HAL open science

On the B-differential of the componentwise minimum of two affine vector 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 vector functions - The full report. Inria de Paris, France & Université de Sherbrooke, Canada. 2024. hal-03872711v3

HAL Id: hal-03872711

https://hal.science/hal-03872711v3

Submitted on 15 Apr 2024 (v3), 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.

Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

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

Jean-Pierre Dussault †, Jean Charles Gilbert ‡ and Baptiste Plaquevent-Jourdain §

(Monday 15th April, 2024, 16:00)

This paper focuses on the description and computation of the B-differential of the componentwise minimum of two affine vector functions. This issue arises in the reformulation of the linear complementarity problem with the Min C-function. The question has many equivalent formulations and we identify some of them in linear algebra, convex analysis and discrete geometry. These formulations are used to state some properties of the B-differential, like its symmetry, condition for its completeness, its connectivity, bounds on its cardinality, *etc.* The set to specify has a finite number of elements, which may grow exponentially with the range space dimension of the functions, so that its description is most often algorithmic. We first present an incremental-recursive approach avoiding to solve any optimization subproblem, unlike several previous approaches. It is based on the notion of matroid circuit and the related introduced concept of stem vector. Next, we propose modifications, adapted to the problem at stake, of an algorithm introduced by Rada and Černý in 2018 to determine the cells of an arrangement in the space of hyperplanes having a point in common. Measured in CPU time on the considered test-problems, the mean acceleration ratios of the proposed algorithms, with respect to the one of Rada and Černý, are in the range 15..31, and this speed-up can exceed 100, depending on the problem and the approach.

Keywords: B-differential • Bipartition of a finite set • C-differential • Complementarity problem • Complexity • Componentwise minimum of functions • Connectivity • Dual approach • Gordan's alternative • Hyperplane arrangement • Matroid circuit • Pointed cone • Schläfli's bound • Stem vector • 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 [30]. It contains additional material, including proofs and comments. The added text is written in dark blue.

†Département d'Informatique, Faculté des Sciences, Université de Sherbrooke, Québec, Canada (e-mail: `Jean-Pierre.Dussault@Usherbrooke.ca`). ORCID 0000-0001-7253-7462.

‡Inria Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France (e-mail: `Jean-Charles.Gilbert@inria.fr`) and Département de Mathématiques, Faculté des Sciences, Université de Sherbrooke, Québec, Canada. ORCID 0000-0002-0375-4663.

§Département de Mathématiques, Faculté des Sciences, Université de Sherbrooke, Québec, Canada (e-mail: `Baptiste.Plaquevent-Jourdain@USherbrooke.ca`) and 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

1 Introduction **3**

2 Background **5**

3 Equivalent problems **9**

3.1 B-differential of the minimum of two affine functions 9

3.2 Linear algebra problems 10

3.2.1 Signed feasibility of strict inequality systems 10

3.2.2 Orthants encountered by the null space of a matrix 12

3.3 Convex analysis problems 15

3.3.1 Pointed cones by vector inversions 15

3.3.2 Linearly separable bipartitions of a finite set 16

3.4 Discrete geometry: hyperplane arrangements 19

4 Description of the B-differential **21**

4.1 Some properties of the B-differential 21

4.2 Cardinality of the B-differential 26

4.2.1 Winder's formula 26

4.2.2 Bounds 29

4.3 Particular configurations 34

4.4 A glance at the C-differential 35

5 Computation of the B-differential **37**

5.1 Computation of a single Jacobian 37

5.2 Computation of all the Jacobians 38

5.2.1 Brute-force algorithm 38

5.2.2 Incremental-recursive algorithms 39

5.2.3 An algorithm using stem vectors 41

5.2.4 Linear optimization problem and stem vector 42

5.2.5 Improvements of the RC and STEM algorithms 44

5.2.6 ISF algorithm 47

5.2.7 Complexity 48

5.2.8 Numerical experiments 49

5.2.9 A numerical example 56

6 Discussion **57**

Acknowledgments **57**

References **58**

1 Introduction

Let $\mathbb{E}$ and $\mathbb{F}$ be two real vector spaces of finite dimensions $n := \dim \mathbb{E}$ and $m := \dim \mathbb{F}$. The *B-differential* (B for Bouligand [70]) *at* $x \in \mathbb{E}$ of a function $H : \mathbb{E} \to \mathbb{F}$ is the set denoted and defined by

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

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 [38, 41, 48, 67]) and this property has the consequence that the B-differential of a locally Lipschitz function is nonempty and bounded everywhere [20]. The B-differential is an intermediate set used to define the C-differential (C for Clarke [20]) of H at x, which is denoted and defined by

$$\partial_C H(x) := \operatorname{co} \partial_B H(x), \quad (1.2)$$

where $\operatorname{co} S$ denotes the convex hull of a set S [16, 49, 71]. Both intervene in the specification of conditions ensuring the local convergence of the semismooth Newton algorithm [64, 65, 77], which can be a motivation for being interested in that concept.

In this paper, we focus on the description of the B-differential of H at x when $H : \mathbb{R}^n \to \mathbb{R}^m$ 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 by

$$H(x) = \min(Ax + a, Bx + b), \quad (1.3)$$

where the minimum operator "min" acts componentwise (for two vectors u, $v \in \mathbb{R}^m$ and $i \in [1:m] := \{1, \ldots, m\}$: $[\min(u, v)]_i := \min(u_i, v_i)$). This function is usually nonsmooth. A motivation to look at the B-differential of that function H comes from the fact that, when $m = n$ and H is given by (1.3), as explained below, the equation

$$H(x) = 0 \quad (1.4)$$

is a reformulation of the *balanced* [29] *Linear Complementarity Problem* (LCP)

$$0 \leqslant (Ax + a) \perp (Bx + b) \geqslant 0. \quad (1.5)$$

This system expresses the fact that a point $x \in \mathbb{R}^n$ is sought such that $Ax + a \geqslant 0$, $Bx + b \geqslant 0$ and $(Ax + a)^\mathsf{T}(Bx + b) = 0$ (the superscript "$^\mathsf{T}$" is used here and below to denote vector or matrix transposition). Problem (1.5) is a special case of the so-called (*extended*) *vertical LCP*, which uses more than two matrices and vectors in its formulation [22, 81, 87]. In the *standard LCP*, A is the identity matrix and $a = 0$ [23, 58].

The reformulation (1.4) of (1.5) is based on the fact that, for two real numbers α and β, $\min(\alpha, \beta) = 0$ if and only if $\alpha \geqslant 0$, $\beta \geqslant 0$ and $\alpha\beta = 0$ [1, 61]. This reformulation serves as the basis for a number of solving methods and investigations [1, 8–10, 27–29, 39, 50, 53, 60–62]. If (1.5) stands alone, it is appropriate to have $m = n$, but (1.5) may be part of a system with other constraints to satisfy [11, 55, 56], in which case $m \leqslant n$. In the computation of the B-differential of the Min function (1.3), m and n may be unrelated. Note that there are many other ways of reformulating problem (1.5) as a nonsmooth system of equations. It is frequent to use the *Fischer function*, whose B-differential is computed in [40]. The function H in (1.3)

has been less studied and used than the Fischer function, although it has various advantages: it is piecewise affine (but has more nondifferentiability kinks), the local convergence of a semi-smooth Newton algorithm using it can be established under weaker assumptions and may be finitely locally convergent for linear complementarity problems [39; § 9.2].

Occasionally, we shall refer to the nonlinear version of the above problem, in which a function $\tilde{H} : \mathbb{E} \to \mathbb{R}^m$ is defined at $x \in \mathbb{E}$ by

$$\tilde{H}(x) := \min(F(x), G(x)), \tag{1.6}$$

where F and $G : \mathbb{E} \to \mathbb{R}^m$ are two functions and the "min" operator still acts componentwise. The equation $\tilde{H}(x) = 0$ is then a reformulation of the complementarity problem "$0 \leqslant F(x) \perp G(x) \geqslant 0$".

As a first general remark, let us quote the fact that the B-differential of H cannot be deduced from the knowledge of the B-differential of its scalar components $H_i : x \in \mathbb{E} \to H_i(x) \in \mathbb{R}$, for $i \in [1:m]$, which is trivial in the present context (lemma 2.1(4)). Indeed, it is known that [20; proposition 2.6.2(e)]

$$\partial_B H(x) \subseteq \partial_B^{\times} H(x) := \partial_B H_1(x) \times \cdots \times \partial_B H_m(x), \tag{1.7}$$

but equality in this inclusion may not hold (see [39; § 7.1.15], counter-example 2.3 and almost all the examples and test-cases below). Therefore, all the components of H must be taken into account simultaneously.

The B-differential of H at x is a finite set, made of Jacobians whose ith row is $A_{i,:}$ or $B_{i,:}$ (proposition 2.2). Consequently, its cardinality can be exponential in m and it occurs that its full mathematical description is a tricky task, essentially when there are many indices i for which $(Ax + a)_i = (Bx + b)_i$ and $A_{i,:} \neq B_{i,:}$, a situation that makes H nondifferentiable (lemma 2.1). Then, a rich panorama of configurations appears, which is barely glimpsed in this contribution.

The paper starts with a background section (section 2), which recalls a basic property of the minimum of two functions (lemma 2.1) and gives us a first perception of the structure of the B-differential of the function H, in particular its finite nature (proposition 2.2). A useful technical lemma is also presented (lemma 2.6).

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 origin in common (section 3.4).

Section 4 gives some properties of the B-differential of H, recalls Winder's formula of its cardinality, to which an analytic proof mimicking the original one is given, provides some lower and upper bounds on this one, proves necessary and sufficient conditions so that two extreme configurations occur and highlights two links between the B-differential and C-differential.

Section 5 presents algorithms for computing one (section 5.1) or all (section 5.2) the Jacobians of $\partial_B H(x)$. In the latter case, the algorithms construct a tree incrementally and recursively (section 5.2.2), as proposed by Rada and Černý [66]. On the one hand (section 5.2.3), an algorithm based on the notion of matroid circuit of the matrix V expressing the "derivative gap" is proposed; it has the nice feature of requiring no linear optimization problem (LOP) to solve. On the other hand (section 5.2.5), various modifications of the algorithm of Rada and Černý [66] are proposed with the goal of decreasing the number of LOPs to solve. The

complexity of one proposed algorithm is analyzed, bounding the number linear optimization problems to solve by a multiple of the cardinality of the B-differential. Numerical experiments are reported (section 5.2.8), showing that the proposed algorithms significantly improve the performance of the Rada and Černý method, with mean (resp. median) acceleration ratios in the range 15..31 (resp. 5..20), measured by the computing time. This speed-up exceeds 100, for some algorithms and test-problems.

An abridged version of this report can be found in [30].

Notation

We denote by $|S|$ the number of elements of a set S (i.e., its *cardinality*). The *power set* of a set S is denoted by $\mathfrak{P}(S)$. The set of *bipartitions* (I, J) of a set K is denoted by $\mathfrak{B}(K)$: $I \cup J = K$ and $I \cap J = \varnothing$. 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 $\operatorname{sgn} : \mathbb{R} \multimap \mathbb{R}$ defined by $\operatorname{sgn}(t) = \{1\}$ if $t > 0$, $\operatorname{sgn}(t) = \{-1\}$ if $t < 0$ and $\operatorname{sgn}(0) = [-1, 1]$. We note $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geqslant 0\}$ 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). For a subset S of a vector space, we denote by $\operatorname{vect}(S)$ the subspace spanned by S. The vector of all one's, in a real space whose dimension is given by the context, is denoted by e. The *Hadamard product* of u and $v \in \mathbb{R}^n$ is the vector $u \bullet v \in \mathbb{R}^n$ whose ith component is $u_i v_i$. The *range space* of an $m \times n$ matrix A is denoted by $\mathcal{R}(A)$, its *null space* by $\mathcal{N}(A)$, its *rank* is $\operatorname{rank}(A) := \dim \mathcal{R}(A)$ and its *nullity* is $\operatorname{null}(A) := \dim \mathcal{N}(A) = n - \operatorname{rank}(A)$ by the rank-nullity theorem. The ith row (resp. column) of A is denoted by $A_{i,:}$ (resp. $A_{:,i}$). Transposition operates after a row/column selection: $A^\mathsf{T}_{i,:}$ is a short notation for the column vector $(A_{i,:})^\mathsf{T}$ and $A^\mathsf{T}_{:,i}$ is a short notation for the row vector $(A_{:,i})^\mathsf{T}$. For a vector α, $\operatorname{Diag}(\alpha)$ is the square diagonal matrix with the α_i's on its diagonal.

2 Background

Recall that $F : \mathbb{E} \to \mathbb{F}$ 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 [20], "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 famous lemma recalls a necessary and sufficient condition guaranteeing the differentiability of the minimum of two scalar functions (see [64; 1993, final remarks (1)] and [85; 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 $g : \mathbb{E} \to \mathbb{R}$ be two functions and $h : \mathbb{E} \to \mathbb{R}$ 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 is G-differentiable (resp. differentiable) at x and $h'(x) = g'(x)$.*
3) *If $f(x) = g(x)$, then*

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

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

4) *If f and g are continuously differentiable at x, then*

$$\partial_B h(x) = \begin{cases} \{f'(x)\} & \text{if } f(x) < g(x), \\ \{f'(x), g'(x)\} & \text{if } f(x) = g(x), \\ \{g'(x)\} & \text{if } f(x) > g(x). \end{cases}$$

Proof. 1) This results from the fact that, when $f(x) < g(x)$, $h = f$ near x.

2) Switch f and g in point 1.

3) [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 $d \to 0$,

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

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

4) Let $J \in \partial_B h(x)$. Then, there exists a sequence $\{x_k\} \to x$ such that $x_k \in \mathcal{D}_h$ and $h'(x_k) \to J$.

If $f(x) < g(x)$ (switch f and g to deal with the case where $f(x) > g(x)$), $h(x_k) = f(x_k)$ and $h'(x_k) = f'(x_k)$ for k sufficiently large (continuity of f and g at x), so that $J = f'(x)$ by the continuity of f' at x.

If $f(x) = g(x)$, by points 1-3 and $x_k \in \mathcal{D}_h$, $h'(x_k) \in \{f'(x_k), g'(x_k)\}$. Since $h'(x_k) \to J$, $f'(x_k) \to f'(x)$ and $g'(x_k) \to g'(x)$, we get at the limit that $J \in \{f'(x), g'(x)\}$, which proves the inclusion $\partial_B h(x) \subseteq \{f'(x), g'(x)\}$.

Let us now prove the inclusion $\partial_B h(x) \supseteq \{f'(x), g'(x)\}$ when $f(x) = g(x)$. This one holds if $f'(x) = g'(x)$, since $\partial_B h(x)$ is nonempty and is contained in $\{f'(x)\} = \{g'(x)\}$ by the preceeding argument. Suppose now that $f'(x) \neq g'(x)$, so that there is a direction d such that $f'(x)d < g'(x)d$. Take $y_k = x + t_k d$ with $t_k \downarrow 0$ (resp. $t_k \uparrow 0$). Then, $f(y_k) - g(y_k) = t_k[f'(x)d - g'(x)d] + o(t_k)$ by the differentiability of f and g at x. Therefore, $h(y_k) = f(y_k) < g(y_k)$, $h'(y_k) = f'(y_k)$ (resp. $f(y_k) > g(y_k) = h(y_k)$, $h'(y_k) = g'(y_k)$) and $y_k \in \mathcal{D}_h$ for k sufficiently large. Taking the limit in k shows that $f'(x) \in \partial_B h(x)$ (resp. $g'(x) \in \partial_B h(x)$), as desired. □

The previous lemma shows the relevance of the following index sets, when the differentiability of the function H is at stake:

$$\mathcal{A}(x) := \{i \in [1:m] : (Ax + a)_i < (Bx + b)_i\}, \tag{2.2a}$$

$$\mathcal{B}(x) := \{i \in [1:m] : (Ax + a)_i > (Bx + b)_i\}, \tag{2.2b}$$

$$\mathcal{E}(x) := \{i \in [1:m] : (Ax + a)_i = (Bx + b)_i\}. \tag{2.2c}$$

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

$$\mathcal{E}^{=}(x) := \{i \in \mathcal{E}(x) : A_{i,:} = B_{i,:}\}, \tag{2.2d}$$

$$\mathcal{E}^{\neq}(x) := \{i \in \mathcal{E}(x) : A_{i,:} \neq B_{i,:}\}. \tag{2.2e}$$

To simplify the presentation, we assume in the sequel that

$$\mathcal{E}^{\neq}(x) = [1:p], \tag{2.3}$$

for some $p \in [0:m]$ ($p = 0$ if and only if $\mathcal{E}^{\neq}(x) = \varnothing$).

The next proposition describes the superset $\partial_B^{\times} H(x)$ of $\partial_B H(x)$ given in the right-hand side of (1.7) (see [51; 1998, § 2] in a somehow different context, [25; 2000, before (8)]). This Cartesian product actually reads

$$\begin{aligned} \partial_B^{\times} H(x) := \{J \in \mathcal{L}(\mathbb{E}, \mathbb{R}^m) : \quad & J_{i,:} = A_{i,:}, \text{ if } i \in \mathcal{A}(x), \\ & J_{i,:} = B_{i,:}, \text{ if } i \in \mathcal{B}(x), \\ & J_{i,:} = A_{i,:} = B_{i,:}, \text{ if } i \in \mathcal{E}^{=}(x), \\ & J_{i,:} \in \{A_{i,:}, B_{i,:}\}, \text{ if } i \in \mathcal{E}^{\neq}(x)\}. \end{aligned} \tag{2.4}$$

Proposition 2.2 (superset of $\partial_B H(x)$) *One has*

$$\partial_B H(x) \subseteq \partial_B H_1(x) \times \cdots \times \partial_B H_m(x) = \partial_B^{\times} H(x). \tag{2.5}$$

In particular, $|\partial_B H(x)| \leqslant 2^p$.

Proof. The inclusion in (2.5) is clear since, when $H'(x_k)$ converges to some J, $H_i'(x_k) \to J_{i,:}$, for all $i \in [1:m]$. The equality is also clear as a consequence of lemma 2.1(4).

The last claim is a straightforward consequences of the fact that $J_{i,:}$ can take two different values, $A_{i,:}$ or $B_{i,:}$, only for the indices $i \in \mathcal{E}^{\neq}(x)$ (recall that $|\mathcal{E}^{\neq}(x)| = p$). □

The following counter-example shows that one can have $\partial_B H(x) \neq \partial_B^{\times} H(x)$ and highlights the interest of the B-differential for the convergence of the semismooth Newton algorithm on (1.4).

Counter-example 2.3 Let $n = 2$, $m = 2$, $A = \left(\begin{smallmatrix} -1 & 1 \\ -1 & -1 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ and $a = b = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. One has $\mathcal{A}(0) = \mathcal{B}(0) = \varnothing$, $\mathcal{E}(0) = \mathcal{E}^{\neq}(0) = \{1, 2\}$, $\partial_B H(0) = \{A, B\}$, while $\partial_B^{\times} H(0) = \{A, B, \left(\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right)\}$. This example also shows that all the Jacobians of $\partial_B H(0)$ can be nonsingular, while the Jacobian $\left(\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right)$ of $\partial_B^{\times} H(0)$ is singular and the central Jacobian (4.16), namely $\frac{1}{2}(A + B) = \left(\begin{smallmatrix} 0 & 1 \\ 0 & -1 \end{smallmatrix}\right) \in \partial_C H(0)$, is also singular. Therefore, in this case, H is strongly BD-regular at 0 in the sense of [64; p. 233] and the conditions ensuring the local convergence of the semismooth Newton algorithm are satisfied [64; theorem 3.1]. □

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

Definition 2.4 (complete B-differential) We say that the B-differential of H at $x \in \mathbb{R}^n$ is *complete* if $\partial_B H(x) = \partial_B^{\times} H(x)$ or, equivalently, if $|\partial_B H(x)| = 2^p$. □

Definitions 2.5 (symmetry in $\partial_B H(x)$) For $x \in \mathbb{E}$, we say that the Jacobian $\tilde{J} \in \partial_B^{\times} H(x)$ is *symmetric* to the Jacobian $J \in \partial_B^{\times} H(x)$ if

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

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 $\partial_B H(x)$. □

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 (one could take for $\mathbb{E}$ a subspace of $\mathbb{R}^q$, for some $q \in \mathbb{N}^*$, and the Euclidean scalar product for $\langle \cdot, \cdot \rangle$). It is a refinement of [85; lemma 2.1].

Lemma 2.6 (discriminating covectors) *Suppose that $(\mathbb{E}, \langle \cdot, \cdot \rangle)$ is a Euclidean vector space, $p \in \mathbb{N}^*$ and $v_1, \ldots, v_p$ 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 Ξ the set of vectors $\xi \in \mathbb{E}$ such that $|\{\langle \xi, v_i \rangle : i \in [1:p]\}| = p$ (i.e., $\{\langle \xi, v_i \rangle : i \in [1:p]\}$ has p distinct values in $\mathbb{R}$). We have to show that Ξ 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, ξ_1 maintains strict the inequalities that are strict with ξ_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 ξ_1 than with ξ_0. Pursuing like this, one can finally obtain a vector ξ in Ξ. This vector is arbitrarily close to ξ_0 by taking the ε_i's positive and sufficienty small. □

3 Equivalent problems

The problem of determining the B-differential of the piecewise affine function, that is the minimum (1.3) of two *affine* functions, appears in various contexts, sometimes with non straightforward connections with it (this one is recalled in section 3.1). We review some equivalent formulations in this section (see also [5,7,84] and the references therein) and give a few properties of the B-differential in this piecewise 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 a number of elements exponential in p, which makes its content difficult to specify (in this respect, the particular case where the B-differential is complete is a trivial exception). Some formulations, such as the one related to the arrangement of hyperplanes 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 interesting to mention and keep in mind. They also offer 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, in this paper, as the *original problem*.

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 $H : \mathbb{R}^n \to \mathbb{R}^m$ defined by (1.3). □

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

$$V := (B - A)^{\mathsf{T}}_{\mathcal{E}^{\neq}(x),:} \in \mathbb{R}^{n\times p}. \tag{3.1}$$

Note that, due to their indices in $\mathcal{E}^{\neq}(x) = [1\,{:}\,p]$ and the definition of this index set, the columns of V are nonzero. This matrix may not always have full rank, however.

The following example will accompany us throughout this section.

Example 3.2 (a simple example) Consider the trivial linear complementarity problem $0 \leqslant x \perp (Mx + q) \geqslant 0$ defined by

$$M = \begin{pmatrix} 2 & 0 & 0 \\ -\alpha & 1{+}\beta & 0 \\ -\alpha & -\beta & 1 \end{pmatrix} \qquad \text{and} \qquad q = 0,$$

where $\alpha := -\cos(2\pi/3) = 1/2 > 0$ and $\beta := \sin(2\pi/3) \in (\alpha, 2\alpha)$. Note that $M \in \mathbf{P}$ and that, at the unique solution $x = 0$ to the problem, one has $\mathcal{A}(x) = \mathcal{B}(x) = \mathcal{E}^{=}(x) = \varnothing$ and $\mathcal{E}(x) = \mathcal{E}^{\neq}(x) = [1\,{:}\,3]$, so that $p = 3$ and

$$V = \begin{pmatrix} 1 & -\alpha & -\alpha \\ 0 & \beta & -\beta \\ 0 & 0 & 0 \end{pmatrix}.$$

□

3.2 Linear algebra problems

3.2.1 Signed feasibility of strict inequality systems

We call *sign vector* a vector whose components are +1 or -1. Many proofs below leverage the equivalence between the original problem 3.1 and the following one. The reason is that working on problem 3.3 often allows us to propose shorter proofs. In addition, the algorithms of section 5 all focus on the generation of the sign vectors s forming the set $\mathcal{S}$ in (3.2) below. Recall the definition of the Hadamard product: $(u \bullet v)_i = u_i v_i$.

Problem 3.3 (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

$$\mathcal{S} := \{s \in \{\pm 1\}^p : \ s \bullet (V^\mathsf{T} d) > 0 \text{ holds for some } d \in \mathbb{R}^n\}. \tag{3.2}$$

□

By routine verification, one can see that the sign vectors s in $\mathcal{S}$ for example 3.2 are given by the columns of the matrix S below and possible associated directions d such that $s \bullet (V^\mathsf{T} d) > 0$ are given by the corresponding columns of the matrix D:

$$S = \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 2 & 2 & 2 & -2 & -2 & -2 \\ 2 & 1 & -2 & -2 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.3}$$

The sign vectors $\pm e := \pm(1,1,1)$ are not in $\mathcal{S}$ since $Ve = 0$ (there is not $d_\pm$ such that $(\pm e)\bullet(V^\mathsf{T} d_\pm) > 0$, since this would imply that $0 < \pm e^\mathsf{T} V^\mathsf{T} d_\pm = 0$, a contradiction). Therefore, there are only 6 sign vectors in $\mathcal{S}$ instead of the 8 sign vectors in $\{\pm 1\}^3$.

The link between problems 3.1 and 3.3 is established by the following map:

$$\sigma : J \in \partial_B^\times H(x) \mapsto s \in \{\pm 1\}^p, \text{ where } s_i = \begin{cases} +1 & \text{if } i \in \mathcal{E}^{\neq}(x),\ J_{i,:} = A_{i,:}, \\ -1 & \text{if } i \in \mathcal{E}^{\neq}(x),\ J_{i,:} = B_{i,:}, \end{cases} \tag{3.4a}$$

where we have used the definition (2.3) of p. The map is well defined since $A_{i,:} \neq B_{i,:}$ when $i \in \mathcal{E}^{\neq}(x)$. Furthermore, σ is bijective since two Jacobians in $\partial_B^\times H(x)$ 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, its reverse map is

$$\sigma^{-1} : s \in \{\pm 1\}^p \mapsto J \in \partial_B^\times H(x), \text{ where } J_{i,:} = \begin{cases} A_{i,:} & \text{if } i \in \mathcal{E}^{\neq}(x),\ s_i = +1, \\ B_{i,:} & \text{if } i \in \mathcal{E}^{\neq}(x),\ s_i = -1. \end{cases} \tag{3.4b}$$

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

Proposition 3.4 (bijection $\partial_B H(x) \leftrightarrow \mathcal{S}$) *Let $H : \mathbb{R}^n \to \mathbb{R}^m$ be given by (1.3), x be a point in $\mathbb{R}^n$ such that $p \neq 0$ and V be given by (3.1). Then, the map σ 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 $\sigma(J) \bullet (V^\mathsf{T} d) > 0$.*
2) *If $s \in \{\pm 1\}^p$ and $\exists\, d \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d) > 0$, then $\sigma^{-1}(s) \in \partial_B H(x)$.*

3) *Let* $J \in \partial_B^\times H(x)$. *Then,* $J \in \partial_B H(x) \iff \sigma(J) \bullet (V^\mathsf{T} d) > 0$ *holds for some* $d \in \mathbb{R}^n$.

