



**HAL**  
open science

# On the B-differential of the componentwise minimum of two affine vectorial functions -The full report

Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain

## ► To cite this version:

Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain. On the B-differential of the componentwise minimum of two affine vectorial functions -The full report. 2022. hal-03872711v1

**HAL Id: hal-03872711**

**<https://hal.science/hal-03872711v1>**

Preprint submitted on 25 Nov 2022 (v1), last revised 2 Aug 2024 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the B-differential of the componentwise minimum of two affine vectorial functions – The full report\*

Jean-Pierre DUSSAULT<sup>†</sup>, Jean Charles GILBERT<sup>‡</sup> and Baptiste PLAQUEVENT-JOURDAIN<sup>§</sup>

This paper focuses on the description and the computation of the B-differential of the componentwise minimum of two affine functions with vectorial values. The problem has many equivalent formulations and we mention some of them in linear algebra, convex analysis and discrete geometry. The original problem arises in the reformulation of the linear complementarity problem with the Min C-function, which was our initial motivation. We use the various problem formulations to shed some light on the properties of the B-differential. The set to describe has a finite number of elements, which can be exponential, so that its description is often algorithmic. We propose several modifications, adapted to the problem at stake, of an output-sensitive incremental algorithm introduced by Rada and Černý in 2018 for determining the cells of an arrangement in the space of hyperplanes having a point in common. As shown by numerical experiments, these modifications improve the performance of the algorithm significantly.

**Keywords:** B-differential • Bipartition of a finite set • C-differential • Complementarity problem • Complexity • Componentwise minimum of functions • Connectivity • Hyperplane arrangement • Matrix null space • Orthant • Pointed cone • Strict linear inequalities • Symmetry • Winder’s formula.

**AMS MSC 2020:** 05A18, 05C40, 26A24, 26A27, 46N10, 47A50, 47A63, 49J52, 49N15, 52C35, 65Y20, 65K15, 90C33, 90C46.

---

\*This is an extended version of the paper [24]. It contains more proofs and comments. The added text is written in dark blue.

<sup>†</sup>Département d’Informatique, Faculté des Sciences, Université de Sherbrooke, Canada; E-mail: Jean-Pierre.Dussault@Usherbrooke.ca. [ORCID 0000-0001-7253-7462](#).

<sup>‡</sup>Inria Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France; E-mail: Jean-Charles.Gilbert@inria.fr. [ORCID 0000-0002-0375-4663](#).

<sup>§</sup>Département de Mathématiques, Faculté des Sciences, Université de Sherbrooke, Canada; E-mail: Baptiste.Plaquevent-Jourdain@Usherbrooke.ca. Inria Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France; E-mail: Baptiste.Plaquevent-Jourdain@inria.fr. [ORCID 0000-0001-7055-4568](#).

## Table of contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Background</b>	<b>5</b>
<b>3</b>	<b>Equivalent problems</b>	<b>8</b>
3.1	B-differential of the minimum of two affine functions . . . . .	8
3.2	Linear algebra problems . . . . .	9
3.2.1	Signed feasibility of strict inequality systems . . . . .	9
3.2.2	Orthants encountered by a matrix null space . . . . .	11
3.3	Convex analysis problems . . . . .	12
3.3.1	Pointed cones by vector inversions . . . . .	12
3.3.2	Linearly separable bipartitions of a finite set . . . . .	13
3.4	Discrete geometry: hyperplane arrangement . . . . .	16
<b>4</b>	<b>Description of the B-differential</b>	<b>18</b>
4.1	Some properties of the B-differential . . . . .	18
4.2	Cardinal of the B-differential . . . . .	22
4.2.1	Winder's formula . . . . .	22
4.2.2	Bounds . . . . .	22
4.2.3	Special configurations . . . . .	26
<b>5</b>	<b>Computation of the B-differential</b>	<b>28</b>
5.1	Computation of a single element . . . . .	28
5.2	Computation of all the elements . . . . .	29
5.2.1	Overview . . . . .	29
5.2.2	Rada and Černý's algorithm . . . . .	29
5.2.3	Some improvements . . . . .	31
5.2.4	ISF algorithm . . . . .	33
5.2.5	Complexity . . . . .	33
5.2.6	Numerical results . . . . .	34
5.2.7	A numerical example . . . . .	35
<b>6</b>	<b>References</b>	<b>37</b>

# 1 Introduction

Let  $\mathbb{E}$  and  $\mathbb{F}$  be two real vector spaces of finite dimension  $n := \dim \mathbb{E}$  and  $m := \dim \mathbb{F}$ . The *B-differential at  $x \in \mathbb{E}$*  of a function  $H : \mathbb{E} \rightarrow \mathbb{F}$  is the set denoted and defined by

$$\partial_B H(x) := \{J \in \mathcal{L}(\mathbb{E}, \mathbb{F}) : H'(x_k) \rightarrow J \text{ for a sequence } \{x_k\} \subseteq \mathcal{D}_H \text{ converging to } x\}, \quad (1.1)$$

where  $\mathcal{L}(\mathbb{E}, \mathbb{F})$  is the set of linear (continuous) maps from  $\mathbb{E}$  to  $\mathbb{F}$  and  $\mathcal{D}_H$  is the set of points at which  $H$  is (Fréchet) differentiable (its derivative at  $x$  is denoted by  $H'(x)$ , an element of  $\mathcal{L}(\mathbb{E}, \mathbb{F})$ ). Recall that a locally Lipschitz continuous function is differentiable almost everywhere in the sense of the Lebesgue measure (Rademacher's theorem [28, 30, 35, 49]) and this property has the consequence that the B-differential of a locally Lipschitz function is nonempty everywhere [16]. The B-differential is an intermediate set used to define the C-differential (C for Clarke [16]) of  $H$  at  $x$ , which is denoted and defined by

$$\partial_C H(x) := \text{co } \partial_B H(x),$$

where  $\text{co } S$  is the convex hull of a set  $S$  [13, 36, 51]. Both intervene in the specification of conditions ensuring the local convergence of the semismooth Newton algorithm [46, 47].

In this paper, we are primarily interested in describing the B-differential of  $H$  at  $x$  when  $H : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is the componentwise minimum of two affine functions  $x \mapsto Ax + a$  and  $x \mapsto Bx + b$ , where  $A, B \in \mathbb{R}^{m \times n}$  and  $a, b \in \mathbb{R}^m$ . Hence,  $H$  is defined at  $x \in \mathbb{R}^n$  by

$$H(x) = \min(Ax + a, Bx + b), \quad (1.2)$$

where the minimum operator “min” acts componentwise (for two vectors  $u, v \in \mathbb{R}^m$  and  $i \in [1:m] := \{1, \dots, m\}$ :  $[\min(u, v)]_i := \min(u_i, v_i)$ ). Our motivation to look at the B-differential of that function  $H$  comes from the fact that, when  $m = n$  and  $H$  is given by (1.2), the equation

$$H(x) = 0 \quad (1.3)$$

is a reformulation (see the next paragraph for a clarification) of the *balanced [23] linear complementarity problem (LCP)*

$$0 \leq (Ax + a) \perp (Bx + b) \geq 0. \quad (1.4)$$

This system expresses the fact that a point  $x \in \mathbb{R}^n$  is sought such that  $Ax + a \geq 0$ ,  $Bx + b \geq 0$  and  $(Ax + a)^\top (Bx + b) = 0$  (the superscript “ $\top$ ” is used here and below to denote vector or matrix transposition). Problem (1.4) is a special case of the so-called (*extended*) *vertical LCP*, which uses more than two matrices and vectors in its formulation [18, 57, 61]. In the *standard LCP*,  $A$  is the identity matrix and  $a = 0$  [19, 42].

When  $Ax + a$  and  $Bx + b$  are nonnegative, the orthogonality condition  $(Ax + a)^\top (Bx + b) = 0$  in (1.4) can be written equivalently “ $(Ax + a)_i (Bx + b)_i = 0$  for all  $i \in [1:m]$ ”. Then, that  $x$  solves (1.4) if and only if (1.3) holds, because, for two real numbers  $\alpha$  and  $\beta$ ,  $\min(\alpha, \beta) = 0$  if and only if  $\alpha \geq 0$ ,  $\beta \geq 0$  and  $\alpha\beta = 0$ . Therefore, (1.3) is indeed a reformulation of (1.4) as a nonsmooth equation solving problem [1, 44]. This reformulation serves as the basis for a number of solving methods and investigations [1, 7–9, 21–23, 29, 37, 39, 43–45]. If (1.4) stands alone, it is appropriate to have  $m = n$ , but (1.4) may be part of a system with other constraints to satisfy [10, 40, 41], in which case  $m \leq n$ . In the computation of the B-differential of the Min function (1.2),  $m$  and  $n$  may not be related.

When generality does not imply difficulty, we consider the nonlinear version of the above problems. The function  $H : \mathbb{E} \rightarrow \mathbb{R}^m$  is then defined at  $x \in \mathbb{E}$  by

$$H(x) := \min(F(x), G(x)), \quad (1.5)$$

where  $F$  and  $G : \mathbb{E} \rightarrow \mathbb{R}^m$  are two functions and the “min” operator still acts componentwise. Then, the equation (1.3) is a reformulation of the complementarity problem

$$0 \leq F(x) \perp G(x) \geq 0. \quad (1.6)$$

As a first general remark, we quote that fact that the B-differential of  $H$  in (1.5) cannot be deduced from the knowledge of the B-differential of its scalar components  $H_i : x \in \mathbb{E} \rightarrow H_i(x) \in \mathbb{R}$ , for  $i \in [1 : m]$ , since if [16; proposition 2.6.2(e)]

$$\partial_B H(x) \subseteq \partial_B H_1(x) \times \cdots \times \partial_B H_m(x), \quad (1.7)$$

equality in this inclusion may not hold (see [29; § 7.1.15] and all the examples below, for which the B-differential is not complete in the sense of definition 2.3). Therefore, all the components of  $H$  must be taken into account simultaneously.

When  $H$  is given by (1.5), with functions  $F$  and  $G$  that are continuously differentiable at  $x$ , the B-differential at  $x$  of  $H$  is a finite set, made of Jacobians  $J$  whose  $i$ th row is in  $\{F'_i(x), G'_i(x)\}$  (proposition 2.2). Consequently, its cardinal can be exponential and it occurs that its full mathematical description is a tricky task, essentially when there are many indices  $i$  for which  $F_i(x) = G_i(x)$  and  $F'_i(x) \neq G'_i(x)$ , a situation that makes  $H$  nondifferentiable (lemma 2.1), and when no “qualification conditions” hold ([15; 2011] and proposition 4.2). Then, a rich panorama of configurations appears, which is barely glimpsed in this paper.

The paper starts with a background section (section 2), which recalls a basic property of the minimum of two functions (lemma 2.1) and specifies the structure of the B-differential of the function  $H$  in the general case (1.5), in particular its finite nature (proposition 2.2).

In section 3, it is shown that the problem of computing  $\partial_B H(x)$  has a rich panel of equivalent formulations, related to various areas of mathematics. We have quoted two forms of the problem in *linear algebra*, which are dual to each other (section 3.2), two equivalent problems in *convex analysis* (section 3.3) and a last equivalent problem, which arises in *computational discrete geometry* and deals with the arrangement of hyperplanes having the zero point in common (section 3.4).

Section 4 gives some properties of the B-differential of  $H$  in (1.2), which is the minimum of two affine functions, recalls Winder’s formula of its cardinal and provides some lower and upper bounds on this one.

Section 5 presents the adaptation of the Rada and Černý algorithm [48] to our framework and proposes several modifications. The complexity of the proposed algorithm is analyzed in terms of the number linear optimization problems to solve. Numerical experiments are reported, showing that the modifications improve the performance of the algorithm significantly.

An abridged version of this report can be found in [24].

## Notation

We denote by  $|S|$  the number of elements of a set  $S$  (i.e., its *cardinal*). The *power set* of a set  $S$  is denoted by  $\mathcal{P}(S)$ . The sets of nonzero natural and real numbers are denoted by  $\mathbb{N}^*$  and  $\mathbb{R}^*$ , respectively. The *sign of a real number* is the multifunction  $\text{sgn} : \mathbb{R} \multimap \mathbb{R}$  defined by  $\text{sgn}(t) = \{1\}$  if  $t > 0$ ,  $\text{sgn}(t) = \{-1\}$  if  $t < 0$  and  $\text{sgn}(0) = [-1, 1]$ . We set

$\mathbb{R}_+^n := \{x \in \mathbb{R}^n : x \geq 0\}$  and  $\mathbb{R}_{++}^n := \{x \in \mathbb{R}^n : x > 0\}$  (strict inequalities must also be understood componentwise; hence  $x > 0$  means  $x_i > 0$  for all indices  $i$ ). The vector of all one's, in a real space whose dimension is given by the context, is denoted by  $e$ . The  $i$ th row (resp. column) of a matrix  $A$  is denoted by  $A_{i,:}$  (resp.  $A_{:,i}$ ). Transposition operates after a row/column selection:  $A_{i,:}^\top$  is a short notation for the column vector  $(A_{i,:})^\top$  and  $A_{:,i}^\top$  is a short notation for the row vector  $(A_{:,i})^\top$ . For a vector  $\alpha$ ,  $\text{Diag}(\alpha)$  is the square matrix with the  $\alpha_i$ 's on its diagonal.

## 2 Background

Recall that  $F : \mathbb{E} \rightarrow \mathbb{F}$  is said to be (*Fréchet*) *differentiable* at  $x$  if  $F(x+d) = F(x) + Ld + o(\|d\|)$  for some  $L \in \mathcal{L}(\mathbb{E}, \mathbb{F})$ , in which case one denotes by  $F'(x) = L$  the *derivative* of  $F$  at  $x$ . It is said that  $F$  is *Gâteaux-differentiable* (or *G-differentiable*) at  $x$  if its *directional derivative at  $x$  along  $d \in \mathbb{E}$* , namely  $F'(x; d) := \lim_{t \downarrow 0} [F(x+td) - F(x)]/t$ , exists for all  $d \in \mathbb{E}$  and is linear in  $d$ ; this linear map is then also denoted by  $F'(x)$ . A differentiable function is G-differentiable. We say below that  $F$  is *continuously differentiable at  $x$*  if it is differentiable near  $x$  (like in [16], “near” means here and below “in a neighborhood of” in the topological sense) and if its derivative is continuous at  $x$ . If  $F$  is continuously differentiable at  $x$ , then  $\partial_B F(x) = \{F'(x)\}$ .

The next well known lemma recalls a necessary and sufficient condition guaranteeing the differentiability of the minimum of two scalar functions (see [46; 1993, final remarks (1)] and [15; 2011, theorem 2.1], for the differentiability); it will be frequently used. We give it a proof that includes the G-differentiability property.

**Lemma 2.1 (differentiability of the Min function)** *Let  $f$  and  $g : \mathbb{E} \rightarrow \mathbb{R}$  be two functions and  $h : \mathbb{E} \rightarrow \mathbb{R}$  be defined by  $h(\cdot) := \min(f(\cdot), g(\cdot))$ . Suppose that  $f$  and  $g$  are G-differentiable (resp. differentiable) at a point  $x \in \mathbb{E}$ .*

- 1) *If  $f(x) < g(x)$ , then  $h$  is G-differentiable (resp. differentiable) at  $x$  and  $h'(x) = f'(x)$ .*
- 2) *If  $f(x) = g(x)$ , then*

$$h \text{ is G-differentiable (resp. differentiable) at } x \iff f'(x) = g'(x).$$

*In this case,  $h'(x) = f'(x) = g'(x)$ .*

*Proof.* 1) This results from the fact that, when  $f$  and  $g$  are G-differentiable and  $d \in \mathbb{E}$ ,  $h(x+td) = f(x+td)$  for small  $t > 0$  and, when  $f$  and  $g$  are differentiable,  $h = f$  near  $x$ .

2) [G-differentiability] Suppose first that  $f$  and  $g$  are G-differentiable at  $x$ . Since  $f(x) = g(x)$ , one has for any  $d \in \mathbb{R}^n$ :

$$h'(x; d) = \min(f'(x)d, g'(x)d). \tag{2.1}$$

[ $\Rightarrow$ ] Since  $f$ ,  $g$  and  $h$  are G-differentiable at  $x$ ,  $d \in \mathbb{R}^n \mapsto (f(x; d), g(x; d), h'(x; d))$  is linear. Then, using (2.1):

$$h'(x; d) = -h'(x; -d) = -\min(f'(x; -d), g'(x; -d)) = \max(f'(x; d), g'(x; d)).$$

Hence  $\min(f'(x)d, g'(x)d) = \max(f'(x)d, g'(x)d)$  or  $f'(x)d = g'(x)d$ . Since  $d$  is arbitrary, it follows that  $f'(x) = g'(x) = h'(x)$ .

[ $\Leftarrow$ ] If  $f'(x) = g'(x)$ , then one has from (2.1):  $h'(x; d) = f'(x)d$  for all  $d \in \mathbb{R}^n$ . Therefore,  $h'(x; d)$  is linear in  $d$ , implying that  $h$  is G-differentiable at  $x$  and  $h'(x) = f'(x)$ .

[Differentiability] Suppose now that  $f$  and  $g$  are differentiable at  $x$ . If  $h$  is differentiable at  $x$ , it is also G-differentiable at  $x$  and, by the first part of the proof,  $h'(x) = f'(x) = g'(x)$ . Conversely, if  $f'(x) = g'(x)$ , one has for  $d \rightarrow 0$ ,

$$\begin{aligned} h(x+d) &= \min(f(x+d), g(x+d)) \\ &= \min(f(x) + f'(x)d + o(\|d\|), g(x) + g'(x)d + o(\|d\|)) \\ &= f(x) + f'(x)d + \min(o(\|d\|), o(\|d\|)) \quad [f(x) = g(x), f'(x) = g'(x)] \\ &= f(x) + f'(x)d + o(\|d\|). \end{aligned}$$

Therefore  $h$  is differentiable at  $x$  and  $h'(x) = f'(x)$ .  $\square$

In the rest of section, we assume that  $H$  is given by (1.5).

The previous lemma shows the relevance of the following index sets, when the differentiability of the function  $H$  in (1.5) is at stake:

$$\mathcal{E}(x) := \{i \in [1:m] : F_i(x) = G_i(x)\}, \quad (2.2a)$$

$$\mathcal{F}(x) := \{i \in [1:m] : F_i(x) < G_i(x)\}, \quad (2.2b)$$

$$\mathcal{G}(x) := \{i \in [1:m] : F_i(x) > G_i(x)\}. \quad (2.2c)$$

The lemma also shows that it is meaningful to distinguish the indices  $i \in \mathcal{E}(x)$  for which  $F'_i(x) = G'_i(x)$  from those for which  $F'_i(x) \neq G'_i(x)$ :

$$\mathcal{E}^=(x) := \{i \in \mathcal{E}(x) : F'_i(x) = G'_i(x)\}, \quad (2.2d)$$

$$\mathcal{E}^\neq(x) := \{i \in \mathcal{E}(x) : F'_i(x) \neq G'_i(x)\}. \quad (2.2e)$$

The next proposition describes a superset of  $\partial_B H(x)$ , which results in part from the inclusion (1.7) (see [38; 1998, §2] in a somehow different context, [20; 2000, before (8)]). This superset is defined and denoted by

$$\begin{aligned} \bar{\partial}_B H(x) := \{J \in \mathcal{L}(\mathbb{E}, \mathbb{R}^m) : & J_{i,:} = F'_i(x), \text{ if } i \in \mathcal{F}(x), \\ & J_{i,:} = F'_i(x) = G'_i(x), \text{ if } i \in \mathcal{E}^=(x), \\ & J_{i,:} \in \{F'_i(x), G'_i(x)\}, \text{ if } i \in \mathcal{E}^\neq(x), \\ & J_{i,:} = G'_i(x), \text{ if } i \in \mathcal{G}(x)\}. \end{aligned} \quad (2.3)$$

**Proposition 2.2 (superset of  $\partial_B H(x)$ )** *If  $F$  and  $G$  are continuously differentiable at  $x \in \mathbb{E}$  and  $H$  is given by (1.5), then,  $\partial_B H(x) \subseteq \bar{\partial}_B H(x)$ . In particular,  $|\partial_B H(x)| \leq 2^{|\mathcal{E}^\neq(x)|}$ .*

*Proof.* Let  $J \in \partial_B H(x)$ . By (1.7),  $J_{i,:} \in \partial_B H_i(x)$ , for any  $i \in [1:m]$ . Fix  $i \in [1:m]$ .

