x) - b]_+}{d_*\left(0_n, \partial g(x)\right)};$$

(ii) Assume that there exists $\widehat{x} \in \mathbb{R}^n$ (called a Slater point) such that $g(\widehat{x}) < b$. Then, for any $x \in \mathbb{R}^n$,

$$d_S(x) \leq \frac{[g(x) - b]_+}{d_*\left(0_n, \partial g\left(P_S(x)\right)\right)},$$

where $P_S(x)$ is the metric projection set of $x$ on $S$ with respect to the norm $\|\cdot\|$.

As a consequence of this result, upper and lower estimates for the $p$-distance ($p \geq 2$) to feasibility are proven in [D, Corollary 27] when $\mathbb{R}^n$ is equipped with the Euclidean norm. When confined to the linear setup, those bounds tighten attaining equality: $d_p\left(\overline{b}, \Theta_c\right) = \left\|\overline{d}\right\|_p$ for a certain $\overline{d} \in \mathbb{R}^m$ (see [D, Corollary 28] for details). The next theorem provides an operative procedure to compute such a distance. In it $[y]_+$, $y \in \mathbb{R}^m$ stands for the positive part coordinate by coordinate; i.e., $[y]_+ := \left([y_i]_+\right)_{i=1,\ldots,m}$ and set $\mathcal{A}$ appearing below is defined at the beginning of [D, Section 5.1].

**Theorem 20** *The following conditions are equivalent:*

(i) $\left(x^0, h^0\right) \in \mathcal{A} \times \left\{\overline{d}\right\}$;

(ii) $\left(x^0, h^0\right)$ is a solution of the system, in the variable $(x, h)$,

$$\begin{cases} \left[Ax - \overline{b}\right]_+ = h, \\ A'h = 0_n. \end{cases}$$

(iii) $\left(x^0, h^0\right)$ is an optimal solution of the quadratic problem, in the variable

$(x, h)$,

$$\begin{aligned} \min \quad & \langle h, h \rangle \\ s.t. \quad & Ax \leq \overline{b} + h, \\ & h \geq 0_m. \end{aligned}$$

# 5 | Conclusions and perspectives

## 5.1 English

The previous chapter presents a selection of the main original contributions of the thesis, conceived just as an overview. It has been stated in a direct style for the sake of brevity and with the intention of avoiding unnecessary redundancies with the appendices containing the complete publications. As a quick summary, we provide point-based expressions for the Lipschitz upper semicontinuity modulus of $\mathcal{F}$, $\mathcal{F}^{op}$ and $\mathcal{F}^{op}_{\bar{c}}$, as well as the Hoffman modulus at $\overline{b}$ of $\mathcal{F}$ and the global Hoffman constant for $\mathcal{F}$ (in the semi-infinite case) and $\mathcal{F}^{op}_{\bar{c}}$ (for $\mathcal{F}^{op}$ it is $+\infty$ except a trivial situation). Moreover, two new concepts of (metric) subregularity, specifically robust and continuous subregularity, have been introduced and the radius with respect to the first of them has been computed. Finally, we have investigated operators which are simultaneously paramonotone and bimonotone, and presented two applications to simultaneous projections and feasibility problems.

There are some open problems directly connected with the achievements of the thesis, which we list below:

- **The semi-infinite case.** To the best of our knowledge, there are no operative expressions of $\operatorname{clm} \mathcal{F}$ for semi-infinite systems without the (quite restrictive) regularity condition introduced in [25]. Any advance in the calmness modulus could have direct repercussions in semilocal Lipschitz-type moduli.
- **The convex case.** Each convex inequality turns our to be equivalent (in the sense of having the same solution set) to a linear semi-infinite inequality

system by means of the so-called standard linearization via the Fenchel-Legendre conjugate. In this way a convex system under RHS perturbations gives rise to a larger linear system with the same type of perturbations and a relatively good structure (depending on the quality of convex functions involved).

- **Moduli for fully perturbed systems.** Perturbing both sides ($a$ and $b$) of linear inequality systems entails notable differences with the context of RHS perturbations. To start with, the graph of the feasible set mapping is no longer convex and hence Theorem 1 does not apply. A first interesting problem to solve is determining the Lipschitz upper semicontinuity modulus of the feasible set mapping in the case when the nominal feasible set is bounded. We have some conjectures.

- **Hoffman modulus of the argmin mapping $\mathcal{F}^{op}$ at $\left(\overline{c}, \overline{b}\right)$.** [B, Example 5.1] exhibits a situation where $\text{Lipusc}\,\mathcal{F}^{op}\left(\overline{c}, \overline{b}\right) < \text{Hof}\,\mathcal{F}^{op}\left(\overline{c}, \overline{b}\right)$ but we do not have a point-based expression for the latter. Perhaps it would be interesting to start with $\mathcal{F}^{op}_{\overline{c}}$. To this respect, the concept of *break steps* introduced in [E, Section 5] could be a key ingredient.

- **Radius of continuous regularity.** The problem of finding a point-based formula for such a radius remains open. Some hints for future research can be traced out from [C, Example 4], which illustrates some of the difficulties that may arise in this search and the wide casuistry that appears.

- **Radii for fully perturbed systems.** A key feature for defining the radius of a (set-valued) mapping with respect to a given property is the fact that the perturbed mapping should belong to the same class as the original one. In the case of the feasible set mapping of a linear inequality system under RHS perturbations, the perturbed mapping corresponds to another system with a different LHS. But if both sides ($a$ and $b$) are regarded as parameters, then the perturbed mapping corresponds to a quadratic system (see [9, Section 5] for details). Therefore, the problem of finding an appropriate framework for studying the radius of metric regularity or different types of subregularity of linear systems under full perturbations arises.

## 5.2 Spanish

El capítulo anterior presenta una selección de las principales aportaciones originales de la tesis, concebida sólo como una visión general. Se ha expuesto en un estilo directo en aras de la brevedad y con la intención de evitar redundancias innecesarias con los apéndices que contienen las publicaciones completas. A modo de resumen rápido, proporcionamos expresiones de tipo "point-based" para el módulo de Lipschitz upper semicontinuity de $\mathcal{F}$, $\mathcal{F}^{op}$ y $\mathcal{F}^{op}_{\bar{c}}$, así como el módulo de Hoffman en $\overline{b}$ de $\mathcal{F}^{op}$ y la constante global de Hoffman para $\mathcal{F}$ (en el caso semi-infinito) y $\mathcal{F}^{op}_{\bar{c}}$ (para $\mathcal{F}^{op}$ es $+\infty$ salvo una situación trivial). Además, dos nuevos conceptos de subregularidad (métrica), específicamente subregularidad robusta y continua, han sido introducidos y el radio con respecto al primero de ellos ha sido calculado. Por último, hemos investigado operadores que son simultáneamente paramonótonos y bimonótonos, y presentado dos aplicaciones a proyecciones simultáneas y problemas de factibilidad.

Hay algunos problemas abiertos directamente relacionados con los logros de la tesis, que enumeramos a continuación:

- **El caso semi-infinito.** Hasta donde nosotros sabemos, no existen expresiones operativas de $\operatorname{clm}\mathcal{F}$ para sistemas semi-infinitos sin la condición de regularidad (bastante restrictiva) introducida en [25]. Cualquier avance en el módulo de calmness podría tener repercusiones directas en los módulos semilocales de tipo Lipschitz.

- **El caso convexo.** Cada desigualdad convexa resulta ser equivalente (en el sentido de tener el mismo conjunto solución) a un sistema lineal semiinfinito por medio de la llamada linealización estándar mediante la conjugada de Fenchel-Legendre. De este modo, un sistema convexo bajo perturbaciones RHS da lugar a un sistema lineal mayor con el mismo tipo de perturbaciones y una estructura relativamente buena (dependiendo de la calidad de las funciones convexas implicadas).

- **Módulos para sistemas bajo perturbaciones totales.** Perturbar ambos lados ($a$ y $b$) de sistemas de desigualdades lineales implica diferencias notables con respecto al contexto de las perturbaciones RHS. Para empezar, el grafo de la multifunción conjunto factible ya no es convexo y por lo tanto el Teorema 1 no se aplica. Un primer problema interesante por resolver es la determinación del módulo de Lipschitz upper semicontinuity de la multifunción conjunto factible en el caso de que el conjunto factible nominal esté acotado. Tenemos algunas conjeturas.

- **Módulo de Hoffman de la multifunción argmin $\mathcal{F}^{op}$ en** $(\overline{c}, \overline{b})$. [B, Ejemplo 5.1] exhibe una situación en la que $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{c}, \overline{b}) < \operatorname{Hof} \mathcal{F}^{op}(\overline{c}, \overline{b})$, pero no tenemos una expresión general para esta último. Tal vez sería interesante empezar con $\mathcal{F}^{op}_{\overline{c}}$. A este respecto, el concepto de *break steps* introducido en [E, Sección 5] podría ser un ingrediente clave.

- **Radio de regularidad continua.** El problema de encontrar una fórmula de tipo "point-based" para dicho radio sigue abierto. Algunas pistas para su futura investigación se pueden seguir a partir de [C, Ejemplo 4], que ilustra algunas de las dificultades que pueden surgir en esta búsqueda y la amplia casuística que aparece.

- **Radios para sistemas bajo perturbaciones totales.** Una característica clave para definir el radio de una aplicación (conjunto-valuada) con respecto a una propiedad dada es el hecho de la multifunción perturbada debe pertenecer a la misma clase que la original. En el caso de la multifunción conjunto factible de un sistema de desigualdades lineales bajo perturbaciones RHS, la aplicación perturbada corresponde a otro sistema con un miembro izquierdo de las desigualdades diferente. Pero si ambos lados ($a$ y $b$) se consideran como parámetros, entonces la multifunción perturbada corresponde a un sistema cuadrático (véase [9, Sección 5] para más detalles). Por lo tanto, se plantea el problema de encontrar un marco apropiado para estudiar el radio de regularidad métrica o diferentes tipos de subregularidad de sistemas lineales bajo perturbaciones totales.

# Bibliography

[A] J. Camacho, M. J. Cánovas, and J. Parra. From calmness to Hoffman constants for linear semi-infinite inequality systems. *SIAM Journal on Optimization*, 32 (2022), 2859–2878.

[B] J. Camacho, M. J. Cánovas, and J. Parra. Lipschitz upper semicontinuity in linear optimization via local directional convexity. *Optimization*, 2022. https://doi.org/10.1080/02331934.2022.2057851

[C] J. Camacho, M. J. Cánovas, M.A. López, and J. Parra. Robust and continuous metric subregularity for linear inequality systems. *Computational Optimization and Applications*, 2022. https://doi.org/10.1007/s10589-022-00437-0

[D] J. Camacho, M. J. Cánovas, J. E. Martínez-Legaz, and J. Parra. Feasibility problems via paramonotone operators in a convex setting. 2023. https://doi.org/10.48550/arXiv.2307.00979

[E] J. Camacho, M. J. Cánovas, H. Gfrerer, and J. Parra. Hoffman constant of the argmin mapping in linear optimization. 2023. https://doi.org/10.48550/arXiv.2307.01034

[1] D. Azé and J. N. Corvellec. On the sensitivity analysis of Hoffman constants for systems of linear inequalities. *SIAM Journal on Optimization*, 12 (2002), 913–927.

[2] H. H. Bauschke and P. L. Combettes. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York, 2011.

[3] H. H. Bauschke, X. Wang, and L. Yao. On Borwein–Wiersma decompositions of monotone linear relations. *SIAM Journal on Optimization*, 20 (2010), 2636–2652.

[4] J. M. Borwein and J. D. Vanderwerff. Convex Functions. Cambridge University Press, Cambridge, MA, 2010.

[5] R. E. Burck. An iterative solution of a variational inequality for certain monotone operators in Hilbert space. *Bulletin of the American Mathematical Society*, 81 (1975), 890–892.

[6] R. S. Burachik and A. N. Iusem. Set-valued mappings and enlargement of monotone operators. Springer, New York, 2009.

[7] J. V. Burke and P. Tseng. A unified analysis of Hoffman's bound via Fenchel duality. *SIAM Journal on Optimization*, 6 (1996), 265–282,.

[8] M. J. Cánovas, A. L. Dontchev, M. A. López, and J. Parra. Metric regularity of semi-infinite constraint systems. *Mathematical Programming*, 104 (2005), 329-–346.

[9] M. J. Cánovas, F. J. Gómez-Senent, and J. Parra. Regularity modulus of arbitrarily perturbed linear inequality systems. *Journal of Mathematical Analysis and Applications*, 343 (2008), 315–327.

[10] M. J. Cánovas, R. Henrion, M. A. López, and J. Parra. Outer limit of subdifferentials and calmness moduli in linear and nonlinear programming. *Journal of Optimization Theory and Applications*, 169 (2016), 925-–952.

[11] M. J. Cánovas, M. A. López, J. Parra, and F. J. Toledo. Calmness of the feasible set mapping for linear inequality systems. *Set-Valued and Variational Analysis*, 22 (2014), 375–389.

[12] Y. Censor, A. N. Iusem, and S. A. Zenios. An interior point method with Bregman functions for the variational inequality problem with paramonotone operators. *Mathematical Programming Series A*, 81 (1998), 373–400.

[13] A. L. Dontchev, H. Gfrerer, A. Y. Kruger, and J. V. Outrata. The radius of metric subregularity. *Set-Valued and Variational Analysis*, 28 (2020), 451—473.

[14] A. L. Dontchev, A. Lewis, and R. T. Rockafellar. The radius of metric regularity. *Transactions of the American Mathematical Society*, 355 (2002), 493-–517.

[15] A. L. Dontchev and R. T. Rockafellar. Implicit Functions and Solution Mappings: A View from Variational Analysis. Springer, New York, 2009.

[16] S. Gass and T. Saaty. Parametric objective function (Part 2)-Generalization. *Journal of the Operations Research Society of America*, 3 (1955), 395–401.

[17] M. A. Goberna and M. A. López. Linear Semi-Infinite Optimization. John Wiley & Sons, Chichester (UK), 1998.

[18] A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49 (1952), 263–265.

[19] A. N. Iusem. On some properties of paramonotone operators. *Journal of Convex Analysis*, 5 (1998), 269–278.

[20] D. Klatte and B. Kummer. Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications. *Nonconvex Optimization and its Applications*, 60, Kluwer Academic, Dordrecht, The Netherlands, 2002.

[21] D. Klatte and G. Thiere. Error bounds for solutions of linear equations and inequalities. *ZOR Mathematical Methods of Operations Research*, 41 (1996), 191–214.

[22] A. Kruger, H. V. Ngai, and M. Théra. Stability of error bounds for convex constraint systems in Banach spaces. *SIAM Journal on Optimization*, 20 (2010), 3280–3296.

[23] A. Y. Kruger, M. A. López, and M. A. Théra. Perturbation of error bounds. *Mathematical Programming Series B*, 168 (2018), 533-–554.

[24] W. LI. Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs. *SIAM Journal on Control and Optimization*, 32 (1994), 140–153.

[25] M. H. Li, K. W. Meng, and X. Q. Yang. On error bound moduli for locally Lipschitz and regular functions. *Mathematical Programming Series A*, 171 (2018), 463–487.

[26] B. S. Mordukhovich. Variational Analysis and Generalized Differentiation, I: Basic Theory. Springer, Berlin, 2006.

[27] J. Peña, J. C. Vera, and L. F. Zuluaga. New characterizations of Hoffman constants for systems of linear constraints. *Mathematical Programming*, 187 (2021), 79–109.

[28] R. R. Phelps. Convex Functions, Monotone Operators and Differentiability. 2nd ed., Springer-Verlag, Berlin, 1993.

[29] R. T. Rockafellar and R. J-B. Wets. Variational Analysis. Springer, Berlin, 1998.

[30] T. Saaty and S. Gass. Parametric objective function (Part 1). *Journal of the Operations Research Society of America*, 2 (1954), 316–319.

[31] S. Simons. From Hahn-Banach to Monotonicity. Springer-Verlag, Berlin, 2008.

[32] C. Zălinezcu. Convex Analysis in General Vector Spaces. World Scientific Publishing, Singapore, 2002.

# Appendices

# A From calmness to Hoffman constants for linear semi-infinite inequality systems.

SIAM J. OPTIM.
Vol. 32, No. 4, pp. 2859–2878





# FROM CALMNESS TO HOFFMAN CONSTANTS FOR LINEAR SEMI-INFINITE INEQUALITY SYSTEMS*

J. CAMACHO†, M. J. CÁNOVAS†, AND J. PARRA†

**Abstract.** In this paper we focus on different—global, semilocal, and local—versions of Hoffman-type inequalities expressed in a variational form. In a first stage our analysis is developed for generic multifunctions between metric spaces, and we finally deal with the feasible set mapping associated with linear semi-infinite inequality systems (finitely many variables and possibly infinitely many constraints) parameterized by their right-hand sides. The Hoffman modulus is shown to coincide with the Lipschitz upper semicontinuity modulus and the supremum of calmness moduli when confined to multifunctions with a convex graph and closed images in a reflexive Banach space, which is the case for our feasible set mapping. Moreover, for this particular multifunction a formula—involving only the system's left-hand side—of the global Hoffman constant is derived, providing a generalization to our semi-infinite context of finite counterparts developed in the literature. In the particular case of locally polyhedral systems, the paper also provides a point-based formula for the (semilocal) Hoffman modulus in terms of the calmness moduli at certain feasible points (extreme points when the nominal feasible set contains no lines), yielding a practically tractable expression for finite systems.

**Key words.** Hoffman constants, Lipschitz upper semicontinuity, calmness, linear inequality systems, feasible set mapping

**MSC codes.** 90C31, 49J53, 90C34, 15A39, 90C05

**DOI.** 10.1137/21M1418228

**1. Introduction.** Concerning finite linear inequality systems parameterized by their right-hand sides, the celebrated Hoffman lemma [10] is a result of *global* nature as far as it works for *any* parameter making the system consistent and *any* point of the Euclidean space. We can also find in the literature related *semilocal* results as far as they work around a nominal (given) parameter and any point in the Euclidean space, leading to the concept of a Hoffman constant at this parameter (see, e.g., Azé and Corvellec [2] and Zălinescu [27]). In this paper we relate these global and semilocal Hoffman constants with the *local* concept of a *calmness modulus*, which involves parameters and points, both around nominal ones. Our analysis is developed in a first step in the context of generic multifunctions then subsequently moves to the particular case of the *feasible set mapping* associated with a parameterized linear semi-infinite inequality system

$$\sigma(b) := \{a_t'x \leq b_t,\ \ t \in T\}, \tag{1}$$

where $T$ is a compact metric space, $t \mapsto a_t \in \mathbb{R}^n$ is a fixed continuous function from $T$ to $\mathbb{R}^n$, and $b \equiv (b_t)_{t\in T} \in C(T,\mathbb{R})$ is the parameter to be perturbed, $C(T,\mathbb{R})$ being the space of continuous functions from $T$ to $\mathbb{R}$. We are considering column-vectors, and the prime stands for transposition, so $x'y$ denotes the usual inner product of $x$ and $y$ in $\mathbb{R}^n$. In this parametric context, the feasible set mapping $\mathcal{F} : C(T,\mathbb{R}) \rightrightarrows \mathbb{R}^n$

*Received by the editors May 6, 2021; accepted for publication (in revised form) July 12, 2022; published electronically November 17, 2022.
https://doi.org/10.1137/21M1418228
**Funding:** This work was partially supported by grant PGC2018-097960-B-C21 from MICINN, Spain, and ERDF, "A Way to Make Europe," European Union, and by grant PROMETEO/2021/063 from Generalitat Valenciana, Spain.

†Center of Operations Research, Miguel Hernández University of Elche, 03202 Elche (Alicante), Spain (j.camacho@umh.es, canovas@umh.es, parra@umh.es).

is given by

$$\mathcal{F}(b) := \{x \in \mathbb{R}^n \mid a_t' x \leq b_t,\ \ t \in T\}. \tag{2}$$

With respect to the topology, $\mathbb{R}^n$ is equipped with an arbitrary norm, $\|\cdot\|$, with *dual norm* given by $\|u\|_* = \max_{\|x\|\leq 1} |u'x|$, and the parameter space $C(T, \mathbb{R})$ is endowed with the supremum norm $\|b\|_\infty := \max_{t\in T} |b_t|$.

The particular case when $T$ is finite is included in this framework, in which case $\mathcal{F}$ coincides with the polyhedral mapping considered in [10], and the Hoffman lemma reads as the existence of some constant $\kappa \geq 0$ such that, for all $x \in \mathbb{R}^n$ and all $b \in \operatorname{dom} \mathcal{F}$ (the domain of $\mathcal{F}$),

$$d(x, \mathcal{F}(b)) \leq \kappa \max_{t\in T} [a_t' x - b_t]_+, \tag{3}$$

where $[\alpha]_+ := \max\{\alpha, 0\}$ is the positive part of $\alpha \in \mathbb{R}$. This result is of a global nature as far as it involves all points $x \in \mathbb{R}^n$ and all $b \in \operatorname{dom} \mathcal{F}$. Since $\max_{t\in T} [a_t' x - b_t]_+ = d\left(b, \mathcal{F}^{-1}(x)\right)$, inequality (3) can be written in a variational form, as is done in the following paragraph for a generic multifunction.

Given a multifunction $\mathcal{M} : Y \rightrightarrows X$ between metric spaces with both distances being denoted by $d$, we say that the (global) Hoffman property holds if there exists a constant $\kappa \geq 0$ such that

$$d(x, \mathcal{M}(y)) \leq \kappa d\left(y, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in X \text{ and all } y \in \operatorname{dom} \mathcal{M}, \tag{4}$$

where $d(x, \Omega) := \inf\{d(x, \omega) \mid \omega \in \Omega\}$ for $x \in X$ and $\Omega \subset X$, with $\inf \emptyset := +\infty$, so that $d(x, \emptyset) = +\infty$. Since this paper is concerned with nonnegative constants, we use the convention $\sup \emptyset := 0$. Here $\operatorname{dom} \mathcal{M}$ is the domain of $\mathcal{M}$ (recall that $y \in \operatorname{dom} \mathcal{M} \Leftrightarrow \mathcal{M}(y) \neq \emptyset$) and $\mathcal{M}^{-1}$ denotes the inverse mapping of $\mathcal{M}$ (i.e., $y \in \mathcal{M}^{-1}(x) \Leftrightarrow x \in \mathcal{M}(y)$).

Now we write a semilocal version of (4) by fixing $y = \overline{y}$. $\mathcal{M}$ is said to be *Hoffman stable at* $\overline{y} \in \operatorname{dom} \mathcal{M}$ if there exists $\kappa \geq 0$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\left(\overline{y}, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in X. \tag{5}$$

When inequality (5) is only required to be satisfied in a neighborhood of $\overline{x} \in \mathcal{M}(\overline{y})$ we are dealing with the *calmness* of $\mathcal{M}$ at $(\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}$, the graph of $\mathcal{M}$. Formally, the calmness of $\mathcal{M}$ at $(\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}$, or equivalently the *metric subregularity* of $\mathcal{M}^{-1}$ at $(\overline{x}, \overline{y})$ (cf. [7, Theorem 3H.3 and Exercise 3H.4]), is satisfied when there exist a constant $\kappa \geq 0$ and a neighborhood $U$ of $\overline{x}$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\left(\overline{y}, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in U. \tag{6}$$

The infimum of constants $\kappa$ appearing in (4), (5), and (6) are called, respectively, the *global Hoffman constant* of $\mathcal{M}$, the *Hoffman modulus* of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$, and the *calmness modulus* of $\mathcal{M}$ at $(\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}$. The three constants are denoted, respectively, by $\operatorname{Hof} \mathcal{M}$, $\operatorname{Hof} \mathcal{M}(\overline{y})$, and $\operatorname{clm} \mathcal{M}(\overline{y}, \overline{x})$ and, as a consequence of the definitions, they may be written as follows:

$$\begin{aligned}
&\operatorname{Hof} \mathcal{M} = \sup_{(y,x)\in(\operatorname{dom} \mathcal{M})\times X} \frac{d(x, \mathcal{M}(y))}{d\left(y, \mathcal{M}^{-1}(x)\right)}, \\
&\operatorname{Hof} \mathcal{M}(\overline{y}) = \sup_{x\in X} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)},\ \overline{y} \in \operatorname{dom} \mathcal{M}, \\
&\operatorname{clm} \mathcal{M}(\overline{y}, \overline{x}) = \limsup_{x\to\overline{x}} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)},\ \ (\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M},
\end{aligned} \tag{7}$$

under the convention $\frac{0}{0} := 0$, where $\limsup$ is understood as the supremum (indeed, maximum) of all possible sequential upper limits (i.e., with $(y, x)$ being replaced with elements of sequences $\{(y_r, x_r)\}_{r\in\mathbb{N}}$ converging to $(\overline{y}, \overline{x})$ as $r \to \infty$).

Now we describe the main contributions of the paper. Clearly

$$\operatorname{Hof} \mathcal{M} = \sup_{\overline{y}\in\operatorname{dom}\mathcal{M}} \operatorname{Hof} \mathcal{M}(\overline{y}),$$

and we wonder if a similar relationship between $\operatorname{Hof}\mathcal{M}(\overline{y})$ and the supremum of all calmness moduli $\operatorname{clm}\mathcal{M}(\overline{y}, x)$, with $x \in \mathcal{M}(\overline{y})$, works. Section 3 is devoted to this question, and Theorem 4 gives a positive answer when $\operatorname{gph}\mathcal{M}$ is convex and $\mathcal{M}(\overline{y})$ is closed, $Y$ being a normed space and $X$ being a reflexive Banach space. Some examples show that the convexity assumption is not superfluous. Moreover, some intermediate constants such as the *Lipschitz upper semicontinuity modulus* are also considered.

With respect to mapping $\mathcal{F}$ our focus is on formulae only involving the system's coefficients for $\operatorname{Hof}\mathcal{F}$ and $\operatorname{Hof}\mathcal{F}(\overline{b})$, which are established in Theorems 5 and 6, respectively. The first one extends to the current semi-infinite framework previous results on finite linear systems (see, e.g., Burke and Tseng [4, Theorem 8], Klatte and Thiere [13, Theorem 2.7], and Peña, Vera, and Zuluaga [18, Formula (3)]); for comparative purposes, some details are gathered in section 2. Theorem 6 provides a formula for $\operatorname{Hof}\mathcal{F}(\overline{b})$ in terms of the $a_t$'s, the $\overline{b}_t$'s, and some feasible points in the case when our system $\sigma\left(\overline{b}\right)$ is *locally polyhedral.* Specifically, from the referred Theorem 4, we have that

$$\operatorname{Hof}\mathcal{F}(\overline{b}) = \sup_{x\in\mathcal{F}(\overline{b})} \operatorname{clm}\mathcal{F}\left(\overline{b}, x\right),$$

and Theorem 6 refines this expression by reducing the supremum to a smaller set (which turns out to be finite when $T$ also is). Then, making use of the expression for $\operatorname{clm}\mathcal{F}\left(\overline{b}, \overline{x}\right)$ established in Li, Meng, and Yang [16] (recalled in Theorem 3), we derive the announced point-based formula for $\operatorname{Hof}\mathcal{F}(\overline{b})$. Here we use the term "point-based" to emphasize the fact that the expression for $\operatorname{Hof}\mathcal{F}(\overline{b})$ does not involve parameters different from $\overline{b}$ or points outside $\mathcal{F}(\overline{b})$. An alternative expression for $\operatorname{Hof}\mathcal{F}(\overline{b})$ appealing to points outside $\mathcal{F}(\overline{b})$ is given in [2, Theorem 2.6] (recalled in Theorem 2). We point out the fact that Theorem 6 yields a particularly tractable procedure for computing $\operatorname{Hof}\mathcal{F}(\overline{b})$ when $T$ is finite.

In summary, the structure of the paper is as follows: Section 2 introduces the necessary notation and gathers some preliminary results. Section 3 analyzes the relationships among different semilocal versions of Hoffman- and Lipschitz-type properties for generic multifunctions and their moduli (Lipschitz-type properties are widely analyzed in the monographs [7, 12, 17, 22]). Section 3 also provides illustrative counterexamples. Section 4 is focused on $\operatorname{Hof}\mathcal{F}$ and $\operatorname{Hof}\mathcal{F}(\overline{b})$, the latter in the case of locally polyhedral systems. Before establishing the announced formula for $\operatorname{Hof}\mathcal{F}(\overline{b})$ some technical geometrical results are proved. The paper finishes with a short section of conclusions and perspectives.

**2. Preliminaries.** Given $S \subset \mathbb{R}^k$, $k \in \mathbb{N}$, we denote by $\operatorname{conv}S$, $\operatorname{cone}S$, and $\operatorname{span}S$ the *convex hull,* the *conical convex hull*, and the *linear hull* of $S$, respectively. It is assumed that $\operatorname{cone}S$ always contains the zero-vector $0_k$, in particular $\operatorname{cone}(\emptyset) = \{0_k\}$.

Moreover, $S^\circ$ denotes the (negative) *polar* of $S$ given by

$$S^\circ := \left\{u \in \mathbb{R}^k \mid u'x \leq 0 \text{ for all } x \in S\right\}$$

($S^\circ = \mathbb{R}^k$ if $S = \emptyset$). From the topological side, $\mathrm{int}S$, $\mathrm{cl}S$, and $\mathrm{bd}S$ stand, respectively, for the (topological) interior, closure, and boundary of $S$. For a nonempty convex set $C \subset \mathbb{R}^k$, $O^+C$ denotes its *recession cone* given by

$$O^+C := \left\{d \in \mathbb{R}^k \mid u + \alpha d \in C \text{ for all } u \in C \text{ and all } \alpha \geq 0\right\},$$

while $\mathrm{end}\, C$ denotes its end set (introduced in [11]) defined as

$$\mathrm{end}\, C := \{u \in \mathrm{cl}\, C \mid \nexists \mu > 1 \text{ such that } \mu u \in \mathrm{cl}\, C\}.$$

Moreover, $\mathrm{extr}C$ stands for the set of extreme points of $C$. Recall that $x \in \mathrm{extr}C$ if $x \in C$ and it cannot be expressed as a convex combination of two points of $C \backslash \{x\}$. In any metric space $(Z, d)$, the *closed* ball centered at $z \in Z$ with radius $r > 0$ is denoted by $B(z, r)$, whereas $B(S, r) := \{z \in Z \mid d(z, S) \leq r\}$, for $S \subset Z$, denotes the $r$*-enlargement* of $S$.

For comparative purposes, the next theorem gathers some results in the literature on $\mathrm{Hof}\, \mathcal{F}$ when confined to finite linear systems, where $C(T, \mathbb{R}) \equiv \mathbb{R}^m$ for some $m \in \mathbb{N}$. It is adapted to our current notation and to our choice of norms. The first two expressions come from [18, formulae (3) and (4)] (see also [13, Theorem 2.7] when $\mathbb{R}^n$ is endowed with the Euclidean norm), while the third one can be derived from [4, Theorem 8], where a dual approach is followed. The last one appeals to the set

$$W_2 := \left\{y \in \mathbb{R}^m_+ \mid \{a_t,\ t \in \mathrm{supp}\,(y)\} \text{ lin. indep.}\right\},$$

where $\mathbb{R}^m_+$ is formed by the vectors of $\mathbb{R}^m$ having nonnegative coordinates and $\mathrm{supp}\,(y) := \{t \in \{1, \ldots, m\} \mid y_t \neq 0\}$ is the support of $y$; indeed $W_2$ is considered a subset of the dual space of $\mathbb{R}^m$, which we are identifying with $\mathbb{R}^m$ itself.

THEOREM 1. *Consider the feasible set mapping $\mathcal{F}$ defined in* (2) *and assume that $T$ is finite. We have*

$$\mathrm{Hof}\, \mathcal{F} = \max_{\substack{J \subset T \\ 0_n \notin \mathrm{conv}\{a_t,\ t \in J\}}} d_*\left(0_n, \mathrm{conv}\{a_t,\ t \in J\}\right)^{-1} \tag{8}$$

$$= \max_{\substack{J \subset T,\ \mathrm{rank}\, A_J = \mathrm{rank}\, A \\ \{a_t,\ t \in J\} \text{ lin. indep.}}} d_*\left(0_n, \mathrm{conv}\{a_t,\ t \in J\}\right)^{-1} \tag{9}$$

$$= \sup\left\{\|y\|_1 \mid y \in W_2,\ \|A'y\|_* = 1\right\}, \tag{10}$$

*where $A_J$ and $A$ stand for the matrices whose rows are $a_t'$, with $t \in J$ and $t \in T$, respectively, and $d_*$ stands for the distance associated with the dual norm $\|\cdot\|_*$.*

*Proof.* According to [18, formula (3)] and the subsequent comments therein, to establish (8) we only have to prove that condition $0_n \notin \mathrm{conv}\{a_t,\ t \in J\}$ is equivalent to the consistency of system $\{a_t'x < 0,\ t \in J\}$, and this follows, for instance, from equivalence (iv) $\Leftrightarrow$ (v) in [9, Theorem 6.1]. Equality (9) comes from [18, formula (4)] with the trivial observation that instead of all linearly independent $\{a_t,\ t \in J\}$, with $J \subset T$, we can confine ourselves to those which are maximal with respect to the inclusion order. Indeed, the result also follows from (8), since the sufficiency of considering those $\{a_t,\ t \in J\}$ which are linearly independent comes from [2, Lemma 3.1].

Formula (10) comes from [4, Theorem 8]. Let us comment that we can, alternatively, see the relationship between the second and the third expression by observing that, for any $y \in \mathbb{R}_+^m$, $y \neq 0_m$,

$$\frac{1}{\|y\|_1} A' y = \frac{1}{\|y\|_1} \sum_{i=1}^{m} y_t a_t \in \operatorname{conv}\{a_t,\ t \in \operatorname{supp}(y)\},$$

and that $\|A'y\|_* = 1$ is equivalent to $\|y\|_1 = \left\| \frac{1}{\|y\|_1} A'y \right\|_*^{-1}$. □

Generalizations of Hoffman constants to infinite-dimensional spaces or to convex functions playing the role of the distance function can be found in [4]. Many other authors have contributed to the study of Hoffman constants and their relationship with other concepts (as Lipschitz constants). Additional references can be obtained from the reference list of the papers mentioned above as well as [2] and [27], among others. At this moment we also cite Belousov and Andronov [3], Li [15], and Robinson [20].

The following theorem provides formulae for $\operatorname{Hof}\mathcal{F}(\overline{b})$, with $\overline{b} \in \operatorname{dom}\mathcal{F}$, and $\operatorname{clm}\mathcal{F}\left(\overline{b}, \overline{x}\right)$, with $\left(\overline{b}, \overline{x}\right) \in \operatorname{gph}\mathcal{F}$ through points outside $\mathcal{F}(\overline{b})$. They appeal to the supremum function $f_b : \mathbb{R}^n \to \mathbb{R}$, with $b \in C(T, \mathbb{R})$, given by

$$f_b(x) := \sup_{t \in T} (a_t' x - b_t) \quad \text{for } x \in \mathbb{R}^n,$$

which is known to be convex on $\mathbb{R}^n$. For each $x \in \mathbb{R}^n$, we consider the subset of indices

$$J_b(x) = \{t \in T \mid a_t' x - b_t = f_b(x)\}.$$

The well-known Valadier's formula works by virtue of the Ioffe–Tikhomirov theorem (see, e.g., [26, Theorem 2.4.18]), yielding

$$\partial f_b(x) = \operatorname{conv}\{a_t,\ t \in J_b(x)\},$$

where $\partial f_b(x)$ stands for the usual subdifferential of convex analysis (see, e.g., [21]).

THEOREM 2. *The following statements hold:*

(i) [2, Theorem 2.6] *For any* $\overline{b} \in \operatorname{dom}\mathcal{F}$, *one has*

$$\begin{aligned}\operatorname{Hof}\mathcal{F}(\overline{b}) &= \sup_{f_{\overline{b}}(x) > 0} d_*\left(0_n, \partial f_{\overline{b}}(x)\right)^{-1} \\ &= \sup_{f_{\overline{b}}(x) > 0} d_*\left(0_n, \operatorname{conv}\left\{a_t,\ t \in J_{\overline{b}}(x)\right\}\right)^{-1}.\end{aligned}$$

(ii) [14, Theorem 1] *For any* $\left(\overline{b}, \overline{x}\right) \in \operatorname{gph}\mathcal{F}$,

$$\begin{aligned}\operatorname{clm}\mathcal{F}(\overline{b}, \overline{x}) &= \limsup_{x \to \overline{x},\ f_{\overline{b}}(x) > 0} d_*\left(0_n, \partial f_{\overline{b}}(x)\right)^{-1} \\ &= \limsup_{x \to \overline{x},\ f_{\overline{b}}(x) > 0} d_*\left(0_n, \operatorname{conv}\left\{a_t,\ t \in J_{\overline{b}}(x)\right\}\right)^{-1}.\end{aligned}$$

*Remark* 1. Observe that $\overline{b} \in \operatorname{dom} \mathcal{F}$ and $f_{\overline{b}}(x) > 0$ mean that $\sigma\left(\overline{b}\right)$ is consistent (it has some feasible solution) but $x \notin \mathcal{F}(\overline{b})$); in this case, $0_n \notin \operatorname{conv}\{a_t,\ t \in J_{\overline{b}}(x)\}$, since $x$ is not a global minimizer of the convex function $f_{\overline{b}}$. Actually, [2, Theorem 2.6] is formulated in terms of $(\operatorname{Hof} \mathcal{F}(\overline{b}))^{-1}$, which is called there the *condition number of* $f_{\overline{b}}$ *at level* 0*;* in the terminology of [14], observe that $(\operatorname{clm}\mathcal{F}(\overline{b}, \overline{x}))^{-1}$ is the *error bound modulus* (also known as *conditioning rate* [19]) *of* $f_{\overline{b}}$ *at* $\overline{x}$.

The following theorem is devoted to the computation of $\operatorname{clm}\mathcal{F}(\overline{b}, \overline{x})$, $(\overline{b}, \overline{x}) \in \operatorname{gph} \mathcal{F}$, through a point-based formula (expressed exclusively in terms of the system's coefficients and the nominal point $\overline{x}$). Now we introduce some extra notation. Given a fixed $\overline{b} \in \operatorname{dom} \mathcal{F}$, for any $x \in \mathcal{F}(\overline{b})$, we consider (for simplicity, since there will be no ambiguity, we omit the dependence on $\overline{b}$)

$$T(x) := \left\{t \in T \mid a_t' x - \overline{b}_t = 0\right\},$$

the *subset of active indices* of system $\sigma\left(\overline{b}\right)$ at $x$; i.e., $T(x) = J_{\overline{b}}(x)$ if $f_{\overline{b}}(x) = 0$, while $T(x) = \emptyset$ if $f_{\overline{b}}(x) < 0$ (i.e., if $x$ is a strict solution—*Slater point*—of the system). Let $A(x)$ be the corresponding *active cone* at $x$; i.e.,

$$A(x) := \operatorname{cone}\{a_t,\ t \in T(x)\}$$

(recall that $A(x) = \{0_n\}$ if $T(x) = \emptyset$). We also consider the family $\mathcal{D}(x)$ of subsets $D \subset T(x)$ such that system

$$\left\{\begin{array}{ll} a_t' d = 1, & t \in D, \\ a_t' d < 1, & t \in T(x) \setminus D \end{array}\right\}, \tag{11}$$

is consistent (in the variable $d \in \mathbb{R}^n$); i.e., $\{a_t, t \in D\}$ is contained in some hyperplane which leaves $\{0_n\} \cup \{a_t, t \in T(x) \setminus D\}$ on one of its two associated open half-spaces. With this notation, the next theorem generalizes the corresponding finite version established in [6, Theorem 4]. It appeals to the following *regularity condition* at $\overline{x}$: "There exists a neighborhood $W$ of $\overline{x}$ such that

$$\mathcal{F}(\overline{b}) \cap W = \left(\overline{x} + A(\overline{x})^\circ\right) \cap W.\text{"} \tag{12}$$

Observe that this condition is held at all points of polyhedral sets and, for instance, at the vertex of the ice cream cone.

THEOREM 3 ([16, Corollary 2.1, Remark 2.3, and Corollary 3.2]). *Let* $\overline{x} \in \mathcal{F}(\overline{b})$ *such that* $f_{\overline{b}}(\overline{x}) = 0$ *and assume that the regularity condition* (12) *is held at* $\overline{x}$*. Then*

$$\operatorname{clm} \mathcal{F}\left(\overline{b}, \overline{x}\right) = d_*\left(0_n, \operatorname{end}\partial f_{\overline{b}}(\overline{x})\right)^{-1} = \sup_{D \in \mathcal{D}(\overline{x})} d_*\left(0_n, \operatorname{conv}\{a_t,\ t \in D\}\right)^{-1}. \tag{13}$$

*Remark* 2. Although condition (12) is not superfluous for establishing the first equality in (13) as [16, Example 3.3] shows (see also Example 4), the second equality does work for semi-infinite systems (1) without any additional condition. Indeed, from [16, Corollary 2.1 and Remark 2.3] we can deduce

$$\cup_{D \in \mathcal{D}(\overline{x})} \operatorname{conv}\{a_t,\ t \in D\} \subset \operatorname{end}\partial f_{\overline{b}}(\overline{x}) \subset \operatorname{cl}\left(\cup_{D \in \mathcal{D}(\overline{x})} \operatorname{conv}\{a_t,\ t \in D\}\right). \tag{14}$$

**3. From calmness to Hoffman constants for a generic multifunction.** The purpose of this section is to analyze the relationship among different Hoffman-and Lipschitz-type properties, including the known Lipschitz upper semicontinuity that goes back to the classical work of Robinson [20]. At the beginning of this section $\mathcal{M} : Y \rightrightarrows X$ is a generic multifunction between metric spaces $Y$ and $X$. Later we will need further structure. To start with, observe that alternatively to (5) we can write the Hoffman stability of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$ in terms of the existence of $\kappa \geq 0$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\,(y, \overline{y}) \text{ for all } (y, x) \in \operatorname{gph} \mathcal{M},$$

while the calmness of $\mathcal{M}$ at $(\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}$, introduced in (6) in terms of the (equivalent) metric subregularity of $\mathcal{M}^{-1}$, is written as the existence of neighborhoods $V$ of $\overline{y}$ and $U$ of $\overline{x}$ along with a constant $\kappa \geq 0$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\,(y, \overline{y}) \text{ for all } (y, x) \in (V \times U) \cap \operatorname{gph} \mathcal{M}.$$

Moreover, the following equalities constitute well-known alternative expressions to (7) for the corresponding moduli

$$\text{(15)} \qquad \begin{aligned} \operatorname{Hof} \mathcal{M}(\overline{y}) &= \sup_{(y,x) \in \operatorname{gph} \mathcal{M}} \frac{d(x, \mathcal{M}(\overline{y}))}{d\,(y, \overline{y})}, \\ \operatorname{clm} \mathcal{M}\,(\overline{y}, \overline{x}) &= \limsup_{\substack{(y,x) \to (\overline{y}, \overline{x}) \\ (y,x) \in \operatorname{gph} \mathcal{M}}} \frac{d(x, \mathcal{M}(\overline{y}))}{d\,(y, \overline{y})}. \end{aligned}$$

Recall that $\mathcal{M}$ is said to be *Lipschitz upper semicontinuous* at $\overline{y} \in \operatorname{dom} \mathcal{M}$ if there exists a neighborhood $V$ of $\overline{y}$ along with a constant $\kappa \geq 0$ such that

$$\text{(16)} \qquad d(x, \mathcal{M}(\overline{y})) \leq \kappa d\,(y, \overline{y}) \text{ for all } (y, x) \in (V \times \mathbb{R}^n) \cap \operatorname{gph} \mathcal{M}.$$

Here we borrow the terminology from [12] or [24], although this property, introduced in [20] as *upper Lipschitz continuity*, has been also popularized as *outer Lipschitz continuity* (see [7]). Equivalently, (16) may be written as $e(\mathcal{M}(y), \mathcal{M}(\overline{y})) \leq \kappa d\,(y, \overline{y})$ for all $y \in V$, where $e\,(A, B) := \sup_{x \in A} d\,(x, B)$ is the *Hausdorff excess* of $A$ over $B$, with $A, B \subset X$. The associated *Lipschitz upper semicontinuity modulus*, denoted by $\operatorname{Lipusc} \mathcal{M}(\overline{y})$, is defined as the infimum of constants $\kappa$ satisfying (16) for some associated $V$.

In the next definition, given $\overline{y} \in \operatorname{dom} \mathcal{M}$ and $\varepsilon > 0$, the mapping $\mathcal{M}_\varepsilon : Y \rightrightarrows X$ is defined by

$$\mathcal{M}_\varepsilon\,(y) := \mathcal{M}\,(y) \cap B\,(\mathcal{M}\,(\overline{y})\,, \varepsilon) \text{ for } y \in Y.$$

(For simplicity in the notation we obviate the dependence of $\mathcal{M}_\varepsilon$ on $\overline{y}$.)

DEFINITION 1. *Given* $\overline{y} \in \operatorname{dom} \mathcal{M}$, *we say that* $\mathcal{M}$ *is* uniformly calm *at* $\overline{y}$ *if there exists a neighborhood* $V$ *of* $\overline{y}$ *along with* $\varepsilon > 0$ *and* $\kappa \geq 0$ *such that*

$$\text{(17)} \qquad d(x, \mathcal{M}(\overline{y})) \leq \kappa d\,(y, \overline{y}) \text{ for all } y \in V \text{ and all } x \in \mathcal{M}_\varepsilon\,(y)\,,$$

*or, equivalently, if* $\mathcal{M}_\varepsilon$ *is Lipschitz upper semicontinuous at* $\overline{y}$ *for some* $\varepsilon > 0$.

The corresponding modulus naturally appears. Specifically, we define the *modulus of uniform calmness of* $\mathcal{M}$ at $\overline{y}$, denoted by $\operatorname{uclm} \mathcal{M}(\overline{y})$, as the infimum of constants $\kappa$ satisfying (17) for some associated $V$ and $\varepsilon > 0$. It is straightforward to check that

$$\text{(18)} \qquad \operatorname{uclm} \mathcal{M}(\overline{y}) = \inf_{\varepsilon > 0} \operatorname{Lipusc} \mathcal{M}_\varepsilon(\overline{y}).$$

Roughly speaking, the uniform calmness of $\mathcal{M}$ at $\overline{y}$ entails the calmness of $\mathcal{M}$ at any $(\overline{y}, x)$ for all $x \in \mathcal{M}(\overline{y})$ with the same calmness constant $\kappa$, the same neighborhood $V$ of $\overline{y}$, and a common radius $\varepsilon$ for all neighborhoods of points $x \in \mathcal{M}(\overline{y})$, say $U_x := B(x, \varepsilon)$. Example 1 below shows that the calmness of $\mathcal{M}$ at $(\overline{y}, x)$ for all $x \in \mathcal{M}(\overline{y})$ does not ensure the uniform calmness of $\mathcal{M}$ at $\overline{y}$.

As it occurs with the calmness property, the uniform calmness turns out to be equivalent to a certain metric regularity–type property, showing that neighborhood $V$ in Definition 1 is redundant. The key fact is that points $x \in \mathcal{M}(y)$ which are required to satisfy (17) are those which are sufficiently close to $\mathcal{M}(\overline{y})$. This comment, which was already pointed out for polyhedral multifunctions in [20] (see the corollary after Proposition 1 therein), is formalized in the following proposition.

PROPOSITION 1. *Let $\overline{y} \in \operatorname{dom} \mathcal{M}$. For any $\kappa > 0$, the following conditions are equivalent:*

(i) *There exist a neighborhood $V$ of $\overline{y}$ and $\varepsilon > 0$ such that* (17) *holds.*

(ii) *There exists $\varepsilon > 0$ such that* (5) *holds when restricted to those $x \in B(\mathcal{M}(\overline{y}), \varepsilon)$.*

*Proof.* Let us establish the nontrivial implication "(i) $\Rightarrow$ (ii)." Consider $V$ and $\varepsilon$ as in statement (i). Take $\varepsilon_1 > 0$ such that $B(\overline{y}, \varepsilon_1) \subset V$ and define $\varepsilon_2 := \min\{\varepsilon, \kappa\varepsilon_1\} > 0$. Let us see that (ii) holds for $\varepsilon_2 > 0$. Take $x \in B(\mathcal{M}(\overline{y}), \varepsilon_2)$ and consider $y \in \mathcal{M}^{-1}(x)$. Now, we distinguish between two cases:

If $d(y, \overline{y}) \leq \varepsilon_1$, then $y \in V$ and, since we also have $x \in B(\mathcal{M}(\overline{y}), \varepsilon)$ (recall that $\varepsilon_2 \leq \varepsilon$), from (i) we conclude the aimed inequality $d(x, \mathcal{M}(\overline{y})) \leq \kappa d(y, \overline{y})$.

Otherwise, if $d(y, \overline{y}) \geq \varepsilon_1$, then $d(x, \mathcal{M}(\overline{y})) \leq \varepsilon_2 \leq \kappa\varepsilon_1 \leq \kappa d(y, \overline{y})$. ∎

*Remark* 3. The statement of Proposition 1 does not hold for $\kappa = 0$. To see this, take $\mathcal{M} : \mathbb{R} \longrightarrow \mathbb{R}$ (single-valued) given by $\mathcal{M}(y) := \max\{0, y - 1\}$ and let $\overline{y} = 0$. Clearly (i) holds for $V = ]-1, 1[$ and $\kappa = 0$, whereas (ii) works for $\varepsilon > 0$ if and only if $\kappa \geq \varepsilon / (1 + \varepsilon)$.