Proof. The properties 1, 2 and 3 in the statement of the proposition are straightforward consequences of the bijectivity of $\sigma : \partial_B H(x) \to \mathcal{S}$. Now, the discussion before the proposition has shown that $\sigma : \partial_B^\times H(x) \mapsto \{\pm 1\}^p$ is a bijection. Therefore, $\sigma : \partial_B H(x) \mapsto \{\pm 1\}^p$ is injective and it suffices to prove that

$$\sigma(\partial_B H(x)) = \mathcal{S}. \tag{3.5a}$$

[$\subseteq$ or point 1] Let $J \in \partial_B H(x)$. We have to show that $s := \sigma(J) \in \mathcal{S}$, which means that one can find a $d \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d) > 0$. By $J \in \partial_B H(x)$, there exists a sequence $\{x_k\} \subseteq \mathcal{D}_H$ converging to x such that

$$H'(x_k) \to J. \tag{3.5b}$$

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 $k \in \mathcal{K}$:

$$(Ax_k + a)_{\mathcal{A}_0} < (Bx_k + b)_{\mathcal{A}_0} \qquad \text{and} \qquad (Ax_k + a)_{\mathcal{B}_0} > (Bx_k + b)_{\mathcal{B}_0}. \tag{3.5c}$$

Now, fix $k \in \mathcal{K}$ and set $d := x_k - x$. Since $(Ax + a)_i = (Bx + b)_i$ for $i \in \mathcal{E}^{\neq}(x)$, one deduces from (3.5c) that

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

Recalling the definitions of V in (3.1) and $\mathcal{S}$ in (3.2), 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 σ, ($J_{i,:} = A_{i,:}$ for $i \in \mathcal{A}_0$) and ($J_{i,:} = B_{i,:}$ for $i \in \mathcal{B}_0$). This is indeed the case, since by (3.5c), 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 (3.5b) to conclude.

[$\supseteq$ or point 2] Let $s \in \mathcal{S}$. We have to find a $J \in \partial_B H(x)$ such that $\sigma(J) = s$, that is, which satisfies for $i \in [1:p]$:

$$(J_{i,:} = A_{i,:} \text{ if } s_i = +1) \quad \text{and} \quad (J_{i,:} = B_{i,:} \text{ if } s_i = -1). \tag{3.5d}$$

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

$$s \bullet (V^\mathsf{T} d) > 0. \tag{3.5e}$$

Take a real sequence $\{t_k\} \downarrow 0$ and define the sequence $\{x_k\} \subseteq \mathbb{R}^n$ by

$$x_k := x + t_k d.$$

Then, $x_k \to x$. We claim that, for k sufficiently large, $x_k \in \mathcal{D}_H$ and $H'(x_k)$ is a constant matrix J satisfying (3.5d), which will conclude the proof. Let $i \in [1:m]$.

- If $i \in \mathcal{A}(x)$, $(Ax_k + a)_i < (Bx_k + b)_i$ for k large, so that $x_k \in \mathcal{D}_H$ and $H_i'(x_k) = A_{i,:}$.
- If $i \in \mathcal{B}(x)$, $(Ax_k + a)_i > (Bx_k + b)_i$ for k large, so that $x_k \in \mathcal{D}_H$ and $H_i'(x_k) = B_{i,:}$.
- If $i \in \mathcal{E}^{=}(x)$, then $A_{i,:} = B_{i,:}$, so that $x_k \in \mathcal{D}_H$ and $H_i'(x_k) = A_{i,:} = B_{i,:}$.

- If $i \in \mathcal{E}^{\neq}(x)$, subtract side by side $(Ax_k + a)_i = (Ax + a)_i + t_k A_{i,:} d$ and $(Bx_k + b)_i = (Bx + b)_i + t_k B_{i,:} d$, use $(Ax + a)_i = (Bx + b)_i$ and next (3.5e) to get

$$(Bx_k + b)_i - (Ax_k + b)_i = t_k(B_{i,:} - A_{i,:})d = t_k V_{i,:}^{\mathsf{T}} d \begin{cases} > 0 & \text{if } s_i = +1, \\ < 0 & \text{if } s_i = -1. \end{cases}$$

Hence, $x_k \in \mathcal{D}_H$, ($H_i'(x_k) = A_{i,:}$ if $s_i = +1$) and ($H_i'(x_k) = B_{i,:}$ if $s_i = -1$). □

Equivalence 3.5 (B-differential ↔ signed feasibility of strict inequality systems) The equivalence between the original problem 3.1 and the signed feasibility of strict inequality system problem 3.3 is a consequence of the previous proposition with V given by (3.1), which shows the bijectivity of the map $\sigma : \partial_B H(x) \to \mathcal{S}$ defined by (3.4a). Therefore, knowing σ by its definition (3.4), determining $\partial_B H(x)$ or $\mathcal{S}$ are equivalent problems. □

3.2.2 Orthants encountered by the null space of a matrix

Recall the definition of $\mathcal{S}$ in (3.2), which is associated with some matrix $V \in \mathbb{R}^{n\times p}$ with nonzero columns, which may or not come from (3.1). The equivalent form of problem 3.3 (hence of problem 3.1 when V is defined by (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 subsets of $[1:p]$ (i.e., $\mathcal{I} \subseteq \mathfrak{P}([1:p])$), which refers to a collection of orthants of $\mathbb{R}^p$, those encountered by the null space of V. This equivalence will play a major part in the conception of the algorithms in section 5.2, in particular, but not only, in an algorithm describing the *complementary set* of $\partial_B H(x)$, which is interesting when $|\partial_B^{\times} H(x) \setminus \partial_B H(x)|$ is small. The concept of *stem vector*, defined in the second part of this section, has proven useful in this regard. The equivalence rests on a duality concept through Gordan's alternative.

Problem 3.6 (orthants encountered by the null space of a matrix) 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 $\mathbb{R}^p$:

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

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

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

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 3.3 and 3.6 is obtained thanks to the following bijection

$$\imath : s \in \{\pm 1\}^p \to \imath(s) := \{i \in [1:p] : s_i = +1\} \in \mathfrak{P}([1:p]), \tag{3.6}$$

whose reverse map is $\imath^{-1} : I \in \mathfrak{P}([1:p]) \to s \in \{\pm 1\}^p$, 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 [44; 1873]: for a matrix $A \in \mathbb{R}^{m\times n}$,

$$\boxed{\exists\, x \in \mathbb{R}^n : Ax > 0 \qquad \Longleftrightarrow \qquad \nexists\, \alpha \in \mathbb{R}_+^m \setminus \{0\} : A^{\mathsf{T}}\alpha = 0.} \tag{3.7}$$

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

Proof. Let $s \in \{\pm 1\}^p$ and set $I := \imath(s) = \{i \in [1:p] : s_i = +1\}$. Define $A := \mathrm{Diag}(s)V^\mathsf{T}$ to make the link with Gordan's alternative (3.7). One has the equivalences

$$\begin{aligned} s \in \mathcal{S}^c \quad &\iff \quad \nexists\, x \in \mathbb{R}^n : Ax > 0 \qquad \text{[definition of } \mathcal{S} \text{ in (3.2)]} \\ &\iff \quad \exists\, \alpha \in \mathbb{R}^m_+ \setminus \{0\} : A^\mathsf{T}\alpha = 0 \qquad \text{[Gordan's alternative (3.7)]} \\ &\iff \quad \exists\, \alpha \in \mathbb{R}^m_+ \setminus \{0\} : s \bullet \alpha \in \mathcal{N}(V) \\ &\iff \quad \mathcal{N}(V) \cap \mathcal{O}^p_I \neq \{0\} \qquad \text{[see below]} \\ &\iff \quad I \in \mathcal{I} \qquad \text{[definition of } \mathcal{I}\text{]}. \end{aligned} \tag{3.8}$$

The implication "$\Rightarrow$" in (3.8) is due to the fact that $s \bullet \alpha$ is nonzero and belongs to both $\mathcal{N}(V)$ and $\mathcal{O}^p_I$. The reverse implication "$\Leftarrow$" in (3.8) is due to the fact that there is a nonzero $y \in \mathcal{N}(V) \cap \mathcal{O}^p_I$, implying that $\alpha := s \bullet y$ is nonzero and $\geqslant 0$ and is such that $s \bullet \alpha = y \in \mathcal{N}(V)$.

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

Equivalence 3.8 ($\mathcal{S}^c \leftrightarrow \mathcal{I}$) The equivalence between problems 3.3 and 3.6 is a consequence of the bijectivity of $\imath : \mathcal{S}^c \to \mathcal{I}$, established in proposition 3.7: to determine $\mathcal{S}$, it suffices to determine $\mathcal{S}^c = \imath^{-1}(\mathcal{I})$, hence to determine $\mathcal{I}$, and vice versa. □

In example 3.2, one has $\mathcal{N}(V) = \mathbb{R}e$, which only encounters the orthants $\mathcal{O}^3_\varnothing$ and $\mathcal{O}^3_{[1:3]}$ outside the origin; hence $\mathcal{I} = \{\varnothing, [1:3]\}$. We have seen that $\mathcal{S}^c = \{\pm(1,1,1)\}$ for this problem. Clearly, $\imath$ maps $\mathcal{S}^c$ onto $\mathcal{I}$ bijectively, as claimed in proposition 3.7.

Recall that the *nullity* of a matrix A, denoted by $\mathrm{null}(A)$, is the dimension of its null space. Let us introduce the following collection of index sets (from now on, J usually denotes a set of indices rather than a Jacobian matrix):

$$\mathcal{C} := \{J \subseteq [1:p] : J \neq \varnothing,\ \ \mathrm{null}(V_{:,J}) = 1, V_{:,J_0} \text{ is injective if } J_0 \subsetneq J\}, \tag{3.9}$$

where "$\subsetneq$" is used to denote strict inclusion. In the terminology of the *vector matroid* formed by the columns of V and its subsets made of linearly independent columns [59; proposition 1.1.1], the elements of $\mathcal{C}$ are called the *circuits* of the matroid [59; proposition 1.3.5(*iii*)]. The particular expression (3.9) of the circuit set is interesting in the present context, since it readily yields the following implication:

$$J \in \mathcal{C} \quad \Longrightarrow \quad \text{any nonzero } \alpha \in \mathcal{N}(V_{:,J}) \text{ has none zero component.} \tag{3.10}$$

From (3.9) and (3.10), one can associate with $J \in \mathcal{C}$ a pair of sign vectors $\pm\tilde{s} \in \{\pm 1\}^J$ by $\tilde{s} := \mathrm{sgn}(\alpha)$ for some nonzero $\alpha \in \mathcal{N}(V_{:,J})$; the sign vectors $\pm\tilde{s}$ do not depend on the chosen $\alpha \in \mathcal{N}(V_{:,J}) \setminus \{0\}$ since $\mathrm{null}(V_{:,J}) = 1$. We call such a sign vector a *stem vector*, because of proposition 3.10 below, which shows that any $s \in \mathcal{S}^c$ can be generated from such a stem vector.

Definition 3.9 (stem vector) A *stem vector* is a sign vector $\tilde{s} = \mathrm{sgn}(\alpha)$, where $\alpha \in \mathcal{N}(V_{:,J})$ for some $J \in \mathcal{C}$. □

Note that there are twice as many stem vectors as circuits and that the stem vectors do not have all the same size.

The matrix V in example 3.2 has $J = [1:3]$ as single circuit. Since $Ve = 0$, the associated stem vectors are $\pm e = \pm(1,1,1)$. The next proposition now confirms that $\pm(1,1,1)$ are the only elements of $\mathcal{S}^c$.

Proposition 3.10 (generating $\mathcal{S}^c$ from the stem vectors) *For $s \in \{\pm 1\}^p$,*

$$s \in \mathcal{S}^c \quad \Longleftrightarrow \quad s_J = \tilde{s} \text{ for some } J \subseteq [1:p] \text{ and some stem vector } \tilde{s}. \tag{3.11}$$

Proof. [$\Rightarrow$] The index set $J \subseteq [1:p]$ in the right-hand side of (3.11) can be determined as one satisfying the following two properties:

$$\{d \in \mathbb{R}^n : s_j v_j^\mathsf{T} d > 0 \text{ for all } j \in J\} = \varnothing, \tag{3.12a}$$

$$\forall\, J_0 \subsetneq J,\ \{d \in \mathbb{R}^n : s_j v_j^\mathsf{T} d > 0 \text{ for all } j \in J_0\} \neq \varnothing. \tag{3.12b}$$

To determine such a J, start with $J = [1:p]$, which verifies (3.12a), since $s \in \mathcal{S}^c$. Next, remove an index j from $[1:p]$ if (3.12a) holds for $J = [1:p] \setminus \{j\}$. Pursuing the elimination of indices j in this way, one arrives to an index set J satisfying (3.12a) and $\{d \in \mathbb{R}^n : s_j v_j^\mathsf{T} d > 0$ for all $j \in J \setminus \{j_0\}\} \neq \varnothing$ for all $j_0 \in J$. Then, (3.12b) clearly holds. We claim that, for a J satisfying (3.12a) and (3.12b), s_J is a stem vector, which will conclude the proof of the implication.

To stick to definition 3.9, we start by showing that J is a matroid circuit. By (3.12a), $J \neq \varnothing$. By Gordan's alternative (3.7), (3.12a) and (3.12b) read

$$\exists\, \alpha \in \mathbb{R}^J_+ \setminus \{0\} \text{ such that } \textstyle\sum_{j \in J} s_j v_j \alpha_j = 0, \tag{3.12c}$$

$$\forall\, J_0 \subsetneq J,\ \nexists\, \alpha' \in \mathbb{R}^{J_0}_+ \setminus \{0\} \text{ such that } \textstyle\sum_{j \in J_0} s_j v_j \alpha'_j = 0. \tag{3.12d}$$

From these properties, one deduces that $\alpha > 0$ and that $\operatorname{null}(V_{:,J}) \geqslant 1$. To show that $\operatorname{null}(V_{:,J}) = 1$, we proceed by contradiction. Suppose that there is a nonzero $\alpha'' \in \mathbb{R}^J$ that is not colinear with α and that verifies $\sum_{j \in J} s_j v_j \alpha''_j = 0$. One can assume that $t := \max\{\alpha''_j/\alpha_j : j \in J\}$ is > 0 (take $-\alpha''$ otherwise). Set $J_0 := \{j \in J : \alpha''_j/\alpha_j < t\}$. By the non-colinearity of α and α'', on the one hand, and the definition of t, on the other hand, one has $\varnothing \subsetneq J_0 \subsetneq J$. Furthermore, $\alpha' := \alpha - \alpha''/t \geqslant 0$, $\alpha'_j > 0$ for $j \in J_0$ and $\alpha'_j = 0$ for $j \in J \setminus J_0$. Therefore, $\sum_{j \in J_0} s_j v_j \alpha'_j = \sum_{j \in J} s_j v_j \alpha'_j = 0$, yielding a contradiction with (3.12d).

To show that $J \in \mathcal{C}$, we still have to prove that $V_{:,J_0}$ is injective when $J_0 \subsetneq J$. Equivalently, it suffices to show that any $\beta \in \mathcal{N}(V_{:,J})$ with some zero component vanishes. We proceed by contradiction. If there is a $\beta \in \mathcal{N}(V_{:,J}) \setminus \{0\}$ with a zero component, $s_J \bullet \alpha$ and β would be two linearly independent vectors in $\mathcal{N}(V_{:,J})$ (since $s_J \bullet \alpha$ has no zero component), contradicting $\operatorname{null}(V_{:,J}) = 1$.

Now, since $s_J = \operatorname{sgn}(s_J \bullet \alpha)$, since $s_J \bullet \alpha \in \mathcal{N}(V_{:,J})$ by (3.12c) and since J is a matroid circuit of V, s_J is a stem vector.

[$\Leftarrow$] Since s_J is a stem vector, it follows that $s_J := \operatorname{sgn}(\alpha)$ for some $\alpha \in \mathbb{R}^J$ with nonzero components that satisfies $V_{:,J}\alpha = 0$. Then, there is no $d \in \mathbb{R}^n$ such that $s_J \bullet (V_{:,J}^\mathsf{T} d) > 0$ (otherwise, $(s_J \bullet \alpha \bullet s_J) \bullet (V_{:,J}^\mathsf{T} d) > 0$, because $s_J \bullet \alpha > 0$, or $\alpha \bullet (V_{:,J}^\mathsf{T} d) > 0$, implying that $0 = \alpha^\mathsf{T}(V_{:,J}^\mathsf{T} d) > 0$, a contradiction). Hence, there exists certainly no $d \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d) > 0$. This implies that $s \in \mathcal{S}^c$. $\square$

To determine the stem vectors, which are based on the matroid circuits of V defined by (3.9), one has to select subsets of columns of V forming a rank one matrix, whose strict subsets form injective matrices. Actually, this last condition can be simplified by the following property.

Proposition 3.11 (matroid circuit detection) *Suppose that $I \subseteq [1:p]$ is such that $\mathrm{null}(V_{:,I}) = 1$ and that $\alpha \in \mathcal{N}(V_{:,I}) \setminus \{0\}$. Then, $J := \{i \in I : \alpha_i \neq 0\}$ is a matroid circuit of V and the unique one included in I.*

Proof. 1) Let us show that J is a matroid circuit.

Since $\alpha \neq 0$, one has $J \neq \varnothing$.

Let us show that $\mathrm{null}(V_{:,J}) = 1$. Since $J \subseteq I$, one has $\mathrm{null}(V_{:,J}) \leqslant \mathrm{null}(V_{:,I}) = 1$. Furthermore, $\alpha_J \in \mathcal{N}(V_{:,J}) \setminus \{0\}$ implies that $\mathrm{null}(V_{:,J}) \geqslant 1$.

Now, let $J_0 \subsetneq J$ and suppose that $V_{:,J_0}\beta = 0$. We have to show that $\beta = 0$. Since $V_{:,J}(\beta, 0_{J\setminus J_0}) = 0$, it follows that $(\beta, 0_{J\setminus J_0}) \in \mathcal{N}(V_{:,J})$, which is of dimension 1, so that $(\beta, 0_{J\setminus J_0})$ is colinear to α. Since the components of α are $\neq 0$, we get that $\beta = 0$.

2) Let us now show that J is the unique matroid circuit of V included in I.

Let J' be a matroid circuit of V included in I. Then $\mathrm{null}(V_{:,J'}) = 1$ and there is a nonzero $\alpha' \in \mathcal{N}(V_{:,J'})$. By (3.10), α' has nonzero components. Furthermore, $(\alpha', 0_{I\setminus J'}) \in \mathcal{N}(V_{:,I})$, which has unit dimension and contains α. Therefore, α and $(\alpha', 0_{I\setminus J'})$ are colinear. Since the components of α are $\neq 0$, we get that $J' = J$. □

3.3 Convex analysis problems

The formulation of the original problem 3.1 in the form of the convex analysis problems 3.12 and 3.15 below may be useful to highlight some properties of $\partial_B H(x)$, thanks to the tools of that discipline. We take this point of view to introduce the notion of extremality (definition 4.6) and to propose other proofs of propositions 4.5 and 4.18 below.

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 $K \cap (-K) = \{0\}$ [16; 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 3.12 (pointed cones by vector inversions) Let be given two positive integers n and $p \in \mathbb{N}^*$ and p vectors $v_1, \ldots, v_p \in \mathbb{R}^n \setminus \{0\}$. It is requested to determine all the sign vectors $s \in \{\pm 1\}^p$ such that $\mathrm{cone}\{s_i v_i : i \in [1:p]\}$ is pointed. □

The equivalence between the original problem 3.1 and problem 3.12 is obtained thanks to the next proposition, which gives another property ("cone pointedness") that is equivalent to those in (3.7) and that is adapted to the present concern. See also [45; theorem 2.3.29].

Proposition 3.13 (pointed polyhedral cone) *For a finite collection of nonzero vectors* $\{w_i : i \in [1:p]\} \subseteq \mathbb{R}^n$*, the following properties are equivalent:*

(i) $\operatorname{cone}\{w_i : i \in [1:p]\}$ *is pointed,*

(ii) $\nexists\, \alpha \in \mathbb{R}^p_+ \setminus \{0\} : \sum_{i\in[1:p]} \alpha_i w_i = 0,$

(iii) $\exists\, d \in \mathbb{R}^n,\ \forall\, i \in [1:p] : w_i^\mathsf{T} d > 0.$

Proof. The equivalence $(ii) \Leftrightarrow (iii)$ follows from Gordan's alternative (3.7) (with $A = V^\mathsf{T}$), so that it remains to prove $(i) \Leftrightarrow (ii)$. Set $K := \operatorname{cone}\{w_i : i \in [1:p]\}$.

$[(i) \Rightarrow (ii)]$ One can assume that $p \geqslant 2$, 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 $\sum_{i\in[1:p]} \alpha_i w_i = 0$. Without loss of generality, one can assume that $\alpha_1 \neq 0$. Set $z := \alpha_1 w_1 \in K \setminus \{0\}$. One also has $-z = \sum_{i\in[2:p]} \alpha_i w_i \in K$. 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, $z = \sum_{i\in[1:p]} \alpha'_i w_i$ and $-z = \sum_{i\in[1:p]} \alpha''_i w_i$ for some α' and $\alpha'' \in \mathbb{R}^p_+ \setminus \{0\}$. Adding the two identities side by side, we get $\sum_{i\in[1:p]} \alpha_i w_i = 0$, with $\alpha := \alpha' + \alpha''$. Since $\alpha \in \mathbb{R}^p_+ \setminus \{0\}$, this contradicts (ii). □

Equivalence 3.14 (signed linear system feasibility ↔ pointed cone by vector inversion) The equivalence $(i) \Leftrightarrow (iii)$ of the previous proposition shows that the set $\mathcal{S}$ defined by (3.2) is also given by

$$\mathcal{S} = \{s \in \{\pm 1\}^p : \ \operatorname{cone}\{s_i v_i : i \in [1:p]\} \text{ is pointed}\}. \tag{3.13}$$

To put it in words, denoting by $v_1, \ldots, v_p$ the columns of the matrix V defined by (3.1), the signed feasibility problem 3.3 is equivalent to problem 3.12. □

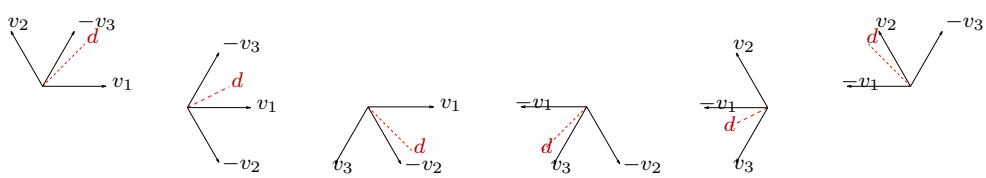

Figure 3.1: The figure is related to the linear complementarity problem defined by example 3.2: the v_i's are the columns of the matrix V (their third zero components are not represented). Each of the 6 sets of vectors plots the 3 vectors $\{s_i v_i : i \in [1:3]\}$, for each of the 6 sign vectors $s \in \mathcal{S}$ (given by the columns of the matrix S in (3.3)), as well as a direction d (given by the columns of D in (3.3), dashed lines) such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1:3]$. Each conic hull of these vectors, namely $\operatorname{cone}\{s_i v_i : i \in [1:3]\}$, is pointed. The conic hulls of $\{v_1, v_2, v_3\}$ and $\{-v_1, -v_2, -v_3\}$ are both the space of dimension 2, hence there are not pointed, which confirms the fact that $(1, 1, 1)$ and $(-1, -1, -1)$ are not in $\mathcal{S}$.

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 shortly considered by Zaslavsky [86; 1975, § 6A]. This enumeration problem appears in the study of neural networks [83]. Baldi and Vershynin [7] make

the connection with *homogeneous linear threshold functions* and highlight its impact in deep learning [6, 74].

Problem 3.15 (linearly separable bipartitioning) Let be given an affine space $\mathbb{A}$ and $p \in \mathbb{N}^*$ vectors $\bar{v}_1, \ldots, \bar{v}_p \in \mathbb{A}$. Let $\mathbb{A}_0 := \mathbb{A} - \mathbb{A}$ be the vector space parallel to $\mathbb{A}$, endowed with a scalar product $\langle \cdot, \cdot \rangle$. It is requested to find all the ordered 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

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

Of course, if (I, J) is an appropriate ordered bipartition to which a separating covector ξ corresponds, then (J, I) is also an appropriate ordered bipartition with separating covector $-\xi$. Therefore, only half of the appropriate ordered bipartitions (I, J) must be identified, a fact that is related to the symmetry of $\partial_B H(x)$ (proposition 4.1). Figure 3.2 shows the solution to

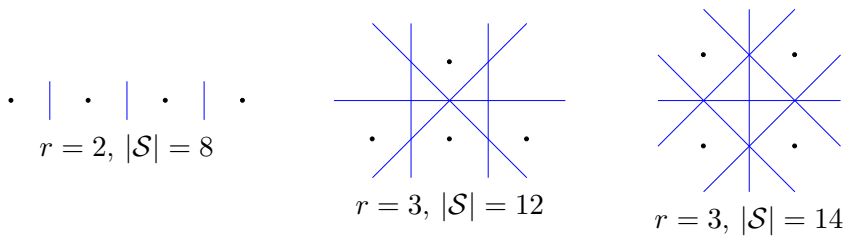

Figure 3.2: Linearly separable bipartitions of a set of $p = 4$ points $\bar{v}_i$ in $\mathbb{R}^2$ (the dots in the figure). Possible separating hyperplanes are the drawn lines. We have not represented any separating line associated with the partition $(\varnothing, [1:p])$ or $([1:p], \varnothing)$, so that $|\mathcal{S}| = 2(n_s + 1)$, where n_s is the number of represented separating lines. We have set $r := \dim(\mathrm{vect}\{\bar{v}_1, \ldots, \bar{v}_p\}) + 1$.

this problem by drawing the separating hyperplanes $\{\bar{v} \in \mathbb{A} : \xi^\mathsf{T} \bar{v} = t\}$ corresponding to some separating covector ξ and some $t \in \mathbb{R}$, for three examples with $p = 4$. 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 $r := \dim(\mathrm{vect}\{\bar{v}_1, \ldots, \bar{v}_p\}) + 1$, 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 bounds 8 (if $r = 2$) and 14 (if $r = 3$) given by propositions 4.11 and 4.15.

The equivalence between the linearly separable bipartitioning problem 3.15 of this section and the vector inversion problem 3.12 (hence, with the original problem 3.1) is grounded on the following construction and proposition.

