Reliable Computing (2007) 13:325–349
DOI: 10.1007/s11155-007-9037-6

Extension of the Hansen-Bliek Method to Right-Quantified Linear Systems

GILLES CHABERT
Projet Coprin, INRIA, 2004 route des Lucioles, 06902 Sophia Antipolis, France, e-mail: gilles.chabert@sophia.inria.fr

and

ALEXANDRE GOLDSZTEJN
University of Central Arkansas, Conway, Arkansas, USA, e-mail: AGoldsztejn@uca.edu

(Received: 10 May 2006; accepted: 6 January 2007)

Abstract. The problem of finding the smallest box enclosing the united solution set of a linear interval system, also known as the "interval hull" problem, was proven to be NP-hard. However, Hansen, Bliek, and others subsequently, have provided a polynomial-time solution in the case of systems preconditioned by the midpoint inverse matrix.

Based upon a similar approach, this paper deals with the interval hull problem in the context of AE-solution sets, where parameters may be given different quantifiers. A polynomial-time algorithm is proposed for computing the hull of AE-solution sets where parameters involved in the matrix are constrained to be existentially quantified. Such AE-solution sets are called *right-quantified solution sets*. They have recently been shown to be of practical interest.

1. Introduction

Since the early times of interval analysis [10], one fundamental object of study has been the solution set of a linear system $Ax = b$, where A (resp. b) is known to range within an interval matrix $\mathbf{A}$ (resp. an interval vector $\mathbf{b}$).

The formal notation $\mathbf{A}x = \mathbf{b}$ is often used to refer to the family of linear systems generated by $\mathbf{A}$ and $\mathbf{b}$, and traditionally, by x satisfies $\mathbf{A}x = \mathbf{b}$, we mean that

$$(\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b). \tag{1.1}$$

Thus, the solution set under study can be formally defined as

$$\Sigma(\mathbf{A}, \mathbf{b}) := \{x \in \mathbb{R} \mid x \text{ satisfies } (1.1)\}.$$

It is referred to as the *united solution set*. Since the shape of this set can be very complicated, the purpose has been to approximate it outwardly by an interval vector (also called *a box*). The smaller this box, the more precise the result, and there is a unique smallest enclosing box called the *interval hull* of $\Sigma(\mathbf{A}, \mathbf{b})$, and usually denoted by $\square\Sigma(\mathbf{A}, \mathbf{b})$.

Although finding the interval hull is much simpler than finding the united solution set on its own, it is still an NP-hard problem [8]. We must actually content ourselves with computing an outer estimation of the interval hull, and common methods for this aim are Gauss-Seidel, Gauss elimination or Krawczyk [12].

However, in the case of midpoint inverse preconditioned systems, the problem of finding the interval hull is not NP-hard. This result was shown in 1992 independently by Hansen [7] and Bliek [2], who both gave an explicit formula involving polynomial-time computations for the subclass of matrices obtained after preconditioning. Rohn [15] showed that only the inversion of a single scalar matrix was required to compute this hull. Further, Ning, Kearfott [13] and Neumaier [11] extended the applicability scope of this method to H-matrices.

As a preconditioning step is usually necessary for the other methods mentioned before to behave well [7], [12], the Hansen-Bliek formula is a valuable alternative.

In this paper we extend the Hansen-Bliek method to a more general situation, where the semantics of intervals is enriched.

Indeed, the model described by (1.1) turns out not to be adequate for many practical problems with uncertainties in the parameters [16], though they are linear by nature. The reason is the lack of flexibility in the quantifiers associated with intervals. Indeed, universal quantifiers are more appropriate to represent a range of values we need to *control*. For instance, one may rather look for all x such that

$$(\forall A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b) \tag{1.2}$$

or

$$(\forall b \in \mathbf{b})(\exists A \in \mathbf{A})(Ax = b). \tag{1.3}$$

The set of points satisfying (1.2) is called the *tolerable solution set*, while the set of points satisfying (1.3) is called the *controllable solution set*. More complicated cases may also be of interest: a specific quantifier can be associated independently to each entry of **A** and **b**. However, in all these situations, universally quantified parameters precede the existentially quantified ones in the resulting formulae. This constraint on the order of quantifiers is summed up by "AE" (**A**ll-**E**xists). In this context, the united solution set $\Sigma(\mathbf{A}, \mathbf{b})$ defined before becomes a particular case. It is rather denoted by $\Sigma_{\exists,\exists}(\mathbf{A}, \mathbf{b})$.