COROLLARY 1. *Let $\overline{y} \in \operatorname{dom} \mathcal{M}$. We have the following:*

(i) *$\mathcal{M}$ is uniformly calm at $\overline{y}$ if and only if there exist $\varepsilon > 0$ and $\kappa \geq 0$ such that*

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\left(\overline{y}, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in B(\mathcal{M}(\overline{y}), \varepsilon). \tag{19}$$

(ii) *The modulus of uniform calmness can be expressed as follows:*

$$\operatorname{uclm} \mathcal{M}(\overline{y}) = \inf\{\kappa \geq 0 \mid \exists \varepsilon > 0 \text{ such that } (19) \text{ holds}\}.$$

*Proof.* Both (i) and (ii) come from the fact that uniform calmness at $\overline{y}$ with associated elements $V, \varepsilon > 0$, and $\kappa \geq 0$ in (17) entails the same property with $V, \varepsilon > 0$, and $\widetilde{\kappa} > \kappa$. Hence the conclusions follow straightforwardly from Proposition 1. ∎

Next we provide characterizations of $\operatorname{Lipusc} \mathcal{M}(\overline{y})$ and $\operatorname{uclm} \mathcal{M}(\overline{y})$ in terms of certain upper limits, which allow for a better understanding of these concepts and a clear relationship among all moduli introduced in the paper.

PROPOSITION 2. *Let $\mathcal{M} : Y \rightrightarrows X$ be a multifunction between metric spaces, and let $\overline{y} \in \operatorname{dom} \mathcal{M}$. Then*

(i) $\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \limsup_{y \to \overline{y}} \left( \sup_{x \in \mathcal{M}(y)} \frac{d(x, \mathcal{M}(\overline{y}))}{d(y, \overline{y})} \right)$;

(ii) $\operatorname{uclm} \mathcal{M}(\overline{y}) = \limsup_{d(x, \mathcal{M}(\overline{y})) \to 0} \frac{d(x, \mathcal{M}(\overline{y}))}{d(\overline{y}, \mathcal{M}^{-1}(x))}$.

*Proof.* (i) For the sake of simplicity, let us denote by $s$ the right-hand side of (i) and

$$K := \{\kappa \geq 0 \mid \exists V \text{ neighborhood of } \overline{y} \text{ verifying (16)}\}. \tag{20}$$

We start by establishing inequality "$\leq$". Since $\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \inf K$, we can write $\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \lim_{r\to\infty} \kappa_r$ for some $\{\kappa_r\} \subset K$. For each $r$ take a neighborhood $V_r$ associated with $\kappa_r$ according to (20) and define

$$\overline{\kappa}_r := \sup_{y\in V_r\cap B(\overline{y},1/r)} \left( \sup_{x\in\mathcal{M}(y)} \frac{d(x,\mathcal{M}(\overline{y}))}{d\left(y,\overline{y}\right)} \right) \leq \kappa_r.$$

By definition $\overline{\kappa}_r \in K$, having $V_r \cap B\left(\overline{y}, 1/r\right)$ as an associated neighborhood, so that we have $\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \lim_{r\to\infty} \overline{\kappa}_r$.

Finally, for each $r$, consider any $y_r \in V_r \cap B\left(\overline{y}, 1/r\right)$ such that

$$\overline{\kappa}_r - \frac{1}{r} \leq \sup_{x\in\mathcal{M}(y_r)} \frac{d(x,\mathcal{M}(\overline{y}))}{d\left(y_r,\overline{y}\right)} \leq \overline{\kappa}_r.$$

Obviously, $\{y_r\}_{r\in\mathbb{N}}$ converges to $\overline{y}$, and then

$$\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \lim_{r\to\infty} \sup_{x\in\mathcal{M}(y_r)} \frac{d(x,\mathcal{M}(\overline{y}))}{d\left(y_r,\overline{y}\right)} \leq s.$$

In order to prove "$\geq$" in (i), we may assume the nontrivial case $s > 0$ and write

$$s = \lim_{r\to\infty} \sup_{x\in\mathcal{M}(\widetilde{y}_r)} \frac{d(x,\mathcal{M}(\overline{y}))}{d\left(\widetilde{y}_r,\overline{y}\right)}$$

for some $\{\widetilde{y}_r\}_{r\in\mathbb{N}}$ converging to $\overline{y}$. It is clear that we may replace $\{\widetilde{y}_r\}_{r\in\mathbb{N}}$ with a suitable subsequence (denoted as the whole sequence for simplicity) such that $\widetilde{y}_r \in V_r$, and then

$$s \leq \lim_{r\to\infty} \kappa_r = \operatorname{Lipusc} \mathcal{M}(\overline{y}).$$

(ii) The procedure is analogous to the previous one by considering

$$\widehat{K} = \{\kappa \geq 0 \mid \exists\varepsilon > 0 \text{ such that (19) holds}\}. \qquad \square$$

As a direct consequence of the expressions in (15) for $\operatorname{clm} \mathcal{M}\left(\overline{y}, \overline{x}\right)$ and $\operatorname{Hof} \mathcal{M}\left(\overline{y}\right)$, together with (18) and the previous proposition, we conclude the following corollary. Observe that the smaller $\varepsilon > 0$, the smaller $\operatorname{Lipusc} \mathcal{M}_\varepsilon\left(\overline{y}\right)$, and $\operatorname{Lipusc} \mathcal{M}\left(\overline{y}\right)$ corresponds to $\varepsilon = +\infty$.

COROLLARY 2. *Let $\overline{y} \in \operatorname{dom} \mathcal{M}$. We have*

$$\sup_{x\in\mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}\left(\overline{y}, x\right) \leq \operatorname{uclm} \mathcal{M}\left(\overline{y}\right) \leq \operatorname{Lipusc} \mathcal{M}\left(\overline{y}\right) \leq \operatorname{Hof} \mathcal{M}\left(\overline{y}\right). \tag{21}$$

*Remark* 4. The previous corollary yields (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv), where
(i) $\mathcal{M}$ is Hoffman stable at $\overline{y}$;
(ii) $\mathcal{M}$ is Lipschitz upper semicontinuous at $\overline{y}$;
(iii) $\mathcal{M}$ is uniformly calm at $\overline{y}$;
(iv) $\mathcal{M}$ is calm at every $(\overline{y}, x) \in \operatorname{gph} \mathcal{M}$.

The next three examples show that all converse implications in the previous remark may fail for a suitable multifunction.

*Example* 1. Let $\mathcal{M} : \mathbb{R} \rightrightarrows \mathbb{R}$ be given by $\mathcal{M}(y) = \{h_r(y),\ r \in \mathbb{N}\}$, where

$$h_r(y) = \begin{cases} r + y & \text{if} \quad y \leq \frac{1}{r}, \\ r + \frac{1}{r} + r\left(y - \frac{1}{r}\right) & \text{if} \quad y > \frac{1}{r}. \end{cases}$$

For $\overline{y} = 0$, it is easy to check that $\operatorname{clm} \mathcal{M}(\overline{y}, x) = 1$ for all $x \in \mathcal{M}(\overline{y})$. Hence, $\sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}(\overline{y}, x) = 1$. Nevertheless, it is impossible to find $\varepsilon > 0$ that meets the conditions for uniform calmness; i.e., $\operatorname{uclm} \mathcal{M}(\overline{y}) = +\infty$. More specifically, take $\varepsilon_r := r^{-1} + r^{-1/2}$ for all $r \in \mathbb{N}$, $r \geq 8$ (to ensure $\varepsilon_r < 1/2$) and consider $y_r := r^{-1} + r^{-3/2}$ and $x_r := h_r(y_r) = r + r^{-1} + r^{-1/2} \in \mathcal{M}_{\varepsilon_r}(y_r)$. Then

$$\frac{d(x_r, \mathcal{M}(0))}{d(y_r, 0)} = \frac{r^{-1} + r^{-1/2}}{r^{-1} + r^{-3/2}} \to +\infty \text{ as } r \to +\infty.$$

*Example* 2. Consider $\mathcal{M} : \mathbb{R} \longrightarrow \mathbb{R}$ (single-valued) given by $\mathcal{M}(y) = 0$ if $y \leq 0$ and $\mathcal{M}(y) = 1$ if $y > 0$. It is clear that $\mathcal{M}$ is uniformly calm at $\overline{y} = 0$ (take $\varepsilon = 1/2$) but not Lipschitz upper semicontinuous by just considering $y_r = 1/r$ for $r \in \mathbb{N}$.

*Example* 3. Let $\mathcal{M} : \mathbb{R} \rightrightarrows \mathbb{R}$ be given by

$$\mathcal{M}(y) = [0, 1] \text{ if } y < 0, \qquad \mathcal{M}(y) = [0, +\infty[ \text{ if } y \geq 0.$$

It is clear that $\mathcal{M}$ is Lipschitz upper semicontinuous, with zero modulus, at any $y \in \mathbb{R}$. Nevertheless, it is not Hoffman stable at any $\overline{y} < 0$.

The next theorem establishes that all inequalities in (21) become equalities under the convexity of $\operatorname{gph} \mathcal{M}$ together with the closedness of $\mathcal{M}(\overline{y})$, provided that $Y$ is a normed space and $X$ is a reflexive Banach space. As an obvious consequence, all properties in Remark 4 become equivalent in such a case. First, we include two lemmas.

LEMMA 1. *Let $X$ be a normed space and $\emptyset \neq C \subset X$ be a closed set. Take any $x \in X$ and assume that there exists a best approximation, $\overline{x}$, of $x$ in $C$. Then $\overline{x}$ is a best approximation of $x_\lambda := (1 - \lambda)\overline{x} + \lambda x$ in $C$ for all $\lambda \in [0, 1]$.*

*Proof.* Reasoning by contradiction, suppose that for some $\lambda \in [0, 1]$ there exists $\hat{x} \in C$ such that $\|\hat{x} - x_\lambda\| < \|\overline{x} - x_\lambda\|$. Then

$$\begin{aligned} \|\hat{x} - x\| &\leq \|\hat{x} - x_\lambda\| + \|x_\lambda - x\| < \|\overline{x} - x_\lambda\| + \|x_\lambda - x\| \\ &= \lambda \|\overline{x} - x\| + (1 - \lambda)\|\overline{x} - x\| = \|\overline{x} - x\|, \end{aligned}$$

which contradicts the fact that $\overline{x}$ is a best approximation of $x$ in $C$. □

In the next result $X$ is assumed to be a reflexive Banach space in order to ensure the existence of best approximations on nonempty closed convex sets; see, e.g., [26, Theorem 3.8.1].

LEMMA 2. *Let $\mathcal{M} : Y \rightrightarrows X$ be a multifunction between a normed space $Y$ and a reflexive Banach space $X$, and assume that* $\operatorname{gph} \mathcal{M}$ *is a nonempty convex set. Let $\overline{y} \in \operatorname{dom} \mathcal{M}$ and suppose that $\mathcal{M}(\overline{y})$ is closed. Consider any $(y, x) \in \operatorname{gph} \mathcal{M}$ and let $\overline{x}$ be a best approximation of $x$ in $\mathcal{M}(\overline{y})$. Then*

$$\frac{d(x, \mathcal{M}(\overline{y}))}{d(y, \overline{y})} \leq \operatorname{clm} \mathcal{M}(\overline{y}, \overline{x}).$$

*Proof.* By the convexity assumption, for each $\lambda \in [0,1]$,

$$(y_\lambda, x_\lambda) := (1-\lambda)(\overline{y}, \overline{x}) + \lambda (y, x) \in \operatorname{gph} \mathcal{M}.$$

According to Lemma 1, $\overline{x}$ is also a best approximation of $x_\lambda$ in $\mathcal{M}(\overline{y})$ for each $\lambda \in [0,1]$. Therefore,

$$\frac{d(x, \mathcal{M}(\overline{y}))}{d(y, \overline{y})} = \frac{\|x - \overline{x}\|}{\|y - \overline{y}\|} = \frac{\|x_\lambda - \overline{x}\|}{\|y_\lambda - \overline{y}\|} = \frac{d(x_\lambda, \mathcal{M}(\overline{y}))}{d(y_\lambda, \overline{y})} \quad \text{for all } \lambda \in ]0,1].$$

Since, letting $\lambda \to 0$, we have $(y_\lambda, x_\lambda) \to (\overline{y}, \overline{x})$, by the definition of the calmness modulus (recall (15)) we conclude that

$$\operatorname{clm} \mathcal{M}(\overline{y}, \overline{x}) \geq \limsup_{\lambda \to 0} \frac{d(x_\lambda, \mathcal{M}(\overline{y}))}{d(y_\lambda, \overline{y})} = \frac{d(x, \mathcal{M}(\overline{y}))}{d(y, \overline{y})}. \qquad \square$$

THEOREM 4. *Let $\mathcal{M} : Y \rightrightarrows X$, with $Y$ being a normed space and $X$ being a reflexive Banach space, and assume that* $\operatorname{gph}\mathcal{M}$ *is a nonempty convex set. Let $\overline{y} \in \operatorname{dom}\mathcal{M}$ with $\mathcal{M}(\overline{y})$ closed. Then one has*

$$\sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}(\overline{y}, x) = \operatorname{uclm} \mathcal{M}(\overline{y}) = \operatorname{Lipusc} \mathcal{M}(\overline{y}) = \operatorname{Hof} \mathcal{M}(\overline{y}).$$

*Proof.* We only have to prove $\operatorname{Hof} \mathcal{M}(\overline{y}) \leq \sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}(\overline{y}, x)$, according to (21).

Take any $(\widetilde{y}, \widetilde{x}) \in \operatorname{gph} \mathcal{M}$ and let $\overline{x}$ be a best approximation of $\widetilde{x}$ in $\mathcal{M}(\overline{y})$. Lemma 2 ensures that

$$\frac{d(\widetilde{x}, \mathcal{M}(\overline{y}))}{d(\widetilde{y}, \overline{y})} \leq \operatorname{clm} \mathcal{M}(\overline{y}, \overline{x}) \leq \sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}(\overline{y}, x).$$

Then, recalling (15), we conclude that

$$\operatorname{Hof} \mathcal{M}(\overline{y}) = \sup_{(\widetilde{y}, \widetilde{x}) \in \operatorname{gph} \mathcal{M}} \frac{d(\widetilde{x}, \mathcal{M}(\overline{y}))}{d(\widetilde{y}, \overline{y})} \leq \sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}(\overline{y}, x). \qquad \square$$

We finish this section by observing that the global Hoffman constant for the whole graph can be larger than the Hoffman modulus for a specific $\overline{y}$. Just consider $\mathcal{M} : \mathbb{R} \rightrightarrows \mathbb{R}$ given by

$$\mathcal{M}(y) = ]-\infty, y] \text{ if } y < 0, \qquad \mathcal{M}(y) = ]-\infty, 0] \text{ if } y \geq 0.$$

Then clearly $\operatorname{Hof} \mathcal{M}(\overline{y}) = 1$ if $\overline{y} < 0$ and $\operatorname{Hof} \mathcal{M}(\overline{y}) = 0$ if $\overline{y} \geq 0$, so that $\operatorname{Hof} \mathcal{M} = 1$.

## 4. Hoffman and calmness moduli for linear semi-infinite inequality systems.

This section aims to obtain expressions for $\operatorname{Hof} \mathcal{F}$ and $\operatorname{Hof} \mathcal{F}(\overline{b})$, $\overline{b} \in \operatorname{dom} \mathcal{F}$, in terms of the system's data. These expressions are established in Theorems 5 and 6, respectively. The first result generalizes Theorem 1 to the current semi-infinite framework, while the second provides an alternative expression to Theorem 2(i), via points inside $\mathcal{F}(\overline{b})$, for locally polyhedral systems. In the case of finite linear systems Theorem 6 is particularly useful as far as it establishes an implementable procedure for computing $\operatorname{Hof} \mathcal{F}(\overline{b})$.

THEOREM 5. *Consider* $\mathcal{F} : C(T, \mathbb{R}) \rightrightarrows \mathbb{R}^n$ *defined in* (2). *We have*

$$\operatorname{Hof} \mathcal{F} = \sup_{\substack{J \subset T \text{ compact} \\ 0_n \notin \operatorname{conv}\{a_t,\ t \in J\}}} d_* \left(0_n, \operatorname{conv}\{a_t,\ t \in J\}\right)^{-1}.$$

*Proof.* It is clear that $\operatorname{Hof} \mathcal{F} = \sup_{b \in \operatorname{dom} \mathcal{F}} \operatorname{Hof} \mathcal{F}(b)$, and applying Theorem 2 we have

$$\operatorname{Hof} \mathcal{F} = \sup_{b \in \operatorname{dom} \mathcal{F}} \sup_{x \notin \mathcal{F}(b)} d_* \left(0_n, \operatorname{conv}\{a_t,\ t \in J_b(x)\}\right)^{-1}. \tag{22}$$

Hence, inequality "$\leq$" comes from (22) taking into account that $b \in \operatorname{dom} \mathcal{F}$ and $x \notin \mathcal{F}(b)$ imply $0_n \notin \operatorname{conv}\{a_t,\ t \in J_b(x)\}$ (recall Remark 1). Also take into account that each $J_b(x)$ is compact since it is closed in $T$ as far as $J_b(x)$ is the preimage of $\{f_b(x)\}$ by the continuous function $t \mapsto a_t' x - b_t$.

Let us prove the converse inequality "$\geq$". Observe that for $J = \emptyset$ we have $d_* \left(0_n, \operatorname{conv}\{a_t,\ t \in J\}\right)^{-1} = d_* (0_n, \emptyset)^{-1} = 0$. Fix a nonempty compact set $\widehat{J} \subset T$ such that $0_n \notin \operatorname{conv}\left\{a_t,\ t \in \widehat{J}\right\}$ and let us define $\widehat{b} \in C(T, \mathbb{R})$ such that

$$\widehat{J} = J_{\widehat{b}}(\widehat{x}) \text{ for some } \widehat{x} \notin \mathcal{F}\left(\widehat{b}\right),\ \widehat{b} \in \operatorname{dom} \mathcal{F}.$$

First, by separation, since $0_n \notin \operatorname{conv}\left\{a_t,\ t \in \widehat{J}\right\}$, there exists $0_n \neq \widehat{x} \in \mathbb{R}^n$, such that

$$a_t' \widehat{x} \geq \widehat{x}' \widehat{x} \text{ for all } t \in \widehat{J},$$

where $\widehat{x}$ is the best approximation of $0_n$ in the compact set $\operatorname{conv}\left\{a_t,\ t \in \widehat{J}\right\}$ with respect to the Euclidean norm in $\mathbb{R}^n$. Define

$$\widehat{b}_t := \max\{a_t' \widehat{x}, \tfrac{1}{2} \widehat{x}' \widehat{x}\} - \varphi(t) \tfrac{1}{2} \widehat{x}' \widehat{x},\ t \in T,$$

where

$$\varphi(t) = 1 - d\left(t, \widehat{J}\right) \text{ for all } t \in T.$$

Observe that $\widehat{b} \in \operatorname{dom} \mathcal{F}$ since $\widehat{b}_t \geq \frac{1}{2}(1 - \varphi(t)) \widehat{x}' \widehat{x} \geq 0$ for all $t \in T$ and, for instance, $0_n \in \mathcal{F}(\widehat{b})$. On the other hand, $\widehat{x} \notin \mathcal{F}(\widehat{b})$ since

$$a_t' \widehat{x} - \widehat{b}_t = a_t' \widehat{x} - \left(a_t' \widehat{x} - \varphi(t) \tfrac{1}{2} \widehat{x}' \widehat{x}\right) = \tfrac{1}{2} \widehat{x}' \widehat{x} > 0 \text{ if } t \in \widehat{J}.$$

Finally, observe that

$$a_t' \widehat{x} - \widehat{b}_t \leq a_t' \widehat{x} - a_t' \widehat{x} + \varphi(t) \tfrac{1}{2} \widehat{x}' \widehat{x} < \tfrac{1}{2} \widehat{x}' \widehat{x} \text{ whenever } t \in T \setminus \widehat{J}.$$

So

$$\widehat{J} = \left\{t \in T \mid a_t' \widehat{x} - \widehat{b}_t = f_{\widehat{b}}(\widehat{x})\right\},$$

in other words, $\widehat{J} = J_{\widehat{b}}(\widehat{x})$, which finishes the proof. ∎

*Remark* 5. Theorem 5 is the only result in this paper which uses the fact that $T$ is assumed to be a compact *metric* space. The rest of the results work for $T$ being a compact Hausdorff space, which is the framework of the so-called *continuous systems* in [9].

The rest of this section is focused on $\operatorname{Hof}\mathcal{F}\left(\bar{b}\right)$, provided that $\bar{b}\in\operatorname{dom}\mathcal{F}$. To start with, as a consequence of Theorem 4, we always have

$$\operatorname{Hof}\mathcal{F}\left(\bar{b}\right)=\sup_{x\in\mathcal{F}\left(\bar{b}\right)}\operatorname{clm}\mathcal{F}\left(\bar{b},x\right)=\sup_{x\in\operatorname{bd}\mathcal{F}\left(\bar{b}\right)}\operatorname{clm}\mathcal{F}\left(\bar{b},x\right),\ \bar{b}\in\operatorname{dom}\mathcal{F}, \tag{23}$$

where the last equality comes from the fact that $\operatorname{clm}\mathcal{F}\left(\bar{b},x\right)=0$ when $x\in\operatorname{int}\mathcal{F}\left(\bar{b}\right)$ (the trivial case $\operatorname{bd}\mathcal{F}\left(\bar{b}\right)=\emptyset$, equivalently $\mathcal{F}\left(\bar{b}\right)=\mathbb{R}^n$, is included; recall $\sup\emptyset:=0$). From now on we are devoted to refining (23) by replacing $\operatorname{bd}\mathcal{F}\left(\bar{b}\right)$ with a smaller subset. The concluding result is Theorem 6. First, we establish some technical results.

Proposition 3. *Let $x^1,x^2\in\operatorname{bd}\mathcal{F}\left(\bar{b}\right)$ such that $T\left(x^1\right)\subset T\left(x^2\right)$. Then*
(i) $\operatorname{end}\partial f_{\bar{b}}\left(x^1\right)\subset\operatorname{end}\partial f_{\bar{b}}\left(x^2\right)$;
(ii) *if the regularity condition* (12) *is held at $x^i$, $i=1,2$, then*

$$\operatorname{clm}\mathcal{F}\left(\bar{b},x^1\right)\leq\operatorname{clm}\mathcal{F}\left(\bar{b},x^2\right).$$

*Proof.* (i) First, $x^i\in\operatorname{bd}\mathcal{F}\left(\bar{b}\right)$ implies $f_{\bar{b}}\left(x^i\right)=0$, and so $T\left(x^i\right)\neq\emptyset$, $i=1,2$, by the compactness of $T$ together with the continuity of $t\mapsto\binom{a_t}{b_t}$. Recall that $\partial f_{\bar{b}}\left(x^i\right)=\operatorname{conv}\left\{a_i,\ i\in T\left(x^i\right)\right\}$, $i=1,2$, and hence $\partial f_{\bar{b}}\left(x^1\right)\subset\partial f_{\bar{b}}\left(x^2\right)$.

Assume, arguing by contradiction, that there exists $a\in\operatorname{end}\partial f_{\bar{b}}\left(x^1\right)\backslash\operatorname{end}\partial f_{\bar{b}}\left(x^2\right)$. Since, by compactness, $\operatorname{end}\partial f_{\bar{b}}\left(x^1\right)\subset\partial f_{\bar{b}}\left(x^1\right)\subset\partial f_{\bar{b}}\left(x^2\right)$, we have $a\in\partial f_{\bar{b}}\left(x^2\right)\backslash\operatorname{end}\partial f_{\bar{b}}\left(x^2\right)$. Then we have $\lambda a\in\partial f_{\bar{b}}\left(x^2\right)$ for some $\lambda>1$ and can write

$$\lambda a=\sum_{t\in T(x^1)}\lambda_t a_t+\sum_{t\in T(x^2)\backslash T(x^1)}\lambda_t a_t \tag{24}$$

for some $\{\lambda_t\}_{t\in T(x^2)}\subset\mathbb{R}_+$ such that $\left\{\lambda_t\mid\lambda_t\neq0,\ t\in T\left(x^2\right)\right\}$ is a finite set.

On the other hand, consider $d:=x^1-x^2$ and observe that

$$\left\{\begin{array}{l} a_t'd=0,\ t\in T\left(x^1\right),\\ a_t'd=a_t'x^1-a_t'x^2<\bar{b}_t-\bar{b}_t=0,\ t\in T\left(x^2\right)\backslash T\left(x^1\right).\end{array}\right.$$

Then, multiplying (with the inner product) both members of (24) by $d$, we deduce

$$0=\lambda a'd=\sum_{t\in T(x^2)\backslash T(x^1)}\lambda_t a_t'd,$$

which yields $\lambda_t=0$ for all $t\in T\left(x^2\right)\backslash T\left(x^1\right)$. So, we attain the contradiction $\lambda a=\sum_{t\in T(x^1)}\lambda_t a_t\in\partial f_{\bar{b}}\left(x^1\right)$.

Statement (ii) follows straightforwardly from Theorem 3. ∎

The following example shows that the regularity condition assumed in statement (ii) of the previous proposition is not superfluous. The example comes from modifying Example 1 in [6] (revisited in [16, Example 3.3]).

*Example* 4. Let us consider the system, in $\mathbb{R}^2$ endowed with the Euclidean norm, given by

$$\sigma\left(\bar{b}\right):=\left\{\begin{array}{rl} t(\cos t)x_1+t(\sin t)x_2\leq t, & t\in[0,\pi],\\ x_1\leq1, & t=4,\\ -x_1-x_2\leq1, & t=5\end{array}\right\};$$

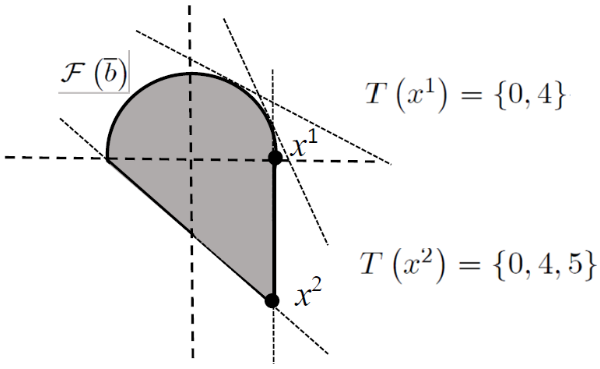


FIG. 1. *Illustration of Example* 4.

i.e., $T := [0,\pi] \cup \{4,5\}$, $a_t := t(\cos t, \sin t)'$ for $t \in [0,\pi]$, $a_4 := (1,0)'$, and $a_5 := (-1,-1)'$; $\overline{b} \in C([0,\pi] \cup \{4,5\}, \mathbb{R})$ is given by $\overline{b}_t = t$, $t \in [0,\pi]$, $\overline{b}_4 = 1$, and $\overline{b}_5 = 1$. Consider the feasible points $x^1 = (1,0)'$ and $x^2 = (1,-2)'$. The feasible set of $\sigma(\overline{b})$ is represented in Figure 1.

As proved in [6, Example 1], we have that

$$\operatorname{clm} \mathcal{F}(\overline{b}, x^1) = +\infty.$$

Alternatively, we can apply Theorem 2(ii) with sequence $x^r = \left(1 + \frac{1}{r}\right)\binom{\cos \frac{1}{r}}{\sin \frac{1}{r}}$. It is clear that the regularity condition (12) is not satisfied at $x^1$. Indeed $(0,1)' \in A(x^1)^\circ = \operatorname{cone}\{(1,0)'\}^\circ = \mathbb{R}_- \times \mathbb{R}$, but $x^1 + \varepsilon(0,1)' \notin \mathcal{F}(\overline{b})$ for any $\varepsilon > 0$. Moreover,

$$\partial f_{\overline{b}}(x^1) = \operatorname{conv}\{(0,0)', (1,0)'\}$$

and $\operatorname{end}\partial f_{\overline{b}}(x^1) = \{(1,0)'\}$. Hence, $\operatorname{clm} \mathcal{F}(\overline{b}, x^1) \neq d_*\left(0_2, \operatorname{end}\partial f_{\overline{b}}(x^1)\right)^{-1}$.

With respect to point $x^2$, one easily sees that condition (12) is satisfied, where $A(x^2)^\circ = \{u \in \mathbb{R}^2 \mid -u_1 - u_2 \leq 0, u_1 \leq 0\}$. In this case, $\partial f_{\overline{b}}(x^2) = \operatorname{conv}\{(0,0)', (1,0)', (-1,-1)'\}$. Hence, from Theorem 3 we have

$$\begin{aligned} \operatorname{clm} \mathcal{F}(\overline{b}, x^2) &= d_*\left(0_2, \operatorname{end}\partial f_{\overline{b}}(x^2)\right)^{-1} \\ &= d_*\left(0_2, \operatorname{conv}\{(1,0)', (-1,-1)'\}\right)^{-1} = \sqrt{5}. \end{aligned}$$

PROPOSITION 4. *Let $C$ be a nonempty closed convex subset of $\mathbb{R}^n$ different from a singleton with* $\operatorname{extr} C \neq \emptyset$ *and let* $x^0 \in C \backslash \operatorname{extr} C$. *Then there exist* $y^0 \in \operatorname{extr} C$, $z^0 \in C$, *and* $\mu \in ]0,1[$ *such that* $x^0 = (1-\mu) y^0 + \mu z^0$.

*Proof.* The assumption $\operatorname{extr} C \neq \emptyset$ is equivalent to the fact that $C$ contains no lines (i.e., its lineality space is $\{0_n\}$). According to [21, Corollary 14.6.1], this is also equivalent to $\operatorname{int}(O^+C)^\circ \neq \emptyset$, recalling that $O^+C$ is the recession cone of $C$. Pick $0_n \neq u \in \operatorname{int}(O^+C)^\circ$ and consider

$$K := C \cap \{x \in \mathbb{R}^n \mid u'x \geq u'x^0 - 1\}.$$

Let us see that $K$ is bounded, i.e., $O^+K = \{0_n\}$ (see [21, Theorem 8.4]). Reasoning by contradiction, assume the existence of $0_n \neq v \in O^+K$. Then $x^0 + \lambda v \in K$ and, accordingly, $u'\left(x^0+\lambda v\right) \geq u'x^0 - 1$ for all $\lambda > 0$. Letting $\lambda \to +\infty$ we obtain $u'v \geq 0$. On the other hand, $v \in O^+C$ and, for $\alpha > 0$ small enough, we have $u + \alpha v \in \left(O^+C\right)^\circ$, yielding the contradiction $0 \geq (u+\alpha v)' v \geq \alpha v'v$.

Once we know that $K$ is a nonempty convex compact set, by applying the Minkowski–Carathéodory theorem (see, e.g., [23, Theorem 8.11]), we have $K = \operatorname{conv}(\operatorname{extr} K)$ and can write

$$x^0 = \sum_{i=1}^{k} \lambda_i x^i \tag{25}$$

with $\left\{x^1,\ldots,x^k\right\} \subset \operatorname{extr} K$ being affinely independent, $\sum_{i=1}^k \lambda_i = 1$, and $\lambda_i > 0$ for all $i = 1,\ldots,k$. Clearly it is not restrictive to assume $u'x^1 \geq u'x^0$, which easily entails $x^1 \in \operatorname{extr} C$. More specifically, if $x^1$ were a midpoint of distinct points in $C$, we could replace these points with others in the same segment verifying $u'x \geq u'x^0 - 1$, and hence these points would be in $K$, contradicting $x^1 \in \operatorname{extr} K$.

On the other hand, by applying [9, Theorem A.7], (25) entails that $x^0$ is in the relative interior of conv $\left\{x^1,\ldots,x^k\right\}$ (i.e., the interior relative to the affine hull of these points), and then $z^0 := x^1 + \beta\left(x^0 - x^1\right) \in \operatorname{conv}\left\{x^1,\ldots,x^k\right\} \subset C$ for a small enough $\beta > 1$. Finally, let us write

$$x^0 = \left(1 - \tfrac{1}{\beta}\right) x^1 + \tfrac{1}{\beta} z^0,$$

which provides the desired result with $y^0 = x^1$ and $\mu = \frac{1}{\beta}$. ◻

The following theorem appeals to locally polyhedral (LOP, in brief) systems. Recall that given $\overline{b} \in \operatorname{dom} \mathcal{F}$, $\sigma\left(\overline{b}\right)$ is a LOP system if and only if

$$D\left(\mathcal{F}\left(\overline{b}\right), \overline{x}\right) = A(\overline{x})^\circ \quad \text{for all } \overline{x} \in \mathcal{F}\left(\overline{b}\right), \tag{26}$$

where $D\left(\mathcal{F}\left(\overline{b}\right), \overline{x}\right)$ denotes the *cone of feasible directions* of $\mathcal{F}\left(\overline{b}\right)$ at $\overline{x}$; i.e., $d \in D\left(\mathcal{F}\left(\overline{b}\right), \overline{x}\right)$ if there exists $\varepsilon > 0$ such that $\overline{x} + \alpha d \in \mathcal{F}\left(\overline{b}\right)$ for all $\alpha \in [0,\varepsilon]$. See [1] for a comprehensive analysis of LOP systems (see also [9]). At this moment we recall a characterization of LOP systems in terms of the regularity condition (26) which can be derived from Corollary 3.3 in [16].

LEMMA 3 (see [16, Corollary 3.3]). *Let $\overline{b} \in \operatorname{dom} \mathcal{F}$. The following conditions are equivalent:*

(i) $D\left(\mathcal{F}\left(\overline{b}\right), \overline{x}\right) = A(\overline{x})^\circ$ *for all* $\overline{x} \in \mathcal{F}\left(\overline{b}\right)$.

(ii) *The regularity condition* (12) *is held at any* $\overline{x} \in \mathcal{F}\left(\overline{b}\right)$.

From now on we consider the set

$$\mathcal{E}\left(\overline{b}\right) := \operatorname{extr}\left(\mathcal{F}\left(\overline{b}\right) \cap \operatorname{span}\{a_t,\ t \in T\}\right), \text{ with } \overline{b} \in \operatorname{dom} \mathcal{F}. \tag{27}$$

Observe that $\mathcal{E}\left(\overline{b}\right)$ is always a nonempty and finite set when $T$ is finite; moreover,

$$\mathcal{E}\left(\overline{b}\right) = \operatorname{extr} \mathcal{F}\left(\overline{b}\right) \Leftrightarrow \operatorname{extr} \mathcal{F}\left(\overline{b}\right) \neq \emptyset;$$

in fact, $\operatorname{extr} \mathcal{F}\left(\overline{b}\right) \neq \emptyset$ if and only if $\mathcal{F}\left(\overline{b}\right)$ does not contain any line, which is equivalent to the fact that $\operatorname{span}\{a_t,\ t \in T\} = \mathbb{R}^n$. This construction is inspired by that of [15, p. 142], and is used in [8] to compute the calmness modulus of the optimal value function of finite linear optimization problems.

THEOREM 6. *Let* $\overline{b} \in \operatorname{dom} \mathcal{F}$ *and assume that* $\sigma\left(\overline{b}\right)$ *is a LOP system. Then*

$$\operatorname{Hof} \mathcal{F}\left(\overline{b}\right) = \sup_{x \in \mathcal{E}\left(\overline{b}\right)} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right) = \sup_{x \in \mathcal{E}\left(\overline{b}\right)} \sup_{D \in \mathcal{D}(x)} d_*\left(0_n, \operatorname{conv}\{a_t,\ t \in D\}\right)^{-1}.$$

*Proof.* To start with, we recall (23):

$$\operatorname{Hof} \mathcal{F}\left(\overline{b}\right) = \sup_{x \in \operatorname{bd}\mathcal{F}\left(\overline{b}\right)} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right).$$

Since $\mathcal{E}\left(\overline{b}\right) \subset \operatorname{bd}\mathcal{F}\left(\overline{b}\right)$, the inequality $\operatorname{Hof} \mathcal{F}\left(\overline{b}\right) \geq \sup_{x \in \mathcal{E}\left(\overline{b}\right)} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right)$ follows trivially.

Let us see that $\operatorname{Hof} \mathcal{F}\left(\overline{b}\right) \leq \sup_{x \in \mathcal{E}\left(\overline{b}\right)} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right)$. Specifically, let us prove that for every $x \in \operatorname{bd}\mathcal{F}\left(\overline{b}\right)$ there exists $\widetilde{x} \in \mathcal{E}\left(\overline{b}\right)$ such that $\operatorname{clm} \mathcal{F}\left(\overline{b}, x\right) \leq \operatorname{clm} \mathcal{F}\left(\overline{b}, \widetilde{x}\right)$.

Fix arbitrarily $x \in \operatorname{bd}\mathcal{F}\left(\overline{b}\right)$ and write $x = y + z$, where $y \in \operatorname{span}\{a_t,\ t \in T\}$ and $z \in \{a_t,\ t \in T\}^{\perp}$ (the orthogonal subspace to $\{a_t,\ t \in T\}$). Since $a_t' x = a_t' y$ for all $t \in T$, $y \in \operatorname{bd}\mathcal{F}\left(\overline{b}\right)$ and

$$T(x) = T(y).$$

Moreover, applying Lemma 3, we have that the regularity condition (12) is held at both $x$ and $y$. Hence, Proposition 3(ii) yields

$$\operatorname{clm} \mathcal{F}\left(\overline{b}, x\right) = \operatorname{clm} \mathcal{F}\left(\overline{b}, y\right). \tag{28}$$

Let us denote

$$C = \mathcal{F}\left(\overline{b}\right) \cap \operatorname{span}\{a_t,\ t \in T\},$$

which satisfies $\operatorname{extr} C \neq \emptyset$. If $y \in \operatorname{extr} C = \mathcal{E}\left(\overline{b}\right)$, we are done. Otherwise, if $y \in C \backslash \operatorname{extr} C$, we can apply Proposition 4 and conclude the existence of $\widetilde{x} \in \operatorname{extr} C$, $\widetilde{z} \in C$, and $\mu \in ]0, 1[$ such that $y = (1-\mu)\widetilde{x} + \mu\widetilde{z}$. Observe that

$$T(y) \subset T(\widetilde{x}),$$

since $a_t' y = b_t$ implies $(1-\mu) a_t'\widetilde{x} + \mu a_t'\widetilde{z} = b_t$, which entails $a_t'\widetilde{x} = a_t'\widetilde{z} = b_t$ (because both $\widetilde{x}, \widetilde{z} \in \mathcal{F}\left(\overline{b}\right)$). So, we conclude with the desired inequality

$$\operatorname{clm} \mathcal{F}\left(\overline{b}, y\right) \leq \operatorname{clm} \mathcal{F}\left(\overline{b}, \widetilde{x}\right),$$

which together with (28) yields

$$\operatorname{clm} \mathcal{F}\left(\overline{b}, x\right) \leq \operatorname{clm} \mathcal{F}\left(\overline{b}, \widetilde{x}\right), \text{ with } \widetilde{x} \in \mathcal{E}\left(\overline{b}\right).$$

In this way, the first equality of the current theorem is established. Finally, observe that the second one comes from Theorem 3. ∎

**4.1. On the finite case.** This subsection gathers some specifics on finite linear systems. Thus, throughout this subsection we assume that $T$ is finite, in which case, for a fixed $\left(\overline{b}, \overline{x}\right) \in \operatorname{gph} \mathcal{F}$, $\mathcal{D}(\overline{x})$ is also finite and clearly

$$\cup_{D \in \mathcal{D}(\overline{x})} \operatorname{conv}\{a_t,\ t \in D\} = \operatorname{end}\partial f(\overline{x})$$

is a closed set; moreover, $\mathcal{E}\left(\overline{b}\right)$ is also finite and $\operatorname{clm}\mathcal{F}\left(\overline{b},\overline{x}\right)$ and $\operatorname{Hof}\mathcal{F}\left(\overline{b}\right)$ can be computed through the implementable computations:

$$\begin{aligned}\operatorname{clm}\mathcal{F}\left(\overline{b},\overline{x}\right) &= \max_{D\in\mathcal{D}(\overline{x})} d_*\left(0_n, \operatorname{conv}\{a_t,\ t\in D\}\right)^{-1},\\ \operatorname{Hof}\mathcal{F}\left(\overline{b}\right) &= \max_{x\in\mathcal{E}\left(\overline{b}\right)} \operatorname{clm}\mathcal{F}\left(\overline{b},x\right).\end{aligned}$$

In addition, as a consequence of Theorem 6, we can write

$$\operatorname{Hof}\mathcal{F} = \max_{b\in\operatorname{dom}\mathcal{F}} \operatorname{Hof}\mathcal{F}(b) = \max_{b\in\operatorname{dom}\mathcal{F}}\max_{x\in\mathcal{E}(b)} \operatorname{clm}\mathcal{F}(b,x).$$

Indeed, if the maximum in (9) in Theorem 1 is attained at $J\subset T$ such that $\operatorname{rank} A_J = \operatorname{rank} A$ and $\{a_t,\ t\in J\}$ is linearly independent, we have

$$\operatorname{Hof}\mathcal{F} = \operatorname{Hof}\mathcal{F}\left(b^J\right) = \operatorname{clm}\mathcal{F}\left(b^J, 0_n\right),$$

where $b^J$ is defined as $b_t^J = 0$ if $t\in J$ and $b_t^J = 1$ otherwise.

Finally, we observe that Proposition 3(i) admits a refinement in this finite case, which is written in the following result.

PROPOSITION 5. *Let* $x^1, x^2 \in \operatorname{bd}\mathcal{F}\left(\overline{b}\right)$ *such that* $T\left(x^1\right)\subset T\left(x^2\right)$*. Then* $\mathcal{D}\left(x^1\right)\subset \mathcal{D}\left(x^2\right)$.

*Proof.* Given $D\in\mathcal{D}\left(x^1\right)$, let us see that $D\in\mathcal{D}\left(x^2\right)$. First, consider $d := x^1 - x^2$ and observe that

$$\begin{cases} a_t'd = 0,\ t\in T\left(x^1\right),\\ a_t'd = a_t'x^1 - a_t'x^2 < \overline{b}_t - \overline{b}_t = 0,\ t\in T\left(x^2\right)\backslash T\left(x^1\right).\end{cases}$$

Now, recalling (11), the fact that $D\in\mathcal{D}\left(x^1\right)$ ensures the existence of $\overline{d}\in\mathbb{R}^n$ such that

$$\begin{cases} a_t'\overline{d} = 1,\ t\in D,\\ a_t'\overline{d} < 1,\ t\in T\left(x^1\right)\backslash D.\end{cases}$$

For every $\alpha > 0$, we consider a new vector $d_\alpha := \overline{d} + \alpha d$; observe that

$$\begin{cases} a_t'd_\alpha = a_t'\overline{d} + \alpha a_t'd = 1,\ t\in D,\\ a_t'd_\alpha = a_t'\left(\overline{d}+\alpha d\right) < 1,\ t\in T\left(x^1\right)\backslash D.\end{cases}$$

Since $a_t'd < 0$ for $t\in T\left(x^2\right)\backslash T\left(x^1\right)$, we can choose $\alpha$ large enough (any $\alpha > \max_{t\in T(x^2)\backslash T(x^1)} \frac{a_t'\overline{d}-1}{-a_t'd}$ will do) to make $a_t'\left(\overline{d}+\alpha d\right) < 1$ for all $t\in T\left(x^2\right)\backslash T\left(x^1\right)$. This proves $D\in\mathcal{D}\left(x^2\right)$. ∎

The following example shows that the previous proposition does not hold in the semi-infinite framework.

*Example* 5. Let us consider the system, in $\mathbb{R}^2$ endowed with the Euclidean norm, given by

$$\sigma\left(\overline{b}\right) := \left\{\ (1+t\cos t)\,x_1 + (t\sin t)\,x_2 \le 0,\ t\in\left[0,\tfrac{\pi}{2}\right]\ \right\},$$

and take $x^1 = (0,-1)'$ and $x^2 = (0,0)'$. Then $T(x^1) = \{0\} \subset \left[0, \frac{\pi}{2}\right] = T(x^2)$. We have

$$\{0\} \in \mathcal{D}(x^1) \setminus \mathcal{D}(x^2).$$

To check that $\{0\} \notin \mathcal{D}(x^2)$ observe that the system, in the variable $d = (d_1, d_2)' \in \mathbb{R}^2$,

$$\left\{ \ d_1 = 1, \quad (1 + t\cos t)\, d_1 + (t \sin t)\, d_2 < 1, \quad t \in \left]0, \tfrac{\pi}{2}\right] \ \right\}$$

is inconsistent.

**5. Conclusions and perspectives.** We have analyzed different properties oriented to quantify the global, semilocal, and local Hoffman behavior of set-valued mappings between metric spaces, where by "semilocal" we mean the study of the whole image set with respect to parameter perturbations (a similar use of this term can be found, for instance, in [25, Definition 2.1]), yielding to the known Lipschitz upper semicontinuity when the study is concentrated around a nominal parameter. Local properties, such as calmness, are focused on the behavior of the multifunction around a fixed element of its graph. The corresponding moduli are analyzed. Both Hoffman stability (5) and uniform calmness (17) constitute intermediate steps between calmness and global Hoffman properties. All these semilocal properties are shown to be equivalent (and with the same rate/modulus) for convex-graph multifunctions taking closed values in a reflexive Banach space (Theorem 4). This is the case of the feasible set mapping, $\mathcal{F}$, associated with a continuous linear semi-infinite inequality system parameterized with respect to the right-hand side. At this moment, let us comment that paper [5] analyzes the upper Lipschitz behavior of the optimal set mapping, $\mathcal{F}^{op}$, in finite linear programming, which does not have a convex graph. Appealing to a certain concept of directional convexity introduced in that paper, [5] establishes a counterpart for the optimal set mapping of formula

$$\operatorname{Lipusc} \mathcal{F}(\bar{b}) = \sup_{x \in \mathcal{F}(\bar{b})} \operatorname{clm} \mathcal{F}\left(\bar{b}, x\right).$$

However, it is shown there that the Hoffman and Lipschitz upper semicontinuity moduli do not coincide when applied to $\mathcal{F}^{op}$ at a nominal parameter.

For this feasible set mapping we succeed in giving the following formula for the global Hoffman constant (Theorem 5), which extends to the current semi-infinite framework some previous results for finite systems:

$$\operatorname{Hof} \mathcal{F} = \sup_{\substack{J \subset T \text{ compact} \\ 0_n \notin \operatorname{conv}\{a_t,\ t \in J\}}} d_*\left(0_n, \operatorname{conv}\{a_t,\ t \in J\}\right)^{-1}.$$

With respect to the semilocal measure, $\operatorname{Hof} \mathcal{F}\left(\bar{b}\right)$, when confined to locally polyhedral systems (which includes finite systems), Theorem 6 provides a point-based formula involving exclusively some feasible points and the nominal data $a_t$'s and $\bar{b}_t$'s:

$$\operatorname{Hof} \mathcal{F}\left(\bar{b}\right) = \sup_{x \in \mathcal{E}(\bar{b})} \sup_{D \in \mathcal{D}(x)} d_*\left(0_n, \operatorname{conv}\{a_t,\ t \in D\}\right)^{-1}, \tag{29}$$

where $\mathcal{E}\left(\bar{b}\right)$ is defined as in (27). When $T$ is finite (and hence $\mathcal{E}\left(\bar{b}\right)$ and each $\mathcal{D}(x)$ also are), the previous expression yields a specially computable procedure. It provides an alternative approach to the one given in [2] via points outside the feasible set:

$$\operatorname{Hof} \mathcal{F}(\bar{b}) = \sup_{x \notin \mathcal{F}(\bar{b})} d_* \left(0_n, \operatorname{conv}\left\{a_t,\ t \in J_{\bar{b}}(x)\right\}\right)^{-1}.$$

The problem of finding an expression for $\operatorname{Hof} \mathcal{F}\left(\bar{b}\right)$ in the line of (29) for not locally polyhedral systems remains as open problem. A crucial step here is to extend Theorem 3 about the calmness modulus (traced out from [16]) to more general semi-infinite systems.

**Acknowledgments.** The authors are indebted to the anonymous referees and the associate editor for their valuable critical comments, which have definitely improved the original version of the paper.

## REFERENCES

[1] J. E. Anderson, M. A. Goberna, and M. A. López, *Locally polyhedral linear inequality systems*, Linear Algebra Appl., 270 (1998), pp. 231–253.
[2] D. Azé and J.-N. Corvellec, *On the sensitivity analysis of Hoffman constants for systems of linear inequalities*, SIAM J. Optim., 12 (2002), pp. 913–927, https://doi.org/10.1137/S1052623400375853.
[3] E. G. Belousov and V. G. Andronov, *On exact Lipschitz and Hoffman constants for systems of linear inequalities*, Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet., 47 (1999), pp. 28–32 (in Russian); English translation in Moscow Univ. Comput. Math. Cybernet., 4 (1999), pp. 35–41.
[4] J. V. Burke and P. Tseng, *A unified analysis of Hoffman's bound via Fenchel duality*, SIAM J. Optim., 6 (1996), pp. 265–282, https://doi.org/10.1137/0806015.
[5] J. Camacho, M. J. Cánovas, and J. Parra, *Lipschitz upper semicontinuity in linear optimization via local directional convexity*, Optimization, (2022), https://doi.org/10.1080/02331934.2022.2057851.
[6] M. J. Cánovas, M. A. López, J. Parra, and F. J. Toledo, *Calmness of the feasible set mapping for linear inequality systems*, Set-Valued Var. Anal., 22 (2014), pp. 375–389.
[7] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer, New York, 2009.
[8] M. J. Gisbert, M. J. Cánovas, J. Parra, and F. J. Toledo, *Calmness of the optimal value in linear programming*, SIAM J. Optim., 28 (2018), pp. 2201–2221, https://doi.org/10.1137/17M112333X.
[9] M. A. Goberna and M. A. López, *Linear Semi-infinite Optimization*, John Wiley & Sons, Chichester, UK, 1998.
[10] A. J. Hoffman, *On approximate solutions of systems of linear inequalities*, J. Research Nat. Bur. Standards, 49 (1952), pp. 263–265.
[11] H. Hu, *Characterizations of the strong basic constraint qualifications*, Math. Oper. Res., 30 (2005), pp. 956–965.
[12] D. Klatte And B. Kummer, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, Nonconvex Optim. Appl. 60, Kluwer Academic, Dordrecht, The Netherlands, 2002.
[13] D. Klatte and G. Thiere, *Error bounds for solutions of linear equations and inequalities*, Z. Oper. Res., 41 (1995), pp. 191–214.
[14] A. Kruger, H. Van Ngai, and M. Théra, *Stability of error bounds for convex constraint systems in Banach spaces*, SIAM J. Optim., 20 (2010), pp. 3280–3296, https://doi.org/10.1137/100782206.
[15] W. Li, *Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs*, SIAM J. Control Optim., 32 (1994), pp. 140–153, https://doi.org/10.1137/S036301299222723X.
[16] M. H. Li, K. W. Meng, and X. Q. Yang, *On error bound moduli for locally Lipschitz and regular functions*, Math. Program. Ser. A, 171 (2018), pp. 463–487.
[17] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I: *Basic Theory*, Springer, Berlin, 2006.
[18] J. Peña, J. C. Vera, and L. F. Zuluaga, *New characterizations of Hoffman constants for systems of linear constraints*, Math. Program., 187 (2021), pp. 79–109.