If  $i \in \mathcal{F}(x)$  (resp.  $i \in \mathcal{G}(x)$ ), it follows that  $F_i(x) < G_i(x)$  (resp.  $F_i(x) > G_i(x)$ ), so that  $H_i$  is differentiable at  $x$  and one has  $\partial_B H_i(x) = \{F'_i(x)\}$  (resp.  $\partial_B H_i(x) = \{G'_i(x)\}$ ). Therefore,  $J_{i,:} = F'_i(x)$  (resp.  $J_{i,:} = G'_i(x)$ ), in agreement with  $\partial_B H(x) \subseteq \bar{\partial}_B H(x)$ .

Suppose now that  $i \in \mathcal{E}(x)$ . By  $J_{i,:} \in \partial_B H_i(x)$ , there exists a sequence  $\{x_k\} \subseteq \mathcal{D}_{H_i}$  converging to  $x$  such that

$$H'_i(x_k) \rightarrow J_{i,:}. \quad (2.4a)$$

By extracting a subsequence if needed, one can assume that one of the following three properties holds for all  $k$ :

$$F_i(x_k) < G_i(x_k), \quad (2.4b)$$

$$F_i(x_k) = G_i(x_k), \quad (2.4c)$$

$$F_i(x_k) > G_i(x_k). \quad (2.4d)$$

Let us examine each of these cases.

- If (2.4b) (resp. (2.4d)) holds,  $H'_i(x_k) = F'_i(x_k)$  (resp.  $H'_i(x_k) = G'_i(x_k)$ ), so that the continuity of  $F'_i$  (resp.  $G'_i$ ) at  $x$ ,  $x_k \rightarrow x$  and (2.4a) yield  $J_{i,:} = F'_i(x)$  (resp.  $J_{i,:} = G'_i(x)$ ), in agreement with  $\partial_B H(x) \subseteq \bar{\partial}_B H(x)$ .
- Suppose now that (2.4c) holds. Since  $x_k \in \mathcal{D}_H$ , one must have  $F'_i(x_k) = G'_i(x_k)$  by lemma 2.1. Again, the continuity of  $F'_i$  and  $G'_i$  at  $x$ ,  $x_k \rightarrow x$  and (2.4a) yield  $J_{i,:} = F'_i(x) = G'_i(x)$ , in agreement with  $\partial_B H(x) \subseteq \bar{\partial}_B H(x)$ .

The last claim is a straightforward consequences of the fact that  $J_{i,:}$  can take two different values,  $F'_i(x)$  or  $G'_i(x)$ , only for the indices  $i \in \mathcal{E}^\neq(x)$ .  $\square$

The previous proposition shows that  $\partial_B H(x)$  is a finite set. It also naturally leads to the next definition, which is related to the inclusion (1.7).

**Definition 2.3 (complete B-differential)** We shall say that the B-differential of the function  $H$  given by (1.5) is *complete* if  $\partial_B H(x) = \bar{\partial}_B H(x)$  or, equivalently, if  $|\partial_B H(x)| = 2^{|\mathcal{E}^\neq(x)|}$ .  $\square$

**Definitions 2.4 (symmetry in  $\partial_B H(x)$ )** For the map  $H$  given by (1.5) and  $x \in \mathbb{E}$ , we say that the Jacobian  $\tilde{J} \in \bar{\partial}_B H(x)$  is *symmetric* to the Jacobian  $J \in \bar{\partial}_B H(x)$  if

$$\tilde{J}_{i,:} = \begin{cases} F'_i(x) & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } J_{i,:} = G'_i(x), \\ G'_i(x) & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } J_{i,:} = F'_i(x). \end{cases}$$

The B-differential  $\partial_B H(x)$  itself is said to be *symmetric* if each Jacobian  $J \in \partial_B H(x)$  has its symmetric Jacobian  $\tilde{J}$  in  $\partial_B H(x)$ .  $\square$

We shall use several times the following lemma, which, for the sake of generality, is written in a slightly more abstract formalism than the one we need below.

**Lemma 2.5 (discriminating covectors)** *Suppose that  $(\mathbb{E}, \langle \cdot, \cdot \rangle)$  is a Euclidean vector space, that  $p \in \mathbb{N}^*$  and that  $v_1, \dots, v_p$  are  $p$  distinct vectors of  $\mathbb{E}$ . Then, the set of vectors  $\xi \in \mathbb{E}$  such that  $|\{\langle \xi, v_i \rangle\}_{i \in [1:p]}| = p$  is dense in  $\mathbb{E}$ .*

*Proof.* Denote by  $\Xi$  the set of vectors  $\xi \in \mathbb{E}$  such that  $|\{\langle \xi, v_i \rangle\}_{i \in [1:p]}| = p$  (i.e.,  $\{\langle \xi, v_i \rangle\}_{i \in [1:p]}$  is formed of  $p$  distinct values in  $\mathbb{R}$ ). We have to show that  $\Xi$  is dense in  $\mathbb{E}$ .

Take  $\xi_0 \notin \Xi$ , so that  $\langle \xi_0, v_i \rangle = \langle \xi_0, v_j \rangle$  for some  $i \neq j$  in  $[1:p]$ . By continuity of the scalar product, for any  $\varepsilon_0 > 0$  sufficiently small, the vector  $\xi_1 := \xi_0 - \varepsilon_0(v_i - v_j)$  guarantees

$$\langle \xi_1, v_{i_1} \rangle < \langle \xi_1, v_{i_2} \rangle$$



for all  $i_1$  and  $i_2 \in [1:p]$  such that  $\langle \xi_0, v_{i_1} \rangle < \langle \xi_0, v_{i_2} \rangle$  (in other words,  $\xi_1$  maintains strict the inequalities that are strict with  $\xi_0$ ). In addition

$$\langle \xi_1, v_i \rangle - \langle \xi_1, v_j \rangle = \underbrace{\langle \xi_0, v_i - v_j \rangle}_{=0} - \underbrace{\varepsilon_0 \|v_i - v_j\|^2}_{>0} < 0.$$

Therefore, one gets one more strict inequality with  $\xi_1$  than with  $\xi_0$ . Pursuing like this, one can finally obtain a vector  $\xi$  in  $\Xi$ . This vector is arbitrarily close to  $\xi_0$  by taking the  $\varepsilon_i$ 's positive and sufficiently small. The density result follows.  $\square$

### 3 Equivalent problems

The problem of determining the B-differential of the piecewise affine function, that is the minimum (1.2) of two *affine* functions, appears in various contexts, sometimes with non straight-forward connections with this original problem (this one is recalled in section 3.1). We review some of these problems in this section (for other formulations or connections to the present problem, see [5, 6, 59] and the references therein) and give a few properties of the B-differential in this affine case. As suggested by proposition 2.2, these problems have an enumeration nature, since a finite list of mathematical objects has to be determined. This list may have an exponential number of elements, which makes its contents difficult to specify (to this respect, the particular case where the B-differential is complete is a trivial exception). Some problems, such as the one related to the arrangement of subspaces containing the origin (section 3.4), have been extensively explored, others much less. Each formulation sheds a particular light on the problem and is therefore, as such, interesting to mention and keep in mind. It also offers the possibility of introducing new algorithmic approaches to describe the B-differential.

#### 3.1 B-differential of the minimum of two affine functions

The problem of this section was already presented in the introduction and is sometimes referred to as the *original problem* below.

**Problem 3.1 (B-differential of the minimum of two affine functions)** Let be given two positive integers  $n$  and  $m \in \mathbb{N}^*$ , two matrices  $A, B \in \mathbb{R}^{m \times n}$  and two vectors  $a, b \in \mathbb{R}^m$ . It is requested to compute the B-differential at some  $x \in \mathbb{R}^n$  of the function  $H : \mathbb{R}^n \rightarrow \mathbb{R}^m$  defined by (1.2).  $\square$

Since there is no functions  $F$  and  $G$  in the affine case, the index sets in (2.2) are mnemonically renamed as follows:

$$\mathcal{E}(x) := \{i \in [1:m] : (Ax + a)_i = (Bx + b)_i\}, \quad (3.1a)$$

$$\mathcal{A}(x) := \{i \in [1:m] : (Ax + a)_i < (Bx + b)_i\}, \quad (3.1b)$$

$$\mathcal{B}(x) := \{i \in [1:m] : (Ax + a)_i > (Bx + b)_i\}, \quad (3.1c)$$

$$\mathcal{E}^=(x) := \{i \in \mathcal{E}(x) : A_{i,:} = B_{i,:}\}, \quad (3.1d)$$

$$\mathcal{E}^{\neq}(x) := \{i \in \mathcal{E}(x) : A_{i,:} \neq B_{i,:}\}. \quad (3.1e)$$

When  $\mathcal{E}^{\neq}(x) \neq \emptyset$ , the rows of  $B - A$  with indices in  $\mathcal{E}^{\neq}(x)$  will play a key role below. We denote its transpose by

$$V := (B - A)_{\mathcal{E}^{\neq}(x),:}^T \in \mathbb{R}^{n \times |\mathcal{E}^{\neq}(x)|}. \quad (3.2)$$

Note that, due to their indices in  $\mathcal{E}^{\neq}(x)$  and the definition of this index set, the columns of  $V$  are nonzero. This matrix may not have full rank, however.

## 3.2 Linear algebra problems

### 3.2.1 Signed feasibility of strict inequality systems

Many proofs below use the equivalence between the original problem 3.1 and the following one. The reason is that working on problem 3.2 often allows us to propose shorter proofs.

**Problem 3.2 (signed feasibility of strict inequality systems)** Let be given two positive integers  $n$  and  $p \in \mathbb{N}^*$  and a matrix  $V$  in  $\mathbb{R}^{n \times p}$  with nonzero columns. It is requested to determine the set

$$\mathcal{S} := \{s \in \{\pm 1\}^p : s \cdot V^\top d > 0 \text{ is feasible for } d \in \mathbb{R}^n\}, \quad (3.3)$$

where the binary operator “ $\cdot$ ” denotes the Hadamard product of vectors (for  $u$  and  $v \in \mathbb{R}^p$ ,  $u \cdot v \in \mathbb{R}^p$  is the vector of  $\mathbb{R}^p$ , whose  $i$ th component is  $u_i v_i$ ).  $\square$

To make the link between this problem and the B-differential problem 3.1, let us start by adapting the definition (2.3) of the superset  $\bar{\partial}_B H(x)$  of the B-differential  $\partial_B H(x)$  to the context of section 3.1:

$$\begin{aligned} \bar{\partial}_B H(x) := \{J \in \mathbb{R}^{m \times n} : & J_{i,:} = A_{i,:}, \text{ if } i \in \mathcal{A}(x), \\ & J_{i,:} = A_{i,:} = B_{i,:}, \text{ if } i \in \mathcal{E}^-(x), \\ & J_{i,:} \in \{A_{i,:}, B_{i,:}\}, \text{ if } i \in \mathcal{E}^\neq(x), \\ & J_{i,:} = B_{i,:}, \text{ if } i \in \mathcal{B}(x)\}. \end{aligned} \quad (3.4)$$

In the correspondence between problems 3.1 and 3.2, the matrix  $V$  in (3.3) is the one in (3.2), so that  $p := |\mathcal{E}^\neq(x)|$ , an integer that is supposed to be nonzero. In this correspondence, we also assume that the indices labelling the components of the sign vectors  $s \in \{\pm 1\}^{|\mathcal{E}^\neq(x)|}$  and the columns of  $V \in \mathbb{R}^{n \times |\mathcal{E}^\neq(x)|}$  are those of  $\mathcal{E}^\neq(x)$ . The link between the two problems is established by the following map

$$\sigma : J \in \bar{\partial}_B H(x) \mapsto s \in \{\pm 1\}^{|\mathcal{E}^\neq(x)|}, \text{ where } s_i = \begin{cases} +1 & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } J_{i,:} = A_{i,:}, \\ -1 & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } J_{i,:} = B_{i,:}. \end{cases} \quad (3.5a)$$

This is a bijection since two Jacobians in  $\bar{\partial}_B H(x)$  only differ by their rows with index in  $\mathcal{E}^\neq(x)$  and that these rows can take any of the values  $A_{i,:}$  or  $B_{i,:}$ . Actually, the reverse map of  $\sigma$  is

$$\sigma^{-1} : s \in \{\pm 1\}^{|\mathcal{E}^\neq(x)|} \mapsto J \in \bar{\partial}_B H(x), \text{ where } J_{i,:} = \begin{cases} A_{i,:} & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } s_i = +1, \\ B_{i,:} & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } s_i = -1. \end{cases} \quad (3.5b)$$

The question that arises now is whether  $\sigma$  is still a bijection between  $\partial_B H(x)$  and  $\mathcal{S}$ .

**Proposition 3.3 (bijection  $\partial_B H(x) \leftrightarrow \mathcal{S}$ )** *Let  $H : \mathbb{R}^n \rightarrow \mathbb{R}^m$  be given by (1.2),  $x$  be a point in  $\mathbb{R}^n$  such that  $\mathcal{E}^\neq(x) \neq \emptyset$  and  $V$  be given by (3.2). Then, the map  $\sigma$  is a bijection from  $\partial_B H(x)$  onto  $\mathcal{S}$ . In particular, the following properties hold.*

- 1) *If  $J \in \partial_B H(x)$ , then  $\exists d \in \mathbb{R}^n$  such that  $\sigma(J) \cdot V^\top d > 0$ .*
- 2) *If  $s \in \{\pm 1\}^{|\mathcal{E}^\neq(x)|}$  and  $\exists d \in \mathbb{R}^n$  is such that  $s \cdot V^\top d > 0$ , then  $\sigma^{-1}(s) \in \partial_B H(x)$ .*
- 3) *Let  $J \in \bar{\partial}_B H(x)$ . Then,  $J \in \partial_B H(x) \iff \sigma(J) \cdot V^\top d > 0$  is feasible for  $d \in \mathbb{R}^n$ .*

*Proof.* The properties 1, 2 and 3 in the statement of the proposition are straightforward consequences of the bijectivity of  $\sigma : \partial_B H(x) \rightarrow \mathcal{S}$ . Now, the discussion before the proposition

has shown that  $\sigma : \bar{\partial}_B H(x) \mapsto \{\pm 1\}^{|\mathcal{E}^\neq(x)|}$  is a bijection. Therefore,  $\sigma : \partial_B H(x) \mapsto \{\pm 1\}^{|\mathcal{E}^\neq(x)|}$  is injective and it suffices to prove that

$$\sigma(\partial_B H(x)) = \mathcal{S}. \quad (3.6a)$$

[ $\subseteq$ ] Let  $J \in \partial_B H(x)$ . Then, there exists a sequence  $\{x_k\} \subseteq \mathcal{D}_H$  converging to  $x$  such that

$$H'(x_k) \rightarrow J. \quad (3.6b)$$

For  $i \in \mathcal{E}^\neq(x)$ , one cannot have  $(Ax_k + a)_i = (Bx_k + b)_i$ , since  $A_{i,:} \neq B_{i,:}$  would imply that  $x_k \notin \mathcal{D}_H$  (lemma 2.1). Therefore, one can find a subsequence  $\mathcal{K}$  of indices  $k$  and a partition  $(\mathcal{A}_0, \mathcal{B}_0)$  of  $\mathcal{E}^\neq(x)$  such that for all  $k \in \mathcal{K}$ :

$$(Ax_k + a)_{\mathcal{A}_0} < (Bx_k + b)_{\mathcal{A}_0}, \quad (3.6c)$$

$$(Ax_k + a)_{\mathcal{B}_0} > (Bx_k + b)_{\mathcal{B}_0}. \quad (3.6d)$$

Now, fix  $k \in \mathcal{K}$ , set  $d := x_k - x$ . Since  $(Ax + a)_i = (Bx + b)_i$  for  $i \in \mathcal{E}^\neq(x)$ , one deduces from (3.6c) and (3.6d) that

$$(B - A)_{\mathcal{A}_0} d > 0 \quad \text{and} \quad (B - A)_{\mathcal{B}_0} d < 0.$$

Recalling the definitions of  $V$  in (3.2) and of  $\mathcal{S}$  in (3.3), we see that, to conclude the proof of the membership  $\sigma(J) \in \mathcal{S}$ , it suffices to show that  $[\sigma(J)]_{\mathcal{A}_0} = +1$  and  $[\sigma(J)]_{\mathcal{B}_0} = -1$ , or equivalently, by the definition of  $\sigma$ ,  $(J_{i,:} = A_{i,:}$  for  $i \in \mathcal{A}_0$ ) and  $(J_{i,:} = B_{i,:}$  for  $i \in \mathcal{B}_0$ ). This is indeed the case, since by (3.6c) and (3.6d), for all  $k \in \mathcal{K}$ , one has  $(H'_{i,:}(x_k) = A_{i,:}$  for  $i \in \mathcal{A}_0$ ) and  $(H'_{i,:}(x_k) = B_{i,:}$  for  $i \in \mathcal{B}_0$ ); now, use the convergence (3.6b) to conclude.

[ $\supseteq$ ] Let  $s \in \mathcal{S}$ . We have to find a  $J \in \partial_B H(x)$  such that  $\sigma(J) = s$ . Since  $\sigma(J)$  only depends on  $J_{\mathcal{E}^\neq(x),:}$ , we only have to determine the rows  $J_{i,:}$  of  $J$  for  $i \in \mathcal{E}^\neq(x)$ .

Since  $s \in \mathcal{S}$ , there is a  $d \in \mathbb{R}^n$  such that

$$s \cdot V^\top d > 0. \quad (3.6e)$$

Define the sequence  $\{x_k\} \subseteq \mathbb{R}^n$  by

$$x_k := x + t_k d + \zeta_k,$$

where  $t_k \downarrow 0$ , while the  $\zeta_k$ 's are (small) perturbation vectors in  $\mathbb{R}^n$  such that  $\zeta_k = o(t_k)$  and  $x_k \in \mathcal{D}_H$  (this is possible by Rademacher's theorem). Then  $x_k \rightarrow x$ . We claim that  $H'(x_k)$  is an appropriate constant matrix  $J$  for  $k$  sufficiently large. Let us examine the row  $i$  of  $H'(x_k)$ .

- If  $i \in \mathcal{A}(x)$ ,  $(Ax_k + a)_i < (Bx_k + b)_i$  for  $k$  large, so that  $H'_i(x_k) = A_{i,:}$ , which converges.
- If  $i \in \mathcal{B}(x)$ ,  $(Ax_k + a)_i > (Bx_k + b)_i$  for  $k$  large, so that  $H'_i(x_k) = B_{i,:}$ , which converges.
- If  $i \in \mathcal{E}^=(x)$ ,  $H'_i(x_k) \in \{A_{i,:}, B_{i,:}\}$ , by the differentiability of  $H$  at  $x_k$  (by construction,  $x_k \in \mathcal{D}_H$ ). Since  $A_{i,:} = B_{i,:}$  for  $i \in \mathcal{E}^=(x)$ , one has  $H'_i(x_k) = A_{i,:} = B_{i,:}$ , which converges.
- If  $i \in \mathcal{E}^\neq(x)$ , one uses

$$(Ax_k + a)_i = (Ax + a)_i + t_k A_{i,:} d + o(t_k),$$

$$(Bx_k + b)_i = (Bx + b)_i + t_k B_{i,:} d + o(t_k).$$

Subtracting side by side and using  $(Ax + a)_i = (Bx + b)_i$  yield

$$(Bx_k + b)_i - (Ax_k + a)_i = t_k (B_{i,:} - A_{i,:}) d + o(t_k).$$

Then, using (3.6e) and the definition (3.2) of  $V$ , one gets for  $k$  sufficiently large:

- if  $s_i = +1$ , then  $(Bx_k + b)_i > (Ax_k + b)_i$  for  $k$  large, so that  $H'_i(x_k) = A_{i,:}$ ,
- if  $s_i = -1$ , then  $(Bx_k + b)_i < (Ax_k + b)_i$  for  $k$  large, so that  $H'_i(x_k) = B_{i,:}$ .