Construction 3.16 1) Let be given two integers n and $p \in \mathbb{N}^*$ and p nonzero vectors $v_1, \ldots, v_p \in \mathbb{R}^n$ such that $K := \mathrm{cone}\{v_k : k \in [1:p]\}$ is a pointed cone. From proposition 3.13, there is a direction $d \in \mathbb{R}^n$ such that

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

Define

$$\mathbb{A} := \{\bar{v} \in \mathbb{R}^n : d^\mathsf{T} \bar{v} = 1\}, \qquad \mathbb{A}_0 := \mathbb{A} - \mathbb{A} = \{v \in \mathbb{R}^n : d^\mathsf{T} v = 0\},$$
$$\forall\, k \in [1:p]: \quad \bar{v}_k := v_k / (v_k^\mathsf{T} d) \in \mathbb{A}.$$

2) For a given bipartition (I, J) of $[1:p]$, define

$$K_I := \operatorname{cone}\{v_i : i \in I\} \qquad \text{and} \qquad K_J := \operatorname{cone}\{v_j : j \in J\}, \tag{3.14a}$$

$$C_I := K_I \cap \mathbb{A} \qquad \text{and} \qquad C_J := K_J \cap \mathbb{A}, \tag{3.14b}$$

with the convention $K_\varnothing = \{0\}$ and $C_\varnothing = \varnothing$. □

Proposition 3.17 (pointed cone after vector inversions) *Adopt the construction 3.16 and take a partition (I, J) of $[1:p]$. Then, the following properties are equivalent:*

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

Proof. $[(i) \Rightarrow (ii)]$ We show the contrapositive. If there is $v \in (K_I \cap K_J) \setminus \{0\}$, then $-v \in (-K_I) \subseteq \operatorname{cone}((-K_I) \cup K_J)$ and $v \in K_J \subseteq \operatorname{cone}((-K_I) \cup K_J)$. Therefore, $\operatorname{cone}((-K_I) \cup K_J)$ is not pointed.

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

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

$$C_I \text{ is nonempty, convex and compact.}$$

Indeed, since C_I is nonempty (it contains the vectors $\bar{v}_i$ for $i \in I \neq \varnothing$), 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 [71; p. 61]), namely $C_I^\infty = K_I \cap \mathbb{A}_0$, is reduced to $\{0\}$ [71; theorem 8.4]. This is indeed the case since $v^\mathsf{T} d > 0$ for all $v \in K_I \setminus \{0\}$. For the same reason,

$$C_J \text{ is nonempty, convex and compact.}$$

Now, since $C_I \cap C_J = \varnothing$ by (iii), one can strictly separate the convex sets C_I and C_J in $\mathbb{A}$ [71; corollary 11.4.2]: there exists $\xi \in \mathbb{A}_0$ such that $\xi^\mathsf{T} v < \xi^\mathsf{T} w$, for all $v \in C_I$ and all $w \in C_J$. This shows that (iv) holds.

$[(iv) \Rightarrow (i)]$ Since $\operatorname{cone}((-K_I) \cup K_J) = \operatorname{cone}(\{-v_i : i \in I\} \cup \{v_j : j \in J\})$, by proposition 3.13, it suffices to find $d_{(I,J)} \in \mathbb{R}^n$ such that

$$\Big(-v_i^\mathsf{T} d_{(I,J)} > 0, \quad \forall\, i \in I\Big) \qquad \text{and} \qquad \Big(v_j^\mathsf{T} d_{(I,J)} > 0, \quad \forall\, j \in J\Big). \tag{3.15}$$

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

$$\max_{i\in I} \frac{\xi^\mathsf{T} v_i}{v_i^\mathsf{T} d} < -\cot\theta < \min_{j\in J} \frac{\xi^\mathsf{T} v_j}{v_j^\mathsf{T} d}. \tag{3.16}$$

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

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

Therefore, (3.15) is satisfied with $d_{(I,J)} := (\cos\theta)d + (\sin\theta)\xi$. □

One can now establish the link between the pointed cone problem of section 3.3.1 (problem 3.12) and the linearly separable bipartitioning problem (problem 3.15).

Equivalence 3.18 (pointed cone $\leftrightarrow$ linearly separable bipartitioning) Let be given a matrix $V \in \mathbb{R}^{n\times p}$ with nonzero columns denoted by $v_1, \ldots, v_p$ and take $s \in \mathcal{S}$, which is nonempty. By (3.13), $\mathrm{cone}\{s_i v_i : i \in [1:p]\}$ is pointed. Use the construction 3.16(1) with $v_i \curvearrowright s_i v_i$.

For $\tilde{s} \in \{\pm 1\}^p$, define a partition (I, J) of $[1:p]$ by

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

Define also K_I and K_J by (3.14a) with $v_i \curvearrowright s_i v_i$. We claim that

$$\mathrm{cone}\{\tilde{s}_i v_i : i \in [1:p]\} \text{ is pointed} \quad \Longleftrightarrow \quad \exists\, \xi \in \mathbb{A}_0 : \max_{i\in I} \xi^{\mathsf{T}} \bar{v}_i < \min_{j\in J} \xi^{\mathsf{T}} \bar{v}_j. \tag{3.17}$$

Indeed, one has

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

where we have used the equivalence $(i) \Leftrightarrow (iv)$ of proposition 3.17 ($v_i \curvearrowright s_i v_i$).

The equivalence (3.17) establishes the expected equivalence between the pointed cone problem 3.12 (in which one looks for all the $\tilde{s} \in \{\pm 1\}^p$ such that $\mathrm{cone}\{\tilde{s}_i v_i : i \in [1:p]\}$ is pointed) and the linearly separable bipartitioning problem 3.15 of the vectors $\bar{v}_i = s_i v_i / (s_i v_i^{\mathsf{T}} d) = v_i / (v_i^{\mathsf{T}} d)$, $i \in [1:p]$, where d is associated with the pointed cone $\mathrm{cone}\{s_i v_i : i \in [1:p]\}$ by the equivalence $(i) \Leftrightarrow (iii)$ of proposition 3.13. □

3.4 Discrete geometry: hyperplane arrangements

The equivalent problem examined in this section has a long history, going back at least to the XIXth century [69, 80]. More recently, it appears in *Computational Discrete Geometry* (the discipline has many other names), under the name of *hyperplane arrangements*. Contributions to this problem, or a more general version of it, with a discrete mathematics point of view, have been reviewed in [2, 36, 46, 47, 78]. It has many applications [18, 37, 76]. From an algorithmic point of view, the algorithms developed in this domain can immediately be used to compute $\mathcal{S}$ defined by (3.2) or $\partial_B H(x)$ defined by (1.1) and (1.3).

Problem 3.19 (arrangement of hyperplanes containing the origin) Let be given two positive integers n and $p \in \mathbb{N}^*$ and p nonzero vectors $v_1, \ldots, v_p \in \mathbb{R}^n$. Consider the hyperplanes containing the origin:

$$\mathcal{H}_i := \{d \in \mathbb{R}^n : v_i^{\mathsf{T}} d = 0\}. \tag{3.18}$$

Figure 3.3 illustrates problem 3.19 for the linear complementarity problem 3.2. It is requested to list the regions of $\mathbb{R}^n$ that are separated by these hyperplanes, which are the connected components of $\mathbb{R}^n \setminus (\bigcup_{i\in[1:p]} \mathcal{H}_i)$. Such a region is called a *cell* or a *chamber*, depending on the authors [2, 5, 75]. More specifically, let us define the half-spaces

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

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

$$\mathfrak{C} := \left\{ (I_+, I_-) \in \mathfrak{B}([1:p]) : (\cap_{i \in I_+} \mathcal{H}_i^+) \cap (\cap_{i \in I_-} \mathcal{H}_i^-) \neq \varnothing \right\}, \tag{3.19}$$

where $\mathfrak{B}([1:p])$ denotes the set of bipartitions of $[1:p]$. □

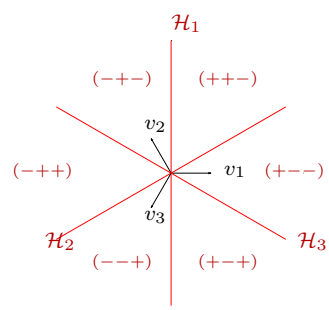

Figure 3.3: Illustration of problem 3.19 (arrangement of hyperplanes containing the origin) for the 3 vectors that are the columns on the matrix V in example 3.2 (since the last components of these v_i's vanish, only the first two ones are represented above). The hyperplanes $\mathcal{H}_i$ are defined by (3.18). The regions to determine are represented by the sign vectors here denoted $(s_1 s_2 s_3)$ with $s_i = \pm$: if $d \in \mathbb{R}^2$ belongs to the region $(s_1 s_2 s_3)$, then $s_i = +$ if $v_i^\mathsf{T} d > 0$ and $s_i = -$ if $v_i^\mathsf{T} d < 0$. We see that there are only $6 = 2p$ regions among the $8 = 2^p$ possible ones; the regions $(+++)$ and $(---)$ are missing, which reflects the fact that $+v_1 + v_2 + v_3 = 0$ and $-v_1 - v_2 - v_3 = 0$ (see problem 3.6).

The link between problem 3.19 and the signed feasibility of strict linear inequality systems of section 3.2.1 is obtained from the bijection

$$\eta : (I_+, I_-) \in \mathfrak{B}([1:p]) \mapsto s \in \{\pm 1\}^p, \text{ where } s_i = \begin{cases} +1 & \text{if } i \in I_+, \\ -1 & \text{if } i \in I_- \end{cases} \tag{3.20}$$

and the setting $V = \begin{pmatrix} v_1 & \cdots & v_p \end{pmatrix}$, whose columns are nonzero by assumption, here and in section 3.2.1. Recall the definition (3.2) of the set of sign vectors $\mathcal{S}$.

Proposition 3.20 (bijection $\mathfrak{C} \leftrightarrow \mathcal{S}$) *For the matrix $V \in \mathbb{R}^{n \times p}$, with nonzero columns v_i's, the map η given by (3.20) is a bijection from $\mathfrak{C}$ onto $\mathcal{S}$.*

Proof. Let $(I_+, I_-) \in \mathfrak{B}([1:p])$ and $s := \eta((I_+, I_-))$. Then,

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

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

Equivalence 3.21 (signed linear system feasibility $\leftrightarrow$ hyperplane arrangement) The equivalence between problems 3.3 and 3.19 follows from the bijection of the map $\eta : \mathfrak{C} \to \mathcal{S}$ claimed in proposition 3.20. □

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 given by (1.3) and $x \in \mathbb{R}^n$. This description is often carried out in terms of the matrix V defined by (3.1), whose p columns are denoted by $v_1, \ldots, v_p \in \mathbb{R}^n$ and are nonzero by construction. When the properties are given for $\mathcal{S}$, one may have $p \geqslant n$ and the referenced matrix $V \in \mathbb{R}^{n\times p}$ is *assumed* to have nonzero columns, which implies that $\mathcal{S} \neq \varnothing$. Some properties of $\partial_B H(x)$ are given in section 4.1, including those that are useful in [33]. Section 4.2 deals with the cardinality $|\partial_B H(x)|$ of the B-differential. Section 4.3 analyzes more precisely two particular configurations. Section 4.4 highlights two links between the B-differential and the C-differential of H.

Besides their theoretical relevance, the properties of the B-differential of H given in this section will also be useful to design the algorithms presented in section 5 and to check the correctness of their implementation.

4.1 Some properties of the B-differential

Let us start with a basic property of $\partial_B H(x)$, which is its symmetry in the sense of definitions 2.5. This property has been observed by many in other contexts [2; § 1.1.4]. The equivalence 3.5 allows us to give a straightforward proof. It is useful for the algorithms since it implies that only half of the B-differential has to be computed.

Proposition 4.1 (symmetry of $\partial_B H(x)$) *Suppose that $p > 0$. Then, the B-differential $\partial_B H(x)$ is symmetric and $|\partial_B H(x)|$ is even.*

Proof. Let $J \in \partial_B H(x)$ and $s := \sigma(J)$, where σ is defined by (3.4). By proposition 3.4, $s \bullet (V^\mathsf{T} d) > 0$ holds for some d. Now, $(-s) \bullet (V^\mathsf{T} d) > 0$ obviously also holds for some d (take the opposite of the previous d as solution), so that $-s \in \mathcal{S}$. By the definition of the bijection $\sigma : \partial_B H(x) \to \mathcal{S}$, 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. □

We now give a necessary and sufficient condition ensuring the completeness of $\partial_B H(x)$ in the sense of definition 2.4. The condition was shown to be sufficient in [85; corollary 2.1(i)] for the nonlinear case (1.6), using a different proof, but we shall see in [33] that it is an easy consequence of that property in the affine case (1.3). Thanks to the equivalence 3.5, the present proof is short. This property is also useful in the development of algorithms, as a test that these must pass: $|\partial_B H(x)| = 2^p$ if and only if $V \in \mathbb{R}^{n\times p}$ is injective.

Proposition 4.2 (completeness of the B-differential) *The B-differential $\partial_B H(x)$ of H at x is complete if and only if the matrix $V \in \mathbb{R}^{n\times p}$ in (3.1) is injective. Hence, this property can hold only if $p \leqslant n$.*

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

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

By Gordan's alternative (3.7), it follows that there is no $d \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d) > 0$. By (3.2), this implies that $s \notin \mathcal{S}$. According to the equivalence 3.5, $\sigma^{-1}(s) \notin \partial_B H(x)$, showing that the B-differential is not complete.

[$\Leftarrow$] Assume the injectivity of V. Let $s \in \{\pm 1\}^p$. Since V^T is surjective, the system $V^\mathsf{T} d = s$ holds for some $d \in \mathbb{R}^n$. For this d, $s \bullet (V^\mathsf{T} d) = e$, so that $s \bullet (V^\mathsf{T} d) > 0$ holds for some $d \in \mathbb{R}^n$, which implies that the selected s is in $\mathcal{S}$. We have shown that $\mathcal{S} = \{\pm 1\}^p$ or that $\partial_B H(x) = \sigma^{-1}(\{\pm 1\}^p)$ (σ^{-1} is defined by (3.4b)) is complete. □

We focus now on the connectivity of $\partial_B H(x)$, a notion that is more easily presented in terms of $\mathcal{S} \subseteq \{\pm 1\}^p$ but that can be transferred straightforwardly to $\partial_B H(x)$ by the bijection σ defined in (3.4). This property was implicitly used, for instance, in the algorithms proposed by Avis, Fukuda and Sleumer [5, 75] for hyperplane arrangements.

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 $\mathrm{co}\{\pm 1\}^p$ can be joined by a single edge). □

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 $s^0, \ldots, s^l \in S$ 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. □

Proposition 4.5 (connectivity of the B-differential) *The set $\mathcal{S}$ defined by (3.2) 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 $l := \sum_{i \in [1\,:\,p]} |\tilde{s}_i - s_i|/2 \leqslant p$ 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 (3.2), for any $s \in \mathcal{S} \neq \varnothing$, one can find $d \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d) > 0$, implying that $s_i = s_j$ (resp. $s_i = -s_j$). Therefore, one cannot find a path in $\mathcal{S}$ joining $s \in \mathcal{S}$ and $-s \in \mathcal{S}$ (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 (3.2) of $\mathcal{S}$, one can find d and $\tilde{d} \in \mathbb{R}^n$ such that

$$s \bullet (V^\mathsf{T} d) > 0 \qquad \text{and} \qquad \tilde{s} \bullet (V^\mathsf{T} \tilde{d}) > 0.$$

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

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

Since a small modification of $\tilde{d}$ preserves the inequality $\tilde{s} \bullet (V^\mathsf{T} \tilde{d}) > 0$ and since the vectors $\bar{v}_i$'s are all different, lemma 2.6 tells us that, at the cost of a small change of ξ, one can assume that

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

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

$$v_i^\mathsf{T}\xi \neq 0, \quad \forall\, i \in [1:p]. \tag{4.2b}$$

Then, one can set $t_i := -(v_i^\mathsf{T} d)/(v_i^\mathsf{T}\xi)$, for $i \in [1:p]$, which are p distinct values by (4.2a). It results that the following equivalent expressions hold for all $i \in [1:p]$:

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

For each $i \in [1:p]$, we are interested in the change of sign of $v_i^\mathsf{T}[(1-t)d + t\tilde{d}]$ 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., $v_i^\mathsf{T} d$ and $v_i^\mathsf{T}\tilde{d}$ have opposite signs). Therefore, the number of $t_i \in (0,1)$ is equal to $l = \sum_{i\in[1:p]} |\tilde{s}_i - s_i|/2 \leqslant p$. Let us denote them by

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

When $t \in (0,1)$ crosses a $t_j \in (0,1)$, a single $v_i^\mathsf{T}[(1-t)d + t\tilde{d}]$, 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

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

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 $l \leqslant p$ in $\mathcal{S}$, namely $s^{i_0} = s$, s^{i_1}, …, $s^{i_l} = \tilde{s}$, joining s to $\tilde{s}$. This proves the implication. □

This connectivity property is also stated in [2; section 1.10.4] as a simple observation 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 3.15). This one consists in dragging a separating hyperplane, thus separating successively from 1 to l points from the others in a set of p points lying in an affine space. Let us give the details.

For $k \in [1:p]$, define $\nu^k \in \{\pm 1\}^p$ by

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

Hence, "$\nu^k\bullet$" applied to a vector reverses the sign of its kth component.

Another proof of proposition 4.5[$\Leftarrow$]. 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 $\tilde{s}$.

The affine space $\mathbb{A}$ in which points $\bar{v}_i$ are going to be seperated is given by the construction 3.16(1) with $v_i \curvearrowright s_i v_i$. This one yields a direction d, the affine space $\mathbb{A} := \{\bar{v} : d^\mathsf{T}\bar{v} = 1\}$, its associated vector space $\mathbb{A}_0 = \{v : d^\mathsf{T} v = 0\}$ and the vectors $\bar{v}_i := s_i v_i/(s_i v_i^\mathsf{T} d) = v_i/(v_i^\mathsf{T} d) \in \mathbb{A}$ for $i \in [1:p]$.

Let us now define a separating covector $\xi \in \mathbb{A}_0$. For this purpose, one introduces the following bipartition (I, J) of $[1:p]$:

$$I := \{i \in [1:p] : \tilde{s}_i = -s_i\} \qquad \text{and} \qquad J := \{j \in [1:p] : \tilde{s}_j = s_j\}. \tag{4.4a}$$

Clearly, $|I| = l$ (the number given in the statement of the proposition) and $|J| = p - l$. Let us now apply the construction 3.16(2) to get the cones K_I and K_J. Since s and $\tilde{s} \in \mathcal{S}$,

(3.13) tells us that $\mathrm{cone}\{s_i v_i : i \in [1:p]\}$ and $\mathrm{cone}\{\tilde{s}_i v_i : i \in [1:p]\}$ are pointed. Now, $\mathrm{cone}\{s_i v_i : i \in [1:p]\} = \mathrm{cone}(K_I \cup K_J)$ and $\mathrm{cone}\{\tilde{s}_i v_i : i \in [1:p]\} = \mathrm{cone}((-K_I) \cup K_J)$. The pointedness of $\mathrm{cone}(K_I \cup K_J)$ and $\mathrm{cone}((-K_I) \cup K_J)$ and the implication $(i) \Rightarrow (iv)$ of proposition 3.17 imply that there exists a covector $\xi \in \mathbb{A}_0$ such that $\xi^\mathsf{T} \bar{v}_i < \xi^\mathsf{T} \bar{v}_j$ 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 ξ and, by lemma 2.6, ξ can be chosen such that all the $\xi^\mathsf{T} \bar{v}_i$ for $i \in [1:p]$ 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

$$\xi^\mathsf{T} \bar{v}_{i_1} < \xi^\mathsf{T} \bar{v}_{i_2} < \cdots < \xi^\mathsf{T} \bar{v}_{i_l} < \xi^\mathsf{T} \bar{v}_{j_1} < \xi^\mathsf{T} \bar{v}_{j_2} < \cdots < \xi^\mathsf{T} \bar{v}_{j_{p-l}}. \tag{4.4b}$$

Let us now define the path in $\mathcal{S}$ from s to $\tilde{s}$. For $k \in [0:l]$, define $s^{i_k} \in \{\pm 1\}^p$ as follows

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

where ν^k is defined by (4.3). We claim that

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

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

The fact that $(s^{i_0},\, s^{i_1},\, \ldots,\, s^{i_l})$ 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 $(s^{i_0},\, s^{i_1},\, \ldots,\, s^{i_l})$ joins s to $\tilde{s}$. It remains to show that the s^{i_k}'s are in $\mathcal{S}$. Define, for $k \in [1:l]$:

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

By (4.4b), the hyperplane $\{\bar{v} \in \mathbb{A} : \xi^\mathsf{T} \bar{v} = \alpha_k\}$ 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 (3.14b), $C_{I_k} \cap C_{J_k} = \varnothing$. By the implication $(iii) \Rightarrow (i)$ of proposition 3.17, this implies that $\mathrm{cone}((-K_{I_k}) \cup K_{J_k})$ is pointed. By the implication $(i) \Rightarrow (iii)$ of proposition 3.13, the system

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

has a solution $\tilde{d} \in \mathbb{R}^n$. By (3.2), this amounts to saying that $s^{i_k} = (-s_{I_k}, s_{J_k})$ is in $\mathcal{S}$, as expected. □

Definition 4.6 (extremality in $\{\pm 1\}^p$) A point $\bar{v}_k$ of a finite set $\overline{\mathfrak{V}} := \{\bar{v}_i \in \mathbb{R}^n : i \in [1:p]\}$ of $\mathbb{R}^n$ is said to be an *extreme point* of $\overline{\mathfrak{V}}$ if $\bar{v}_k \notin \mathrm{co}\{\bar{v}_i : i \in [1:p] \setminus \{k\}\}$. □

The next proposition shows 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 $\nu^k \bullet s$, where ν^k is defined by (4.3) and $k \in [1:p]$ is such that $\bar{v}_k$ is an extreme point of the set of vectors $\bar{v}_i$ ($i \in [1:p]$) defined by the construction 3.16(1), with $v_i \curvearrowright s_i v_i$. Note that these vectors $\bar{v}_i$'s depend on $s \in \mathcal{S}$ by this construction.

Proposition 4.7 (adjacency and extremality) *Let $s \in \mathcal{S}$ and adopt the construction 3.16(1) with $v_i \curvearrowright s_i v_i$, which yields the set $\overline{\mathfrak{V}} := \{\bar{v}_i : i \in [1:p]\}$. For some $k \in [1:p]$, let ν^k be defined by (4.3). Then, the following properties are equivalent:*

(i) $\nu^k \bullet s \in \mathcal{S}$,

(ii) $\bar{v}_k$ *is an extreme point of* $\overline{\mathfrak{V}}$.

Proof. Since $s \in \mathcal{S}$, (3.13) implies that $\operatorname{cone}\{s_i v_i : i \in [1:p]\}$ is pointed. Define $\tilde{s} := \nu^k \bullet s$. We have the following equivalences

$$\begin{aligned}
\tilde{s} \in \mathcal{S} \quad &\iff \quad \operatorname{cone}\{\tilde{s}_i v_i : i \in [1:p]\} \text{ is pointed} \qquad [(3.13)] \\
&\iff \quad \bar{v}_k \notin \operatorname{co}\{\bar{v}_i : i \in [1:p] \setminus \{k\}\} \qquad [(i) \Leftrightarrow (iii) \text{ in proposition 3.17}] \\
&\iff \quad \bar{v}_k \text{ is an extreme point of } \overline{\mathfrak{V}} \qquad [\text{definition}].
\end{aligned}$$

□

For $k \in [1:p]$, we introduce

$$\mathcal{S}_k := \{s \in \{\pm 1\}^k : \ \exists\, d \in \mathbb{R}^n \text{ such that } s_i v_i^\mathsf{T} d > 0 \text{ for } i \in [1:k]\}. \tag{4.5}$$

We also note $\mathcal{S}_k^c := \{\pm 1\}^k \setminus \mathcal{S}_k$. Hence $\mathcal{S} = \mathcal{S}_p$ and $\mathcal{S}^c = \mathcal{S}_p^c$. Point 1 of the next proposition will be used to motivate an improvement of algorithm 5.6 in section 5.2.5 and its points 2 and 3 will be used to get the equivalence in proposition 4.18, related to a fan arrangement.

Proposition 4.8 (incrementation) 1) *If $s \in \mathcal{S}_k^c$, then $(s, \pm 1) \in \mathcal{S}_{k+1}^c$. In particular, $|\mathcal{S}_{k+1}^c| \geqslant 2|\mathcal{S}_k^c|$.*

2) *If $v_{k+1} \notin \operatorname{vect}\{v_1, \dots, v_k\}$, then, $(s, \pm 1) \in \mathcal{S}_{k+1}$ for all $s \in \mathcal{S}_k$. In particular, $|\mathcal{S}_{k+1}| = 2|\mathcal{S}_k|$ and $|\mathcal{S}_{k+1}^c| = 2|\mathcal{S}_k^c|$.*

3) *If v_{k+1} is not colinear to any of the vectors $v_1, \dots, v_k$, then, $[(s, \pm 1)$ and $(-s, \pm 1) \in \mathcal{S}_{k+1}$ for one $s \in \mathcal{S}_k]$ and $[(s', +1)$ or $(s', -1) \in \mathcal{S}_{k+1}$ for any $s' \in \mathcal{S}_k]$. In particular, $|\mathcal{S}_{k+1}| \geqslant |\mathcal{S}_k| + 2$.*

Proof. 1) If $s \in \mathcal{S}_k^c$, there is no $d \in \mathbb{R}^n$ such that $s_i v_i^\mathsf{T} d > 0$ for $i \in [1:k]$. Therefore, there is no $d \in \mathbb{R}^n$ such that ($s_i v_i^\mathsf{T} d > 0$ for $i \in [1:k]$) and $\pm v_{k+1}^\mathsf{T} d > 0$. Therefore, $(s, \pm 1) \in \mathcal{S}_{k+1}^c$. This implies that $|\mathcal{S}_{k+1}^c| \geqslant 2|\mathcal{S}_k^c|$.