So the quantified solution sets we consider here are called *AE solution sets*. They have been thoroughly studied since the 90's [5], [9], [16], [17], and the problem of finding the interval hull (i.e., the optimal outer box) lies in the same terms as for the united solution set.

In this paper we extend the Hansen-Bliek method to compute the interval hull for a preconditioned system $\mathbf{A}x = \mathbf{b}$, with **b** arbitrarily quantified (**A** remains existentially quantified). Such AE-solution sets will be called *right-quantified solution sets*. Hence, the controllable solution set is a canonic instance of the AE-solution sets we can handle. Figure 1 depicts an example of a controllable solution set.

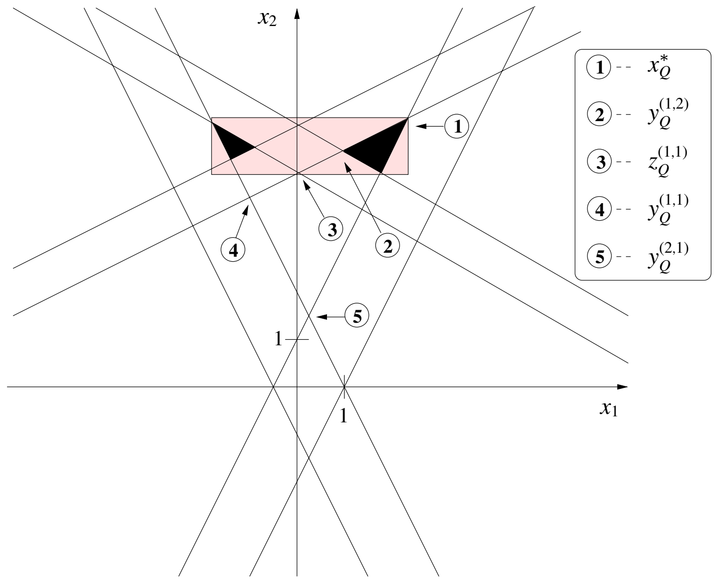

Figure 1. The controllable solution set of Example 1.1 (in black), and its hull (in gray). Annotations will illustrate our discussion in Sections 4 and 5: Points 1 to 5 are example of *candidate points*. The subscript Q designates the top right orthant ($x_1 \geq 0$, $x_2 \geq 0$).

EXAMPLE 1.1.

$$\begin{pmatrix} [1,1] & [-0.5,0.5] \\ [-0.5,0.5] & [1,1] \end{pmatrix} x = \begin{pmatrix} [-0.5,1] \\ [4.5,5.5] \end{pmatrix}.$$

It is remarkable fact that a right-quantified solution set may be disconnected, as shown on Figure 1.

The theory of AE-solution sets is out of the scope of this paper, and the reader need not be familiar with it. He/she simply has to admit two results. First, the right-quantified solution set of an interval linear problem has an equivalent expression in pure scalar arithmetic (see Proposition 3.1). Second, he/she must know that preconditioning is also compatible with AE-solution sets (i.e., it does not lose any solution). This process still makes **A** close to the identity matrix, and transforms a right-quantified solution set into another (wider) right-quantified solution set.

In a word, we propose here a Hansen-Bliek-like method for computing the interval hull of any right-quantified solution set of $\mathbf{A}x = \mathbf{b}$, with $\mathbf{A}$ being close to the identity, i.e., both centered around the identity and satisfying $\rho(\text{rad}\,(\mathbf{A})) < 1$ (ρ is the spectral radius).