We have shown that  $H'(x_k) = J$  for  $k$  sufficiently large, where

$$J_{i,:} = \begin{cases} A_{i,:} & \text{if } i \in \mathcal{A}(x), \\ A_{i,:} = B_{i,:} & \text{if } i \in \mathcal{E}^=(x), \\ A_{i,:} & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } s_i = +1, \\ B_{i,:} & \text{if } i \in \mathcal{E}^\neq(x) \text{ and } s_i = -1, \\ B_{i,:} & \text{if } i \in \mathcal{B}(x). \end{cases}$$

Since, by (3.5),  $\sigma(J) = s$ , the proof is concluded.  $\square$

**Equivalence 3.4 (B-differential  $\leftrightarrow$  signed linear system feasibility)** The equivalence between the original problem 3.1 and the signed linear system feasibility problem 3.2 is a consequence of the previous proposition with  $V$  given by (3.2), which shows the bijectivity of the map  $\sigma : \partial_B H(x) \rightarrow \mathcal{S}$  defined by (3.5a). Therefore, knowing  $\sigma$  by its definition (3.5), determining  $\partial_B H(x)$  or  $\mathcal{S}$  are equivalent problems.  $\square$

### 3.2.2 Orthants encountered by a matrix null space

Recall the definition of  $\mathcal{S}$  in (3.3), which is associated with some matrix  $V \in \mathbb{R}^{n \times p}$  with nonzero columns. The equivalent form of problem 3.2 (hence of problem 3.1) introduced in this section is based on a bijection between the *complementary set* of  $\mathcal{S}$  in  $\{\pm 1\}^p$ , denoted  $\mathcal{S}^c := \{\pm 1\}^p \setminus \mathcal{S}$ , and a collection  $\mathcal{I}$  of subintervals of  $[1:p]$  (hence  $\mathcal{I} \subseteq \mathcal{P}([1:p])$ ), which refers to a collection of orthants of  $\mathbb{R}^p$ , those encountered by the null space of  $V$ . Therefore, this equivalence could make it possible to introduce an algorithm that describes the *complementary set* of  $\partial_B H(x)$ , which is interesting when  $|\overline{\partial_B H(x)} \setminus \partial_B H(x)|$  is small. The equivalence has a dual nature in the sense that it is based on Gordan's alternative.

**Problem 3.5 (orthants encountered by a matrix null space)** Let be given two positive integers  $n$  and  $p \in \mathbb{N}^*$  and a matrix  $V$  in  $\mathbb{R}^{n \times p}$  with nonzero columns. Associate with  $I \subseteq [1:p]$  the following orthant of  $\mathbb{R}^p$  :

$$\mathcal{O}_I^p := \{y \in \mathbb{R}^p : y_I \geq 0, y_{I^c} \leq 0\},$$

where  $I^c := [1:p] \setminus I$ . It is requested to determine the set

$$\mathcal{I} := \{I \subseteq [1:p] : \mathcal{N}(V) \cap \mathcal{O}_I^p \neq \{0\}\}. \quad \square$$

Note that, if  $I \in \mathcal{I}$ , then  $I^c \in \mathcal{I}$  (because  $y \in (\mathcal{N}(V) \cap \mathcal{O}_I^p) \setminus \{0\}$  implies that  $-y \in (\mathcal{N}(V) \cap \mathcal{O}_{I^c}^p) \setminus \{0\}$ ), so that  $|\mathcal{I}|$  is even, just like  $|\mathcal{S}|$  and  $|\mathcal{S}^c|$  (see proposition 4.1).

The equivalence between problems 3.2 and 3.5 is obtained thanks to the following bijection

$$\iota : s \in \{\pm 1\}^p \rightarrow \iota(s) := \{i \in [1:p] : s_i = +1\} \in \mathcal{P}([1:p]), \quad (3.7)$$

whose reverse map is  $\iota^{-1} : I \in \mathcal{P}([1:p]) \rightarrow s \in \{\pm 1\}^p$ , where  $s_i = +1$  if  $i \in I$  and  $s_i = -1$  if  $i \notin I$ . As announced above, this equivalence relies on Gordan's theorem of the alternative [32; 1873]: for a matrix  $A \in \mathbb{R}^{m \times n}$ ,

$$\boxed{\exists x \in \mathbb{R}^n : Ax > 0} \iff \nexists \alpha \in \mathbb{R}_+^m \setminus \{0\} : A^T \alpha = 0. \quad (3.8)$$

More precisely, the equivalence between problems 3.2 and 3.5 is based on the following bijection, which may be viewed as an interpretation of Gordan's alternative.

**Proposition 3.6 (bijection  $\mathcal{S}^c \leftrightarrow \mathcal{I}$ )** *The map  $\iota$  defined by (3.7) is a bijection from  $\mathcal{S}^c$  onto  $\mathcal{I}$ .*

*Proof.* Let  $s \in \{\pm 1\}^p$  and set  $I := \iota(s) = \{i \in [1:p] : s_i = +1\}$ . Define  $A := \text{Diag}(s)V^T$  to make the link with Gordan's alternative (3.8). One has the equivalences

$$\begin{aligned}
s \in \mathcal{S}^c &\iff \nexists x \in \mathbb{R}^n : Ax > 0 && \text{[definition of } \mathcal{S} \text{ in (3.3)]} \\
&\iff \exists \alpha \in \mathbb{R}_+^m \setminus \{0\} : A^T \alpha = 0 && \text{[Gordan's alternative (3.8)]} \\
&\iff \exists \alpha \in \mathbb{R}_+^m \setminus \{0\} : s \cdot \alpha \in \mathcal{N}(V) \\
&\iff \mathcal{N}(V) \cap \mathcal{O}_I^p \neq \{0\} && \text{[see below]} \\
&\iff I \in \mathcal{I} && \text{[definition of } \mathcal{I}\text{].}
\end{aligned} \tag{3.9}$$

The implication “ $\Rightarrow$ ” in (3.9) is due to the fact that  $s \cdot \alpha$  is nonzero and belongs to both  $\mathcal{N}(V)$  and  $\mathcal{O}_I^p$ . The reverse implication “ $\Leftarrow$ ” in (3.9) is due to the fact that there is a nonzero  $y \in \mathcal{N}(V) \cap \mathcal{O}_I^p$ , implying that  $\alpha := s \cdot y$  is nonzero and  $\geq 0$  and is such that  $s \cdot \alpha = y \in \mathcal{N}(V)$ .

Since  $\iota : \{\pm 1\}^p \rightarrow \mathcal{P}([1:p])$  is a bijection, the above equivalences show that  $\iota$  is also a bijection from  $\mathcal{S}^c$  onto  $\mathcal{I}$ .  $\square$

**Equivalence 3.7 ( $\mathcal{S}^c \leftrightarrow \mathcal{I}$ )** The equivalence between problems 3.2 and 3.5 is a direct consequence of the bijectivity of  $\iota : \mathcal{S}^c \rightarrow \mathcal{I}$ , established in proposition 3.6: to determine  $\mathcal{S}$ , it suffices to determine  $\mathcal{S}^c = \iota^{-1}(\mathcal{I})$ , hence to determine  $\mathcal{I}$ .  $\square$

### 3.3 Convex analysis problems

The formulation of the original problem 3.1 in the form of the convex analysis problems 3.8 and 3.11 below may be useful to highlight some properties of  $\partial_B H(x)$ , thanks to the tools of this discipline. We take this point of view to introduce the notion of extremality (definition 4.6), to propose another proof of proposition 4.5 and to prove (4.11).

#### 3.3.1 Pointed cones by vector inversions

Recall that a *convex cone*  $K$  of  $\mathbb{R}^n$  is a convex set verifying  $\mathbb{R}_{++}K \subseteq K$  (or, more explicitly,  $tx \in K$  when  $t > 0$  and  $x \in K$ ). A *closed* convex cone  $K$  is said to be *pointed* if  $K \cap (-K) = \{0\}$  [13; p. 54], which amounts to saying that  $K$  does not contain a line (i.e., an affine subspace of dimension one) or that  $K$  has no nonzero direction  $z$  such that  $-z \in K$ . For  $P \subseteq \mathbb{R}^n$ , we also denote by “cone  $P$ ” the smallest *convex* cone containing  $P$ .

**Problem 3.8 (pointed cones by vector inversions)** Let be given two positive integers  $n$  and  $p \in \mathbb{N}^*$  and  $p$  nonzero vectors  $v_1, \dots, v_p \in \mathbb{R}^n$ . It is requested to determine all the sign vectors  $s \in \{\pm 1\}^p$  such that  $\text{cone}\{s_i v_i : i \in [1:p]\}$  is a pointed cone.  $\square$

The equivalence between the original problem 3.1 and this problem 3.8 is obtained thanks to the next proposition, which gives another property (“cone pointedness”) that is equivalent to those in (3.8) and that is adapted to the present concern.

**Proposition 3.9 (pointed polyhedral cone)** *Let be given a finite collection of nonzero vectors  $\{v_i : i \in [1:p]\} \subseteq \mathbb{R}^n$ . Then, the following properties are equivalent:*

- (i)  $\text{cone}\{v_i : i \in [1:p]\}$  is pointed,
- (ii)  $\nexists \alpha \in \mathbb{R}_+^p \setminus \{0\} : \sum_{i \in [1:p]} \alpha_i v_i = 0$ ,
- (iii)  $\exists d \in \mathbb{R}^n, \forall i \in [1:p] : v_i^\top d > 0$ .

*Proof.* The equivalence (ii)  $\Leftrightarrow$  (iii) follows directly from Gordan's alternative (3.8) (with  $A = V^\top$ ), so that it remains to prove (i)  $\Leftrightarrow$  (ii). Set  $K := \text{cone}\{v_i : i \in [1:p]\}$ .

[(i)  $\Rightarrow$  (ii)] One can assume that  $p \geq 2$ , since when  $p = 1$ , both (i) and (ii) hold. We prove the contrapositive. Assume that there is an  $\alpha \in \mathbb{R}_+^p \setminus \{0\}$  such that  $\sum_{i \in [1:p]} \alpha_i v_i = 0$ . Without loss of generality, one can assume that  $\alpha_1 \neq 0$ . Set  $z := \alpha_1 v_1$  which is in  $K \setminus \{0\}$ . One also has  $-z = \sum_{i \in [2:p]} \alpha_i v_i$ , showing that  $-z \in K$ . Hence,  $K$  is not pointed.

[(ii)  $\Rightarrow$  (i)] We prove the contrapositive. If  $K$  is not pointed, there exists a nonzero vector  $z \in K \cap (-K)$ . Therefore,

$$z = \sum_{i \in [1:p]} \alpha'_i v_i \quad \text{and} \quad -z = \sum_{i \in [1:p]} \alpha''_i v_i$$

for some  $\alpha'$  and  $\alpha'' \in \mathbb{R}_+^p \setminus \{0\}$ . Adding the two identities side by side, we get  $\sum_{i \in [1:p]} \alpha_i v_i = 0$ , with  $\alpha := \alpha' + \alpha''$ . Since  $\alpha \in \mathbb{R}_+^p \setminus \{0\}$ , this contradicts (ii).  $\square$

**Equivalence 3.10 (signed linear system feasibility  $\leftrightarrow$  pointed cone by vector inversion)** The equivalence (i)  $\Leftrightarrow$  (iii) of the previous proposition shows that the set  $\mathcal{S}$  defined by (3.3) is also given by

$$\mathcal{S} = \{s \in \{\pm 1\}^p : \text{cone}\{s_i v_i : i \in [1:p]\} \text{ is pointed}\}. \quad (3.10)$$

To put it in words, supposing that  $p := |\mathcal{E}^\neq(x)|$  is nonzero and denoting by  $v_1, \dots, v_p$  the columns of the matrix  $V$  defined by (3.2), the original problem of section 3.1 is equivalent to problem 3.8.  $\square$

### 3.3.2 Linearly separable bipartitions of a finite set

This section extends section 3.3.1 and adopts its concepts and notation. The point of view presented in this section was also considered by Zaslavsky [60; 1975, § 6A].

**Problem 3.11 (linearly separable bipartitioning)** Let  $p \geq 1$  and  $n \geq 2$  be two integers. Let be given an affine space  $\mathbb{A}$  of dimension  $n - 1$  in  $\mathbb{R}^n$  and  $p$  vectors  $\bar{v}_1, \dots, \bar{v}_p \in \mathbb{A}$  ( $\bar{v}_i$  is the "normalized"  $v_i$  appearing in the other problems). Let  $\mathbb{A}_0 := \mathbb{A} - \mathbb{A}$  be the vector space parallel to  $\mathbb{A}$ . It is requested to find all the bipartitions (i.e., the partitions made of two subsets)  $(I, J)$  of  $[1:p]$  for which there exists a vector  $\xi \in \mathbb{A}_0$  (also called *separating covector* below) such that

$$\forall i \in I, \forall j \in J : \quad \xi^\top \bar{v}_i < \xi^\top \bar{v}_j. \quad \square$$

Of course, if  $(I, J)$  is an appropriate partition to which a separating covector  $\xi$  corresponds, then  $(J, I)$  is also an appropriate partition with separating covector  $-\xi$ . Therefore, only half of the appropriate partitions  $(I, J)$  must be identified, a fact that is related to the symmetry

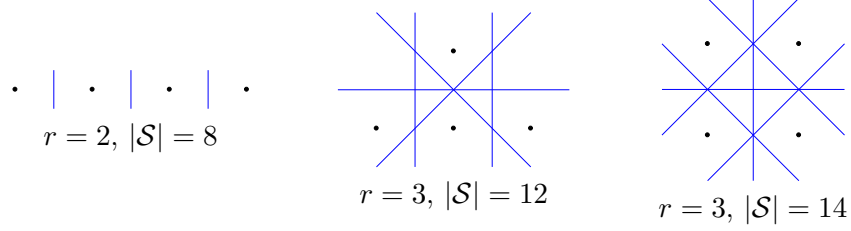


Figure 3.1: Linearly separable bipartitions of a set of  $p = 4$  points  $\bar{v}_i$  in  $\mathbb{R}^2$  (the dots in the figure). Possible separating hyperplanes are the drawn lines. We have not represented any separating line associated with the partition  $(\emptyset, [1:p])$  or  $([1:p], \emptyset)$ , so that  $|\mathcal{S}| = 2(n_s + 1)$ , where  $n_s$  is the number of represented separating lines. We have set  $r := \dim(\text{vect}\{\bar{v}_1, \dots, \bar{v}_p\}) + 1$ .

of  $\partial_B H(x)$  (proposition 4.1). Figure 3.1 shows the solution to this problem by drawing the separating hyperplanes  $\{\bar{v} \in \mathbb{A} : \xi^\top \bar{v} = \alpha\}$  corresponding to some separating covector  $\xi$  and some  $\alpha \in \mathbb{R}$ , for three examples with  $n - 1 = 2$  and  $p = 4$ . Since it will be shown that  $|\mathcal{S}|$  is the number of these searched linearly separable bipartitions, this one is denoted that way in the figure. Obviously,  $|\mathcal{S}|$  not only depends on  $p$  and  $r$ , but it also depends on the arrangement of the  $\bar{v}_i$ 's in the affine space  $\mathbb{A}$ . We also see that  $|\mathcal{S}|$  cannot take all the even values (proposition 4.1) between its lower bound  $2p = 8$  and its upper bound 14 given by propositions 4.8 and 4.11.

The equivalence between the linearly separable bipartitioning problem 3.11 of this section and the vector inversion problem 3.8 (hence, with the original problem 3.1) is grounded on the following proposition.

**Proposition 3.12 (pointed cone after vector inversions)** *Suppose that  $p$  is an integer  $\geq 2$ , that  $\{v_k\}_{k \in [1:p]} \subseteq \mathbb{R}^n \setminus \{0\}$  and that  $K := \text{cone}\{v_k : k \in [1:p]\}$  is a pointed cone, in which case there is a  $d \in \mathbb{R}^n$  such that*

$$\|d\| = 1 \quad \text{and} \quad (\forall k \in [1:p] : v_k^\top d > 0).$$

For all  $k \in [1:p]$ , set

$$\bar{v}_k := \frac{v_k}{v_k^\top d}.$$

Let  $(I, J)$  be a partition of  $[1:p]$  with nonempty index sets  $I$  and  $J$ . Define

$$\begin{aligned} \mathbb{A} &:= \{v \in \mathbb{R}^n : v^\top d = 1\}, & \mathbb{A}_0 &:= \{v \in \mathbb{R}^n : v^\top d = 0\}, \\ K_I &:= \text{cone}\{v_i : i \in I\}, & K_J &:= \text{cone}\{v_j : j \in J\}, \\ C_I &:= K_I \cap \mathbb{A} & \text{and} & C_J := K_J \cap \mathbb{A}. \end{aligned}$$

Then, the following properties are equivalent:

- (i)  $\text{cone}((-K_I) \cup K_J)$  is pointed,
- (ii)  $K_I \cap K_J = \{0\}$ ,
- (iii)  $C_I \cap C_J = \emptyset$ ,
- (iv) there exists a vector  $\xi \in \mathbb{A}_0$  such that  $\max_{i \in I} \xi^\top \bar{v}_i < \min_{j \in J} \xi^\top \bar{v}_j$ .



*Proof.* By proposition 3.9, since  $K$  is pointed, there is indeed a vector  $d \in \mathbb{R}^n$ , such that  $v_k^\top d > 0$  for  $k \in [1:p]$ . It can be supposed that  $d$  has unit norm.

[(i)  $\Rightarrow$  (ii)] We show the contrapositive. If there is a nonzero vector  $v \in K_I \cap K_J$ , then

$$-v \in (-K_I) \subseteq \text{cone}((-K_I) \cup K_J) \quad \text{and} \quad v \in K_J \subseteq \text{cone}((-K_I) \cup K_J).$$

Therefore,  $\text{cone}((-K_I) \cup K_J)$  is not pointed.

[(ii)  $\Rightarrow$  (iii)] One has  $\emptyset = \mathbb{A} \cap \{0\} = \mathbb{A} \cap K_I \cap K_J$  [(ii)] =  $(\mathbb{A} \cap K_I) \cap (\mathbb{A} \cap K_J) = C_I \cap C_J$ .

[(iii)  $\Rightarrow$  (iv)] We claim that

$C_I$  is nonempty, convex and compact.

Indeed, since  $C_I$  is nonempty (it contains the vectors  $\bar{v}_i$  for  $i \in I \neq \emptyset$ ), convex (because  $K_I$  and  $\mathbb{A}$  are convex) and closed (because  $K_I$  and  $\mathbb{A}$  are closed), it suffices to show that  $C_I$  is bounded or that its asymptotic cone (or recession cone in [51; p. 61]), namely  $C_I^\infty = K_I \cap \mathbb{A}_0$ , is reduced to  $\{0\}$  [51; theorem 8.4]. This is indeed the case since  $v^\top d > 0$  for all  $v \in K_I \setminus \{0\} \subseteq K \setminus \{0\}$ . For the same reason,

$C_J$  is nonempty, convex and compact.