2) Let P be the orthogonal projector on $\operatorname{vect}\{v_1, \dots, v_k\}^\perp$ for the Euclidean scalar product. By assumption, $\mathrm{P}\, v_{k+1} \neq 0$. Let $s \in \mathcal{S}_k$, so that there is a direction $d \in \mathbb{R}^n$ such that $s_i v_i^\mathsf{T} d > 0$ for $i \in [1:k]$. For any $t \in \mathbb{R}$ and $i \in [1:k]$, the directions $d_\pm := d \pm t\, \mathrm{P}\, v_{k+1}$ verify $s_i v_i^\mathsf{T} d_\pm = s_i v_i^\mathsf{T} d > 0$ (because $v_i^\mathsf{T}\, \mathrm{P}\, v_{k+1} = 0$). In addition, for $t > 0$ sufficiently large, one has $\pm v_{k+1}^\mathsf{T} d_\pm = \pm v_{k+1}^\mathsf{T} d + t \| \mathrm{P}\, v_{k+1} \|^2 > 0$ (because $\mathrm{P}^2 = \mathrm{P}$ and $\mathrm{P}^\mathsf{T} = \mathrm{P}$). We have shown that both $(s, +1)$ and $(s, -1)$ are in $\mathcal{S}_{k+1}$. Therefore, $|\mathcal{S}_{k+1}| \geqslant 2|\mathcal{S}_k|$.

Now, $|\mathcal{S}_k| + |\mathcal{S}_k^c| = 2^k$, $|\mathcal{S}_{k+1}| + |\mathcal{S}_{k+1}^c| = 2^{k+1}$ and $|\mathcal{S}_{k+1}^c| \geqslant 2|\mathcal{S}_k^c|$ by point 1. Therefore, one must have $|\mathcal{S}_{k+1}| = 2|\mathcal{S}_k|$ and $|\mathcal{S}_{k+1}^c| = 2|\mathcal{S}_k^c|$.

3) We claim that one can find a direction $d \in \mathbb{R}^n$ such that

$$\Big(\forall\, i \in [1:k] : \ v_i^\mathsf{T} d \neq 0\Big) \qquad \text{and} \qquad v_{k+1}^\mathsf{T} d = 0. \tag{4.6}$$

Indeed, let $\mathbb{E} := \{d \in \mathbb{R}^n : v_{k+1}^\mathsf{T} d = 0\}$ and P be the orthogonal projector on $\mathbb{E}$ for the Euclidean scalar product. By lemma 2.6, one can find a direction $d \in \mathbb{E}$ (hence $v_{k+1}^\mathsf{T} d = 0$) such that $|\{(\mathrm{P}\, v_i)^\mathsf{T} d : i \in [1:k+1]\}| = |\{\mathrm{P}\, v_i : i \in [1:k+1]\}|$. Since $\mathrm{P}\, v_{k+1} = 0$ and $\mathrm{P}\, v_i \neq 0$ for $i \in [1:k]$ (because the v_i's are not colinear with v_{k+1}), one has $(\mathrm{P}\, v_i)^\mathsf{T} d \neq 0$ for $i \in [1:k]$. Since, $0 \neq (\mathrm{P}\, v_i)^\mathsf{T} d = v_i^\mathsf{T}\, \mathrm{P}\, d = v_i^\mathsf{T} d$, (4.6) follows.

Taking $s_i := \mathrm{sgn}(v_i^\mathsf{T} d)$ for $i \in [1:k]$, one deduces from (4.6) that there is a direction $d \in \mathbb{R}^n$ such that

$$\Big(\forall\, i \in [1:k] :\ s_i v_i^\mathsf{T} d > 0\Big) \qquad \text{and} \qquad v_{k+1}^\mathsf{T} d = 0.$$

It follows that, for $\varepsilon > 0$ sufficiently small, the directions $d_\pm := d \pm \varepsilon v_{k+1}$ satisfy

$$\Big(\forall\, i \in [1:k] :\ s_i v_i^\mathsf{T} d_\pm > 0\Big) \qquad \text{and} \qquad \pm\, v_{k+1}^\mathsf{T} d_\pm > 0.$$

This means that $(s, \pm 1) \in \mathcal{S}_{k+1}$. By symmetry (proposition 4.1), one also has $(-s, \pm 1) \in \mathcal{S}_{k+1}$, so that we have found 4 vectors in $\mathcal{S}_{k+1}$. Now, since, for any $s' \in \mathcal{S}_k \backslash \{\pm s\}$ (in number $|\mathcal{S}_k| - 2$), either $(s', +1) \in \mathcal{S}_{k+1}$ or $(s', -1) \in \mathcal{S}_{k+1}$, it follows that $|\mathcal{S}_{k+1}| \geqslant 4 + (|\mathcal{S}_k| - 2) = |\mathcal{S}_k| + 2$. $\square$

4.2 Cardinality of the B-differential

Information on the cardinality of $\partial_B H(x)$ can be useful to check the correctness of the number of elements computed by the algorithms presented in section 5.2.

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 (3.2), (3.19) and (3.6), 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 3.2). Many contributions have been done on this subject; the earliest we cite dates from 1826 [2–4, 21, 36, 46, 69, 78, 80, 82, 86]. The formula (4.7) for $|\partial_B H(x)|$ is due to Winder [84; 1966] (see [79] for a recent and short review) and reads for the matrix V with nonzero columns given by (3.1)

$$\boxed{|\partial_B H(x)| = \sum_{I \subseteq [1:p]} (-1)^{\mathrm{null}(V_{:,I})},} \tag{4.7}$$

where $\mathrm{null}(V_{:,I})$ is the nullity of $V_{:,I}$ and the term in the right-hand side corresponding to $I = \varnothing$ is 1 (one takes the convention that $\mathrm{null}(V_{:,\varnothing}) = 0$). Note that, in this formula, the columns of V can be colinear with each other. 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 $\partial_B H(x)$.

We give below a proof of (4.7) that follows the same line of reasoning as the one of Winder [84], but that is more analytic in that it uses the sign vectors introduced in section 3.2.1 rather than geometric arguments. The proof uses the following lemma of general interest. For $k \in [1:p-1]$, one defines

$$\begin{aligned} V_k &:= V_{:,[1:k]}, \\ \mathcal{S}(V_k) &:= \{s \in \{\pm 1\}^k : \exists\, d \in \mathbb{R}^n \text{ such that } s_i v_i^\mathsf{T} d > 0, \text{ for } i \in [1:k]\} \end{aligned}$$

and denotes by P_{k+1} the orthogonal projector on $v_{k+1}^\perp$. Below, P_{k+1} is also the matrix representation of the projector, hence $\mathrm{P}_{k+1}\, V_k$ is the product of two matrices. The term "descendant"

used in the following proposition will find an explanation in the algorithmic part of the paper, in section 5.2.3.

Proposition 4.9 (sign vector with two descendants) *Suppose that* $s \in \mathcal{S}(V_k)$. *Then,*

$$\begin{aligned}
&(s,+1) \text{ and } (s,-1) \in \mathcal{S}(V_{k+1}) \\
&\iff \exists\, d \in \mathbb{R}^n : s_i v_i^\mathsf{T} d > 0, \text{ for } i \in [1:k], \text{ and } v_{k+1}^\mathsf{T} d = 0, && (4.8a)\\
&\iff s \in \mathcal{S}(\mathrm{P}_{k+1} V_k). && (4.8b)
\end{aligned}$$

Proof. One has the following equivalences ((4.9a) and (4.9b) are justified afterwards):

$$\begin{aligned}
&(s,+1) \text{ and } (s,-1) \in \mathcal{S}(V_{k+1}) \\
&\iff \begin{cases} \exists\, d_+ \in \mathbb{R}^n : s_i v_i^\mathsf{T} d_+ > 0, \text{ for } i \in [1:k], \text{ and } +v_{k+1}^\mathsf{T} d_+ > 0 \\ \exists\, d_- \in \mathbb{R}^n : s_i v_i^\mathsf{T} d_- > 0, \text{ for } i \in [1:k], \text{ and } -v_{k+1}^\mathsf{T} d_- > 0 \end{cases} \\
&\iff \exists\, d \in \mathbb{R}^n : s_i v_i^\mathsf{T} d > 0, \text{ for } i \in [1:k], \text{ and } v_{k+1}^\mathsf{T} d = 0 && (4.9a)\\
&\iff \exists\, d \in \mathbb{R}^n : s_i (\mathrm{P}_{k+1} v_i)^\mathsf{T} d > 0, \text{ for } i \in [1:k] && (4.9b)\\
&\iff s \in \mathcal{S}(\mathrm{P}_{k+1} V_k).
\end{aligned}$$

The equivalence in (4.9a) is shown as follows:

[⇒] $d = (-v_{k+1}^\mathsf{T} d_-) d_+ + (v_{k+1}^\mathsf{T} d_+) d_- \in v_{k+1}^\perp$ is appropriate;
[⇐] take $d_\pm = d \pm \varepsilon v_{k+1}$ for a sufficiently small $\varepsilon > 0$.

The equivalence in (4.9b) is shown as follows:

[⇒] $d = \mathrm{P}_{k+1}\, d$ (since $v_{k+1}^\mathsf{T} d = 0$) and $\mathrm{P}_{k+1}^\mathsf{T} = \mathrm{P}_{k+1}$ (since P_{k+1} is an orthogonal projector), so that $0 < s_i v_i^\mathsf{T} d = s_i v_i^\mathsf{T} (\mathrm{P}_{k+1}\, d) = s_i (\mathrm{P}_{k+1}\, v_i)^\mathsf{T} d$;
[⇐] take $d := \mathrm{P}_{k+1}\, d_0$ in (4.9a), where d_0 is the d given by (4.9b). □

Knowing a direction d such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1:k]$, the equivalence (4.8a) questions the validity of the following equivalence

$$(s,+1) \text{ and } (s,-1) \in \mathcal{S}(V_{k+1}) \quad \overset{?}{\iff} \quad s_i v_i^\mathsf{T} (\mathrm{P}_{k+1}\, d) > 0, \text{ for } i \in [1:k].$$

The implication "⇐" is certainly true by the implication "⇐" in (4.8a) (with $d \curvearrowright \mathrm{P}_{k+1}\, d$), but the reverse implication "⇒" is not true in general, as shown by the following counterexample.

Counter-example 4.10 (no double descendant test with $\mathbf{P}_{k+1}\, d$) Suppose that $k = 1$, $v_1 = \binom{1}{1}$, $v_2 = \binom{0}{1}$, $s = +1$ and $d = \binom{-1}{2}$. One has indeed $s v_1^\mathsf{T} d = 1 > 0$. Now, $(s, \pm 1) \in \mathcal{S}(V_2)$ since $v_1^\mathsf{T} d_+ = 1 > 0$ and $v_2^\mathsf{T} d_+ = 2 > 0$ with $d_+ = d = \binom{-1}{2}$ and $v_1^\mathsf{T} d_- = 1 > 0$ and $-v_2^\mathsf{T} d_- = 1 > 0$ with $d_- = \binom{2}{-1}$. However, $\mathrm{P}_2\, d = \binom{-1}{0}$, so that $v_1^\mathsf{T}\, \mathrm{P}_2\, d = -1 < 0$. □

Proof of (4.7) *along Winder's proof* [84]. The proof is by induction on p. The result is true for $p = 1$, since then the right-hand side of (4.7) is 2 (I can be $\varnothing$ and $\{1\}$ and, in each of these cases, $\mathrm{null}(V_{:,I}) = 0$), which is indeed the number of regions seperated by the hyperplane $v_1^\perp$.

Suppose now that the identity (4.7) holds for V_k, with $k \in [1:p-1]$, and let us prove it for V_{k+1}. We do this in three steps.

1) Observe first that

$$|\mathcal{S}(V_{k+1})| = |\mathcal{S}(V_k)| + |\mathcal{S}(\mathrm{P}_{k+1} V_k)|, \tag{4.10a}$$

where P_{k+1} is the orthogonal projector on $v_{k+1}^{\perp}$. Indeed, if $s \in \mathcal{S}(V_k)$, then $(s, +1)$ and/or $(s, -1) \in \mathcal{S}(V_{k+1})$. According to proposition 4.9, the number of times both $(s, +1)$ and $(s, -1)$ are in $\mathcal{S}(V_{k+1})$ is $|\mathcal{S}(\mathrm{P}_{k+1} V_k)|$ and this brings $2|\mathcal{S}(\mathrm{P}_{k+1} V_k)|$ elements to $|\mathcal{S}(V_{k+1})|$. As a result, the number of times only one descendant $(s, +1)$ or $(s, -1)$ is in $\mathcal{S}(V_{k+1})$ is $|\mathcal{S}(V_k)| - |\mathcal{S}(\mathrm{P}_{k+1} V_k)|$ and this brings $|\mathcal{S}(V_k)| - |\mathcal{S}(\mathrm{P}_{k+1} V_k)|$ elements to $|\mathcal{S}(V_{k+1})|$. We conclude that $|\mathcal{S}(V_{k+1})| = (2|\mathcal{S}(\mathrm{P}_{k+1} V_k)|) + (|\mathcal{S}(V_k)| - |\mathcal{S}(\mathrm{P}_{k+1} V_k)|) = |\mathcal{S}(V_k)| + |\mathcal{S}(\mathrm{P}_{k+1} V_k)|$.

2) We claim now that we can assume that, for any $i \in [1:k]$, v_{k+1} is not colinear to v_i or, equivalently, $\mathrm{P}_{k+1} v_i \neq 0$. Suppose indeed that v_{k+1} is colinear to v_i, for some $i \in [1:k]$. One has

$$\sum_{I \subseteq [1:k+1]} (-1)^{\mathrm{null}(V_{:,I})} = \sum_{I \subseteq [1:k]} (-1)^{\mathrm{null}(V_{:,I})} + \sum_{\substack{I \subseteq [1:k+1] \\ i \notin I \\ k+1 \in I}} (-1)^{\mathrm{null}(V_{:,I})} + \sum_{\substack{I \subseteq [1:k+1] \\ i \in I \\ k+1 \in I}} (-1)^{\mathrm{null}(V_{:,I})}.$$

The second term in the right-hand side is equal to the first one, since one has the same vectors (up to a multiplicative factor) by taking $I \subseteq [1:k+1]$ with $i \notin I$ and $k+1 \in I$ or by taking $I \subseteq [1:k]$. Furthermore, the last term in the right-hand side is equal to

$$\sum_{I \subseteq [1:k]} (-1)^{|I|+1-\mathrm{rank}(V_{:,I})} = -\sum_{I \subseteq [1:k]} (-1)^{\mathrm{null}(V_{:,I})}.$$

Therefore, the last two terms in the right-hand side of the first identity cancel each other. In conclusion adding a vector v_{k+1} that is colinear to a previous v_i does not modify the right-hand side of (4.7), which is the expected behavior since such a vector does not modify $\mathcal{S}(V_k)$ and can be discarded.

3) By the induction assumption and the fact that V_k and $(\mathrm{P}_{k+1} V_k)$ have k nonzero columns (by assumption and point 2), one has

$$|\mathcal{S}(V_k)| = \sum_{I \subseteq [1:k]} (-1)^{\mathrm{null}(V_{:,I})} \qquad \text{and} \qquad |\mathcal{S}(\mathrm{P}_{k+1} V_k)| = \sum_{I \subseteq [1:k]} (-1)^{\mathrm{null}((\mathrm{P}_{k+1} V)_{:,I})}.$$

Therefore, due to (4.10a), the proof will be complete if we establish that

$$\sum_{I \subseteq [1:k]} (-1)^{\mathrm{null}((\mathrm{P}_{k+1} V)_{:,I})} = \sum_{\substack{I \subseteq [1:k+1] \\ k+1 \in I}} (-1)^{\mathrm{null}(V_{:,I})}. \tag{4.10b}$$

Let us show that, for $I \subseteq [1:k]$,

$$\mathcal{N}(V_{:,I \cup \{k+1\}}^{\mathsf{T}}) = \mathcal{N}(V_{:,I}^{\mathsf{T}} \mathrm{P}_{k+1}) \cap v_{k+1}^{\perp}. \tag{4.10c}$$

[$\subseteq$] If $V_{:,I \cup \{k+1\}}^{\mathsf{T}} d = 0$, one has $V_{:,I}^{\mathsf{T}} d = 0$ and $v_{k+1}^{\mathsf{T}} d = 0$. Therefore $d \in v_{k+1}^{\perp}$ and $d = \mathrm{P}_{k+1} d$. It follows that $V_{:,I}^{\mathsf{T}} \mathrm{P}_{k+1} d = 0$, meaning that $d \in \mathcal{N}(V_{:,I}^{\mathsf{T}} \mathrm{P}_{k+1})$.

[$\supseteq$] If $V_{:,I}^{\mathsf{T}} \mathrm{P}_{k+1} d = 0$ and $d \in v_{k+1}^{\perp}$, one has $\mathrm{P}_{k+1} d = d$ and $v_{k+1}^{\mathsf{T}} d = 0$. Therefore $V_{:,I \cup \{k+1\}}^{\mathsf{T}} d = 0$.

Taking the orthogonal of both sides of (4.10c), one gets

$$\mathcal{R}(V_{:,I\cup\{k+1\}}) = \mathcal{R}(\mathrm{P}_{k+1}\, V_{:,I}) + \mathbb{R}v_{k+1}.$$

Since $\mathcal{R}(\mathrm{P}_{k+1}\, V_{:,I})$ and $\mathbb{R}v_{k+1}$ are orthogonal to each other, one deduces successively that

$$\begin{aligned}\operatorname{rank}(V_{:,I\cup\{k+1\}}) &= \operatorname{rank}(\mathrm{P}_{k+1}\, V_{:,I}) + 1,\\ |I\cup\{k+1\}| - \operatorname{null}(V_{:,I\cup\{k+1\}}) &= |I| - \operatorname{null}(\mathrm{P}_{k+1}\, V_{:,I}) + 1,\\ \operatorname{null}(V_{:,I\cup\{k+1\}}) &= \operatorname{null}((\mathrm{P}_{k+1}\, V)_{:,I}).\end{aligned}$$

Using this last equality, we see that the identity (4.10b) holds. □

4.2.2 Bounds

When p is large, computing the cardinality $|\partial_B H(x)|$ from (4.7) by evaluating the 2^p ranks $\operatorname{rank}(V_{:,I})$ for $I \subseteq [1:p]$ could be excessively expensive. Therefore, having simple-to-compute lower and upper bounds on $|\partial_B H(x)|$ may be useful in some circumstances, including theoretical ones. Actually, such bounds can be obtained very easily, using one of the formulations of the problem given in section 3. Proposition 4.11 gives elementary lower and upper bounds, while proposition 4.15 reinforces the upper bound, thanks to a lower semicontinuity argument (proposition 4.12). Necessary and sufficient conditions ensuring equality in the left-hand side or right-hand side inequalities in the next proposition are given in section 4.3.

Proposition 4.11 (lower and upper bounds on $|\partial_B H(x)|$) *For V given by (3.1) and $r := \operatorname{rank}(V)$, one has $\max(2p, 2^r) \leqslant 2^r + 2(p-r) \leqslant |\partial_B H(x)| \leqslant 2^p$.*

Proof. The first inequality is clear since $p \geqslant r \geqslant 1$ and $2r \leqslant 2^r$.

Consider the second inequality. One can assume that the first r columns of V are linearly independent, so that $|\mathcal{S}_r| = 2^r$ (notation (4.5) and proposition 4.8(2)). Next, by proposition 4.8(3), $|\mathcal{S}_{r+1}| \geqslant 2^r + 2$. By induction, the given lower bound holds for $|\mathcal{S}_p| = |\mathcal{S}| = |\partial_B H(x)|$.

The upper bound was already mentioned in proposition 2.2. □

Proposition 4.15 below provides a refinement of the upper bound given by proposition 4.11. The next proposition will be useful for this purpose. Recall that a function $\varphi : x \in \mathbb{T} \to \varphi(x) \in \mathbb{R}$, defined on a topological space $\mathbb{T}$, is said to be *lower semicontinuous* if, for any $x \in \mathbb{T}$ and any $\varepsilon > 0$, there is a neighborhood $\mathcal{V}$ of x such that, for all $\tilde{x} \in \mathcal{V}$, one has $\varphi(\tilde{x}) \leqslant \varphi(x) + \varepsilon$. 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 σ is defined by (3.4).

Proposition 4.12 (lower semicontinuity of $|\partial_B H(x)|$) *Suppose that the set $\mathcal{S}$, defined by (3.2), is viewed as a function of $V \in \mathbb{R}^{n\times p}$. Then, $\mathcal{S}(V) \subseteq \mathcal{S}(\tilde{V})$ for $\tilde{V}$ near V in $\mathbb{R}^{n\times p}$. In particular, $V \in \mathbb{R}^{n\times p} \mapsto |\mathcal{S}(V)| \in \mathbb{N}^*$ is lower semicontinuous.*

Proof. By the definition (3.2) of $\mathcal{S}(V)$, for all $s \in \mathcal{S}(V)$, there is a $d_s \in \mathbb{R}^n$ such that $s \bullet (V^\mathsf{T} d_s) > 0$. Clearly, one still has $s \bullet (\tilde{V}^\mathsf{T} d_s) > 0$, for $\tilde{V}$ near V. Since $\mathcal{S}(V)$ is finite, there

is a neighborhood $\mathcal{V}$ of V, such that, for $\tilde{V} \in \mathcal{V}$ and $s \in \mathcal{S}(V)$, there is a $d \in \mathbb{R}^n$ such that $s \bullet (\tilde{V}^\mathsf{T} d) > 0$ or $s \in \mathcal{S}(\tilde{V})$. 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 $|\mathcal{S}(V)| \leqslant |\mathcal{S}(\tilde{V})|$ for $\tilde{V}$ near V. The lower semicontinuity of $V \mapsto |\mathcal{S}(V)|$ follows. □

Proposition 4.2 establishes a necessary and sufficient condition to have completeness of $\partial_B H(x)$. Here follows a less restrictive assumption, called *general position*, which is equivalent to have equality in (4.13) below. In connection with this assumption, it is worth noting that, for a matrix $V \in \mathbb{R}^{n \times p}$ of rank r, one has

$$\forall\, I \subseteq [1:p]: \quad \operatorname{rank}(V_{:,I}) \leqslant \min(|I|, r). \tag{4.11}$$

Definition 4.13 (general position) The vectors $v_1, \ldots, v_p \in \mathbb{R}^n$ are said to be in *general position*, if the matrix $V := \begin{pmatrix} v_1 & \cdots & v_p \end{pmatrix}$ verifies

$$\forall\, I \subseteq [1:p]: \quad \operatorname{rank}(V_{:,I}) = \min(|I|, r), \tag{4.12}$$

where $r := \operatorname{rank}(V)$. □

In the matroid terminology, the vector matroid formed by the columns of V in general position is said to be uniform [59; example 1.2.7]. The general position notion is used by Winder [84] when $r = n$; this is not a restriction since one can work in $\mathcal{R}(V)$ rather than in $\mathbb{R}^n$, observe that the regions in $\mathcal{R}(V)$ are the regions in $\mathbb{R}^n$, projected on $\mathcal{R}(V)$, and that the assumption is satisfied in $\mathcal{R}(V)$. The notion is also specific to arrangements of hyperplanes having a point in common; it has an adapted formulation otherwise [78]. Example of vectors in general position are those in the left-hand side and right-hand side panes in figure 3.2 (the points are the normalized vectors $\bar{v}_i$'s so that the v_i's are actually in $\mathbb{R}^3$); note that in the first case $2 = r < n = 3$. Those in the middle pane are not in general position. This is due to the fact that $r := \operatorname{rank}(V) = 3$ while for the 3 bottom vectors, with indices in I say, one has $\min(|I|, r) - \operatorname{rank}(V_{:,I}) = 3 - 2 \neq 0$.

The proof of proposition 4.15 will use the following lemma, for which we give a detailed proof.

Lemma 4.14 (intentional perturbation) *Let be given a normed space $(\mathbb{E}, \|\cdot\|)$ of dimention $n \in \mathbb{N}^*$ and $q \in \mathbb{N}^*$ vectors $v_1, \ldots, v_q \in \mathbb{E}$ generating a subspace of dimension q_0. Then, for all $\varepsilon > 0$ and all $r \in [q_0 : n]$, one can find vectors $\tilde{v}_1, \ldots, \tilde{v}_q \in \mathbb{E}$, such that*

1) *for $i \in [1:q]$, $\|\tilde{v}_i - v_i\| \leqslant \varepsilon$,*
2) $\operatorname{rank}(\tilde{v}_1 \ \cdots \ \tilde{v}_q) = \min(q, r)$.

Proof. Let $\mathbb{E}_0 := \operatorname{vect}\{v_i : i \in [1:q]\}$, which of dimention q_0. One can suppose that $q_0 < \min(q, r)$, since $q_0 \leqslant \min(q, r)$ and if $q_0 = \min(q, r)$, one can take $\tilde{v}_i = v_i$ and the result is proved. Next, one can find a partition (I, D) of $[1:q]$ such that $\{v_i : i \in I\}$ are linearly independent and $\{v_i : i \in D\} \subseteq \mathbb{E}_0$; hence $|I| = q_0$ and $|D| = q - q_0$. Finally, one can also take a partition (D_0, D_1) of D, with $|D_1| = \min(q, r) - q_0$, which is positive and $\leqslant q - q_0 = |D|$. Hence $|D_0| = |D| - |D_1| = (q - q_0) - (\min(q, r) - q_0) = \max(0, q - r)$.

Since $|D_1| \leqslant r - q_0 \leqslant n - q_0 = \dim \mathbb{E}_0^\perp$, one can find $|D_1|$ linearly independent unit vectors $u_i \in \mathbb{E}_0^\perp$, labeled by $i \in D_1$. For the $\varepsilon > 0$ given in the statement of the lemma, the vectors $\tilde{v}_i$ are defined by

$$\tilde{v}_i := \begin{cases} v_i & \text{for } i \in I \cup D_0, \\ v_i + \varepsilon u_i & \text{for } i \in D_1. \end{cases}$$

These vectors satisfy point 1 in the statement of the lemma, since $\|u_i\| = 1$ for $i \in D_1$. Note also that the vectors $\{\tilde{v}_i : i \in I \cup D_1\}$ are linearly independent. Indeed, if $\sum_{i \in I \cup D_1} \alpha_i \tilde{v}_i = 0$ for some $\alpha \in \mathbb{R}^{I \cup D_1}$, one has

$$\sum_{i \in I \cup D_1} \alpha_i v_i + \varepsilon \sum_{i \in D_1} \alpha_i u_i = 0.$$

Since the two terms in the left-hand side of the previous identity are perpendicular, one has $\sum_{i \in I \cup D_1} \alpha_i v_i = 0$ and $\sum_{i \in D_1} \alpha_i u_i = 0$. The latter condition implies that $\alpha_{D_1} = 0$ and the former now implies that $\alpha_I = 0$. Note also that $\{\tilde{v}_i : i \in D_0\} = \{v_i : i \in D_0\} \subseteq \operatorname{vect}\{v_i : i \in I\}$, so that $\{\tilde{v}_i : i \in [1\,{:}\,q]\} \subseteq \operatorname{vect}\{\tilde{v}_i : i \in I \cup D_1\}$. One concludes that $\operatorname{vect}\{\tilde{v}_i : i \in [1\,{:}\,q]\} = \operatorname{vect}\{\tilde{v}_i : i \in I \cup D_1\}$, which is of dimension $q_0 + (\min(q, r) - q_0) = \min(q, r)$, so that point 2 also holds. □

Note that the vectors $\{\tilde{v}_i : i \in [1\,{:}\,p]\}$ given by the previous lemma are not in general position, while $\{v_i : i \in [1\,{:}\,p]\}$ can be in general position. For example, take $\mathbb{E} = \mathbb{R}^3$, then the vectors $v_1 = e_1$, $v_2 = e_2$, $v_2 = e_1 + e_2$, $v_1 = e_1 - e_2$ are in general position, while $\tilde{v}_1 = e_1$, $\tilde{v}_2 = e_2$, $\tilde{v}_2 = e_1 + e_2 + \varepsilon e_3$, $\tilde{v}_1 = e_1 - e_2$ satisfy the conclusion of the lemma, but are not in general position.