[19] J.-P. Penot, *Error bounds, calmness and their applications in nonsmooth analysis*, in Nonlinear Analysis and Optimization. II. Optimization, Contemp. Math. 514, Israel Math. Conf. Proc., AMS, Providence, RI, 2010, pp. 225–247.
[20] S. M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Progr. Study, 14 (1981), pp. 206–214.
[21] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[22] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
[23] B. Simon, *Convexity: An Analytic Viewpoint*, Cambridge Tracts Math. 187, Cambridge University Press, New York, 2011.
[24] A. Uderzo, *On the quantitative solution stability of parameterized set-valued inclusions*, Set-Valued Var. Anal., 29 (2021), pp. 425–451.
[25] N. D. Yen, J.-C. Yao, and B. T. Kien, *Covering properties at positive-order rates of multifunctions and some related topics*, J. Math. Anal. Appl., 338 (2008), pp. 467–478.
[26] C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.
[27] C. Zălinescu, *Sharp estimates for Hoffman's constant for systems of linear inequalities and equalities*, SIAM J. Optim., 14 (2003), pp. 517–533, https://doi.org/10.1137/S1052623402403505.

# B Lipschitz upper semicontinuity in linear optimization via local directional convexity.

OPTIMIZATION
https://doi.org/10.1080/02331934.2022.2057851

Taylor & Francis
Taylor & Francis Group



# Lipschitz upper semicontinuity in linear optimization via local directional convexity

J. Camacho, M. J. Cánovas and J. Parra

Center of Operations Research, Miguel Hernández University of Elche, Elche, Alicante, Spain

**ABSTRACT**
This work is focussed on computing the Lipschitz upper semicontinuity modulus of the argmin mapping for canonically perturbed linear programs. The immediate antecedent can be traced out from Camacho J et al. [2022. From calmness to Hoffman constants for linear semi-infinite inequality systems. Available from: https://arxiv.org/pdf/2107.10000v2.pdf], devoted to the feasible set mapping. The aimed modulus is expressed in terms of a finite amount of calmness moduli, previously studied in the literature. Despite the parallelism in the results, the methodology followed in the current paper differs notably from Camacho J et al. [2022] as far as the graph of the argmin mapping is not convex; specifically, a new technique based on a certain type of local directional convexity is developed.

**ARTICLE HISTORY**
Received 22 July 2021
Accepted 18 February 2022

**KEYWORDS**
Lipschitz upper semicontinuity; calmness; argmin mapping; linear programming

**MATHEMATICS SUBJECT CLASSIFICATIONS**
90C31; 49J53; 49K40; 15A39; 90C05

## 1. Introduction and motivation

The main goal of the present paper is to compute the upper semicontinuity modulus of the optimal set (argmin) mapping associated with the parameterized linear programming (LP, in brief) problem given by

$$\pi: \quad \begin{array}{ll} \text{minimize} & c'x \\ \text{subject to} & a_t'x \leq b_t, \quad t \in T := \{1, 2, \ldots, m\}, \end{array} \tag{1}$$

where $x \in \mathbb{R}^n$ is the decision variable, regarded as a column-vector, the prime stands for transposition, $a_t \in \mathbb{R}^n$ is fixed for each $t \in T$, and the pair $(c, b) \in \mathbb{R}^n \times \mathbb{R}^m$, with $b = (b_t)_{t \in T} \in \mathbb{R}^m$, is the parameter to be perturbed around a nominal one $(\bar{c}, \bar{b}) \in \mathbb{R}^n \times \mathbb{R}^m$. In this way we are dealing with the so-called *canonical perturbations* (tilt perturbations of the objective function together with the right-hand side – RHS – of the constraints). We consider the *feasible set* and

**CONTACT** J. Parra parra@umh.es

the *optimal set* mappings $\mathcal{F} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and $\mathcal{F}^{op} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined as

$$\mathcal{F}(b) := \left\{ x \in \mathbb{R}^n : a_t' x \leq b_t \text{ for all } t \in T \right\}, \tag{2}$$

$$\mathcal{F}^{op}(\pi) := \arg\min \left\{ c' x : x \in \mathcal{F}(b) \right\}. \tag{3}$$

From now on we identify each problem $\pi$ with the corresponding parameter $(c, b)$; accordingly, $\overline{\pi} \equiv (\overline{c}, \overline{b})$ denotes the nominal problem. The space of variables, $\mathbb{R}^n$, is endowed with an arbitrary norm $\|\cdot\|$, whose dual norm is denoted by $\|\cdot\|_*$, i.e. $\|u\|_* = \max_{\|x\| \leq 1} |u'x|$. The parameter space $\mathbb{R}^n \times \mathbb{R}^m$ is endowed with the norm $\|(c, b)\| := \max\{\|c\|_*, \|b\|_\infty\}$(since $c$ is identified with the linear functional $x \mapsto c'x$), where $\|b\|_\infty := \max_{t \in T} |b_t|$.

The current paper is mainly oriented to the computation of the Lipschitz upper semicontinuity modulus of $\mathcal{F}^{op}$ at the nominal parameter $\overline{\pi}$, denoted by $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi})$ following [1]; see Section 2 for the formal definitions. At this moment let us comment that $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi})$ provides a semi-local measure of the stability (in fact, a rate of deviation) of the optimal set around the nominal problem $\overline{\pi}$. The term 'semi-local' refers to the fact that only parameters $\pi$ around $\overline{\pi}$ are considered, while all elements of $\mathcal{F}^{op}(\pi)$ are taken into account. A point-based formula (only depending on the nominal data $(\overline{c}, \overline{b})$) for the aimed $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi})$ is provided in Theorem 4.2 (see also Theorem 4.1) in terms of a finite amount of calmness moduli (of a local nature) of certain feasible set mappings coming from adding new constraints to the system in (1). These calmness moduli can be computed through the point-based formula given in [2, Theorem 4] and recalled in [3, Theorem 3]. The reader is referred to monographs on variational analysis as [4–7] for details about calmness and other Lipschitz-type properties and to [8] for other stability criteria in linear optimization.

The theory of parametric linear optimization goes back to the early 1950s (see, e.g. [9, 10]). Some years later a systematic development of stability theory in LP with canonical perturbations emerged. One direction of research was the analysis of semicontinuity and Lipschitz continuity properties based on approaches from variational analysis like Berge's theory or Hoffman's error bounds. In the current parametric context both $\mathcal{F}$ and $\mathcal{F}^{op}$ are polyhedral multifunctions; i.e. their graphs, $\operatorname{gph} \mathcal{F}$ and $\operatorname{gph} \mathcal{F}^{op}$, are finite unions of convex polyhedra. In fact $\operatorname{gph} \mathcal{F}$ is a convex polyhedral cone, while the union of polyhedra constituting $\operatorname{gph} \mathcal{F}^{op}$ comes from considering the different choices of subsets of active indices involved in the Karush–Kuhn–Tucker (KKT) optimality conditions. Hence, as a consequence of a classical result by Robinson [11], both $\mathcal{F}$ and $\mathcal{F}^{op}$ are Lipschitz upper semicontinuous (see Section 2.2 for the formal definition) at any $\overline{b}$ and $\overline{\pi}$, respectively, provided that $\mathcal{F}(\overline{b})$ and $\mathcal{F}^{op}(\overline{\pi})$ are nonempty. The current paper borrows the terminology from [5] or [1], although the Lipschitz upper semicontinuity property, introduced in [11] as *upper Lipschitz continuity*, has been also popularized under the name *outer Lipschitz continuity* (see the reference book [4]). In the context of RHS perturbations (where only $b$ is perturbed, $c$ remains fixed) a well-known result establishes that both $\mathcal{F}$ and $\mathcal{F}^{op}$ are Lipschitz continuous relative to their domains (see, e.g. [[12, p. 232],[13, Chapter IX (Sec. 7)],[4, Chapter 3C]]). At this moment we point out that the case of 'fully perturbed' problems, when all data ($c$, $a_t$ and $b_t$, $t \in T$) are perturbed, entails notable differences regarding the Lipschitz upper semicontinuity as it is emphasized in Remark 2.3.

The immediate antecedent to this work can be found in [3, Theorems 4 and 6], where the Lipschitz upper semicontinuity modulus of $\mathcal{F}$ at $\overline{b}$, $\operatorname{Lipusc} \mathcal{F}(\overline{b})$, is analysed.

**Remark 1.1:** There exists a striking resemblance between the formula of $\operatorname{Lipusc} \mathcal{F}(\overline{b})$ obtained from [3, Theorems 4 and 6] and the new one, established in Theorem 4.2, of $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi})$. Just to show the similar appearance, here we gather both formulae:

$$\operatorname{Lipusc} \mathcal{F}\left(\overline{b}\right) = \max_{x \in \mathcal{E}\left(\overline{b}\right)} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right),$$

$$\operatorname{Lipusc} \mathcal{F}^{op}\left(\overline{\pi}\right) = \max_{x \in \mathcal{E}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}\left(\overline{\pi}, x\right),$$

where $\operatorname{clm} \mathcal{F}(\overline{b}, x)$ denotes the calmness modulus (see again Section 2 for the definition) of $\mathcal{F}$ at $(\overline{b}, x) \in \operatorname{gph} \mathcal{F}$ and $\operatorname{clm} \mathcal{F}^{op}(\overline{\pi}, x)$ the corresponding calmness modulus of $\mathcal{F}^{op}$ at $(\overline{\pi}, x) \in \operatorname{gph} \mathcal{F}^{op}$, and $\mathcal{E}(\overline{b})$ and $\mathcal{E}^{op}(\overline{\pi})$ are two nonempty finite subsets of extreme points of certain subsets of $\mathcal{F}(\overline{b})$ and $\mathcal{F}^{op}(\overline{\pi})$ introduced in (11) and (16), respectively. Despite these formal similarities between the two results, let us emphasize that they both follow different methodologies, mainly due to the fact that $\operatorname{gph} \mathcal{F}$ is convex (hence the last part of Theorem 2.1 below applies), while $\operatorname{gph} \mathcal{F}^{op}$ is not, even when fixing $\overline{c}$ and allowing only for RHS perturbations. Indeed, as commented above, $\operatorname{gph} \mathcal{F}^{op}$ is a finite union of convex polyhedra.

To overcome the drawback coming from the lack of convexity of $\operatorname{gph} \mathcal{F}^{op}$, the paper appeals to a weaker form of this property. Specifically, Theorem 3.1 shows that a certain local directional convexity property of the graph of $\mathcal{F}^{op}$ is preserved; indeed, when parameter $c$ remains fixed ($c = \overline{c}$) and $b$ is perturbed in the way $\overline{b} + \mu d$ for a fixed direction $d \in \mathbb{R}^n$ and a small $\mu \geq 0$. In order to illustrate this idea, let us consider the following example, where $\operatorname{conv} X$ stands for the convex hull of $X \subset \mathbb{R}^n$.

**Example 1.1:** Let us consider the problem in $\mathbb{R}^2$

$$\begin{array}{ll} \text{minimize} & c_1 x_1 + c_2 x_2 \\ \text{subject to} & x_1 \leq b_1, \quad x_2 \leq b_2, \; x_1 + x_2 \leq b_3, \quad -\frac{1}{2}x_1 + \frac{1}{2}x_2 \leq b_4, \end{array} \tag{4}$$

with

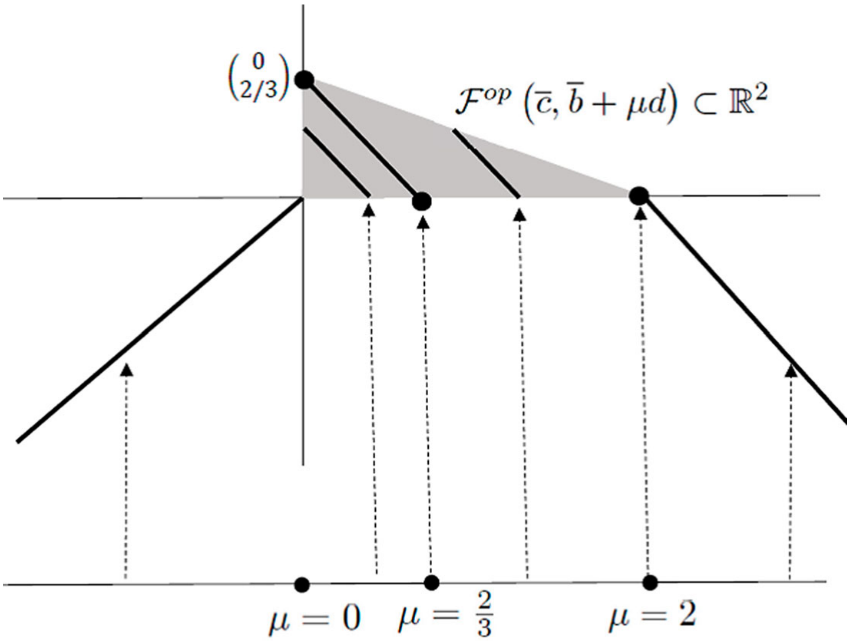


**Figure 1.** Local directional convexity in Example 1.1.

$$c = \bar{c} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{and} \quad b = (0,0,0,1)' + \mu\,(1,1,1,-1)',\ \mu \in \mathbb{R};$$

in other words, we are perturbing $\bar{b} = (0,0,0,1)'$ *in the directions* $\pm d$ with $d = (1,1,1,-1)'$.

It can be easily checked that (see Figure 1)

$$\mathcal{F}^{op}\left(\bar{c}, \bar{b}+\mu d\right) = \begin{cases} \left\{\begin{pmatrix} \mu \\ \mu \end{pmatrix}\right\} & \text{if } \mu \leq 0, \\ \operatorname{conv}\left\{\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \mu \end{pmatrix}\right\} & \text{if } 0 \leq \mu \leq \frac{2}{3}, \\ \operatorname{conv}\left\{\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} -1+\frac{3}{2}\mu \\ 1-\frac{1}{2}\mu \end{pmatrix}\right\} & \text{if } \frac{2}{3} \leq \mu \leq 2, \\ \left\{\begin{pmatrix} \mu \\ 2-\mu \end{pmatrix}\right\} & \text{if } \mu \geq 2. \end{cases}$$

In particular $\left((\bar{c}, \bar{b}-d), \begin{pmatrix} -1 \\ -1 \end{pmatrix}\right)$ and $\left((\bar{c}, \bar{b}+d), \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ belong to $\operatorname{gph} \mathcal{F}^{op}$, while the middle point $\left((\bar{c}, \bar{b}), \begin{pmatrix} 0 \\ -1/2 \end{pmatrix}\right)$ does not.

Now we summarize the structure of the paper. Section 2 contains some notation and preliminary results used later on. Section 3 formalizes the announced local directional convexity of $\operatorname{gph} \mathcal{F}^{op}$. The main result of this section, Theorem 3.1, is applied in Section 4 to establish Lemma 4.1, which constitutes a key step in the process of computing $\operatorname{Lipusc} \mathcal{F}^{op}(\bar{\pi})$, leading to Theorem 4.2. Finally, Section 5 includes some conclusions and perspectives. In particular, the so-called Hoffman stability modulus is recalled, which is known to coincide with the Lipschitz upper semicontinuity modulus under the convexity of the graph. This is the case of $\mathcal{F}$, but no longer of $\mathcal{F}^{op}$ (see Example 5.1). The reader is referred to [3, 14–18] for extra details about Hoffman constants from a global (instead of semi-local) approach.

## 2. Preliminaries

This section introduces some necessary notation and results which are used later on. Given $X \subset \mathbb{R}^p$, $p \in \mathbb{N}$, we use the standard notation cone$X$ and span$X$ for the conical convex hull and linear hull of $X$ respectively, with the convention cone $\emptyset$ = span $\emptyset = \{0_p\}$ (the zero vector of $\mathbb{R}^p$). Provided that $X$ is convex, extr$X$ stands for the set of extreme points of $X$.

Consider a generic multifunction $\mathcal{M} : Y \rightrightarrows X$ between metric spaces $Y$ and $X$ (with both distances denoted by d). Recall that the graph and the domain of $\mathcal{M}$ are respectively given by $(y, x) \in \operatorname{gph} \mathcal{M} \Leftrightarrow x \in \mathcal{M}(y)$ and $y \in \operatorname{dom} \mathcal{M} \Leftrightarrow \mathcal{M}(y) \neq \emptyset$. Mapping $\mathcal{M}$ is said to be *calm* at $(\bar{y}, \bar{x}) \in \operatorname{gph} \mathcal{M}$ if there exist a constant $\kappa \geq 0$ and neighbourhoods $V$ of $\bar{y}$ and $U$ of $\bar{x}$ such that

$$\mathrm{d}(x, \mathcal{M}(\bar{y})) \leq \kappa\, \mathrm{d}\left(y, \bar{y}\right) \quad \text{for all } y \in V \text{ and all } x \in \mathcal{M}(y) \cap U, \tag{5}$$

where $\mathrm{d}(x, \Omega) := \inf\{\mathrm{d}(x, \omega) : \omega \in \Omega\}$, provided that $x \in X$ and $\Omega \subset X$, with the usual convention $\inf \emptyset := +\infty$ and $\mathrm{d}(x, \emptyset) = +\infty$. Since we are concerned with nonnegative constants, we understand that $\sup \emptyset := 0$. It is well-known (see, e.g. [4, Theorem 3H.3 and Exercise 3H.4]) that neighbourhood $V$ appearing in the definition of calmness is redundant; formally, the calmness of $\mathcal{M}$ at $(\bar{y}, \bar{x}) \in \operatorname{gph} \mathcal{M}$ is equivalent to the existence of a constant $\kappa \geq 0$ and a (possibly smaller) neighbourhood $U$ of $\bar{x}$ such that

$$\mathrm{d}(x, \mathcal{M}(\bar{y})) \leq \kappa\, \mathrm{d}\left(\bar{y}, \mathcal{M}^{-1}(x)\right) \quad \text{for all } x \in U. \tag{6}$$

The latter property is known as the *metric subregularity* of $\mathcal{M}^{-1}$ at $(\bar{x}, \bar{y})$ (recall that $y \in \mathcal{M}^{-1}(x) \Leftrightarrow x \in \mathcal{M}(y)$). The calmness modulus of $\mathcal{M}$ at $(\bar{y}, \bar{x}) \in \operatorname{gph} \mathcal{M}$, denoted by $\operatorname{clm} \mathcal{M}(\bar{y}, \bar{x})$, is the infimum of constants $\kappa$ such that (5) (equivalently (6)) holds for some associated neighbourhoods; this modulus can be written as:

$$\operatorname{clm} \mathcal{M}\left(\bar{y}, \bar{x}\right) = \limsup_{\substack{(y,x)\to(\bar{y},\bar{x}) \\ x\in\mathcal{M}(y)}} \frac{\mathrm{d}(x, \mathcal{M}(\bar{y}))}{\mathrm{d}\left(y, \bar{y}\right)} = \limsup_{x\to\bar{x}} \frac{\mathrm{d}(x, \mathcal{M}(\bar{y}))}{\mathrm{d}\left(\bar{y}, \mathcal{M}^{-1}(x)\right)}, \tag{7}$$

where $\frac{0}{0} := 0$ and lim sup is understood as the supremum (maximum, indeed) of all possible sequential upper limits (i.e. with $(y, x)$ being replaced with elements of sequences $\{(y_r, x_r)\}_{r\in\mathbb{N}}$ converging to $(\bar{y}, \bar{x})$ as $r \to \infty$).

### 2.1. Lipschitz upper semicontinuity of multifunctions

The current work is focussed on a semi-local Lipschitz-type property: A multifunction $\mathcal{M}$ is said to be *Lipschitz upper semicontinuous* (see, e.g. [[1, Definition 2.1(iii)],[3, Section 3]]) at $\overline{y} \in \operatorname{dom} \mathcal{M}$ if there exist a constant $\kappa \geq 0$ and a neighbourhood $V$ of $\overline{y}$ such that

$$\mathrm{d}(x, \mathcal{M}(\overline{y})) \leq \kappa\, \mathrm{d}\left(y, \overline{y}\right) \quad \text{for all } y \in V \text{ and all } x \in \mathcal{M}(y). \tag{8}$$

The associated Lipschitz upper semicontinuity modulus at $\overline{y} \in \operatorname{dom} \mathcal{M}$, denoted by $\operatorname{Lipusc} \mathcal{M}(\overline{y})$, is defined as the infimum of constants $\kappa$ satisfying (8) for some associated $V$. Clearly $V$ is not redundant here. The following result provides a limiting expression for $\operatorname{Lipusc} \mathcal{M}(\overline{y})$.

**Proposition 2.1 ([3, Proposition 2(i)]):** *Let $\mathcal{M} : Y \rightrightarrows X$ be a multifunction between metric spaces and let $\overline{y} \in \operatorname{dom} \mathcal{M}$, then*

$$\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \limsup_{y \to \overline{y}} \left( \sup_{x \in \mathcal{M}(y)} \frac{\mathrm{d}(x, \mathcal{M}(\overline{y}))}{\mathrm{d}\left(y, \overline{y}\right)} \right).$$

The following result establishes the relationship between calmness and Lipschitz upper semicontinuity for generic multifunctions. See Section 5 for additional comments including the so-called Hoffman stability, defined in (22).

**Theorem 2.1 ([3, Corollary 2 and Theorem 4]):** *Let $\mathcal{M} : Y \rightrightarrows X$ be a multifunction between metric spaces and let $\overline{y} \in \operatorname{dom} \mathcal{M}$. We have*

$$\operatorname{Lipusc} \mathcal{M}\left(\overline{y}\right) \geq \sup_{x \in \mathcal{M}(\overline{y})} \operatorname{clm} \mathcal{M}\left(\overline{y}, x\right). \tag{9}$$

*Moreover, equality holds in (9) if $Y$ is a normed space, $X$ is a reflexive Banach space, $\operatorname{gph} \mathcal{M}$ is a nonempty convex set, and $\mathcal{M}(\overline{y})$ is closed.*

**Remark 2.1:** (i) Observe that the previous theorem relates local and semi-local Lipschitz-type measures for multifunctions.

(ii) The convexity assumption in the previous theorem is not superfluous for establishing equality in (9), as it was shown in [3, Example 2].

### 2.2. Lipschitz upper semicontinuity of the feasible set mapping

This subsection deals with the feasible set mapping $\mathcal{F} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ introduced in (2), which has a closed convex graph and, so, Theorem 2.1 allows us to write

$$\operatorname{Lipusc} \mathcal{F}(\overline{b}) = \sup_{x \in \mathcal{F}(\overline{b})} \operatorname{clm} \mathcal{F}\left(\overline{b}, x\right). \tag{10}$$

Going further, the next theorem provides a more computable expression for $\operatorname{Lipusc} \mathcal{F}(\overline{b})$ as far as it involves a finite amount of calmness moduli. It appeals to the following set of extreme points:

$$\mathcal{E}\left(b\right):=\operatorname{extr}\left(\mathcal{F}(b)\cap\operatorname{span}\left\{a_t,\ t\in T\right\}\right),\quad b\in\operatorname{dom}\mathcal{F}. \tag{11}$$

For details about this construction the reader is addressed to [19, p. 142]. In addition to the following theorem, set $\mathcal{E}(b)$ is also a key tool in [20] to provide a point-based formula for the calmness modulus of the optimal value function in linear optimization. This set is known to be always nonempty and finite.

**Theorem 2.2 ([3, See Theorems 4 and 6]):** *Let $\overline{b}\in\operatorname{dom}\mathcal{F}$. Then*

$$\operatorname{Lipusc}\mathcal{F}(\overline{b})=\max_{x\in\mathcal{E}\left(\overline{b}\right)}\operatorname{clm}\mathcal{F}\left(\overline{b},x\right). \tag{12}$$

In the sequel, for any $(b,x)\in\operatorname{gph}\mathcal{F}$, $T_b(x)$ represents the set of active indices at $x$, defined as

$$T_b\left(x\right):=\left\{t\in T:\ a_t'x=b_t\right\}.$$

In particular, it is appealed to in the definition of the following family of sets appearing in the next remark: given $(b,x)\in\operatorname{gph}\mathcal{F}$, $\mathcal{D}_b(x)$ is formed by all subsets $D\subset T_b(x)$ such that system

$$\left\{\begin{array}{ll} a_t'd=1, & t\in D,\\ a_t'd<1, & t\in T_b(x)\setminus D \end{array}\right\}$$

is consistent (in the variable $d\in\mathbb{R}^n$); in other words, $\{a_t,\ t\in D\}$ lives in some hyperplane which leaves the remaining coefficient vectors $a_t, t\in T_b(x)\setminus D$ and the origin, $0_n$, in one of its two associated open half-spaces.

**Remark 2.2:** It is worth mentioning that any calmness modulus, $\operatorname{clm}\mathcal{F}(\overline{b},x)$, at any $x\in\mathcal{E}(\overline{b})$ appearing in (12) can be computed through the point-based formula given in [2, Theorem 4] (see also [3, Theorem 3], as stated in the introduction), which is recalled here for completeness:

$$\operatorname{clm}\mathcal{F}(\overline{b},\overline{x})=\max_{D\in\mathcal{D}_{\overline{b}}(\overline{x})}d_*\left(0_n,\operatorname{conv}\left\{a_t,\ t\in D\right\}\right)^{-1},\quad\left(\overline{b},\overline{x}\right)\in\operatorname{gph}\mathcal{F}, \tag{13}$$

where $d_*$ represents the distance associated with the dual norm $\|\cdot\|_*$.

**Remark 2.3:** The fact of considering RHS perturbations is crucial in our analysis. Observe that in the case when the $a_t$'s are also perturbed the corresponding feasible set mapping is no longer Lipschitz upper semicontinuous in general. Just consider the example in $\mathbb{R}^2$ with only one constraint: $\widetilde{\mathcal{F}}(a,b)=\{x\in\mathbb{R}^2\mid a_1x_1+a_2x_2\le b\}$, $a=(a_1,a_2)'\in\mathbb{R}^2$, $b\in\mathbb{R}$. Then, for each $r=1,2,\ldots$, we have

$$\sup_{x\in\widetilde{\mathcal{F}}\left(\left(1,-\frac{1}{r}\right),0\right)}\frac{\mathrm{d}(x,\widetilde{\mathcal{F}}((1,0)',0))}{r^{-1}}\geq\lim_{k\to\infty}\frac{\mathrm{d}(\binom{k}{rk},\widetilde{\mathcal{F}}\left((1,0)',0\right))}{r^{-1}}=+\infty,$$

which entails $\operatorname{Lipusc}\widetilde{\mathcal{F}}((1,0)',0)=+\infty$. Roughly speaking, the previous situation arises from the semi-local nature of this property (i.e. the fact that it involves the whole feasible sets associated with perturbed parameters). Indeed, other variational properties of local character (dealing with parameters and points near the nominal ones) do not change so drastically. For instance, this is the case of the calmness property (see, [2, Theorem 5]).

## 3. Local directional convexity of $\mathcal{F}^{op}$

This is an instrumental section oriented to tackle our problem of computing Lipusc $\mathcal{F}^{op}(\overline{\pi})$. As gph $\mathcal{F}^{op}$ is not convex, we are not allowed to apply the last part of Theorem 2.1, and this fact entails notable differences with respect to the methodology followed in [3]. Specifically, this section is devoted to establish a certain directional-type convexity of the graph of $\mathcal{F}^{op}$ around $\overline{b}$ with a fixed $\overline{c}$, where only local directional RHS perturbations of the constraints are considered; see Figure 1 for a geometrical motivation. Formally, associated with our nominal problem $\overline{\pi} = (\overline{c}, \overline{b})$, any scalar $\varepsilon > 0$, and any direction $d \in \mathbb{R}^m$ with $\|d\|_\infty = 1$, we consider the *local directional optimal set mapping* $\mathcal{F}^{op}_{\overline{\pi},d,\varepsilon} : [0, \varepsilon] \rightrightarrows \mathbb{R}^n$ given by

$$\mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}(\mu) = \mathcal{F}^{op}\left(\overline{c}, \overline{b} + \mu d\right), \quad \mu \in [0, \varepsilon]. \tag{14}$$

The following lemmas constitute key tools for establishing Theorem 3.1. To start with, we introduce some extra notation: given $\pi = (c, b) \in \operatorname{dom} \mathcal{F}^{op}$, $\mathcal{M}_\pi$ denotes the so-called *minimal KKT subsets of indices* at $\pi$, defined as

$$\mathcal{M}_\pi := \left\{ D \subset T_b(x) \;\middle|\; \begin{array}{l} -c \in \operatorname{cone}\{a_t, t \in D\} \\ D \text{ is minimal for the inclusion order} \end{array} \right\}, \tag{15}$$

provided that $x$ is any optimal point of $\pi$. For convenience, sometimes $\mathcal{M}_\pi$ is alternatively denoted by $\mathcal{M}_{c,b}$ for $\pi = (c, b) \in \operatorname{dom} \mathcal{F}^{op}$. Let us observe that $\mathcal{M}_\pi$ is correctly defined since the expression in (15) indeed does not depend on $x$ as it was commented in [20, Remark 2]. Finally, we introduce the counterpart of $\mathcal{E}(b)$ when dealing with optimization problems,

$$\mathcal{E}^{op}(\pi) := \operatorname{extr}\left(\mathcal{F}^{op}(\pi) \cap \operatorname{span}\{a_t, \ t \in T\}\right), \quad \pi \in \operatorname{dom} \mathcal{F}^{op}. \tag{16}$$

The reader is referred to [20, Section 2.2] for additional details about this set of extreme points. Standard arguments of linear optimization yield $\mathcal{E}^{op}(\pi) = \mathcal{F}^{op}(\pi) \cap \mathcal{E}(b)$ for $\pi = (c, b) \in \operatorname{dom} \mathcal{F}^{op}$.

**Lemma 3.1 ([20, Lemma 2]):** *Let $\{\pi^r\}_{r\in\mathbb{N}} \subset \operatorname{dom} \mathcal{F}^{op}$ converge to $\overline{\pi}$. Then*

$$\emptyset \neq \operatorname{Limsup}_r \mathcal{E}^{op}\left(\pi^r\right) \subset \mathcal{E}^{op}\left(\overline{\pi}\right),$$

*where* Limsup$_r$ *stands for the Painlevé–Kuratowski upper/outer limit as $r \to \infty$ (see, e.g. [6, 7]).*

**Lemma 3.2:** *Let* $(\bar{c}, \bar{b}) \in \operatorname{dom} \mathcal{F}^{op}$. *Then there exists* $\varepsilon > 0$ *such that for every* $b \in \operatorname{dom} \mathcal{F}$ *with* $\|b - \bar{b}\|_\infty \leq \varepsilon$ *we have* $\mathcal{M}_{\bar{c},b} \subset \mathcal{M}_{\bar{c},\bar{b}}$.

***Proof:*** Reasoning by contradiction, suppose that there exists a sequence $\{b^r\}_{r\in\mathbb{N}}$ such that $\operatorname{dom} \mathcal{F} \ni b^r \to \bar{b}$ and $D_r \in \mathcal{M}_{\bar{c},b^r} \setminus \mathcal{M}_{\bar{c},\bar{b}}$ for all $r \in \mathbb{N}$. Since $D_r \subset T$ (finite) for all $r$, we may assume (by taking a subsequence if necessary) that $\{D_r\}_{r\in\mathbb{N}}$ is constant, say $D_r = D$ for all $r$.

Applying Lemma 3.1, let $\bar{x} \in \operatorname{Limsup}_r \mathcal{E}^{op}(\bar{c}, b^r)$ and, without loss of generality (for an appropriate subsequence, without relabelling), write $\bar{x} = \lim_r x^r$, for some $x^r \in \mathcal{E}^{op}(\bar{c}, b^r)$, $r = 1, 2, \ldots$ For each $r$, $D \in \mathcal{M}_{\bar{c},b^r}$ entails $D \subset T_{b^r}(x^r)$, which yields $D \subset T_{\bar{b}}(\bar{x})$. Moreover, $-\bar{c} \in \operatorname{cone}\{a_t,\ t \in D\}$ and $D$ is minimal with respect to this property since it is for any $r$ (recall $D_r = D$). Hence, we attain the contradiction $D \in \mathcal{M}_{\bar{c},\bar{b}}$. ∎

**Theorem 3.1:** *Let* $\bar{\pi} = (\bar{c}, \bar{b}) \in \operatorname{dom} \mathcal{F}^{op}$ *and* $\varepsilon > 0$ *be as in Lemma 3.2. Then* $\operatorname{gph} \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}$ *is convex for all* $d \in \mathbb{R}^m$ *with* $\|d\|_\infty = 1$.

***Proof:*** Let $(\mu_0, x^0), (\mu_1, x^1) \in \operatorname{gph} \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}$ with $0 \leq \mu_0 \leq \mu_1 \leq \varepsilon$. Let us see that $(\mu_\lambda, x^\lambda) := (1 - \lambda)(\mu_0, x^0) + \lambda(\mu_1, x^1) \in \operatorname{gph} \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}$ for $\lambda \in ]0, 1[$. If $\mu_0 = \mu_1$, then $x^0$ and $x^1$ belong to the same convex set $\mathcal{F}^{op}(\bar{c}, \bar{b} + \mu_0 d)$, and hence clearly $x^\lambda \in \mathcal{F}^{op}(\bar{c}, \bar{b} + \mu_0 d) = \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(\mu_\lambda)$for all $\lambda \in ]0, 1[$.

From now on, let us assume $\mu_0 < \mu_1$. First observe that $x^\lambda \in \mathcal{F}(\bar{b} + \mu_\lambda d)$ for all $\lambda \in ]0, 1[$ because of the convexity of $\operatorname{gph} \mathcal{F}$. We distinguish two cases:

*Case 1* $\mu_0 = 0$. Fix any $\lambda \in ]0, 1[$. Observe that in this case $\mu_\lambda = \lambda\mu_1$. Let us prove that $x^\lambda \in \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(\mu_\lambda)$. Take $D_1 \in \mathcal{M}_{\bar{c},\bar{b}+\mu_1 d} \subset \mathcal{M}_{\bar{c},\bar{b}}$ (because of Lemma 3.2 and the choice of $\varepsilon$). In particular, $D_1 \subset T_{\bar{b}+\mu_1 d}(x^1) \cap T_{\bar{b}}(x^0)$. Therefore $D_1 \subset T_{\bar{b}+\lambda\mu_1 d}(x^\lambda)$ since, for any $t \in D_1$,

$$a_t' x^\lambda = (1 - \lambda)\, a_t' x^0 + \lambda a_t' x^1 = (1 - \lambda)\, \bar{b}_t + \lambda \left( \bar{b}_t + \mu_1 d_t \right) = \bar{b}_t + \lambda\mu_1 d_t.$$

Hence, as we are not perturbing $\bar{c}$, KKT optimality conditions ensure $x^\lambda \in \mathcal{F}^{op}(\bar{c}, \bar{b} + \lambda\mu_1 d) = \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(\mu_\lambda)$.

*Case 2* $\mu_0 > 0$. Fix again any $\lambda \in ]0, 1[$ and let us see that $x^\lambda \in \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(\mu_\lambda)$.

Start by choosing an arbitrary $\bar{x} \in \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(0) = \mathcal{F}^{op}(\bar{c}, \bar{b})$ and define

$$\tilde{x}^1 := \left(1 - \frac{\mu_0}{\mu_1}\right) \bar{x} + \frac{\mu_0}{\mu_1} x^1.$$

Reasoning as in the previous case, with $\bar{x}$ and $\frac{\mu_0}{\mu_1}$ playing the role of $x^0$ and $\lambda$, respectively, we deduce

$$\tilde{x}^1 \in \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}\left(\frac{\mu_0}{\mu_1}\mu_1\right) = \mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}(\mu_0).$$

Appealing to the convexity of the previous optimal set, define

$$\tilde{x}^{\alpha} := (1-\alpha)\, x^0 + \alpha \tilde{x}^1 \in \mathcal{F}^{op}_{\pi,d,\varepsilon}\,(\mu_0) \quad \forall\, \alpha \in [0,1].$$

A routine computation yields the existence of scalars $\alpha \in [0,1]$ and $\beta \geq 1$ such that

$$x^{\lambda} - \overline{x} = \beta \left(\tilde{x}^{\alpha} - \overline{x}\right).$$

Indeed, they are unique and their explicit expressions are

$$\alpha = \frac{\lambda\mu_1}{(1-\lambda)\,\mu_0 + \lambda\mu_1}, \qquad \beta = \frac{(1-\lambda)\,\mu_0 + \lambda\mu_1}{\mu_0}.$$

Since $\tilde{x}^{\alpha} \in \mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}(\mu_0)$, there exists $D_\alpha \in \mathcal{M}_{\overline{c},\overline{b}+\mu_0 d} \subset \mathcal{M}_{\overline{c},\overline{b}}$, in particular, $D_\alpha \subset T_{\overline{b}+\mu_0 d}(\tilde{x}^{\alpha}) \cap T_{\overline{b}}(\overline{x})$. Hence, for any $t \in D_\alpha$, taking into account the fact that $\beta\mu_0 = \mu_\lambda$, one has

$$\begin{aligned} a_t' x^{\lambda} &= a_t'\overline{x} + \beta a_t' \left(\tilde{x}^{\alpha} - \overline{x}\right) \\ &= \overline{b}_t + \beta \left(\overline{b}_t + \mu_0 d_t - \overline{b}_t\right) = \overline{b}_t + \beta\mu_0 d_t = \overline{b}_t + \mu_\lambda d_t. \end{aligned}$$

In this way, $D_\alpha \subset T_{\overline{b}+\mu_\lambda d}(x^{\lambda})$ and again the KKT optimality conditions yield $x^{\lambda} \in \mathcal{F}^{op}(\overline{c}, \overline{b} + \mu_\lambda d)$. In other words, $(\mu_\lambda, x^{\lambda}) \in \operatorname{gph} \mathcal{F}^{op}_{\overline{\pi},d,\varepsilon}$. ■

The following example illustrates the previous results.

**Example 3.1 (Example 1.1 revisited):** Let us consider the parameterized problem (4) in $\mathbb{R}^2$ with $T = \{1,2,3,4\}$ and let us see that the statement of Lemma 3.2 fulfils by taking $0 < \varepsilon < \frac{2}{5}$. First observe that

$$\left.\begin{array}{c} \|b - \overline{b}\|_\infty < \frac{2}{5} \\ x \in \mathcal{F}^{op}(\overline{c}, b) \end{array}\right\} \Longrightarrow T_b\,(x) \subset \{1,2,3\}. \tag{17}$$

Reasoning by contradiction, assume that there exists $(b,x) \in \mathbb{R}^4 \times \mathbb{R}^2$ such that $x \in \mathcal{F}^{op}(\overline{c}, b)$, $\|b - \overline{b}\|_\infty < \frac{2}{5}$ and $4 \in T_b(x)$. It is clear that $\{4\} \subsetneq T_b(x)$ according to KKT conditions ($-\overline{c} \notin \operatorname{cone}\{a_4\}$). Then, by distinguishing cases we attain a contradiction: assume that $1 \in T_b(x)$, then $x_1 = b_1$ and $-\frac{1}{2}x_1 + \frac{1}{2}x_2 = b_4$, which yields $x_2 = b_1 + 2b_4 > -\frac{2}{5} + 2(1 - \frac{2}{5}) = \frac{4}{5}$ which contradicts $x_2 \leq b_2 < \frac{2}{5}$. Hence $1 \notin T_b(x)$. Then, necessarily $3 \in T_b(x)$ (since $-\overline{c} \notin \operatorname{cone}\{a_2, a_4\}$), but again $x_1 + x_2 = b_3$ and $-\frac{1}{2}x_1 + \frac{1}{2}x_2 = b_4$ yield the contradiction $\frac{2}{5} > x_2 = \frac{1}{2}b_3 + b_4 > -\frac{1}{5} + \frac{3}{5} = \frac{2}{5}$.

From (17) one easily derives

$$\mathcal{M}_{\bar{c},b} \subset \mathcal{M}_{\bar{c},\bar{b}} = \{\{1,2\},\{3\}\}, \quad \text{whenever } \|b-\bar{b}\|_\infty < \frac{2}{5}. \tag{18}$$

Indeed, given any $\|b-\bar{b}\|_\infty < \frac{2}{5}$ one can check

$$\mathcal{M}_{\bar{c},b} = \begin{cases} \{\{1,2\},\{3\}\}, & \text{if } b_1+b_2=b_3, \\ \{\{1,2\}\}, & \text{if } b_1+b_2<b_3, \\ \{\{3\}\}, & \text{if } b_1+b_2>b_3. \end{cases}$$

Therefore, for any $d \in \mathbb{R}^4$, with $\|d\|_\infty = 1$, and $\mu \in [0, \frac{2}{5}[$, we have

$$\mathcal{F}^{op}\left(\bar{c}, \bar{b}+\mu d\right) = \begin{cases} \left\{\mu\begin{pmatrix} d_1 \\ d_2\end{pmatrix}\right\}, & \text{if } d_1+d_2 \leq d_3, \\ \mu \operatorname{conv}\left\{\begin{pmatrix} d_1 \\ d_3-d_1\end{pmatrix}, \begin{pmatrix} d_3-d_2 \\ d_2\end{pmatrix}\right\}, & \text{if } d_1+d_2>d_3, \end{cases} \tag{19}$$

which clearly entails the convexity of gph $\mathcal{F}^{op}_{\bar{\pi},d,\varepsilon}$ for each $d \in \mathbb{R}^4$ with $\|d\|_\infty = 1$ and each $0 < \varepsilon < \frac{2}{5}$.

Finally, observe that (17) is no longer true for $\varepsilon = \frac{2}{5}$ since

$$T_{\bar{b}+\frac{2}{5}(-1,1,-1,-1)'}\left(\left(-\frac{4}{5}, \frac{2}{5}\right)'\right) = \{2,3,4\}.$$

However, (18) still holds by replacing $\frac{2}{5}$ with $\frac{1}{2}$, since the only way to preclude $\mathcal{M}_{\bar{c},b} \subset \mathcal{M}_{\bar{c},\bar{b}}$ is having $\{1,4\} \in \mathcal{M}_{\bar{c},b}$, which implies $\|b-\bar{b}\|_\infty \geq \frac{1}{2}$; although the description of $\mathcal{F}^{op}(\bar{c}, \bar{b}+\mu d)$ would be different from that of (19) when $\mu > \frac{2}{5}$.

## 4. Lipschitz upper semicontinuity of the optimal set mapping

This section tackles the final goal of the current paper: the computation of the Lipschitz upper semicontinuity modulus for the optimal set mapping $\mathcal{F}^{op}$ introduced in (3). First, let us see that perturbations of $c$ are redundant when looking for the aimed modulus. Formally, we consider $\mathcal{F}^{op}_{\bar{c}} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ given by

$$\mathcal{F}^{op}_{\bar{c}}(b) = \mathcal{F}^{op}(\bar{c}, b), \quad b \in \mathbb{R}^m.$$

**Proposition 4.1 ([20, Lemma 4]):** *There exists $\varepsilon > 0$ such that*

$$\mathcal{F}^{op}(\pi) \subset \mathcal{F}^{op}(\bar{c}, b),$$

*whenever $\pi \equiv (c,b), \bar{\pi} \equiv (\bar{c}, \bar{b}) \in$ dom $\mathcal{F}^{op}$ satisfy $\|\pi - \bar{\pi}\| < \varepsilon$.*

**Corollary 4.1:** *Let* $\overline{\pi} \equiv (\overline{c}, \overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$. *We have*

$$\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) = \operatorname{Lipusc} \mathcal{F}^{op}_{\overline{c}}\left(\overline{b}\right) = \limsup_{b \to \overline{b}} \left( \sup_{x \in \mathcal{F}^{op}(\overline{c}, b)} \frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d\left(b, \overline{b}\right)} \right). \tag{20}$$

***Proof:*** Inequality $\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) \geq \operatorname{Lipusc} \mathcal{F}^{op}_{\overline{c}}(\overline{b})$ follows directly from the definitions. Appealing to Propositions 2.1 and 4.1 we have

$$\begin{aligned}
\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) &= \limsup_{\pi \to \overline{\pi}} \left( \sup_{x \in \mathcal{F}^{op}(\pi)} \frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d(\pi, \overline{\pi})} \right) \\
&\leq \limsup_{\pi \to \overline{\pi}} \left( \sup_{x \in \mathcal{F}^{op}(\overline{c}, b)} \frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d(\pi, \overline{\pi})} \right) \\
&\leq \limsup_{b \to \overline{b}} \left( \sup_{x \in \mathcal{F}^{op}(\overline{c}, b)} \frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d\left(b, \overline{b}\right)} \right) \\
&= \operatorname{Lipusc} \mathcal{F}^{op}_{\overline{c}}\left(\overline{b}\right). \qquad \blacksquare
\end{aligned}$$

From now on, appealing to the local directional convexity of the graph of $\mathcal{F}^{op}$ under RHS perturbations, we adapt some arguments used in [3] to the current setting. The following technical lemma uses Theorem 3.1 to provide an upper bound on the variation rate of optimal solutions with respect to RHS perturbations.

**Lemma 4.1:** *Let* $\overline{\pi} = (\overline{c}, \overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$. *Let* $\varepsilon > 0$ *be as in Lemma 3.2, take* $((\overline{c}, b), x) \in \operatorname{gph} \mathcal{F}^{op}$ *with* $\|b - \overline{b}\|_\infty \leq \varepsilon$ *and let* $p(x)$ *be a projection (a best approximation) of* $x$ *in* $\mathcal{F}^{op}(\overline{\pi})$. *Then*

$$\frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d\left(b, \overline{b}\right)} \leq \operatorname{clm} \mathcal{F}^{op}\left(\overline{\pi}, p(x)\right).$$

***Proof:*** The case $b = \overline{b}$ is trivial from $0/0 := 0$. Assume $b \neq \overline{b}$ and let $d := \frac{b - \overline{b}}{\|b - \overline{b}\|_\infty}$. With the notation (14), $p(x) \in \mathcal{F}^{op}_{\overline{\pi}, d, \varepsilon}(0)$ and $x \in \mathcal{F}^{op}_{\overline{\pi}, d, \varepsilon}(\|b - \overline{b}\|_\infty)$, which entails, for each $\lambda \in [0, 1]$, by applying Theorem 3.1,

$$x_\lambda := (1 - \lambda) p(x) + \lambda x \in \mathcal{F}^{op}_{\overline{\pi}, d, \varepsilon}\left(\lambda \|b - \overline{b}\|_\infty\right) = \mathcal{F}^{op}\left(\overline{c}, (1 - \lambda)\overline{b} + \lambda b\right).$$

Moreover, from [3, Lemma 1] we have that $p(x)$ is also a best approximation of $x_\lambda$ in $\mathcal{F}^{op}(\overline{\pi})$; i.e. $\|x_\lambda - p(x)\| = d(x_\lambda, \mathcal{F}^{op}(\overline{\pi}))$, $\lambda \in [0, 1]$. Consequently,

$$\frac{d\left(x_\lambda, \mathcal{F}^{op}(\overline{\pi})\right)}{\|\lambda\left(b - \overline{b}\right)\|_\infty} = \frac{\lambda \|x - p(x)\|}{\lambda \|b - \overline{b}\|_\infty} = \frac{d\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{d\left(b, \overline{b}\right)}, \quad \text{whenever } \lambda \in \left]0, 1\right].$$

Hence, letting $\lambda \to 0$ we obtain the aimed inequality. ■

**Proposition 4.2:** *Let $\overline{\pi} = (\overline{c}, \overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$, then*

$$\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) = \sup_{z \in \mathcal{F}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}(\overline{\pi}, z).$$

***Proof:*** Inequality '$\geq$' comes from Theorem 2.1. Let us prove the nontrivial inequality $\leq$. Take an $\varepsilon > 0$ as in Lemma 3.2 (and Theorem 3.1). We can write (recall Corollary 4.1)

$$\begin{aligned}\operatorname{Lipusc} \mathcal{F}^{op}\left(\overline{c}, \overline{b}\right) &= \limsup_{b \to \overline{b}} \left( \sup_{x \in \mathcal{F}^{op}_{\overline{c}}(b)} \frac{\mathrm{d}\left(x, \mathcal{F}^{op}(\overline{\pi})\right)}{\mathrm{d}\left(b, \overline{b}\right)} \right) \\ &\leq \limsup_{\substack{b \to \overline{b} \\ \|b - \overline{b}\| \leq \varepsilon}} \left( \sup_{x \in \mathcal{F}^{op}_{\overline{c}}(b)} \operatorname{clm} \mathcal{F}^{op}\left(\overline{\pi}, p(x)\right) \right),\end{aligned}$$

where we have applied Lemma 4.1 and, as in that lemma, for each $x \in \mathcal{F}^{op}_{\overline{c}}(b)$, with $\|b - \overline{b}\| \leq \varepsilon$, $p(x) \in \mathcal{F}^{op}(\overline{\pi})$ is a projection of $x$ in $\mathcal{F}^{op}(\overline{\pi})$.