Now, since  $C_I \cap C_J = \emptyset$  by (iii), one can strictly separate the convex sets  $C_I$  and  $C_J$  in  $\mathbb{A}$  [51; corollary 11.4.2]: there exists  $\xi \in \mathbb{A}_0$  such that  $\xi^\top v < \xi^\top w$ , for all  $v \in C_I$  and all  $w \in C_J$ . This shows that (iv) holds.

[(iv)  $\Rightarrow$  (i)] Since  $\text{cone}((-K_I) \cup K_J) = \text{cone}(\{-v_i : i \in I\} \cup \{v_j : j \in J\})$ , by proposition 3.9, it suffices to find  $d_{(I,J)} \in \mathbb{R}^n$  such that

$$\left(-v_i^\top d_{(I,J)} > 0, \quad \forall i \in I\right) \quad \text{and} \quad \left(v_j^\top d_{(I,J)} > 0, \quad \forall j \in J\right). \quad (3.11)$$

By (iv) and the fact that  $\theta \in (0, \pi) \rightarrow \cot \theta \in \mathbb{R}$  is surjective, one can determine  $\theta \in (0, \pi)$  such that

$$\max_{i \in I} \frac{\xi^\top v_i}{v_i^\top d} < -\cot \theta < \min_{j \in J} \frac{\xi^\top v_j}{v_j^\top d}. \quad (3.12)$$

Since  $\sin \theta > 0$  for  $\theta \in (0, \pi)$  and since  $v_k^\top d > 0$  for all  $k \in [1:p]$ , this is equivalent to

$$\max_{i \in I} v_i^\top [(\cos \theta)d + (\sin \theta)\xi] < 0 < \min_{j \in J} v_j^\top [(\cos \theta)d + (\sin \theta)\xi].$$

Therefore, (3.11) is satisfied with  $d_{(I,J)} := (\cos \theta)d + (\sin \theta)\xi$ . □

The previous proposition has introduced the following construction, which will be used several times below.

**Construction 3.13** 1) Let be given two integers  $n$  and  $p \in \mathbb{N}^*$ , the nonzero vectors  $v_1, \dots, v_p \in \mathbb{R}^n$  and  $s \in \mathcal{S}$ . By the definition (3.3) of  $\mathcal{S}$ , there is a vector  $d \in \mathbb{R}^n$  such that  $s_i v_i^\top d > 0$  for all  $i \in [1:p]$ . Then, define

$$\begin{aligned} \mathbb{A} &:= \{\bar{v} \in \mathbb{R}^n : d^\top \bar{v} = 1\}, & \mathbb{A}_0 &:= \mathbb{A} - \mathbb{A} = \{\bar{v} \in \mathbb{R}^n : d^\top \bar{v} = 0\}, \\ \mathcal{P} &:= \{\bar{v}_i := v_i / (v_i^\top d) = s_i v_i / (s_i v_i^\top d) : i \in [1:p]\} \subseteq \mathbb{A}. \end{aligned}$$



2) For a given partition  $(I, J) \in \mathcal{P}([1:p])$ , define

$$K_I := \text{cone}\{s_i v_i : i \in I\} \quad \text{and} \quad K_J := \text{cone}\{s_j v_j : j \in J\}, \quad (3.13a)$$

$$C_I := K_I \cap \mathbb{A} \quad \text{and} \quad C_J := K_J \cap \mathbb{A}. \quad (3.13b)$$

□

One can now establish the link between the vector inversion problem of section 3.3.1 (problem 3.8) and the linearly separable bipartitioning problem of this section (problem 3.11).

**Equivalence 3.14 (vector inversion  $\leftrightarrow$  linearly separable bipartitioning)** Let be given a matrix  $V \in \mathbb{R}^{n \times p}$  with nonzero columns denoted by  $v_1, \dots, v_p \in \mathbb{R}^n$  and take  $\bar{s} \in \mathcal{S}$ , which is nonempty. By (3.3), there is a direction  $d \in \mathbb{R}^n$  such that  $\bar{s}_i v_i^\top d > 0$ , implying that  $\text{cone}\{\bar{s}_i v_i : i \in [1:p]\}$  is a pointed cone (proposition 3.9). Use the construction 3.13(1) associated with these  $v_i$ 's and this particular  $\bar{s}$  to define an affine space  $\mathbb{A}_0$  in  $\mathbb{R}^n$ .

For  $s \in \{\pm 1\}^p$ , set

$$I := \{i \in [1:p] : s_i \bar{s}_i = -1\} \quad \text{and} \quad J := \{i \in [1:p] : s_i \bar{s}_i = +1\}.$$

Then,

$$\text{cone}\{s_i v_i : i \in [1:p]\} \text{ is a pointed cone} \iff \exists \xi \in \mathbb{A}_0 : \max_{i \in I} \xi^\top \bar{v}_i < \min_{j \in J} \xi^\top \bar{v}_j. \quad (3.14)$$

Indeed, with  $K_I$  and  $K_J$  defined by (3.13a) (with  $s_i \curvearrowright \bar{s}_i$ ),

$$\begin{aligned} & \text{cone}\{s_i v_i : i \in [1:p]\} \text{ is a pointed cone} \\ & \iff \text{cone}\{s_i \bar{s}_i (\bar{s}_i v_i) : i \in [1:p]\} \text{ is a pointed cone} \\ & \iff \text{cone}((-K_I) \cup K_J) \text{ is a pointed cone} \\ & \iff \exists \xi \in \mathbb{A}_0 : \max_{i \in I} \xi^\top \bar{v}_i < \min_{j \in J} \xi^\top \bar{v}_j, \end{aligned}$$

where we have used the equivalence (i)  $\Leftrightarrow$  (iv) of the previous proposition (in which  $v_i$  is supposed to play the role of  $\bar{s}_i v_i$ ).

The equivalence (3.14) establishes the expected equivalence between the vector inversion problem 3.8 (in which one looks for all the  $s \in \{\pm 1\}^p$  such that  $\text{cone}\{s_i v_i : i \in [1:p]\}$  is a pointed cone) and the linearly separable bipartitioning problem 3.11 of the vectors  $\bar{v}_1 = v_1/(v_1^\top d), \dots, \bar{v}_p = v_p/(v_p^\top d)$ , where  $d$  is associated with the pointed cone  $\text{cone}\{\bar{s}_i v_i : i \in [1:p]\}$  by the equivalence (i)  $\Leftrightarrow$  (iii) of proposition 3.9. □

### 3.4 Discrete geometry: hyperplane arrangement

The equivalent problem examined in this section has a long history, going back at least to the XIXth century [50, 56]. More recently, it appears in *Computational Discrete Geometry* (the discipline has many other names), under the name of *hyperplane arrangement*. Contributions to this problem, or a more general version of it, with a discrete mathematics point of view, has been reviewed in [2, 27, 33, 34, 55]. It has applications in robotics [54],  $\lambda$ -matrices, halfspatial range estimation, Voronoi diagrams, degeneracy tests, minimum measure simplices [26; §4] and matrix rank computation [14].

**Problem 3.15 (arrangement of hyperplanes containing the origin)** Let be given two positive integers  $n$  and  $p \in \mathbb{N}^*$  and  $p$  nonzero vectors  $v_1, \dots, v_p \in \mathbb{R}^n$ . Consider the hyperplanes containing the origin:

$$\mathcal{H}_i := \{d \in \mathbb{R}^n : v_i^\top d = 0\}. \quad (3.15)$$

It is requested to list the regions of  $\mathbb{R}^n$  that are separated by these hyperplanes. Such a region is called a *cell* or a *chamber*, depending on the authors [2,5,53]. More specifically, let us define the half-spaces

$$\mathcal{H}_i^+ := \{d \in \mathbb{R}^n : v_i^\top d > 0\} \quad \text{and} \quad \mathcal{H}_i^- := \{d \in \mathbb{R}^n : v_i^\top d < 0\}.$$

The problem is to determine the following set of open sectors or cells of  $\mathbb{R}^n$ , indexed by the bipartitions  $(I_+, I_-)$  of  $[1 : p]$ :

$$\mathfrak{C} := \{(I_+, I_-) \in \mathcal{P}([1 : p]) : (\cap_{i \in I_+} \mathcal{H}_i^+) \cap (\cap_{i \in I_-} \mathcal{H}_i^-) \neq \emptyset\}. \quad (3.16)$$

□

The link between problem 3.15 and the signed feasibility of strict linear inequality systems of section 3.2.1 is obtained from the bijection

$$\eta : (I_+, I_-) \in \mathcal{P}([1 : p]) \mapsto s \in \{\pm 1\}^p, \text{ where } s_i = \begin{cases} +1 & \text{if } i \in I_+, \\ -1 & \text{if } i \in I_- \end{cases} \quad (3.17a)$$

and the setting

$$V = (v_1 \cdots v_p), \quad (3.17b)$$

whose columns are nonzero by assumption here and in section 3.2.1. Recall the definition (3.3) of the set of sign vectors  $\mathcal{S}$ .

**Proposition 3.16 (bijection  $\mathfrak{C} \leftrightarrow \mathcal{S}$ )** For the matrix  $V$  given by (3.17b), with nonzero columns  $v_i$ 's, the map  $\eta$  given by (3.17a) is a bijection from  $\mathfrak{C}$  onto  $\mathcal{S}$ .

*Proof.* Let  $(I_+, I_-) \in \mathcal{P}([1 : p])$  and  $s \in \{\pm 1\}^p$  be in correspondence through the bijection  $\eta$ :  $s = \eta((I_+, I_-))$ . Then,

$$\begin{aligned} (I_+, I_-) \in \mathfrak{C} &\iff \exists d \in (\cap_{i \in I_+} \mathcal{H}_i^+) \cap (\cap_{i \in I_-} \mathcal{H}_i^-) \\ &\iff \exists d \in \mathbb{R}^n : (v_i^\top d > 0 \text{ for } i \in I_+) \text{ and } (v_i^\top d < 0 \text{ for } i \in I_-) \\ &\iff \exists d \in \mathbb{R}^n : s \cdot V^\top d > 0 \quad [(3.17b)] \\ &\iff s \in \mathcal{S}. \end{aligned}$$

These equivalences show the bijectivity of  $\eta$  from  $\mathfrak{C}$  onto  $\mathcal{S}$ . □

**Equivalence 3.17 (signed linear system feasibility  $\leftrightarrow$  hyperplane arrangement)** The equivalence between problem 3.15 and problem 3.2 follows from the bijection of the map  $\eta : \mathfrak{C} \rightarrow \mathcal{S}$  claimed in proposition 3.16. □

## 4 Description of the B-differential

This section gives some elements of description of the B-differential  $\partial_B H(x)$ , when  $H$  is the piecewise affine function  $H$  given by (1.2) and  $x \in \mathbb{R}^n$  (we assume that  $n \in \mathbb{N}^*$ ). This description is often carried out in terms of the matrix  $V$  defined by (3.2), whose  $p := |\mathcal{E}^\neq(x)|$  columns are denoted by  $v_1, \dots, v_p \in \mathbb{R}^n$  (we assume that  $p \in \mathbb{N}^*$ ) and are assumed to be nonzero. Some properties of  $\partial_B H(x)$  are given in section 4.1, including those that are useful in [25]. Section 4.2 deals with the cardinal  $|\partial_B H(x)|$  of the B-differential, that is its number of elements.

### 4.1 Some properties of the B-differential

Let us start by a basic property of  $\partial_B H(x)$  in the affine case, which is its symmetry in the sense of definitions 2.4. The equivalence 3.4 allows us to give a straightforward proof.

**Proposition 4.1 (symmetry of  $\partial_B H(x)$ )** *When  $H$  is given by (1.2) and  $x \in \mathbb{R}^n$ ,  $\partial_B H(x)$  is symmetric. In particular,  $|\partial_B H(x)|$  is even.*

*Proof.* Let  $J \in \partial_B H(x)$  and  $s := \sigma(J)$ , where  $\sigma$  is defined by (3.5). By proposition 3.3,  $s \cdot V^\top d > 0$  is feasible for  $d$ . Now,  $(-s) \cdot V^\top d > 0$  is obviously feasible for  $d$  (take the opposite of the previous  $d$  as solution), so that  $-s \in \mathcal{S}$ . By the definition of the bijection  $\sigma : \partial_B H(x) \rightarrow \mathcal{S}$ , we see that  $J = \sigma^{-1}(s)$  and  $\tilde{J} := \sigma^{-1}(-s)$  are symmetric to each other in  $\partial_B H(x)$ . This shows the symmetry of  $\partial_B H(x)$ . It follows immediately that  $|\partial_B H(x)|$  is even.  $\square$

We now give a necessary and sufficient condition ensuring the completeness of  $\partial_B H(x)$  in the sense of definition 2.3. The condition was shown to be sufficient in [15; corollary 2.1(i)] for the nonlinear case (1.5), using a different proof, but we shall see in [25] that it is an easy consequence of that property in the affine case (1.2). Thanks to the equivalence 3.4, the present proof is short.

**Proposition 4.2 (completeness of the B-differential)** *The B-differential  $\partial_B H(x)$  is complete if and only if the matrix  $V \in \mathbb{R}^{n \times p}$  in (3.2) is injective. Hence, this property can hold only if  $p \leq n$ .*

*Proof.* [ $\Rightarrow$ ] We show the contrapositive. Assume that  $V$  is not injective, so that  $V\alpha = 0$  for some nonzero  $\alpha \in \mathbb{R}^p$ . With  $s \in \text{sgn}(\alpha)$ , one can write

$$\sum_{i \in [1:p]} |\alpha_i| s_i v_i = 0.$$

By the implication (iii)  $\Rightarrow$  (ii) of proposition 3.9 (or Gordan's alternative), it follows that there is no  $d \in \mathbb{R}^n$  such that  $s \cdot V^\top d > 0$ . By (3.3), this implies that  $s \notin \mathcal{S}$ . According to the equivalence 3.4,  $\sigma^{-1}(s) \notin \partial_B H(x)$ , showing that the B-differential is not complete.

[ $\Leftarrow$ ] Assume the injectivity of  $V$ . Let  $s \in \{\pm 1\}^p$ . Since  $V^\top$  is surjective, the system  $V^\top d = s$  is feasible for  $d \in \mathbb{R}^n$ . For this  $d$ ,  $s \cdot V^\top d = e$ , so that  $s \cdot V^\top d > 0$  is feasible for  $d \in \mathbb{R}^n$  and the selected  $s$  is in  $\mathcal{S}$ . We have shown that  $\mathcal{S} = \{\pm 1\}^p$  or that  $\partial_B H(x) = \sigma^{-1}(\{\pm 1\}^p)$  ( $\sigma^{-1}$  is defined by (3.5b)) is complete.  $\square$

We focus now on the connectivity of  $\partial_B H(x)$ , a notion that is more easily presented in  $\{\pm 1\}^p$  but that can be transferred straightforwardly to  $\partial_B H(x)$  by the bijection  $\sigma$  defined in (3.5).

**Definition 4.3 (adjacency in  $\{\pm 1\}^p$ )** Two sign vectors  $s^1$  and  $s^2 \in \{\pm 1\}^p$  are said to be *adjacent* if they differ by a single component (i.e., the vertices  $s^1$  and  $s^2$  of the cube  $\text{co}\{\pm 1\}^p$  can be joined by a single edge).  $\square$

**Definitions 4.4 (connectivity in  $\{\pm 1\}^p$ )** A *path* of length  $l$  in a subset  $S$  of  $\{\pm 1\}^p$  is a finite set of sign vectors  $s^0, \dots, s^l \in S$  such that  $s^i$  and  $s^{i+1}$  are adjacent for all  $i \in [0 : l - 1]$ ; in which case the path is said to be *joining*  $s^0$  to  $s^l$ . One says that a subset  $S$  of  $\{\pm 1\}^p$  is *connected* if any pair of points of  $S$  can be joined by a path in  $S$ .  $\square$

**Proposition 4.5 (connectivity of the B-differential)** *The set  $\mathcal{S}$  defined by (3.3) is connected if and only if  $V$  has no colinear columns. In this case, any points  $s$  and  $\tilde{s}$  of  $\mathcal{S}$  can be joined by a path of length  $l := \sum_{i \in [1 : p]} |\tilde{s}_i - s_i|/2 \leq p$  in  $\mathcal{S}$ .*

*Proof.*  $[\Rightarrow]$  We prove the contrapositive. Suppose that the columns  $v_i$  and  $v_j$  of  $V$  are colinear:  $v_j = \alpha v_i$ , for some  $\alpha \in \mathbb{R}^*$ . Assume that  $\alpha > 0$  (resp.  $\alpha < 0$ ). By (3.3), for any  $s \in \mathcal{S}$ , one can find  $d \in \mathbb{R}^n$  such that  $s \cdot V^T d > 0$ , implying that  $s_i = s_j$  (resp.  $s_i = -s_j$ ). Therefore, one cannot find a path joining  $s \in \mathcal{S}$  and  $-s$ , which also belongs to  $\mathcal{S}$  by proposition 4.1, since one would have to change the two components with index in  $\{i, j\}$  and that these components must be changed simultaneously for the sign vectors in  $\mathcal{S}$  (while the adjacency property along a path prevents from changing more than one sign at a time).

$[\Leftarrow]$  Let  $s$  and  $\tilde{s} \in \mathcal{S}$ . It suffices to show that there is a path of length  $l$  in  $\mathcal{S}$  joining  $s$  to  $\tilde{s}$ . By the expression (3.3) of  $\mathcal{S}$ , one can find  $d$  and  $\tilde{d} \in \mathbb{R}^n$  such that

$$s \cdot V^T d > 0 \quad \text{and} \quad \tilde{s} \cdot V^T \tilde{d} > 0.$$

Note that, since the vectors  $\{v_i : i \in [1 : p]\}$  are not colinear, by assumption, the vectors  $\{\bar{v}_i := v_i / (v_i^T d) : i \in [1 : p]\}$  are all different. Set

$$\xi := \tilde{d} - d.$$

Since a small modification of  $\tilde{d}$  preserves the inequality  $\tilde{s} \cdot V^T \tilde{d} > 0$  and since the vectors  $\bar{v}_i$ 's are all different, lemma 2.5 tells us that, at the cost of a small change of  $\xi$ , one can assume that

$$|\{v_i^T \xi / (v_i^T d) : i \in [1 : p]\}| = p. \quad (4.2a)$$

Since, one could have added  $v_0 = 0$  to the list of vectors  $v_i$ 's, one can also assume that

$$v_i^T \xi \neq 0, \quad \forall i \in [1 : p]. \quad (4.2b)$$

Then, one can set  $t_i := -(v_i^T d) / (v_i^T \xi)$ , for  $i \in [1 : p]$ , which are  $p$  distinct values by (4.2a). It follows that

$$v_i^T d + t_i v_i^T \xi = 0 \quad \text{or} \quad (1 - t_i) v_i^T d + t_i v_i^T \tilde{d} = 0 \quad \text{or} \quad v_i^T [(1 - t_i) d + t_i \tilde{d}] = 0. \quad (4.2c)$$

For each  $i \in [1 : p]$ , we are interested in the change of sign of  $v_i^T [(1 - t) d + t \tilde{d}]$  when  $t$  goes through the interval  $(0, 1)$  (i.e., when  $(1 - t) d + t \tilde{d}$  goes through the relative interior of the

segment  $[d, \tilde{d}]$ . From the middle identity in (4.2c), we see that  $t_i \in (0, 1)$  if and only if  $s_i \tilde{s}_i = -1$  (i.e.,  $v_i^\top d$  and  $v_i^\top \tilde{d}$  have opposite signs). Therefore, the number of  $t_i \in (0, 1)$  is equal to  $l = \sum_{i \in [1:p]} |\tilde{s}_i - s_i|/2 \leq p$ . Let us denote them by

$$0 < t_{i_1} < \dots < t_{i_l} < 1.$$

When  $t \in (0, 1)$  crosses a  $t_j \in (0, 1)$ , a single  $v_i^\top [(1-t)d + t\tilde{d}]$ , for  $i \in [1:p]$ , changes its sign (since all the  $t_j$ 's are different, see the last identity in (4.2c)). Therefore, there are sign vectors  $s^{i_j} \in \{\pm 1\}^p$ , for  $j \in [1:l]$ , such that

$$s^{i_j} \cdot V^\top [(1-t)d + t\tilde{d}] > 0, \quad \text{for } t \in (t_{i_j}, t_{i_{j+1}}),$$

and each of these sign vectors is different from the previous one by a single component (they are adjacent in the sense of definition 4.4). Therefore, we have defined a path of length  $l \leq p$  in  $\mathcal{S}$ , namely  $s^{i_0} = s, s^{i_1}, \dots, s^{i_l} = \tilde{s}$ , joining  $s$  to  $\tilde{s}$ . This proves the proposition.  $\square$

This connectivity property is also proved in [2; section 1.10.4] with a very different point of view, related to graph theory. One can also give a proof of the implication “ $\Leftarrow$ ” of proposition 4.5, using linearly separable bipartitioning (problem 3.11), but it does not have the conceptual simplicity of the previous one. For  $k \in [1:p]$ , we define  $\nu^k \in \{\pm 1\}^p$  by

$$\nu_i^k := \begin{cases} -1 & \text{if } i = k, \\ +1 & \text{otherwise.} \end{cases} \quad (4.3)$$

Hence, “ $\nu^k \cdot$ ” applied to a vector reverses the sign of its  $k$ th component.