Equality in the upper estimate (4.13) of the next proposition was shown by Winder [84; 1966, corollary] when the columns of V are in general position and $r = n$, thanks to the identity (4.7). Long before him, the Swiss mathematician Ludwig Schläfli [73; p. 211] established the identity under the same assumptions, before 1852 [73; p. 174], without reference to (4.7), which was probably not known at that time. Note that equality does not hold in (4.13) for the middle configuration in figure 3.2 since $|\partial_B H(x)| = 12$, while the right-hand side of (4.13) reads $2[\binom{3}{0} + \binom{3}{1} + \binom{3}{2}] = 14$ (we have seen that the vectors in this pane are not in general position). The bound (4.13) is also useful to check the behavior of the algorithms for test-cases in which the columns of V are in general position. This is likely to be so for randomly generated V, and it was verified by all our random test-cases in section 5.2.8(B.1).

Proposition 4.15 (upper bound on $|\partial_B H(x)|$) *For V given by (3.1) and $r := \operatorname{rank}(V)$, one has*

$$|\partial_B H(x)| \leqslant 2 \sum_{i \in [0\,{:}\,r-1]} \binom{p-1}{i}, \tag{4.13}$$

with equality if and only if (4.12) holds.

Proof. 1) The proof of the implication "(4.12) ⇒ (4.13) with equality" is established in [84; corollary], using the identity (4.7). For the reader's convenience, we give a proof with our notation. One has

$$|\mathcal{S}| \;=\; \sum_{I \subseteq [1\,{:}\,p]} (-1)^{\operatorname{null}(V_{:,I})} \qquad \text{[Winder's formula (4.7)]}$$

$$
\begin{aligned}
&= \sum_{\substack{I\subseteq[1:p]\\|I|\leqslant r}} (-1)^{\mathrm{null}(V_{:,I})} + \sum_{\substack{I\subseteq[1:p]\\|I|>r}} (-1)^{\mathrm{null}(V_{:,I})} \\
&= \sum_{\substack{I\subseteq[1:p]\\|I|\leqslant r}} 1 + \sum_{\substack{I\subseteq[1:p]\\|I|>r}} (-1)^{|I|-r} \qquad [\mathrm{rank}(V_{:,I}) = \min(|I|, r)] \\
&= \sum_{i\in[0:r]} \binom{p}{i} + \sum_{i\in[r+1:p]} (-1)^{i-r} \binom{p}{i} \\
&= \sum_{i\in[0:p]} \binom{p}{i} - 2 \sum_{i\in\{r+1,r+3,\ldots\}} \binom{p}{i} \\
&= 2\left[\sum_{i\in[0:p-1]} \binom{p-1}{i} - \sum_{i\in\{r+1,r+3,\ldots\}} \binom{p}{i}\right], \qquad (4.14a)
\end{aligned}
$$

where the first sum in the brakets has the value 2^{p-1} and the last sum in the brakets is zero if $r = p$. If $r < p$ and $p - r$ is odd, the last sum in (4.14a) has at least the term $\binom{p}{p} = 1$. If $r < p$ and $p - r$ is even, then $r + 1 \leqslant p - 1$. With these particular cases in mind, one can evaluate the last sum in (4.14a) as follows

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

Using (4.14a), we get immediately (4.13) with equality.

2) Let us now show that (4.13) holds. Below, we systematically identify $\partial_B H(x)$ and $\mathcal{S}$, thanks to the equivalence 3.5. We also note $\mathcal{S} \equiv \mathcal{S}(V)$ to stress the dependence of $\mathcal{S}$ on V. Let β be the right-hand side of (4.13). We proceed by contradiction, assuming that there is a matrix $V \in \mathbb{R}^{n\times p}$ of rank r such that

$$
|\mathcal{S}(V)| > \beta. \qquad (4.14b)
$$

It certainly suffices to show that one can find a matrice $\tilde{V} \subseteq \mathbb{R}^{n\times p}$ of rank r arbitrarily close to V that satisfies

$$
|\mathcal{S}(\tilde{V})| = \beta, \qquad (4.14c)
$$

since then one would have the expected contradiction with the lower semicontinuity of $V \mapsto |\mathcal{S}(V)|$ ensured by proposition 4.12:

$$
|\mathcal{S}(\tilde{V})| = \beta < |\mathcal{S}(V)|.
$$

To find $\tilde{V}$ of rank r arbitrarily close to V verifying (4.14c), we proceed as follows. Since (4.14b) holds, the first part of the proof implies that V does not satify (4.12). Our goal is to construct from V a matrix $\tilde{V}$ of rank r arbitrarily close to V with columns in general position. Then, $\tilde{V}$ satisfies (4.14c) by the first part of the proof.

In view of (4.11) and since V does not satisfy (4.12), there is some $I \subseteq [1:p]$ such that $\operatorname{rank}(V_{:,I}) < \min(|I|, r)$. By lemma 4.14, with $\mathbb{E} = \mathcal{R}(V)$, $q := |I|$, the v_i's being the columns of $V_{:,I}$ and $r = \operatorname{rank}(V) \in [:]$, one can get an arbitrarily small perturbation $\tilde{V}_{:,I}$ of $V_{:,I}$, such that $\operatorname{rank}(\tilde{V}_{:,I}) = \min(|I|, r)$ and $\mathcal{R}(\tilde{V}_{:,I}) \subseteq \mathcal{R}(V)$. Next, one forms $\tilde{V} \in \mathbb{R}^{n\times p}$ by setting $\tilde{V}_{:,I^c} = \tilde{V}_{:,I^c}$, so that $\tilde{V}$ is as close to V as desired and verifies $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$. The perturbation $\tilde{V}_{:,I}$ of $V_{:,I}$ can also perturb $V_{:,I'}$ for other index sets $I' \subseteq [1:p]$. However, one has $\operatorname{rank}(\tilde{V}_{:,I'}) \leqslant \min(|I'|, r)$ by (4.11). Now, by the property of the rank, which can only increase in a neighborhood of a given matrix, if the perturbation taken above is sufficiently small, one has $\operatorname{rank}(V_{:,I'}) \leqslant \operatorname{rank}(\tilde{V}_{:,I'}) \leqslant \min(|I'|, r)$ for any $I' \subseteq [1:p]$. Therefore, $\operatorname{rank}(V_{:,I'}) = \min(|I'|, r)$ implies that $\operatorname{rank}(\tilde{V}_{:,I'}) = \min(|I'|, r)$. As a result, the modification of V into $\tilde{V}$ described above increases by at least one the number of intervals $I' \subseteq [1:p]$ such that $\operatorname{rank}(\tilde{V}_{:,I'}) = \min(|I'|, r)$. Since the number of such intervals is finite, proceeding similarly with all the nonempty index sets $I'' \subseteq [1:p]$ such that $\operatorname{rank}(\tilde{V}_{:,I''}) < \min(|I''|, r)$, one finally obtains a matrix $\tilde{V}$, arbitrarily close to V, such that (4.12) holds: $\operatorname{rank}((\tilde{V})_{:,I}) = \min(|I|, r)$ for all $I \subseteq [1:p]$.

3) One still has to show that "(4.13) with equality $\Rightarrow$ (4.12)". We proceed by contradiction, assuming that (4.13) holds with equality for $\partial_B H(x) \equiv \mathcal{S}(V)$, but that (4.12) does not hold. By (4.11), there exists $I \subseteq [1:p]$ such that

$$\operatorname{rank}(V_{:,I}) < \min(|I|, r). \tag{4.14d}$$

Let $\beta = |\mathcal{S}(V)|$ be the right-hand side of (4.13). It certainly suffices to show that, thanks to (4.14d), one can find a matrix $\tilde{V} \in \mathbb{R}^{n\times p}$ such that $\operatorname{rank}(\tilde{V}) \leqslant r$ and $|\mathcal{S}(\tilde{V})| > \beta$, since this would be in contradiction with what has been shown in part 2 of the proof. This matrix $\tilde{V}$ is obtained by perturbing V. By proposition 4.12, if the perturbation is sufficiently small, 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 (4.14d) implies that one can find an index set $J \subseteq I$ such that

$$V_{:,J} \text{ is not injective} \qquad \text{and} \qquad |J| \leqslant r. \tag{4.14e}$$

Indeed, if $|I| \leqslant r$, one can take $J = I$ to satisfy (4.14e), since $\operatorname{rank}(V_{:,I}) < |I|$ by (4.14d), so that $V_{:,I}$ is not injective. If $|I| > r$, then $\operatorname{rank}(V_{:,I}) < r$ by (4.14d), which implies that any $J \subseteq I$ such that $|J| = r$ satisfies (4.14e).

Since $V_{:,J}$ is not injective, one can find $\alpha_J \in \mathbb{R}^J \setminus \{0\}$ such that

$$0 = \sum_{j\in J} \alpha_j v_j = \sum_{j\in J} \tilde{s}_j |\alpha_j| v_j,$$

for some $\tilde{s}_J \in \{\pm 1\}^J$ satisfying $\tilde{s}_j \in \operatorname{sgn}(\alpha_j)$ for all $j \in J$. Then, by Gordan's alternative (3.7),

$$\nexists\, d \in \mathbb{R}^n : \quad \tilde{s}_j v_j d > 0, \quad \text{for all } j \in J.$$

This implies that there is no $s \in \mathcal{S}(V)$ such that $s_J = \tilde{s}_J$. To conclude the proof, it suffices now to show that one can construct an arbitrarily small perturbation $\tilde{V}$ of V, such that $\mathcal{R}(\tilde{V}) \subseteq \mathcal{R}(V)$ and with an $s \in \mathcal{S}(\tilde{V})$ satisfying $s_J = \tilde{s}_J$.

Let $J^c := [1:p] \setminus J$. By (4.14e), $|J| \leqslant r \leqslant n$ so that one can find vectors $\{\tilde{v}_j : j \in [1:p]\}$, such that $\tilde{v}_j = v_j$ for $j \in J^c$, the vectors $\{\tilde{v}_j : j \in J\}$ are linearly independent, $\tilde{v}_j - v_j$ is arbitrarily small and $\{\tilde{v}_j : j \in [1:p]\} \subseteq \mathcal{R}(V)$. Since the vectors $\{\tilde{v}_j : j \in J\}$ are linearly independent, one can find a direction $d_0 \in \mathbb{R}^n$ such that $\tilde{v}_j^\mathsf{T} d_0 = \tilde{s}_j$ for $j \in J$, hence

$$\tilde{s}_j \tilde{v}_j^\mathsf{T} d_0 > 0, \quad \forall\, j \in J. \tag{4.14f}$$

Set $\tilde{s}_j = 1$ for $j \in J^c$. Let d be a discriminating covector given by lemma 2.6 (there denoted ξ) for the vectors $\{0\} \cup \{\tilde{s}_i v_i : i \in [1:p]\}$ sufficiently close to d_0. It results that $\tilde{s}_j \tilde{v}_j^\mathsf{T} d > 0$ for $j \in J$ (by (4.14f)) and that $\tilde{s}_j \tilde{v}_j^\mathsf{T} d \neq 0$ for $j \in J^c$. Finally, we see that the sign vector $s \in \{\pm 1\}^p$ defined by $s_i = \operatorname{sgn}(\tilde{v}_i^\mathsf{T} d)$ for all $i \in [1:p]$ is in $\mathcal{S}(\tilde{V})$ and satisfies $s_J = \tilde{s}_J$, as desired. □

Corollary 4.16 (stability of the sign vector set) *The sign vector set $\mathcal{S} \subseteq \{\pm 1\}^p$ defined by (3.2) is unchanged by small variations of the matrix $V \in \mathbb{R}^{n \times p}$ preserving its rank, provided the columns $v_1, \ldots, v_p \in \mathbb{R}^n$ of V are in general position in the sense of definition 4.13.*

Proof. If $\tilde{V}$ is near V, $\mathcal{S}(V) \subseteq \mathcal{S}(\tilde{V})$ by proposition 4.12. If the columns of V are in general position, proposition 4.15 tells us that $|\mathcal{S}(V)| = \beta$, where β is the right-hand side of Schläfli's bound (4.13) with $r = \operatorname{rank}(V)$. Now, by the fact that $\operatorname{rank}(\tilde{V}) = r$, proposition 4.15 ensures that $|\mathcal{S}(\tilde{V})| \leqslant \beta$. Therefore, one must have $\mathcal{S}(\tilde{V}) = \mathcal{S}(V)$. □

4.3 Particular configurations

We consider in this section some particular matrices $V \in \mathbb{R}^{n \times p}$ given by (3.1), which may be useful to get familiar with the B-differential of H. For these V's, $|\partial_B H(x)|$ can be computed easily. We consider two matrices V with the property that $r := \operatorname{rank}(V)$ takes the value 2 or p; they yield the lower and upper bounds on $|\partial_B H(x)|$ given by proposition 4.11. The lower bound $2p$ applies to the left-hand side pane of figure 3.2. As shown by the intermediate pane in figure 3.2, however, with $2 < r < p$, $|\partial_B H(x)|$ does not only depend on r.

Proposition 4.17 (injective matrix) *The matrix $V \in \mathbb{R}^{n \times p}$ given by (3.1) is injective if and only if $|\partial_B H(x)| = 2^p$.*

Proof. Indeed, by proposition 4.2, the B-differential $\partial_B H(x)$ is complete (meaning that it is equal to $\partial_B^\times H(x)$, given by (2.4)) if and only if V is injective. Clearly, the completeness of $\partial_B H(x)$ is equivalent to $|\partial_B H(x)| = 2^p$. □

More algebraically, this result can be deduced from Winder's formula (4.7) and the bound (4.13). [⇒] If V is injective, then, for any $I \subseteq [1:p]$, one has $\operatorname{null}(V_{:,I}) = 0$. Therefore, there are 2^p terms in the right-hand side of (4.7), 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 (4.13), meaning the V is injective.

The next result applies to the left-hand side example of figure 3.2.

Proposition 4.18 (fan arrangement) *If $p \geqslant 2$ and the vectors v_i's are not two by two colinear, one has $\operatorname{rank}(V) = 2$ if and only if $|\partial_B H(x)| = 2p$.*

Proof. [⇒] A short proof leverages Schläfli's bound (4.13) with equality. Since the v_i's are not two by two colinear, one has for any $I \subseteq [1:p]$:

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

Therefore (4.12) holds. By proposition 4.15, this implies that equality holds in (4.13), that is, with $r := \operatorname{rank}(V) = 2$: $|\partial_B H(x)| = 2\sum_{i\in[0:1]}\binom{p-1}{i} = 2p$.

[$\Leftarrow$] If $|\partial_B H(x)| = 2p$, proposition 4.11 yields $2p \leqslant \max(2p, 2^r) \leqslant 2^r + 2(p-r) \leqslant 2p$, so that equality holds in these inequalities. By the last one, $2^r = 2r$, which only occurs for $r \in \{1, 2\}$. Since $p \geqslant 2$ and the vectors are not colinear, one has $r = 2$. □

Another proof of proposition 4.18[$\Rightarrow$] *using the bipartitioning of section* 3.3.2. Let $s \in \mathcal{S}$, so that $\operatorname{cone}\{s_i v_i : i \in [1:p]\}$, called the original cone in this proof, is pointed by (3.13). Adopt the construction 3.16 (with $s_i \curvearrowright s_i v_i$). Since $r = 2$, the vectors $\bar{v}_i$'s are arranged along a line and one can assume that $\bar{v}_1, \ldots, \bar{v}_p$ 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 3.17, this amounts to identifying the partitions $(I, J) \in \mathfrak{B}([1:p])$ 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 resulting 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 $\{\bar{v}_1, \ldots, \bar{v}_k\}$ and $\{\bar{v}_{p-k+1}, \ldots, \bar{v}_p\}$ 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 yield a pointed cone, is

$$\sum_{k\in[1:p-1]} \left[\binom{p}{k} - 2\right] = (2^p - 2) - 2(p-1) = 2^p - 2p, \tag{4.15}$$

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 $I = \varnothing$ 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 $|\partial_B H(x)| = 2p$. □

Another proof of proposition 4.18[$\Leftarrow$]. The $\operatorname{rank}(V) =: r$ cannot be 1, since $p \geqslant 2$ and the v_i's are not two by two colinear. We proceed by contradiction, assuming the $r > 2$. Then, one can find $k \in [2:p-1]$ such that $\dim \operatorname{vect}\{v_1, \ldots, v_k\} = 2$ and $\dim \operatorname{vect}\{v_1, \ldots, v_{k+1}\} = 3$. For any $k \in [1:p]$, denote by $\mathcal{S}_k$ the set defined by (4.5). By the first part of the proof, $|\mathcal{S}_k| = 2k$. Since $v_{k+1} \notin \{v_1, \ldots, v_k\}$, proposition 4.8(2) tells us that $|\mathcal{S}_{k+1}| = 4k$. Now, for $j \in [k+2:p]$, proposition 4.8(3) tells us that $|\mathcal{S}_j| \geqslant |\mathcal{S}_{j-1}| + 2$. As a result, we get $|\mathcal{S}_p| \geqslant 4k + 2(p-k-1) = 2p + 2(k-1) \geqslant 2p + 2$ (since $k \geqslant 2$), which contradicts the assumption $|\mathcal{S}_p| = 2p$. □

4.4 A glance at the C-differential

The section presents two links between the B-differential and the C-differential of the function H given by (1.3). The first proposition tells us that, whilst $\partial_C H(x)$ can be obtained from $\partial_B H(x)$ by taking its convex hull (it is its definition (1.2)), the latter can be obtained from the former by taking its extreme points.

Proposition 4.19 (a link with the C-differential) $\partial_B H(x) = \operatorname{ext} \partial_C H(x)$.

Proof. Observe first that, since $\mathcal{S}$ given by (3.2) is contained in $\{\pm 1\}^p$, one has $\mathcal{S} = \operatorname{ext}(\operatorname{co} \mathcal{S})$. To get the result, it suffices now to carry this identity into $\mathbb{R}^{p\times n}$ thanks to the affine map $\tau : \mathbb{R}^p \to \mathbb{R}^{p\times n}$ defined at $s \in \mathbb{R}^p$ by

$$\tau(s) = \frac{1}{2}\left[(I - \operatorname{Diag}(s))B_{\mathcal{E}^{\neq}(x),:} + (I + \operatorname{Diag}(s))A_{\mathcal{E}^{\neq}(x),:}\right].$$

The restriction of τ to $\mathcal{S}$ is $\tau|_{\mathcal{S}} = \sigma^{-1}$, defined by (3.4b). Furthermore, τ is injective, since $A_{i,:} \neq B_{i,:}$ for $i \in \mathcal{E}^{\neq}(x)$. Therefore, by applying τ to both sides of the identity $\mathcal{S} = \operatorname{ext}(\operatorname{co} \mathcal{S})$, one gets

$$\begin{aligned} \tau(\mathcal{S}) &= \operatorname{ext}(\tau(\operatorname{co} \mathcal{S})) && \text{[injectivity of } \tau \text{ [42; prop. 2.12(2)]]} \\ &= \operatorname{ext}(\operatorname{co}(\tau(\mathcal{S}))) && \text{[affinity of } \tau \text{ [42; prop. 2.5(1)]].} \end{aligned}$$

The result now follows from the fact that $\tau(\mathcal{S}) = \sigma^{-1}(\mathcal{S}) = \partial_B H(x)$ (proposition 3.4) and $\partial_C H(x) = \operatorname{co} \partial_B H(x)$. □

The second proposition restates theorem 2.2 of Xiang and Chen [85; 2011], which applies to the more general nonlinear function (1.6). The interest of this restatement comes from its proof that is short, thanks to the use of the symmetry of the B-differential (proposition 4.1), and from the fact that proposition 4.20 can be used, straightforwardly, to recover Xiang and Chen's central C-Jacobian of $\tilde{H}$, given by (1.6); see [33]. Recall the notation (2.2) of the index sets.

Proposition 4.20 (the central C-Jacobian) *One has $J \in \partial_C H(x)$ for the Jacobian whose ith row, $i \in [1:m]$, is defined by*

$$J_{i,:} = \begin{cases} A_{i,:} & \text{if } i \in \mathcal{A}(x), \\ \frac{1}{2}[A_{i,:} + B_{i,:}] & \text{if } i \in \mathcal{E}(x), \\ B_{i,:} & \text{if } i \in \mathcal{B}(x). \end{cases} \tag{4.16}$$

Proof. Let $M \in \partial_B H(x)$, which is known to be nonempty. By proposition 2.2, $M_{i,:} = A_{i,:}$ for $i \in \mathcal{A}(x)$, $M_{i,:} = B_{i,:}$ for $i \in \mathcal{B}(x)$ and $M_{i,:} = A_{i,:} = B_{i,:}$ for $i \in \mathcal{E}^{=}(x)$. By the symmetry of $\partial_B H(x)$ (proposition 4.1), M' defined by $M'_{:,i} = M_{:,i}$ if $i \in \mathcal{A}(x) \cup \mathcal{E}^{=}(x) \cup \mathcal{B}(x)$ and by

$$M'_{i,:} = \begin{cases} B_{i,:} & \text{if } i \in \mathcal{E}^{\neq}(x) \text{ and } M_{i,:} = A_{i,:} \\ A_{i,:} & \text{if } i \in \mathcal{E}^{\neq}(x) \text{ and } M_{i,:} = B_{i,:} \end{cases}$$

is also in $\partial_B H(x)$. Therefore, $J = (M + M')/2$ is in $\operatorname{co} \partial_B H(x) = \partial_C H(x)$, by (1.2). This is the formula of J given in the statement of the proposition. □

Instead of taking $J_{1/2} := \frac{1}{2}(M + M')$ in the preceeding proof, one could also have taken $J_t := (1-t)M + tM'$, which is also in $\operatorname{co} \partial_B H(x) = \partial_C H(x)$ for any $t \in [0, 1]$. The inconvenient of this latter choice, when $t \neq 1/2$, is that M is usually not known. In particular, it is not necessarily known whether $M_{i,:}$ may be $A_{i,:}$ or $B_{i,:}$, for a particular $i \in \mathcal{E}^{\neq}(x)$, while J_t depends on this value when $t \neq 1/2$. In contrast, $J_{1/2}$ has an explicit formula that does not require the knowledge of the value of $M_{i,:}$ for $i \in \mathcal{E}^{\neq}(x)$.

5 Computation of the B-differential

This section describes techniques for computing a single Jacobian (section 5.1) or all the Jacobians (section 5.2) of the B-differential $\partial_B H(x)$, in exact arithmetic, when H is the piecewise affine function given by (1.3). The complexity of a proposed algorithm for computing $\partial_B H(x)$ 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 checking the compatibillity of a linear system with inequalities. The algorithms are presented as tools for computing the sign vector set $\mathcal{S} \equiv \mathcal{S}(V)$, defined by (3.2) from a matrix $V \in \mathbb{R}^{n\times p}$, which makes them appropriate, even when $p > n$. When V is defined by (3.1), one has $p \leqslant n$ and the equivalence 3.5 tells us that $\mathcal{S}$ is then in bijection with $\partial_B H(x)$, so that the algorithms actually compute Jacobians of the B-differential $\partial_B H(x)$. The piece of software ISF has been written to test the algorithms [34, 35].

5.1 Computation of a single Jacobian

An interest of the problem equivalence highlighted in proposition 3.4(3) is to provide a method to find rapidly an element of $\partial_B H(x)$, which complements Qi's [64; 1993, final remarks (1)]. It is shown in [33], that this method extends to the computation of an element of the B-differential in the nonlinear case, i.e., when H is the function $\tilde{H}$ given by (1.6). The method is based on the following algorithm, which associates with p nonzero vectors $v_1, \ldots, v_p$, which may be identical or colinear, a direction d such that $v_i^\mathsf{T} d \neq 0$ for all $i \in [1:p]$; it is a variant of the technique used in the proof of [85; lemma 2.1]. When the v_i's are also distinct, the direction d can also be derived from lemma 2.6, by adding the vector $v_0 = 0$.

Algorithm 5.1 (computes $d \in \mathbb{R}^n$ such that $v_i^\mathsf{T} d \neq 0$ for all i)
Let be given p nonzero vectors $v_1, \ldots, v_p$ in $\mathbb{R}^n$ and take $d \in \mathbb{R}^n \setminus \{0\}$.
Repeat:

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

Explanation. In step 3, any sufficiently small $t > 0$ is appropriate (the proof of [85; lemma 2.1] computes bounds explicitely), since $(v_j^\mathsf{T} d)(v_j^\mathsf{T}[d+tv_i])$ is positive for $t = 0$. The new direction d set in step 4 is such that $v_i^\mathsf{T}(d + tv_i) = t\|v_i\|^2 > 0$, so that this direction makes at least one more $v_j^\mathsf{T} d$ nonzero than the previous one. This implies that the algorithm finds an appropriate direction in at most p loops. □

The next procedure uses a direction d computed by algorithm 5.1 to obtain a single element of $\partial_B H(x)$. Recall that the map σ is defined by (3.4a) and is a bijection from $\partial_B H(x)$ onto $\mathcal{S}$, defined by (3.2) (proposition 3.4).

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

Let H be given by (1.3), $x \in \mathbb{R}^n$ and suppose that $p \neq 0$.

1. Compute $V \in \mathbb{R}^{n\times p}$ by (3.1) and denote its columns by $v_1, \ldots, v_p \in \mathbb{R}^n$.
2. By algorithm 5.1, compute $d \in \mathbb{R}^n$ such that $v_i^\mathsf{T} d \neq 0$ for all $i \in [1:p]$.
3. Define $s \in \mathcal{S}$ by $s_i := \operatorname{sgn}(v_i^\mathsf{T} d)$, for $i \in [1:p]$.
4. Then, $\sigma^{-1}(s) \in \partial_B H(x)$.