Consequently,

$$\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) \leq \sup_{z \in \mathcal{F}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}(\overline{\pi}, z).$$

■

Next we show that the supremum in the previous proposition may be reduced to a finite subset of points, indeed to points in $\mathcal{E}^{op}(\overline{c}, \overline{b})$. To do that we appeal to the following theorem.

**Theorem 4.1 ([21, Corollary 4.1]):** *Let $(\overline{\pi}, \overline{x}) \in \operatorname{gph} \mathcal{F}^{op}$ with $\overline{\pi} = (\overline{c}, \overline{b})$. Then*

$$\operatorname{clm} \mathcal{F}^{op}\left(\left(\overline{c}, \overline{b}\right), \overline{x}\right) = \max_{D \in \mathcal{M}_{\overline{\pi}}} \operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), \overline{x}\right), \tag{21}$$

*where, for each $D \in \mathcal{M}_{\overline{\pi}}$, $\mathcal{L}_D : \mathbb{R}^m \times \mathbb{R}^D \rightrightarrows \mathbb{R}^n$ is defined by*

$$\mathcal{L}_D(b, d) := \left\{ x \in \mathbb{R}^n : a_t' x \leq b_t,\ t = 1, \ldots, m;\ -a_t' x \leq d_t,\ t \in D \right\},$$

*and $\overline{b}_D := (\overline{b}_t)_{t \in D}$.*

**Remark 4.1:** Observe that each $\mathcal{L}_D$ is nothing else but a feasible set mapping of the same type as $\mathcal{F}$ but associated with an enlarged system. Consequently, Remark 2.2 also applies here for computing each $\operatorname{clm} \mathcal{L}_D((\overline{b}, -\overline{b}_D), \overline{x})$, and therefore $\operatorname{clm} \mathcal{F}^{op}((\overline{c}, \overline{b}), \overline{x})$.

Finally, by gathering the previous results of this section, we can establish our main goal in the following theorem. We point out that this theorem provides an implementable procedure for computing the aimed Lipusc $\mathcal{F}^{op}(\overline{\pi})$ as far as it is written in terms of a finite amount of calmness moduli of feasible set mappings (as the previous theorem says) and these calmness moduli can be computed via formula (13).

**Theorem 4.2:** *Let* $\overline{\pi} \in \operatorname{dom} \mathcal{F}^{op}$, *then*

$$\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) = \max_{x \in \mathcal{E}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}(\overline{\pi}, x).$$

***Proof:*** Starting from the equalities of Proposition 4.2 and Theorem 4.1, we can write

$$\begin{aligned}
\operatorname{Lipusc} \mathcal{F}^{op}(\overline{\pi}) &= \sup_{x \in \mathcal{F}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}(\overline{\pi}, x) \\
&= \sup_{x \in \mathcal{F}^{op}(\overline{\pi})} \max_{D \in \mathcal{M}_{\overline{\pi}}} \operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), x\right) \\
&= \max_{D \in \mathcal{M}_{\overline{\pi}}} \sup_{x \in \mathcal{L}_D\left(\overline{b}, -\overline{b}_D\right)} \operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), x\right) \\
&= \max_{D \in \mathcal{M}_{\overline{\pi}}} \max_{x \in \mathcal{E}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), x\right) \\
&= \max_{x \in \mathcal{E}^{op}(\overline{\pi})} \operatorname{clm} \mathcal{F}^{op}\left(\left(\overline{c}, \overline{b}\right), x\right).
\end{aligned}$$

In the third equality we have used the fact that $\mathcal{F}^{op}(\overline{c}, \overline{b}) = \mathcal{L}_D(\overline{b}, -\overline{b}_D)$ for all $D \in \mathcal{M}_{\overline{\pi}}$ (see [21, Proposition 4.1]), while the fourth comes from Theorems 2.1 and 2.2 by taking Remark 4.1 into account, together with the fact that the role played by $\mathcal{E}(\overline{b})$ in Theorem 2.2 is now played by

$$\operatorname{extr}\left(\mathcal{L}_D\left(\overline{b}, -\overline{b}_D\right) \cap \operatorname{span}\{a_t,\ t \in T\}\right) = \mathcal{E}^{op}(\overline{\pi}), \quad \text{for all } D \in \mathcal{M}_{\overline{\pi}}.$$

The last equality comes again from Theorem 4.1. ■

## 5. Conclusions and perspectives

At this moment we recall the parallelism between both formulae of Lipusc $\mathcal{F}(\overline{b})$ and Lipusc $\mathcal{F}^{op}(\overline{\pi})$, for $\overline{\pi} = (\overline{c}, \overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$, pointed out in Remark 1.1. The first one was established in [3] by strongly appealing to the convexity of gph $\mathcal{F}$, while a local directional convexity property of gph $\mathcal{F}^{op}$ has been used here to develop the second one. Indeed, [3] analyses another Lipschitz-type property, called there Hoffman stability, which is commented in the next subsection.

### 5.1. On the Hoffman stability of the optimal set

We say that $\mathcal{M} : Y \rightrightarrows X$, with $Y$ and $X$ being metric spaces, is *Hoffman stable at* $\bar{y}$ if there exists $\kappa \geq 0$ such that $\mathrm{d}(x, \mathcal{M}(\bar{y})) \leq \kappa\, \mathrm{d}(y, \bar{y})$ for all $y \in \operatorname{dom} \mathcal{M}$ and all $x \in \mathcal{M}(y)$, or, equivalently, if

$$\mathrm{d}(x, \mathcal{M}(\bar{y})) \leq \kappa\, \mathrm{d}\left(\bar{y}, \mathcal{M}^{-1}(x)\right) \quad \text{for all } x \in X. \tag{22}$$

The associated Hoffman modulus at $\bar{y} \in \operatorname{dom} \mathcal{M}$, $\operatorname{Hof} \mathcal{M}(\bar{y})$, is the infimum of constants $\kappa$ satisfying (22) and may be expressed as

$$\operatorname{Hof} \mathcal{M}(\bar{y}) = \sup_{(y,x) \in \operatorname{gph} \mathcal{M}} \frac{\mathrm{d}(x, \mathcal{M}(\bar{y}))}{\mathrm{d}\left(y, \bar{y}\right)} = \sup_{x \in X} \frac{\mathrm{d}(x, \mathcal{M}(\bar{y}))}{\mathrm{d}\left(\bar{y}, \mathcal{M}^{-1}(x)\right)}.$$

Theorem 4 in [3], again appealing to the convexity of $\operatorname{gph} \mathcal{F}$, establishes the equality

$$\operatorname{Lipusc} \mathcal{F}\left(\bar{b}\right) = \operatorname{Hof} \mathcal{F}\left(\bar{b}\right),$$

which is no longer true for our optimal set mapping $\mathcal{F}^{op}$, neither for $\mathcal{F}_{\bar{c}}^{op}$ (recall that $\operatorname{gph} \mathcal{F}^{op}$ and $\operatorname{gph} \mathcal{F}_{\bar{c}}^{op}$ are not convex), as the following example shows.

**Example 5.1:** Consider the parameterized problem of Example 1.1 with $\mathbb{R}^2$ being endowed with the Euclidean norm, and consider the following nominal parameters:

$$\bar{c} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{and} \quad \bar{b} = (-1, -1, -1, 1)'.$$

The reader can easily check that

$$\left\| (c, b) - \left(\bar{c}, \bar{b}\right) \right\| < \frac{1}{3} \Rightarrow \mathcal{F}^{op}(c, b) = \left\{ \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\}.$$

Thus, an ad hoc routine calculation gives

$$\operatorname{Lipusc} \mathcal{F}^{op}\left(\bar{c}, \bar{b}\right) = \sqrt{2}.$$

Of course Theorem 4.2 can be alternatively used for this computation.

Now, let us perturb $\bar{b}$ by considering $b_\mu := \bar{b} + \mu(-1, 0, 0, -1)'$ for $\mu \geq 2/3$. Then it is easy to check that $\begin{pmatrix} -1-\mu \\ 1-3\mu \end{pmatrix} \in \mathcal{F}^{op}(\bar{c}, b_\mu)$ for all $\mu \geq 2/3$ and, accordingly,

$$\operatorname{Hof} \mathcal{F}^{op}\left(\bar{c}, \bar{b}\right) \geq \operatorname{Hof} \mathcal{F}_{\bar{c}}^{op}\left(\bar{b}\right) \geq \lim_{\mu \to +\infty} \frac{\left\| \begin{pmatrix} -1-\mu \\ 1-3\mu \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\|}{\left\| b_\mu - \bar{b} \right\|_\infty}$$

$$= \lim_{\mu \to +\infty} \frac{\sqrt{\mu^2 + (3\mu - 2)^2}}{\mu} = \sqrt{10}.$$

The computation of Hof $\mathcal{F}^{op}(\bar{c},\bar{b})$ and Hof $\mathcal{F}_{\bar{c}}^{op}(\bar{b})$ remains as an open problem for further research.

### 5.2. Some repercussions on the stability of the optimal value

Let us finish the paper with some comments about perspectives on the behaviour of the optimal value function $\vartheta : \mathbb{R}^n \times \mathbb{R}^m \to [-\infty, +\infty]$, given by

$$\vartheta(\pi) := \inf\{c'x : x \in \mathcal{F}(b)\}, \quad \pi = (c,b) \in \mathbb{R}^n \times \mathbb{R}^m,$$

(with the convention $\vartheta(\pi) := +\infty$ when $\mathcal{F}(b) = \emptyset$); in our context of finite LP problems $\vartheta(\pi)$ is finite if and only if $\pi \in \mathrm{dom}\mathcal{F}^{op}$. Moreover, a well-known result in the literature establishes the continuity of $\vartheta$ restricted to its domain (see, e.g. [22, Theorem 4.5.2] for a proof based on the Berge's theory); i.e. the continuity of $\vartheta^R := \vartheta\mid_{\mathrm{dom}\mathcal{F}^{op}}$. Going further and regarding the subject of the current paper, the calmness modulus of $\vartheta^R$ provides a quantitative measure of the stability of the optimal value (indeed, a rate of variation with respect to perturbations of the data). Specifically, for $\bar{\pi} \in \mathrm{dom}\mathcal{F}^{op}$ this calmness modulus is given by

$$\mathrm{clm}\vartheta^R(\bar{\pi}) = \limsup_{\substack{\pi\to\bar{\pi} \\ \pi\in\mathrm{dom}\mathcal{F}^{op}}} \frac{|\vartheta(\pi) - \vartheta(\bar{\pi})|}{\|\pi - \bar{\pi}\|}.$$

This calmness modulus is analysed in [20], where point-based formulae for this quantity are provided in two stages: firstly, under RHS perturbations and, in a second stage, under canonical perturbations. That paper is focussed on a dual approach and formulae obtained there involve the maximum and minimum norms of dual optimal solutions (vectors of KKT multipliers) associated with the minimal KKT subsets of indices, $\mathcal{M}_{\bar{\pi}}$.

At this moment we point out the fact that appealing to Lipusc $\mathcal{F}^{op}(\bar{\pi})$ we may follow a primal approach to the estimation of $\mathrm{clm}\vartheta^R(\bar{\pi})$. The situation is particularly easy in the case of RHS perturbations as commented in the next lines: given $(\bar{c},b), (\bar{c},\bar{b}) \in \mathrm{dom}\mathcal{F}^{op}$ we have

$$|\vartheta(\bar{c},b) - \vartheta\left(\bar{c},\bar{b}\right)| = |\bar{c}'x - \bar{c}'p(x)| \leq \|\bar{c}\|_* \,\mathrm{d}\left(x, \mathcal{F}^{op}(\bar{\pi})\right), \tag{23}$$

where $x \in \mathbb{R}^n$ is any optimal solution of $(\bar{c},b)$ and $p(x) \in \mathbb{R}^n$ is any projection of $x$ on $\mathcal{F}^{op}(\bar{\pi})$. From (23), we obtain

$$\limsup_{\substack{b\to\bar{b} \\ (\bar{c},b)\in\mathrm{dom}\mathcal{F}^{op}}} \frac{|\vartheta(\bar{c},b) - \vartheta(\bar{\pi})|}{\|\pi - \bar{\pi}\|} \leq \|\bar{c}\|_* \,\mathrm{Lipusc}\,\mathcal{F}^{op}(\bar{\pi}).$$

Adding perturbations of $c$ and trying to reproduce inequalities of the form (23) in the context of canonical perturbations yields a different scenario, where the size of primal optimal solutions could play an important role, but this lies out of the scope of the present paper.

## Acknowledgments

The authors wish to thank the anonymous referees for their valuable critical comments, which have definitely improved the original version of the paper.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research has been partially supported by Grant PGC2018-097960-B-C21 from MICINN , Spain, and ERDF, 'A way to make Europe', European Union, and grant PROMETEO/2021/063 from Generalitat Valenciana, Spain.

## References

[1] Uderzo A. On the quantitative solution stability of parameterized set-valued inclusions. Set-Valued Var Anal. 2021;29(2):425–451.
[2] Cánovas MJ, López MA, Parra J, et al. Calmness of the feasible set mapping for linear inequality systems. Set-Valued Var Anal. 2014;22(2):375–389.
[3] Camacho J, Cánovas MJ, Parra J. From calmness to Hoffman constants for linear semi-infinite inequality systems. Available from: https://arxiv.org/pdf/2107.10000v2.pdf
[4] Dontchev AL, Rockafellar RT. Implicit functions and solution mappings: a view from variational analysis. New York: Springer; 2009.
[5] Klatte D, Kummer B. Nonsmooth equations in optimization: regularity, calculus, methods and applications. Dordrecht: Kluwer Academic; 2002. (Nonconvex optimization and its applications; vol. 60).
[6] Mordukhovich BS. Variational analysis and generalized differentiation, I: basic theory. Berlin: Springer; 2006.
[7] Rockafellar RT, Wets RJ-B. Variational analysis. Berlin: Springer; 1998.
[8] Goberna MA, López MA. Linear semi-infinite optimization. Chichester: Wiley; 1998.
[9] Gass S, Saaty T. Parametric objective function (part 2)-generalization. J Oper Res Soc Am. 1955;3:395–401.
[10] Saaty T, Gass S. Parametric objective function (part 1). J Oper Res Soc Am. 1954;2:316–319.
[11] Robinson SM. Some continuity properties of polyhedral multifunctions. Math Progr Study. 1981;14:206–214.
[12] Klatte D. Lipschitz continuity of infima and optimal solutions in parametric optimization: the polyhedral case. In: Guddat J, Jongen HTh, Kummer B, et al., editors. Parametric optimization and related topics. Berlin: Akademie; 1987. p. 229–248.
[13] Dontchev A, Zolezzi T. Well-posed optimization problems. Berlin: Springer; 1993. (Lecture Notes in Mathematics; vol. 1543.
[14] Azé D, Corvellec J-N. On the sensitivity analysis of Hoffman constants for systems of linear inequalities. SIAM J Optim. 2002;12(4):913–927.
[15] Hoffman AJ. On approximate solutions of systems of linear inequalities. J Res Natl Bur Stand. 1952;49(4):263–265.
[16] Klatte D, Thiere G. Error bounds for solutions of linear equations and inequalities. Z Oper Res. 1995;41:191–214.

[17] Peña J, Vera JC, Zuluaga LF. New characterizations of Hoffman constants for systems of linear constraints. Math Program. 2021;187(1–2):79–109.

[18] Zălinescu C. Sharp estimates for Hoffman's constant for systems of linear inequalities and equalities. SIAM J Optim. 2003;14(2):517–533.

[19] Li W. Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs. SIAM J Control Optim. 1994;32(1):140–153.

[20] Gisbert MJ, Canovas MJ, Parra J, et al. Calmness of the optimal value in linear programming. SIAM J Optim. 2018;28(3):2201–2221.

[21] Cánovas MJ, Henrion R, López MA, et al. Outer limit of subdifferentials and calmness moduli in linear and nonlinear programming. J Optim Theory Appl. 2016;169(3):925–952.

[22] Bank B, Guddat J, Klatte D, et al. Non-linear parametric optimization. Berlin: Akademie; 1982 and Birkhäuser: Basel; 1983.

# C Robust and continuous metric subregularity for linear inequality systems.

Computational Optimization and Applications
https://doi.org/10.1007/s10589-022-00437-0

# Robust and continuous metric subregularity for linear inequality systems

**J. Camacho**[1] · **M. J. Cánovas**[1] · **M. A. López**[2,3] · **J. Parra**[1]

Received: 4 June 2022 / Accepted: 15 November 2022



## Abstract

This paper introduces two new variational properties, robust and continuous metric subregularity, for finite linear inequality systems under data perturbations. The motivation of this study goes back to the seminal work by Dontchev, Lewis, and Rockafellar (2003) on the radius of metric regularity. In contrast to the metric regularity, the unstable continuity behavoir of the (always finite) metric subregularity modulus leads us to consider the aforementioned properties. After characterizing both of them, the radius of robust metric subregularity is computed and some insights on the radius of continuous metric subregularity are provided.


**Keywords** Radius of metric subregularity · Linear inequality systems · Calmness · Feasible set mapping

**Mathematics Subject Classification** 90C31 · 49J53 · 15A39 · 90C05

---

To the memory of Asen L. Dontchev.

---

This research has been partially supported by Grant PGC2018-097960-B-C2(1,2) from MICINN, Spain, and ERDF, "A way to make Europe", European Union, and Grant PROMETEO/2021/063 from Generalitat Valenciana, Spain.

---

✉ J. Parra
parra@umh.es

J. Camacho
j.camacho@umh.es

M. J. Cánovas
canovas@umh.es

M. A. López
marco.antonio@ua.es

[1] Center of Operations Research, Miguel Hernández University of Elche, 03202 Elche, Alicante, Spain

[2] Department of Mathematics, University of Alicante, 03071 Alicante, Spain

[3] CIAO, Federation University, Ballarat, Australia

Published online: 16 December 2022

## 1 Introduction

In this paper we firstly analyze continuity properties of the *modulus of metric subregularity* for linear inequality systems. This analysis motivates the introduction of new properties named as *robust* and *continuous metric subregularity*. Hereafter we frequently omit the word 'metric' for simplicity. We are particularly concerned with the radius (a sort of distance to ill-posedness) with respect to both properties, as well as with the connection with the *modulus of robust subregularity*. This topic is framed in the broader paradigm

$$\text{radius of } \mathcal{P} = \frac{1}{\text{modulus of } \mathcal{P}} \tag{1}$$

for some stability property, $\mathcal{P}$, which has been widely studied in different contexts (cf. [11, 12, 20]). We also draw the reader's attention to paper [4], devoted to the metric regularity of the inverse feasible set mapping for linear semi-infinite inequality systems (see [14]), where equality (1) holds. We advance that relation (1) does not always hold for the properties analyzed in the present paper.

We deal with (finite) linear inequality systems in $\mathbb{R}^n$ of the form

$$\sigma = \{a_t' x \leq b_t,\ t \in T = \{1, \dots, m\}\}, \tag{2}$$

where $x \in \mathbb{R}^n$ is the decision variable, regarded as a column-vector, and the prime represents transposition. System $\sigma$ will be identified with the pair of coefficient functions $(a, b)$, where $a = \left(a_t\right)_{t \in T} \in (\mathbb{R}^n)^T$ and $b = \left(b_t\right)_{t \in T} \in \mathbb{R}^T \equiv \mathbb{R}^m$. For the sake of simplicity in the notation we will identify $(\mathbb{R}^n)^T$ with $\mathbb{R}^{n \times m}$, so that function $a : t \mapsto a_t$ will be regarded as a matrix whose $t$-th column is $a_t$. In this way system $\sigma$ may be abbreviated as $a'x \leq b$. The space $\mathbb{R}^n$ is equipped with an arbitrary norm $\|\cdot\|$, while $\|\cdot\|_*$ stands for its dual norm, given by $\|u\|_* := \max_{\|x\|=1} |u'x|$, whose associated distance is denoted by $d_*$, and $\mathbb{R}^T$ is endowed with the supremum norm, $\|b\|_\infty := \max_{t \in T} |b_t|$.

In this framework, system $\sigma$ may be rewritten as the generalized equation

$$\mathcal{G}_a(x) := a'x + \mathbb{R}^m_+ \ni b, \tag{3}$$

where $\mathcal{G}_a : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $\mathbb{R}^m_+$ stands for the subset of elements of $\mathbb{R}^m$ with nonnegative coordinates. For each $a \in \mathbb{R}^{n \times m}$, the inverse multifunction

$$\mathcal{F}_a := \mathcal{G}_a^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n,$$

given by $x \in \mathcal{F}_a(b) \Leftrightarrow b \in \mathcal{G}_a(x)$, is nothing else but the *feasible set mapping* of system $\sigma$ under right-hand side perturbations.

Throughout the paper we work with a fixed consistent system denoted by $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ and a fixed $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$. We refer to $\overline{a}$, $\overline{b}$ and $\overline{x}$ as the nominal data. Given any property $\mathcal{P}$ of $\mathcal{G}_{\overline{a}}$ fulfilled at the nominal $\left(\overline{x}, \overline{b}\right) \in \operatorname{gph}\mathcal{G}_{\overline{a}}$ (where gph stands for graph), the radius of $\mathcal{P}$-stability at that point is defined as

$$\operatorname{rad}_{\mathcal{P}}\mathcal{G}_{\overline{a}}(\overline{x},\overline{b}) := \inf_{g\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)}\left\{\|g\| \;\middle|\; \mathcal{G}_{\overline{a}}+g \text{ does not have } \mathcal{P} \text{ at } \left(\overline{x},\overline{b}+g\left(\overline{x}\right)\right)\right\}, \tag{4}$$

where $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ stands for the space of linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ endowed with the norm subordinated to the norms under consideration in these spaces. This definition of radius is inspired by the one given in [12, Definition 1.4] for the metric regularity property in more general contexts; see also [11] for the property of metric subregularity. In order to adapt this concept to our current notation, let us identify a linear function $g$ with the matrix $g\in\mathbb{R}^{n\times m}$ such that $g(x)$ reads as $g'x$. In this way, denoting by $g_t$ the $t$-th column of $g$, we have

$$\|g\| = \max_{\|x\|=1}\|g'x\|_\infty = \max_{\|x\|=1}\max_{t\in T}\left|g_t'x\right| = \max_{t\in T}\left\|g_t\right\|_*.$$

***Remark 1*** Observe that $\left(\mathcal{G}_{\overline{a}}+g\right)(x) = (\overline{a}+g)'x+\mathbb{R}^m_+$ for all $x\in\mathbb{R}^n$. In other words,

$$\mathcal{G}_{\overline{a}}+g = \mathcal{G}_{\overline{a}+g}. \tag{5}$$

In this way, linear perturbations of $\mathcal{G}_{\overline{a}}$ translate into left-hand side (LHS, in brief) perturbations of the linear inequality system (2). Hence, assuming that $\overline{\sigma}\equiv\left(\overline{a},\overline{b}\right)$ satisfies a certain stability property $\mathcal{P}$, roughly speaking, $\operatorname{rad}_{\mathcal{P}}\mathcal{G}_{\overline{a}}(\overline{x},\overline{b})$ provides the infimum size of LHS perturbations of $\overline{\sigma}$ which cause failure of property $\mathcal{P}$ at the same point $\overline{x}$ with parameter $\overline{b}+g'\overline{x}$.

As already commented in [11, Example 1.1], when $\mathcal{P}$ is the metric subregularity property, then $\operatorname{rad}_{\mathcal{P}}\mathcal{G}_{\overline{a}}(\overline{x},\overline{b})$ is $+\infty$ as, for any $a\in\mathbb{R}^{n\times m}$, $\mathcal{G}_a$ is always metrically subregular at any $(x,b)\in\operatorname{gph}\mathcal{G}_a$ (this fact follows from the classical work of Robinson [24]). Therefore, the associated modulus is always finite, but not necessarily zero (in which case (1) fails). Indeed, the *subregularity modulus* of $\mathcal{G}_{\overline{a}}$ at $(\overline{x},\overline{b})$ is known to coincide with the *calmness modulus* of $\mathcal{F}_{\overline{a}}$ at $(\overline{b},\overline{x})$ which is computed through an implementable formula in Theorem 2; see Sect. 2 for further details. This comment motivates the fact of considering a different (more restrictive) property $\mathcal{P}$ which is not satisfied at all $((a,b),x)$ with $(x,b)\in\operatorname{gph}\mathcal{G}_a$. In this way, the continuous/robust subregularity come into play.

It is important to emphasize the practical repercussions of the metric subregularity property (and its counterparts in terms of *calmness* and *local error bounds*), for instance with respect to the convergence of algorithms. Just observe that finding a solution of our generalized equation $\mathcal{G}_a(x)\ni b$, with $b$ sufficiently close to the nominal $\overline{b}$, might be considerably difficult, whereas the residual (in our case, $\max_{t\in T}\left[a_t'x-\overline{b}_t\right]_+$, where $[\alpha]_+$ represents the positive part of $\alpha\in\mathbb{R}$) is much easier to compute or estimate. Hence, the metric subregularity of $\mathcal{G}_a$ at $(\overline{x},\overline{b})$ with constant $\kappa$ (see Sect. 2 for the definition) ensures the existence of such a solution whose distance to $\overline{x}$ is no longer than $\kappa$ times the residual. In particular, if we know an estimate for the rate of convergence of the residual to zero, then we can evaluate the rate of convergence of a sequence of approximate solutions to an exact solution. Two specific applications of calmness modulus are given in [6, Section 5] to the computation of some constants related to the convergence of certain optimization methods. The first one is focused on a particular procedure described in [19, Section 3.1] for a descent method, and the second deals with a concrete regularization scheme for mathematical programs with complementarity constraints introduced in [17]. In [23] we can find several references on the algorithmic repercussions of Hoffman constants (of a global nature, in contrast with the local character of calmness) as well as other related error bounds in establishing convergence properties of a variety of modern convex optimization algorithms.

Concerning interior-point methods, in [5, Section 4] the well-known *central path* construction associated with a linear programming problem is considered. If $\{(x(\mu), y(\mu), z(\mu)), \text{ for } \mu > 0\}$ denotes such a path and $\Lambda$ is the primal-dual solution set for the original problem (corresponding to $\mu = 0$), then, under appropriate hypotheses, [5, Theorem 4.1] shows that

$$d((x(\mu), y(\mu), z(\mu)), \Lambda) \leq \kappa\mu$$

for $\mu$ small enough, where $\kappa$ is directly related with the calmness modulus of a suitable feasible set mapping defined in terms of the nominal problem's data, so that constant $\kappa$ can be computed through an implementable procedure as it involves only fixed elements. A closely related problem is tackled in [1, Corollary 3], where an application to the convergence of a certain path-following algorithmic scheme, also in terms of calmness constants, is developed.

Aside the importance of the regularity concepts themselves, the study of related radii is also relevant. As already mentioned in [11, Section 5], the radius of nonsingularity of matrices is ultimately related to their condition number, and preconditioning is a highly efficient tool for enhancing computations in numerical linear algebra. In that paper the authors also suggest that different radius expressions could be utilized in procedures for conditioning of problems of feasibility and optimization. For a wider insight on conditioning, see [2].

The present paper is structured as follows: Sect. 2 sets up the necessary notation and preliminary results. Section 3 deals with the continuity behavior of the subregularity modulus of linear inequality systems under LHS perturbations, which is analyzed in two steps. First, Theorem 3 sheds light on the stability of the end set of polyhedra. As a consequence of this result, the continuity of the subregularity modulus is characterized in Theorem 4. In Sect. 4 we introduce the properties of robust and continuous subregularity and characterize them in Theorem 5 and Corollary 2 , respectively. Section 5 computes the radius of robust subregularity (Theorem 6) and gives some insights on the radius of continuous subregularity (see Example 4). The paper finishes with a section of conclusions and future research.

## 2 Preliminaries

Firstly, let us give some definitions and notations used along the paper. Given $S \subset \mathbb{R}^\ell$, $\ell \in \mathbb{N}$, conv$S$ denotes the convex hull of $S$. From the topological side, int$S$, cl$S$, and bd$S$ stand for the interior, the closure, and the boundary of $S$, respectively. Additionally, if $S$ is convex, its *end set* (introduced in [15]) is defined as

$$\operatorname{end} S := \{u \in \operatorname{cl} S \mid \nexists \mu > 1 \text{ such that } \mu u \in \operatorname{cl} S\}.$$

Here we recall the lower/inner and upper/outer limit of sets in the Painlevé-Kuratowski sense (cf. [22, p. 13], see also [25, p. 152]). Given two metric spaces $X$ and $A$ and a family of subsets of $X$, $\{X_a\}_{a \in A}$, we say $x \in \operatorname{Lim\,inf}_{a \to \overline{a}} X_a$ if for each sequence $\{a^r\}_{r \in \mathbb{N}}$ converging to $\overline{a}$ there exist $r_0 \in \mathbb{N}$ and $\{x^r\}_{r \geq r_0}$ verifying $x^r \in X_{a^r}$ for all $r \geq r_0$ and $\lim_{r \to \infty} x^r = x$. Regarding the outer limit, $x \in \operatorname{Lim\,sup}_{a \to \overline{a}} X_a$ if $x = \lim_{r \to \infty} x^r$ with $x^r \in X_{a^r}$ for some sequence $\{a^r\}_{r \in \mathbb{N}}$ converging to $\overline{a}$.

A set-valued mapping $\mathcal{M} : X \rightrightarrows Y$ between metric spaces (with both distances denoted by $d$) is said to be (metrically) *subregular* at $(\overline{x}, \overline{y}) \in \operatorname{gph} \mathcal{M}$ if there exist a constant $\kappa \geq 0$ together with a neighborhood $U$ of $\overline{x}$ such that

$$d(x, \mathcal{M}^{-1}(\overline{y})) \leq \kappa d(\overline{y}, \mathcal{M}(x)) \quad \text{for all } x \in U. \tag{6}$$

Here $d(x, C) := \inf_{y \in C} d(x, y)$ denotes the point-to-set distance, with $d(x, \emptyset) = +\infty$. Throughout the paper we assume $1/0 = +\infty$ and $1/(+\infty) = 0$. The infimum of constants $\kappa$ in (6), over the set of all possible $(\kappa, U)$ is called the *subregularity modulus* of $\mathcal{M}$ at $(\overline{x}, \overline{y})$ and it is denoted by $\operatorname{subreg} \mathcal{M}(\overline{x}, \overline{y})$.

The subregularity property of $\mathcal{M}$ at $(\overline{x}, \overline{y}) \in \operatorname{gph} \mathcal{M}$ is known to be equivalent to the calmness of its inverse $\mathcal{M}^{-1}$ at $(\overline{y}, \overline{x})$ and it is also known that $\operatorname{subreg} \mathcal{M}(\overline{x}, \overline{y})$ coincides with the calmness modulus of $\mathcal{M}^{-1}$ at $(\overline{y}, \overline{x})$ (cf. [13, 16, 18, 22, 25]).

Our focus is on mapping $\mathcal{G}_a$, with $a \in \mathbb{R}^{n \times m}$, given in (3), where a point-based formula (in terms of the given data) for its subregularity modulus is known (see Theorem 2). More specifically, given our nominal data $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ and $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, such expression of $\operatorname{subreg} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)$ appeals to the set of active indices at $\overline{x}$,

$$T_{\overline{\sigma}}(\overline{x}) := \left\{t \in T \mid \overline{a}_t' \overline{x} = \overline{b}_t\right\},$$

and involves the family $\mathcal{D}_{\overline{a}}$ (introduced in [9, Section 4] under the name $\mathcal{D}(\overline{x})$) of subsets $D \subset T_{\overline{\sigma}}(\overline{x})$ such that system

$$\left\{ \begin{array}{l} \overline{a}_t' d = 1, \ t \in D, \\ \overline{a}_t' d < 1, \ t \in T_{\overline{\sigma}}(\overline{x}) \setminus D, \end{array} \right\} \tag{7}$$

is consistent in the variable $d \in \mathbb{R}^n$. Observe that if $D \in \mathcal{D}_{\overline{a}}$ and $d$ is such a solution, then $\{\overline{a}_t, \ t \in D\}$ is contained in the hyperplane $\{z \in \mathbb{R}^n \mid d'z = 1\}$, which leaves $\{0_n\} \cup \{\overline{a}_t, \ t \in T_{\overline{\sigma}}(\overline{x}) \setminus D\}$ on one of its two associated open half-spaces.

Another key tool in the present paper is the family of sets $\mathcal{D}_{\overline{a}}^0$ (see [7, Section 3.2]) formed by all $D \subset T_{\overline{\sigma}}(\overline{x})$ such that system

$$\left\{\begin{array}{l}\overline{a}_t'd = 0,\ t \in D,\\ \overline{a}_t'd < 0,\ t \in T_{\overline{\sigma}}(\overline{x}) \setminus D,\end{array}\right\} \tag{8}$$

has nonzero solutions in the variable $d \in \mathbb{R}^n$. Now, if $D \in \mathcal{D}^0_{\overline{a}}$ and $d \in \mathbb{R}^n \backslash \{0_n\}$ satisfies (8), then the hyperplane $\{z \in \mathbb{R}^n \mid d'z = 0\}$ contains $\{0_n\} \cup \{\overline{a}_t,\ t \in D\}$ and leaves $\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x}) \setminus D\}$ on one of its two associated open half-spaces.

**Theorem 1** *Let* $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$. *Then*

$$\operatorname{end}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\} = \bigcup_{D \in \mathcal{D}_{\overline{a}}} \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\}. \tag{9}$$

***Proof*** It is a direct consequence of [21, Corollary 2.1 and Remark 2.3]. □

**Theorem 2** *Let* $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$. *Then*

$$\begin{aligned}\operatorname{subreg} \mathcal{G}_{\overline{a}}(\overline{x}, \overline{b}) &= d_*\left(0_n, \operatorname{end}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}\right)^{-1}\\ &= \max_{D \in \mathcal{D}_{\overline{a}}} d_*\left(0_n,\ \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\}\right)^{-1}.\end{aligned}$$

***Proof*** For $\overline{x} \in \operatorname{bd} \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$ the result follows from [9, Theorem 4] together with Theorem 1. If $\overline{x} \in \operatorname{int} \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, then $\mathcal{D}_{\overline{a}} = \{\emptyset\}$ and

$$0 = \operatorname{subreg} \mathcal{G}_{\overline{a}}(\overline{x}, \overline{b}) = d_*\left(0_n, \emptyset\right)^{-1}.$$

□

***Remark 2*** (On semi-infinite systems) For the sake of completeness, let us comment on some facts which may arise when the set $T$ indexing the constraints is infinite. To start with, in the case when $T$ is a compact metric space and $t \mapsto \left(a_t, b_t\right)$ is continuous on $T$, the set $T_{\overline{\sigma}}(\overline{x})$ is also compact and [21, Corollary 2.1 and Remark 2.3] ensures, denoting $\mathcal{B}(\overline{x}) := \bigcup_{D \in \mathcal{D}_{\overline{a}}} \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\}$, that

$$\mathcal{B}(\overline{x}) \subset \operatorname{end}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\} \subset \operatorname{cl}\mathcal{B}(\overline{x}),$$

hence,

$$d_*\left(0_n, \operatorname{end}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}\right) = \inf_{D \in \mathcal{D}_{\overline{a}}} d_*\left(0_n,\ \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\}\right),$$

which generalizes to this continuous semi-infinite case the second equality in Theorem 2. The first equality in Theorem 2 holds under the following regularity condition (see [21, Corollary 2.1, Remark 2.3 and Corollary 3.2]): "There exists a neighborhood $W$ of $\overline{x}$ such that

$$\mathcal{F}_{\overline{a}}(\overline{b}) \cap W = \left(\overline{x} + (\operatorname{cone}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\})^\circ\right) \cap W\text{''}, \tag{10}$$

where $X^{\circ}$ denotes the (negative) polar of $X$. Observe that this condition is held at all points of polyhedral sets and, for instance, at the vertex of the ice-cream cone. Indeed, the fulfilment of the condition (10) at all points of $\mathcal{F}_{\overline{a}}(\overline{b})$ is equivalent the fact that system $\overline{\sigma}$ is locally polyhedral (see [21, Corollary 3.3] and also [14, Section 5.2]). To the authors knowledge, the exact computation of subreg $\mathcal{G}_{\overline{a}}(\overline{x}, \overline{b})$ for more general semi-infinite systems via a point-based formula (in terms exclusively of the nominal data $\overline{a}, \overline{b}, \overline{x}$) remains as an open problem.

## 3 On the continuity of the subregularity modulus

Given the nominal data $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ and $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, we follow the perturbation structure of [12, p. 496]. In other words, we are considering arbitrary LHS perturbations $a - \overline{a}$ of $\overline{\sigma}$ and, in order to preserve feasibility of $\overline{x}$, the corresponding right-hand side perturbations are given by $\overline{b} + \left(a - \overline{a}\right)'\overline{x}$. In this way, (5) with $g = a - \overline{a}$ shifts $\left(\overline{x}, \overline{b}\right)$ to $\left(\overline{x}, \overline{b} + \left(a - \overline{a}\right)'\overline{x}\right)$. In the sequel, it will be useful to note that, denoting the set of active indices of system (2) at $x \in \mathcal{F}_a(b)$ by $T_{(a,b)}(x) := \left\{t \in T \mid a_t' x = b_t\right\}$, we have

$$T_{\left(a, \overline{b}+\left(a-\overline{a}\right)'\overline{x}\right)}\left(\overline{x}\right) = T_{\overline{\sigma}}\left(\overline{x}\right) \text{ for all } a \in \mathbb{R}^{n\times m}. \tag{11}$$

Now we introduce the function $\mathcal{S} : \mathbb{R}^{n\times m} \to \mathbb{R}$ given by

$$\mathcal{S}(a) := \operatorname{subreg} \mathcal{G}_a(\overline{x}, \overline{b} + \left(a - \overline{a}\right)'\overline{x}),\ a \in \mathbb{R}^{n\times m}. \tag{12}$$

In order to simplify the notation, in $\mathcal{S}(a)$ we omit the dependence on the nominal data $\overline{a}$, $\overline{b}$, and $\overline{x}$. Taking Theorem 2 and equality (11) into account, the end set of $\operatorname{conv}\left\{a_t, t \in T_{\overline{\sigma}}\left(\overline{x}\right)\right\}$ constitutes a crucial ingredient in the computation of $\mathcal{S}(a)$ for any $a \in \mathbb{R}^{n\times m}$. The following subsection is devoted to analyzing the stability behavior of this end set under perturbations of the $a_t$'s; this is a subject of independent interest.

### 3.1 Stability of the end set of polyhedra

This subsection is intended to be self-contained as far as our statements on $\left\{a_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right)\right\}$ could be given for any finite family in $\mathbb{R}^n$, not necessarily coming from a linear inequality system. In this way, set $T_{\overline{\sigma}}\left(\overline{x}\right)$ could be replaced by any finite index set. Accordingly, throughout this subsection we consider a finite index set $I$. For each $a = \left(a_t\right)_{t\in I} \in (\mathbb{R}^n)^I$, we define

$$E(a) := \operatorname{end} \operatorname{conv}\left\{a_t,\ t \in I\right\}, \tag{13}$$

and the families $\mathcal{D}_a$ and $\mathcal{D}_a^0$ coming from replacing in (7) and (8), respectively, $T_{\overline{\sigma}}\left(\overline{x}\right)$ by $I$ and $\overline{a}$ by $a$. Recall that, from Theorem 1,

$$E(a) = \bigcup_{D \in \mathcal{D}_a} \operatorname{conv}\{a_t,\ t \in D\},\ \text{for each } a \in (\mathbb{R}^n)^I. \tag{14}$$

The following lemma provides the Painlevé-Kuratowski upper/outer limit of $\mathcal{D}_a \subset 2^I$ (the subsets of $I$), with $a$ approaching $\overline{a}$; in it, the finite set $2^I$ is endowed with the discrete topology.

**Lemma 1** *Given* $\overline{a} \in (\mathbb{R}^n)^I$, *we have*:

(i) $\operatorname*{Lim\,sup}_{a \to \overline{a}} \bigcup_{D \in \mathcal{D}_a} \operatorname{conv}\{a_t,\ t \in D\} = \bigcup_{D \in \operatorname*{Lim\,sup}_{a \to \overline{a}} \mathcal{D}_a} \operatorname{conv}\{\overline{a}_t,\ t \in D\}$;

(ii) $\operatorname*{Lim\,sup}_{a \to \overline{a}} \mathcal{D}_a = \{S \subset I \mid \exists D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}^0_{\overline{a}} \text{ with } S \subset D\}$.

***Proof*** $(i)$ Let $u \in \operatorname{Lim\,sup}_{a \to \overline{a}} \bigcup_{D \in \mathcal{D}_a} \operatorname{conv}\{a_t,\ t \in D\}$ be written as $u = \lim_{r \to \infty} u^r$ with $u^r = \sum_{t \in D_r} \lambda_t^r a_t^r$, $\sum_{t \in D_r} \lambda_t^r = 1$, $\lambda_t^r \geq 0$ for all $t \in D_r$, for certain $D_r \in \mathcal{D}_{a^r}$ associated with some sequence $a^r \to \overline{a}$. Since $D_r \subset I$ (finite) for all $r$, it is not restrictive to assume (by taking a suitable subsequence) that $\{D_r\}_{r \in \mathbb{N}}$ is constant, say $D_r = D$, and $\{\lambda_t^r\}_r$ converges to some $\lambda_t \geq 0$ for each $t \in D$, hence $\sum_{t \in D} \lambda_t = 1$ and $u = \sum_{t \in D} \lambda_t \overline{a}_t$, with

$$D \in \operatorname*{Lim\,sup}_{r \to \infty} \mathcal{D}_{a^r} \subset \operatorname*{Lim\,sup}_{a \to \overline{a}} \mathcal{D}_a.$$

Now, let us prove '$\supset$'. Take $u = \sum_{t \in \widetilde{D}} \lambda_t \overline{a}_t$ with $\sum_{t \in \widetilde{D}} \lambda_t = 1$, $\lambda_t \geq 0$ for all $t \in \widetilde{D}$ and $\widetilde{D} \in \operatorname*{Lim\,sup}_{a \to \overline{a}} \mathcal{D}_a$. Then, there exists $a^r \to \overline{a}$ with $\widetilde{D} \in \mathcal{D}_{a^r}$ for all $r$, which entails

$$\bigcup_{D \in \mathcal{D}_{a^r}} \operatorname{conv}\{a_t^r,\ t \in D\} \ni \sum_{t \in \widetilde{D}} \lambda_t a_t^r \to u.$$

Accordingly, $u \in \operatorname{Lim\,sup}_{a \to \overline{a}} \bigcup_{D \in \mathcal{D}_a} \operatorname{conv}\{a_t,\ t \in D\}$.

$(ii)$ We start by proving the inclusion '$\supset$'. Let $S \subset I$ be such that $S \subset D$ for some $D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}^0_{\overline{a}}$ . If $D \in \mathcal{D}_{\overline{a}}$, take $p = 1$, otherwise ($D \in \mathcal{D}^0_{\overline{a}}$) take $p = 0$. In any case, let $d \in \mathbb{R}^n \setminus \{0_n\}$ be such that

$$\begin{cases} \overline{a}_t' d = p,\ t \in D, \\ \overline{a}_t' d < p,\ t \in I \setminus D. \end{cases}$$

Define the sequence by

$$a_t^r := \begin{cases} \overline{a}_t + \frac{1}{r} d, & t \in S, \\ \overline{a}_t, & t \in I \setminus S, \end{cases}$$

so that, denoting by $\|\cdot\|_2$ the Euclidean norm,

$$\left(a_t^r\right)' \left(p + \frac{1}{r} \|d\|_2^2\right)^{-1} d \begin{cases} = 1,\ t \in S, \\ < 1,\ t \in I \setminus S, \end{cases}$$

for all $r \in \mathbb{N}$; i.e., in both cases ($p = 0$ or $p = 1$), $S \in \mathcal{D}_{a^r}$ for all $r \in \mathbb{N}$. Therefore, $S \in \limsup_{a \to \overline{a}} \mathcal{D}_a$.

Let us prove the converse inclusion, '$\subset$'. Take any $S \in \operatorname{Lim\,sup}_{a \to \overline{a}} \mathcal{D}_a$ and assume the non-trivial case $S \neq \emptyset$. There exists some sequence $a^r \to \overline{a}$ such that $S \in \mathcal{D}_{a^r}$ for all $r \in \mathbb{N}$. Hence, for each $r$ there exists an associated $d^r \in \mathbb{R}^n \backslash \{0_n\}$ such that

$$\begin{cases} \left(a_t^r\right)' d^r = 1, \text{ if } t \in S, \\ \left(a_t^r\right)' d^r < 1, \text{ if } t \in I \setminus S. \end{cases} \tag{15}$$

If $\{d^r\}_{r \in \mathbb{N}}$ is bounded, then we can take a subsequence (denoted as the whole sequence for simplicity) converging to some $d \in \mathbb{R}^n$. Since $\left(a_t^r\right)' d^r = 1$ for $t \in S \neq \emptyset$, we conclude $\overline{a}_t' d = 1$ for those $t$, which entails $d \neq 0_n$, i.e.,

$$S \subset D := \left\{ t \in I \mid \overline{a}_t' d = 1 \right\}$$

(the inclusion may be strict) and $D \in \mathcal{D}_{\overline{a}}$.

In the case when $\{d^r\}_{r \in \mathbb{N}}$ is unbounded we may assume (by taking an appropriate subsequence if necessary) that $\|d^r\| \to +\infty$ and $\frac{d^r}{\|d^r\|} \to d \in \mathbb{R}^n$ with $\|d\| = 1$. Then, dividing both sides of (15) by $\|d^r\|$ and letting $r \to +\infty$ we obtain

$$S \subset D := \left\{ t \in I \mid \overline{a}_t' d = 0 \right\}$$

(the inclusion may be strict again) and $D \in \mathcal{D}_{\overline{a}}^0$. In any case $S \subset D$, with $D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0$, and the proof is complete. □

**Theorem 3** *Let $\overline{a} \in (\mathbb{R}^n)^I$. We have*

(i) $\displaystyle \operatorname*{Lim\,inf}_{a \to \overline{a}} E(a) = \bigcup_{D \in \mathcal{D}_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t \in D\} = E\left(\overline{a}\right)$;

(ii) $\displaystyle \operatorname*{Lim\,sup}_{a \to \overline{a}} E(a) = \bigcup_{D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0} \operatorname{conv}\{\overline{a}_t,\ t \in D\} \supset E\left(\overline{a}\right)$.

***Proof*** ($i$) The second equality is established Theorem 1 and it is clear from the definition that $\operatorname{Lim\,inf}_{a \to \overline{a}} E(a) \subset E\left(\overline{a}\right)$ as $E$ is closed-valued. In order to prove the converse inclusion, take $D \in \mathcal{D}_{\overline{a}}$ and $u = \sum_{t \in D} \lambda_t \overline{a}_t$ for some $\lambda = \left(\lambda_t\right)_{t \in D} \in \mathbb{R}_+^D$ with $\sum_{t \in D} \lambda_t = 1$ and let $d \in \mathbb{R}^n$ with $\overline{a}_t' d = 1$ for all $t \in D$ and $\overline{a}_t' d < 1$ for all $t \in I \backslash D$. Taking any $\{a^r\}_{r \in \mathbb{N}} \subset (\mathbb{R}^n)^I$ converging to $\overline{a}$, define, for each $r$, $w^r := \sum_{t \in D} \lambda_t a_t^r \in \operatorname{conv}\left\{a_t^r,\ t \in I\right\}$. Then

$$(w^r)' d \to u' d = \sum_{t \in D} \lambda_t \overline{a}_t' d = 1 \text{ as } r \to \infty.$$

On the other hand, for a fixed $r$ and any $v \in \operatorname{conv}\left\{a_t^r,\ t \in I\right\}$, writing $v = \sum_{t \in I} \gamma_t a_t^r$ with $\gamma \in \mathbb{R}_+^I$ and $\sum_{t \in I} \gamma_t = 1$, one has

$$v'd \leq \left( \sum_{t \in I} \gamma_t \overline{a}_t' d \right) + \left( \sum_{t \in I} \gamma_t \left\| a_t^r - \overline{a}_t \right\|_* \|d\| \right) \leq 1 + \left\| a^r - \overline{a} \right\| \|d\|;$$

in particular, $(w^r)'d \leq 1 + \left\| a^r - \overline{a} \right\| \|d\|$. This entails, for each $r \in \mathbb{N}$ such that $(w^r)'d > 0$, the existence of

$$\mu_r \in \left[ 1, \frac{1 + \left\| a^r - \overline{a} \right\| \|d\|}{(w^r)'d} \right]$$

such that $u^r := \mu_r w^r \in E(a^r)$. Consequently,

$$u = \lim_{r \to \infty} u^r \in \operatorname{Lim\,inf}_{r \to \infty} E(a^r).$$

Since $\{a^r\}_{r \in \mathbb{N}}$ has been arbitrarily chosen, one has $u \in \operatorname{Lim\,inf}_{a \to \overline{a}} E(a)$.

$(ii)$ comes from the Lemma 1 together with (14). □

***Remark 3*** $(i)$ Theorem 3 $(i)$ establishes the lower semicontinuity in the sense of Berge of mapping $E$ at $\overline{a}$ (equivalently, the inner semicontinuity at $\overline{a}$). Here we do not have an analogous result to Lemma 1 $(ii)$; more specifically, $\operatorname{Lim\,inf}_{a \to \overline{a}} \mathcal{D}_a$ may be strictly contained in $\mathcal{D}_{\overline{a}}$. For instance, for the family

$$\left\{ \overline{a}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \overline{a}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \overline{a}_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

one has $\operatorname{Lim\,inf}_{a \to \overline{a}} \mathcal{D}_a = \{\emptyset, \{1\}, \{3\}\}$, whereas $\mathcal{D}_{\overline{a}} = \{\emptyset, \{1\}, \{3\}, \{1,2,3\}\}$.