*Another proof of proposition 4.5.* We only consider the implication “ $\Leftarrow$ ”. Let  $s$  and  $\tilde{s} \in \mathcal{S}$ . It suffices to show that there is a path of length  $l$  in  $S$  joining  $s$  to  $\tilde{s}$ .

Let us apply the construction 3.13(1-2) to introduce  $\mathbb{A}, \mathbb{A}_0, \{\bar{v}_i : i \in \mathcal{E}^\neq(x)\}, K_I, K_J, C_I$  and  $C_J$ , from the vectors  $\{v_i : i \in \mathcal{E}^\neq(x)\}, s \in \mathcal{S}$  and  $(I, J) \in \mathcal{P}(\mathcal{E}^\neq(x))$  defined by

$$I := \{i \in \mathcal{E}^\neq(x) : \tilde{s}_i = -s_i\} \quad \text{and} \quad J := \{j \in \mathcal{E}^\neq(x) : \tilde{s}_j = s_j\}. \quad (4.4a)$$

Clearly,  $|I| = l$  (the number given in the statement of the proposition) and  $|J| = p - l$ . Since  $s$  and  $\tilde{s} \in \mathcal{S}$ , (3.10) tells us that  $\text{cone}\{s_i v_i : i \in \mathcal{E}^\neq(x)\}$  and  $\text{cone}\{\tilde{s}_i v_i : i \in \mathcal{E}^\neq(x)\}$  are pointed cones. Now,  $\text{cone}\{s_i v_i : i \in \mathcal{E}^\neq(x)\} = \text{cone}(K_I \cup K_J)$  and  $\text{cone}\{\tilde{s}_i v_i : i \in \mathcal{E}^\neq(x)\} = \text{cone}((-K_I) \cup K_J)$ . The pointedness of  $\text{cone}(K_I \cup K_J)$  and  $\text{cone}((-K_I) \cup K_J)$  and the implication  $(i) \Rightarrow (iv)$  of proposition 3.12 imply that there exists a covector  $\xi \in \mathbb{A}_0$  such that  $\xi^\top \bar{v}_i < \xi^\top \bar{v}_j$  for all  $i \in I$  and  $j \in J$ . By their strictness, these inequalities, in finite number, are not modified by a small perturbation of  $\xi$  and, by lemma 2.5,  $\xi$  can be chosen such that all the  $\xi^\top \bar{v}_i$  for  $i \in \mathcal{E}^\neq(x)$  are distinct (the  $\bar{v}_i$  are all distinct by the assumption on the non-colinearity of the  $v_i$ 's). If the indices in  $I$  are denoted by  $i_k$ , for  $k \in [1:l]$  and those in  $J$  are denoted by  $j_k$ , for  $k \in [1:p-l]$ , one can assume that

$$\xi^\top \bar{v}_{i_1} < \xi^\top \bar{v}_{i_2} < \dots < \xi^\top \bar{v}_{i_l} < \xi^\top \bar{v}_{j_1} < \xi^\top \bar{v}_{j_2} < \dots < \xi^\top \bar{v}_{j_{p-l}}. \quad (4.4b)$$

For  $k \in [0:l]$ , define  $s^{i_k} \in \{\pm 1\}^p$  as follows

$$s^{i_0} := s \quad \text{and} \quad s^{i_k} := \nu^{i_k} \cdot s^{i_{k-1}}, \quad \text{for } k \in [1:l],$$

where  $\nu^k$  is defined by (4.3). We claim that

$$s^{i_0}, s^{i_1}, \dots, s^{i_l} \text{ is a path of length } l \text{ in } S,$$

that  $s = s^{i_0}$  and that  $\tilde{s} = s^{i_l}$ . This claim will prove the proposition.

The fact that  $(s^{i_0}, s^{i_1}, \dots, s^{i_l})$  is a path of length  $l$  in  $\{\pm 1\}^p$  is clear since  $s^{i_{k+1}}$  is obtained from  $s^{i_k}$  by changing a single of its components ( $s^{i_k}$  and  $s^{i_{k+1}}$  are adjacent in the sense of definition 4.4). Furthermore  $s = s^{i_0}$  by definition and  $\tilde{s} = s^{i_l}$  since  $s^{i_l}$  is obtained from  $s$  by changing the sign of all its components with index in  $I$  (definition of the  $i_k$ 's). Hence, the path  $(s^{i_0}, s^{i_1}, \dots, s^{i_l})$  joins  $s$  to  $\tilde{s}$ . It remains to show that the  $s^{i_k}$ 's are in  $\mathcal{S}$ . Define, for  $k \in [1:l]$ :

$$\alpha_k := \frac{\xi^\top \bar{v}_{i_k} + \xi^\top \bar{v}_{i_{k+1}}}{2}, \quad I_k := \{i_1, \dots, i_k\} \quad \text{and} \quad J_k := [1:p] \setminus I_k.$$

By (4.4b), the hyperplane  $\{\bar{v} \in \mathbb{A} : \xi^\top \bar{v} = \alpha_k\}$  separates the vectors  $\{\bar{v}_i : i \in I_k\}$  and  $\{\bar{v}_j : j \in J_k\}$  in  $\mathbb{A}$ . Therefore, with the notation (3.13a)-(3.13b),  $C_{I_k} \cap C_{J_k} = \emptyset$ . By the implication (iii)  $\Rightarrow$  (i) of proposition 3.12, this implies that  $\text{cone}((-K_{I_k}) \cup K_{J_k})$  is pointed. By the implication (i)  $\Rightarrow$  (iii) of proposition 3.9, the system

$$\begin{cases} -s_i v_i^\top \tilde{d} > 0 & \text{for } i \in I_k, \\ s_i v_i^\top \tilde{d} > 0 & \text{for } i \in J_k \end{cases}$$

has a solution  $\tilde{d} \in \mathbb{R}^n$ . By (3.3), this amounts to saying that  $s^{i_k} = (-s_{I_k}, s_{J_k})$  is in  $\mathcal{S}$ , as expected.  $\square$

**Definition 4.6 (extremality in  $\{\pm 1\}^p$ )** A point  $\bar{v}_k$  of a finite set  $\mathfrak{V} := \{\bar{v}_i : i \in [1:p]\}$  is said to be an *extreme point* of  $\mathfrak{V}$  if  $\bar{v}_k \notin \text{co}\{\bar{v}_i : i \in [1:p] \setminus \{k\}\}$ . This is equivalent to saying that  $\bar{v}_k$  is an extreme point (in the convex analysis sense [51; § 18]) of the convex hull  $\text{co } \mathfrak{V}$ .  $\square$

**Proposition 4.7 (adjacency and extremality)** *Adopt the construction 3.13(1) associated with the nonzero vectors  $v_1, \dots, v_p \in \mathbb{R}^n$  and  $s \in \mathcal{S}$ . Set  $\mathfrak{V} := \{\bar{v}_i : i \in [1:p]\}$ . For some  $k \in [1:p]$ , let  $\nu^k$  be defined by (4.3). Then, the following properties are equivalent:*

- (i)  $\nu^k \cdot s \in \mathcal{S}$ ,
- (ii)  $\bar{v}_k$  is an extreme point of  $\mathfrak{V}$ .

*Proof.* Since  $s \in \mathcal{S}$ , (3.10) implies that  $\text{cone}\{s_i v_i : i \in [1:p]\}$  is pointed. Define  $\tilde{s} := \nu^k \cdot s$ . We have the following equivalences

$$\begin{aligned} \tilde{s} \in \mathcal{S} &\iff \text{cone}\{\tilde{s}_i v_i : i \in [1:p]\} \text{ is pointed} && [(3.10)] \\ &\iff \bar{v}_k \notin \text{co}\{\bar{v}_i : i \in [1:p] \setminus \{k\}\} && [(i) \Leftrightarrow (iii) \text{ in proposition 3.12}] \\ &\iff \bar{v}_k \text{ is an extreme point of } \mathfrak{V} && [\text{definition}]. \end{aligned} \quad \square$$

To put it another way, the previous proposition tells us that the sign vectors (resp. the Jacobians in  $\partial_B H(x)$ ) that are adjacent to a given  $s \in \mathcal{S}$  (resp. to a given Jacobian  $\sigma^{-1}(s) \in \partial_B H(x)$ ) are those of the form  $\nu^k \cdot s$  with  $k \in [1:p]$  such that  $\bar{v}_k$  is an extreme point of  $\mathfrak{V}$ .

## 4.2 Cardinal of the B-differential

### 4.2.1 Winder's formula

Giving the exact number of elements in  $\partial_B H(x)$ , that is  $|\partial_B H(x)| = |\mathcal{S}| = |\mathfrak{C}| = 2^p - |\mathcal{S}^c| = 2^p - |\mathcal{I}|$ , with the notation (3.3), (3.16) and (3.7), is a tricky task, even in the present affine case, since it subtly depends on the arrangement of the vectors  $v_i$ 's in the space (see figure 3.1). Many contributions have been done on this subject; the earliest we cite dates from 1826 [2–4, 17, 27, 33, 50, 55, 56, 58, 60]. The formula for  $|\partial_B H(x)|$  that is appropriate to our setting is due to Winder [59; 1966] and reads

$$|\partial_B H(x)| = \sum_{I \subseteq [1:p]} (-1)^{|I| - \text{rank}(V, I)}, \quad (4.5)$$

where the term in the right-hand side corresponding to  $I = \emptyset$  is 1 (since  $|\emptyset| = 0$  and one takes the convention that  $\text{rank}(V, \emptyset) = 0$ ). This amazing expression, with its only algebraic nature, potentially made of positive and negative terms, is explicit but, to our knowledge, has not been at the origin of a method to list the elements of  $\partial_B H(x)$ .

### 4.2.2 Bounds

When  $p$  is large, computing the cardinal  $|\partial_B H(x)|$  from (4.5) by evaluating the  $2^p$  ranks  $\text{rank}(V, I)$  for  $I \subseteq [1:p]$  could be excessively expensive. Therefore, having lower and upper bounds on  $|\partial_B H(x)|$  may happen to be useful in some circumstances. Actually, simple bounds can be obtained very easily, using one of the formulations of the problem given in section 3. Proposition 4.8 gives a first example of simple lower and upper bounds, while proposition 4.9 reinforces the lower bound and proposition 4.11 reinforces the upper bound, by taking into account the rank of  $V$ .

**Proposition 4.8 (lower and upper bounds on  $|\partial_B H(x)|$ )** *One has*

$$2p \leq |\partial_B H(x)| \leq 2^p.$$

*Proof.* Consider first the lower bound. Let  $s \in \mathcal{S}$ , which is nonempty (since  $\partial_B H(x) \neq \emptyset$  when  $H$  is Lipschitz continuous, see also algorithm 5.2). Then,  $-s \in \mathcal{S}$  by symmetry (proposition 4.1). By connectivity (proposition 4.5), there is path  $s^0, s^1, \dots, s^p$  of length  $p$ , joining  $s^0 := s$  to  $s^p := -s$ . Since  $\{-s^i : i \in [0:p-1]\} \cap \{s^j : j \in [0:p-1]\} = \emptyset$  (indeed, one cannot have  $-s^i = s^j$  for some  $i$  and  $j \in [0:p-1]$ , because  $s^i$  and  $s^j$  have some identical components),  $\mathcal{S}$  contains at least the  $2p$  sign vectors  $\{-s^i : i \in [0:p-1]\} \cup \{s^i : i \in [0:p-1]\}$ . By the bijectivity of  $\sigma$  in (3.5a),  $\partial_B H(x)$  contains at least  $2p$  Jacobians.

The upper bound results from the inclusion  $\partial_B H(x) \subseteq \bar{\partial}_B H(x)$  (proposition 2.2) and  $|\bar{\partial}_B H(x)| = 2^p$  by the definition (2.3) of  $\bar{\partial}_B H(x)$ .  $\square$

**Proposition 4.9 (lower bound on  $|\partial_B H(x)|$ )** *Suppose that  $H$  is given by (1.2) and*

that  $V$ , given by (3.2), has nonzero columns. Set  $r := \text{rank}(V)$ . Then,

$$2^r \leq |\partial_B H(x)|.$$

*Proof.* Let  $I \subseteq [1:p]$  be such that  $|I| = r$  and  $V_{:,I}$  is injective (this is possible by definition of the rank). By the surjectivity of  $V_{:,I}^\top$ , for any  $s^I \in \{\pm 1\}^{|I|}$ , one can find  $d^I \in \mathbb{R}^n$  such that

$$V_{:,I}^\top d^I = s^I. \quad (4.6a)$$

Now, using lemma 2.5 on the vectors  $\{0, v_1, \dots, v_p\}$ , one can find  $\xi \in \mathbb{R}^n$  such that (if there are two identical  $v_i$ 's, the values  $v_i^\top \xi$  will be identical and nonzero):

$$v_i^\top \xi \neq 0, \quad \forall i \in [1:p]. \quad (4.6b)$$

We claim that

$$\forall t > 0 \text{ sufficiently small, } \forall i \in [1:p]: \quad v_i^\top (d^I + t\xi) \neq 0. \quad (4.6c)$$

Indeed, this is clear for the indices  $i \in I$  by (4.6a). Therefore, if the claim is false, one would have  $v_i^\top (d^I + t\xi) = 0$  for some  $i \in [1:p] \setminus I$  and a sequence of positive  $t$ 's tending to zero; this would imply that  $v_i^\top d^I = 0$  and next  $v_i^\top \xi = 0$ , which contradicts (4.6b).

From (4.6c), one deduces that one can find a  $d_t := d^I + t\xi$  such that

$$(\text{sgn}(v_i^\top d_t) = \text{sgn}(s_i^I), \quad \forall i \in I) \quad \text{and} \quad (v_i^\top d_t \neq 0, \quad \forall i \in [1:p] \setminus I).$$

As a result  $s = \text{sgn}(V_{:,I}^\top d_t) \in \mathcal{S}$ . Since there are at least  $2^r$  different such  $s$ 's (one for each choice of  $s^I \in \{\pm 1\}^{|I|}$ ),  $|\mathcal{S}| \geq 2^{|I|} = 2^r$ . By the equivalence 3.4,  $|\partial_B H(x)| \geq 2^r$ .  $\square$

Recall that a function  $\varphi : x \in \mathbb{M} \rightarrow \varphi(x) \in \mathbb{R}$ , defined on a metric space  $\mathbb{M}$ , is said to be *lower semicontinuous* if, for any  $x \in \mathbb{M}$  and any sequence  $\{x_k\}$  converging to  $x$ , one has  $\varphi(x) \leq \liminf_{k \rightarrow \infty} \varphi(x_k)$ . It is known that the rank of a matrix can only increase in the neighborhood of a given matrix, which implies its lower semicontinuity. The next lemma shows that the same property holds for  $|\mathcal{S}| \in \mathbb{N}^*$ , viewed as a function of  $V$ . Recall that the bijection  $\sigma$  is defined by (3.5).

**Lemma 4.10 (lower semicontinuity of  $|\partial_B H(x)|$ )** *Suppose that  $\mathcal{S} = \sigma(\partial_B H(x))$  is viewed as a function of  $V \in \mathbb{R}^{n \times p}$  given by (3.2). Then,  $\mathcal{S}(V) \subseteq \mathcal{S}(\tilde{V})$  for  $\tilde{V}$  near  $V$  in  $\mathbb{R}^{n \times p}$ . In particular,  $|\partial_B H(x)| \in \mathbb{N}^*$  is a lower semicontinuous function of  $V \in \mathbb{R}^{n \times p}$ .*

*Proof.* Suppose that  $s \in \mathcal{S}(V)$ . Then, by the definition (3.3) of  $\mathcal{S}$ ,  $s \cdot V^\top d > 0$  is feasible for  $d \in \mathbb{R}^n$ . Clearly, it follows that, for  $\tilde{V}$  near  $V$ ,  $s \cdot \tilde{V}^\top d > 0$  is also feasible for  $d \in \mathbb{R}^n$ . Since  $\mathcal{S}$  is finite, there is a neighborhood of  $V$  such that, for any  $\tilde{V}$  in this neighborhood and any  $s \in \mathcal{S}$ ,  $s \cdot \tilde{V}^\top d > 0$  is also feasible for  $d \in \mathbb{R}^n$ . We have shown that  $\mathcal{S}(V) \subseteq \mathcal{S}(\tilde{V})$  for  $\tilde{V}$  near  $V$ .

As a direct consequence of this inclusion, we have that  $|\mathcal{S}(V)| \leq |\mathcal{S}(\tilde{V})|$  for  $\tilde{V}$  near  $V$ . The lower semicontinuity of  $V \mapsto |\partial_B H(x)|$  now follows from the fact that  $|\partial_B H(x)| = |\mathcal{S}|$ .  $\square$



Equality in the upper estimate (4.7) below was shown by Winder [59; corollary] when (4.8) holds. When this condition is verified, the vectors  $v_i$ 's are said to be in *general position* (note that this condition is stable for small perturbations of  $V$ , which preserve the range condition  $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$ ). Example of vectors in general position are those in the left-hand side and right-hand side pictures in figure 3.1 (recall that the points are the normalized vectors  $\bar{v}_i$ 's so that the  $v_i$ 's are in  $\mathbb{R}^3$ ). Those in the middle picture are not in general position. This is due to the fact that  $r := \text{rank}(V) = 3$  while for the 3 bottom vectors, with indices in  $I$  say, one has  $\min(|I|, r) - \text{rank}(V_{:,I}) = 3 - 2 \neq 0$ . We also observe that equality does not hold in (4.7) for this middle configuration since  $|\partial_B H(x)| = 12$ , while the right-hand side of (4.7) reads  $2[\binom{3}{0} + \binom{3}{1} + \binom{3}{2}] = 14$ .