Explanation. When $p = 0$, $\partial_B H(x) = \partial_B^\times H(x)$ contains a single Jacobian that is given by (2.4), which explains why algorithm 5.2 focuses on the case when $p > 0$. The sign vector s computed in step 3 is such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1:p]$, so that it is indeed in $\mathcal{S}$ and, by proposition 3.4, $\sigma^{-1}(s)$ is a Jacobian in $\partial_B H(x)$. □

5.2 Computation of all the Jacobians

This section presents several basic algorithms, and some more efficient variants, for computing all the B-differential of H. They use the notion of $\mathcal{S}$-tree presented in section 5.2.2(A). The brute force algorithm of section 5.2.1 is out of the game due to its lack of efficiency, but it can be used to get a relatively safe list of all the elements of $\partial_B H(x)$. The second algorithm is grounded on the notion of stem vector (section 3.2.2) and is described in section 5.2.3. The third algorithm is the outcome of a series of improvements brought to an algorithm by Rada and Černý [66; 2018] (section 5.2.2(B)) 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 when the hyperplanes contain zero (see section 3.4). The improvements are detailed in section 5.2.5 and the resulting algorithm is described in section 5.2.6. Finally, numerical experiments are presented in section 5.2.8 to compare the efficiency of the algorithms.

Algorithms for listing the elements of the finite set $\partial_B H(x)$ can be designed by looking at one of the various forms of the problem, those described in section 3 and others [5]; this is what we shall do. Most algorithms we have found in the scientific literature take the point of view of hyperplane arrangements of section 3.4 and can 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). One can quote the contributions by Bieri and Nef [13; 1982], Edelsbrunner, O'Rourke and Seidel [37; 1986], Avis and Fukuda [5; 1996], improved by Sleumer [75; 1998], and, more recently, Rada and Černý [66; 2018], which is described in section 5.2.2(B). See also [32].

5.2.1 Brute-force algorithm

We call *brute-force algorithm* the one that uses the sign system feasibility formulation of section 3.2.1 and examines all the sign vectors $s \in \{\pm 1\}^p$ that make the system $s \bullet (V^\mathsf{T} d) > 0$ feasible for some $d \in \mathbb{R}^n$, where V is defined by (3.1). Recall the map σ defined by (3.4).

Algorithm 5.3 (brute-force) For each $s \in \{\pm 1\}^p$, determine whether $s \bullet (V^\mathsf{T} d) > 0$ holds for some $d \in \mathbb{R}^n$. If so, select s as one element of $\mathcal{S}$ or, equivalently, $\sigma^{-1}(s)$ as one element of $\partial_B H(x)$.

Since the feasibility of $s \bullet (V^\mathsf{T} d) > 0$ in d is equivalent to the one of $s \bullet (V^\mathsf{T} d) \geqslant e$, this problem can be solved by considering the feasible linear optimization problem (LOP) $\min_{(d,t)\in\mathbb{R}^n\times\mathbb{R}}\{t \in$

$[-1,\infty]) : s\bullet(V^\mathsf{T}d)+te \geqslant 0\}$ (one has $s \in \mathcal{S}$ if and only if the problem has a feasible point (d,t) with $t < 0$, hence the minimization should not be undertaken to completion). Algorithm 5.3 requires to solve *approximately* 2^p LOPs with $n+1$ variables and p constraints. We use this expensive algorithm in the numerical experiments to have a reference list of the elements of $\mathcal{S}$ for each test-problem.

5.2.2 Incremental-recursive algorithms

The algorithms described in this section are incremental in the sense that the considered sign vectors have their length increased by one at each step. Furthermore, the algorithms explore the $\mathcal{S}$-tree described in subsection A below by recursive procedures, whose names are recognizable by their suffix "-REC". All the procedures end by returning to their calling program.

A. The $\mathcal{S}$-tree

A common feature of the algorithms considered in this paper is the construction of the $\mathcal{S}$-tree described below, incrementally and recursively. This idea was probably introduced by Rada and Černý [66; 2018].

The level k of the $\mathcal{S}$-tree is formed of a set of sign vectors denoted by

$$\mathcal{S}_k^1 := \{s \in \mathcal{S}_k : s_1 = +1\}, \tag{5.1}$$

where $\mathcal{S}_k$ is the subset of $\{\pm 1\}^k$ defined by (4.5). In particular, the level 1 or root of the $\mathcal{S}$-tree contains the unique sign vector $+1 \in \{\pm 1\}^1$. The $\mathcal{S}$-tree has p levels, where p is the number of vectors v_i, or columns of the given matrix $V \in \mathbb{R}^{n\times p}$. Note that there is no reason to compute $\{s \in \mathcal{S} : s_1 = -1\}$ since this part of $\mathcal{S}$ is equal to $-\{s \in \mathcal{S} : s_1 = 1\}$ by the symmetry property of $\mathcal{S}$ (proposition 4.1). In order to avoid the memorization of the elements of $\mathcal{S}_k^1$, the $\mathcal{S}$-tree is constructed by a *depth-first search*, which can be schematized as follows.

Algorithm 5.4 (STREE (V)) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$, having nonzero columns.

1. Execute the recursive procedure STREE-REC$(V, +1)$.

Algorithm 5.5 (STREE-REC (V, s)) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$, having nonzero columns, and a sign vector $s \in \mathcal{S}_k^1$ for some $k \in [1:p]$.

1. If $k = p$, print s and return.
2. If $(s, +1) \in \mathcal{S}_{k+1}^1$, execute STREE-REC$(V, (s, +1))$.
3. If $(s, -1) \in \mathcal{S}_{k+1}^1$, execute STREE-REC$(V, (s, -1))$.

The method used to determine whether $(s, \pm 1)$ is in $\mathcal{S}_{k+1}^1$ depends on the specific algorithm and may or may not use a direction d intervening in (4.5). Note that, as emphasized in proposition 4.8(3), at least one of the sign vectors $(s, +1)$ and $(s, -1)$ belongs to $\mathcal{S}_{k+1}^1$ (maybe both). It is justified not to explore the $\mathcal{S}$-tree below an $(s, \pm 1)$ that is not in $\mathcal{S}_{k+1}^1$, since then $(s, \pm 1, s') \notin \mathcal{S}$ for any $s' \in \{\pm 1\}^{p-k-1}$. By construction, the algorithm STREE prints all the elements of $\mathcal{S}_p^1 \equiv \mathcal{S}^1 := \{s \in \mathcal{S} : s_1 = +1\}$ in step 1 of the STREE-REC procedure.

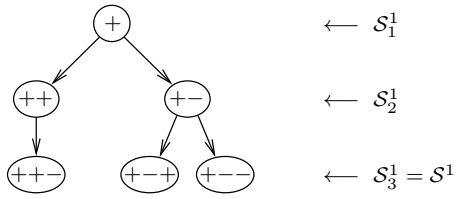

Figure 5.1: Half of the $\mathcal{S}$-tree for example 3.2 (the other half is obtained by swapping the $+$'s and the $-$'s). Top-down arrows indicate descendence; the sign sets $\mathcal{S}_k^1$ are defined by (5.1).

B. Rada and Černý's algorithm

The algorithm proposed by Rada and Černý [66; 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 of hyperplanes containing all zero (see section 3.4), which is the case when V result from (3.1) in the computation of the B-differential $\partial_B H(x)$. We also use the linear algebra language of section 3.2.1, viewing the problem as the one of determining the set $\mathcal{S}$ defined by (3.1)); in contrast, the language used in [66] is more geometric. The algorithm builds the $\mathcal{S}$-tree of the previous section A and, for each $s \in \mathcal{S}_k^1$, it solves a single problem (LOP) to determine whether $(s, +1)$ or $(s, -1)$ is in $\mathcal{S}_{k+1}^1$.

The RC algorithm succeeds in solving only one LOP to determine whether $(s, +1)$ and $(s, -1)$ are in $\mathcal{S}_{k+1}^1$, at the node $s \in \mathcal{S}_k^1$, thanks to the memorization of a direction d such that $s \bullet (V_k^\mathsf{T} d) > 0$ (we note $V_k := V_{:,[1:k]}$). Indeed, one has

$$\begin{array}{lcl} v_{k+1}^\mathsf{T} d < 0 & \Longrightarrow & (s, -1) \in \mathcal{S}_{k+1}^1, \\ v_{k+1}^\mathsf{T} d > 0 & \Longrightarrow & (s, +1) \in \mathcal{S}_{k+1}^1, \end{array}$$

and one of these two cases takes place if we exclude the case where $v_{k+1}^\mathsf{T} d = 0$. In [66; Algorithm 1], the case where $v_{k+1}^\mathsf{T} d = 0$ is not dealt with completely since $(s, +1)$ is declared to belong to $\mathcal{S}_{k+1}^1$ in that case, while it is clear that $(s, -1)$ is also in $\mathcal{S}_{k+1}^1$. Indeed, 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 $v_{k+1}^\mathsf{T} d \simeq 0$, so that both $(s, \pm 1)$ are accepted in $\mathcal{S}_{k+1}^1$ in that case. This choice may be at the origin of the differences that one observes in table 5.1 below between the statistics of the original RC algorithm in [66] and those of our implementation.

Next, when $(s, s_{k+1}) \in \{\pm 1\}^{k+1}$ is observed to belong to $\mathcal{S}_{k+1}^1$, the question of whether $(s, -s_{k+1})$ also belongs to $\mathcal{S}_{k+1}^1$ arises. In the RC algorithm, the answer to this question is obtained by solving a LOP similar to

$$\left\{ \begin{array}{l} \min_{(d,t) \in \mathbb{R}^n \times \mathbb{R}} \ t \\ s_i v_i^\mathsf{T} d \geqslant 1, \quad \forall\, i \in [1:k] \\ -s_{k+1} v_{k+1}^\mathsf{T} d \geqslant -t \\ t \geqslant -1. \end{array} \right. \tag{5.2}$$

When $s \in \mathcal{S}_k^1$, this problem is feasible (take d satisfying $s_i v_i^\mathsf{T} d \geqslant 1$, for all $i \in [1:k]$, and t sufficiently large) and bounded (its optimal value is $\geqslant -1$), so that it has a solution [14, 15, 19, 43]. Solving these LOPs is a time consuming part of the algorithms and in the numerical experiments of section 5.2.8, in particular in table 5.2, following [66], we measure the efficiency of the algorithms by the number of LOPs they solve.

One can now formally describe our version of the RC algorithm (the change is in step 2 of the RC-REC algorithm, which is not considered in the original RC algorithm).

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

1. Execute the recursive procedure RC-REC$(V, v_1, +1)$.

Algorithm 5.7 (RC-REC (V, d, s)) 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 $k \in [1:p]$, such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1:k]$.

1. If $k = p$, print s and return.
2. If $v_{k+1}^\mathsf{T} d \simeq 0$, then
   2.1. Execute RC-REC$(V, d_+, (s, +1))$, where $d_+ := d + t_+ v_{k+1}$ with $t_+ > 0$ chosen in the nonempty open interval
   $$\left(0, \min_{\substack{i\in[1:k]\\ s_i v_i^\mathsf{T} v_{k+1}<0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}}\right).$$
   2.2. Execute RC-REC$(V, d_-, (s, -1))$, where $d_- := d + t_- v_{k+1}$ with $t_- < 0$ chosen in the nonempty open interval
   $$\left(\max_{\substack{i\in[1:k]\\ s_i v_i^\mathsf{T} v_{k+1}>0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}},\ 0\right).$$
3. Else $s_{k+1} := \operatorname{sgn}(v_{k+1}^\mathsf{T} d)$.
   3.1. Execute RC-REC$(V, d, (s, s_{k+1}))$.
   3.2. Solve the LOP (5.2) and denote by (d, t) a solution.
   If $t = -1$, execute RC-REC$(V, d, (s, -s_{k+1}))$.

In steps 2.1 and 2.2, the minimum and maximum are supposed to be infinite if their feasible set is empty. One can check that the directions $d_\pm$ computed in steps 2.1 and 2.2 are such that $s_i v_i^\mathsf{T} d_\pm > 0$ for $i \in [1:k+1]$ and $s_{k+1} = \pm 1$, provided $|v_{k+1}^\mathsf{T} d|$ is sufficiently small, which justifies the recursive call to RC-REC with the given arguments. The test $v_{k+1}^\mathsf{T} d \simeq 0$ done at the beginning of step 2 is supposed to take into account floating point arithmetic; admittedly it is not very rigorous, but the algorithm is designed to be as close as possible to the original RC algorithm in [66]; a more careful treatment of this situation is presented in section 5.2.5(B). The most time-consuming part of the RC algorithm comes from the possible numerous LOPs to solve in step 3.2 of RC-REC.

5.2.3 An algorithm using stem vectors

When $s \in \mathcal{S}_k$, it is conceptually easy to check whether $(s, \pm 1)$ is in $\mathcal{S}_{k+1}$, provided a list of all the stem vectors associated with V is known. Indeed, by proposition 3.10, if no subvector

of $(s,+1)$ (resp. $(s,-1)$) is a stem vector, then $(s,+1)$ (resp. $(s,-1)$) belongs to $\mathcal{S}_{k+1}$. Note also that, because any $s \in \mathcal{S}_k$ has at least one descendant in the $\mathcal{S}$-tree (proposition 4.8(3)), if it is observed that $(s,+1) \notin \mathcal{S}_{k+1}$, then, necessarily, $(s,-1) \in \mathcal{S}_{k+1}$. This observation prevents the algorithm from checking whether $(s,-1)$ contains a stem vector, which is a time consuming operation when the list of stem vectors is large. For future reference, we formalize this algorithm below.

Algorithm 5.8 (STEM (V)) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$, having nonzero columns.

1. Compute all the stem vectors associated with V.
2. Execute the recursive procedure STEM-REC$(V,+1)$.

Algorithm 5.9 (STEM-REC (V,s)) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$, having nonzero columns and a sign vector $s \in \{\pm 1\}^k$ for some $k \in [1:p]$.

1. If $k = p$, print s and return.
2. If no subvector of $(s,+1)$ is a stem vector, execute STEM-REC$(V,(s,+1))$.
3. If $(s,+1) \notin \mathcal{S}_{k+1}$ or no subvector of $(s,-1)$ is a stem vector, execute STEM-REC$(V,(s,-1))$.

This algorithm is improved below, as the option AD_4 of the ISF algorithm (see paragraphs A and D of section 5.2.5).

Note that, this algorithm need not generate directions d satisfying $s \bullet (V_k^\mathsf{T} d) > 0$, like the RC algorithm and need not solve any linear optimization problem. Nevertheless, regarding the computation time, the algorithm has two bottlenecks that we now describe.

The first bottleneck comes from the fact that the algorithm must compute all the stem vectors (or the set $\mathcal{C}$ of matroid circuits in (3.9)) associated with V. This is usually an expensive operation [52, 57, 68]. For example, if V is randomly generated and of rank r, like in the test-cases `data_rand_*` in the experiments of section 5.2.8, any selection of r columns of V is likely to form an independent set of vectors, so that $\mathcal{C}$ is likely to be the sets of column indices of size $r+1$. In this case, the number of circuits is likely to be the combination $\binom{p}{r+1}$ (and it is actually that number, see section 5.2.8(B.1)), which can be exponential in p (this number is bounded below by $2^{p/2}/(p+1)$ if p is even and $r+1 = p/2$ [24; (11.52)]). In the implemented ISF code, numerically tested in section 5.2.8, only the sets of columns whose cardinality is in $[3:r+1]$ are examined (since any group of two columns of V is supposed to be linearly independent and a group of $r+2$ columns or more is of nullity $\geqslant 2$, hence such group cannot form a matroid circuit; see (3.9)).

The second bottleneck is linked to the detection of a stem vector is the current sign vectors $(s,\pm 1)$. This operation requires to examine the long list of stem vectors, which is a time consuming operation.

We shall see in the numerical experiments of section 5.2.8 that algorithm 5.8 is generally the fastest, provided the number of stem vectors is not too large.

5.2.4 Linear optimization problem and stem vector

The property described in this section will be useful for the improvement D_2 of the ISF algorithm, described in section 5.2.5(D). It shows that a stem vector can be obtained easily

from the dual solution of the linear optimization (LOP) (5.2), when $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$. Consider indeed the LOP (5.2) and denote by (d, t) one of its solutions (these have been shown to exist). Then, either $t \geqslant 0$ (equivalently, $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$) or $t = -1$ (equivalently, $(s, -s_{k+1}) \in \mathcal{S}_{k+1}$).

Let σ_i, $i \in [1:k+1]$, be the multipliers associated with the first $k+1$ constraints of (5.2) and τ be the multiplier associated with its last constraint. Then, the Lagrangian dual of (5.2) reads [12, 14, 15, 42]

$$\left\{\begin{array}{l} \max_{(\sigma,\tau)\in\mathbb{R}^{k+1}\times\mathbb{R}} \ \sum_{i\in[1:k]} \sigma_i - \tau \\ \sigma \geqslant 0 \\ \tau \geqslant 0 \\ \sigma_{k+1} + \tau = 1 \\ \sigma_{k+1} s_{k+1} v_{k+1} = \sum_{i\in[1:k]} \sigma_i s_i v_i. \end{array}\right. \equiv \left\{\begin{array}{l} \max_{\sigma\in\mathbb{R}^{k+1}} \ \sum_{i\in[1:k+1]} \sigma_i - 1 \\ \sigma \geqslant 0 \\ \sigma_{k+1} \leqslant 1 \\ \sigma_{k+1} s_{k+1} v_{k+1} = \sum_{i\in[1:k]} \sigma_i s_i v_i, \end{array}\right. \tag{5.3}$$

where the second form of the dual is obtained by eliminating τ from the first form. By strong duality in linear optimization, the dual problems in (5.3) are feasible, have a solution and have the same optimal value as the primal problem. Let $(\sigma, \tau) \in \mathbb{R}^{k+1} \times \mathbb{R}$ be a dual solution. Then, $(s, -s_{k+1}) \in \mathcal{S}_{k+1}$ if and only if $t = -1$ if and only if $\sum_{i\in[1:k]} \sigma_i = 0$ and $\sigma_{k+1} = 0$. We have shown that

$$(s, -s_{k+1}) \in \mathcal{S}_{k+1} \quad \Longleftrightarrow \quad \sigma = 0.$$

Therefore, $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$ if and only if $\sigma \neq 0$ if and only if $\sigma_{k+1} = 1$ (if $\sigma_{k+1} = 0$, one can make the dual objective value as large as desired by multiplying σ by a factor going to $+\infty$; if $\sigma_{k+1} \in (0, 1)$, the dual objective would by increased by replacing σ by σ/σ_{k+1}; in both cases the optimality of σ would be contradicted) if and only if $\tau = 0$. We have shown that

$$(s, -s_{k+1}) \notin \mathcal{S}_{k+1} \quad \Longleftrightarrow \quad s_{k+1} v_{k+1} \in \operatorname{cone}\{s_i v_i : i \in [1:k]\}.$$

The next proposition shows how a matroid circuit can be detected from the dual solution σ when $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$.

Proposition 5.10 (matroid circuit detection) *Suppose that $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$ and that (σ, τ) is a solution to the dual problem in the left-hand side of (5.3) located at an extreme point of its feasible set. Then, $\{i \in [1:k+1] : \sigma_i > 0\}$ is a matroid circuit of V.*

Proof. We have seen that $\sigma_{k+1} = 1$ and $\tau = 0$ when $(s, -s_{k+1}) \notin \mathcal{S}_{k+1}$. The fact that $(\sigma, 0)$ is an extreme point of the feasible set of the problem in the left-hand side of (5.3) implies that the vectors [19, 42]

$$\left\{ \begin{pmatrix} 0 \\ s_i v_i \end{pmatrix}_{i\in[1:k],\ \sigma_i>0}, \begin{pmatrix} 1 \\ -s_{k+1} v_{k+1} \end{pmatrix} \right\} \text{ are linearly independent.}$$

In particular, the vectors

$$\{s_i v_i : i \in [1:k],\ \sigma_i > 0\} \text{ are linearly independent.}$$

Since $s_{k+1} v_{k+1} = \sum_{i\in[1:k]} \sigma_i s_i v_i$, it follows that

$$\{s_i v_i : i \in [1:k+1],\ \sigma_i > 0\} \text{ has nullity one.}$$

The conclusion of the proposition follows from proposition 3.11. □

Recall that the dual-simplex algorithm finds a dual solution at an extreme point of the dual feasible set. For this reason, we use this approach in the ISF algorithm with option D_2 (see section 5.2.5(D)).

5.2.5 Improvements of the RC and STEM algorithms

This section presents several modifications of the RC algorithm and one modification of the STEM algorithm that significantly improve their performance. The modifications are indicated by the letters A, B, C and D, with reference to the sections where they are introduced. Additional numeric indices specify variants of the D option. The version AD_4 (modifications A and D_4) can be considered as an improvement of the new algorithm 5.8.

A. Taking the rank of V into account

Instead of starting with the vector $s = +1$, one can take into account the rank $r := \operatorname{rank}(V)$ to determine 2^r initial vectors s, hence avoiding to solve linear optimization problems (LOPs) to determine these initial s's. This is especially useful when $p - r$ is small. In particular, when $p = r$, $\mathcal{S}$ is straightforwardly determined.

The algorithm selects $r := \operatorname{rank}(V)$ 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 $P \in \{0,1\}^{p\times p}$ is a permutation matrix, $Q \in \mathbb{R}^{n\times n}$ is orthogonal (i.e., $Q^\mathsf{T} Q = I_n$) 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 $v_1, \ldots, v_r$ are linearly independent (in practice, the vectors are symbolically reordered by using the permutation matrix P). By proposition 4.2 and with the notation (4.5):

$$\mathcal{S}_r = \{\pm 1\}^r. \tag{5.4}$$

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

$$d_s = Q_r R_r^{-\mathsf{T}} s \tag{5.5}$$

is such that $s \bullet (V_{:,[1\,:\,r]}^\mathsf{T} d_s) = e > 0$, as desired.

For each $s \in \mathcal{S}_r$ and the associated d_s given by (5.5), the modified algorithm 5.6 runs the recursive function RC-REC(V, d_s, s) (see algorithm 5.12 below).

B. Special handling of the case where $v_{k+1}^\mathsf{T} d \simeq 0$

The equivalence (4.8a) shows that the existence of a direction d such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1\,:\,k]$ and $v_{k+1}^\mathsf{T} d = 0$ is a necessary and sufficient condition for $s \in \mathcal{S}_k$ to have two descendants. This property underlies the following modification.

Directions $d_\pm := d + t_\pm v_{k+1}$ ensuring that $(s, \pm 1) \bullet (V_{k+1}^\mathsf{T} d_\pm) > 0$ can be computed not only when $v_{k+1}^\mathsf{T} d \simeq 0$ like in step 2 of the RC-REC algorithm 5.7, but also when $v_{k+1}^\mathsf{T} d$ is in the interval specified by (5.6) below. Note that the left-hand side in (5.6) is negative and the right-hand side is positive (this can be seen by multiplying numerators and denominators by s_i and by using $s_i v_i^\mathsf{T} d > 0$ for all $i \in [1\,:\,k]$), so that these inequalities are verified when $v_{k+1}^\mathsf{T} d = 0$. With the additional flexibility that (5.6) offers, the ISF algorithm can sometimes avoid solving a significant number of LOPs of the form (5.2).

Proposition 5.11 (two descendants without optimization) *Suppose that* $s \in \{\pm 1\}^k$ *verifies* $s \bullet (V_k^\mathsf{T} d) > 0$, *that* $v_{k+1} \neq 0$ *and that*

$$\max_{\substack{i \in [1:k] \\ s_i v_i^\mathsf{T} v_{k+1} > 0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}} < \frac{-v_{k+1}^\mathsf{T} d}{\|v_{k+1}\|^2} < \min_{\substack{i \in [1:k] \\ s_i v_i^\mathsf{T} v_{k+1} < 0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}}. \tag{5.6}$$

1) *The direction* $d_+ := d + t_+ v_{k+1}$ *verifies* $s \bullet (V_k^\mathsf{T} d_+) > 0$ *and* $v_{k+1}^\mathsf{T} d_+ > 0$ *if and only if* t_+ *is in the nonempty open interval*

$$\left(\frac{-v_{k+1}^\mathsf{T} d}{\|v_{k+1}\|^2}, \quad \min_{\substack{i \in [1:k] \\ s_i v_i^\mathsf{T} v_{k+1} < 0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}} \right). \tag{5.7a}$$

2) *The direction* $d_- := d + t_- v_{k+1}$ *verifies* $s \bullet (V_k^\mathsf{T} d_-) > 0$ *and* $-v_{k+1}^\mathsf{T} d_- > 0$ *if and only if* t_- *is in the nonempty open interval*

$$\left(\max_{\substack{i \in [1:k] \\ s_i v_i^\mathsf{T} v_{k+1} > 0}} \frac{-v_i^\mathsf{T} d}{v_i^\mathsf{T} v_{k+1}}, \quad \frac{-v_{k+1}^\mathsf{T} d}{\|v_{k+1}\|^2} \right). \tag{5.7b}$$

Proof. Clearly, (5.6) implies that the open intervals (5.7a) and (5.7b) are nonempty.

The direction $d_+ := d + t_+ v_{k+1}$ verifies $s \bullet (V_k^\mathsf{T} d_+) > 0$ and $v_{k+1}^\mathsf{T} d_+ > 0$ if and only if

$$\Big(s_i v_i^\mathsf{T} (d + t_+ v_{k+1}) > 0, \quad \forall\, i \in [1:k] \Big) \qquad \text{and} \qquad v_{k+1}^\mathsf{T} (d + t_+ v_{k+1}) > 0. \tag{5.8a}$$

With the first inequality in (5.6), this is equivalent to taking t_+ in the interval (5.7a).

The direction $d_- := d + t_- v_{k+1}$ verifies $s \bullet (V_k^\mathsf{T} d_-) > 0$ and $-v_{k+1}^\mathsf{T} d_- > 0$ if and only if

$$\Big(s_i v_i^\mathsf{T} (d + t_- v_{k+1}) > 0, \quad \forall\, i \in [1:k] \Big) \qquad \text{and} \qquad -\, v_{k+1}^\mathsf{T} (d + t_- v_{k+1}) > 0. \tag{5.8b}$$

With the second inequality in (5.6), this is equivalent to taking t_- in the interval (5.7b). □

C. Changing the order of the vectors v_i's

Each node s of the $\mathcal{S}$-tree described in section 5.2.2(A) has one or two descendants: $(s, +1)$ and/or $(s, -1)$. Since there is at most one LOP solved per node of the $\mathcal{S}$-tree, decreasing the number of nodes should decrease the number of LOPs to solve, which significantly count in the computing time. To reach that goal, one can try to get as much as possible at the top of the tree the nodes having a single descendant. As shown below, this can be achieved by changing the order in which the vectors v_i's, the columns of V, are considered in the *depth-first search* of the tree; previously, the order was imposed by the modification A, taking into account the rank of V. As we shall see, a new order is not fixed once and for all, but is determined during

the construction of the $\mathcal{S}$-tree, is reconsidered at each node and depends on the path going from the root of the $\mathcal{S}$-tree to its leaves.