$(ii)$ The union in Theorem 3 $(ii)$ could be confined to those $D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0$ which are maximal with respect to the inclusion order. Moreover, the inclusion '$\supset$' may be strict as Example 1 below shows.

### 3.2 Lower and upper semicontinuity of the subregularity modulus

Let us consider the nominal data $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ and $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$. From now on in the paper, for each $a \in \mathbb{R}^{n \times m}$, we consider the end set defined in (13) in the particular case $I = T_{\overline{\sigma}}(\overline{x})$; i.e.,

$$E(a) := \operatorname{end} \operatorname{conv}\left\{ a_t,\ t \in T_{\overline{\sigma}}(\overline{x}) \right\}. \tag{16}$$

Observe that the index set $T_{\overline{\sigma}}(\overline{x})$ does not vary as $a$ varies by virtue of (11). From Theorem 2 we can write

$$\mathcal{S}(a) = d_*\left(0_n, E(a)\right)^{-1}, \text{ for any } a \in \mathbb{R}^{n \times m}. \tag{17}$$

**Theorem 4** *Let* $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$. *Then:*

(i) *$\mathcal{S}$ is lower semicontinuous at $\overline{a}$; i.e.*,

$$\liminf_{a\to\overline{a}} \mathcal{S}(a) = \left[ d_*\left( 0_n, \bigcup_{D\in\mathcal{D}_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t\in D\} \right) \right]^{-1} = \mathcal{S}(\overline{a});$$

(ii) *We have*

$$\limsup_{a\to\overline{a}} \mathcal{S}(a) = \left[ d_*\left( 0_n, \bigcup_{D\in\mathcal{D}_{\overline{a}}\cup\mathcal{D}^0_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t\in D\} \right) \right]^{-1} \geq \mathcal{S}(\overline{a}).$$

***Proof*** (*i*) Since $E$ is inner semicontinuous at $\overline{a}$, by [25, Proposition 5.11(b)] we have that $d_*\left(0_n, E(\cdot)\right)$ is upper semicontinuous at $\overline{a}$ and, accordingly, $d_*\left(0_n, E(\cdot)\right)^{-1}$ is lower semicontinuous at $\overline{a}$.

(*ii*)

Appealing to (17), we may write

$$\begin{aligned}
\limsup_{a\to\overline{a}} \mathcal{S}(a) &= \limsup_{a\to\overline{a}} d_*\left(0_n, E(a)\right)^{-1} \\
&= \left( \liminf_{a\to\overline{a}} d_*\left(0_n, E(a)\right) \right)^{-1} \\
&= d_*\left(0_n,\ \operatorname{Lim\ sup}_{a\to\overline{a}} E(a)\right)^{-1} \\
&= d_*\left(0_n, \bigcup_{D\in\mathcal{D}_{\overline{a}}\cup\mathcal{D}^0_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t\in D\}\right)^{-1}.
\end{aligned}$$

where the third equality follows from [25, Exercise 4.8] and the last one comes from Theorem 3(*ii*). □

**Corollary 1** *If* $\operatorname{Lim\ inf}_{a\to\overline{a}} E(a) = \operatorname{Lim\ sup}_{a\to\overline{a}} E(a) = E\left(\overline{a}\right)$, *i.e., if $E$ is continuous in the Painlevé-Kuratowski sense, then $\mathcal{S}$ is continuous at $\overline{a}$.*

***Proof*** As in the proof of statement (*ii*) in Theorem 4, and applying the current assumption, we have

$$\limsup_{a\to\overline{a}} \mathcal{S}(a) = d_*\left(0_n,\ \operatorname{Lim\ sup}_{a\to\overline{a}} E(a)\right)^{-1} = d_*\left(0_n, E\left(\overline{a}\right)\right)^{-1} = \mathcal{S}\left(\overline{a}\right).$$

□

***Remark 4*** Observe that:

(i) $\mathcal{S}$ may fail to be upper semicontinuous at $\overline{a}$, i.e., one can have

$$\limsup_{a\to\overline{a}} \mathcal{S}(a) > \mathcal{S}(\overline{a})$$

and the 'continuity gap' can be finite (Example 1) or infinite (Example 2).

(ii) The sufficient condition for the continuity of $\mathcal{S}$ given in Corollary 1 is not necessary. Just replace $\overline{a}_3$ in Example 1 below with $\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$.

***Example 1*** Consider the nominal system in $\mathbb{R}^2$, endowed with the Euclidean norm,

$$\overline{\sigma} := \left\{x_2 \leq 0,\ 2x_2 \leq 0,\ x_1 + 2x_2 \leq 0\right\},$$

and take $\overline{x} = 0_2$. One easily checks from Theorem 3(*ii*) that $\operatorname*{Lim\,sup}_{a \to \overline{a}} E(a) = \operatorname{conv}\left\{\begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\} \bigcup \operatorname{conv}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}\right\}$ while $E\left(\overline{a}\right) = \operatorname{conv}\left\{\begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$. Regarding function $\mathcal{S}$, from Theorem 4(*ii*), one has

$$\limsup_{a \to \overline{a}} \mathcal{S}(a) = 1 > \mathcal{S}(\overline{a}) = 1/2.$$

***Example 2*** Consider the nominal system in $\mathbb{R}^2$, endowed with the Euclidean norm,

$$\overline{\sigma} := \left\{x_2 \leq 0,\ -x_2 \leq 0\right\},$$

and take $\overline{x} = 0_2$. Again from Theorem 3(*ii*), we have $\operatorname{Lim\,sup}_{a \to \overline{a}} E(a) = \operatorname{conv}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right\}$ while $E\left(\overline{a}\right) = \left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix}\right\}$ and

$$\limsup_{a \to \overline{a}} \mathcal{S}(a) = +\infty,\ \mathcal{S}(\overline{a}) = 1.$$

## 4 Robust and continuous subregularity

Starting from the fact that

$$\limsup_{a \to \overline{a}} \mathcal{S}(a) \geq \mathcal{S}\left(\overline{a}\right) = \liminf_{a \to \overline{a}} \mathcal{S}(a),$$

and taking into account that the inequality above may be strict, this section is firstly devoted to characterizing the finiteness of the continuity gap, i.e., to characterize the condition $\limsup_{a \to \overline{a}} \mathcal{S}(a) < +\infty$, through an alternative (in principle, simpler) condition to the one which can be derived from the explicit formula of Theorem 4(*ii*). In a second stage, we provide a new approach to $\limsup_{a \to \overline{a}} \mathcal{S}(a)$ which allows us to interpret this quantity as a modulus of a robust-type metric subregularity property.

To start with, appealing to the definitions of $\mathcal{D}_{\overline{a}}$ and $\mathcal{D}_{\overline{a}}^0$, recall (7)-(8), one easily checks that

$$\bigcup_{D \in \mathcal{D}_{\overline{a}} \cup \mathcal{D}_{\overline{a}}^0} \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\} \subset \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right)\right\},$$

and the following proposition is an immediate consequence of this inclusion together with Theorem 4(*ii*).

**Proposition 1** *We have*

$$\limsup_{a\to\overline{a}} \mathcal{S}(a) \leq d_*\left(0_n, \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}\right)^{-1}. \tag{18}$$

The following example shows that the inequality in Proposition 1 may be strict.

***Example 3*** [9, Example 4] Let us consider the nominal system, in $\mathbb{R}^2$ endowed with the Euclidean norm,

$$\left\{x_1 \leq 0,\ x_2 \leq 0,\ x_1 + x_2 \leq 0\right\}$$

(associated with indices 1, 2, and 3, respectively) and the nominal solution $\overline{x} = 0_2$; in other words, recalling that each $\overline{a}_t$ is regarded as a column-vector,

$$\overline{a} = \left(\ \overline{a}_1\ \overline{a}_2\ \overline{a}_3\ \right) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \overline{b} = 0_3,$$

which entails $T_{\overline{\sigma}}(\overline{x}) = T = \{1,2,3\}$. Then, after a routine computation, we can show that $\|a-\overline{a}\| \leq \frac{1}{2\sqrt{2}}$ implies

$$\operatorname{end}\operatorname{conv}\left\{a_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\} = \operatorname{conv}\left\{a_1,a_3\right\} \cup \operatorname{conv}\left\{a_2,a_3\right\}. \tag{19}$$

Observe that the condition '$\|a-\overline{a}\| \leq \frac{1}{2\sqrt{2}}$' is not superfluous to ensure (19); indeed, if we take the unitary vector $u = \frac{1}{\sqrt{2}}\binom{1}{1}$ and, for any $\mu > \frac{1}{2\sqrt{2}}$, we consider the perturbed matrix $a^\mu = \left(\ \overline{a}_1 + \mu u\ \ \overline{a}_2 + \mu u\ \ \overline{a}_3 - \mu u\ \right)$, then we obtain $\operatorname{end}\operatorname{conv}\left\{a_t^\mu,\ t=1,2,3\right\} = \operatorname{conv}\left\{a_1,a_2\right\}$. Moreover, $\|a-\overline{a}\| \leq \frac{1}{2\sqrt{2}}$ also implies $d_*\left(0_2, \operatorname{bd}\operatorname{conv}\left\{a_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}\right) = d_*\left(0_2, \operatorname{conv}\left\{a_1,a_2\right\}\right)$. In particular,

$$\begin{aligned}\limsup_{a\to\overline{a}} \mathcal{S}(a) &= d_*\left(0_2, \operatorname{conv}\left\{\overline{a}_1,\overline{a}_3\right\} \cup \operatorname{conv}\left\{\overline{a}_2,\overline{a}_3\right\}\right)^{-1} = 1 \\ &< d_*\left(0_2, \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}\right)^{-1} = \sqrt{2}.\end{aligned}$$

In spite of not having equality in (18), $d_*\left(0_n, \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}\right)$ can be used to characterize the finiteness of $\limsup_{a\to\overline{a}} \mathcal{S}(a)$, as the following theorem establishes.

**Theorem 5** *Given* $\overline{\sigma} \equiv \left(\overline{a},\overline{b}\right)$ *and* $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, *the following statements are equivalent*:

(i) $\limsup_{a\to\overline{a}} \mathcal{S}(a)$ *is finite*;

(ii) $0_n \notin \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}$;

(iii) *There exist constants* $\kappa \geq 0$ *and* $\varepsilon > 0$ *along with a neighborhood* $U$ *of* $\overline{x}$ *such that*

$$d\left(x, \mathcal{F}_a\left(\overline{b} + (a - \overline{a})'\overline{x}\right)\right) \leq \kappa d\left(\overline{b} + (a - \overline{a})'\overline{x}, \mathcal{G}_a(x)\right) \tag{20}$$

for all $x \in U$ and all $a \in \mathbb{R}^n$ such that $\|a - \overline{a}\| < \varepsilon$.

Moreover, $\limsup_{a \to \overline{a}} \mathcal{S}(a)$ coincides with the infimum of constants $\kappa$ over the triplets $(\kappa, \varepsilon, U)$ satisfying (20) .

***Proof*** $(i) \Leftrightarrow (ii)$ Implication '$\Leftarrow$' is a direct consequence of Proposition 1. In order to prove the converse implication, assume reasoning by contradiction that $0_n \in \operatorname{bd} \operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$. By separation, consider $d \in \mathbb{R}^n \backslash \{0_n\}$ such that $\overline{a}_t' d \leq 0$ for all $t \in T_{\overline{\sigma}}(\overline{x})$, which necessarily satisfies $D := \left\{t \in T \mid \overline{a}_t' d = 0\right\} \neq \emptyset$. Consider an arbitrary $\varepsilon > 0$ and let $a_t^\varepsilon := \overline{a}_t + \varepsilon d$ for all $t \in T$. For $\widetilde{d} := \left(\varepsilon d' d\right)^{-1} d$ we clearly have $\left(a_t^\varepsilon\right)' \widetilde{d} = 1$ for all $t \in D$ and $\left(a_t^\varepsilon\right)' \widetilde{d} < 1$ for all $t \in T_{\overline{\sigma}}(\overline{x}) \backslash D$. Consequently, taking (11) into account, $D \in \mathcal{D}_{a^\varepsilon}$. Moreover, the fact that $0_n \in \operatorname{bd} \operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$ and the definition of $D$ easily imply $0_n \in \operatorname{conv}\left\{\overline{a}_t,\ t \in D\right\}$ and, then, $\varepsilon d \in \operatorname{conv}\left\{a_t^\varepsilon,\ t \in D\right\}$. Accordingly, recalling Theorem 2, we attain the contradiction

$$\mathcal{S}(a^\varepsilon) \geq d_*\left(0_n, \operatorname{conv}\left\{a_t^\varepsilon,\ t \in D\right\}\right)^{-1} \geq \left(\varepsilon \|d\|_*\right)^{-1} \to +\infty \text{ as } \varepsilon \downarrow 0.$$

$(i) \Leftrightarrow (iii)$ Implication '$\Leftarrow$' comes from the fact that any $\kappa \geq 0$ as in the statement is a subregularity constant for $\mathcal{G}_a$ at $\left(\overline{x}, \overline{b} + (a - \overline{a})'\overline{x}\right)$; in other words, taking $\kappa$ and $\varepsilon$ as in $(iii)$, we have

$$\mathcal{S}(a) \leq \kappa, \text{ whenever } \|a - \overline{a}\| < \varepsilon,$$

entailing $\limsup_{a \to \overline{a}} \mathcal{S}(a) \leq \kappa$. In order to prove the converse implication assume that $\limsup_{a \to \overline{a}} \mathcal{S}(a)$ is finite and take any $\kappa > \limsup_{a \to \overline{a}} \mathcal{S}(a)$. Let us prove that there exists a neighborhood $U$ of $\overline{x}$ along with $\varepsilon > 0$ such that (20) holds for all $x \in U$ and all $a \in \mathbb{R}^n$ with $\|a - \overline{a}\| < \varepsilon$. To do this we appeal to [10, Theorem 3], which shows –adapted to our current notation– that each $\mathcal{S}(a)$ is indeed a subregularity constant itself with an associated neighborhood $U_a$, which –see formula (8) in that paper–, taking into account (11) and the 'slack relationship' $\left[\overline{b}_t + \left(a_t - \overline{a}_t\right)'\overline{x}\right] - a_t'\overline{x} = \overline{b}_t - \overline{a}_t'\overline{x}$, is given by

$$U_a := \left\{ x \in \mathbb{R}^n \;\middle|\; \|x - \overline{x}\| < \delta_a := \inf_{t \notin T_{\overline{\sigma}}(\overline{x}),\ a_t \neq 0_n} \frac{\overline{b}_t - \overline{a}_t'\overline{x}}{2\|a_t\|_*} \right\}, \tag{21}$$

with the convention $\inf \emptyset := +\infty$. First, we analyze the case $\{t \in T \backslash T_{\overline{\sigma}}(\overline{x}) \mid \overline{a}_t \neq 0_n\} \neq \emptyset$. Now define $\rho := \min_{t \notin T_{\overline{\sigma}}(\overline{x})} \left(\overline{b}_t - \overline{a}_t'\overline{x}\right) > 0$ (recalling the finiteness of $T$) and take any

$$0 < \delta_1 < \min\left\{ \frac{\rho}{2\delta_{\overline{a}}}, \min\left\{ \left\|\overline{a}_t\right\|_* \mid t \notin T_{\overline{\sigma}}\left(\overline{x}\right),\ \overline{a}_t \neq 0_n \right\} \right\}.$$

Assume $\|a-\overline{a}\| < \delta_1$; then, $a_t \neq 0_n$ whenever $t \notin T_{\overline{\sigma}}\left(\overline{x}\right)$ and $\overline{a}_t \neq 0_n$. If, for some $t \notin T_{\overline{\sigma}}\left(\overline{x}\right)$, we have$a_t \neq 0_n$ and $\overline{a}_t = 0_n$, then,

$$\frac{\overline{b}_t - \overline{a}_t'\overline{x}}{2\left\|a_t\right\|_*} \geq \frac{\overline{b}_t - \overline{a}_t'\overline{x}}{2\delta_1} \geq \frac{\rho}{2\delta_1} \geq \delta_{\overline{a}}.$$

Therefore, $\|a-\overline{a}\| < \delta_1$ implies

$$\delta_a \geq \min_{t\notin T_{\overline{\sigma}}(\overline{x}),\ \overline{a}_t\neq 0_n} \frac{\overline{b}_t - \overline{a}_t'\overline{x}}{2\left(\left\|\overline{a}_t\right\|_* + \delta_1\right)} =: \delta_2.$$

Finally, take any $\delta_3 > 0$ satisfying $\|a-\overline{a}\| < \delta_3 \Rightarrow \mathcal{S}(a) \leq \kappa$. Then,

$$\left.\begin{array}{c} \|a-\overline{a}\| < \varepsilon := \min\left\{\delta_1, \delta_3\right\} \\ \|x-\overline{x}\| < \delta_2 \end{array}\right\} \Rightarrow (20) \text{ holds},$$

which establishes (*iii*) in this case.

Consider now the case when $T_{\overline{\sigma}}\left(\overline{x}\right) = T$. In this case, $\delta_a = +\infty$ for all $a \in \mathbb{R}^{n\times m}$. Hence, (20) holds whenever $\|a-\overline{a}\| < \delta_3$ (defined as above) and $x \in \mathbb{R}^n$. Finally, assume $\overline{a}_t = 0$ for all $t \in T\backslash T_{\overline{\sigma}}\left(\overline{x}\right) \neq \emptyset$. This entails $\overline{b}_t > 0$ for all $t \in T\backslash T_{\overline{\sigma}}\left(\overline{x}\right)$. Define $\rho := \min\left\{\overline{b}_t,\ t \in T\backslash T_{\overline{\sigma}}\left(\overline{x}\right)\right\}$. Then $\delta_a \geq \frac{\rho}{2\|a-\overline{a}\|}$ for all $a \in \mathbb{R}^{n\times m}$. Accordingly, (20) holds whenever $\|a-\overline{a}\| < \delta_3$ and $\|x-\overline{x}\| < \frac{\rho}{2\delta_3}$.

Moreover, reviewing the previous argument and observing that the proof is valid for any $\kappa > \limsup_{a\to\overline{a}} \mathcal{S}(a)$, we conclude that $\limsup_{a\to\overline{a}} \mathcal{S}(a)$ coincides with the infimum of constants $\kappa$ over the triplets $(\kappa, \varepsilon, U)$ satisfying (20). □

**Definition 1** Given system $\overline{\sigma} \equiv (\overline{a}, \overline{b})$ and $\overline{x} \in \mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, we say that

(i) $\mathcal{G}_{\overline{a}}$ is *robustly subregular* at $\left(\overline{x}, \overline{b}\right)$ if any of the three equivalent conditions of Theorem 5 holds. Regarding Theorem 5(*iii*), the infimum of constants $\kappa$ over the triplets $(\kappa, \varepsilon, U)$ satisfying (20) is called the *robust subregularity modulus* of $\mathcal{G}_{\overline{a}}$ at $\left(\overline{x}, \overline{b}\right)$ and will be denoted by $\operatorname{rob} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)$. As stated there,

$$\operatorname{rob} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) = \limsup_{a\to\overline{a}} \mathcal{S}(a). \tag{22}$$

(ii) $\mathcal{G}_{\overline{a}}$ is *continuously subregular* at $\left(\overline{x}, \overline{b}\right)$ if $\mathcal{S}$ is continuous at $\overline{a}$.

***Remark 5*** Condition (*iii*) in Theorem 5 looks like a kind of uniform regularity property with respect to $a$. Since the term uniform calmness has been already introduced in [3, Definition 1] (to be applied to the feasible set mapping $\mathcal{F}_{\overline{a}}$) with another meaning –uniformly with respect to $x$ in $\mathcal{F}_{\overline{a}}\left(\overline{b}\right)$–, we have preferred here the term robust.

See the comment preceding Corollary 3 below. Also observe that [4, Theorem 2.1], adapted to our current setting, entails that the metric regularity property of $\mathcal{G}_{\overline{a}}$ at $\left(\overline{x},\overline{b}\right)$ is characterized as $0_n \notin \operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}$. Indeed, [4, Corollary 3.2] provides the following expression for the the modulus of metric regularity:

$$\operatorname{reg}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right) = 1/d_*\left(0_n,\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}\right).$$

Accordingly, if $0_n \in \operatorname{int}\operatorname{conv}\left\{\overline{a}_t,\ t\in T_{\overline{\sigma}}(\overline{x})\right\}$, then $\mathcal{G}_{\overline{a}}$ is robustly regular but not metrically regular at $\left(\overline{x},\overline{b}\right)$.

**Corollary 2** *For the nominal data $\overline{\sigma}\equiv(\overline{a},\overline{b})$and $\overline{x}\in\mathcal{F}_{\overline{a}}\left(\overline{b}\right)$, the following statements are equivalent*:

- (i) $\mathcal{G}_{\overline{a}}$ *is continuously subregular at* $\left(\overline{x},\overline{b}\right)$;
- (ii) $\operatorname{rob}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)=\mathcal{S}(\overline{a})$;
- (iii) *It holds*

$$0\neq d_*\left(0_n,\bigcup_{D\in\mathcal{D}_{\overline{a}}}\operatorname{conv}\{\overline{a}_t,\ t\in D\}\right)=d_*\left(0_n,\bigcup_{D\in\mathcal{D}_{\overline{a}}\cup\mathcal{D}_{\overline{a}}^0}\operatorname{conv}\{\overline{a}_t,\ t\in D\}\right).$$

***Proof*** The proof comes straightforwardly from (22) and Theorem 4. □

## 5 Radii

In this section we formally introduce the radii announced in the introduction and succeed to compute one of them and give some hints on the other.

Following the notation introduced in (4), let us denote by $\operatorname{rad}_{\text{rob}}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)$ and $\operatorname{rad}_{\text{cont}}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)$ the radius of robust subregularity and continuous subregularity, respectively, of $\mathcal{G}_{\overline{a}}$ at $\left(\overline{x},\overline{b}\right)$. As a direct consequence of the definitions, continuous subregularity implies robust subregularity, and, hence,

$$\operatorname{rad}_{\text{rob}}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)\geq\operatorname{rad}_{\text{cont}}\mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right). \tag{23}$$

The next technical lemma provides a quite standard result that could be given with more generality. We state it as we need it, in $\mathbb{R}^n$ endowed with the dual norm $\|\cdot\|_*$.

**Lemma 2** *For* $i = 1, 2$, *let* $C_i = \operatorname{conv}\left\{ u_j^i,\ j = 1, \ldots, m \right\} \subset \mathbb{R}^n$. *Assume that for some* $u_0 \in \mathbb{R}^n$ *we have* $d_*\left(u_0, \operatorname{bd} C_1\right) = \delta > 0$ *and* $\max_{1 \leq j \leq m} \left\| u_j^1 - u_j^2 \right\|_* \leq \varepsilon < \delta$. *Then*

$$d_*\left(u_0, \operatorname{bd} C_2\right) \geq \delta - \varepsilon.$$

***Proof*** Firstly we consider the case $u_0 \notin \operatorname{cl} C_1$ and assume, reasoning by contradiction, that $d_*\left(u_0, \operatorname{bd} C_2\right) = \left\| u_0 - w_2 \right\|_* < \delta - \varepsilon$ for some $w_2 \in C_2$. Let us write $w_2 = \sum_{j=1}^m \lambda_j u_j^2$, with $\lambda_j \geq 0$ for all $j$ and $\sum_{j=1}^m \lambda_j = 1$. Take $w_1 := \sum_{j=1}^m \lambda_j u_j^1 \in C_1$. Then

$$\left\| u_0 - w_1 \right\|_* \leq \left\| u_0 - w_2 \right\|_* + \sum_{j=1}^m \lambda_j \left\| u_j^2 - u_j^1 \right\|_* < \delta - \varepsilon + \varepsilon = \delta,$$

contradicting the fact that $d_*\left(u_0, \operatorname{bd} C_1\right) = d_*\left(u_0, C_1\right) = \delta$.

Secondly, consider the case $u_0 \in \operatorname{int} C_1$, so that $u_0 + \delta B_* \subset C_1$. Then we will prove that $u_0 + (\delta - \varepsilon) B_* \subset C_2$. The argument here is similar to that of [8, Lemma 6], which we sketch here for completeness. Assume by contradiction that there exists $\widetilde{w}_2 \in \left(u_0 + (\delta - \varepsilon) B_*\right) \backslash C_2$. Then we can strictly separate $\widetilde{w}_2$ and $C_2$, so that there exists $p \in \mathbb{R}^n$ such that

$$p' \widetilde{w}_2 < p' u_j^2 \text{ for all } j = 1, \ldots, m. \tag{24}$$

Take $z \in \mathbb{R}^n$ with $\|z\|_* = \varepsilon$ and $p'z = \|p\| \|z\|_*$. Then

$$\left\| \widetilde{w}_2 - z - u_0 \right\|_* \leq \left\| \widetilde{w}_2 - u_0 \right\|_* + \|z\|_* \leq \delta - \varepsilon + \varepsilon = \delta,$$

entailing $\widetilde{w}_2 - z \in C_1$. Thus write $\widetilde{w}_2 - z = \sum_{j=1}^m \widetilde{\lambda}_j u_j^1$, with $\widetilde{\lambda}_j \geq 0$ for all $j$ and $\sum_{j=1}^m \widetilde{\lambda}_j = 1$. Therefore, recalling (24), we attain the contradiction

$$p' \widetilde{w}_2 - p'z = \sum_{j=1}^m \widetilde{\lambda}_j p' u_j^2 + \sum_{j=1}^m \widetilde{\lambda}_j p' \left( u_j^1 - u_j^2 \right) > p' \widetilde{w}_2 - \|p\| \varepsilon = p' \widetilde{w}_2 - p'z.$$

□

**Theorem 6** *Assume that* $\mathcal{G}_{\overline{a}}$ *is robustly subregular at* $\left(\overline{x}, \overline{b}\right)$. *Then*

$$\operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) = d_*\left(0_n,\ \operatorname{bd} \operatorname{conv}\left\{ \overline{a}_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right) \right\}\right).$$

***Proof*** Write $\delta := d_*\left(0_n, \operatorname{bd} \operatorname{conv}\left\{ \overline{a}_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right) \right\}\right)$ and pick any $a \in \mathbb{R}^n$ with $\left\| a - \overline{a} \right\| < \delta$. Then Lemma 2 entails

$$d_*\left(0_n, \operatorname{bd} \operatorname{conv}\left\{ a_t,\ t \in T_{\overline{\sigma}}\left(\overline{x}\right) \right\}\right) \geq \delta - \left\| a - \overline{a} \right\|,$$

which in particular implies $0_n \notin \operatorname{bd}\operatorname{conv}\left\{a_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$. Therefore, Theorem 5 yields that $\mathcal{G}_a$ is robustly subregular at $\left(\overline{x}, \overline{b} + (a - \overline{a})'\overline{x}\right)$. This, together with (5), proves $\operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) \geq \delta$.

In order to prove the converse inequality, take $u \in \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$ with $\|u\|_* = \delta$. Let us show that $0_n \in \operatorname{bd}\operatorname{conv}\left\{\overline{a}_t - u,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$. It is clear that $0_n \in \operatorname{conv}\left\{\overline{a}_t - u,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$ since, if $u = \sum_{t \in T} \lambda_t \overline{a}_t$, with $\lambda_t \geq 0$ for all $t \in T$ and $\sum_{t \in T} \lambda_t = 1$, then $0_n = \sum_{t \in T} \lambda_t \left(\overline{a}_t - u\right)$. Moreover, a very similar calculation shows that if there existed $\varepsilon > 0$ with $\varepsilon B_* \subset \operatorname{conv}\left\{\overline{a}_t - u,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$, i.e. $0_n \in \operatorname{int}\operatorname{conv}\left\{\overline{a}_t - u,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$, then we would have the contradiction $u + \varepsilon B_* \subset \operatorname{conv}\left\{\overline{a}_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}$. Now Theorem 5 ensures that $\mathcal{G}_{(\overline{a}_t - u)_{t \in T}}$ is not robustly subregular at $\left(\overline{x}, \left(\overline{b}_t - u'\overline{x}\right)_{t \in T}\right)$. Accordingly, $\operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) \leq \delta$. □

As an immediate consequence of Theorem 6, together with Proposition 1, (22), and Example 3, we obtain the following result. Observe that this is the opposite inequality to that obtained in general for the radius of metric regularity in [12, Theorem 1.5] (where radius $\geq 1/\text{modulus}$). The last part of this result, which comes from Theorem 5(ii) together with the definition of robust subregularity, asserts that if $\mathcal{G}_{\overline{a}}$ is robustly subregular at $\left(\overline{x}, \overline{b}\right)$, then the robust subregularity radius is positive. Otherwise, the term 'robust' would sound inappropriate.

**Corollary 3** *One has*

$$\operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) \leq \frac{1}{\operatorname{rob} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)},$$

*and the inequality may be strict. Moreover,* $\operatorname{rob} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) < +\infty$ *implies* $\operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) > 0$.

The next example shows that inequality (23) may be strict, as well as provides some hints for the study of the radius of continuous subregularity.

***Example 4*** Let us consider the nominal system, in $\mathbb{R}^2$ endowed with the Euclidean norm,

$$\overline{\sigma} = \left\{x_1 + 2x_2 \leq 0,\ x_1 + 4x_2 \leq 0,\ 6x_1 + 5x_2 \leq 0\right\}$$

(associated with indices 1, 2 and 3, respectively) and the nominal solution $\overline{x} = 0_2$. Hence, $T_{\overline{\sigma}}(\overline{x}) = \{1, 2, 3\}$. Let us check that

$$\operatorname{rad}_{\text{cont}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) = \frac{1}{\sqrt{10}} < \operatorname{rad}_{\text{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) = \sqrt{5},$$

where the last equality comes from Theorem 6. Indeed, writing $\overline{\sigma} \equiv \left(\overline{a}, \overline{b}\right)$, the minimum perturbation size from $\overline{a}$ making the perturbed $a$ have proportional $a_1$ and $a_2$ (i.e., making $a_1$ and $a_2$ belong to some straight line in $\mathbb{R}^2$ passing through the origin), obtained by computing, with the well-known Ascoli formula, the distance from $\overline{a}_1$ and $\overline{a}_2$ to such line, is $1/\sqrt{10}$. This minimum perturbation size is attained at the following system $\sigma_\mu \equiv (a^\mu, b^\mu)$ for $\mu = 1/\sqrt{10}$, where the perturbed $a_1^{1/\sqrt{10}}$ and $a_2^{1/\sqrt{10}}$ are the only possible ones with the aimed property and the perturbed $a_3^{1/\sqrt{10}}$ is irrelevant (so that we have kept it as in the nominal system), and where we have followed the criterion $b^\mu = \overline{b} + \left(a^\mu - \overline{a}\right)'\overline{x}$. Define

$$\sigma_\mu = \begin{cases} \left(1 - \frac{3}{\sqrt{10}}\mu\right)x_1 + \left(2 + \frac{1}{\sqrt{10}}\mu\right)x_2 \le 0, \\ \left(1 + \frac{3}{\sqrt{10}}\mu\right)x_1 + \left(4 - \frac{1}{\sqrt{10}}\mu\right)x_2 \le 0, \\ 6x_1 + 5x_2 \le 0. \end{cases}$$

Thus, we have:

$$\begin{aligned} \|a - \overline{a}\| \le \frac{1}{\sqrt{10}} \Rightarrow \mathcal{S}(a) &= d_*\left(0_2, \text{end conv}\left\{a_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}\right)^{-1} \\ &= d_*\left(0_2, \text{conv}\left\{a_2, a_3\right\}\right)^{-1} = \|a_2\|_*^{-1}. \end{aligned}$$

Hence, $\mathcal{S}$ is continuous in the open ball centered at $\overline{a}$ with radius $1/\sqrt{10}$. On the other hand,

$$\text{end conv}\left\{a_1^\mu, a_2^\mu, a_3^\mu\right\} = \begin{cases} \text{conv}\left\{a_2^\mu, a_3^\mu\right\}, & \text{if } 0 \le \mu \le \frac{1}{\sqrt{10}}, \\ \text{conv}\left\{a_1^\mu, a_2^\mu\right\} \cup \text{conv}\left\{a_2^\mu, a_3^\mu\right\}, & \text{if } \frac{1}{\sqrt{10}} < \mu \le \frac{5\sqrt{10}}{11}, \end{cases}$$

which entails

$$\mathcal{S}(a^\mu) = \begin{cases} \left\|a_2^\mu\right\|_*^{-1}, \text{ if } 0 \le \mu \le 1/\sqrt{10}, \\ \left\|a_1^\mu\right\|_*^{-1}, \text{ if } 1/\sqrt{10} < \mu \le 5\sqrt{10}/11. \end{cases}$$

Consequently, $\lim_{\mu \to \left(1/\sqrt{10}\right)^+} \mathcal{S}(a^\mu) = \left\|a_1^{1/\sqrt{10}}\right\|_*^{-1} = \frac{\sqrt{10}}{7} > \mathcal{S}\left(a^{1/\sqrt{10}}\right) = \left\|a_2^{1/\sqrt{10}}\right\|_*^{-1} = \frac{\sqrt{10}}{13}$. Putting all together, we have $\text{rad}_{\text{cont}}\, \mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right) = \frac{1}{\sqrt{10}}$.

## 6 Conclusions and further research

The following diagram is intended to provide a complete picture of the main results of this work. In it, for each $a \in \mathbb{R}^{n\times m}$, $\mathcal{S}(a)$ and $E(a)$ represent the subregularity modulus and the end set defined in (12) and (16), respectively.

Our starting background is:

- Theorem 1 (from [21, Corollary 2.1 and Remark 2.3] ), which establishes

$$E(a) = \bigcup_{D\in\mathcal{D}_{\overline{a}}} \operatorname{conv}\{a_t,\ t\in D\},\ a\in\mathbb{R}^{n\times m}.$$

- Theorem 2 (derived from [9, Theorem 4]), which yields

$$\mathcal{S}(a) = d_*\left(0_n, E(a)\right)^{-1},\ a\in\mathbb{R}^{n\times m}.$$

Hereafter, $(I)$ and $(II)$ are used as abbreviations as follows:

$$(I) := \left[d_*\left(0_n, \bigcup_{D\in\mathcal{D}_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t\in D\}\right)\right]^{-1},$$

$$(II) := \left[d_*\left(0_n, \bigcup_{D\in\mathcal{D}_{\overline{a}}\cup\mathcal{D}^0_{\overline{a}}} \operatorname{conv}\{\overline{a}_t,\ t\in D\}\right)\right]^{-1}.$$

The next diagram summarizes the main results of this paper, all of them being new except equality $\mathcal{S}(\overline{a}) = d_*\left(0_n, E(\overline{a})\right)^{-1} = (I)$.

DIAGRAM OF RESULTS ON THE STABLE BEHAVIOR OF $\mathcal{S}$

$$\begin{aligned}
\liminf_{a\to\overline{a}} \mathcal{S}(a) &= \mathcal{S}(\overline{a}) = d_*\left(0_n, E(\overline{a})\right)^{-1} = (I)\\
&\le \underbrace{\limsup_{a\to\overline{a}} \mathcal{S}(a)}_{\substack{\| \\ \operatorname{rob} \mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)}} = d_*\left(0_n,\ \operatorname{Lim\ sup}_{a\to\overline{a}} E(a)\right)^{-1} = (II)\\
&\le \underbrace{d_*\left(0_n, \operatorname{bd}\operatorname{conv}\{a_t,\ t\in T_{\overline{\sigma}}(\overline{x})\}\right)^{-1}}_{\substack{\| \\ [\operatorname{rad}_{\operatorname{rob}} \mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)]^{-1}}}\\
&\le [\operatorname{rad}_{\operatorname{cont}} \mathcal{G}_{\overline{a}}\left(\overline{x},\overline{b}\right)]^{-1}.
\end{aligned}$$

Finally, let us point out some remarkable facts:

- Examples 1 and 2 show that the gap with respect to the first inequality of the diagram may be finite or infinite, respectively. Examples 3 and 4 show that the second and the third inequalities may be strict.

- Regarding the second inequality, from Theorem 5, the gap cannot be infinite, as condition $d_*\left(0_n, \operatorname{bd}\operatorname{conv}\left\{a_t,\ t \in T_{\overline{\sigma}}(\overline{x})\right\}\right) = 0$ is equivalent to $\limsup_{a \to \overline{a}} \mathcal{S}(a) = +\infty$.
- The modulus and radius of robust subregularity, $\operatorname{rob}\mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)$ and $\operatorname{rad}_{\operatorname{rob}}\mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)$, are computed through point-based formulae (only involving the nominal data $\left(\overline{a}, \overline{b}, \overline{x}\right)$, not appealing to elements in a neighborhood).
- The problem of finding a point-based formula for the radius of continuous subregularity, $\operatorname{rad}_{\operatorname{cont}}\mathcal{G}_{\overline{a}}\left(\overline{x}, \overline{b}\right)$, remains as an open problem; Example 4 provides some hints for future research, as far as it illustrates some of the difficulties which may arise in the computation of this radius.

**Acknowledgements** The authors are indebted to the anonymous referees and the Coordinating Editor for their valuable critical comments, which have definitely improved the original version of the paper.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Data availability and conflict of interest statement** This work is of theoretical nature and has not analyzed or generated any datasets. The authors have no conflicts of interest to declare in relation to this article.



## References

1. Argáez, C., Cánovas, M.J., Parra, J.: Calmness of linear constraint systems under structured perturbations with an application to the path-following scheme. Set-Valued Var. Anal. **29**, 839–860 (2021)
2. Bürgisser, P., Cucker, F.: Condition. The geometry of numerical algorithms, grundlehren der mathematischen wissenschaften fundamental principles of mathematical sciences. Springer, New York (2013)
3. Camacho, J., Cánovas, M. J., Parra, J.: From calmness to Hoffman constants for linear inequality systems, SIAM. J. Optim. **32**, 2859–2878 (2022)
4. Cánovas, M.J., Dontchev, A.L., López, M.A., Parra, J.: Metric regularity of semi-infinite constraint systems. Math. Progr. **104B**, 329–346 (2005)
5. Cánovas, M.J., Hall, J.A.J., López, M.A., Parra, J.: Calmness of partially perturbed linear systems with an application to the central path. Optimization **68**, 465–483 (2019)
6. Cánovas, M.J., Hantoute, A., Parra, J., Toledo, F.J.: Calmness modulus of fully perturbed linear programs. Math. Progr. **158A**, 267–290 (2016)
7. Cánovas, M.J., Henrion, R., López, M.A., Parra, J.: Outer limit of subdifferentials and calmness moduli in linear and nonlinear programming. J. Optim. Theory Appl. **169**, 925–952 (2016)
8. Cánovas, M.J., López, M.A., Parra, J., Toledo, F.J.: Distance to ill-posedness and the consistency value of linear semi-infinite inequality systems. Math. Progr. **103A**, 95–126 (2005)

9. Cánovas, M.J., López, M.A., Parra, J., Toledo, F.J.: Calmness of the feasible set mapping for linear inequality systems. Set-Valued Var. Anal. **22**, 375–389 (2014)
10. Cánovas, M.J., Parra, J., Rückmann, J.-J., Toledo, F.J.: Point-based neighborhoods for sharp calmness constants in linear programming. Set-Valued Var. Anal. **25**, 757–772 (2017)
11. Dontchev, A.L., Gfrerer, H., Kruger, A.Y., Outrata, J.V.: The radius of metric subregularity. Set-Valued Var. Anal. **28**, 451–473 (2020)
12. Dontchev, A.L., Lewis, A.S., Rockafellar, R.T.: The radius of metric regularity. Trans. Am. Math. Soc. **355**, 493–517 (2003)
13. Dontchev, A.L., Rockafellar, R.T.: Implicit functions and solution mappings: a view from variational analysis. Springer, New York (2009)
14. Goberna, M.A., López, M.A.: Linear semi-infinite optimization. Wiley, Chichester (UK) (1998)
15. Hu, H.: Characterizations of the strong basic constraint qualifications. Math. Op. Res. **30**, 956–965 (2005)
16. Ioffe, A.D.: Variational analysis of regular mappings. theory and applications. Springer monographs in mathematics. Springer, Cham (2017)
17. Kadrani, A., Dussault, J.-P., Benchakroun, A.: A new regularization scheme for mathematical programs with complementarity constraints. SIAM J. Optim. **20**, 78–103 (2009)
18. Klatte, D., Kummer, B.: Nonsmooth equations in optimization: regularity, calculus, methods and applications, nonconvex optim. appl. 60, Kluwer academic, Dordrecht, The Netherlands, (2002)
19. Klatte, D., Kummer, B.: Optimization methods and stability of inclusions in Banach spaces. Math. Progr. **117B**, 305–330 (2009)
20. Kruger, A.Y., López, M.A., Théra, M.A.: Perturbation of error bounds. Math. Progr. Ser. B **168**, 533–554 (2018)
21. Li, M.H., Meng, K.W., Yang, X.Q.: On error bound moduli for locally Lipschitz and regular functions. Math. Progr. Ser. A **171**, 463–487 (2018)
22. Mordukhovich, B.S.: Variational analysis and generalized differentiation, i: basic theory. Springer, Berlin (2006)
23. Peña, J., Vera, J.C., Zuluaga, L.F.: New characterizations of Hoffman constants for systems of linear constraints. Math. Progr. **187**, 79–109 (2021)
24. Robinson, S.M.: Some continuity properties of polyhedral multifunctions. Math. Progr. Study **14**, 206–214 (1981)
25. Rockafellar, R.T., Wets, R.J.-B.: Variational analysis. Springer, Berlin (1998)

# D Feasibility problems via paramonotone operators in a convex setting.

# Feasibility problems via paramonotone operators in a convex setting*

J. Camacho † M.J. Cánovas† J.E. Martínez-Legaz ‡

J. Parra†

**Abstract**

This paper is focused on some properties of paramonotone operators on Banach spaces and their application to certain feasibility problems for convex sets in a Hilbert space and convex systems in the Euclidean space. In particular, it shows that operators that are simultaneously paramonotone and bimonotone are constant on their domains, and this fact is applied to tackle two particular situations. The first one, closely related to simultaneous projections, deals with a finite amount of convex sets with an empty intersection and tackles the problem of finding the smallest perturbations (in the sense of translations) of these sets to reach a nonempty intersection. The second is focused on the distance to feasibility; specifically, given an inconsistent convex inequality system, our goal is to compute/estimate the smallest right-hand side perturbations that reach feasibility. We advance that this work derives lower and upper estimates of such a distance, which become the exact value when confined to linear systems.

**Key words.** Distance function, convex inequalities, distance to feasibility, paramonotone operators, displacement mapping

**Mathematics Subject Classification:** 47N10, 47H05, 52A20, 90C31, 49K40

*This research has been partially supported by Grant PGC2018-097960-B-C21 from MICINN, Spain, and ERDF, "A way to make Europe", European Union, and Grant PROMETEO/2021/063 from Generalitat Valenciana, Spain. The third author has also been partially supported by the Severo Ochoa Programme for Centres of Excellence in R&D [SEV-2015-0563]. He is affiliated with MOVE (Markets, Organizations and Votes in Economics).

†Center of Operations Research, Miguel Hernández University of Elche, 03202 Elche (Alicante), Spain (j.camacho@umh.es, canovas@umh.es, parra@umh.es).

‡Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, and BGSMath, Barcelona, Spain (JuanEnrique.Martinez.Legaz@uab.cat).

## 1 Introduction

The present paper is focused on paramonotone operators with applications to certain feasibility problems for convex sets in a Hilbert space and convex inequality systems in $\mathbb{R}^n$. To start with, we recall some basic properties of operators in Banach spaces. Let $X$ be a real Banach space, with topological dual $X^*$, and denote by $\langle \cdot, \cdot \rangle$ the corresponding canonical pairing. A set-valued operator $T : X \rightrightarrows X^*$ is said to be *monotone* if

$$\langle x - y, x^* - y^* \rangle \geq 0 \text{ whenever } (x, x^*), (y, y^*) \in \text{gph}T,$$

where $\text{gph}T := \{(x, x^*) \in X \times X^* : x^* \in T(x)\}$ is the graph of $T$. In the case when both $T$ and $-T$ are monotone, then $T$ is called *bimonotone.* If $T$ is monotone and, in addition, gph$T$ is maximal in the sense of inclusion order, it is said to be *maximally monotone.* A well-known example of maximally monotone operator is the subdifferential operator of a proper, lower semicontinuous (lsc, for short), convex function $f : X \to ]-\infty, +\infty]$, denoted by $\partial f$ (see Section 2 for details). Monotone operators are fundamental tools of nonlinear analysis and optimization; see, e.g., the books [1, 6, 7, 18, 20, 22, 23]. A monotone operator $T$ is called *paramonotone* if the following implication holds:

$$\left. \begin{array}{c} (x, x^*), (y, y^*) \in \text{gph}T \\ \langle x - y, x^* - y^* \rangle = 0 \end{array} \right\} \Rightarrow (x, y^*), (y, x^*) \in \text{gph}T.$$

The term paramonotonicity was introduced in [12] (although the condition was previously presented in [9] without a name). The initial motivation for the introduction of paramonotone operators comes from its crucial role regarding interior point methods for variational inequalities (see again [9] and [12], and also [14]). Some important examples of paramonotone operators are gathered in Section 2. At this moment, let us mention that subdifferentials of proper lsc convex functions enjoy this property (see [14, Proposition 2.2] in the Euclidean space and [3, Fact 3.1] for its extension to Banach spaces).

Looking at the applications of Sections 4 and 5, we are interested in operators of the form $T_1 \cap (-T_2)$, where $T_1, T_2 : X \rightrightarrows X^*$ are paramonotone, which are also paramonotone and, additionally, bimonotone; this fact entails that $T_1 \cap (-T_2)$ is constant on its domain (as shown in Corollary 9); recall that the domain of an operator $T$ is given by $\text{dom}T := \{x \in X \mid T(x) \neq \emptyset\}$. Observe that

$$\text{dom}(T_1 \cap (-T_2)) = \{x \in X \mid 0 \in (T_1 + T_2)(x)\}, \tag{1}$$

which, in the particular case $T_i = \partial f_i$, $i = 1, 2$, where the $f_i$'s are proper, lsc and convex, is known to coincide with

$$\arg\min (f_1 + f_2), \tag{2}$$

i.e., with the set of (global) minima of $f_1 + f_2$, provided that a regularity condition ensuring $\partial f_1 + \partial f_2 = \partial(f_1 + f_2)$ is satisfied. These comments easily generalize to the sum of a finite number of functions (see Section 3 for details) and are applied to particular problems of the form

$$\underset{x\in X}{\text{minimize}} \quad \sum_{i=1}^{m} f_i(x), \tag{3}$$

where all $f_i$'s are proper lsc convex functions on $X$.

Now we present two applications discussed in the paper. The first one is developed in a Hilbert space $X$ whose norm, associated with the corresponding inner product, is denoted by $\|\cdot\|$. It deals with a finite number of nonempty closed convex sets $S_1, S_2, ..., S_m$ such that $\cap_{i=1}^m S_i = \emptyset$ and is focused on the optimization problem given by

$$\underset{x\in X}{\text{minimize}} \quad \sum_{i=1}^{m} \alpha_i d(x, S_i)^p, \tag{4}$$

where $\alpha_i > 0$, $i = 1, ..., m$; without loss of generality we assume $\sum_{i=1}^{m} \alpha_i = 1$, $p \geq 1$ and $d(x, S_i)$ denotes the distance from point $x$ to set $S_i$, $i = 1, ..., m$. The following proposition establishes that (4) is equivalent to the problem:

$$\begin{array}{ll} \text{minimize} & \|u\|_{\alpha,p} \\ \text{subject to} & \cap_{i=1}^m (S_i + u_i) \neq \emptyset, \\ & u = (u_1, ..., u_m) \in X^m, \end{array} \tag{5}$$

where $\|u\|_{\alpha,p}$ denotes the weighted $p$-norm in space $X^m$ defined as

$$\|u\|_{\alpha,p} = \left(\sum_{i=1}^{m} \alpha_i \|u_i\|^p\right)^{1/p}. \tag{6}$$

This equivalence was already observed in [2, Section 4] for Euclidean spaces; we include a proof for the sake of completeness.

**Proposition 1** *A point $\overline{u} = (\overline{u}_1, ..., \overline{u}_m) \in X^m$ is an optimal solution to (5) if and only if there exist an optimal solution $\overline{x}$ to (4) such that $\overline{u}_i = \overline{x} - P_i(\overline{x})$, $i = 1, ..., m$, with $P_i(\overline{x})$ being the best approximation of $\overline{x}$ in $S_i$.*

**Proof.** Let $\overline{u} = (\overline{u}_1, ..., \overline{u}_m) \in X^m$ be an optimal solution to (5), and take $\overline{x} \in \cap_{i=1}^m (S_i + \overline{u}_i)$. There exist $s_i \in S_i$, $i = 1, ..., m$, such that $\overline{x} = s_i + \overline{u}_i$. For every $x \in X$, we have

$$\begin{aligned} \sum_{i=1}^{m} \alpha_i d(\overline{x}, S_i)^p \leq \sum_{i=1}^{m} \alpha_i \|\overline{x} - s_i\|^p = \sum_{i=1}^{m} \alpha_i \|\overline{u}_i\|^p \leq \sum_{i=1}^{m} \alpha_i \|x - P_i(x)\|^p \\ = \sum_{i=1}^{m} \alpha_i d(x, S_i)^p. \end{aligned} \tag{7}$$

To justify the latter inequality, observe that $\cap_{i=1}^{m}\left(S_i + x - P_i\left(x\right)\right) \neq \emptyset$, because from the equalities $x = P_i\left(x\right) + x - P_i\left(x\right),\ i = 1, ..., m$, it immediately follows that $x \in \cap_{i=1}^{m}\left(S_i + x - P_i\left(x\right)\right)$. Thus, $\overline{x}$ is an optimal solution to (4). Furthermore, setting $x = \overline{x}$, we also deduce that $d\left(\overline{x}, S_i\right) = \left\|\overline{x} - s_i\right\|,\ i = 1, ..., m$, that is, $s_i = P_i\left(\overline{x}\right)$, so that $\overline{u}_i = \overline{x}_i - s_i = \overline{x}_i - P_i\left(\overline{x}\right)$.