**Proposition 4.11 (upper bound on  $|\partial_B H(x)|$ )** *Suppose that  $H$  is given by (1.2) and that  $V$ , given by (3.2), has nonzero columns. Let  $r := \text{rank}(V)$ . Then,*

$$|\partial_B H(x)| \leq 2 \sum_{i \in [0:r-1]} \binom{p-1}{i} \quad (4.7)$$

*with equality if and only if*

$$\forall I \subseteq [1:p] : \quad \text{rank}(V_{:,I}) = \min(|I|, r). \quad (4.8)$$

*Proof.* 1) The proof of the implication “(4.8)  $\Rightarrow$  (4.7) with equality” is given in [59; corollary]. For the reader's convenience we give a proof with our notation. One has

$$\begin{aligned} |\mathcal{S}| &= \sum_{I \subseteq [1:p]} (-1)^{|I| - \text{rank}(V_{:,I})} \quad [\text{Winder's formula (4.5)}] \\ &= \sum_{\substack{I \subseteq [1:p] \\ |I| \leq r}} (-1)^{|I| - \text{rank}(V_{:,I})} + \sum_{\substack{I \subseteq [1:p] \\ |I| > r}} (-1)^{|I| - \text{rank}(V_{:,I})} \\ &= \sum_{\substack{I \subseteq [1:p] \\ |I| \leq r}} 1 + \sum_{\substack{I \subseteq [1:p] \\ |I| > r}} (-1)^{|I| - r} \quad [\text{rank}(V_{:,I}) = \min(|I|, r)] \\ &= \sum_{i \in [0:r]} \binom{p}{i} + \sum_{i \in [r+1:p]} (-1)^{i-r} \binom{p}{i} \\ &= \sum_{i \in [0:p]} \binom{p}{i} - 2 \sum_{i \in \{r+1, r+3, \dots\}} \binom{p}{i} \\ &= 2 \left[ \sum_{i \in [0:p-1]} \binom{p-1}{i} - \sum_{i \in \{r+1, r+3, \dots\}} \binom{p}{i} \right], \end{aligned} \quad (4.9a)$$

where the last sum in the brackets is zero if  $r+1 > p$ . If  $r+1 \leq p$  and  $p-r$  is odd, the last sum in (4.9a) has at least the term  $\binom{p}{p} = 1$ . If  $r+1 \leq p$  and  $p-r$  is even, then  $r+1 \leq p-1$ .

With these particular cases in mind, one can evaluate the last sum in (4.9a) as follows

$$\begin{aligned} \sum_{i \in \{r+1, r+3, \dots\}} \binom{p}{i} &= \begin{cases} \sum_{i \in \{r+1, r+3, \dots, p-2\}} \left[ \binom{p-1}{i-1} + \binom{p-1}{i} \right] + 1 & \text{if } p-r \text{ is odd} \\ \sum_{i \in \{r+1, r+3, \dots, p-1\}} \left[ \binom{p-1}{i-1} + \binom{p-1}{i} \right] & \text{if } p-r \text{ is even} \end{cases} \\ &= \sum_{i \in [r:p-1]} \binom{p-1}{i}. \end{aligned}$$

Using (4.9a), we get immediately (4.7) with equality.

2) Let us now show (4.7) by contradiction, assuming that it does not hold for  $H$  defined by (1.2) and some matrix  $V \in \mathbb{R}^{n \times p}$  of rank  $r$  given by (3.2):

$$|\mathcal{S}(V)| > \beta, \quad (4.9b)$$

where we have written  $|\partial_B H(x)| = |\mathcal{S}(V)|$  to highlight the dependence of the B-differential in  $V$  and  $\beta$  is the right-hand side of (4.7). It certainly suffices to show that one can find a sequence  $\{V_k\} \subseteq \mathbb{R}^{n \times p}$  converging to  $V$  that satisfies

$$|\mathcal{S}(V_k)| = \beta, \quad (4.9c)$$

since then one would have the expected contradiction with the lower semicontinuity of  $V \mapsto |\mathcal{S}(V)|$  ensured by lemma 4.10:

$$\liminf_{k \rightarrow \infty} |\mathcal{S}(V_k)| = \beta < |\mathcal{S}(V)|.$$

To find  $V_k$  arbitrarily close to  $V$  verifying (4.9c), we proceed as follows. If (4.9b) holds, the first part of the proof implies that  $V$  does not satisfy (4.8). Our goal is to construct from  $V$  a matrix  $V_k$  arbitrarily close to  $V$  that satisfies (4.8) with  $V = V_k$ , hence (4.9c), by the first part of the proof. We also arrange for  $\mathcal{R}(V_k) \subseteq \mathcal{R}(V)$ , in order to ensure  $\text{rank}(V_k) \leq r$ .

Note first that  $\text{rank}(V_{:,I}) \leq |I|$  and  $\text{rank}(V_{:,I}) \leq \text{rank}(V) =: r$  for all  $I \subseteq [1:p]$ , so that  $\min(|I|, r) - \text{rank}(V_{:,I})$  is nonnegative. Now, suppose that for some  $I \subseteq [1:p]$  one has  $\min(|I|, r) - \text{rank}(V_{:,I}) > 0$ . We consider two complementary cases.

- If  $|I| < r$ , then, for an arbitrary small perturbation of the vectors  $v_i \rightsquigarrow \tilde{v}_i$ , with  $i \in I$ , one can get the  $\tilde{v}_i$ 's linearly independent in  $\mathcal{R}(V)$ . If one takes  $\tilde{v}_i = v_i$  for  $i \notin I$ , the matrix  $\tilde{V}$  formed of the vectors  $\tilde{v}_i$ 's verifies  $\text{rank}(\tilde{V}_{:,I}) = |I| = \min(|I|, r)$ .
- If  $|I| \geq r$ , then, for an arbitrary small perturbation of the vectors  $v_i \rightsquigarrow \tilde{v}_i$ , with  $i \in I$ , one can get the  $\tilde{v}_i$ 's generate  $\mathcal{R}(V)$ , which is of dimension  $r$ . If one takes  $\tilde{v}_i = v_i$  for  $i \notin I$ , the matrix  $\tilde{V}$  formed of the vectors  $\tilde{v}_i$ 's verifies  $\text{rank}(\tilde{V}_{:,I}) = r = \min(|I|, r)$ .

By the property of the rank, which can only increase in a neighborhood of a given matrix, if the perturbations taken above are sufficiently small, one can only increase  $\text{rank}(V_{:,I'})$  for all the other subsets  $I' \subseteq [1:p]$  (in finite number). In addition, since the perturbed matrix  $\tilde{V}$  is such that  $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$ , one has  $\text{rank}(\tilde{V}_{:,I'}) \leq \min(|I'|, r)$ , with equality if one already has  $\text{rank}(V_{:,I'}) = \min(|I'|, r)$ . By proceeding similarly with the next nonempty index sets  $I'' \subseteq [1:p]$  such that  $\text{rank}(\tilde{V}_{:,I''}) < \min(|I''|, r)$ , one finally obtain a matrix  $V_k$  such that (4.8) holds, that is  $\text{rank}((V_k)_{:,I}) = \min(|I|, r)$  for all  $I \subseteq [1:p]$ . By taking smaller and smaller perturbations of  $V$ , one also has  $V_k \rightarrow V$ .

3) One still has to show that “(4.7) with equality  $\Rightarrow$  (4.8)”. We proceed by contradiction, assuming that there is an index set  $I \subseteq [1 : p]$  such that

$$\text{rank}(V_{:,I}) < \min(|I|, r). \quad (4.9d)$$

Let  $\beta$  be the right-hand side of (4.7). It certainly suffices to show that there is a perturbation  $\tilde{V} \in \mathbb{R}^{n \times p}$  of  $V$  such that  $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$  and  $|\mathcal{S}(\tilde{V})| > \beta$  (since then we would have a matrix  $\tilde{V} \in \mathbb{R}^{n \times p}$  of rank  $r$  for which the bound (4.7) does not hold, in contradiction with what has been shown in part 2 of the proof). By lemma 4.10, if the perturbed matrix  $\tilde{V}$  is sufficiently close to  $V$ , one has  $\mathcal{S}(V) \subseteq \mathcal{S}(\tilde{V})$ , so that it suffices to show that  $\mathcal{S}(\tilde{V})$  contains a sign vector  $s$  that is not in  $\mathcal{S}(V)$ .

We claim that (4.9d) implies that one can find an index set  $J \subseteq I$  such that

$$V_{:,J} \text{ is not injective} \quad \text{and} \quad |J| \leq r. \quad (4.9e)$$

Indeed, if  $|I| \leq r$ , one can take  $J = I$  to satisfy (4.9e), since  $\text{rank}(V_{:,I}) < |I|$  implies that  $V_{:,I}$  is not injective. If  $|I| > r$ , then  $\text{rank}(V_{:,I}) < r$ , which implies that any  $J \subseteq I$  such that  $|J| = r$  satisfies (4.9e).

Since  $V_{:,J}$  is not injective,  $V_{:,J}\alpha_J = 0$  for some nonzero  $\alpha_J \in \mathbb{R}^{|J|}$ . Take  $\tilde{s}_J \in \{\pm 1\}^{|J|}$  satisfying  $\tilde{s}_i \in \text{sgn}(\alpha_i)$  for all  $i \in J$ . Hence,

$$(V_{:,J} \text{Diag}(\tilde{s}_J))(\tilde{s}_J \cdot \alpha_J) = 0.$$

Since  $\tilde{s}_J \cdot \alpha_J \in \mathbb{R}_+^{|J|} \setminus \{0\}$ , Gordan’s alternative (3.8) yields

$$\nexists d \in \mathbb{R}^n : \quad \tilde{s}_J \cdot V_{:,J}^T d > 0.$$

This implies that there is no  $s \in \mathcal{S}(V)$  such that  $s_J = \tilde{s}_J$ . We are going to construct an arbitrary small perturbation  $\tilde{V}$  of  $V$  with an  $s \in \mathcal{S}(\tilde{V})$  satisfying  $s_J = \tilde{s}_J$ , which will conclude the proof.

Let  $J^c := [1 : p] \setminus J$ . By (4.9e),  $|J| \leq r \leq n$  so that one can find a matrix  $\tilde{V} \in \mathbb{R}^{n \times p}$  such that  $\tilde{V}_{:,J^c} = V_{:,J^c}$ ,  $\tilde{V}_{:,J}$  is injective,  $\tilde{V}_{:,J} - V_{:,J}$  is arbitrary small and  $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$ . The surjectivity of  $\tilde{V}_{:,J}^T$  implies that there is a  $d_0 \in \mathbb{R}^n$  such that  $\tilde{V}_{:,J}^T d_0 = \tilde{s}_J$ , hence

$$\tilde{s}_J \cdot \tilde{V}_{:,J}^T d_0 > 0.$$

Now, set  $d := d_0 + \varepsilon \tilde{d}$ , where  $\varepsilon > 0$  is sufficiently small and  $\tilde{d}$  is a discriminating vector given by lemma 2.5. We see that  $\tilde{s}_i \tilde{V}_{:,i}^T d > 0$  for  $i \in J$  and  $\tilde{V}_{:,i}^T d \neq 0$  for  $i \in J^c$ . Finally, we obtain that  $s \in \{\pm 1\}^p$  defined by  $s_i = \text{sgn}(\tilde{V}_{:,i}^T d)$  for all  $i \in [1 : p]$  is in  $\mathcal{S}(\tilde{V})$  and satisfies  $s_J = \tilde{s}_J$ , as desired.  $\square$

### 4.2.3 Special configurations

We consider in this section some particular matrices  $V$  given by (3.2), which may be interesting to familiarize with the B-differential of  $H$ . In these cases,  $|\partial_B H(x)|$  can be computed easily. They can usually be obtained from Winder’s formula (4.5), although some other proofs are also possible. We consider two case in which the rank  $r := \text{rank}(V)$  takes the value 2 or  $p$ , which yield the lower and upper bounds on  $|\partial_B H(x)|$  given by proposition 4.8. As shown by figure 3.1, however,  $|\partial_B H(x)|$  not only depends on  $r$ .

**Injective matrix**

$$V \text{ is injective} \iff |\partial_B H(x)| = 2^p. \quad (4.10)$$

*Proof.* Indeed, by proposition 4.2, the B-differential  $\partial_B H(x)$  is complete (meaning that it is equal to  $\bar{\partial}_B H(x)$ , which is given by (2.3)) if and only if  $V$  is injective. Clearly, the completeness of  $\partial_B H(x)$  is equivalent to  $|\partial_B H(x)| = 2^p$ . Then, (4.10) follows.  $\square$

More algebraically, the implication “ $\Rightarrow$ ” can be deduced from Winder’s formula (4.5). If  $V$  is injective, then, for any  $I \subseteq [1 : p]$ , one has  $|I| = \text{rank}(V_{:,I})$ . Therefore, there are  $2^p$  terms in the right-hand side of (4.5), each of them of value 1. This yields  $|\partial_B H(x)| = 2^p$ . Conversely, if  $|\partial_B H(x)| = 2^p$ , one must have  $r = p$  in (4.7), meaning the  $V$  is injective.

**Fan arrangement**

$$\text{The } v_i \text{'s are not 2 by 2 colinear and } \text{rank}(V) = 2 \implies |\partial_B H(x)| = 2p. \quad (4.11)$$

*Proof.* A short proof uses equality in the Winder bound (4.7). Since the  $v_i$ 's are not 2 by 2 colinear, one has for any  $I \subseteq [1 : p]$ :

$$\text{rank}(V_{:,I}) = \begin{cases} |I| & \text{if } |I| \leq 2 \\ 2 & \text{if } |I| > 2. \end{cases}$$

Therefore (4.8) holds. By proposition 4.11, this implies that equality holds in (4.7), that is, with  $r = 2$ :  $|\partial_B H(x)| = 2 \sum_{i \in [0 : 1]} \binom{p-1}{i} = 2p$ .  $\square$

This result applies to the left-hand side example of figure 3.1.

*Another proof of (4.11) using the bipartitioning of section 3.3.2.* Let  $s \in \mathcal{S}$ , so that  $\text{cone}\{s_i v_i : i \in [1 : p]\}$ , called the original cone in this proof, is pointed by (3.10). Adopt the construction 3.13(1-2) associated with this sign vector  $s$ , as well as its notation. Since  $r = 2$ , the vectors  $\bar{v}_i$ 's are arranged along a line and one can assume that  $\bar{v}_1, \dots, \bar{v}_p$  follow each other in that order along this line (the  $\bar{v}_i$ 's are all different since the  $v_i$ 's are not colinear). The proposed proof consists in determining the complementary set of  $\mathcal{S}$  in  $\{\pm 1\}^p$ . By proposition 3.12, this amounts to identifying the partitions  $(I, J) \in \mathcal{P}([1 : p])$  such that  $K_I \cap K_J \neq \{0\}$  (then, the inversion of the vectors  $\{v_i\}_{i \in I}$  does not preserve the pointedness of the original cone, implying that  $(-s_I, s_J) \notin \mathcal{S}$  and  $\sigma^{-1}((-s_I, s_J)) \notin \partial_B H(x)$ ).

Clearly, any group of  $k$  vectors, with  $k \in [1 : p-1]$ , that is not one of the sets  $\{\bar{v}_1, \dots, \bar{v}_k\}$  and  $\{\bar{v}_{p-k+1}, \dots, \bar{v}_p\}$  cannot be linearly separated from the other vectors and there are

$$\binom{p}{k} - 2$$

such groups. Hence, the total number of groups of vectors, whose inversion does not preserve the pointedness of the origin cone, is

$$\sum_{k \in [1 : p-1]} \left[ \binom{p}{k} - 2 \right] = 2^p - 2 - 2(p-1) = 2^p - 2p, \quad (4.12)$$

where we have used  $\sum_{k \in [0 : p]} \binom{p}{k} = 2^p$  and  $\binom{p}{0} = \binom{p}{p} = 1$ . Since there are  $2^p$  subsets  $I$  of  $[1 : p]$  (including  $I = \emptyset$  and  $I = [1 : p]$ , describing the cases where there is no change of sign and  $p$  changes of signs, respectively), the total number of sign changes that preserve the pointedness of the original cone is  $2^p - (2^p - 2p) = 2p$ , which yields (4.11).  $\square$

## 5 Computation of the B-differential

This section describes techniques to compute a single element (section 5.1) or all the elements (section 5.2) of the B-differential  $\partial_B H(x)$ , in exact arithmetics, when  $H$  is the piecewise affine function given by (1.2). The complexity of the algorithm for the latter case is also analyzed. Let us mention that, once the B-differential is known, it is possible to verify whether a particular Jacobian  $J$  is in the C-differential  $\partial_C H(x)$  by checking whether  $J$  is a convex combination of the elements of  $\partial_B H(x)$ , which can be realized by solving a linear system with inequalities.

### 5.1 Computation of a single element

An interest of the problem equivalence highlighted in proposition 3.3 is to provide a method to find rapidly an element of  $\partial_B H(x)$ , which simplifies a little the one proposed by Liquan Qi [46; 1993, final remarks (1)] (or is just another option). It is shown in [25], that this method extends to the computation of an element of the B-differential in the nonlinear case, i.e., when  $H$  is given by (1.5). The method is based on the following algorithm, which associates with  $p$  nonzero vectors  $v_1, \dots, v_p$ , a direction  $d$  such that  $v_i^\top d \neq 0$  for all  $i \in [1:p]$ ; it is a variant of the technique used in the proof of [15; lemma 2.1]. It can also be derived from lemma 2.5, by adding the vector  $v_0 = 0$ .

**Algorithm 5.1 (computes  $d \in \mathbb{R}^n$  such that  $v_i^\top d \neq 0$  for all  $i$ )**

Data:  $p$  nonzero vectors  $v_1, \dots, v_p$  in  $\mathbb{R}^n$ .

Return: a direction  $d$  such that  $v_i^\top d \neq 0$  for all  $i \in [1:p]$ .

Take an arbitrary nonzero direction  $d \in \mathbb{R}^n$ .

Repeat:

1. If  $I := \{i \in [1:p] : v_i^\top d = 0\} = \emptyset$ , exit.
2. Let  $i \in I$ .
3. Take  $t > 0$  sufficiently small such that, for all  $j \notin I$ ,  $(v_j^\top d)(v_j^\top [d + tv_i]) > 0$ .
4. Update  $d := d + tv_i$ .

*Explanation.* In step 3, any sufficiently small  $t > 0$  is appropriate (the proof of [15; lemma 2.1] computes bounds explicitly), since  $(v_j^\top d)(v_j^\top [d + tv_i])$  is positive for  $t = 0$ . The new direction  $d$  set in step 4 is such that  $v_i^\top (d + tv_i) = t\|v_i\|^2 > 0$ , so that this direction makes at least one more  $v_j^\top d$  nonzero than the previous one. This implies that the algorithm finds an appropriate direction in at most  $p$  loops.  $\square$

The next procedure uses a direction  $d$  computed by algorithm 5.1 to obtain a single element of  $\partial_B H(x)$  when  $H$  is given by (1.2). Recall that the map  $\sigma$  is defined by (3.5) and is a bijection from  $\partial_B H(x)$  onto  $\mathcal{S}$ , defined by (3.3) (see proposition 3.3).

**Algorithm 5.2 (computes a single Jacobian in  $\partial_B H(x)$ )**

Let  $H$  be given by (1.2),  $x \in \mathbb{R}^n$  and suppose that  $p := |\mathcal{E}^\neq(x)| \neq 0$ .

1. Compute the nonzero vectors  $v_1, \dots, v_p$  in  $\mathbb{R}^n$  as the columns of the matrix  $V$  in (3.2).
2. By algorithm 5.1, compute a direction  $d \in \mathbb{R}^n$  such that  $v_i^\top d \neq 0$  for all  $i \in [1:p]$ .

3. Define  $s \in \mathcal{S}$  by  $s_i := \text{sgn}(v_i^\top d)$ , for  $i \in [1 : p]$ .
4. Then,  $\sigma^{-1}(s) \in \partial_B H(x)$ .

*Explanation.* When  $p = 0$ ,  $\partial_B H(x) = \bar{\partial}_B H(x)$  contains a single Jacobian that is given by (3.4), which explains why algorithm 5.2 focuses on the case when  $p > 0$ . In step 3,  $\text{sgn}(t)$  is the sign of  $t \in \mathbb{R} \setminus \{0\}$  ( $= +1$  if  $t > 0$ ,  $= -1$  if  $t < 0$ ). The sign vector  $s$  computed in step 3 is such that  $s \cdot V^\top d > 0$ , so that it is indeed in  $\mathcal{S}$  (defined in (3.3)) and, by proposition 3.3,  $\sigma^{-1}(s)$  is a Jacobian in  $\partial_B H(x)$ .  $\square$

## 5.2 Computation of all the elements

We present in this section preliminary results on several improvements to the Rada and Černý [48; 2018] algorithm for computing the cells of a hyperplane arrangement, which is known to be an equivalent problem to the one of computing the B-differential of  $H$  in (1.2), when the hyperplanes contain zero (see section 3.4).