To implement this strategy, one associates with each node $s \in \mathcal{S}_k^1$ of the $\mathcal{S}$-tree, $k \in [1:p-1]$, the list of vectors considered so far at that node, denoted by $T_s := \{i_1, \ldots, i_k\} \subseteq [1:p]$. Hence, we have to choose the next vector $v_{i_{k+1}}$ be selecting an index i_{k+1} in $T_s^c := [1:p] \setminus T_s$. Now, a natural idea is to restrict the set of possible indices to T_s^b, the set of indices j of T_s^c for which one of the intervals (5.7a) or (5.7b), with $v_{k+1} \equiv v_j$, is empty (implying that the technique used in the modification B will not give two descendants), if there is such an index, or T_s^c otherwise. To determine the index in T_s^b, we take

$$i_{k+1} = \arg\max_{i \in T_s^b} \frac{|v_i^\mathsf{T} d|}{\|v_i\|}, \tag{5.9}$$

which favors the vectors v_i for which $|v_i^\mathsf{T} d|/\|v_i\|$ is away from zero.

As table 5.2 indicates (section 5.2.8(C.3)), this modification has a significant impact on the decrease of the number of LOPs to solve.

D. Using stem vectors

We present in this section various modifications that use the concept of *stem vector*, introduced in the second part of section 3.2.2. These stem vectors are used to detect infeasible sign vectors, i.e., elements of $\mathcal{S}^c$, thanks to proposition 3.10. If $s \in \mathcal{S}_k^1$ and $(s, s_{k+1}) \in \mathcal{S}^c$ for $s_{k+1} \in \{\pm 1\}$, s has no descendant in $\mathcal{S}$ along (s, s_{k+1}), so that this part of the $\mathcal{S}$-tree does not need to be explored. From this point of view, computing all the stem vectors looks attractive, but, to our knowledge, this is a time consuming process, so that this option is not necessarily the most efficient one. The modifications presented below use more and more stem vectors, whose computation requires more and more time.

- D$_1$) Natural candidates as stem vectors are those obtained from the matroid circuits I made of $r+1$ columns of V ($r = \text{rank}(V)$) formed of the r linear independent columns selected by the QR factorization of section 5.2.5(A) and one of the remaining $p-r$ columns of V. By proposition 3.11, such I contains exactly one circuit. Therefore, one detects in this way $p-r$ circuits and $2(p-r)$ stem vectors. This is not much compared to the total number of stem vectors, which may depend exponentially on p, so that the number of infeasible sign vectors detected by these stem vectors is usually relatively small (see table 5.2).

- D$_2$) With this option, when a LOP (5.2) is solved at a certain node $s \in \mathcal{S}_k^1$ to see whether (s, s_{k+1}) belongs to $\mathcal{S}_{k+1}^1$, for $s_{k+1} \in \{\pm 1\}$, the dual solution is used to determine a matroid circuit, as shown by proposition 5.10. For this purpose, the ISF code solves the LOP with the dual-simplex algorithm, so that the computed dual solution is at a vertex of the dual feasible set.

- D$_3$) With this option, all the stem vectors are computed, before running the recursive process that builds the $\mathcal{S}$-tree. At each note $s \in \mathcal{S}_k^1$, the algorithm still computes a direction $d \in \mathbb{R}^n$ such that $s_i v_i^\mathsf{T} d > 0$ for all $i \in T_s$ (the set of vector indices considered so far at s). The advantage of this direction is to allow the algorithm to use the beneficial modifications B and C and to easily determine one or two signs $s_{k+1} \in \{\pm 1\}$ such that $(s, s_{k+1}) \in \mathcal{S}_{k+1}^1$. If a single sign $s_{k+1} \in \{\pm 1\}$ is selected, the stem vectors can decide whether $(s, -s_{k+1}) \in \mathcal{S}_{k+1}^1$. If this is the case, this option D$_3$ has the inconvenient of still requiring to solve a LOP to get a direction associated with $(s, -s_{k+1})$. These LOPs

(5.2) have an optimal value -1 and should not be solved exactly. Indeed, as soon as a feasible direction d for (5.2) gives a negative value to the objective of the problem, one could stop solving it, since this d verifies $s_i v_i^\mathsf{T} d > 0$ for all $i \in T_{(s,-s_{k+1})}$. We have not implemented that inexact solve of the LOPs, by lack of flexibility of the solver `Linprog` in Matlab.

D$_4$) Like with the option D$_3$, all the stem vectors are computed, before running the recursive process that builds the $\mathcal{S}$-tree. But now, unlike with option D$_3$, the algorithm computes no direction $d \in \mathbb{R}^n$. When option A is also activated, the resulting approach can be viewed as an improvement of the algorithm 5.8 (STEM) presented in section 5.2.3.

Note that, knowing all the stem vectors, one could compute the complementary set $\mathcal{S}^c$ rather easily by completing with ± 1 the unspecified components of the stem vectors. Next, $\mathcal{S}$ could be obtained from $\mathcal{S}^c$ by taking its complementary set in $\{\pm 1\}^p$, but a straightforward implementation of this last operation looks rather expensive, so that we have not experimented it numerically.

5.2.6 ISF algorithm

We have named ISF (for Incremental Signed Feasibility) the algorithm that improves the RC algorithm 5.6 or the STEM algorithm 5.8 with the enhancements described in section 5.2.5. For the purpose of precision and reference, we formally state it in this section. It would be cumbersome and confusing, hence inappropriate, to mention all the options in its description, in particular because all of them have been specified separately in the previous section. As an example of algorithm, we provide a description with the options ABCD$_2$. It starts with a hat procedure ISF, similar to that of the RC algorithm but with the additional easy determination of $\mathcal{S}_r$ (modification A) and the computation of some stem vectors (modification D$_1$). Then, the hat procedure calls the recursive procedure ISF-REC.

Algorithm 5.12 (ISF (V), with options ABCD$_2$) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$, having nonzero columns.

1. Compute the QR factorization of V. Let $r = \text{rank}(V)$ and $T_r := \{i_1, \dots, i_r\}$ be the indices of r selected linear independent columns of V.
2. Compute the $p - r$ matroid circuits containing T_r (see option D$_1$).
3. For each $s \in \mathcal{S}_r$, given by (5.4), and its associated d_s, given by (5.5), call the recursive procedure ISF-REC(V, T_r, d_s, s).

Algorithm 5.13 (ISF-REC (V, T, d, s), with options BCD$_2$) Let be given $V \in \mathbb{R}^{n\times p}$, with n and $p \in \mathbb{N}^*$ of rank r, having nonzero columns v_i, T a selection of k columns of V (with $k \in [r\,{:}\,p]$), a direction $d \in \mathbb{R}^n$ and a sign vector $s \in \{\pm 1\}^k$ for some $k \in [r\,{:}\,p]$. It is assumed that $s_i v_i^\mathsf{T} d > 0$ for all $i \in T$.

1. If $k = p$, print s and return.
2. Determine the index $i_{k+1} \in [1\,{:}\,p] \setminus T$ of the next vector to consider by option C and set $T_+ := T \cup \{i_{k+1}\}$.
3. If (5.6) holds (with $[1\,{:}\,k]$ changed into T and $k+1$ into i_{k+1}), then

3.1. Execute ISF-REC$(V, T_+, d_+, (s, +1))$, where $d_+ := d + t_+ v_{i_{k+1}}$ and t_+ is chosen in the nonempty open interval (5.7a).

3.2. Execute ISF-REC$(V, T_+, d_-, (s, -1))$, where $d_- := d + t_- v_{i_{k+1}}$ and t_- chosen in the nonempty open interval (5.7b).

4. Else $s_{k+1} := \operatorname{sgn}(v_{i_{k+1}}^\mathsf{T} d)$.

4.1. Execute ISF-REC$(V, T_+, d, (s, s_{k+1}))$.

4.2. If $(s, -s_{k+1})$ contains a stem vector, return.

4.3. Solve the LOP (5.2) (with $[1:k]$ changed into T and $k+1$ into i_{k+1}) by the dual-simplex algorithm and denote by (d, t) a solution.

4.3.1. If $t = -1$, execute ISF-REC$(V, T_+, d, (s, -s_{k+1}))$.

4.3.2. Else, use the dual solution to store two more stem vectors by option D_2.

5.2.7 Complexity

This section briefly analyzes the complexity of the ISF(A) algorithm, i.e., with only option A of section 5.2.5.A, knowing that, in practice, the other options improve the performance of the algorithm. This one is similar to algorithm 5.12 from which the options BCD_2 have been discarded. As shown by the next proposition, algorithm ISF(A) 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 (recall that this one may be exponential in p). The computation effort is measured in terms of the number of LOPs that must be solved in the worst case. The next result is very similar to [66; theorem 3.2], although its proof is slightly different. The given upper bound on the number of LOPs to solve takes also into account steps 1 and 3 of algorithm 5.12, 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 when $p \gg r$.

Proposition 5.14 (complexity of algorithm ISF(A)) *Suppose that $V \in \mathbb{R}^{n\times p}$ has no colinear columns and denote by r its rank. Then, the number of linear optimization problems solved by algorithm* ISF(A) *does not exceed*

$$\min\big(2^p - 2^r, (p-r)|\mathcal{S}|\big). \tag{5.10}$$

Proof. In its steps 1 and 3, algorithm 5.12 specifies the 2^r sign vectors associated with r linearly independent columns of V, which may be assumed to be $v_1, \ldots, v_r$. This step does not require to solve any LOP. One has

$$|\mathcal{S}_r| = 2^r. \tag{5.11a}$$

Next, algorithm 5.13 considers the remaining $p - r$ columns $v_{r+1}, \ldots, v_p$ of V recursively. Let us show that

$$\text{for } k \in [r+1:p]: \quad |\mathcal{S}_{k-1}| \subseteq |\mathcal{S}_k| \subseteq |\mathcal{S}|. \tag{5.11b}$$

The first inclusion in (5.11b) comes from the fact that algorithm builds a sign vector in $\mathcal{S}_k$ for each sign vector in $\mathcal{S}_{k-1}$, in any of its step 3 or 4 (proposition 4.8(3)). The second inclusion in (5.11b) 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 LOPs to solve is bounded by $|\mathcal{S}_{k-1}|$, since at most one LOP is solved (in step 4.3 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 upper bound in proposition 4.11 and the fact that there are $k-1$ vectors in $V_{:,[1:k-1]}$; while the second bound comes from (5.11b)). Therefore, the total number of LOPs solved by the algorithm is bounded by

$$\min(2^r, |\mathcal{S}|) + \min(2^{r+1}, |\mathcal{S}|) + \cdots + \min(2^{p-1}, |\mathcal{S}|) \leqslant \min\big(2^p - 2^r, (p-r)|\mathcal{S}|\big).$$

This is the announced bound (5.10). □

5.2.8 Numerical experiments

We present in tables 5.1, 5.2 and 5.3 the results obtained by running the algorithms 5.8 and 5.12 (with several variants) on a small number of problems and compare it with our implementation of the RC algorithm 5.6, simulating algorithm 1 (IE) in [66].

A. Computer and problem presentation

The implementations have been done in Matlab (version "`9.11.0.1837725 (R2021b) Update 2`") on a `MacBookPro18,2/10cores` (parallelism is not implemented however) with the system `macOS Monterey`, version `12.6.1`. The linear optimization problem solver is `linprog`.

Computation in ISF is done in floating point numbers, so that numerical roundoff errors may occur. To deal with this difficulty, the code uses various tolerances, for instance, to detect almost identical normalized vectors (columns of V), to identify nonzero components of circuits, *etc.* The Julia code described in [32], which deals with more general hyperplane arrangements, offers the user the possibility of requiring a computation in rational numbers, so as to have a computation in exact arithmetic.

We have assessed the codes on randomly generated problems (function `rand` in Matlab, names prefixed by `rand` and `srand`) and problems adapted/taken from [66] (names prefixed by `rc`) and [17] (names prefixed by `bek`). Their relevant features are given in table 5.1 and their specifications are now given.

- The `rand-n-p-r` problems have their data formed of a randomly generated matrix $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ with prescribed rank `r`.
- For the problems `srand-n-p-q`, the first `n` columns of $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ form the identity matrix and the last $\mathtt{p}-\mathtt{n} > 0$ columns have `q` nonzero random integer elements ($0 < \mathtt{q} \leqslant \mathtt{p}-\mathtt{n}$), randomly positioned.
- The matrix $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ of problem `rc-2d-n-p` is formed of 4 blocs: $V_{1:2,1:\mathtt{n}-2} = 0$, $V_{3:\mathtt{n},\mathtt{n}-1:p} = 0$, and the remaining blocks have random integer data.
- The problems `rc-perm-n` refer to the hyperplane arrangements that are called *permutahedron* in [66]: the matrix $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ is such that $V_{:,[1:\mathtt{n}]}$ is the identity matrix and $V_{:,[\mathtt{n}:\mathtt{p}]}$ is a Coxeter matrix [63] (each column is of the form $e_i - e_j$ for some $i \neq j$ in $[1:\mathtt{n}]$, where e_k is the kth basis vector of $\mathbb{R}^{\mathtt{n}}$).
- The problems `rc-ratio-n-p-r` refer to the problems that are controlled by a degeneracy ratio ρ in [66]: the first `n` columns of the matrix $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ are randomly generated, while the other $\mathtt{p}-\mathtt{n} > 0$ columns can either (with a probability ρ) be linear combination of the previously generated columns or randomly generated.
- The problems `bek-threshold-n` refer to the *threshold arrangements* in [17; § 6.2]: for $n \geqslant 2$, each column of $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ is formed of the components of $(1, w)$ where $w \in \mathbb{R}^{n-1}$ are all the

vectors of $\{0,1\}^{n-1}$ (hence $p = 2^{n-1}$). This arrangement appears in the study of neural networks [83].

- The problems `bek-resonance-n` refer to the *resonance arrangements* in [17; § 6.3]: the columns of $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ are all the nonzero vectors with components in $\{0,1\}$ (hence $p = 2^n - 1$). Note that, for this arrangement, the number of chambers (i.e., $|\mathcal{S}|$ in our notation) is only known for $n \leqslant 9$. Our approach, which does not use the particular structure of this arrangement, can get $|\mathcal{S}|$ in a reasonable time on a laptop for $n \leqslant 6$, which is to be compared to $n \leqslant 9$ in [17]. See [54] for applications.
- The problems `bek-crosspolytope-n` refer to the *cross-polytope arrangements* in [17; § 6.4]: for $n \geqslant 2$, each column of $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ is formed of the components of $(1, w)$ where $w \in \mathbb{R}^{n-1}$ are all the $\pm e_i$ for $i \in [1\,{:}\,n-1]$; hence $p = 2(n-1)$. For these problems, one numerically observes that $|\mathcal{S}| = 2^1 3^{n-1} - 2^{n-1}$ for $n \leqslant 12$ (this observation is made for $n \leqslant 21$ in [17]).
- The problems `bek-demicube-n` refer to the *demicube arrangements* in [17; § 6.6]: the columns of $V \in \mathbb{R}^{\mathtt{n}\times\mathtt{p}}$ are the components of $(1, w)$ where $w \in \{w' \in \{0,1\}^{n-1} : \sum_i w'_i \text{ is odd}\}$.

We have retained 3 problems per family, the most difficult that ISF can solve in a reasonable time for the `rc-perm` and `bek` families. These test-problems are available on Software Heritage [35].

B. Observations on table 5.1

The dimensions n, p and r of the problems are given in columns 2-4 of table 5.1. Column 5 gives the number ς of matroid circuits of V, which is known to be bounded by $\varsigma_{\max} := \binom{p}{r+1}$ ($= 0$ if $r = p$) [26; 2006, theorem 2.1], whose value is given in column 6. In columns 7 and 8, one finds the cardinality $|\partial_B H(x)| = |\mathcal{S}|$ of the B-differential $\partial_B H(x)$ and the Schläfli upper bound (the right-hand side of (4.13)). The codes will be compared on the number of linear optimization problems (LOPs) they 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 LOP solver. A first example of comparison is given in columns 9–11 of table 5.1, where one finds the number of LOPs solved by the original RC algorithm and the simulated RC algorithm implemented in the ISF code, as well as the difference between these two numbers. The latter code will be used next, in the comparison with its improved versions, both regarding the LOP counters (table 5.2) and the CPU times (table 5.3).

1) The randomly generated problems `rand` are likely to provide vectors v_i's (the columns of V) in general position, in the sense of definition 4.13. This can be seen indirectly on the numbers in table 5.1.
   - It is known from proposition 4.15 that (4.12) implies equality in (4.13). This equality indeed holds, as we can observe by comparing columns 7 and 8.
   - The same phenomenon occurs with the bound $\varsigma_{\max}$, which is reached by ς if and only if the vectors are in general position [26; 2006, theorem 2.1].
   - Incidentally, one can compute mentally Schläfli's bound β when p is even and $r = p/2$. In that case, the right-hand side of (4.13) reads

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

     This is what one observes in [31; table 5.1]; for example, when $p = 8$ and $r = 4$, one has $\beta = 128$, which is indeed 2^{8-1}.
2) One observes in [31; table 5.1] that when $r = 2$, one has $|\partial_B H(x)| = 2p$ (proposition 4.18).

Problem	n	p	r	ς	$\varsigma_{\max}$	$\lvert\partial_B H(x)\rvert$	Schläfli's bound	LOPs solved in: Original RC	LOPs solved in: Simulated RC	LOPs solved in: Difference
`rand-8-15-7`	8	15	7	6435	6435	12952	12952	9908	9907	1
`rand-9-16-8`	9	16	8	11440	11440	32768	32768	22821	22818	3
`rand-10-17-9`	10	17	9	19448	19448	78406	78406	50643	50642	1
`srand-8-20-2`	8	20	8	540	167960	24544	188368	28748	28620	128
`srand-8-20-4`	8	20	8	84390	167960	157192	188368	136133	135566	567
`srand-8-20-6`	8	20	8	159702	167960	186430	188368	167545	167262	283
`rc-2d-20-6`	6	20	6	560	77520	512	33328	1936	1927	9
`rc-2d-20-7`	7	20	7	455	125970	960	87592	3392	3343	49
`rc-2d-20-8`	8	20	8	364	167960	1792	188368	5888	5855	33
`rc-perm-6`	6	21	6	1172	116280	5040	43400	10417	9346	1071
`rc-perm-7`	7	28	7	8018	4292145	40320	795188	99155	90169	8986
`rc-perm-8`	8	36	8	62814	94143280	362880	17463696	1036897	953009	83888
`rc-ratio-20-5-7`	5	20	5	34556	38760	8470	10072	13798	13785	13
`rc-ratio-20-6-7`	6	20	6	56184	77520	26194	33328	32993	32980	13
`rc-ratio-20-7-7`	7	20	7	112576	125970	76790	87592	82751	82738	13
`bek-threshold-4`	4	8	5	20	28	104	128	88	87	1
`bek-threshold-5`	5	16	5	1348	8008	1882	3882	2758	2757	1
`bek-threshold-6`	6	32	6	353616	3365856	94572	412736	248522	248521	1
`bek-resonance-4`	4	15	4	638	3003	370	940	705	635	70
`bek-resonance-5`	5	31	5	100091	736281	11292	63862	37766	36311	1455
`bek-resonance-6`	6	63	6	(1)	553270671	1066044	14137242	6272462	6164040	108422
`bek-crosspolytope-11`	11	20	11	45	125970	117074	709044	111442	86526	24916
`bek-crosspolytope-12`	12	22	12	55	497420	352246	2802584	339958	260601	79357
`bek-crosspolytope-13`	13	24	13	66	1961256	1058786	11092764	1032162	788970	243192
`bek-demicube-5`	5	8	5	6	28	146	198	106	99	7
`bek-demicube-6`	6	16	6	460	11440	3756	9888	4752	4719	33
`bek-demicube-7`	7	32	7	324640	10518300	291558	1885298	678453	674663	3790

Table 5.1: Description of the test-problems and comparison of the "original RC algorithm in [66]", written in Python, and the "simulated RC algorithm 5.6", written in Matlab: "(n, p, r, ς)" are the features of the problem ($V \in \mathbb{R}^{n\times p}$ is of rank r and has ς circuits, this last number being known to be bounded by $\varsigma_{\max}$), "$|\partial_B H(x)|$" is the cardinality of the B-differential of H given by (1.3), "Schläfli's bound" is the right-hand side of (4.13), "Original RC" gives the number of linear optimization problems (LOPs) solved by the original piece of software in Python of Rada and Černý [66], "Simulated RC" gives the number of LOPs solved by the implementation in the Matlab code ISF of the Rada and Černý algorithm (see algorithm 5.6), "Difference" is the difference between the two previous columns. Note (1): computer crash after several weeks of computation.

3) The number of matroid circuits, given in the column labeled by ς, depends on the determination of the nonzero elements of the normalized vector $\alpha \in \mathcal{N}(V_{:,I}) \setminus \{0\}$ for the selected index set I (proposition 3.11). This operation is sensitive to a threshold value that is set to $10^5\varepsilon$, where $\varepsilon > 0$ is the machine epsilon; smaller values for this threshold have occasionally given larger numbers of matroid circuits. In other words, due to the floating point calculation, there is no certainty that the given number of circuits is the one that would be obtained in exact arithmetic. With a computation in rational numbers, this difficulty is avoided [32].
4) A comparison between the "Original RC code" in Python and its "Simulated RC code" in Matlab shows that the latter is slightly more effective in terms of the number of LOPs solved. This is probably due to the special treatment in step 2 of the case where $v_{k+1}^{\mathsf{T}} d \simeq 0$ in algorithm 5.7, which is not considered in the original code.

C. Observations on table 5.2

Table 5.2 shows the effect of the modifications discussed in section 5.2.5 on the number of LOPs solved, which significantly counts in the computing time. This will lead us to select three algorithms, those which bring the best profit on the LOP counter. The columns labeled "Ratio" show the acceleration ratio with respect to the simulated RC code in terms of LOPs, that is the ratio of the LOP counter of the considered algorithm divided by the LOP counter of the simulated RC algorithm. On the last two lines of the table, one finds the mean and median values of these acceleration ratios, which may be viewed as a summary of the effect of the considered modification. These mean/median values must be taken with caution when a solver fails to solve a problem as is the case with ISF(ABCD$_3$) and ISF(AD$_4$) on problem `bek-resonance-6`.

1) The modification A, proposed in section 5.2.5(A), which uses the QR factorization to get r linearly independent columns of V, does not bring a large benefit ("Ratio" is close to 1) and sometimes increases the number of LOPs to solve. The benefit is not important since it "only" prevents $\sum_{i\in[0:r-1]} 2^i = 2^r - 1$ nodes from running the LOP solver, which is usually a small fraction of the total number of nodes of the $\mathcal{S}$-tree. One also observes that the number of solved LOPs may increase (acceleration ratio < 1), which is sometimes due to the fact that the number 2^{r-1} of nodes at level r with modification A is larger than the one without modification A, which contributes to increase the total number of nodes of the constructed $\mathcal{S}$-tree and, therefore, tends to increase the number of LOPs to solve. Furthermore, the order in which the vectors are considered without/with modification A is not identical, which has also an impact on the number of solved LOPs (see section 5.2.5(C)).
2) The modification B, proposed in section 5.2.5(B), which is able to detect two descendants of an $\mathcal{S}$-tree node, without solving any LOP, has a significant impact on the total number of these problems. We see, indeed, that the (mean, median) acceleration ratio is raised to (1.28, 1.17).

	Number of linear optimization problems (LOPs) solved and acceleration ratio (Ratio) for various options														
	Simulated	ISF (A)		ISF (AB)		ISF (ABC)		ISF ($ABCD_1$)		ISF ($ABCD_2$)		ISF ($ABCD_3$)		ISF (AD_4)	
Problem	RC	LOP	Ratio	LOP	Ratio	LOP	Ratio	LOP	Ratio	LOP	Ratio	LOP	Ratio	LOP	Ratio
rand-8-15-7	9907	9844	1.01	7641	1.30	5210	1.90	5199	1.91	4355	2.27	3638	2.72	0	—
rand-9-16-8	22818	22691	1.01	17586	1.30	13046	1.75	13023	1.75	11185	2.04	9943	2.29	0	—
rand-10-17-9	50642	50387	1.01	38167	1.33	28849	1.76	28839	1.76	25370	2.00	23266	2.18	0	—
srand-8-20-2	28620	28620	1.00	20207	1.42	6668	4.29	5535	5.17	2881	9.93	2851	10.04	0	—
srand-8-20-4	135566	136027	1.00	113493	1.19	60066	2.26	59267	2.29	45569	2.97	42445	3.19	0	—
srand-8-20-6	167262	167351	1.00	137450	1.22	77800	2.15	77752	2.15	62694	2.67	54980	3.04	0	—
rc-2d-20-6	1927	1904	1.01	1680	1.15	912	2.11	688	2.80	40	48.17	0	—	0	—
rc-2d-20-7	3343	3296	1.01	2912	1.15	2208	1.51	1792	1.87	52	64.29	0	—	0	—
rc-2d-20-8	5855	5760	1.02	4992	1.17	2752	2.13	1984	2.95	28	209.11	0	—	0	—
rc-perm-6	9346	9280	1.01	7898	1.18	2076	4.50	1836	5.09	92	101.59	61	153.21	0	—
rc-perm-7	90169	90094	1.00	79049	1.14	17230	5.23	16558	5.45	960	93.93	855	105.46	0	—
rc-perm-8	953009	952597	1.00	856597	1.11	160781	5.93	158989	5.99	9766	97.58	9393	101.46	0	—
rc-ratio-20-5-7	13669	15341	0.89	14028	0.97	7108	1.92	7064	1.94	3644	3.75	2467	5.54	0	—
rc-ratio-20-6-7	32883	35882	0.92	31992	1.03	17797	1.85	17505	1.88	10669	3.08	8765	3.75	0	—
rc-ratio-20-7-7	82447	81428	1.01	72272	1.14	47798	1.72	47748	1.73	30442	2.71	25841	3.19	0	—
bek-threshold-4	87	79	1.10	54	1.61	46	1.89	37	2.35	26	3.35	16	5.54	0	—
bek-threshold-5	2757	2884	0.96	2399	1.15	1270	2.17	1180	2.34	502	5.49	370	3.75	0	—
bek-threshold-6	248521	261728	0.95	236027	1.05	71963	3.45	70410	3.53	21339	11.65	19184	3.19	0	—
bek-resonance-4	635	672	0.94	546	1.16	171	3.71	138	4.60	31	20.48	0	—	0	—
bek-resonance-5	36311	37607	0.97	34056	1.07	6700	5.42	6569	5.53	1141	31.82	810	44.83	0	—
bek-resonance-6	6164040	6269410	0.98	5956586	1.03	760930	8.10	760457	8.11	155555	39.63	(1)	—	0	—
bek-crosspolytope-11	86526	110418	0.78	58954	1.47	17569	4.92	15265	5.67	6085	14.22	6049	14.30	0	—
bek-crosspolytope-12	260601	337910	0.77	182575	1.43	46900	5.56	41780	6.24	18785	13.87	18740	13.91	0	—
bek-crosspolytope-13	788970	1028066	0.77	560013	1.41	124828	6.32	113564	6.95	57299	13.77	57244	13.78	0	—
bek-demicube-5	99	90	1.10	33	3.00	24	4.12	12	8.25	3	33.00	0	—	0	—
bek-demicube-6	4719	4761	0.99	3659	1.29	1882	2.51	1741	2.71	665	7.10	588	8.03	0	—
bek-demicube-7	674663	704553	0.96	623160	1.08	175870	3.84	175595	3.84	60876	11.08	58333	11.57	0	—
Mean			0.97		1.28		3.45		3.88		31.54		24.52		—
Median			1.00		1.17		2.51		2.95		11.65		5.54		—

Table 5.2: Evaluation of the efficiency of the solvers by the number of LOPs they solve: A (taking the rank of V into account), B (special handling of the case where $v_{k+1}^{\mathsf{T}} d \simeq 0$), C (changing the order of the vectors v_i's by taking i_{k+1} by (5.9)), D_1 (pre-computation of $2(p-r)$ stem vectors after the QR factorization), D_2 (D_1 and 2 additional stem vectors computed after solving a LOP, whose optimal value is nonnegative), D_3 (all the stem vectors are first computed and, for $(s, \pm 1) \in \mathcal{S}_{k+1}$, a LOP is solved to get a handle d), D_4 (all the stem vectors are first computed and no LOP is solved). The "Ratio" (acceleration ratio) columns give for each considered problem the ratio (*LOPs of the considered* ISF *version*)/(*LOPs of simulated* RC). Note (1): interruption of the run after several days of computation. The Mean/Median rows give the mean and medan values of the ratios.