Conversely, let $\overline{x}$ be an optimal solution to (4), $u$ be a feasible solution to (5), and take $x \in \cap_{i=1}^{m}\left(S_i + u_i\right)$. Then, there exist $s_i \in S_i$, $i = 1, ..., m$, such that $x = s_i + u_i$, and we have

$$\begin{aligned} \|u\|_{\alpha,p}^{p} &= \sum_{i=1}^{m} \alpha_i \left\|u_i\right\|^p = \sum_{i=1}^{m} \alpha_i \left\|x - s_i\right\|^p \geq \sum_{i=1}^{m} \alpha_i d\left(x, S_i\right)^p \geq \sum_{i=1}^{m} \alpha_i d\left(\overline{x}, S_i\right)^p \\ &= \sum_{i=1}^{m} \alpha_i \left\|\overline{x} - P_i\left(\overline{x}\right)\right\|^p, \end{aligned} \tag{8}$$

which shows that the point $\overline{u} = \left(\overline{u}_1, ..., \overline{u}_m\right)$, with $\overline{u}_i := \overline{x}_i - P_i\left(\overline{x}\right),\ i = 1, ..., m$, is an optimal solution to (5). ■

According to Proposition 1, problem (4) is equivalent to that of finding the smallest translations of the sets $S_i$ that achieve a nonempty intersection.

The second application, developed in Section 5, deals with convex inequality systems in $\mathbb{R}^n$ parameterized with respect to the right-hand side (RHS, in brief),

$$\sigma\left(b\right) := \left\{g_i(x) \leq b_i,\ i = 1, \ldots, m\right\}, \tag{9}$$

where $x \in \mathbb{R}^n$ is the vector of decision variables and, for each $i \in 1, \ldots, m$, $g_i : \mathbb{R}^n \to \mathbb{R}$ is a (finite-valued) convex function on $\mathbb{R}^n$, and $(b_i)_{i=1,\ldots,m} \equiv b \in \mathbb{R}^m$. Taking a nominal $\overline{b} \in \mathbb{R}^m$ such that $\sigma\left(\overline{b}\right)$ is inconsistent (i.e., there is no $x \in \mathbb{R}^n$ satisfying all inequalities of $\sigma\left(\overline{b}\right)$), our aim is to estimate the distance in $\mathbb{R}^m$ endowed with any $p$-norm, with $p \geq 2$, from $\overline{b}$ to the set of parameters $b$ such that $\sigma\left(b\right)$ is consistent. This *distance to feasibility* can be computed by solving the following problem:

$$\underset{x \in X}{\text{minimize}} \quad \sum_{i=1}^{m} \left[g_i\left(x\right) - b_i\right]_+^p, \tag{10}$$

which also adapts to the format of (3). Sharper results are presented for linear systems when $p = 2$.

At this point, we summarize the structure of the paper. Section 2 gathers some background on convex sets, convex functions, and monotone operators, which is appealed to in the remaining sections. Section 3 explores some new properties of paramonotone operators and, in particular, analyzes the simultaneous fulfilment of paramonotonicity and bimonotonicity. The problem of simultaneous projections -see (4) and (5)- is tackled in Section 4, while the distance to feasibility of convex systems under RHS perturbations is dealt with in Section 5.

## 2 Preliminaries

Let $X$ be a real Banach space and $f : X \to ]-\infty, +\infty]$ be a proper lsc convex function. We denote by $\mathrm{dom} f := \{x \in X \mid f(x) < +\infty\}$ the domain of function $f$. Recall that the subdifferential operator of $f$, $\partial f : X \rightrightarrows X^*$, assigns to each $x \in \mathrm{dom} f$ the (possibly empty) set $\partial f(x)$ formed by all $x^* \in X^*$ (called subgradients) such that

$$f(y) - f(x) \geq \langle y - x, x^* \rangle, \text{ for all } y \in X.$$

When $x \notin \mathrm{dom} f$ we define $\partial f(x) := \emptyset$; in this way the domain of the set-valued mapping, $\mathrm{dom}\partial f$, is always contained in $\mathrm{dom} f$. Associated with $f$, its Fenchel conjugative function $f^* : X^* \to ]-\infty, +\infty]$ is given by

$$f^*(x^*) = \sup\{\langle x, x^* \rangle - f(x) \mid x \in X\}.$$

Recall that the Young-Fenchel inequality writes as $f^*(x^*) + f(x) \geq \langle x, x^* \rangle$ for all $x \in X$.

For completeness, we gather in the following theorem some well-known results about $\partial f$ and $f^*$ in Banach spaces used in the paper. They can be traced out from different references dealing with convex analysis in infinite dimensional spaces. Here, we mainly cite the books [6, 15, 17, 23]. From now on, $\mathrm{int} A$ denotes the interior of $A \subset X$ (where, as usual, $\subset$ is understood as $\subseteq$) and the zero vector of $X^*$ is denoted by just 0.

**Theorem 2** *Let $f : X \to ]-\infty, +\infty]$ be a proper lsc convex function. Then we have:*

($i$) [6, Proposition 4.1.5] $f$ *is continuous at* $x$ *if and only if* $x \in \mathrm{int\ dom} f$;

($ii$) [15, Theorem 3.2.15] $\mathrm{int\ dom} f \subset \mathrm{dom}\partial f$;

($iii$) [15, Proposition 3.2.17] $x \in \arg\min f$ *if and only if* $0 \in \partial f(x)$;

($iv$) [15, Exercise 4.2.15] $\partial f$ *is maximally monotone;*

($v$) [3, Fact 3.1] (see [14, Proposition 2.2] for finite dimensions) $\partial f$ *is paramonotone;*

($vi$) [15, Proposition 5.31] *For any* $x \in X$, *we have the equivalence*

$$x^* \in \partial f(x) \Leftrightarrow f^*(x^*) + f(x) = \langle x, x^* \rangle;$$

*(indeed, this statement does not require convexity);*

($vii$) [15, Theorem 3.4.2] (see also [19, Theorem 3]) *Let* $g : X \to ]-\infty, +\infty]$ *be any proper convex function. If* $(\mathrm{int\ dom} f) \cap \mathrm{dom} g \neq \emptyset$, *then we have the subdifferential sum rule*

$$\partial f(x) + \partial g(x) = \partial(f + g)(x), \text{ whenever } x \in \mathrm{dom}\partial f \cap \mathrm{dom}\partial g,$$

*(indeed, '$\supset$' is the nontrivial inclusion, as* $\partial f(x) + \partial g(x) \subset \partial(f+g)(x)$ *comes directly from the definition of subdifferential; moreover, the lower semicontinuity of* $f$ *is not needed).*

Recall that, for an arbitrary monotone operator $T : X \rightrightarrows X^*$, an lsc convex function $h : X \times X^* \to ]-\infty, +\infty]$ is said to be a *representative function* of $T$ if

$$h(x, x^*) \begin{cases} = \langle x, x^* \rangle, & \text{if } (x, x^*) \in \operatorname{gph} T, \\ > \langle x, x^* \rangle, & \text{elsewhere.} \end{cases} .$$

Operators for which a representative function exists are called representable monotone. For a detailed study of representable monotone operators we refer to [16], where this notion was introduced.

**Remark 3** From Theorem $2(vi)$, observe that, if $f : X \to ]-\infty, +\infty]$ is a proper lsc convex function, the function $h_f : X \times X^* \to ]-\infty, +\infty]$ defined by

$$h_f(x, x^*) = f(x) + f^*(x^*), \text{ for } (x, x^*) \in X \times X^* \tag{11}$$

is a representative function of $\partial f$. More generally, every maximally monotone operator is representable, as far as its well-known Fitzpatrick function is a representative function (see, e.g. [6, Section 9.1.2] for details). An easy consequence of this fact is that intersections of arbitrary collections of maximally monotone operators are representable, too. According to [16, Corollary 32], in finite-dimensional spaces only such intersections are representable. This is no longer true in infinite dimensional spaces, as proved in [21, Theorem 11.2].

The rest of this section is devoted to recall some results about metric projections and, in order to ensure existence and uniqueness of the best approximation to closed convex sets, we assume that $X$ is a Hilbert space. Here, $\|\cdot\|$ denotes the norm associated with the corresponding inner product $\langle \cdot, \cdot \rangle$. Given any nonempty closed convex set $S \subset X$, we denote by $P_S : X \to X$ the *metric projection on* $S$, which assigns to each $x \in X$ its (unique) best approximation in $S$, denoted by $P_S(x)$, i.e., $P_S(x)$ is the unique point of $S$ such that

$$\|x - P_S(x)\| = d(x, S) = \min\{\|x - s\| : s \in S\}.$$

(Observe that we write $P_S : X \to X$ instead $P_S : X \rightrightarrows X$ due to its single-valuedness.) It is well-known that function $x \mapsto d(x, S)$, denoted for convenience by $d_S : X \to [0, +\infty[$, is a continuous convex function. Recall that, for a continuous convex function, $f : X \to \mathbb{R}$, applying [6, Corollary 4.2.5 ], we deduce that $f$ is Gâteaux differentiable at a point $x$ if and only if $\partial f(x)$ reduces to a singleton, i.e. $\partial f(x) = \{\nabla f(x)\}$; see [6, Section 2 ] for details. In our applications, the facts that the subdifferentials $\partial d_S(x)$ or $\partial d_S^2(x)$ reduce to a singleton are crucial. Accordingly, condition $(i)$ in the following proposition is stated directly in these terms (instead of Gâteaux differentiability). From now on, $N_S(x)$ denotes the normal cone to $S$ at $x$ which is given by

$$N_S(x) := \{x^* \in X^* \mid \langle s - x, x^* \rangle \leq 0, \ s \in S\}, \tag{12}$$

$B^*$ denotes the closed unit ball in $X^*$ and $\operatorname{bd} S$ the boundary of $S$.

**Proposition 4** *Let $X$ be a Hilbert space and $\emptyset \neq S \subset X$ a closed set. Then, we have*

$(i)$ [6, Corollary 4.2.5 and Theorem 4.5.7] *$S$ is convex if and only if $\partial d_S^2(x)$ is singleton for all $x \in X$; in such a case,*

$$\nabla d_S^2(x) = 2\left(x - P_S(x)\right).$$

$(ii)$ [15, Proposition 4.1.5] (see also [13, Section 1]) *If $S$ is convex, then*

$$\partial d_S(x) = \begin{cases} \{0\}, & \text{if } x \in \operatorname{int} S, \\ N_S(x) \cap B^*, & \text{if } x \in \operatorname{bd} S, \\ \left\{ \|x - P_S(x)\|^{-1} \left(x - P_S(x)\right) \right\}, & \text{if } x \notin S. \end{cases}$$

## 3 On paramonotone and bimonotone operators

This section provides some results, appealed to in Sections 4 and 5, about operators which are simultaneously paramonotone and bimonotone on a real Banach space $X$. To start with, we provide some basic results on these two properties separately.

**Proposition 5** *Let $T : X \rightrightarrows X^*$ be a representable monotone operator. The following statements are equivalent:*

$(i)$ *$T$ is paramonotone;*

$(ii)$ *For any representative function $h$ of $T$, the following implication holds:*

$$\left.\begin{array}{c} (x, x^*), (y, y^*) \in \operatorname{gph} T \\ \langle x - y, x^* - y^* \rangle = 0 \end{array}\right\} \Rightarrow h(x, y^*) + h(y, x^*) = h(x, x^*) + h(y, y^*). \quad (13)$$

$(iii)$ *There exists a representative function $h$ of $T$ such that (13) holds.*

**Proof.** $(i) \Rightarrow (ii)$. Consider any representative function of $T$, $h$, and take $(x, x^*), (y, y^*) \in \operatorname{gph} T$, with $\langle x - y, x^* - y^* \rangle = 0$. Then, the paramonotonicity entails $y^* \in T(x)$ and $x^* \in T(y)$, yielding

$$h(x, y^*) + h(y, x^*) = \langle x, y^* \rangle + \langle y, x^* \rangle = \langle x, x^* \rangle + \langle y, y^* \rangle = h(x, x^*) + h(y, y^*);$$

$(ii) \Rightarrow (iii)$. Straightforward.

$(iii) \Rightarrow (i)$. Let $h$ be a representative function of $T$ satisfying (13). Let $x, y \in X$, $x^* \in T(x)$, $y^* \in T(y)$ and suppose $\langle x - y, x^* - y^* \rangle = 0$. Hence,

$$h(x, y^*) + h(y, x^*) = h(x, x^*) + h(y, y^*) = \langle x, x^* \rangle + \langle y, y^* \rangle = \langle x, y^* \rangle + \langle y, x^* \rangle.$$

Since $h(x, y^*) \geq \langle x, y^* \rangle$ and $h(y, x^*) \geq \langle y, x^* \rangle$, these inequalities actually hold as equalities, yielding $y^* \in T(x)$ and $x^* \in T(y)$. ■

**Remark 6** Observe that the paramonotonicity of the subdifferential operator $\partial f$ of a proper lsc convex function $f$ can be alternatively deduced from Proposition 5. Just consider the representative function $h_f$ introduced in (11), which is separable and, hence, one always has

$$h_f(x, y^*) + h_f(y, x^*) = h_f(x, x^*) + h_f(y, y^*).$$

Other examples of paramonotone operators are mappings of the form $I - A$ where $I$ is the identity mapping and $A$ is nonexpansive (see [4, Theorem 6.1]); see also [14, Section 3 ] for the analysis of paramonotonicity of affine functions in $\mathbb{R}^n$.

**Proposition 7** *For an operator* $T : X \rightrightarrows X^*$, *the following conditions are equivalent:*
$(i)$ $T$ *is bimonotone;*
$(ii)$ $\langle x - y, x^* - y^* \rangle = 0$, *whenever* $(x, x^*), (y, y^*) \in \operatorname{gph} T$;
$(iii)$ *There exist monotone operators* $T_1$ *and* $T_2$ *such that* $T = T_1 \cap (-T_2)$.

**Proof.** $(i) \Leftrightarrow (ii)$ is trivial.
$(i) \Rightarrow (iii)$. Write $T = T \cap (-(-T))$.
$(iii) \Rightarrow (i)$. If $T_1$ is monotone, so is $T$, since $T \subset T_1$. Analogously, since

$$-T = T_2 \cap (-T_1),$$

$-T$ is monotone, too. ■

**Proposition 8** *For a representable monotone operator* $T : X \rightrightarrows X^*$, *the following conditions are equivalent:*
$(i)$ $T$ *is bimonotone;*
$(ii)$ *For any representative function* $h$ *of* $T$, *the following implication holds*

$$h(x, x^*) + h(y, y^*) = \langle x, x^* \rangle + \langle y, y^* \rangle \Rightarrow h(x, x^*) + h(y, y^*) = \langle x, y^* \rangle + \langle y, x^* \rangle. \tag{14}$$

$(iii)$ *There exists a representative function* $h$ *of* $T$ *such that* (14) *holds.*

**Proof.** $(i) \Rightarrow (ii)$. Consider any representative function $h$ of $T$, and assume that

$$h(x, x^*) + h(y, y^*) = \langle x, x^* \rangle + \langle y, y^* \rangle.$$

Hence, $h(x, x^*) = \langle x, x^* \rangle$ and $h(y, y^*) = \langle y, y^* \rangle$, that is, $(x, x^*), (y, y^*) \in \operatorname{gph} T$, yielding $\langle x - y, x^* - y^* \rangle = 0$. Consequently,

$$\langle x, y^* \rangle + \langle y, x^* \rangle = \langle x, x^* \rangle + \langle y, y^* \rangle = h(x, x^*) + h(y, y^*).$$

$(ii) \Rightarrow (iii)$. Straightforward.
$(iii) \Rightarrow (i)$. Let $x, y \in X$, $(x, x^*), (y, y^*) \in \operatorname{gph} T$. We then have

$$h(x, x^*) + h(y, y^*) = \langle x, x^* \rangle + \langle y, y^* \rangle,$$

and hence $\langle x, y^*\rangle + \langle y, x^*\rangle = h(x, x^*) + h(y, y^*) = \langle x, x^*\rangle + \langle y, y^*\rangle$ , from which the equality $\langle x - y, x^* - y^*\rangle = 0$ immediately follows. ■

From now on, symbol '$\perp$' represents orthogonality; specifically, given any subsets $A \subset X$ and $B \subset X^*$, $A \perp B$ means that $\langle x, x^*\rangle = 0$ for any $(x, x^*) \in A \times B$, whereas $A^\perp := \{x^* \in X^* \mid \langle x, x^*\rangle = 0, \text{ for all } x \in A\}$ and $B^\perp := \{x \in X \mid \langle x, x^*\rangle = 0, \text{ for all } x^* \in B\}$.

**Corollary 9** *For $T : X \rightrightarrows X^*$, the following conditions are equivalent:*
($i$) *$T$ is paramonotone and bimonotone*;
($ii$) *$T$ is monotone and constant on its domain*;
($iii$) $(\operatorname{dom} T - \operatorname{dom} T) \perp (\operatorname{range} T - \operatorname{range} T)$ *and* $\operatorname{gph} T = \operatorname{dom} T \times \operatorname{range} T$.

**Proof.** ($i$) $\Rightarrow$ ($ii$). Let $x, y \in \operatorname{dom} T$, and take $x^* \in T(x)$, $y^* \in T(y)$. By bimonotonicity, we have $\langle x - y, x^* - y^*\rangle = 0$. Hence, by paramonotonicity, $y^* \in T(x)$ and $x^* \in T(y)$. This proves that $T(x) = T(y)$.

($ii$) $\Rightarrow$ ($i$). The paramonotonicity of $T$ is an obvious consequence of its being constant on its domain. To prove bimonotonicity, let $(x, x^*), (y, y^*) \in \operatorname{gph} T$. Monotonicity implies $\langle x - y, x^* - y^*\rangle \geq 0$, and we can interchange $x^*$ and $y^*$, since $T(x) = T(y)$; therefore $\langle x - y, x^* - y^*\rangle = 0$.

($ii$) $\Leftrightarrow$ ($iii$) Comes from the fact that $T$ is constant on $\operatorname{dom} T$ if and only if $\operatorname{gph} T = \operatorname{dom} T \times \operatorname{range} T$. ■

**Corollary 10** *Let $T : X \rightrightarrows X^*$ be paramonotone and bimonotone. Then, we have*
($i$) *If $\operatorname{dom} T$ is dense in $X$, then $T$ is single valued*;
($ii$) *If $\operatorname{range} T$ is dense in $X^*$, then $\operatorname{dom} T$ is a singleton*;
($iii$) *$T$ is maximally monotone if and only if $\operatorname{dom} T$ and $\operatorname{range} T$ are closed affine varieties and*

$$\operatorname{dom} T - \operatorname{dom} T = (\operatorname{range} T - \operatorname{range} T)^\perp . \tag{15}$$

**Proof.** ($i$) Assume, reasoning by contradiction, that there exist $(x, x^*)$ and $(x, \widetilde{x}^*)$ in $\operatorname{gph} T$ with $x^* \neq \widetilde{x}^*$ and take $u \in X$ with $\langle u, x^* - \widetilde{x}^*\rangle \neq 0$. Under the current assumption, we can take a sequence $\{x^r\}_{r\in\mathbb{N}} \subset \operatorname{dom} T$ converging to $x + u$. For $r$ large enough we have $\langle x^r - x, x^* - \widetilde{x}^*\rangle \neq 0$ and $x^* \in T(x) = T(x^r)$ because of Corollary 9. This contradicts ($i$) $\Rightarrow$ ($ii$) in Proposition 7.

($ii$) follows analogously to ($i$) by considering $(x, x^*)$ and $(\widetilde{x}, x^*)$ in $\operatorname{gph} T$ with $x \neq \widetilde{x}$, taking again Corollary 9 into account.

($iii$) Assume that $T$ is maximally monotone. Take $x_0 \in \operatorname{dom} T$, $x_0^* \in \operatorname{range} T$, and let $S$ and $S_*$ be the linear subspaces generated by $\operatorname{dom} T - \operatorname{dom} T$ and $\operatorname{range} T - \operatorname{range} T$, respectively. Define $\widehat{T} : X \rightrightarrows X^*$ by

$$\widehat{T}(x) := \begin{cases} x_0^* + S^\perp \text{ if } x \in x_0 + \operatorname{cl} S, \\ \emptyset \qquad \text{otherwise.} \end{cases}$$

We have

$$\operatorname{dom}\widehat{T}-\operatorname{dom}\widehat{T} = (x_0+\operatorname{cl}S)-(x_0+\operatorname{cl}S) = \operatorname{cl}S-\operatorname{cl}S = \operatorname{cl}S = \left(S^{\perp}\right)^{\perp} \tag{16}$$

$$= \left(S^{\perp}-S^{\perp}\right)^{\perp} = \left(\left(x_0^*+S^{\perp}\right)-\left(x_0^*+S^{\perp}\right)\right)^{\perp} \tag{17}$$

$$= \left(\operatorname{range}\widehat{T}-\operatorname{range}\widehat{T}\right)^{\perp},$$

which proves (15) for operator $\widehat{T}$. Moreover,

$$\operatorname{gph}\,\widehat{T} = (x_0+\operatorname{cl}S)\times\left(x_0^*+S^{\perp}\right) = \operatorname{dom}\widehat{T}\times\operatorname{range}\widehat{T}.$$

Therefore, by equivalence $(i) \Leftrightarrow (iii)$ in Corollary 9, the operator $\widehat{T}$ is paramonotone and bimonotone; in particular, $\widehat{T}$ is monotone. On the other hand, by the same equivalence, we have

$$\operatorname{range}T-\operatorname{range}T \subset (\operatorname{dom}T-\operatorname{dom}T)^{\perp} = S^{\perp};$$

hence $S_* \subset S^{\perp}$ and

$$\operatorname{gph}\,T = \operatorname{dom}T\times\operatorname{range}T \subset (x_0+\operatorname{cl}S)\times(x_0^*+\operatorname{cl}S_*) \subset \operatorname{dom}\widehat{T}\times\left(x_0^*+S^{\perp}\right) \tag{18}$$

$$= \operatorname{dom}\widehat{T}\times\operatorname{range}\widehat{T} = \operatorname{gph}\,\widehat{T}.$$

Thus, by the maximal monotonicity of $T$, we have $T=\widehat{T}$, from which we deduce that $\operatorname{dom}T=\operatorname{dom}\widehat{T}=x_0+\operatorname{cl}S$ and $\operatorname{range}T=\operatorname{range}\widehat{T}=x_0^*+S^{\perp}$, thus proving that $\operatorname{dom}T$ and $\operatorname{range}T$ are closed affine varieties.

Let us see the converse implication. Let $(x,x^*)\in X\times X^*$ be such that

$$\langle x-y, x^*-y^*\rangle \geq 0 \text{ for all } (y,y^*)\in\operatorname{gph}\,T = \operatorname{dom}T\times\operatorname{range}T \tag{19}$$

(the latter equality following again from Corollary 9). Since $\operatorname{range}T$ is an affine variety, we can easily prove that $x-y\in(\operatorname{range}T-\operatorname{range}T)^{\perp}=\operatorname{dom}T-\operatorname{dom}T$. More in detail, replace $y^*$ in (19) with $x_0^*\pm\lambda v$ for any given $x_0^*\in\operatorname{range}T$ and any $v\in\operatorname{range}T-\operatorname{range}T$, with $\lambda>0$, then divide both sides of the resulting specification of **(19)** by $\lambda$ and let $\lambda\to+\infty$ to obtain $\langle x-y,\pm v\rangle\geq 0$. Therefore, given that $\operatorname{dom}T$ is an affine variety, we deduce that $x\in\operatorname{dom}T$. Similarly, using that $\operatorname{range}T-\operatorname{range}T=(\operatorname{dom}T-\operatorname{dom}T)^{\perp}$, we obtain that $x^*\in\operatorname{range}T$. Thus, $(x,x^*)\in\operatorname{dom}T\times\operatorname{range}T=\operatorname{gph}\,T$, which proves that $T$ is maximally monotone. ∎

The following propositions involve a finite number of paramonotone operators and are intended to provide a unified framework to deal with the applications of Sections 4 and 5. First, we introduce the following lemma, which has an easy proof.

**Lemma 11** *If $T_1 : X \rightrightarrows X^*$ and $T_2 : X \rightrightarrows X^*$ are paramonotone, then so are $T_1+T_2$ and $T_1\cap(-T_2)$.*

**Proposition 12** *Let $T_i : X \rightrightarrows X^*$, $i = 1, \ldots, m$, be paramonotone operators. Then the intersection mappings*

$$\widetilde{T}_i := T_i \cap \left( -\sum_{j \neq i} T_j \right), \ i = 1, \ldots, m. \tag{20}$$

*are monotone and constant in their common domain*

$$\mathcal{A} := \left\{ x \in X \mid 0 \in \sum_{j=1}^{m} T_j \left( x \right) \right\}.$$

**Proof.** Fix $i \in \{1, \ldots, m\}$. From Lemma 11, $\sum_{j \neq i} T_j$ is paramonotone and, hence, the same lemma establishes that $\widetilde{T}_i$ is paramonotone. Then, from equivalence $(i) \Leftrightarrow (iii)$ in Proposition 7, $\widetilde{T}_i$ is bimonotone. Hence 9 $(ii)$ yields that $\widetilde{T}_i$ is constant in $\mathrm{dom}\widetilde{T}_i$. Finally, one easily sees that $\mathrm{dom}\widetilde{T}_i$ coincides with $\mathcal{A}$. ■

Now, we particularize Proposition 12 by considering finitely many proper lsc convex functions, $f_i : X \to \left]-\infty, +\infty\right]$, $i = 1, \ldots, m$, and the corresponding subdifferential operators $T_i := \partial f_i$, $i = 1, \ldots, m$. We assume the following regularity condition in order to apply the subdifferential sum rule (see Theorem 2 $(vii)$): there exists some index $i_0 \in \{1, \ldots, m\}$, such that

$$\mathrm{dom} f_{i_0} \cap \left( \bigcap_{i \neq i_0} \mathrm{intdom} f_i \right) \neq \emptyset, \tag{21}$$

which is equivalent to the existence of some $\overline{x} \in \cap_{i=1,\ldots,m} \mathrm{dom} f_i$ such that the $m-1$ of the functions $f_i$, $i \in \{1, \ldots, m\} \setminus \{i_0\}$ are continuous at $\overline{x}$ (see Theorem 2 $(i)$).

In this particular case, we are considering the operators

$$\widetilde{T}_i := \partial f_i \cap \left( -\sum_{j \neq i} \partial f_j \right), \ i = 1, \ldots, m, \tag{22}$$

whose the common domain, appealing to statements $(iii)$ and $(vii)$ in Theorem 2, writes as

$$\mathcal{A} = \left\{ x \in X \mid 0 \in \sum_{i=1}^{m} \partial f_i \left( x \right) \right\} = \arg\min \sum_{i=1}^{m} f_i. \tag{23}$$

We summarize the previous comments in the following proposition.

**Proposition 13** *Let $f_i : X \to \left]-\infty, +\infty\right]$, $i = 1, \ldots, m$, be proper lsc convex functions and assume that for some $i_0 \in \{1, \ldots, m\}$ condition (21) holds. Then, operators $\widetilde{T}_i$, $\{1, \ldots, m\}$, defined in (22) are constant on their common domain*

$$\mathcal{A} = \arg\min \sum_{i=1}^{m} f_i.$$

**Remark 14** Proposition 13 can be applied to specific operators in order to derive some classical statements which can be found in the literature, as the one of [8, Lemma 2] involving $\partial f \cap (-N_S)$, and regarding the optimization problem

$$\begin{array}{cc} \textit{minimize} & f(x) \\ \textit{subject to} & x \in C, \end{array}$$

in the case when $f : \mathbb{R}^n \to ]-\infty, +\infty]$ is a proper lsc convex function and $C$ is a closed convex subset of $\mathbb{R}^n$. Specifically, if $S$ denotes the set of optimal solutions of such a problem, [8, Lemma 2] states that $\partial f(x) \cap (-N_C(x))$ is independent of $x \in S$. To derive this statement from Proposition 13, just observe that the normal cone operator, $N_C$ (recall (12)), is paramonotone as it coincides with the subdifferential of the indicator function of $C$.

**Corollary 15** *Under the assumptions of Proposition 13, one has:*
$(i)$ *If for some* $j_0 \in \{1, \ldots, m\}$, *the function* $f_{j_0}$ *is differentiable at* $\overline{x} \in \mathcal{A}$, *then* $\widetilde{T}_{j_0}(x) = \{\nabla f_{j_0}(\overline{x})\}$, *for all* $x \in \mathcal{A}$.
$(ii)$ *If for some* $j_0 \in \{1, \ldots, m\}$, *the function* $f_{j_0}$ *is differentiable on* $\mathcal{A}$, *then* $\nabla f_{i_0}$ *is constant on* $\mathcal{A}$.

**Proof.** $(i)$ follows straightforwardly from Proposition 13, taking into account that if $f_{j_0}$ is differentiable at $\overline{x}$, then $\widetilde{T}_{j_0}(\overline{x}) = \partial f_{j_0}(\overline{x}) = \{\nabla f_{j_0}(\overline{x})\}$ (since $\emptyset \neq \widetilde{T}_{j_0}(\overline{x}) \subset \{\nabla f_{j_0}(\overline{x})\}$), which entails that $\widetilde{T}_{j_0}(x) = \widetilde{T}_{j_0}(\overline{x}) = \{\nabla f_{j_0}(\overline{x})\}$ whenever $x \in \mathcal{A}$.
$(ii)$ comes from $(i)$ since for every $\overline{x}, x \in \mathcal{A}$ we have

$$\{\nabla f_{j_0}(\overline{x})\} = \widetilde{T}_{j_0}(x) \subset \partial f_{j_0}(x) = \{\nabla f_{j_0}(x)\};$$

hence $\nabla f_{j_0}(\overline{x}) = \nabla f_{j_0}(x)$. ■

# 4 Simultaneous projections and displacement mappings

This section is mainly devoted to study the minimal weighted distance to two disjoint non-empty closed and convex subsets $S_1$ and $S_2$ of a Hilbert space $X$. We will denote by $d : X \times X \to \mathbb{R}$ the *distance function* on $X$, i.e., $d(x, y) := \|x - y\|$, and by $d_{S_i} : X \to \mathbb{R}$ the distance function to $S_i$, $i = 1, 2$. We set

$$d(S_1, S_2) := \inf_{s_1 \in S_1, \ s_2 \in S_2} d(s_1, s_2).$$

For arbitrary real numbers $\alpha_1, \alpha_2 > 0$, with $\alpha_1 + \alpha_2 = 1$, and $p \geq 1$, we define

$$\begin{aligned} v(\alpha_1, \alpha_2, p) &:= \inf_{x \in X} \alpha_1 d(x, S_1)^p + \alpha_2 d(x, S_2)^p, \\ \mathcal{A}(\alpha_1, \alpha_2, p) &:= \arg\min \alpha_1 d_{S_1}^p + \alpha_2 d_{S_2}^p. \end{aligned} \tag{24}$$

Observe that $v(\alpha_1, \alpha_2, p)$ and $\mathcal{A}(\alpha_1, \alpha_2, p)$ are the optimal value and the set of optimal solutions of problem (4) for the case of two sets. Notice that $\mathcal{A}(\alpha_1, \alpha_2, p)$ may be empty; consider, e.g., the case when $X := \mathbb{R}^2$, $S_1$ is the convex hull of a branch of a hyperbola and $S_2$ is one of its asymptotes; in this case $v(\alpha_1, \alpha_2, p) = 0$ is not attained.

We denote by $P_1 := P_{S_1}$ and $P_2 := P_{S_2}$ the metric projections over $S_1$ and $S_2$, respectively. We distinguish several cases depending on the values of the power $p$ and parameters $\alpha_1$ and $\alpha_2$. At this moment we advance that in the case when $d_{S_1}^p$ and $d_{S_2}^p$ are differentiable we are able to apply Corollary 15 to derive information about $\mathcal{A}(\alpha_1, \alpha_2, p)$. Going further, Proposition 4($i$) establishes the differentiability of $d_{S_i}^p$ on the whole space $X$ when $p \geq 2$, which allows us to tackle the case of a finite amount of sets.

*Case 1.* $p := 1$, $\alpha_1 \neq \alpha_2$.

Without loss of generality, we assume that $\alpha_1 > \alpha_2$. The following result has a clear geometrical meaning according to Proposition 1.

**Proposition 16** *If* $\alpha_1 > \alpha_2$, *then* $\mathcal{A}(\alpha_1, \alpha_2, 1) = \arg\min_{S_1} d_{S_2}$.

**Proof.** We start by proving that every $x \in X$ satisfies a useful inequality:

$$\begin{aligned} \alpha_1 d\left(P_1(x), S_1\right) + \alpha_2 d\left(P_1(x), S_2\right) &= \alpha_2 d\left(P_1(x), S_2\right) && (25)\\ &\leq \alpha_2\left(\|P_1(x) - x\| + d(x, S_2)\right) && (26)\\ &= \alpha_2\left(d(x, S_1) + d(x, S_2)\right) && (27)\\ &\leq \alpha_1 d(x, S_1) + \alpha_2 d(x, S_2). \end{aligned}$$

Since the latter inequality is strict when $x \notin S_1$, it follows that $\mathcal{A}(\alpha_1, \alpha_2, 1) \subset S_1$. To prove the inclusion $\mathcal{A}(\alpha_1, \alpha_2, 1) \subset \arg\min_{S_1} d_{S_2}$, let $\overline{x} \in \mathcal{A}(\alpha_1, \alpha_2, 1)$ and $x \in S_1$. Since $\overline{x} \in S_1$, we have

$$\alpha_2 d(\overline{x}, S_2) = \alpha_1 d(\overline{x}, S_1) + \alpha_2 d(\overline{x}, S_2) \leq \alpha_1 d(x, S_1) + \alpha_2 d(x, S_2) = \alpha_2 d(x, S_2),$$

which shows that $\overline{x} \in \arg\min_{S_1} d_{S_2}$, thus proving the desired inclusion. For the opposite inclusion, let $\overline{x} \in \arg\min_{S_1} d_{S_2}$ and $x \in X$. Then

$$\begin{aligned} \alpha_1 d(\overline{x}, S_1) + \alpha_2 d(\overline{x}, S_2) &= \alpha_2 d(\overline{x}, S_2) \leq \alpha_2 d\left(P_1(x), S_2\right) && (28)\\ &= \alpha_1 d\left(P_1(x), S_1\right) + \alpha_2 d\left(P_1(x), S_2\right) && (29)\\ &\leq \alpha_1 d(x, S_1) + \alpha_2 d(x, S_2), \end{aligned}$$

which implies that $\overline{x} \in \mathcal{A}(\alpha_1, \alpha_2, 1)$. Therefore $\arg\min_{S_1} d_{S_2} \subset \mathcal{A}(\alpha_1, \alpha_2, 1)$, so the equality in the statement is proved. ■

In the following corollary, $\Pi_1 : S_1 \times S_2 \to S_1$ denotes de projection mapping, defined by $\Pi_1(s_1, s_2) = s_1$.

**Corollary 17** *If* $\alpha_1 > \alpha_2$, *then* $\mathcal{A}(\alpha_1, \alpha_2, 1) = \Pi_1\left(\arg\min_{S_1 \times S_2} d\right)$

**Proof.** Taking into account Proposition 16, we will actually prove the equivalent equality $\arg\min_{S_1} d_{S_2} = \Pi_1 (\arg\min_{S_1\times S_2} d)$. To prove the inclusion $\subset$, let $\overline{s}_1 \in \arg\min_{S_1} d_{S_2}$ and $s_1 \in S_1$. Then, for every $s_2 \in S_2$, we have

$$d(\overline{s}_1, P_2(\overline{s}_1)) = d(\overline{s}_1, S_2) \leq d(s_1, S_2) \leq d(s_1, s_2);$$

hence $(\overline{s}_1, P_2(\overline{s}_1)) \in \arg\min_{S_1\times S_2} d$, implying that $\overline{s}_1 \in \Pi_1 (\arg\min_{S_1\times S_2} d)$, thus proving the desired inclusion. We now proceed to prove the opposite inclusion. Let $\overline{s}_1 \in \Pi_1 (\arg\min_{S_1\times S_2} d)$ and $s_1 \in S_1$. There exists $\overline{s}_2 \in S_2$ such that $(\overline{s}_1, \overline{s}_2) \in \arg\min_{S_1\times S_2} d$, and for every $s_2 \in S_2$ we have

$$d(\overline{s}_1, S_2) \leq d(\overline{s}_1, \overline{s}_2) \leq d(s_1, s_2);$$

taking infimum over $s_2 \in S_2$, this yields $d(\overline{s}_1, S_2) \leq d(s_1, S_2)$, which implies that $\overline{s}_1 \in \arg\min_{S_1} d_{S_2}$. Thus $\Pi_1 (\arg\min_{S_1\times S_2} d) \subset \arg\min_{S_1} d_{S_2}$, and the proof is complete. ■

*Case 2.* $p = 1$, $\alpha_1 = \alpha_2 = \frac{1}{2}$. From now on $]P_1(x), P_2(x)[$ represents the segment of points between $P_1(x)$ and $P_2(x)$, except these two ones; i.e., $]P_1(x), P_2(x)[ := \{(1-\lambda) P_1(x) + \lambda P_2(x) : 0 < \lambda < 1\}$.

**Proposition 18** *One has:*
*(i)* $v(\frac{1}{2}, \frac{1}{2}, 1) = \frac{1}{2} d(S_1, S_2)$,
*(ii)* $\mathcal{A}(\frac{1}{2}, \frac{1}{2}, 1) = \{x \in X : x \in ]P_1(x), P_2(x)[\} \cup \arg\min_{S_1} d_{S_2} \cup \arg\min_{S_2} d_{S_1}$.

**Proof.** $(i)$ For $x \in X$, we have

$$\begin{aligned} d(x, S_1) + d(x, S_2) &= \|x - P_1(x)\| + \|x - P_2(x)\| \geq \|P_1(x) - P_2(x)\| \\ &\geq d(S_1, S_2), \end{aligned} \tag{30}$$

which proves the inequality $\geq$. To prove the opposite inequality, it suffices to observe that, for $s_1 \in S_1$ and $s_2 \in S_2$, we have

$$\begin{aligned} \|s_1 - s_2\| &= d(s_1, S_1) + \|s_1 - s_2\| \geq d(s_1, S_1) + d(s_1, S_2) \\ &\geq 2v(\tfrac{1}{2}, \tfrac{1}{2}, 1). \end{aligned} \tag{31}$$

$(ii)$ Let $x \in \mathcal{A}(\frac{1}{2}, \frac{1}{2}, 1)$. If $x \notin S_1 \cup S_2$, then $x \in ]P_1(x), P_2(x)[$, since otherwise we would have

$$\begin{aligned} d(x, S_1) + d(x, S_2) &= \|x - P_1(x)\| + \|x - P_2(x)\| > \|P_1(x) - P_2(x)\| \\ &\geq d(S_1, S_2), \end{aligned} \tag{32}$$

a contradiction with $(i)$. If $x \in S_1$, then, for any $s_1 \in S_1$ we have

$$d(x, S_2) = d(x, S_1) + d(x, S_2) \leq d(s_1, S_1) + d(s_1, S_2) = d(s_1, S_2),$$

which shows that $x \in \arg\min_{S_1} d_{S_2}$. In the same way, if $x \in S_2$, then $x \in \arg\min_{S_2} d_{S_1}$. We have thus proved the inclusion $\subset$. To prove the opposite inclusion, let $x \in X$ be such that $x \in \left]P_1(x), P_2(x)\right[$ and take $\lambda \in \left]0,1\right[$ such that $x = (1-\lambda)P_1(x) + \lambda P_2(x)$. Combining this equality with the inequalities $\langle s_i - P_i(x), x - P_i(x)\rangle \leq 0$, which hold for every $s_i \in S_i$, we obtain $\langle s_1 - P_1(x), P_2(x) - P_1(x)\rangle \leq 0$ and $\langle s_2 - P_2(x), P_1(x) - P_2(x)\rangle \leq 0$. Adding the latter inequalities, we get $\langle s_2 - s_1 + P_1(x) - P_2(x), P_1(x) - P_2(x)\rangle \leq 0$; hence

$$\|P_1(x) - P_2(x)\|^2 \leq \langle s_1 - s_2, P_1(x) - P_2(x)\rangle \leq \|s_1 - s_2\| \, \|P_1(x) - P_2(x)\|.$$

Therefore, $\|s_1 - s_2\| \geq \|P_1(x) - P_2(x)\|$, and we deduce that

$$d(x, S_1) + d(x, S_2) = \|x - P_1(x)\| + \|x - P_2(x)\| = \|P_1(x) - P_2(x)\| \leq \|s_1 - s_2\|.$$

Since $s_i \in S_i$, $i = 1, 2$, are arbitrarily chosen, we conclude

$$d(x, S_1) + d(x, S_2) \leq d(S_1, S_2),$$

which, by $(i)$, says that $x \in \mathcal{A}(\frac{1}{2}, \frac{1}{2}, 1)$. It remains to prove that

$$\arg\min_{S_1} d_{S_2} \cup \arg\min_{S_2} d_{S_1} \subset \mathcal{A}(\tfrac{1}{2}, \tfrac{1}{2}, 1).$$

For symmetry reasons, it suffices to prove that $\arg\min_{S_1} d_{S_2} \subset \mathcal{A}(\frac{1}{2}, \frac{1}{2}, 1)$, but this inclusion follows from the fact that, for $x \in \arg\min_{S_1} d_{S_2}$, we have

$$\begin{aligned} d(x, S_1) + d(x, S_2) = d(x, S_2) &= \min_{s_1 \in S_1} d(s_1, S_2) = \min_{s_1 \in S_1} \min_{s_2 \in S_2} d(s_1, s_2) \quad (33) \\ &= \min_{s_1 \in S_1,\ s_2 \in S_2} d(s_1, s_2) = d(S_1, S_2). \end{aligned}$$

■

*Case 3.* $p > 1$.

In our current setting, it is known that function $d_{S_i} : X \to \mathbb{R}$ is convex and differentiable outside $S_i$ and for every $x \in X \setminus S_i$ one has (recall Proposition 4$(ii)$)

$$\nabla d_{S_i}(x) = (x - P_i(x)) / \|x - P_i(x)\|. \quad (34)$$

**Theorem 19** *If $p > 1$, then:*

*$(i)$ $\mathcal{A}(\alpha_1, \alpha_2, p) \cap (S_1 \cup S_2) = \emptyset$.*

*$(ii)$ For each $i = 1, 2$, function $d_{S_i}^p$ is differentiable in $\mathcal{A}(\alpha_1, \alpha_2, p)$ and*

$$\nabla d_{S_i}^p(x) = p\,\|x - P_i(x)\|^{p-2}(x - P_i(x)), x \in \mathcal{A}(\alpha_1, \alpha_2, p).$$

*$(iii)$ $\mathcal{A}(\alpha_1, \alpha_2, p)$ coincides with the set of fixed points of*

$$\frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_1 + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_2.$$

**Proof.** $(i)$ It will suffice to prove that $\mathcal{A}(\alpha_1,\alpha_2,p)\cap S_1=\emptyset$. Let $x\in S_1$ and pick $\lambda>0$ such that $\frac{\lambda^p}{1-(1-\lambda)^p}<\frac{\alpha_2}{\alpha_1}$ (this is possible, since $\lim_{\lambda\to 0^+}\frac{\lambda^p}{1-(1-\lambda)^p}=0$). Since

$$d\left((1-\lambda)x+\lambda P_2(x),S_1\right)\leq d\left((1-\lambda)x+\lambda P_2(x),x\right)=\lambda\left\|P_2(x)-x\right\|$$

and

$$d\left((1-\lambda)x+\lambda P_2(x),S_2\right)\leq d\left((1-\lambda)x+\lambda P_2(x),P_2(x)\right)=(1-\lambda)\left\|x-P_2(x)\right\|,$$

we have

$$\begin{aligned}
&\alpha_1 d\left((1-\lambda)x+\lambda P_2(x),S_1\right)^p+\alpha_2 d\left((1-\lambda)x+\lambda P_2(x),S_2\right)^p && (35)\\
&\quad\leq \alpha_1\lambda^p\left\|P_2(x)-x\right\|^p+\alpha_2\left(1-\lambda\right)^p\left\|x-P_2(x)\right\|^p && (36)\\
&\quad=\left(\alpha_1\lambda^p+\alpha_2\left(1-\lambda\right)^p\right)\left\|x-P_2(x)\right\|^p<\alpha_2\left\|x-P_2(x)\right\|^p && (37)\\
&\quad=\alpha_1 d\left(x,S_1\right)^p+\alpha_2 d\left((x,S_2\right)^p,
\end{aligned}$$

which shows that $x\notin\mathcal{A}(\alpha_1,\alpha_2,p)$, thus proving that $\mathcal{A}(\alpha_1,\alpha_2,p)$ and $S_1$ are disjoint.

$(ii)$ is a consequence of $(i)$ taking (34) into account.

$(iii)$ For simplicity of notation, for $x\in\mathcal{A}(\alpha_1,\alpha_2,p)$ and $i=1,2$ we will denote

$$D_i^p(x):=\alpha_i\nabla d_{S_i}^p\left(x\right). \qquad (38)$$

Let $x\in\mathcal{A}(\alpha_1,\alpha_2,p)$. The equality $D_1^p(x)+D_2^p(x)=0$ yields

$$\alpha_1\left\|x-P_1(x)\right\|^{p-2}\left(x-P_1(x)\right)+\alpha_2\left\|x-P_2(x)\right\|^{p-2}\left(x-P_2(x)\right)=0, \qquad (39)$$

from which we deduce that

$$\begin{aligned}
x=&\frac{\alpha_1\left\|x-P_1(x)\right\|^{p-2}}{\alpha_1\left\|x-P_1(x)\right\|^{p-2}+\alpha_2\left\|x-P_2(x)\right\|^{p-2}}P_1(x) && (40)\\
&+\frac{\alpha_2\left\|x-P_2(x)\right\|^{p-2}}{\alpha_1\left\|x-P_1(x)\right\|^{p-2}+\alpha_2\left\|x-P_2(x)\right\|^{p-2}}P_2(x) && (41)\\
=&\frac{\alpha_1}{\alpha_1+\alpha_2\left(\frac{\|x-P_2((x)x)\|}{\|x-P_1(x)\|}\right)^{p-2}}P_1(x)+\frac{\alpha_2}{\alpha_1\left(\frac{\|x-P_1(x)\|}{\|x-P_2(x)\|}\right)^{p-2}+\alpha_2}P_2(x)
\end{aligned}$$

Since condition $D_1^p(x)+D_2^p(x)=0$ implies that $\left\|D_1^p(x)\right\|=\left\|D_2^p(x)\right\|$, that is, $\alpha_1 p\left\|x-P_1(x)\right\|^{p-1}=\alpha_2 p\left\|x-P_2(x)\right\|^{p-1}$, which is equivalent to the equality

$$\frac{\alpha_1}{\alpha_2}=\left(\frac{\|x-P_2(x)\|}{\|x-P_1(x)\|}\right)^{p-1}, \qquad (42)$$

we obtain

$$x = \frac{\alpha_1}{\alpha_1 + \alpha_2 \left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{p-2}{p-1}}} P_1(x) + \frac{\alpha_2}{\alpha_1 \left(\frac{\alpha_2}{\alpha_1}\right)^{\frac{p-2}{p-1}} + \alpha_2} P_2(x) \tag{43}$$

$$= \frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_1(x) + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_2(x).$$

This shows that $x$ is a fixed point of $\frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}}+\alpha_2^{\frac{1}{p-1}}} P_1 + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}}+\alpha_2^{\frac{1}{p-1}}} P_2$.

Conversely, if $x \in X$ is a fixed point of $\frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}}+\alpha_2^{\frac{1}{p-1}}} P_1 + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}}+\alpha_2^{\frac{1}{p-1}}} P_2$, then $x \notin S_1 \cup S_2$. Indeed, otherwise, if, say, $x \in S_1$, then, from the equalities

$$x = \frac{\alpha_1^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_1(x) + \frac{\alpha_2^{\frac{1}{p-1}}}{\alpha_1^{\frac{1}{p-1}} + \alpha_2^{\frac{1}{p-1}}} P_2(x) \tag{44}$$

and $P_1(x) = x$ we would obtain $x = P_2(x) \in S_2$, thus contradicting the assumption that $S_1 \cap S_2 = \emptyset$. Therefore, the functions $d_{S_i}$, $i = 1, 2$, are differentiable at $x$. From (44), it follows that

$$\alpha_1^{\frac{1}{p-1}} (x - P_1(x)) + \alpha_2^{\frac{1}{p-1}} (x - P_2(x)) = 0, \tag{45}$$

from which we deduce (42). Now, using (42), we can rewrite (45) as (39) to obtain the equality $D_1^p(x) + D_2^p(x) = 0$, which shows that $x \in \mathcal{A}(\alpha_1, \alpha_2, p)$. ■

Notice that the set $\mathcal{A}(\frac{1}{2}, \frac{1}{2}, p)$ does not depend on $p$, since, by Theorem 19($iii$), it coincides with the set of fixed points of $\frac{1}{2}(P_1 + P_2)$. Also notice that $\mathcal{A}(\alpha_1, \alpha_2, 2)$ coincides with the set of fixed points of $\alpha_1 P_1 + \alpha_2 P_2$.

The following lemma provides the counterpart of Theorem 19($ii$) for the case $p \geq 2$.