### 5.2.1 Overview

Algorithms for listing all the elements of the finite set  $\partial_B H(x)$ , when  $H$  is given by (1.2), can be designed by looking at one of the various forms of the problem, those described in section 3 and others [5]. Actually, the only algorithms we have found take the point of view of the hyperplane arrangement of section 3.4 and can usually be used for more general arrangements than those needed to describe  $\partial_B H(x)$  (i.e., in which case the hyperplanes pass through zero). Let us quote the contributions by Bieri and Nef [11; 1982], by Edelsbrunner, O'Rourke and Seidel [26; 1986], by Avis and Fukuda [5; 1996], improved by Sleumer [53; 1998].

More recently, Rada and Černý [48; 2018] proposed an incremental/recursive algorithm, which seems to have the best performance, so far. It solves linear optimization problems to determine the presence of certain sign vectors in  $\mathcal{S}$ . The algorithm is briefly described in section 5.2.2. In section 5.2.3, we propose some improvements to their algorithm. The resulting algorithm is presented in section 5.2.4 and its complexity is analyzed in section 5.2.5. Preliminary numerical results are related in section 5.2.6 to show the degree of improvements brought by the modifications proposed in section 5.2.3.

### 5.2.2 Rada and Černý's algorithm

The algorithm proposed by Rada and Černý [48; 2018], which is referenced below as the RC algorithm, deals with the determination of the cells associated with a general hyperplane arrangement. We describe it below for an arrangement that results from the computation of the B-differential  $\partial_B H(x)$ , whose hyperplanes all contains zero (see section 3.4). We also uses the linear algebra language of section 3.2.1, by viewing the problem as the one determining the set  $\mathcal{S}$  of sign vectors  $s \in \{\pm 1\}^p$  such that  $s \cdot V^\top d > 0$  is feasible for  $d \in \mathbb{R}^n$  ( $V$  is the matrix defined by (3.2)); in contrast, the language used in [48] is more geometric.

Note that the RC algorithm in [48] deals with hyperplanes of the form  $\{d \in \mathbb{R}^n : v_i^\top d = 1\}$ , which may not have a point in common. Therefore, the description below is an adaptation of this algorithm to the case of the hyperplanes (3.15), all containing zero.

Let us shorten the notation by setting  $V_k := V_{:, [1:k]}$ , for  $k \in [1:p]$  and denote by  $\mathcal{S}_k$  the set of sign vectors  $s \in \{\pm 1\}^k$  such that  $s \cdot V_k^\top d > 0$  is feasible for  $d \in \mathbb{R}^n$ . The RC algorithm is incremental in the sense that, knowing a sign vector  $s \in \mathcal{S}_k$ , for some  $k \in [1:p-1]$ , it seeks to determine whether  $(s, \pm 1) \in \mathcal{S}_{k+1}$ . Since a direction  $d \in \mathbb{R}^n$  such that  $s \cdot V_k^\top d > 0$  is memorized, it suffices to look at the sign of  $v_{k+1}^\top d$  to see whether  $(s, 1)$  or  $(s, -1)$  is in  $\mathcal{S}_{k+1}$ . Indeed, one has

$$\begin{aligned} v_{k+1}^\top d < 0 &\implies (s, -1) \in \mathcal{S}_{k+1}, \\ v_{k+1}^\top d > 0 &\implies (s, +1) \in \mathcal{S}_{k+1}. \end{aligned}$$

In [48; Algorithm 1], the case where  $v_{k+1}^\top d = 0$  is not dealt with properly since  $(s, +1)$  is declared to belong to  $\mathcal{S}_{k+1}$  in that case. In our implementation of the RC algorithm, we modify slightly  $d$  by adding a small positive or negative multiple of  $v_{k+1}$  to  $d$  when  $v_{k+1}^\top d = 0$ , so that both  $(s, \pm 1)$  are accepted in  $\mathcal{S}_{k+1}$  in that case. One of the proposed modifications below consists in considering more carefully the case where  $v_{k+1}^\top d$  is near zero.

Next, when  $(s, s_{k+1}) \in \{\pm 1\}^{k+1}$  is accepted in  $\mathcal{S}_{k+1}$ , the question of whether  $(s, -s_{k+1})$  is also accepted in  $\mathcal{S}_{k+1}$  arises. In the RC algorithm, the answer to this question is obtained by solving the linear optimization problem

$$\begin{cases} \min_{(d,t) \in \mathbb{R}^n \times \mathbb{R}} t \\ s_i v_i^\top d \geq -t, \quad \forall i \in [1:k] \\ -s_{k+1} v_{k+1}^\top d \geq -t \\ t \geq -1. \end{cases} \quad (5.1)$$

This problem is feasible (take  $d = 0$  and  $t$  sufficiently large) and bounded (its optimal value is  $\geq -1$ ), so that it has a solution [12, 31]. If the optimal  $t$  is zero,  $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$ , otherwise, the optimal  $t$  is  $-1$  and  $(s, -s_{k+1}) \in \mathcal{S}_{k+1}$ . These LO solves are the most expensive part of the algorithm and in the numerical experiments of section 5.2.6, we measure the efficiency of the algorithms by the amount of LO solves they require.

The RC algorithm is recursive since, once the belonging of  $(s, s_{k+1})$  to  $\mathcal{S}_{k+1}$  is determined and  $k \leq p-2$ , the question whether  $(s, s_{k+1}, \pm 1)$  belongs to  $\mathcal{S}_{k+2}$  arises and the answer is obtained by calling the procedure just described recursively.

One can now formally describe our version of the RC algorithm (the change is in step 2, which is not considered in the original algorithm). It starts with an empty sign vector  $s$  and then calls recursively the procedure RCREC. The algorithm does not memorize the sign vectors in  $\mathcal{S}$  but print them in step 1 of the procedure RCREC.

**Algorithm 5.3 (RC)** Let be given  $V \in \mathbb{R}^{n \times p}$ , with  $n$  and  $p \in \mathbb{N}^*$ , having nonzero columns.

1. Set  $d = 0 \in \mathbb{R}^n$  and  $s = \emptyset$ .
2. Execute the recursive procedure RCREC( $V, d, s$ ).

**Algorithm 5.4 (RCREC)** Let be given  $V \in \mathbb{R}^{n \times p}$ , with  $n$  and  $p \in \mathbb{N}^*$ , having nonzero columns, a direction  $d \in \mathbb{R}^n$  and a sign vector  $s \in \{\pm 1\}^k$  for some  $k \in [0:p]$ .

1. If  $k = p$ , print  $s$  and return.
2. If  $v_{k+1}^\top d = 0$ , then



- 2.1. Execute  $\text{RCREC}(V, d_+, (s, +1))$ , where  $d_+ := d + t_+ v_{k+1}$  with  $t_+ > 0$  chosen in the nonempty open interval

$$\left( 0, \min_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} < 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}} \right).$$

- 2.2. Execute  $\text{RCREC}(V, d_-, (s, -1))$ , where  $d_- := d + t_- v_{k+1}$  with  $t_- < 0$  chosen in the nonempty open interval

$$\left( \max_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} > 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}}, 0 \right).$$

3. Else  $s_{k+1} := \text{sgn}(v_{k+1}^\top d)$ .
- 3.1. Execute  $\text{RCREC}(V, d, (s, s_{k+1}))$ .
- 3.2. Solve the LO problem (5.1) and denote by  $(d, t)$  a solution.  
If  $t = -1$ , execute  $\text{RCREC}(V, d, (s, -s_{k+1}))$ .
4. Return.

In step 2.1 and 2.2, the minimum is supposed to be infinite if its feasible set is empty. It is easy to see that the directions  $d_\pm$  computed in steps 2.1 and 2.2 are such that  $s_i v_i^\top d_\pm > 0$  for  $i \in [1:k+1]$  and  $s_{k+1} = \pm 1$ , which justifies the recursive call to  $\text{RCREC}$  with the given arguments. Actually, The recursive algorithm  $\text{RCREC}$  explores a tree, whose nodes are labelled by the sign vectors  $s$ 's of various dimensions, in which each node has one or two descendants.

### 5.2.3 Some improvements

#### *Taking the rank of $V$ into account*

Instead of starting with the empty vector  $s$ , one can take into account the rank  $r := \text{rank}(V)$  to determine  $2^r$  initial vectors  $s$ , hence avoiding to solve a number of LO problems, especially as  $r$  is large.

The algorithm selects  $r := \text{rank}(V)$  linearly independent vectors  $v_i$ , among the columns of  $V \in \mathbb{R}^{n \times p}$ . These vectors can be obtained by a QR factorization of

$$VP = QR,$$

where  $P$  is a permutation matrix,  $Q \in \mathbb{R}^{n \times n}$  is orthogonal (i.e.,  $Q^\top Q = I_n$ ) and  $R \in \mathbb{R}^{n \times p}$  is upper triangular with  $R_{[r+1:n],:} = 0$ . To simplify the presentation, one can assume, without loss of generality, that  $P = I$ , in which case the vectors  $v_1, \dots, v_r$  are linearly independent (in practice, the vectors are symbolically reordered by using the permutation matrix  $P$ ). By proposition 4.2,

$$\mathcal{S}_r = \{\pm 1\}^r. \quad (5.2)$$

Furthermore, for each  $s \in \mathcal{S}_r$ , we have, using  $S := \text{Diag}(s)$ ,  $Q_r := Q_{:, [1:r]}$  and  $R_r := R_{[1:r], [1:r]}$ , that the vector

$$d_s := Q_r (R_r S)^{-\top} e = Q_r R_r^{-\top} S e = Q_r R_r^{-\top} s \quad (5.3)$$



is such that  $s \cdot V_{:, [1:r]}^\top d_s = e > 0$ , as desired.

For each  $s \in \mathcal{S}_r$  and the associated  $d_s$  computed above, the algorithm can run the recursive algorithm  $\text{RCREC}(V, d_s, s)$ .

**Special handling of the case where  $v_{k+1}^\top d \simeq 0$**

Directions  $d_\pm := d + t_\pm v_{k+1}$  ensuring that  $(s, \pm 1) \cdot V_{k+1}^\top d_\pm > 0$  can be computed not only when  $v_{k+1}^\top d = 0$  like in step 2 of the  $\text{RCREC}$  recursive algorithm, but also when  $v_{k+1}^\top d \simeq 0$  in the sense (5.4) below. Note that the left-hand side in (5.4) is negative and the right-hand side is positive (this can be seen by multiplying numerators and denominators by  $s_i$  and by using  $s_i v_i^\top d > 0$  for all  $i \in [1:k]$ ), so that these inequalities are verified when  $v_{k+1}^\top d = 0$ . With the additional flexibility that (5.4) offers, the algorithm sometimes can avoid solving a significant number of LO problems of the form (5.1).

**Lemma 5.5 (two more signs without optimization)** *Suppose that  $s \in \{\pm 1\}^k$  verifies  $s \cdot V_k^\top d > 0$ , that  $v_{k+1} \neq 0$  and that*

$$\max_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} > 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}} < \frac{-v_{k+1}^\top d}{\|v_{k+1}\|^2} < \min_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} < 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}}. \quad (5.4)$$

1) If  $t_+$  is chosen in the nonempty open interval

$$\left( \frac{-v_{k+1}^\top d}{\|v_{k+1}\|^2}, \min_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} < 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}} \right) \quad (5.5)$$

then,  $d_+ := d + t_+ v_{k+1}$  verifies  $s \cdot V_k^\top d_+ > 0$  and  $v_{k+1}^\top d_+ > 0$ .

2) If  $t_-$  is chosen in the nonempty open interval

$$\left( \max_{\substack{i \in [1:k] \\ s_i v_i^\top v_{k+1} > 0}} \frac{-v_i^\top d}{v_i^\top v_{k+1}}, \frac{-v_{k+1}^\top d}{\|v_{k+1}\|^2} \right). \quad (5.6)$$

then,  $d_- := d + t_- v_{k+1}$  verifies  $s \cdot V_k^\top d_- > 0$  and  $-v_{k+1}^\top d_- > 0$ .

*Proof.* Clearly, (5.4) implies that the open intervals (5.5) and (5.6) are nonempty.

One can certainly take  $(s, +1)$  in  $\mathcal{S}_k$  if one can find a  $t_+ \in \mathbb{R}$  satisfying

$$\left( s_i v_i^\top (d + t_+ v_{k+1}) > 0, \quad \forall i \in [1:k] \right) \quad \text{and} \quad v_{k+1}^\top (d + t_+ v_{k+1}) > 0. \quad (5.7a)$$

It is easy to check that a  $t_+$  taken in the interval (5.5) ensures (5.7a).

Similarly, one can certainly take  $(s, -1)$  in  $\mathcal{S}_k$  if one can find a  $t_- \in \mathbb{R}$  satisfying

$$\left( s_i v_i^\top (d + t_- v_{k+1}) > 0, \quad \forall i \in [1:k] \right) \quad \text{and} \quad -v_{k+1}^\top (d + t_- v_{k+1}) > 0. \quad (5.7b)$$

It is easy to check that a  $t_-$  taken in the interval (5.6) ensures (5.7b).  $\square$

### 5.2.4 ISF algorithm

We have named ISF (for Incremental Signed Feasibility), the algorithm that modifies the RC algorithm by incorporating the modifications described in section 5.2.3. For the purpose of precision and reference, we state it in this section. It starts with a hat procedure ISF, similar to that of the RC algorithm but with the easy determination of  $\mathcal{S}_r$ , where  $r := \text{rank}(V)$ . Then the hat procedure calls the recursive procedure ISFREC. The structure of the algorithm is very similar to the one of the RC algorithm 5.3.

**Algorithm 5.6 (ISF)** Let be given  $V \in \mathbb{R}^{n \times p}$ , with  $n$  and  $p \in \mathbb{N}^*$ , having nonzero columns.

1. Compute  $\mathcal{S}_r$  by (5.2).
2. For each  $s \in \mathcal{S}_r$  and its associated  $d_s$ , given by (5.3), call the recursive procedure ISFREC( $V, d_s, s$ ).

**Algorithm 5.7 (ISFREC)** Let be given  $V \in \mathbb{R}^{n \times p}$ , with  $n$  and  $p \in \mathbb{N}^*$ , having nonzero columns, a direction  $d \in \mathbb{R}^n$  and a sign vector  $s \in \{\pm 1\}^k$  for some  $k \in [r:p]$ , where  $r := \text{rank}(V)$ .

1. If  $k = p$ , print  $s$  and return.
2. If (5.4) holds, then
  - 2.1. Execute ISFREC( $V, d_+, (s, +1)$ ), where  $d_+ := d + t_+ v_{k+1}$  with  $t_+$  is chosen in the nonempty open interval (5.5).
  - 2.2. Execute ISFREC( $V, d_-, (s, -1)$ ), where  $d_- := d + t_- v_{k+1}$  with  $t_- < 0$  chosen in the nonempty open interval (5.6).
3. Else  $s_{k+1} := \text{sgn}(v_{k+1}^\top d)$ .
  - 3.1. Execute ISFREC( $V, d, (s, s_{k+1})$ ).
  - 3.2. Solve the LO problem (5.1) and denote by  $(d, t)$  a solution.  
If  $t = -1$ , execute ISFREC( $V, d, (s, -s_{k+1})$ ).
4. Return.

### 5.2.5 Complexity

As shown by the next proposition, the ISF algorithm is *output-sensitive*, in the sense that its computation effort is bounded above by a number proportional to the size  $|\partial_B H(x)|$  of the output (this one may be exponential). This upper bound is also proportional to the computational effort required by the linear optimization solver used by the algorithm.

The next result is very similar to [48; theorem 3.2], although the proof is slightly different. The given upper bound on the number of LO solves takes also into account the step 1 of algorithm 5.6, which is not present in the RC algorithm and from which comes the use of the rank  $r$  of  $V$ . The bound  $2^p - 2^r$  is also new, but is unlikely to be active.

**Proposition 5.8 (complexity of algorithm 5.6)** Let  $V \in \mathbb{R}^{n \times p}$  and  $r := \text{rank}(V)$ .

Then, the number of linear optimization problems solved by the incremental algorithm 5.6 does not exceed

$$\min(2^p - 2^r, (p - r)|\mathcal{S}|). \quad (5.8)$$

*Proof.* In its step 1, algorithm 5.6 specifies the  $2^r$  sign vectors associated with  $r$  linearly independent columns of  $V$ , which may be assumed to be  $v_1, \dots, v_r$ . This step does not require to solve any LO problem. One has

$$|\mathcal{S}_r| = 2^r. \quad (5.9a)$$

Next, algorithm 5.7 considers the remaining  $p - r$  columns  $v_{r+1}, \dots, v_p$  of  $V$  recursively. Let us show that

$$\text{for } k \in [r + 1 : p]: \quad |\mathcal{S}_{k-1}| \subseteq |\mathcal{S}_k| \subseteq |\mathcal{S}|. \quad (5.9b)$$

The first inclusion in (5.9b) comes from the fact that algorithm 5.7 builds a sign vector in  $\mathcal{S}_k$  for each sign vector in  $\mathcal{S}_{k-1}$ , in any of its step 2 or 3. The second inclusion in (5.9b) is deduced by induction and from the fact that  $\mathcal{S}_p = \mathcal{S}$ .

To compute  $\mathcal{S}_k$ , for any  $k \in [r + 1 : p]$ , the number of LO problems to solve is bounded by  $|\mathcal{S}_{k-1}|$ , since at most one LO problem is solved (in step 3 of the algorithm) for each sign vector in  $\mathcal{S}_{k-1}$ . This bound  $|\mathcal{S}_{k-1}|$  is itself bounded by  $\min(2^{k-1}, |\mathcal{S}|)$  (the first bound comes from the fact that there are  $k - 1$  vectors in  $V_{:, [1:k-1]}$ , while the second bound comes from (5.9b)). Therefore, the total number of LO problems solved by the algorithm is bounded by

$$\min(2^r, |\mathcal{S}|) + \min(2^{r+1}, |\mathcal{S}|) + \dots + \min(2^{p-1}, |\mathcal{S}|) \leq \min(2^p - 2^r, (p - r)|\mathcal{S}|).$$

This is the announced bound (5.8). □

## 5.2.6 Numerical results

### *Problem and table presentation*

We present in table 5.1 the result obtained by running the ISF algorithm 5.6 on a small number of problems and compare it with our implementation of the RC algorithm 5.3, simulating the algorithm 1 (IE) in [48]. The implementations have been done in `Matlab` (R2017a).

In addition to academic problems that are not reported here, we have tested the codes on randomly generated data (named `rand` in the first part of the table) and problems adapted from [48] (their name is prefixed by `rc-`). The random data is formed of a randomly generated matrix  $V \in \mathbb{R}^{n \times p}$  with prescribed rank  $r := \text{rank}(V)$ . The problems adapted from [48] are given in the second part of the table.

The dimensions  $n$ ,  $p$  and  $r$  of the problems are given in columns 2-4 of the table. In column 5 and 6, one finds the Winder upper bound (the right-hand side of (4.7)) and the cardinal  $|\partial_B H(x)| = |\mathcal{S}|$  of the B-differential  $\partial_B H(x)$ . The RC and ISF algorithms are compared on the number of LO problems they require to solve, which is a good image of their computation effort, measured independently of the computer used to run the codes and the features of the LO solver. In columns 7 and 8, one finds the number of LO solves required by the RC and ISF algorithms respectively. Column 9 gives the number of LO problems (5.1) that could be solved inexactly (because the optimal value is  $-1$  and this fact can be detected as soon as

the objective value becomes negative along the iterations). When this number is large with respect to the total number of LO solves, this means that an inexact solve of the LO problem can improve the efficiency of the piece of software significantly. Finally, the ratio between the number in columns 7 and 8 represents the gain in efficiency of the ISF algorithm over the RC algorithm, when this value is  $> 1$ .