3) Consider now the modification C, described in section 5.2.5(C), which changes the order in which the vectors v_i's are considered. We use the test-problem `rand-7-13-5` to show its effect in the next table.

	Number of nodes per level													Total
With modifications AB	1	2	4	8	16	31	57	99	163	256	386	562	794	2379
With modifications ABC	1	2	4	8	16	26	43	69	107	168	270	443	794	1951
$\mathcal{S}$-tree levels	1	2	3	4	5	6	7	8	9	10	11	12	13	

The table gives the number of nodes for each level in the $\mathcal{S}$-tree, with the modifications AB and with the modifications ABC. Since $\text{rank}(V) = 5$ for this problem and since the modification A is used in both cases, the number of nodes per level, only starts to differ from level 6 (before that it is equal to 2^{l-1}, where l is the $\mathcal{S}$-tree level). The final level is 13 (since there are $p = 13$ vectors) and its number of leaves is $|\mathcal{S}|/2 = 794$ (an observation from the table above), necessary identical in both cases. The effect of the modification C can be seen on the smaller number of nodes per level and in all the $\mathcal{S}$-tree (rightmost column). This contributes to the decrease of the number of LOPs to solve: the (mean, median) acceleration ratio is raised to (3.45, 2.51).

4) The modifications D, described in section 5.2.5(D), deal with the contribution of the computed stem vectors, whose number increases from modification D_1 ($2(p-r)$ stem vectors after the QR factorization of V), D_2 (more stem vectors from the dual solution of the LOP (5.2) when this one has a nonnegative optimal value), D_3 and D_4 (all the stem vectors).
   - We see that the option D_1 yields already some improvement (less LOPs to solve), but not much, raising the (mean, median) acceleration ratio from (3.45, 2.51) to (3.88, 2.95).
   - The use of the option D_2 is more beneficial since the (mean, median) acceleration ratio now goes up to (31.54, 11.65). We understand this fact to have its origin in the increase in the number of stem vectors detected from the dual solutions of some solved LOP. Note that this last operation does not require much computation time.
   - With option D_3, only the LOPs (5.2) with the optimal value -1 are solved, while, with option D_4, no LOP is solved. The efficiency of these modifications largely depends on the total number 2ς of stem vectors. If this one is not too large, the modifications have an important benefit. Otherwise, it can lead to execution failure, as for problem `bek-resonance-6`, which requires days of computation.

In conclusion of these observations, one could retain the following three solvers for a comparison on their computing time.

- ISF(ABCD$_2$) is the most efficient solver that does not compute all the stem vectors.
- The solvers ISF(ABCD$_3$) and ISF(AD$_4$) cannot be compared with the other solvers on the results of table 5.2 since both use all the stem vectors, so that the time to compute and use these must be taken into account, and ISF(AD$_4$) does not solve any LOP, which is the measure of efficiency in table 5.2.

D. Observations on table 5.3

Measuring the efficiency of the algorithms by the number of LOPs solved during execution, like in table 5.2, is sometimes misleading. If this is the main cost item for some algorithms, it is no longer the case when a large amount of stem vectors is computed. For two reasons. First, the time spent in the computation of these stem vectors is not negligible, far from it, at least in our implementation, in which each of them requires the computation of the nullity

	CPU times (in sec)						
	Simulated	ISF ($ABCD_2$)		ISF ($ABCD_3$)		ISF (AD_4)	
Problem	RC	Time	Ratio	Time	Ratio	Time	Ratio
rand-8-15-7	71.77	33.27	2.16	32.91	2.18	5.62	12.77
rand-9-16-8	151.39	75.45	2.01	82.30	1.84	14.43	10.49
rand-10-17-9	347.32	185.05	1.88	198.18	1.75	55.96	6.21
srand-8-20-2	174.44	16.91	10.32	19.64	8.88	3.66	47.68
srand-8-20-4	832.74	309.15	2.69	450.35	1.85	349.83	2.38
srand-8-20-6	1011.30	483.97	2.09	732.82	1.38	746.49	1.35
rc-2d-20-6	11.01	0.32	34.71	0.25	43.53	0.22	50.95
rc-2d-20-7	19.88	0.50	39.95	0.50	39.68	0.38	52.97
rc-2d-20-8	35.87	0.41	87.97	0.74	48.56	0.63	56.78
rc-perm-6	53.29	0.76	70.05	2.10	25.41	1.90	28.00
rc-perm-7	549.04	7.44	73.78	45.62	12.04	67.10	8.18
rc-perm-8	6171.22	74.93	82.36	1233.80	5.00	3355.22	1.84
rc-ratio-20-5-7	83.34	22.71	3.67	28.36	2.94	18.58	4.49
rc-ratio-20-6-7	202.09	72.04	2.81	101.51	1.99	112.12	1.80
rc-ratio-20-7-7	504.52	247.99	2.03	351.08	1.44	353.15	1.43
bek-threshold-4	0.61	0.18	3.44	0.11	5.46	0.01	74.64
bek-threshold-5	17.43	3.56	4.89	2.83	6.16	0.35	50.40
bek-threshold-6	1758.16	194.75	9.03	4577.26	0.38	6532.56	0.27
bek-resonance-4	3.97	0.22	17.71	0.09	46.12	0.08	48.99
bek-resonance-5	228.41	7.90	28.90	44.78	5.10	183.84	1.24
bek-resonance-6	38296.20	1988.60	19.26	(1)	—	(1)	—
bek-crosspolytope-11	480.07	34.35	13.97	39.27	12.22	7.63	62.95
bek-crosspolytope-12	1579.19	108.76	14.52	124.66	12.67	25.22	62.62
bek-crosspolytope-13	5017.73	322.43	15.56	404.80	12.40	104.22	48.15
bek-demicube-5	0.55	0.02	25.24	0.01	85.73	0.01	108.69
bek-demicube-6	27.38	4.15	6.59	4.09	6.69	0.43	63.82
bek-demicube-7	4310.35	510.25	8.45	2405.08	1.79	6396.66	0.67
Mean			21.71		15.12		31.14
Median			10.32		5.81		20.39

Table 5.3: Evaluation of the efficiency of the solvers by their computing times. The "Ratio" (acceleration ratio) columns give for each considered problem the ratio (*Time of the considered* ISF *version*)/(*Time of simulated* RC). Note (1): interruption of the run after several days of computation. The Mean/Median rows give the mean and medan values of the ratios.

of a matrix and a null space vector. Next, verifying that a sign vector contains a stem vector (proposition 3.10) is also time consuming when there are many stem vectors. Therefore a comparison of the CPU time of the runs is welcome. This is done for a selection of versions of the ISF codes in table 5.3, those selected at the end of section 5.2.8(C). Here are some observations on the statistics of this table.

1) A first observation is that the good behavior of the selected versions of the ISF codes is confirmed, even though the acceleration ratios are not as large as the one based on the number of LOPs solved. This can be explained by the fact that the time spent in solving LOPs is counterbalanced by the handling of stem vectors for the versions ABCD_3 and AD_4. Anyway, one observes that the CPU time acceleration ratios have (mean, median) values in the ranges (15..31, 5..20), which is significant.
2) The most effective combination of code options depends actually on the considered problems. It is difficult to state a rule that would predict which code behaves best because some solvers are better on some phases of the run, but worse on others (the three main phases are the detection of the stem vectors, the execution of LOPs and the search for stem vectors covered by a given sign vector). However, an inductive rule manifests itself: the purely dual method AD_4 is ahead for problems with a reasonable number of stem vectors (or matrix circuits), but can require a too large number of computing time if this number becomes large (this is the case of problems `bek-threshold-6`, `bek-resonance-6` and `bek-demicube-7`). This conclusion could be invalidated if better techniques are used to enumerate and use the stem vectors.

5.2.9 A numerical example

Consider the LCP in standard form (1.5), which reads $0 \leqslant x \perp (Mx+q) \geqslant 0$, 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

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

Since M is a **P**-matrix (i.e., all its principal minors are positive), the problem has a unique solution [72], which is $x = 0$. With the notation (2.2), one has $\mathcal{A}(x) = \mathcal{B}(x) = \varnothing$ and $\mathcal{E}(x) = \mathcal{E}^{\neq}(x) = \{1,2,3\}$, so that V^{T} given by (3.1) reads

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

Note that $p = 3$, while $\operatorname{rank}(V) = 2$, so that $|\partial_B H(x)| = 2p = 6$, by proposition 4.18. The sign vectors $s \in \{\pm 1\}^3$ that make $s \bullet (V^{\mathsf{T}} d) > 0$ feasible for some 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 $\partial_B H(x)$ are obtained for the s's in $\mathcal{S}$ given above by the bijection σ defined by (3.4). One gets a set of 6 Jacobians out of the $2^3 = 8$ Jacobians in

$\partial_B^{\times} H(x)$, namely

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

From the matrix S given above, we see that $\mathcal{S}^c := \{\pm 1\}^3 \backslash \mathcal{S} = \{s, -s\}$, where $s = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}^{\mathsf{T}}$. Observe also that, since

$$\pm \operatorname{Diag}(s) V^{\mathsf{T}} = \pm \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & -1 & -1 \end{pmatrix},$$

there is no $d \in \mathbb{R}^3$ such that $\pm s \bullet (V^{\mathsf{T}} d) > 0$, as expected.

6 Discussion

This paper deals with the description and computation of the B-differential of the componentwise minimum of two affine vector functions. The fact that this problem has many equivalent formulations, some of them being highlighted in section 3, implies that the present contribution has an impact on several domains, including on the description of the arrangement of hyperplanes in the space. To this respect, a singular aspect of this contribution is to propose a dual approach to solve the problem, using some or all the stem vectors, a concept made useful thanks to the convex analysis tool that is Gordan's alternative. Besides this contribution, the paper also brings various improvements of an algorithm of Rada and Černý [66], which was designed to determine the cells of an arrangement of hyperplanes in the space.

Even in the spirit of the methods proposed in this article, there is still room for improvement, in relation to three identified bottlenecks: (i) we have mentioned that with the option D_3, the LOP (5.2) can be solved inexactly, since, in that case, the optimal value is -1, while any negative objective value for a feasible unknown would suffice, but this requires a better tuning of the linear optimization solver, (ii) computing more efficiently all the stem vectors (or matroid circuits) of the matrix V is certainly a source of improvement, (iii) a better algorithm to decide more rapidly that a sign vector contains a stem vector is also welcome. Some of these possible improvements are also linked to a better choice of programming language, probably one using a compilation phase.

This contribution has also various possible extensions. A first one would be to develop a dual approach to the problem of the arrangement in the space of hyperplanes *having no point in common* [32]. Another natural extension would be to see the implications of this work for computing the B-differential of the componentwise minimum of *nonlinear* vector functions [33]. Finally, the possibility to take profit of the computation of the full B-differential of the function H in (1.3) in a Newton-like approach to solve (1.4) is a subject that deserves reflection.

Acknowledgments

We thank Černý and Rada for providing their code and test problems, those used in [66]; part of these were used in the numerical experiments. We also thank the referees for their remarks and recommendations, which have helped us make the paper more readable.

References

[1] Muhamed 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]. 19, 21, 23, 26

[3] Gerald L. Alexanderson, John E. Wetzel (1981). Arrangements of planes in space. *Discrete Mathematics*, 34(3), 219–240. [doi]. 26

[4] Gerald L. Alexanderson, John E. Wetzel (1983). Erratum: "arrangements of planes in space". *Discrete Mathematics*, 45(1), 140. [doi]. 26

[5] David Avis, Komei Fukuda (1996). Reverse search for enumeration. *Discrete Applied Mathematics*, 65(1-3), 21–46. [doi]. 9, 19, 22, 38

[6] Pierre Baldi (2018). Deep learning in biomedical data science. *Annual Review of Biomedical Data Science*, 1, 181–205. [doi]. 17

[7] 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]. 9, 16

[8] Ibtihel Ben Gharbia, Jean Charles Gilbert (2012). Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a P-matrix. *Mathematical Programming*, 134, 349–364. [doi]. 3

[9] Ibtihel Ben Gharbia, Jean Charles Gilbert (2013). An algorithmic characterization of P-matricity. *SIAM Journal on Matrix Analysis and Applications*, 34(3), 904–916. [doi]. 3

[10] Ibtihel Ben Gharbia, Jean Charles 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

[11] Ibtihel Ben Gharbia, Jérôme Jaffré (2014). Gas phase appearance and disappearance as a problem with complementarity constraints. *Mathematics and Computers in Simulation*, 99, 28–36. [doi]. 3

[12] Dimitri P. Bertsekas (1999). *Nonlinear Programming* (second edition). Athena Scientific. 43

[13] Hanspeter Bieri, Walter Nef (1982). A recursive sweep-plane algorithm, determining all cells of a finite division of $\mathbb{R}^d$. *Computing*, 28(3), 189–198. [doi]. 38

[14] J. Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, Claudia Sagastizábal (1997). *Optimisation Numérique – Aspects théoriques et pratiques*. Mathématiques et Applications 27. Springer Verlag, Berlin. [editor]. 40, 43

[15] J. Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, Claudia Sagastizábal (2006). *Numerical Optimization – Theoretical and Practical Aspects* (second edition). Universitext. Springer Verlag, Berlin. [authors] [editor] [doi]. 40, 43

[16] Jonathan Michael Borwein, Adrian S. Lewis (2006). *Convex Analysis and Nonlinear Optimization – Theory and Examples* (second edition). CMS Books in Mathematics 3. Springer, New York. 3, 15

[17] Taylor Brysiewicz, Holger Eble, Lukas Kühne (2023). Computing characteristic polynomials of hyperplane arrangements with symmetries. *Discrete & Computational Geometry*, 70, 1356–1377. 49, 50

[18] 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]. 19

[19] Vaček Chvátal (1983). *Linear Programming*. W.H. Freeman and Company, New York. 40, 43

[20] Frank 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

[21] Kenneth L. Clarkson, Peter W. Shor (1989). Applications of random sampling in computational geometry, II. *Discrete & Computational Geometry*, 4, 387–421. [doi]. 26

[22] Richard Warren Cottle, George Bernard Dantzig (1970). A generalization of the linear complementarity problem. *Journal of Combinatorial Theory*, 8(1), 79–90. 3

[23] Richard Warren Cottle, Jong-Shi Pang, Richard E. Stone (2009). *The Linear Complementarity Problem*. Classics in Applied Mathematics 60. SIAM, Philadelphia, PA, USA. [doi]. 3

[24] Thomas M. Cover, Joy A. Thomas (2006). *Elements of information theory* (second edition). Wiley-Interscience [John Wiley & Sons], Hoboken, NJ. 42

[25] Tecla De Luca, Francisco Facchinei, Christian Kanzow (2000). A theoretical and numerical comparison of some semismooth algorithms for complementarity problems. *Computational Optimization and Applications*, 16, 173–205. [doi]. 7

[26] Gy Dósa, I. Szalkai, C. Laflamme (2006). The maximum and minimum number of circuits and bases of matroids. *Pure Mathematics and Applications*, 15(4), 383–392. 50

[27] Jean-Pierre Dussault, Mathieu Frappier, Jean Charles 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

[28] Jean-Pierre Dussault, Mathieu Frappier, Jean Charles Gilbert (2023). Polyhedral Newton-min algorithms for complementarity problems. *Mathematical Programming* (in revision). [hal-02306526]. 3

[29] Jean-Pierre Dussault, Jean Charles Gilbert (2023). Exact computation of an error bound for the balanced linear complementarity problem with unique solution. *Mathematical Programming*, 199(1-2), 1221–1238. [doi]. 3

[30] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). On the B-differential of the componentwise minimum of two affine vector functions. *Mathematical Programming Computation* (submitted). 1, 5

[31] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). On the B-differential of the componentwise minimum of two affine vector functions – The full report. Research report (version 1). [hal-03872711v1]. 50

[32] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). Primal and dual approaches for the chamber enumeration of hyperplane arrangements. Research report (in preparation). 38, 49, 52, 57

[33] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). Partial description of the B-differential of the componentwise minimum of two vector functions by linearization. Research report (in preparation). 21, 36, 37, 57

[34] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). ISF and BDIFFMIN - Matlab functions for central hyperplane arrangements and the computation of the B-differential of the componentwise minimum of two affine vector functions. Technical report. [hal]. 37

[35] Jean-Pierre Dussault, Jean Charles Gilbert, Baptiste Plaquevent-Jourdain (2023). ISF and BDIFFMIN. [hal] [swh]. 37, 50

[36] Herbert Edelsbrunner (1987). *Algorithms in Combinatorial Geometry*, volume 10 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin. [doi]. 19, 26

[37] Herbert Edelsbrunner, Joseph O'Rourke, Raimund Seidel (1986). Constructing arrangements of lines and hyperplanes with applications. *SIAM Journal on Control*, 15(2), 341–363. [doi]. 19, 38

[38] L.C. Evans, R.F. Gariepy (2015). *Measure Theory and Fine Properties of Functions* (revised edition). CRC Press, Boca Raton. 3

[39] Francisco Facchinei, Jong-Shi Pang (2003). *Finite-Dimensional Variational Inequalities and Complementarity Problems* (two volumes). Springer Series in Operations Research. Springer. 3, 4

[40] Francisco Facchinei, João Soares (1997). A new merit function for nonlinear complementarity problems and a related algorithm. *SIAM Journal on Optimization*, 7(1), 225–247. [doi]. 3

[41] H. Federer (1996). *Geometric Measure Theory*. Grundlehren der mathematischen Wissenschaften 153. Springer-Verlag, New York. 3

[42] Jean Charles Gilbert (2021). *Fragments d'Optimisation Différentiable – Théorie et Algorithmes*. Lecture Notes (in French) of courses given at ENSTA and at Paris-Saclay University, Saclay, France. [hal-03347060, pdf]. 36, 43

[43] Jean Charles Gilbert (2022). *Selected Topics on Continuous Optimization – Version* 2. Lecture notes of the Master-2 "Optmization" at the University Paris-Saclay. [hal]. 40

[44] Paul Gordan (1873). Über die Auflösung linearer Gleichungen mit reellen Coefficienten. *Mathematische Annalen*, 6, 23–28. 12

[45] Rick Greer (1984). *Trees and Hills: Methodology for Maximizing Functions of Systems of Linear Relations*. Mathematical Studies 96, Annals of Discrete Mathematics 22. North-Holland. 15

[46] Branko Grünbaum (1972). *Arrangements and Spreads*. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics 10. AMS, Providence, RI. 19, 26

[47] 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. 19

[48] Juha Heinonen (2005). Lectures on Lipschitz Analysis. Report 100, Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä. 3

[49] Jean-Baptiste Hiriart-Urruty, Claude Lemaréchal (2001). *Fundamentals of Convex Analysis*. Springer. 3

[50] Alexey F. Izmailov, Mikhail V. Solodov (2014). *Newton-Type Methods for Optimization and Variational Problems*. Springer Series in Operations Research and Financial Engineering. Springer. [doi]. 3

[51] Christian Kanzow, Masao 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]. 7

[52] Leonid G. Khachiyan, Endre Boros, Khaled M. Elbassioni, Vladimir A. Gurvich, Kazuhisa Makino (2006). On the complexity of some enumeration problems for matroids. *SIAM Journal on Discrete Mathematics*, 19(4), 966–984. [doi]. 42

[53] Masakazu Kojima, Susumu 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

[54] Lukas Kühne (2023). The universality of the resonance arrangement and its Betti numbers. *Combinatorica*, 43, 277–298. [doi]. 50

[55] Estelle Marchand, Torsten Müller, Peter 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

[56] Estelle Marchand, Torsten Müller, Peter 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

[57] Arnaud Mary, Yann Strozecki (2019). Efficient enumeration of solutions produced by closure operations. *Discrete Mathematics and Theoretical Computer Science*, 21(3), # 22. [doi]. 42

[58] Katta G. Murty (1988). *Linear Complementarity, Linear and Nonlinear Programming* (Internet edition, prepared by Feng-Tien Yu, 1997). Heldermann Verlag, Berlin. 3

[59] James Oxley (2011). *Matroid Theory* (second edition). Oxford Graduate Texts in Mathematics 21. Oxford University Press, Oxford. [doi]. 13, 30

[60] Jong-Shi Pang (1990). Newton's method for B-differentiable equations. *Mathematics of Operations Research*, 15, 311–341. [doi]. 3

[61] Jong-Shi 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

[62] Jong-Shi 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

[63] Alexander Postnikov, Richard P. Stanley (2000). Deformations of Coxeter hyperplane arrangements. *Journal of Combinatorial Theory – Series A*, 91(1-2), 544–597. [doi]. 49

[64] Liqun Qi (1993). Convergence analysis of some algorithms for solving nonsmooth equations. *Mathematics of Operations Research*, 18, 227–244. [doi]. 3, 5, 8, 37

[65] Liqun Qi, Jie Sun (1993). A nonsmooth version of Newton's method. *Mathematical Programming*, 58, 353–367. [doi]. 3

[66] 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, 38, 39, 40, 41, 48, 49, 51, 57

[67] Hans Rademacher (1919). Über partielle und totale differenzierbarkeit. *I. Math. Ann.*, 89, 340–359. 3

[68] Jörg Rambau (2023). Symmetric lexicographic subset reverse search for the enumeration of circuits, cocircuits, and triangulations up to symmetry. Draft 2. 42

[69] 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]. 19, 26

[70] S.M. Robinson (1987). Local structure of feasible sets in nonlinear programming, part III: stability and sensitivity. *Mathematical Programming Study*, 30, 45–66. [doi]. 3

[71] R. Tyrrell Rockafellar (1970). *Convex Analysis*. Princeton Mathematics Ser. 28. Princeton University Press, Princeton, New Jersey. 3, 18

[72] 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]. 56

[73] Ludwig Schläfli (1950). Theorie der vielfachen Kontinuität (in German). In *Gesammelte mathematische Abhandlugen,* Band 1, pages 168–387. Springer, Basel. [doi]. 31

[74] Jürgen Schmidhuber (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. 17

[75] 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]. 19, 22, 38

[76] Nora Helena Sleumer (2000). *Hyperplane arrangements – Construction, visualization and application.* PhD Thesis, Swiss Federal Institute of Technology, Zurich, Switzerland. [doi]. 19

[77] Marek J. Śmietánski (2019). On a new exponential iterative method for solving nonsmooth equations. *Numerical Linear Algebra with Applications*, 25. [doi]. 3

[78] Richard P. Stanley (2007). An introduction to hyperplane arrangements. In Ezra Miller, Victor Reiner, Bernd Sturmfels (editors), *Geometric Combinatorics*, pages 389–496. [doi]. 19, 26, 30

[79] Richard P. Stanley (2021). Enumerative and algebraic combinatorics in the 1960's and 1970's. *Notices of the International Congress of Chinese Mathematicians*, 9(2), 19–38. 26

[80] J. Steiner (1826). Einige Gesetze über die Theilung der Ebene und des Raumes. *Journal für die Reine und Angewandte Mathematik*, 1, 349–364. [doi]. 19, 26

[81] Roman Sznajder, Muddappa Seetharama Gowda (1994). The generalized order linear complementarity problem. *SIAM Journal on Matrix Analysis and Applications*, 15(3), 779–795. [doi]. 3

[82] Michel Las Vergnas (1975). Matroïdes orientables. *C. R. Acad. Sci. Paris Série A-B*, 280, A61–A64. 26

[83] Walter Wenzel, Nihat Ay, Frank Pasemann (2000). Hyperplane arrangements separating arbitrary vertex classes in n-cubes. *Acta Applicandae Mathematicae*, 25(3), 284–306. 16, 50

[84] Robert O. Winder (1966). Partitions of N-space by hyperplanes. *SIAM Journal on Applied Mathematics*, 14(4), 811–818. [doi]. 9, 26, 27, 30, 31

[85] Shuhuang Xiang, Xiaojun Chen (2011). Computation of generalized differentials in nonlinear complementarity problems. *Computational Optimization and Applications*, 50, 403–423. [doi]. 5, 8, 21, 36, 37

[86] 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]. 16, 26

[87] Chao Zhang, Xiaojun Chen, Naihua Xiu (2009). Global error bounds for the extended vertical LCP. *Computational Optimization and Applications*, 42(3), 335–352. [doi]. 3