**Lemma 20** *Take $p \geq 2$, and let $\emptyset \neq S \subset X$ be a closed convex set. The function $d_S^p$ is differentiable in $X$ and we have*

$$\nabla d_S^p(x) = p d_S^{p-2}(x)(x - P_S(x)), \text{ for } x \in X.$$

**Proof.** Just write $d_S^p(x)$ as $\left(d_S^2(x)\right)^{p/2}$ and apply Proposition 4 ($i$). ■

The fact that function $d_S^p$ is differentiable in the whole space $X$ enables us to tackle the case of a finite amount of subsets $S_1, ..., S_m$, with $\cap_{i=1}^m S_i = \emptyset$, $m \in \mathbb{N}$. For simplicity, we use the notation

$$\mathcal{A}(\alpha, p) := \arg\min \sum_{i=1}^m \alpha_i d_{S_i}^p, \tag{46}$$

where $\alpha := (\alpha_1, \alpha_2, ..., \alpha_m)$ , with $\alpha_i > 0$, $i = 1, ..., m$, and $\sum_{i=1}^{m} \alpha_i = 1$. The following theorem gathers the announced application of Corollary 15($ii$).

**Theorem 21** *If $p > 1$ and $m = 2$, or $p \geq 2$, the displacement mappings $I - P_i$, $i = 1, ..., m$ are constant on $\mathcal{A}(\alpha, p)$.*

**Proof.** From Theorem 19($ii$) and Lemma 20 if any of the current cases occurs we have that $d_{S_i}^p$ is differentiable on $\mathcal{A}(\alpha, p)$, for each $i = 1, ..., m$. Hence, by Corollary 15($ii$),

$$\nabla d_{S_i}^p (x) = p \|x - P_i(x)\|^{p-2} (x - P_i (x)) \tag{47}$$

is constant on $\mathcal{A}(\alpha, p)$, $i = 1, ..., m$ (again, $P_i := P_{S_i}$, $i = 1, 2, ..., m$). So,

$$\left\| \nabla d_{S_i}^p (x) \right\| = p \|x - P_i(x)\|^{p-1}$$

is constant on $\mathcal{A}(\alpha, p)$, too, and hence so is $\|x - P_i(x)\|$ . Therefore, from (47), we conclude that $I - P_i$ is constant on $\mathcal{A}(\alpha, p)$, $i = 1, ..., m$. ■

As a consequence of the previous theorem, taking Proposition 1 into account, we derive the following corollary. Roughly speaking, under the current assumptions, the corollary says that the smallest translations of the sets $S_i$ that achieve a nonempty intersection are unique.

**Corollary 22** *If $p > 1$ and $m = 2$, or $p \geq 2$, problem (5) has a unique optimal solution, provided that problem (4) is solvable.*

# 5 Distance to feasibility

This section is focused on the distance to feasibility for convex inequality systems in $\mathbb{R}^n$ under RHS perturbations. In this framework, lower and upper estimates for such a distance are provided in terms of some elements whose existence is guaranteed from Corollary 15. Both estimates coincide when confined to linear systems.

Let us consider the parameterized system,

$$\sigma (b) := \{g_i(x) \leq b_i, \ i = 1, \dots, m\}, \tag{48}$$

where $x \in \mathbb{R}^n$, $(b_i)_{i=1,\dots,m} \equiv b \in \mathbb{R}^m$, and $g_i : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $i = 1, 2, ..., m$. To start with, the space of variables, $\mathbb{R}^n$, is endowed with an arbitrary norm, $\|\cdot\|$, with dual norm $\|\cdot\|_*$ and the associated distances denoted by $d$ and $d_*$, respectively. From Corollary 27 on we consider $\mathbb{R}^n$ equipped with the Euclidean norm, $\|\cdot\|_2$. The space of parameters, $\mathbb{R}^m$, is endowed with any $p$-norm, $\|\cdot\|_p$ , provided that $p \geq 2$, and the associated distance is denoted by $d_p$. We denote by $\Theta_c$ the set of consistent parameters; i.e.,

$$\Theta_c := \{b \in \mathbb{R}^m \mid \sigma (b) \text{ is consistent}\}.$$

Throughout this section we consider a fixed $\overline{b} \in \mathbb{R}^m \setminus \Theta_c$ and our aim is to estimate
$$d_p\left(\overline{b}, \Theta_c\right) = \inf\left\{\left\|\overline{b} - b\right\|_p : b \in \text{ is consistent}\right\},$$
called the distance from $\overline{b}$ to feasibility.

**Proposition 23** *Let $\overline{b} \in \mathbb{R}^m \setminus \Theta_c$, then*
$$d_p\left(\overline{b}, \Theta_c\right)^p = \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p.$$

**Proof.** To establish the inequality '$\leq$', take any $x \in \mathbb{R}^n$ and define
$$b_i := \overline{b}_i + [g_i(x) - \overline{b}_i]_+,\ i = 1, ..., m.$$
One can easily check that $b = (b_i)_{i=1,\ldots,m} \in \Theta_c$ and, hence,
$$d_p\left(\overline{b}, \Theta_c\right)^p \leq d_p\left(\overline{b}, b\right)^p = \sum_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p.$$
Since $x \in \mathbb{R}^n$ has been arbitrarily chosen, then
$$d_p\left(\overline{b}, \Theta_c\right)^p \leq \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p.$$

Let us prove the converse inequality. Take any $b \in \Theta_c$, i.e., such that, for some $\overline{x} \in \mathbb{R}^n$, $g_i(\overline{x}) \leq b_i$, $i = 1, ..., m$; then, $g_i(\overline{x}) - \overline{b}_i \leq b_i - \overline{b}_i$, $i = 1, ..., m$, and so
$$[g_i(\overline{x}) - \overline{b}_i]_+ \leq [b_i - \overline{b}_i]_+ \leq \left|b_i - \overline{b}_i\right|,\ i = 1, ..., m.$$
Hence
$$\inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p \leq \sum_{i=1}^{m} [g_i(\overline{x}) - \overline{b}_i]_+^p \leq \left\|\overline{b} - b\right\|_p^p.$$
Since the previous inequality is held for all $b \in \Theta_c$, then $\inf_{x \in \mathbb{R}^n} \sum\limits_{i=1}^{m} [g_i(x) - \overline{b}_i]_+^p \leq d_p\left(\overline{b}, \Theta_c\right)^p$. ■

The well-known Ascoli formula establishes that the distance from a point $x \in \mathbb{R}^n$ to a half-space $H := \{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq b\}$, with $0_n \neq a \in \mathbb{R}^n$ and $b \in \mathbb{R}$, is given by
$$d_H(x) = \frac{[\langle a, x\rangle - b]_+}{\|a\|_*}. \tag{49}$$
The following result is focused on the extension of (49) to the convex case, where a convex inequality of the form '$g(x) \leq b$' is considered. In this context, the distance from $x \in \mathbb{R}^n$ to the nonempty closed convex set $S := \{x \in \mathbb{R}^n \mid g(x) \leq b\}$, denoted by $d_S(x)$, is lower and upper bounded by quotients involving the residual $[g(x) - b]_+$ and the minimum norm of some subgradients of $g$. Regarding these quotients, we use the convention $\frac{0}{0} := 0$.

**Proposition 24** *Let $g:\mathbb{R}^n \to \mathbb{R}$ be a convex function and $b \in \mathbb{R}$ such that the corresponding sublevel set, $S$, is nonempty. Then we have:*

*$(i)$ For any $x \in \mathbb{R}^n$,*

$$d_S(x) \geq \frac{[g(x)-b]_+}{d_*(0_n, \partial g(x))};$$

*$(ii)$ Assume that there exists $\widehat{x} \in \mathbb{R}^n$ (called a Slater point) such that $g(\widehat{x}) < b$. Then, for any $x \in \mathbb{R}^n$,*

$$d_S(x) \leq \frac{[g(x)-b]_+}{d_*(0_n, \partial g(P_S(x)))},$$

*where $P_S(x)$ is the metric projection set of $x$ on $S$ with respect to the norm $\|\cdot\|$.*

**Proof.** $(i)$ Inequality $g(x) \leq b$ turns out to be equivalent (same solution set, $S$) to its standard linearization by means of the Fenchel conjugate, $g^*$, (see, e.g., [5, Formula (3)]), namely system

$$\{\langle u, x\rangle \leq g^*(u) + b,\ u \in \partial g(\mathbb{R}^n)\}.$$

The distance $d_S(x)$ may be computed by means of [10, Lemma 1], yielding (with the convention $\frac{0}{0} := 0$)

$$\begin{aligned} d_S(x) &= \sup\left\{ \frac{[\langle v, x\rangle - \alpha]_+}{\|v\|_*} \,\middle|\, (v,\alpha) \in \operatorname{conv}\{(u, g^*(u)+b),\ u \in \partial g(\mathbb{R}^n)\}\right\} && (50)\\ &\geq \sup\left\{ \frac{[\langle u, x\rangle - (g^*(u)+b)]_+}{\|u\|_*} \,\middle|\, u \in \partial g(\mathbb{R}^n)\right\} && (51)\\ &\geq \sup\left\{ \frac{[g(x)-b]_+}{\|u\|_*} \,\middle|\, u \in \partial g(x)\right\} && (52)\\ &= \frac{[g(x)-b]_+}{\inf\{\|u\|_* \mid u \in \partial g(x)\}} = \frac{[g(x)-b]_+}{d_*(0_n, \partial g(x))}, \end{aligned}$$

where in the third step we have appealed to the fact that

$$g(x) = g^{**}(x) = \langle u, x\rangle - g^*(u) \Leftrightarrow u \in \partial g(x).$$

$(ii)$ It follows from [11, Lemma 2$(ii)$]. Observe that for $x \in S$ we apply the convention $\frac{0}{0} := 0$, whereas for $x \notin S$ the existence of a Slater point entails that $P_S(x)$ is not a minimizer of $g$ (since $g(P_S(x)) = 0$), and then $d_*(0_n, \partial g(P_S(x))) > 0$. ■

**Remark 25** In many cases it is not difficult to see that

$$b \mapsto \delta(b) := d_*\left(0_n, \partial g\left(g^{-1}(b)\right)\right)$$

is a positive nondecreasing function on the interval $]\inf_{\mathbb{R}^n} g, +\infty[$ (we are assuming the nontrivial case when $g$ is not constant, hence not bounded above). Here $\inf_{\mathbb{R}^n} g$ could be $-\infty$ and $\partial g\left(g^{-1}(b)\right) = \bigcup_{g(y)=b} \partial g(y)$. For instance, if $g(x_1, x_2) = e^{x_1} + e^{x_2}$, with the Euclidean norm in $\mathbb{R}^2$, then $\delta(b) = b/\sqrt{2}$ for $b > 0$. Accordingly, item $(ii)$ in the previous lemma entails $d_S(x) \leq [g(x)-b]_+/\delta(b)$.

**Corollary 26** *Let $\overline{b} \in \mathbb{R}^m \setminus \Theta_c$ and assume that $S_i := \left\{x \in \mathbb{R}^n \mid g_i(x) \leq \overline{b}_i\right\} \neq \emptyset$, $i = 1, ..., m$. Then, the following statements hold:*

$(i)$ *Let $\emptyset \neq C \subset \mathbb{R}^n$ be a closed convex set such that, for each $i \in \{1, ..., m\}$, there exists an upper bound $u_i \geq d_* \left(0_n, \partial g_i \left(x\right)\right)$ for all $x \in C$. Then,*

$$d_p \left(\overline{b}, \Theta_c\right)^p \leq \inf_{x \in C} \sum_{i=1}^{m} \left(u_i\right)^p d_{S_i}^p \left(x\right) = \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} \left(u_i\right)^p d_{S_i}^p \left(x\right) + I_C \left(x\right), \qquad (53)$$

*where $I_C$ is the indicator function of $C$; i.e., $I_C \left(x\right) = 0$ if $x \in C$ and $I_C \left(x\right) = +\infty$ if $x \in \mathbb{R}^n \setminus C$.*

$(ii)$ *Assume that for each $i \in \{1, ..., m\}$ there exists a lower bound $0 < l_i \leq d_* \left(0_n, \partial g_i \left(P_{S_i} \left(x\right)\right)\right)$ for all $x \in \mathbb{R}^n \setminus S_i$. Then,*

$$d_p \left(\overline{b}, \Theta_c\right)^p \geq \inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} \left(l_i\right)^p d_{S_i}^p \left(x\right). \qquad (54)$$

**Proof.** $(i)$ comes straightforwardly from Propositions 23 and 24 $(i)$, taking into account the obvious fact that $\inf_{x \in \mathbb{R}^n} \sum_{i=1}^{m} [g_i \left(x\right) - \overline{b}_i]_+^p \leq \inf_{x \in C} \sum_{i=1}^{m} [g_i \left(x\right) - \overline{b}_i]_+^p$.

$(ii)$ follows immediately from Propositions 23 and 24 $(ii)$. ■

Provided that $C$, $u = \left(u_i\right)_{i=1,\ldots,m}$, $l = \left(l_i\right)_{i=1,\ldots,m}$ satisfy the conditions of the previous corollary, we consider the argmin sets coming from (53) and (54):

$$\mathcal{A}\left(C, u\right) := \arg\min \sum_{i=1}^{m} \left(u_i\right)^p d_{S_i}^p \left(x\right) + I_C \left(x\right),$$

$$\mathcal{A}\left(l\right) := \arg\min \sum_{i=1}^{m} \left(l_i\right)^p d_{S_i}^p \left(x\right).$$

Then we can state another corollary of Proposition 24, appealing also to Corollary 15. Indeed, it brings to light the advantages of appealing to $\mathcal{A}\left(C, u\right)$ and $\mathcal{A}\left(l\right)$, instead of working directly with $\arg\min \sum_{i=1}^{m} [g_i \left(x\right) - \overline{b}_i]_+^p$. The key point is that, in the current case in which $p \geq 2$, each function $d_{S_i}^p$ is differentiable in $\mathbb{R}^n$ (see Lemma 20)), which allows to appeal to Corollary 15, while this is not the case of $[g_i \left(\cdot\right) - \overline{b}_i]_+^p$.

Hereafter in this section we consider that $\mathbb{R}^n$ is endowed with the Euclidean norm $\left\|\cdot\right\|_2$ and $P_S \left(x\right)$ will denote the unique projection point of $x$.on a closed convex set $S$.

**Corollary 27** *Keeping the previous notation, assume that $\mathcal{A}\left(C, u\right)$ and $\mathcal{A}\left(l\right)$ are nonempty. Then, we have that:*

$(i)$ *$d_{S_i}$ is constant on both $\mathcal{A}\left(C, u\right)$ and $\mathcal{A}\left(l\right)$, for each $i = 1, ..., m$;*

$(ii)$ *For each $i = 1, ..., m$, let us denote by $d_i^+$ and $d_i^-$ the constant values of $u_i d_{S_i} \left(\cdot\right)$ and $l_i d_{S_i} \left(\cdot\right)$ on $\mathcal{A}\left(C, u\right)$ and $\mathcal{A}\left(l\right)$, respectively, and let $d^+ = \left(d_i^+\right)_{i=1,\ldots,m}$ and $d^- = \left(d_i^-\right)_{i=1,\ldots,m}$. Then,*

$$d_p \left(\overline{b}, \Theta_c\right) \leq \left\|d^+\right\|_p.$$

*If, in addition, for each $i = 1, ..., m$ there exists $\widehat{x}_i \in \mathbb{R}^n$ such that $g_i(\widehat{x}_i) < \overline{b}_i$, then*

$$d_p\left(\overline{b}, \Theta_c\right) \geq \left\|d^-\right\|_p .$$

**Proof.** $(i)$ Regarding $\mathcal{A}(l)$, the statement coincides with the one of Theorem 21 (in the case when $p \geq 2$) just replacing each $\alpha_i$ with $(l_i)^p$. With respect to $\mathcal{A}(C, u)$, the statement comes from an analogous argument to the one of that theorem, just by adding the nondifferentiable mapping $I_C$. For completeness, we include here a sketch of the proof. Observe that all functions $x \mapsto (u_i)^p\, d_{S_i}^p(x)$ are convex and differentiable in $\mathbb{R}^n$, and $x \mapsto I_C(x)$ is a proper lower semicontinuous convex function from $\mathbb{R}^n$ to $]-\infty, +\infty]$. Hence, the regularity condition (21) is satisfied, yielding

$$\mathcal{A}(C, u) = \left\{x \in \mathbb{R}^n \mid 0_n \in \sum_{i=1}^{m} (u_i)^p\, \nabla d_{S_i}^p(x) + \partial I_C(x)\right\} (\neq \emptyset).$$

From Corollary 15, for each $i = 1, ..., m$, we have that $\nabla d_{S_i}^p$ is constant on $\mathcal{A}(C, u)$, hence $d_{S_i}$ is also constant on $\mathcal{A}(C, u)$ since taking norms we have

$$\left\|\nabla d_{S_i}^p(x)\right\| = \left\|p d_{S_i}^{p-2}(x)\,(x - P_i(x))\right\| = p d_{S_i}^{p-1}(x), \text{ for each } x \in \mathcal{A}(C, u),$$

where $P_i(x)$ denotes the projection of $x$ on $S_i$ (recall again Lemma 20).

$(ii)$ follows immediately from $(i)$ and Corollary 26. ■

## 5.1 Linear systems

This subsection is devoted to the linear case, i.e., where $g_i(x) = \langle a_i, x\rangle$, for some $a_i \in \mathbb{R}^n$, $i = 1, ..., m$. In this particular case, obviously $\partial g_i(x) = \{a_i\}$ for all $x \in \mathbb{R}^n$, $i = 1, ..., m$. Let us consider $\overline{b}$ such that

$$\sigma\left(\overline{b}\right) = \left\{\langle a_i, x\rangle \leq \overline{b}_i,\ i = 1, ..., m\right\} \tag{55}$$

is inconsistent and for each $i$ there exists $\widehat{x}_i \in \mathbb{R}^n$ such that $\langle a_i, \widehat{x}_i\rangle < \overline{b}_i$ (observe that it is always held when $a_i \neq 0_n$ or $\overline{b}_i > 0$). According to the notation of Corollary 26, we can choose:

$$C = \mathbb{R}^n,\ l_i = u_i = \|a_i\|_*,\ i = 1, ..., m.$$

Hence $\mathcal{A}(C, u) = \mathcal{A}(l)$, and $d_i^+ = d_i^-$ for all $i$. Let us denote by $\mathcal{A} := \mathcal{A}(C, u)$ and $\overline{d} := \left(d_i^+\right)_{i=1,...,m}$.

The following corollary follows straightforwardly from Corollary 27.

**Corollary 28** *Under the current assumptions, we have*

$$d_p\left(\overline{b}, \Theta_c\right) = \left\|\overline{d}\right\|_p,$$

*where $d_i^+ = \|a_i\|_*\, d_{S_i}(x) = [\langle a_i, x\rangle - \overline{b}_i]_+$, for all $x \in \mathcal{A}$. Moreover $\sigma\left(\overline{b} + \overline{d}\right)$ is a consistent system nearest to $\sigma\left(\overline{b}\right)$.*

The next result is devoted to provide an operative expression for determining $\overline{d}$ with the Euclidean norm in both the space of variables and the space of parameters. For simplicity all norms are denoted by $\|\cdot\|$, $A$ represents the matrix whose rows are $a_i'$, $i = 1, \ldots, m$, $A'$ denotes its transpose and, for any $y \in \mathbb{R}^m$, $[y]_+$ denotes positive part coordinate by coordinate; i.e.,

$$[y]_+ := \left([y_i]_+\right)_{i=1,\ldots,m}.$$

**Theorem 29** *The following conditions are equivalent:*
*(i)* $\left(x^0, h^0\right) \in \mathcal{A} \times \left\{\overline{d}\right\}$;
*(ii)* $\left(x^0, h^0\right)$ *is a solution of the system, in the variable* $(x, h)$,

$$\begin{cases} \left[Ax - \overline{b}\right]_+ = h, \\ \quad A'h = 0_n. \end{cases} \tag{56}$$

*(iii)* $\left(x^0, h^0\right)$ *is an optimal solution of the quadratic problem, in the variable* $(x, h)$,

$$\begin{array}{ll} \min & \langle h, h \rangle \\ s.t. & Ax \leq \overline{b} + h, \\ & h \geq 0_m. \end{array} \tag{57}$$

**Proof.** Let us see $(i) \Rightarrow (ii)$. Let $\left(x^0, h^0\right) \in \mathcal{A} \times \left\{\overline{d}\right\}$, i.e., $x^0 \in \mathcal{A}$ and $h^0 = \overline{d}$. By Corollary 28, $h_i^0 (= d_i^+) = \left[a_i' x^0 - \overline{b}_i\right]_+$, for all $i$. Moreover, the optimality condition

$$x^0 \in \mathcal{A} := \arg\min \sum_{i=1}^m \|a_i\|^2 d_{S_i}^2(x)$$

is equivalent to

$$0_n = \sum_{i=1}^m \|a_i\|^2 \nabla d_{S_i}^2\left(x^0\right) = 2\sum_{i=1}^m \|a_i\| \, d_{S_i}\left(x^0\right) a_i = 2\sum_{i=1}^m \left[a_i' x^0 - \overline{b}_i\right]_+ a_i; \tag{58}$$

in other words

$$0_n = \sum_{i=1}^m h_i^0 a_i = A' h^0.$$

So, $\left(x^0, h^0\right)$ is a solution of system (56).
$(ii) \Rightarrow (i)$ Let $\left(x^0, h^0\right)$ be a solution of (56); i.e., $h^0 = \left[Ax^0 - \overline{b}\right]_+$ and

$$0_n = \sum_{i=1}^m h_i^0 a_i = A' h^0.$$

Then, by repeating the previous argument of (58), we have

$$0_n = \sum_{i=1}^m \|a_i\|^2 \nabla d_{S_i}^2\left(x^0\right),$$

which means that $x^0 \in \mathcal{A}$. Then, appealing again to Corollary 28, we deduce $h^0 = \overline{d}$.

Now, let us prove $(ii) \Leftrightarrow (iii)$. By the Karush-Kuhn-Tucker (KKT, in brief) conditions, $\left(x^0, h^0\right)$ is an optimal solution of (57) if and only if there exist $\lambda, \mu \in \mathbb{R}^m_+$ such that

$$\left\{ \begin{array}{c} -\begin{pmatrix} 0_n \\ 2h^0 \end{pmatrix} = \begin{pmatrix} A' \\ -I_m \end{pmatrix}\lambda + \begin{pmatrix} 0_{n\times m} \\ -I_m \end{pmatrix}\mu, \\ \left(Ax^0 - \overline{b} - h^0\right)' \lambda = 0, \ -\left(h^0\right)' \mu = 0, \\ Ax^0 - \overline{b} - h^0 \leq 0_m, \ h^0 \geq 0_m. \end{array} \right. \tag{59}$$

So, $A'\lambda = 0_n$, and $h^0 = \frac{\lambda+\mu}{2}$. Moreover, $h_i^0 \mu_i = 0$ for all $i$. Let us see that $\mu = 0_m$. If $h_i^0 = 0$, then $\lambda_i + \mu_i = 0$, which entails $\lambda_i = \mu_i = 0$, while, if $h_i^0 > 0$, then $\mu_i = 0$. Therefore

$$h^0 = \frac{\lambda}{2} \tag{60}$$

and, so,

$$A'h^0 = 0_n.$$

Let us see that $\left[Ax^0 - \overline{b}\right]_+ = h^0$. Observe that $\left(a_i'x^0 - \overline{b}_i - h_i^0\right)\lambda_i = 0$ for all $i$. If $a_i'x^0 - \overline{b}_i < 0$, then $a_i'x^0 - \overline{b}_i - h_i^0 < 0$, thus we have $\lambda_i = 0$ and

$$h_i^0 = \frac{\lambda_i}{2} = 0.$$

If $a_i'x^0 - \overline{b}_i > 0$, then $h_i^0 > 0$ and $\lambda_i > 0$, yielding $a_i'x^0 - \overline{b}_i - h_i^0 = 0$. Finally, if $a_i'x^0 - \overline{b}_i = 0$, then $h_i^0\lambda_i = 0$, and from (60) we have $h_i^0 = 0$. So,

$$\left[a_i'x^0 - \overline{b}_i\right]_+ = h_i^0, \text{ for all } i,$$

and consequently $\left(x^0, h^0\right)$ is a solution of (56).

Reciprocally, if $\left(x^0, h^0\right)$ is a solution of (56) and we consider

$$\lambda = 2h^0 \text{ and } \mu = 0_m,$$

it can be easily seen that $x^0, h^0, \lambda$ and $\mu$ satisfy the KKT conditions (59), and then $\left(x^0, h^0\right)$ is an optimal solution for problem (57). ■

# References


[1] H.H. BAUSCHKE, P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[2] H.H. BAUSCHKE, P. L. COMBETTES, D.R. LUKE, *Finding best approximation pairs relative to closed convex sets in Hilbert spaces*, Journal of Approximation Theory 125 (2004), 178-192.

[3] H.H. BAUSCHE, X. WANG, YAO, *On Borwein–Wiersma decompositions of monotone linear relations*, SIAM J. Optim. 20 (2010), 2636–2652.

[4] H.H. BAUSCHKE, X. WANG, L. YAO, *Rectangularity and paramonotonicity of maximally monotone operators*, Optimization, 63 (2014), 487-504.

[5] G. BEER, M.J. CÁNOVAS, M.A. LÓPEZ, J. PARRA, *Lipschitz modulus of linear and convex inequality systems with the Hausdorff metric*, Math. Program. 189 (2021), pp. 75–98.

[6] J.M. BORWEIN, J.D. VANDERWERFF, *Convex Functions*, Cambridge University Press, Cambridge, MA, 2010.

[7] R. S. BURACHIK, A. N. IUSEM, *Set-valued mappings and enlargement of monotone operators,* Springer, New York, 2009.

[8] J.V. BURKE, M.C. FERRIS, *Characterization of solution sets of convex programs*, Oper. Res. Lett. 10 (1991), 57-60.

[9] R.E. BRUCK, *An iterative solution of a variational inequality for certain monotone operators in Hilbert space*, Bull. AMS 81 (1975), 890–892.

[10] M.J. CÁNOVAS, F.J. GÓMEZ-SENET, J. PARRA, *Regularity modulus of arbitrarily perturbed linear inequality systems*, J. Math. Anal. Appl. 343 (2008), pp. 315-327.

[11] M.J. CÁNOVAS, F.J. GÓMEZ-SENET, J. PARRA, *Linear regularity, equirregularity, and intersection mappings for convex semi-infinite inequality systems*, Math. Program. Ser B, 123 (2010), 33-60.

[12] Y. CENSOR, A.N. IUSEM, S.A. ZENIOS, *An interior point method with Bregman functions for the variational inequality problem with paramonotone operators*, Math. Program. (Series A) 81 (1998), 373–400.

[13] R.B. HOLMES, *Smoothness of certain metric projections on Hilbert space*, Trans. Amer. Math. Soc. 184 (1973), 87-100.

[14] A.N. IUSEM, *On some properties of paramonotone operators*, J. Convex Anal. 5 (1998), 269-278.

[15] R. LUCCHETTI, *Convexity and Well-Posed Problems*, volume 22 of CMS Books in Mathematics. Springer-Verlag, New York, 2006.

[16] J. E. MARTÍNEZ-LEGAZ, B. F. SVAITER, *Monotone operators representable by l.s.c. convex functions*, Set-Valued Anal. 13 (2005), 21–46

[17] B.S. MORDUKHOVICH, N.M. NAM, *Convex Analysis and Beyond. Volume I. Basic theory*, Springer Series in Operations Research and Financial Engineering. Springer-Verlag, Cham, 2022.

[18] R.R. PHELPS, *Convex Functions*, Monotone Operators and Differentiability, 2nd ed., Springer-Verlag, Berlin, 1993.

[19] R.T. ROCKAFELLAR, *Extension of Fenchel' duality theorem for convex functions,* Duke Math. J. 33 (1966), 81-89.

[20] R.T. ROCKAFELLAR, R.J.-B. WETS, *Variational Analysis,* 3rd ed., Springer-Verlag, Berlin, 2009.

[21] S. SIMONS, *Positive sets and monotone sets*, J. Convex Anal. 14 (2007) 297–317.

[22] S. SIMONS, *From Hahn-Banach to Monotonicity*, Springer-Verlag, Berlin, 2008.

[23] C. ZĂLINESCU, *Convex Analysis in General Vector Spaces*, World Scientific Publishing, Singapore, 2002.

# E Hoffman constant of the argmin mapping in linear optimization.

# Hoffman constant of the argmin mapping in linear optimization*

J. Camacho[†] M.J. Cánovas[†] H. Gfrerer[‡] J. Parra[†]

**Abstract**

The main contribution of this paper consists of providing an explicit formula to compute the Hoffman constant of the argmin mapping in linear optimization. The work is developed in the context of right-hand side perturbations of the constraint system as the Hoffman constant is always infinite when we perturb the objective function coefficients, unless the left-hand side of the constraints reduces to zero. In our perturbation setting, the argmin mapping is a polyhedral mapping whose graph is the union of convex polyhedral sets which assemble in a so nice way that global measures of the stability (Hoffman constants) can be computed through semilocal and local ones (as Lipschitz upper semicontinuity and calmness moduli, whose computation has been developed in previous works). Indeed, we isolate this nice behavior of the graph in the concept of well-connected polyhedral mappings and, in a first step, the paper focuses on Hoffman constant for these multifunctions. When confined to the optimal set, some specifics on directional stability are also presented.

**Key words.** Hoffman constants, calmness constants, Lipschitz upper semicontinuity, linear inequality systems, optimal set mapping.

**Mathematics Subject Classification:** 90C31, 49J53, 49K40, 90C05

*This research has been partially supported by Grant PGC2018-097960-B-C21 from MICINN, Spain, and ERDF, ”A way to make Europe”, European Union, and Grant PROMETEO/2021/063 from Generalitat Valenciana, Spain.

[†]Center of Operations Research, Miguel Hernández University of Elche, 03202 Elche (Alicante), Spain (j.camacho@umh.es, canovas@umh.es, parra@umh.es).

[‡]Institute of Computational Mathematics, Johannes Kepler University Linz, A-4040, Linz, Austria (helmut.gfrerer@jku.at).

## 1 Introduction and overview

The present paper is mainly focused on the global measure of the stability called *Hoffman constant* for the optimal set in linear optimization. The terminology 'Hoffman constant' (or Hoffman bound) comes from its counterpart to feasible sets in the context of linear systems (see, e.g., [1, **?**, 12, 18, 20]) and, going further, from the pioneer work of Hoffman [9] establishing the existence of a positive constant $\kappa$ such that the distance of any point in the Euclidean space to the feasible set of any consistent system is bounded above by $\kappa$ time the absolute residual. Hoffman constants play a fundamental role in mathematical programming; in particular, regarding convergence properties of optimization algorithms and sensitivity analysis, among others; see again [18] and references therein (e.g., [14, 16, 17, 21]).

We consider the linear optimization problem

$$\begin{array}{ll} \text{minimize} & c'x \\ \text{subject to} & a_t'x \leq b_t, \;\; t \in T := \{1, 2, ..., m\}, \end{array} \tag{1}$$

where $c, a_t \in \mathbb{R}^n$, $t \in T$, the prime stands for transposition, $x \in \mathbb{R}^n$ is the decision variable, regarded as a column-vector, and $b = (b_t)_{t \in T} \in \mathbb{R}^m$. Throughout the paper the $a_t$'s are fixed. Moreover, except in Section 4.1, vector $c$ also remains (indeed, we write $c = \overline{c}$ to emphasize this fact). So, we mainly deal with the *right-hand side* (*RHS* for short) *perturbation* setting, where $b$ is the parameter to be perturbed around a nominal one, $\overline{b} \in \mathbb{R}^m$. Regarding Section 4.1, it is focused in the framework of the so-called *canonical perturbations*, where both $c$ and $b$ are treated as parameters. Let us mention that direct antecedents to the current work, as [3] and [4], are developed in such a setting. While these previous works deal with local and semilocal stability measures (*calmness* and *Lipschitz upper semicontinuity moduli*), the current one focuses on a global measure and this fact entails notable differences. Here the term 'semilocal' means that we are considering the whole optimal sets associated with parameters in a neighborhood of the nominal one. Indeed, as we will see in Section 4.1, the Hoffman constant for the optimal set mapping of problems (1) under canonical perturbations is always $+\infty$ unless $\{a_t, \; t \in T\}$ reduces to zero, and this fact justifies the choice of our RHS perturbation framework.

For the reader's convenience, the present paper follows the notation of [3], which constitutes an immediate antecedent to the current work: $\mathcal{F} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and $\mathcal{F}^{op} : \mathbb{R}^n \times \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ denote the *feasible* and the *optimal set* *(argmin)* mappings, given by

$$\mathcal{F}(b) := \left\{x \in \mathbb{R}^n : a_t' x \leq b_t \text{ for all } t \in T\right\}, \text{ and} \tag{2}$$

$$\mathcal{F}^{op}(c, b) := \arg\min\left\{c'x : x \in \mathcal{F}(b)\right\}. \tag{3}$$

The optimal set mapping under RHS perturbations (with $c = \overline{c}$ fixed) is defined as

$$\mathcal{F}_{\overline{c}}^{op}(b) := \mathcal{F}^{op}(\overline{c}, b). \tag{4}$$

The reader is addressed to the monograph [8] for a comprehensive development of the stability theory in linear optimization from a qualitative point of view, formalized through continuity properties of the feasible and argmin mappings in different parametric settings, and with an arbitrary (possibly infinite) index set $T$.

At this point, let us specify the main contributions of the paper. Theorem 5 constitutes one of the main final results as it provides a point-based formula (only involving the nominal problem's data) for the Hoffman constant of $\mathcal{F}_{\overline{c}}^{op}$, denoted by $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$. Roughly speaking, $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$ is such a constant that the distance of any point $x \in \mathcal{F}_{\overline{c}}^{op}(b)$ to the optimal set $\mathcal{F}_{\overline{c}}^{op}\left(\overline{b}\right)$ is bounded above by $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$ times the distance between parameters $b$ and $\overline{b}$, provided that $b, \overline{b} \in \operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$ (the domain of $\mathcal{F}_{\overline{c}}^{op}$). Accordingly, $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$ can seen as a Lipschitz constant of a global nature. A crucial intermediate step to this final result consists of expressing $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$ as the maximum of certain semilocal and local measures, namely the *Lipschitz upper semicontinuity* and *calmness* moduli of $\mathcal{F}_{\overline{c}}^{op}$, denoted by $\operatorname{Lipusc}\mathcal{F}_{\overline{c}}^{op}(b)$ and $\operatorname{clm}\mathcal{F}_{\overline{c}}^{op}(b, x)$, respectively, at all $b \in \operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$ and all $x \in \mathcal{F}_{\overline{c}}^{op}(b)$, which have been previously analyzed in [3] and [4], respectively; for completeness, Section 2 recalls the necessary results about $\operatorname{Lipusc}\mathcal{F}_{\overline{c}}^{op}(b)$ which, in turn, are expressed in terms of certain calmness moduli. See, e.g., the monographs [6], [11], [15] and [19] for a wider perspective on Lipschitz-type properties and their moduli.

Section 3 isolates the key properties, held by $\mathcal{F}_{\overline{c}}^{op}$, which are behind the representation of $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$ in terms of $\operatorname{Lipusc}\mathcal{F}_{\overline{c}}^{op}(b)$, $b \in \operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$. These properties are formalized in Definition 2 giving rise to the concept of *well-connected polyhedral mappings*, which are multifunctions of the form $\mathcal{S} := \bigcup_{i\in I}\mathcal{S}_i$, where $I$ is a finite index set such that the graphs of mappings $\mathcal{S}_i$ are convex polyhedral sets assembled in an appropriate way. In particular, a key technical result (Lemma 3) ensures, for each pair, $b, \overline{b} \in \operatorname{dom}\mathcal{S}$, the existence of a subdivision $0 := \mu_0 < \mu_1 ... < \mu_N = 1$ and indices $\{i_1, ..., i_N\} \subset I$ connecting $b$ with $\overline{b}$ in the sense of Definition 3, yielding

$$\mathcal{S}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right) = \mathcal{S}_{i_k}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right), \text{ whenever } \mu \in [\mu_{k-1}, \mu_k].$$

Starting from this technical result, we derive the variational inequality of Theorem 4, which entails the crucial equality of Corollary 1:

$$\begin{aligned}\operatorname{Hof}\mathcal{S} &= \sup\{\operatorname{Lipusc}\mathcal{S}(b) \mid b\in\operatorname{dom}\mathcal{S}\} \\ &= \sup\{\operatorname{clm}\mathcal{S}(b,x) \mid (b,x)\in\operatorname{gph}\mathcal{S}\},\end{aligned} \tag{5}$$

provided that $\mathcal{S}$ is a well-connected polyhedral mapping.

Finally, let us comment that Section 5 is concerned with some characteristics of the optimal set along segments determined by two parameters $b$, $\overline{b}\in\operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$, formally, with some features of $\mathcal{F}_{\overline{c}}^{op}\left(\overline{b}+\mu\left(b-\overline{b}\right)\right)$, provided that $\overline{b},b\in\operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$ and $\mu\in[0,1]$. That section complements some results of Section 3. In particular, Section 5 introduces the concept of *break steps*, which provides a constructive procedure to determine a particular subdivision together with a family of indices connecting $b$ with $\overline{b}$. We advance that this procedure is based on the concept of *minimal KKT* (Karush-Kuhn-Tucker) *sets of indices* introduced in [4] in relation to the calmness modulus of the argmin mapping.

The paper is structured as follows: Section 2 provides some notation and preliminary results used throughout the work. Section 3 gathers the results about a well-connected polyhedral mapping $\mathcal{S}=\bigcup_{i\in I}\mathcal{S}_i$, among which we underline equality (5), which gives rise to its specification to $\mathcal{S}=\mathcal{F}_{\overline{c}}^{op}$ in Theorem 5 of Section 4. Section 5 introduces the concept of break steps and provides the announced constructive procedure to connect two elements $b$, $\overline{b}\in\operatorname{dom}\mathcal{F}_{\overline{c}}^{op}$ yielding the variational inequality of Theorem 6, which underlies the computation of $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$.

## 2 Preliminaries

To begin with, we introduce some definitions and fix the notation used hereafter. Given $X\subset\mathbb{R}^p$, $p\in\mathbb{N}$, we denote by $\operatorname{int}X$, $\operatorname{cl}X$, $\operatorname{conv}X$, $\operatorname{cone}X$, and $\operatorname{span}X$ the interior, the closure, the convex hull, the conical convex hull, and the linear hull of $X$ respectively, with the convention $\operatorname{conv}\emptyset=\emptyset$ and $\operatorname{cone}\emptyset=\operatorname{span}\emptyset=\{0_p\}$ (the zero vector of $\mathbb{R}^p$). Provided that $X$ is convex, $\operatorname{extr}X$ stands for the set of extreme points of $X$.

Now we recall the Lipschitz type properties appealed to in the paper for a multifunction $\mathcal{M}:Y\rightrightarrows X$ between metric spaces, with both distances being denoted by $d$. The *Hoffman property* holds if there exists a constant $\kappa\geq 0$ such that

$$d(x,\mathcal{M}(\widetilde{y}))\leq\kappa d\left(y,\widetilde{y}\right) \text{ for all } y,\widetilde{y}\in\operatorname{dom}\mathcal{M} \text{ and all } x\in\mathcal{M}\left(y\right); \tag{6}$$

equivalently,

$$d(x, \mathcal{M}(\widetilde{y})) \leq \kappa d\left(\widetilde{y}, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in X \text{ and all } \widetilde{y} \in \operatorname{dom} \mathcal{M},$$

where, for $x \in X$ and $\Omega \subset X$, $d(x, \Omega) := \inf \{d(x, \omega) \mid \omega \in \Omega\}$, with the convention $\inf \emptyset := +\infty$ (so that $d(x, \emptyset) = +\infty$), $\operatorname{dom} \mathcal{M}$ is the domain of $\mathcal{M}$ (recall that $y \in \operatorname{dom} \mathcal{M} \Leftrightarrow \mathcal{M}(y) \neq \emptyset$) and $\mathcal{M}^{-1}$ denotes the inverse mapping of $\mathcal{M}$ (i.e. $y \in \mathcal{M}^{-1}(x) \Leftrightarrow x \in \mathcal{M}(y)$).

Regarding semilocal measures, this paper focuses on the *Lipschitz upper semicontinuity* of $\mathcal{M}$ at $\overline{y} \in \operatorname{dom} \mathcal{M}$, which is defined as the existence of a neighborhood $V$ of $\overline{y}$ along with a constant $\kappa \geq 0$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d(y, \overline{y}) \text{ for all } y \in V \text{ and all } x \in \mathcal{M}(y). \tag{7}$$

We also appeal to the *calmness* property, which is a local measure as it considers solutions near a given (nominal) solution $\overline{x}$ and parameters in a neighborhood of the nominal one $\overline{y}$. Specifically, $\mathcal{M}$ is said to be calm at $(\overline{y}, \overline{x}) \in \operatorname{gph}\mathcal{M}$ (the graph of $\mathcal{M}$) if there exist a constant $\kappa \geq 0$ and a neighborhood of $(\overline{y}, \overline{x})$, $V \times U$, such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d(y, \overline{y}) \text{ for all } x \in \mathcal{M}(y) \cap U \text{ and all } y \in V, \tag{8}$$

which is known to be equivalent to the *metric subregularity* (cf. [6, Theorem 3H.3 and Exercise 3H.4]) of $\mathcal{M}^{-1}$ at $(\overline{x}, \overline{y})$ which reads as the existence of $\kappa \geq 0$ and a (possibly smaller) neighborhood $U$ of $\overline{x}$ such that

$$d(x, \mathcal{M}(\overline{y})) \leq \kappa d\left(\overline{y}, \mathcal{M}^{-1}(x)\right) \text{ for all } x \in U\ . \tag{9}$$

The infimum of constants $\kappa$, for some associated neighborhoods, appearing in (6), (7) and (8)-(9) are the *Hoffman constant* of $\mathcal{M}$, the *Lipschitz upper semicontinuity modulus* of $\mathcal{M}$ at $\overline{y}$ and the *calmness modulus* of $\mathcal{M}$ at $(\overline{y}, \overline{x})$, denoted by $\operatorname{Hof} \mathcal{M}$, $\operatorname{Lipusc} \mathcal{M}(\overline{y})$ and $\operatorname{clm} \mathcal{M}(\overline{y}, \overline{x})$, respectively. These three constants may be written as follows (the first and the third come directly from the definitions, while the second is established in [2, Proposition 2]):

$$\begin{aligned}
&\operatorname{Hof} \mathcal{M} = \sup_{(y,x) \in (\operatorname{dom} \mathcal{M}) \times X} \frac{d(x, \mathcal{M}(y))}{d\left(y, \mathcal{M}^{-1}(x)\right)}, \\
&\operatorname{Lipusc} \mathcal{M}(\overline{y}) = \limsup_{y \to \overline{y}} \left( \sup_{x \in \mathcal{M}(y)} \frac{d(x, \mathcal{M}(\overline{y}))}{d(y, \overline{y})} \right), \ \overline{y} \in \operatorname{dom} \mathcal{M}, \\
&\operatorname{clm} \mathcal{M}(\overline{y}, \overline{x}) = \limsup_{x \to \overline{x}} \frac{d(x, \mathcal{M}(\overline{y}))}{d\left(\overline{y}, \mathcal{M}^{-1}(x)\right)}, \ (\overline{y}, \overline{x}) \in \operatorname{gph} \mathcal{M}.
\end{aligned} \tag{10}$$

Here we use the convention $\frac{0}{0} := 0$ and $\limsup_{z \to \overline{z}}$ is understood as the supremum (maximum, indeed) of all possible sequential upper limits for all possible sequences $\{z_r\}_{r \in \mathbb{N}}$ converging to $\overline{z}$ as $r \to \infty$. It is clear that

$$\sup_{(y,x) \in \operatorname{gph} \mathcal{M}} \operatorname{clm} \mathcal{M}(y,x) \leq \sup_{y \in \operatorname{dom} \mathcal{M}} \operatorname{Lipusc} \mathcal{M}(y) \leq \operatorname{Hof} \mathcal{M}. \tag{11}$$

The following theorem, which comes from [2, Theorem 4], provides a sufficient condition to get equalities in (11); Corollary 1 and Theorem 5 constitute the counterpart of this result, in Euclidean spaces, for well-connected polyhedral mappings and for the particular case of $\mathcal{F}_{\overline{c}}^{op}$ , respectively.

**Theorem 1** *Let $\mathcal{M} : Y \rightrightarrows X$, with $Y$ being a normed space and $X$ being a reflexive Banach space. Assume that* $\operatorname{gph} \mathcal{M}$ *is a nonempty convex set and that $\mathcal{M}$ has closed images. Then*

$$\operatorname{Hof} \mathcal{M} = \sup_{y \in \operatorname{dom} \mathcal{M}} \operatorname{Lipusc} \mathcal{M}(y) = \sup_{(y,x) \in \operatorname{gph} \mathcal{M}} \operatorname{clm} \mathcal{M}(y,x). \tag{12}$$

Clearly, $\operatorname{gph} \mathcal{F}$ is a convex set and $\mathcal{F}$ has closed images (indeed, $\operatorname{gph} \mathcal{F}$ is a convex polyhedral set) and, hence, the previous theorem applies for $\mathcal{M} = \mathcal{F}$. Regarding the argmin mapping, $\operatorname{gph} \mathcal{F}^{op}$ is no longer convex; indeed, $\operatorname{gph} \mathcal{F}_{\overline{c}}^{op}$ is also nonconvex in general. This underlies the fact that the analysis of $\operatorname{Hof} \mathcal{F}^{op}$ and $\operatorname{Hof} \mathcal{F}_{\overline{c}}^{op}$ do not rely on Theorem 1; indeed, as announced in Section 1, this analysis constitutes the main goal of the current paper and it is developed in Section 4 appealing to the results of Section 3 about well-connected polyhedral mapping.

The rest of this section is devoted to provide some background on the calmness and Lipschitz upper semicontinuity moduli for multifunction $\mathcal{F}^{op}$. Specifically, Theorems 2 and 3 provide point-based formulae (only depending on the nominal parameter and point) for these two constants. Although $\operatorname{gph} \mathcal{F}^{op}$ is not convex, $\mathcal{F}^{op}$ still satisfies a certain local directional convexity property, which turns out to be crucial for obtaining Theorem 3 (see [3, Theorem 5]).

Let us introduce some notation and fix the topology of the involved spaces. The space of variables, $\mathbb{R}^n$, is endowed with an arbitrary norm $\|\cdot\|$, whose dual norm $\|\cdot\|_*$ is given by $\|u\|_* = \max_{\|x\| \leq 1} |u'x|$. Each element $b$ in the RHS parameter space, $\mathbb{R}^m$, works with the norm $\|b\|_\infty := \max_{t \in T} |b_t|$. The full parameter space, $\mathbb{R}^n \times \mathbb{R}^m$, is endowed with the norm $\|(c,b)\| := \max\{\|c\|_*, \|b\|_\infty\}$. From now on we appeal to the set of active indices at $x \in \mathcal{F}(b)$, defined as

$$T_b(x) := \left\{t \in T : a_t' x = b_t\right\}.$$

**Definition 1** *Given* $(c,b) \in \operatorname{dom} \mathcal{F}^{op}$ *and* $x \in \mathcal{F}^{op}(c,b)$, *the family of* minimal KKT index subsets *at* $((c,b),x)$ *denoted by* $\mathcal{M}_{c,b}(x)$ *–and introduced in [4]–, is defined as the collection of all* $D \subset T_b(x)$ *such that* $D$ *is minimal (with respect to the inclusion order) among those satisfying* $-c \in \operatorname{cone}\{a_t,\, t \in D\}$. *It can be checked that* $\mathcal{M}_{c,b}(x)$ *does not depend on* $x \in \mathcal{F}^{op}(c,b)$ *(cf. [7, Remark 2]), so that it is referred to as the family of* minimal KKT index subsets *at* $(c,b)$, *denoted as* $\mathcal{M}_{c,b}$.

**Lemma 1** [3, Lemma3.2] *Let* $(\overline{c},\overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$. *Then there exists* $\varepsilon > 0$ *such that for every* $b \in \operatorname{dom} \mathcal{F}$ *with* $\|b - \overline{b}\|_\infty \leq \varepsilon$ *we have* $\mathcal{M}_{\overline{c},b} \subset \mathcal{M}_{\overline{c},\overline{b}}$.

For any $D \in \mathcal{M}_{c,b}$ we consider the mapping $\mathcal{L}_D : \mathbb{R}^m \times \mathbb{R}^D \rightrightarrows \mathbb{R}^n$ given by

$$\mathcal{L}_D(b,d) := \left\{x \in \mathbb{R}^n : a_t' x \leq b_t,\, t \in T;\ -a_t' x \leq d_t,\, t \in D\right\}. \qquad (13)$$

**Proposition 1** [4, Proposition 4.1] *Let* $(\overline{c},\overline{b}) \in \operatorname{dom} \mathcal{F}^{op}$. *Then*

$$\mathcal{L}_D\left(\overline{b}, -\overline{b}_D\right) = \mathcal{F}^{op}(\overline{c},\overline{b}) \textit{ for all } D \in \mathcal{M}_{\overline{c},\overline{b}}.$$

The next result provides three different expressions for the calmness modulus of the optimal set mapping $\mathcal{F}^{op}$ at $\left((\overline{c},\overline{b}),\overline{x}\right) \in \operatorname{gph} \mathcal{F}^{op}$. The first two ones come directly from [4, Corollary 4.1]; in the second expression each $\operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), \overline{x}\right)$ can be computed through the concept of *end set* of a convex set $C \subset \mathbb{R}^n$, traced out from [10] and defined as

$$\operatorname{end} C := \{u \in \operatorname{cl} C \mid \nexists \mu > 1 \text{ such that } \mu u \in \operatorname{cl} C\}.$$

The third expression can be seen as a geometrical interpretation of the formula given in [5, Theorem 4] for the calmness modulus of a feasible set mapping. Observe that $\mathcal{L}_D$ is nothing else but the feasible set mapping associated with an extension of the constraint system of (1) in order to force that inequalities indexed by $D$ are held as equalities at the nominal parameter. Here $\overline{b}_D$ means $\left(\overline{b}_t\right)_{t \in D}$.