### Observations

Here are some remarks on the numerical results.

- 1) The randomly generated problems `rand` are likely to provide vectors  $v_i$ 's (the columns of  $V$ ) in general position, in the sense defined before proposition 4.11, which is that condition (4.8) holds. This can be seen indirectly on the numbers in the table. Indeed, it is known from proposition 4.11 that (4.8) implies equality in (4.7). This equality indeed holds, as we can observe by comparing columns 5 and 6.

Incidentally, one can compute mentally Winder's bound  $\beta$  when  $p$  is even and  $r = p/2$ . In that case, the right-hand side of (4.7) reads

$$\beta = 2 \sum_{i \in [0:r-1]} \binom{2r-1}{i} = \sum_{i \in [0:2r-1]} \binom{2r-1}{i} = s^{2r-1} = 2^{p-1}.$$

This is what one observes in the table; for example when  $p = 8$  and  $r = 4$ , one has  $\beta = 128$ , which is indeed  $2^{8-1}$ .

- 2) One also observes that when  $p = r$ , one has  $|\partial_B H(x)| = 2^p$  (see the value (4.10) given for an injective matrix  $V$ ) and, when  $r = 2$ , one has  $|\partial_B H(x)| = 2p$  (see the value (4.11) given for a fan disposition).
- 3) A comparison between the "Original RC code" in `Python` and its "Simulated version" in `Matlab` shows that the latter is slightly more effective, probably due to the special treatment in step 2 of the case where  $v_{k+1}^\top d = 0$  in algorithm 5.4, which is not considered in the original code.
- 4) The main lesson from the table is that the "profit" in the last column is generally significantly larger than 1 (there is only one exception), meaning that the ISF algorithm is significantly faster than the RC algorithm.
- 5) Another observation is that the number of LO problems that could be solved inexactly (last but one column) is very often a large fraction of the total number of LO solves. This means that one could gain much in solving these problems inexactly.  $\square$

### 5.2.7 A numerical example

Consider the LCP in standard form (1.4), which reads  $0 \leq x \perp (Mx + q) \geq 0$ , where  $x \in \mathbb{R}^n$ ,  $M \in \mathbb{R}^{n \times n}$  and  $q \in \mathbb{R}^n$ . Suppose that  $n = 3$  and that  $M$  and  $q$  are given by

$$M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad \text{and} \quad q = 0.$$

Since  $M$  is a **P**-matrix (i.e., all its principal minors are positive), the problem has a unique solution [52], which is  $x = 0$ . With the notation (3.1), one has  $\mathcal{A}(x) = \mathcal{B}(x) = \emptyset$  and  $\mathcal{E}(x) = \mathcal{E}^\neq(x) = \{1, 2, 3\}$ , so that  $V^\top$  given by (3.2) reads

$$V^\top = M - I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Problem	$n$	$p$	$r$	Winder's bound	$ \partial_B H(x) $	LO solves in				
						Original RC	Simulated RC	Total	Inexact	Profit
rand	4	4	4	16	16		28	0	0	1.33
rand	4	8	2	16	16		7	21	0	—
rand	7	8	4	128	128		98	54	19	1.81
rand	7	9	4	186	186		162	102	32	1.59
rand	7	10	5	512	512		381	240	114	1.59
rand	7	11	4	352	352		385	278	68	1.38
rand	7	12	6	2048	2048		1485	1056	594	1.41
rand	7	13	5	1588	1588		1585	1242	450	1.28
rand	7	14	7	8192	8192		5811	4222	2506	1.38
rand	8	15	7	12952	12952		9907	7553	4121	1.31
rand	9	16	8	32768	32768		22818	17365	10930	1.31
rand	10	17	9	78406	78406		50642	39101	27661	1.30
rc-2d-20-4	4	19	4	1976	136	616	545	480	0	1.14
rc-2d-20-5	5	20	5	10072	272	1096	1091	960	0	1.14
rc-2d-20-6	6	20	6	33328	512	1936	1927	1680	0	1.15
rc-2d-20-7	7	20	7	87592	960	3392	3343	2912	0	1.15
rc-2d-20-8	8	20	8	188368	1792	5888	5855	4992	0	1.17
rc-perm-5	5	15	5	2942	720	1211	1049	851	0	1.23
rc-perm-6	6	21	6	43400	5040	10417	9285	7898	1	1.18
rc-perm-8	8	36	8	17463696	362880	1036897	952597	856597	1140	1.11
rc-ratio-20-3-7	3	19	3	344	304	929	928	839	62	1.11
rc-ratio-20-3-9	3	19	3	344	178	539	538	541	27	0.99
rc-ratio-20-4-7	4	20	4	2320	2278	4954	4953	4522	710	1.10
rc-ratio-20-4-9	4	20	4	2320	2016	4393	4392	4231	583	1.04
rc-ratio-20-5-7	5	20	5	10072	8470	13798	13788	12979	3091	1.06
rc-ratio-20-5-9	5	20	5	10072	7826	13798	13797	12220	2467	1.13
rc-ratio-20-6-7	6	20	6	33328	26194	32993	32992	31967	8681	1.03
rc-ratio-20-6-9	6	20	6	33328	26758	39823	39822	36485	9840	1.09
rc-ratio-20-7-7	7	20	7	87592	76790	82751	82750	76158	30725	1.09
rc-ratio-20-7-9	7	20	7	87592	58468	70214	70212	51974	20829	1.35

Table 5.1: Comparison of the “simulated RC algorithm 5.3” and the “proposed ISF algorithm” on a few problems: “ $(n, p, r)$ ” are the dimensions of the problem ( $V \in \mathbb{R}^{n \times p}$  is of rank  $r$ ), “Winder’s bound” is the right-hand side of (4.7), “ $|\partial_B H(x)|$ ” is the cardinal of the B-differential of  $H$  given by (1.2), “Original RC” gives the number of LO solves for the original piece of software in Python of Rada and Černý [48], “Simulated RC” gives the number of LO solves with our implementation in Matlab of the Rada and Černý algorithm (see algorithm 5.3), for the proposed ISF algorithm, “Total” gives the number of LO solves, “Inexact” is the number of LO problems that could be solve inexactly and “Profit” is the ratio between the number of LO solves of the simulated-RC and ISF algorithms.

Note that  $p := |\mathcal{E}^\neq(x)| = 3$ , while  $\text{rank}(V) = 2$ , so that  $|\partial_B H(x)| = 2p = 6$ , by (4.11). The sign vectors  $s \in \{\pm 1\}^3$  that make  $s \cdot V^\top d > 0$  feasible for  $d$  are gathered in the set denoted by  $\mathcal{S}$ , are the columns of the matrix  $S$  below and possible feasible directions  $d \in \mathbb{R}^3$  are the columns of the matrix  $D$ :

$$S = \begin{pmatrix} 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 1 & -1 & 2 & -2 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

The Jacobians of the B-differential  $\partial_B H(x)$ , where  $H$  is given by (1.2), are obtained for the  $s$ 's in  $\mathcal{S}$  given above by the bijection  $\sigma$  defined by (3.5). One gets a set of 6 Jacobians out of the  $2^3 = 8$  Jacobians in  $\bar{\partial}_B H(x)$ , namely

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

## 6 References

- [1] M. Aganagić (1984). Newton's method for linear complementarity problems. *Mathematical Programming*, 28, 349–362. [\[doi\]](#). 3
- [2] Marcelo Aguiar, Swapneel Mahajan (2017). *Topics in hyperplane Arrangements*. Mathematical Surveys and Monographs 226. American Mathematical Society, Providence, RI. [\[doi\]](#). 16, 17, 20, 22
- [3] G.L. Alexanderson, John E. Wetzel (1981). Arrangements of planes in space. *Discrete Mathematics*, 34(3), 219–240. [\[doi\]](#). 22
- [4] G.L. Alexanderson, John E. Wetzel (1983). Erratum: “arrangements of planes in space”. *Discrete Mathematics*, 45(1), 140. [\[doi\]](#). 22
- [5] David Avis, Komei Fukuda (1996). Reverse search for enumeration. *Discrete Applied Mathematics*, 65(1-3), 21–46. [\[doi\]](#). 8, 17, 29
- [6] Pierre Baldi, Roman Vershynin (2019). Polynomial threshold functions, hyperplane arrangements, and random tensors. *SIAM Journal on Mathematics of Data Science*, 1(4), 699–729. [\[doi\]](#). 8
- [7] I. Ben Gharbia, J.Ch. Gilbert (2012). Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a  $P$ -matrix. *Mathematical Programming*, 134, 349–364. [\[doi\]](#). 3
- [8] I. Ben Gharbia, J.Ch. Gilbert (2013). An algorithmic characterization of  $P$ -matricity. *SIAM Journal on Matrix Analysis and Applications*, 34(3), 904–916. [\[doi\]](#). 3
- [9] I. Ben Gharbia, J.Ch. Gilbert (2019). An algorithmic characterization of  $P$ -matricity II: adjustments, refinements, and validation. *SIAM Journal on Matrix Analysis and Applications*, 40(2), 800–813. [\[doi\]](#). 3
- [10] I. Ben Gharbia, J. Jaffré (2014). Gas phase appearance and disappearance as a problem with complementarity constraints. *Mathematics and Computers in Simulation*, 99, 28–36. [\[doi\]](#). 3
- [11] H. Bieri, W. Nef (1982). A recursive sweep-plane algorithm, determining all cells of a finite division of  $\mathbb{R}^d$ . *Computing*, 28(3), 189–198. [\[doi\]](#). 29
- [12] J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, C. Sagastizábal (2006). *Numerical Optimization – Theoretical and Practical Aspects* (second edition). Universitext. Springer Verlag, Berlin. [\[authors\]](#) [\[editor\]](#) [\[doi\]](#). 30
- [13] J.M. Borwein, A.S. Lewis (2006). *Convex Analysis and Nonlinear Optimization – Theory and Examples* (second edition). CMS Books in Mathematics 3. Springer, New York. 3, 12
- [14] Michal Černý, Miroslav Rada, Jaromír Antoch, Milan Hladík (2022). A class of optimization problems motivated by rank estimators in robust regression. *Optimization*, 71(8), 2241–2271. [\[doi\]](#). 16
- [15] X. Chen, S. Xiang (2011). Computation of generalized differentials in nonlinear complementarity problems. *Computational Optimization and Applications*, 50, 403–423. [\[doi\]](#). 4, 5, 18, 28
- [16] F.H. Clarke (1983). *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York. Reprinted in 1990 by SIAM, Classics in Applied Mathematics 5 [\[doi\]](#). 3, 4, 5
- [17] Kenneth L. Clarkson, Peter W. Shor (1995). Applications of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4, 387–421. [\[doi\]](#). 22

- [18] R.W. Cottle, G.B. Dantzig (1970). A generalization of the linear complementarity problem. *Journal of Combinatorial Theory*, 8(1), 79–90. [3](#)
- [19] R.W. Cottle, J.-S. Pang, R.E. Stone (2009). *The Linear Complementarity Problem*. Classics in Applied Mathematics 60. SIAM, Philadelphia, PA, USA. [\[doi\]](#). [3](#)
- [20] T. De Luca, F. Facchinei, C. Kanzow (2000). A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. *Computational Optimization and Applications*, 16, 173–205. [\[doi\]](#). [6](#)
- [21] J.-P. Dussault, M. Frappier, J.Ch. Gilbert (2019). A lower bound on the iterative complexity of the Harker and Pang globalization technique of the Newton-min algorithm for solving the linear complementarity problem. *EURO Journal on Computational Optimization*, 7(4), 359–380. [\[doi\]](#). [3](#)
- [22] J.-P. Dussault, M. Frappier, J.Ch. Gilbert (2019). Polyhedral Newton-min algorithms for complementarity problems. Technical report. [\[hal-02306526, document\]](#). [3](#)
- [23] J.-P. Dussault, J.Ch. Gilbert (2022). Exact computation of an error bound for the balanced linear complementarity problem with unique solution. *Mathematical Programming*. [\[doi\]](#). [3](#)
- [24] J.-P. Dussault, J.Ch. Gilbert, B. Plaquet-Jourdain (2022). On the B-differential of the componentwise minimum of two affine vectorial functions. Research report (to appear). [1](#), [4](#)
- [25] J.-P. Dussault, J.Ch. Gilbert, B. Plaquet-Jourdain (2023). On the B-differential of the componentwise minimum of two vectorial functions. Research report (to appear). [18](#), [28](#)
- [26] H. Edelsbrunner, J. O’Rourke, R. Seidel (1986). Constructing arrangements of lines and hyperplanes with applications. *SIAM Journal on Control*, 15(2), 341–363. [\[doi\]](#). [16](#), [29](#)
- [27] Herbert Edelsbrunner (1987). *Algorithms in Combinatorial Geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin. [\[doi\]](#). [16](#), [22](#)
- [28] L.C. Evans, R.F. Gariepy (2015). *Measure Theory and Fine Properties of Functions* (revised edition). CRC Press, Boca Raton. [3](#)
- [29] F. Facchinei, J.-S. Pang (2003). *Finite-Dimensional Variational Inequalities and Complementarity Problems* (two volumes). Springer Series in Operations Research. Springer. [3](#), [4](#)
- [30] H. Federer (1996). *Geometric Measure Theory*. Grundlehren der mathematischen Wissenschaften 153. Springer-Verlag, New York. [3](#)
- [31] J.Ch. Gilbert (2022). *Selected Topics on Continuous Optimization – Version 2*. Lecture notes of the Master-2 “Optimization” at the University Paris-Saclay. [\[hal\]](#). [30](#)
- [32] P. Gordan (1873). Über die Auflösung linearer Gleichungen mit reellen Coefficienten. *Mathematische Annalen*, 6, 23–28. [11](#)
- [33] Branko Grünbaum (1972). *Arrangements and Spreads*. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics 10. AMS, Providence, RI. [16](#), [22](#)
- [34] Dan Halperin, Micha Sharir (2018). Arrangements. In Jacob E. Goodman, Joseph O’Rourke, Csaba D. Tóth (editors), *Handbook of Discrete and Computational Geometry* (third edition), Discrete Mathematics its Applications, pages 723–762. CRC Press - Taylor & Francis Group. [16](#)
- [35] Juha Heinonen (2005). Lectures on Lipschitz Analysis. Report 100, Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä. [3](#)
- [36] J.-B. Hiriart-Urruty, C. Lemaréchal (2001). *Fundamentals of Convex Analysis*. Springer. [3](#)
- [37] A.F. Izmailov, M.V. Solodov (2014). *Newton-Type Methods for Optimization and Variational Problems*. Springer Series in Operations Research and Financial Engineering. Springer. [\[doi\]](#). [3](#)
- [38] C. Kanzow, M. Fukushima (1998). Solving box constrained variational inequalities by using the natural residual with D-gap function globalization. *Operations Research Letters*, 23(1-2), 45–51. [\[doi\]](#). [6](#)
- [39] M. Kojima, S. Shindo (1986). Extension of Newton and quasi-Newton methods to systems of  $PC^1$  equations. *Journal of Operations Research Society of Japan*, 29, 352–375. [\[doi\]](#). [3](#)
- [40] E. Marchand, T. Müller, P. Knabner (2012). Fully coupled generalised hybrid-mixed finite element approximation of two-phase two-component flow in porous media. Part II: numerical scheme and numerical results. *Computational Geosciences*, 16(3), 691–708. [\[doi\]](#). [3](#)
- [41] E. Marchand, T. Müller, P. Knabner (2013). Fully coupled generalised hybrid-mixed finite element approximation of two-phase two-component flow in porous media. Part I: formulation and properties of the mathematical model. *Computational Geosciences*, 17(2), 431–442. [\[doi\]](#). [3](#)
- [42] K.G. Murty (1988). *Linear Complementarity, Linear and Nonlinear Programming* (Internet edition, prepared by Feng-Tien Yu, 1997). Heldermann Verlag, Berlin. [3](#)



- [43] J.-S. Pang (1990). Newton’s method for B-differentiable equations. *Mathematics of Operations Research*, 15, 311–341. [\[doi\]](#). 3
- [44] J.-S. Pang (1991). A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems. *Mathematical Programming*, 51(1-3), 101–131. [\[doi\]](#). 3
- [45] J.-S. Pang (1995). Complementarity problems. In R. Horst, P.M. Pardalos (editors), *Handbook of Global Optimization*, volume 2 of *Nonconvex Optimization and Its Applications*, pages 271–338. Kluwer, Dordrecht. [\[doi\]](#). 3
- [46] Liqun Qi (1993). Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of Operations Research*, 18, 227–244. [\[doi\]](#). 3, 5, 28
- [47] Liqun Qi, Jie Sun (1993). A nonsmooth version of Newton’s method. *Mathematical Programming*, 58, 353–367. [\[doi\]](#). 3
- [48] Miroslav Rada, Michal Černý (2018). A new algorithm for enumeration of cells of hyperplane arrangements and a comparison with Avis and Fukuda’s reverse search. *SIAM Journal on Discrete Mathematics*, 32(1), 455–473. [\[doi\]](#). 4, 29, 30, 33, 34, 36
- [49] H. Rademacher (1919). Über partielle und totale differenzierbarkeit. *I. Math. Ann.*, 89, 340–359. 3
- [50] Samuel Roberts (1887/88). On the figures formed by the intercepts of a system of straight lines in a plane, and on analogous relations in space of three dimensions. *Proc. London Math. Soc.*, 19, 405–422. [\[doi\]](#). 16, 22
- [51] R.T. Rockafellar (1970). *Convex Analysis*. Princeton Mathematics Ser. 28. Princeton University Press, Princeton, New Jersey. 3, 15, 21
- [52] H. Samelson, R.M. Thrall, O. Wesler (1958). A partition theorem for the Euclidean  $n$ -space. *Proceedings of the American Mathematical Society*, 9, 805–807. [\[editor\]](#). 35
- [53] Nora Helena Sleumer (1998). Output-sensitive cell enumeration in hyperplane arrangements. In *Algorithm theory-SWAT’98 (Stockholm)*, Lecture Notes in Comput. Sci. 1432, pages 300–309. Springer, Berlin. [\[doi\]](#). 17, 29
- [54] Nora Helena Sleumer (2000). *Hyperplane arrangements – Construction, visualization and application*. PhD Thesis, Swiss Federal Institute of Technology, Zurich, Switzerland. [\[doi\]](#). 16
- [55] R.P. Stanley (2007). An introduction to hyperplane arrangements. In Ezra Miller, Victor Reiner, Bernd Sturmfels (editors), *Geometric Combinatorics*, pages 389–496. [\[doi\]](#). 16, 22
- [56] J. Steiner (1826). Einige Gesetze über die Theilung der Ebene und des Raumes. *Journal für die Reine und Angewandte Mathematik*, 1, 349–364. [\[doi\]](#). 16, 22
- [57] R. Sznajder, M.S. Gowda (1994). The generalized order linear complementarity problem. *SIAM Journal on Matrix Analysis and Applications*, 15(3), 779–795. [\[doi\]](#). 3
- [58] Michel Las Vergnas (1975). Matroïdes orientables. *C. R. Acad. Sci. Paris Série A-B*, 280, A61–A64. 22
- [59] Robert O. Winder (1966). Partitions of  $N$ -space by hyperplanes. *SIAM Journal on Applied Mathematics*, 14(4), 811–818. [\[doi\]](#). 8, 22, 24
- [60] Thomas Zaslavsky (1975). Facing up to arrangements: face-count formulas for partitions of space by hyperplanes. *Memoirs of the American Mathematical Society*, Volume 1, Issue 1, Number 154. [\[doi\]](#). 13, 22
- [61] C. Zhang, X. Chen, N. Xiu (2009). Global error bounds for the extended vertical LCP. *Computational Optimization and Applications*, 42(3), 335–352. [\[doi\]](#). 3