**Theorem 2** [4, Corollary 4.1], [5, Theorem 4] *Let* $\left((\overline{c},\overline{b}),\overline{x}\right) \in \operatorname{gph} \mathcal{F}^{op}$. *Then*

$$\begin{aligned}\operatorname{clm} \mathcal{F}^{op}\left((\overline{c},\overline{b}),\overline{x}\right) &= \operatorname{clm} \mathcal{F}^{op}_{\overline{c}}\left(\overline{b},\overline{x}\right) = \max_{D \in \mathcal{M}_{\overline{c},\overline{b}}} \operatorname{clm} \mathcal{L}_D\left(\left(\overline{b}, -\overline{b}_D\right), \overline{x}\right) \\ &= \left(\min_{D \in \mathcal{M}_{\overline{c},\overline{b}}} d_*\left(0_n, \operatorname{end} \operatorname{conv}\left\{a_t,\ t \in T_{\overline{b}}(\overline{x}); -a_t,\ t \in D\right\}\right)\right)^{-1}.\end{aligned}$$

Let us introduce the natural extension for the set of extreme points:

$$\mathcal{E}(b) := \text{extr}\left(\mathcal{F}(b) \cap \text{span}\{a_t,\ t \in T\}\right),\ b \in \text{dom}\,\mathcal{F},$$
$$\mathcal{E}^{op}(c,b) := \text{extr}\left(\mathcal{F}^{op}(c,b) \cap \text{span}\{a_t,\ t \in T\}\right),\ (c,b) \in \text{dom}\,\mathcal{F}^{op}.$$

The reader is addressed to [13, p. 142] and [7, Section 2.2] for details about these constructions. Note that $\mathcal{E}^{op}(c,b) = \mathcal{F}^{op}(c,b) \cap \mathcal{E}(b)$ for $(c,b) \in \text{dom}\,\mathcal{F}^{op}$.

**Theorem 3** [3, Corollary 4.1, Proposition 4.2, and Theorem 4.2] *Let* $\left(\overline{c},\overline{b}\right) \in \text{dom}\,\mathcal{F}^{op}$*, then*

$$\begin{aligned}\text{Lipusc}\,\mathcal{F}^{op}\left(\overline{c},\overline{b}\right) &= \text{Lipusc}\,\mathcal{F}^{op}_{\overline{c}}\left(\overline{b}\right) \\ &= \sup_{x \in \mathcal{F}^{op}(\overline{\pi})} \text{clm}\,\mathcal{F}^{op}\left(\left(\overline{c},\overline{b}\right),x\right) = \max_{x \in \mathcal{E}^{op}(\overline{\pi})} \text{clm}\,\mathcal{F}^{op}\left(\left(\overline{c},\overline{b}\right),x\right).\end{aligned}$$

# 3 Hoffman constant for a class of polyhedral mappings

The main objective of this section is to establish equalities (12) for a certain class of multifunctions, which are introduced in the following definition. In this way, Hoffman constants for such multifunctions are determined by the local behavior of their graph, specifically, by the supremum of calmness moduli at all points of their graphs.

**Definition 2** *Let $I$ be a finite index set and, for each $i \in I$, consider a multifunction $\mathcal{S}_i : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with a convex polyhedral graph. We say that $\mathcal{S} := \bigcup_{i \in I} \mathcal{S}_i$ is a* well-connected polyhedral mapping *if the following properties are fulfilled:*

*(i)* $\text{dom}\,\mathcal{S}$ $\left(= \bigcup_{i \in I} \text{dom}\,\mathcal{S}_i\right)$ *is a convex set in* $\mathbb{R}^m$;

*(ii)* $\mathcal{S}\ |_{\text{dom}\,\mathcal{S}_i} = \mathcal{S}_i$*, for all $i \in I$ (equivalently, $\mathcal{S}_i(b) = \mathcal{S}_j(b)$ whenever $b \in \text{dom}\,\mathcal{S}_i \cap \text{dom}\,\mathcal{S}_j$, $i,j \in I$).*

Along this section we consider $\mathbb{R}^m$ and $\mathbb{R}^n$ endowed with arbitrary norms, both denoted by $\|\cdot\|$ for simplicity. Given a nonempty closed set $C \subset \mathbb{R}^n$, we denote by

$$P_C(x) := \text{argmin}_y\{\|y - x\| \mid y \in C\}$$

the set of best approximations (projections) of $x \in \mathbb{R}^n$ onto $C$. One easily checks (see [2, Lemma 1]) that for any $\widetilde{x} \in P_C(x)$ there holds

$$\widetilde{x} \in P_C(\widetilde{x} + \mu(x - \widetilde{x})),\ \text{for all}\ \mu \in [0,1]. \tag{14}$$

The following lemma constitutes the counterpart of [2, Lemma 1] into our current context; indeed, this result holds for multifunctions with a convex graph and a closed image at the reference point (the closedness of the whole graph is not required there). For completeness, we write a sketch of the proof. Hereafter, we liberalize the notation in the sense that symbol $\overline{b}$ does not refer to a given nominal parameter.

**Lemma 2** *Let $\mathcal{S} = \bigcup_{i \in I} \mathcal{S}_i$ be a well-connected polyhedral mapping. Assume that $b, \overline{b} \in \operatorname{dom} \mathcal{S}_k$ for some $k \in I$. Take $x \in \mathcal{S}(b)$ and any $\overline{x} \in P_{\mathcal{S}(\overline{b})}(x)$. Then*

$$\frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} \leq \operatorname{clm} \mathcal{S}\left(\overline{b}, \overline{x}\right).$$

**Proof.** Since $\operatorname{gph} \mathcal{S}_k$ is a convex polyhedral set,

$$(b_\mu, x_\mu) := (\overline{b}, \overline{x}) + \mu((b, x) - (\overline{b}, \overline{x})) \in \operatorname{gph} \mathcal{S}_k, \text{ for each } \mu \in [0,1].$$

According to (14), $\overline{x} \in P_{\mathcal{S}_k(\overline{b})}(x_\mu)$, for each $\mu \in [0,1]$. Moreover, taking into account that $\mathcal{S}_k\left(\overline{b}\right) = \mathcal{S}\left(\overline{b}\right)$, we have

$$\frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} = \frac{\|x - \overline{x}\|}{\left\|b - \overline{b}\right\|} = \frac{\|x_\mu - \overline{x}\|}{\left\|b_\mu - \overline{b}\right\|} = \frac{d\left(x_\mu, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b_\mu, \overline{b}\right)}, \text{ for all } \mu \in ]0,1].$$

Since $\lim_{\mu \downarrow 0} b_\mu = \overline{b}$, $\lim_{\mu \downarrow 0} x_\mu = \overline{x}$ and, for each $\mu \in ]0,1]$,

$$b_\mu \in \operatorname{dom} \mathcal{S}_k \text{ and } x_\mu \in \mathcal{S}_k(b_\mu) = \mathcal{S}(b_\mu),$$

appealing to the definition of calmness modulus, we obtain the claimed inequality

$$\frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} = \lim_{\mu \to 0} \frac{d\left(x_\mu, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b_\mu, \overline{b}\right)} \leq \operatorname{clm} \mathcal{S}\left(\overline{b}, \overline{x}\right).$$

■

**Definition 3** *Let $\mathcal{S} = \bigcup_{i \in I} \mathcal{S}_i$ be a well-connected polyhedral mapping. Let $b, \overline{b} \in \operatorname{dom} \mathcal{S}$. We call a subdivision $0 =: \mu_0 < \mu_1 < ... < \mu_N := 1$ together with a family of indices $i_1, ..., i_N \in I$ connecting $b$ with $\overline{b}$ if for all $k \in \{1, ..., N\}$ and all $\mu \in [\mu_{k-1}, \mu_k]$ there holds*

$$\overline{b} + \mu\left(b - \overline{b}\right) \in \operatorname{dom} \mathcal{S}_{i_k} \tag{15}$$

*(equivalently, $\mathcal{S}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right) = \mathcal{S}_{i_k}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right)$, whenever $\mu \in [\mu_{k-1}, \mu_k]$).*

**Lemma 3** *Let $\mathcal{S} = \bigcup_{i \in I} \mathcal{S}_i$ be a well-connected polyhedral mapping. For every pair $b, \overline{b} \in \operatorname{dom} \mathcal{S}$ there are $0 =: \mu_0 < \mu_1 < ... < \mu_N := 1$ together with a family of indices $i_1, ..., i_N \in I$ connecting $b$ with $\overline{b}$.*

**Proof.** We will follow a recursive process. Defining $\mu_0 = 0$, assume that for some $k \geq 0$ we have already found $\mu_0 < ... < \mu_k < 1$ and indices $i_1, ..., i_k \in I$ (not any index when $k = 0$) with property (15) whenever $\mu \in [\mu_{k-1}, \mu_k]$. Define $d := b - \overline{b}$ and pick sequences $\nu^j \downarrow \mu_k$ and $i_j \in I$ such that $\overline{b} + \nu^j d \in \operatorname{dom} \mathcal{S}_{i_j}$, where the latter exists for all $\nu^j$ sufficiently small due to convexity of $\operatorname{dom} \mathcal{S} = \bigcup_{i \in I} \operatorname{dom} \mathcal{S}_i$. Since there are only finitely many indices, by possibly passing to a subsequence we can assume that $i_j = i$ for all $j$, for some $i \in I$. Consider the set

$$R_{k+1} := \{\mu \geq 0 \mid \overline{b} + \mu d \in \operatorname{dom} \mathcal{S}_i\}.$$

Since $\operatorname{dom} \mathcal{S}_i$ is a closed convex set as the projection of the convex polyhedral set $\operatorname{gph} \mathcal{S}_i$ onto its first component, the set $R_{k+1}$ is a closed (possibly unbounded) interval. Now define

$$\mu_{k+1} := \begin{cases} \sup R_{k+1} & \text{if } \sup R_{k+1} < 1 \\ 1 & \text{otherwise} \end{cases}$$

and $i_{k+1} := i$. Since $R_{k+1} \ni \nu^j$ for all $j$, $\nu^j \downarrow \mu_k$ and $R_{k+1}$ is closed, we readily obtain $\mu_k \in R_{k+1}$ and $\mu_{k+1} > \mu_k$. Further, if $\sup R_{k+1} < \infty$ then $\sup R_{k+1} = \max R_{k+1}$ and consequently $\mu_{k+1} \in R_{k+1}$. Thus $[\mu_k, \mu_{k+1}] \subset R_{k+1}$ and therefore $\overline{b} + \mu d \in \operatorname{dom} \mathcal{S}_{k+1}$, whenever $\mu \in [\mu_k, \mu_{k+1}]$. We now claim that we can stop the procedure after a finite number $N$ of steps with $\mu_N = 1$ connecting $b$ with $\overline{b}$. Indeed, if the claim did not hold we would construct an infinite sequence of indices $\{i_r\}_{r \in \mathbb{N}} \subset I$ and consequently there would be $1 \leq p < q$ with $\mu_p < \mu_q$ and $i_p = i_q$, implying $R_p = R_q$. By our construction, this yields the contradiction $\mu_p = \mu_q$ and thus our claim holds true. ■

**Theorem 4** *Let $\mathcal{S} = \bigcup_{i \in I} \mathcal{S}_i$ be a well-connected polyhedral mapping. Let $b, \overline{b} \in \operatorname{dom} \mathcal{S}$ with $b \neq \overline{b}$ and consider a subdivision $0 =: \mu_0 < \mu_1 < ... < \mu_N := 1$ together with a family of indices $i_1, ..., i_N \in I$ connecting $b$ with $\overline{b}$. Then, for every $x \in \mathcal{S}(b)$, there exist points $x^k \in \mathcal{S}\left(\overline{b} + \mu_k d\right)$ with $k = 0, ..., N-1$ such that*

$$\begin{aligned} \frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} &\leq \max\{\operatorname{clm} \mathcal{S}\left(\overline{b} + \mu_k\left(b - \overline{b}\right), x^k\right) \mid k = 0, ..., N-1\} \\ &\leq \max\{\operatorname{Lipusc} \mathcal{S}\left(\overline{b} + \mu_k\left(b - \overline{b}\right)\right) \mid k = 0, ..., N-1\}. \end{aligned}$$

**Proof.** The second inequality comes directly from the definitions. We will prove the first one by induction in $N$. First assume that $N = 1$. Take any $x \in \mathcal{S}(b)$. Then both $\overline{b}$ and $b$ belong to $\operatorname{dom}\mathcal{S}_{i_1}$. Let $x^0 = P_{\mathcal{S}(\overline{b})}(x)$ and apply Lemma 2 to obtain the aimed inequality

$$\frac{d\left(x, \mathcal{S}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} \leq \operatorname{clm}\mathcal{S}\left(\overline{b}, x^0\right) =: \gamma_1.$$

Now consider a subdivision $0 =: \mu_0 < \mu_1 < ... < \mu_N := 1$ together with a family of indices $i_1, ..., i_N \in I$ connecting $b$ with $\overline{b}$. Take any $x \in \mathcal{S}(b)$ and consider

$$x^{N-1} = P_{\mathcal{S}_N\left(\overline{b}+\mu_{N-1}\left(b-\overline{b}\right)\right)}(x) = P_{\mathcal{S}\left(\overline{b}+\mu_{N-1}\left(b-\overline{b}\right)\right)}(x).$$

On the one hand, Lemma 2 yields

$$\frac{d\left(x, \mathcal{S}\left(\overline{b} + \mu_{N-1}\left(b - \overline{b}\right)\right)\right)}{d\left(b, \overline{b} + \mu_{N-1}\left(b - \overline{b}\right)\right)} \leq \operatorname{clm}\mathcal{S}\left(\overline{b} + \mu_{N-1}\left(b - \overline{b}\right), x^{N-1}\right),$$

equivalently

$$\left\|x - x^{N-1}\right\| \leq \operatorname{clm}\mathcal{S}\left(\overline{b} + \mu_{N-1}\left(b - \overline{b}\right), x^{N-1}\right)\left\|\left(1 - \mu_{N-1}\right)\left(b - \overline{b}\right)\right\|. \quad (16)$$

On the other hand, we can write

$$d\left(x, \mathcal{S}\left(\overline{b}\right)\right) \leq \left\|x - x^{N-1}\right\| + d\left(x^{N-1}, \mathcal{S}\left(\overline{b}\right)\right). \quad (17)$$

Now, consider

$$\widetilde{b} := \overline{b} + \mu_{N-1}\left(b - \overline{b}\right),$$

subdivision $0 =: \widetilde{\mu}_0 < \widetilde{\mu}_1 < ... < \widetilde{\mu}_{N-1} := 1$, with $\widetilde{\mu}_i := \dfrac{\mu_i}{\mu_{N-1}}$, $i = 1, ..., N-1$, together with the family of indices $i_1, ..., i_{N-1}$. Observe that this subdivision is connecting $\overline{b}$ and $\widetilde{b}$ as

$$\overline{b} + \widetilde{\mu}_i\left(\widetilde{b} - \overline{b}\right) = \overline{b} + \mu_i\left(b - \overline{b}\right),\ i = 1, ..., N-1. \quad (18)$$

Hence, we apply the induction hypothesis with $\overline{b}$ and $\widetilde{b}$ and subdivision $0 =: \widetilde{\mu}_0 < \widetilde{\mu}_1 < ... < \widetilde{\mu}_{N-1} := 1$ to conclude that for $x^{N-1} \in \mathcal{S}\left(\widetilde{b}\right)$ we can find points $x^k \in \mathcal{S}\left(\overline{b} + \widetilde{\mu}_k\left(\widetilde{b} - \overline{b}\right)\right)$ with $k = 0, ..., N-2$ such that

$$\frac{d\left(x^{N-1}, \mathcal{S}\left(\overline{b}\right)\right)}{\left\|\mu_{N-1}\left(b - \overline{b}\right)\right\|} \leq \max\{\operatorname{clm}\mathcal{S}\left(\overline{b} + \widetilde{\mu}_k\left(\widetilde{b} - \overline{b}\right), x^k\right) \mid k = 0, ..., N-2\} =: \widetilde{\gamma}. \quad (19)$$

Then, combining (16), (17), (18) and (19) we conclude,

$$\begin{aligned} d\left(x, \mathcal{S}\left(\overline{b}\right)\right) &\leq \operatorname{clm}\mathcal{S}\left(\widetilde{b}, x^{N-1}\right)\left\|\left(1-\mu_{N-1}\right)\left(b-\overline{b}\right)\right\| + \widetilde{\gamma}\left\|\mu_{N-1}\left(b-\overline{b}\right)\right\| \\ &\leq \max\{\operatorname{clm}\mathcal{S}\left(\overline{b}+\mu_k\left(b-\overline{b}\right), x^k\right) \mid k=0,...,N-1\}\left\|\left(b-\overline{b}\right)\right\|, \end{aligned}$$

as we wanted to prove. ■

**Corollary 1** *Let $\mathcal{S}$ be a well-connected polyhedral mapping. Then*

$$\operatorname{Hof}\mathcal{S} = \sup\{\operatorname{Lipusc}\mathcal{S}\left(b\right) \mid b \in \operatorname{dom}\mathcal{S}\} = \sup\{\operatorname{clm}\mathcal{S}\left(b, x\right) \mid \left(b, x\right) \in \operatorname{gph}\mathcal{S}\}.$$

**Proof.** The inequalities "$\geq$" are always true (recall (11)) and inequalities "$\leq$" follow from Theorem 4. ■

**Remark 1** Apart from conditions $(i)$ and $(ii)$ in Definition 2, the results of this section only require mappings $\mathcal{S}_i$, to satisfy the following conditions (which are obviously held when $\operatorname{gph}\mathcal{S}_i$, $i \in I$, are convex polyhedral sets):

- $\operatorname{gph}\mathcal{S}_i$ is a (nonempty) closed convex set in $\mathbb{R}^m \times \mathbb{R}^n$,
- $\operatorname{dom}\mathcal{S}_i$ is a closed convex set in $\mathbb{R}^m$.

Nevertheless, the section has been written for polyhedral mappings having in mind our application to the argmin mapping under RHS perturbations.

The following example shows that the assumption of Corollary 1 is not superfluous.

**Example 1** Consider mappings $\mathcal{M}_1$, $\mathcal{M}_2 : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$\mathcal{M}_1(y) = \begin{cases} 0 \text{ if } y \leq 0 \\ 1 \text{ if } y > 0 \end{cases} \quad \textit{and } \mathcal{M}_2(y) = \begin{cases} 0 \text{ if } y < 0 \\ [0,1] \text{ if } y \geq 0 \end{cases}$$

Observe that

$$\operatorname{clm}\mathcal{M}_1\left(y, x\right) = 0 \ \forall \left(y, x\right) \in \operatorname{gph}\mathcal{M}_1, \ \operatorname{Lipusc}\mathcal{M}_1\left(0\right) = +\infty\ ,$$

and

$$\operatorname{Lipusc}\mathcal{M}_2\left(y\right) = 0 \ \forall y \in \operatorname{dom}\mathcal{M}, \ \textit{while} \ \operatorname{Hof}\mathcal{M} = +\infty.$$

Clearly $\mathcal{M}_1$ is not polyhedral as $\operatorname{gph}\mathcal{M}_1$ is not a finite union of convex polyhedral sets. Regarding $\mathcal{M}_2$, it is a polyhedral mapping, which can be represented as $\mathcal{M}_2 = \mathcal{M}_{21} \cup \mathcal{M}_{22}$, with

$$\operatorname{gph}\mathcal{M}_{21} = ]-\infty, 0] \times \{0\}, \ \operatorname{gph}\mathcal{M}_{22} = [0, +\infty[ \times [0,1],$$

although condition $(ii)$ in Definition 2 fails, since $0 \in \operatorname{dom}\mathcal{M}_{21} \cap \operatorname{dom}\mathcal{M}_{22}$ but $\mathcal{M}_{21}\left(0\right) \neq \mathcal{M}_{22}\left(0\right)$.

# 4 Hoffman constant of the argmin mapping under RHS perturbations

This section gathers the results about the Hoffman constant for the argmin mapping in linear optimization. It is divided into two subsections depending on the perturbation setting. In fact, as advanced in the introduction, the interesting case is the one of RHS perturbations.

## 4.1 Canonical perturbations

This subsection deals with the optimal set mapping in the context of canonical perturbations, $\mathcal{F}^{op}$, introduced in (3). Indeed, it is oriented to show that the computation of $\operatorname{Hof}\mathcal{F}^{op}$ has no interest as it is infinite unless we are placed in a trivial case formalized in the following proposition.

Recall that, in ordinary (finite) linear programming, optimality is equivalent to primal-dual consistency. In other words,

$$\operatorname{dom}\mathcal{F}^{op} = \left(-\operatorname{cone}\{a_t,\ t\in T\}\right)\times\operatorname{dom}\mathcal{F}. \tag{20}$$

Clearly, $\operatorname{gph}\mathcal{F}$, $\operatorname{dom}\mathcal{F}$, and $\operatorname{dom}\mathcal{F}^{op}$ are convex, while $\operatorname{gph}\mathcal{F}^{op}$ is no longer convex.

**Proposition 2** *We have that*

$$\operatorname{Hof}\mathcal{F}^{op} = \begin{cases} 0 \text{ if } \{a_t,\ t\in T\} = \{0_n\}, \\ +\infty \text{ otherwise.}\end{cases}$$

**Proof.** In the case when $\{a_t,\ t\in T\} = \{0_n\}$, one trivially has

$$\mathcal{F}^{op}(c,b) = \begin{cases} \mathbb{R}^n \text{ if } (c,b)\in\{0_n\}\times\mathbb{R}^m_+, \\ \emptyset \text{ otherwise.}\end{cases}$$

So, it is clear that $\operatorname{Hof}\mathcal{F}^{op} = 0$.

Assume now that $\{a_t,\ t\in T\} \neq \{0_n\}$. Take $\overline{x}\in\mathbb{R}^n$ and define $\overline{b}_t = a_t'\overline{x}+1$ for all $t\in T$. Hence $\overline{x}\in\operatorname{int}\mathcal{F}\left(\overline{b}\right)$ and it cannot be an optimal solution for any $\left(c,\overline{b}\right)\in\operatorname{dom}\mathcal{F}^{op}$ for any $c\neq 0_n$. Fix any $\overline{c}\in -\operatorname{cone}\{a_t,\ t\in T\}\backslash\{0_n\}$. We have $d\left(\overline{x},\mathcal{F}^{op}\left(\overline{c},\overline{b}\right)\right) = d\left(\overline{x},\mathcal{F}^{op}\left(\frac{1}{r}\overline{c},\overline{b}\right)\right) > 0$ for all $r\in\mathbb{N}$.

On the other hand, it is clear that $\overline{x}\in\mathcal{F}^{op}\left(0_n,\overline{b}\right)$. Hence

$$\operatorname{Hof}\mathcal{F}^{op} \geq \lim_{r\to+\infty}\frac{d\left(\overline{x},\mathcal{F}^{op}\left(\frac{1}{r}\overline{c},\overline{b}\right)\right)}{d\left(\left(0_n,\overline{b}\right),\left(\frac{1}{r}\overline{c},\overline{b}\right)\right)} = +\infty.$$

■

### 4.2 RHS perturbations

This subsection is concerned with multifunction $\mathcal{F}_{\overline{c}}^{op}$ defined in (4). We assume that $-\overline{c} \in \text{cone}\{a_t,\, t \in T\}$, which entails, according to well-known arguments in ordinary (finite) linear programming, that

$$\text{dom}\,\mathcal{F}_{\overline{c}}^{op} = \text{dom}\,\mathcal{F},$$

hence $\text{dom}\,\mathcal{F}_{\overline{c}}^{op}$ is a closed and convex set in $\mathbb{R}^m$. On the other hand, $\text{gph}\,\mathcal{F}_{\overline{c}}^{op}$ is no longer convex, although it is a finite union of convex polyhedral sets. Indeed, let us show that $\mathcal{F}_{\overline{c}}^{op}$ is a well-connected polyhedral mapping. To do this, looking at the notation of Definition 2, the role of index set $I$ will be played by the following family:

$$\mathcal{M}_{\overline{c}} := \left\{ D \subset T \;\middle|\; \begin{array}{l} -\overline{c} \in \text{cone}\{a_t,\, t \in D\} \text{ and } D \text{ is} \\ \text{minimal w.r.t. the inclusion order} \end{array} \right\}. \tag{21}$$

**Proposition 3** *Let* $-\overline{c} \in \text{cone}\{a_t,\; t \in T\}$. *Then* $\mathcal{M}_{\overline{c}} = \bigcup_{b \in \text{dom}\,\mathcal{F}} \mathcal{M}_{\overline{c},b}$.

**Proof.** Inclusion '$\supset$' is evident. On the other hand, for any $D \in \mathcal{M}_{\overline{c}}$ we have $\{D\} = \mathcal{M}_{\overline{c},b^D} = \mathcal{M}_{\overline{c},b^D}(0_n)$, where $b_t^D := 0$ if $t \in D$ and $b_t^D := 1$ if $t \in T \backslash D$. ■

Associated to each $D \in \mathcal{M}_{\overline{c}}$ we consider

$$\mathcal{S}_D(b) = \left\{ x \in \mathbb{R}^n : a_t' x \leq b_t,\, t \in T \backslash D;\; a_t' x = b_t,\, t \in D \right\}. \tag{22}$$

Observe that, recalling (13),

$$\mathcal{S}_D(b) = \mathcal{L}_D\left(b, -b^D\right), \text{ whenever } b \in \mathbb{R}^m.$$

**Proposition 4** *Assume that* $-\overline{c} \in \text{cone}\{a_t,\, t \in T\}$. *Then, according to the previous notation, we have:*
*(i)* $\mathcal{F}_{\overline{c}}^{op} = \bigcup_{D \in \mathcal{M}_{\overline{c}}} \mathcal{S}_D$;
*(ii)* $\mathcal{F}_{\overline{c}}^{op} |_{\text{dom}\,\mathcal{S}_D} = \mathcal{S}_D$.
*Accordingly,* $\mathcal{F}_{\overline{c}}^{op}$ *is a well-connected polyhedral mapping.*

**Proof.** Condition $(i)$ comes directly from the well-known Karush-Kuhn-Tucker conditions in linear programming, while condition $(ii)$ can be derived from Proposition 1. Specifically, if $b \in \text{dom}\,\mathcal{S}_D$, then, it is clear that $D \in \mathcal{M}_{\overline{c},b}$ and, hence Proposition 1 yields

$$\mathcal{F}^{op}(\overline{c}, b) = \mathcal{L}_D(b, -b_D) = \mathcal{S}_D(b).$$

Moreover, it is clear that $\operatorname{gph} \mathcal{S}_D$ is a convex polyhedral set and that $\operatorname{dom} \mathcal{S}_D$ is a closed convex set as the projection of this convex polyhedral set, $\operatorname{gph} \mathcal{S}_D$, onto its first component. Consequently, $\mathcal{F}_{\overline{c}}^{op}$ is a well-connected polyhedral mapping. ■

**Theorem 5** *Let* $-\overline{c} \in \operatorname{cone}\{a_t,\ t \in T\}$. *One has*

$$\begin{aligned} \operatorname{Hof} \mathcal{F}_{\overline{c}}^{op} &= \max_{b \in \operatorname{dom} \mathcal{F}} \operatorname{Lipusc} \mathcal{F}^{op}(\overline{c}, b) \\ &= \max_{b \in \operatorname{dom} \mathcal{F}} \max_{x \in \mathcal{E}^{op}(\overline{c}, b)} \operatorname{clm} \mathcal{F}^{op}((\overline{c}, b), x) \\ &= \max_{\substack{D \subset S \subset T \\ D \in \mathcal{M}_{\overline{c}}}} \left\{ d_* \left( 0_n, \operatorname{end} \operatorname{conv}\{a_t,\ t \in S;\ -a_t,\ t \in D\} \right) \right\}^{-1}. \end{aligned}$$

**Proof.** The first equality comes from Corollary 1. The second equality comes from Theorems 2 and 3. To prove the third equality, according to the mentioned theorems, and the fact that the inverse of the minimum of positive numbers is the maximum of the inverses, it will be enough to prove that, for each $D \in \mathcal{M}_{\overline{c}}$ and each $D \subset S \subset T$, there exist $b_{D,S} \in \operatorname{dom} \mathcal{F}$ and $x_{D,S} \in \mathcal{F}^{op}(\overline{c}, b_{D,S})$ such that

$$d_* \left( 0_n, \operatorname{end} \operatorname{conv}\{a_t,\ t \in S;\ -a_t,\ t \in D\} \right)^{-1} = \operatorname{clm} \mathcal{F}^{op}((\overline{c}, b_{D,S}), x_{D,S}),$$

and for having this it suffices to prove that $S = T_{b_{D,S}}(x_{D,S})$, which automatically implies $D \in \mathcal{M}_{\overline{c}, b_{D,S}}$ (because of $D \in \mathcal{M}_{\overline{c}}$). This can be done by just taking

$$x_{D,S} = 0_n \text{ and } b_{D,S}(t) = \begin{cases} 0 & \text{if } t \in S, \\ 1 & \text{if } t \in T \backslash S. \end{cases}$$

■

# 5 Break steps and directional behavior of the optimal set

This section is focused on some features of the optimal set along the segment determined by two elements $\overline{b}$ and $b$ of its domain; i.e., on the behavior of $\mathcal{F}_{\overline{c}}^{op}\left(\overline{b} + \mu\left(b - \overline{b}\right)\right)$, provided that $\overline{b}, b \in \operatorname{dom} \mathcal{F}_{\overline{c}}^{op}$ and $\mu \in [0, 1]$. Specifically, the main contribution of this section is to provide a way of computing a subdivision $0 =: \mu_0 < \mu_1 < ... < \mu_N := 1$ together with a family $D_1, ..., D_N \in \mathcal{M}_{\overline{c}}$ connecting $b$ with $\overline{b}$. Recall that the existence of such subdivision is guaranteed since $\mathcal{F}_{\overline{c}}^{op}$ is a well connected polyhedral mapping (recall Lemma 3).

**Lemma 4** *Let $\overline{b}, b \in \operatorname{dom} \mathcal{F}$. Assume $\mathcal{M}_{\overline{c},\overline{b}} = \mathcal{M}_{\overline{c},b} \neq \emptyset$. Then,*

$$\mathcal{M}_{\overline{c},\overline{b}+\mu\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}} \text{ for all } \mu \in [0,1].$$

**Proof.** First observe that $\overline{b} + \mu\left(b - \overline{b}\right) \in \operatorname{dom} \mathcal{F}$ for all $\mu \in ]0,1[$ because of the convexity of $\operatorname{dom} \mathcal{F}$. Fix arbitrarily $\overline{x} \in \mathcal{F}_{\overline{c}}^{op}\left(\overline{b}\right)$ and $x \in \mathcal{F}_{\overline{c}}^{op}(b)$ and consider the convex combination

$$\left(b^{\mu}, x^{\mu}\right) := \left(\overline{b}, \overline{x}\right) + \mu\left(b - \overline{b}, x - \overline{x}\right) \in \operatorname{gph} \mathcal{F},\ \mu \in ]0,1[.$$

It is immediate from the definitions that

$$T_{b^{\mu}}\left(x^{\mu}\right) = T_{\overline{b}}\left(\overline{x}\right) \cap T_{b}\left(x\right),\ \text{for all } \mu \in ]0,1[. \tag{23}$$

Just observe that, for any $t \in T$,

$$a_t' x^{\mu} - b_t^{\mu} = (1-\mu)\left(a_t' \overline{x} - \overline{b}_t\right) + \mu\left(a_t' x - b_t\right) \leq 0, \tag{24}$$

and equality holds if and only if $a_t'\overline{x} - \overline{b}_t = a_t' x - b_t = 0$. Take any $\mu \in ]0,1[$ and let us show that $\mathcal{M}_{\overline{c},b^{\mu}} = \mathcal{M}_{\overline{c},\overline{b}}$.

Observe that any $D \in \mathcal{M}_{\overline{c},\overline{b}} = \mathcal{M}_{\overline{c},b}$ is contained in $T_{b^{\mu}}\left(x^{\mu}\right)$ because of (23), and clearly $D$ is minimal among the subsets of $T_{b^{\mu}}\left(x^{\mu}\right)$ satisfying $-\overline{c} \in \operatorname{cone}\{a_t,\ t \in D\}$, since this minimality happens, for instance, in the subsets of $T_{\overline{b}}\left(\overline{x}\right)$. In other words, $D \in \mathcal{M}_{\overline{c},b^{\mu}}$.

In order to check the converse inclusion, take any $D \in \mathcal{M}_{\overline{c},b^{\mu}}$, in particular $D \in T_{b^{\mu}}\left(x^{\mu}\right)$ and hence $D \in T_{\overline{b}}\left(\overline{x}\right)$. Moreover, the minimality of $D$ among the subsets of $T_{\overline{b}}\left(\overline{x}\right)$ comes from the minimality over the subsets of $T_{b^{\mu}}\left(x^{\mu}\right)$. ■

**Definition 4** *Given $\overline{b}, b \in \operatorname{dom} \mathcal{F}$, $\overline{b} \neq b$, we define the* break step set *between $\overline{b}$ and $b$ by:*

$$\mathcal{B}\left(\overline{b}, b\right) := \left\{\mu \in ]0,1[ \,\middle|\, \exists \nu_r \to \mu\ ,\ \mathcal{M}_{c,\overline{b}+\nu_r\left(b-\overline{b}\right)} \subsetneqq \mathcal{M}_{c,\overline{b}+\mu\left(b-\overline{b}\right)}\ \forall r \in \mathbb{N}\right\}.$$

**Remark 2** Because of Lemma 1, in the previous definitions we might replace "$\subsetneqq$" with "$\neq$".

**Proposition 5** *Given $\overline{b}, b \in \operatorname{dom} \mathcal{F}$, $\overline{b} \neq b$, we have:*

*(i) $\mathcal{B}\left(\overline{b}, b\right)$ is a finite set (possibly empty).*

*(ii) If $\mathcal{B}\left(\overline{b}, b\right) = \{\mu_1, \mu_2, ..., \mu_N\}$, with $0 =: \mu_0 < \mu_1 < \mu_2 < ... < \mu_N < \mu_{N+1} := 1$, then*

$$\mu \longmapsto \mathcal{M}_{c,\overline{b}+\mu\left(b-\overline{b}\right)} \text{ is constant on } ]\mu_{i-1}, \mu_i[ \text{ for any } i = 1, 2, ..., N+1.$$

*(Take $N = 0$ when $\mathcal{B}\left(\overline{b}, b\right) = \emptyset$, in which case we just have $\mathcal{M}_{\overline{c},\overline{b}+\mu\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}}$ for all $\mu \in [0,1]$).*

*(iii) Let $\mathcal{B}\left(\overline{b}, b\right) = \{\mu_1, \mu_2, ..., \mu_N\}$, with $0 := \mu_0 < \mu_1 < ... < \mu_N < \mu_{N+1} := 1$ and, for each $i = 1, ..., N+1$, fix any*

$$D_i \in \mathcal{M}_{\overline{c},\overline{b}+\mu\left(b-\overline{b}\right)},\ \mu \in ]\mu_{i-1}, \mu_i[.$$

*Then subdivision $0 := \mu_0 < \mu_1 ... \mu_N < \mu_{N+1} = 1$ together with the family $D_1, ..., D_{N+1} \in \mathcal{M}_{\overline{c}}$ is connecting $b$ with $\overline{b}$.*

**Proof.** $(i)$ Reasoning by contradiction, assume that $\mathcal{B}\left(\overline{b}, b\right)$ is infinite, take a sequence of scalars $\{\mu_r\}_{r\in\mathbb{N}} \subset \mathcal{B}\left(\overline{b}, b\right)$ and consider the corresponding sequence of subsets $\left\{\mathcal{M}_{\overline{c},\overline{b}+\mu_r\left(b-\overline{b}\right)}\right\}_{r\in\mathbb{N}}$. Since $\mathcal{M}_{\overline{c},\overline{b}+\mu_r\left(b-\overline{b}\right)} \subset \mathcal{M}_{\overline{c}}$, for all $r$, and $\mathcal{M}_{\overline{c}}$ is finite, there exist infinitely many repeated subsets in $\left\{\mathcal{M}_{\overline{c},\overline{b}+\mu_r\left(b-\overline{b}\right)}\right\}_{r\in\mathbb{N}}$. In particular, we can take three break steps $\mu_{r_1} < \mu_{r_2} < \mu_{r_3}$ with the same minimal KKT set of indices. Then, since $\mathcal{M}_{\overline{c},\overline{b}+\mu_{r_1}\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}+\mu_{r_3}\left(b-\overline{b}\right)}$, applying the previous lemma (with $\overline{b} + \mu_{r_1}\left(b - \overline{b}\right)$ and $\overline{b} + \mu_{r_3}\left(b - \overline{b}\right)$ playing the role of $\overline{b}$ and $b$), $\mathcal{M}_{\overline{c},\overline{b}+\mu\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}+\mu_{r_1}\left(b-\overline{b}\right)}$ for all $\mu \in [\mu_{r_1}, \mu_{r_3}]$, which contradicts the fact that $\mu_{r_2}$ is also a break step between $\overline{b}$ and $b$.

$(ii)$ Fix any $i \in \{1, ..., N+1\}$ and let us see that $\mathcal{M}_{c,\overline{b}+\mu\left(b-\overline{b}\right)}$ is constant on $]\mu_{i-1}, \mu_i[$. Arguing by contradiction, assume that there exist $\overline{\mu}$ and $\widetilde{\mu}$ with $\mu_{i-1} < \overline{\mu} < \widetilde{\mu} < \mu_i$ such that $\mathcal{M}_{\overline{c},\overline{b}+\overline{\mu}\left(b-\overline{b}\right)} \neq \mathcal{M}_{\overline{c},\overline{b}+\widetilde{\mu}\left(b-\overline{b}\right)}$. Define

$$\alpha := \sup\left\{\mu > 0 \mid \mathcal{M}_{\overline{c},\overline{b}+\mu\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}+\overline{\mu}\left(b-\overline{b}\right)}\right\}.$$

Observe that $\alpha > \overline{\mu}$ since $\overline{\mu} \notin \mathcal{B}\left(\overline{b}, b\right)$ (which comes from the definition of break step together with Remark **??**). Indeed, $\alpha \geq \mu_i$ (contradicting the existence of $\widetilde{\mu}$) since, if we had $\alpha < \mu_i$ we would attain a contradiction by distinguishing two cases:

*Case 1*: If $\mathcal{M}_{\overline{c},\overline{b}+\alpha\left(b-\overline{b}\right)} = \mathcal{M}_{\overline{c},\overline{b}+\overline{\mu}\left(b-\overline{b}\right)}$ by definition of supremum, there would exist a decreasing sequence of scalars $\nu_j \downarrow \alpha$ with $\mathcal{M}_{\overline{c},\overline{b}+\nu_j\left(b-\overline{b}\right)} \neq \mathcal{M}_{\overline{c},\overline{b}+\alpha\left(b-\overline{b}\right)}$. Then we would have $\alpha \in \mathcal{B}\left(\overline{b}, b\right)$ which represents a contradiction (observe that $\mu_{i-1} < \alpha < \mu_i$).

*Case 2*: If $\mathcal{M}_{\overline{c},\overline{b}+\alpha\left(b-\overline{b}\right)} \neq \mathcal{M}_{\overline{c},\overline{b}+\overline{\mu}\left(b-\overline{b}\right)}$, again by the definition of supremum, there would exists an increasing sequence $\nu_j \uparrow \alpha$ with $\mathcal{M}_{\overline{c},\overline{b}+\nu_j\left(b-\overline{b}\right)} \neq \mathcal{M}_{\overline{c},\overline{b}+\alpha\left(b-\overline{b}\right)}$. Again, we would attain the contradiction $\alpha \in \mathcal{B}\left(\overline{b}, b\right)$.

$(iii)$ is a direct consequence of $(i)$ and $(ii)$ together with Proposition 4. Specifically, $\mathcal{F}_{\overline{c}}^{op} = \bigcup_{D \in \mathcal{M}_{\overline{c}}} \mathcal{S}_D$, subsets $D_1, ..., D_N \in \mathcal{M}_{\overline{c}}$ and

$$\overline{b} + \mu\left(b - \overline{b}\right) \operatorname{dom} \mathcal{S}_{D_i} \text{ for all } \mu \in [\mu_{i-1}, \mu_i],$$

since $D_i \in \mathcal{M}_{\overline{c}, \overline{b} + \mu\left(b - \overline{b}\right)}$ for all $\mu \in ]\mu_{i-1}, \mu_i[$, and taking the closedness of $\operatorname{dom} \mathcal{S}_{D_i}$ (recall Definition 3). ■

The following result is a direct consequence of Theorem 4 and Propositions 4 and 5.

**Theorem 6** *Let $b, \overline{b} \in \operatorname{dom} \mathcal{F}$, $b \neq \overline{b}$, and consider the set of break steps $\mathcal{B}\left(\overline{b}, b\right) = \{\mu_1, \mu_2, ..., \mu_N\}$, with $0 := \mu_0 < \mu_1 < \mu_2 < ... < \mu_N < \mu_{N+1} := 1$. Then for every $x \in \mathcal{F}_{\overline{c}}^{op}(b)$ one has*

$$\frac{d\left(x, \mathcal{F}_{\overline{c}}^{op}\left(\overline{b}\right)\right)}{d\left(b, \overline{b}\right)} \leq \max\{\operatorname{Lipusc} \mathcal{F}_{\overline{c}}^{op}\left(\overline{b} + \mu_k d\right) \mid k = 0, ..., N\}.$$

# 6 Conclusions, perspectives and examples

The main contributions of the current work are, on the one hand, to ensure the fulfilment of the following equalities

$$\operatorname{Hof} \mathcal{S} = \sup\{\operatorname{Lipusc} \mathcal{S}(b) \mid b \in \operatorname{dom} \mathcal{S}\} = \sup\{\operatorname{clm} \mathcal{S}(b, x) \mid (b, x) \in \operatorname{gph} \mathcal{S}\}, \tag{25}$$

provided that $\mathcal{S}$ be a well-connected polyhedral mapping, which is the case of the argmin mapping $\mathcal{F}_{\overline{c}}^{op}$ (under RHS perturbations). In this way, the global Lipschitzian behavior of optimal solutions is characterized through the local one. On the other hand, in the particular case when $\mathcal{S} = \mathcal{F}_{\overline{c}}^{op}$ we can go further to derive a point based expression for such constant:

$$\operatorname{Hof} \mathcal{F}_{\overline{c}}^{op} = \max_{\substack{D \subset S \subset T \\ D \in \mathcal{M}_{\overline{c}}}} \{d_*\left(0_n, \operatorname{end} \operatorname{conv}\{a_t,\ t \in S;\ -a_t,\ t \in D\}\right)\}^{-1}.$$

There is another type of Hoffman constants located at a fixed point of the domain of the multifunction under consideration. Following the terminology of [2], the *Hoffman modulus* of $\mathcal{S}$ *at* $\overline{b} \in \operatorname{dom} \mathcal{S}$ is defined as

$$\operatorname{Hof} \mathcal{S}(\overline{b}) := \sup_{(b,x) \in \operatorname{gph} \mathcal{S}} \frac{d(x, \mathcal{S}(\overline{b}))}{d\left(b, \overline{b}\right)}. \tag{26}$$

It is clear that we can add the following equality to (25):

$$\operatorname{Hof} \mathcal{S} = \sup_{b \in \operatorname{dom} \mathcal{S}} \operatorname{Hof} \mathcal{S}(b).$$

In the case when $\operatorname{gph}\mathcal{S}$ is a convex polyhedral set $\operatorname{Hof}\mathcal{S}(\overline{b}) = \operatorname{Lipusc}\mathcal{S}\left(\overline{b}\right)$ (see [2, Theorem 4 ] for a more general framework where this equality holds), which is the case of the feasible set mapping $\mathcal{S}$. However, this is not the case of a general, even well-connected, polyhedral mapping as the following simple example shows.

**Example 2** Consider the mapping $\mathcal{S} : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$\mathcal{S}(y) = \begin{cases} 0 \text{ if } y \leq 0 \\ y \text{ if } y > 0 \end{cases}$$

Observe that

$$\operatorname{Lipusc}\mathcal{S}\,(-1) = 0 \textit{ while } \operatorname{Hof}\mathcal{S}(-1) = \operatorname{Hof}\mathcal{S} = 1.$$

Another feature about $\operatorname{Hof}\mathcal{S}(\overline{b})$ is that the supremum in (26) may be attained or not. In the previous example, it is not attained, while, for instance, for the new mapping given by $\widetilde{\mathcal{S}}(y) = 0$, *if* $y \leq 0$, $\widetilde{\mathcal{S}}(y) = y$, if $y \in [0, 1]$, $\widetilde{\mathcal{S}}(y) = 1$, if $y \geq 1$, we have

$$\operatorname{Hof}\widetilde{\mathcal{S}} = 1 > \operatorname{Hof}\widetilde{\mathcal{S}}(-1) = \frac{|\mathcal{S}\,(1) - \mathcal{S}\,(-1)|}{1-(-1)} = \frac{1}{2} > \operatorname{Lipusc}\widetilde{\mathcal{S}}\,(-1) = 0.$$

Moreover, in contrast to Lipschitz upper semicontinuity moduli or to the Hoffman constant, the Hoffman modulus $\operatorname{Hof}\widetilde{\mathcal{S}}(-1)$ cannot be expressed in terms of calmness moduli, since the only possible calmness moduli in this example are either 0 or 1.

All the previous situations may also happen in general for mapping $\mathcal{F}_{\overline{c}}^{op}$. Specifically, $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}(\overline{b})$ may be strictly in between $\operatorname{Lipusc}\mathcal{F}_{\overline{c}}^{op}\left(\overline{b}\right)$ and $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}$, and the supremum defining $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}(\overline{b})$ may be attained or not. Moreover, are not enough to express $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}(\overline{b})$.

While we already have (Theorems 2 and 3) point-based formulae (only involving the nominal data) for computing $\operatorname{Lipusc}\mathcal{F}_{\overline{c}}^{op}\left(\overline{b}\right)$ and $\operatorname{clm}\mathcal{F}_{\overline{c}}^{op}\left(\overline{b}, x\right)$, the computation of $\operatorname{Hof}\mathcal{F}_{\overline{c}}^{op}(\overline{b})$ remains as open problem.

# References


[1] D. AZÉ, J.-N. CORVELLEC, *On the sensitivity Analysis of Hoffman constants for systems of linear inequalities*, SIAM J. Optim. 12 (2002), pp. 913-927.

[2] J. CAMACHO, M. J. CÁNOVAS, J. PARRA, *From calmness to Hoffman constants for linear semi-infinite inequality systems*, SIAM J, Optim. 32 (2022), pp. 2859-2878.

[3] J. CAMACHO, M. J. CÁNOVAS, J. PARRA, *Lipschitz upper semicontinuity in linear optimization via local directional convexity*, Optimization, published online 06 Apr 2022, DOI:10.1080/02331934.2022.2057851

[4] M. J. CÁNOVAS, R. HENRION, M. A. LÓPEZ, J. PARRA, *Outer limit for subdifferentials and calmness moduli in linear and nonlinear programming*, J. Optim. Theory Appl. 169 (2016), pp. 925-952.

[5] M. J. CÁNOVAS, M. A. LÓPEZ, J. PARRA, F. J. TOLEDO, *Calmness of the feasible set mapping for linear inequality systems*, Set-Valued Var. Anal. 22 (2014), pp. 375–389.

[6] A. L. DONTCHEV, R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer, New York, 2009.

[7] M.J. GISBERT, M.J. CANOVAS, J. PARRA, F.J. TOLEDO, *Calmness of the optimal value in linear programming.* SIAM J. Optim. 28 (2018), pp. 2201 - 2221.

[8] M. A. GOBERNA, M. A. LÓPEZ, *Linear Semi-Infinite Optimization*, John Wiley & Sons, Chichester (UK), 1998.

[9] A.J. HOFFMAN, *On approximate solutions of systems of linear inequalities.* J. Res. Natl. Bur. Stand. 49 (1952), pp. 263–265.

[10] H. HU, *Characterizations of the strong basic constraint qualifications.* Math Oper Res. 30 (2005), pp. 956-965.

[11] D. KLATTE, B. KUMMER, *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, Nonconvex Optim. Appl. 60, Kluwer Academic, Dordrecht, The Netherlands, 2002.

[12] D. KLATTE, G. THIERE, *Error bounds for solutions of linear equations and inequalities* Z. Oper. Res. 41 (1995), pp. 191-214.

[13] W. LI, *Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs.* SIAM J. Control Optim. 32 (1994), pp. 140–153.

[14] Z, LUO, P. TSENG, *Error bounds and convergence analysis of feasible descent methods: a general approach.* Ann. Oper. Res. 46(1), 157–178 (1993)

[15] B. S. MORDUKHOVICH, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin, 2006.

[16] I. NECOARA, Y. NESTEROV, F. GLINEAUR, *Linear convergence of first order methods for non-strongly convex optimization.* Math. Program. 175, 69–107 (2019)

[17] J. PEÑA, D. RODRÍGUEZ, *Polytope conditioning and linear convergence of the Frank–Wolfe algorithm.* Math. Oper. Res. 44, 1–18 (2019)

[18] J. PEÑA, J.C. VERA, L.F. ZULUAGA, *New characterizations of Hoffman constants for systems of linear constraints*, Math. Program. 187 (2021), pp. 79–109.

[19] R. T. ROCKAFELLAR, R. J-B. WETS, *Variational Analysis*, Springer, Berlin, 1998.

[20] C. ZĂLINESCU, *Sharp estimates for Hoffman's constant for systems of linear inequalities and equalities*, SIAM J. Optim. 14 (2003), pp. 517-533.

[21] Z. ZHOU, A. SO, *A unified approach to error bounds for structured convex optimization problems.* Math. Program. 165(2), 689–728 (2017)