2. Main Result

Let us start with some conventions.

If $\mathbf{x}$ denotes an interval, $\overline{\mathbf{x}}$ its upper bound and $\underline{\mathbf{x}}$ its lower bound, we define $\text{mid}\,\mathbf{x} := (\underline{\mathbf{x}} + \overline{\mathbf{x}}) / 2$ and $\text{rad}\,\mathbf{x} := (\overline{\mathbf{x}} - \underline{\mathbf{x}}) / 2$. These definitions are extended component-wise to vectors and matrices of intervals.

Let $\mathbf{A}x = \mathbf{b}$ be an interval linear problem with $\text{mid}\,\mathbf{A} = I$, and $\rho(\text{rad}\,(\mathbf{A})) < 1$. We will denote Σ the right-quantified solution set of $\mathbf{A}x = \mathbf{b}$, with $\mathbf{A}$ being existentially quantified everywhere and $\mathbf{b}$ arbitrarily quantified. Hence, if we assume (with no loss of generality) that the s first components of $\mathbf{b}$ are universally quantified, and the $(n - s)$ other components not, we have:

$$\Sigma := \{x \in \mathbb{R}^n \mid (\forall b_1 \in \mathbf{b}_1) \ldots (\forall b_s \in \mathbf{b}_s) \; (\exists A \in \mathbf{A})(\exists b_{s+1} \in \mathbf{b}_{s+1}) \ldots (\exists b_n \in \mathbf{b}_n) \; Ax = (b_1, \ldots, b_n)^T\}.$$

Note that we constrain universal quantifiers to precede the existential ones in the formula. Let us fix now quantifiers for $\mathbf{b}_1, \ldots, \mathbf{b}_n$ once for all.

We denote $\Delta := \text{rad}\,\mathbf{A}$, $b := \text{mid}\,\mathbf{b}$ and δ the vector such that

$$\forall i \in [1..n] \qquad \delta_i := \begin{cases} \text{rad}\,\mathbf{b}_i & \text{if } \mathbf{b}_i \text{ is existentially quantified,} \\ -\text{rad}\,\mathbf{b}_i & \text{if } \mathbf{b}_i \text{ is universally quantified.} \end{cases}$$

Now, as $\rho(\Delta) < 1$, the scalar matrix $I - \Delta$ is an M-matrix [12]. Therefore, as a classical result, $I - \Delta$ is regular and inverse positive. We will denote by M the inverse of $(I - \Delta)$, and by m_{ij} the entries of M. Hence, $\forall (i, j)$, $m_{ij} \geq 0$.

In the following, we borrow notation from Rohn [15].

Let us fix $k \in [1..n]$, and consider

$$\begin{aligned} x^* &:= M(|b| + \delta), \\ \tilde{x}_k &:= \begin{cases} x_k^* & \text{if } b_k \geq 0, \\ x_k^* + 2m_{kk}b_k & \text{otherwise,} \end{cases} \\ \underset{\sim}{x}_k &:= \begin{cases} -x_k^* & \text{if } b_k \leq 0, \\ -x_k^* + 2m_{kk}b_k & \text{otherwise,} \end{cases} \\ \nu_k &= \frac{1}{2m_{kk} - 1}. \end{aligned}$$

Define also $\hat{x}_k$ in the following way:

$$\begin{aligned} &\text{If } (\forall i \neq k \;\; m_{ik} = 0) && \text{then} \quad \hat{x}_k = -\infty, \\ & && \text{else} \quad \hat{x}_k = \max\{\theta_{ki}, i \neq k \text{ and } m_{ik} \neq 0\} \end{aligned}$$

with

$$\theta_{ki} := \alpha_{ki}|b_i| + \sum_{j=1, j \neq k}^{n} (m_{kj} - m_{ij}\alpha_{ki})(|b_j| + \delta_j) \qquad \text{with} \qquad \alpha_{ki} = m_{kk} / m_{ik}.$$

Next, compute $\underline{x}_k$ and $\overline{x}_k$ with the following scheme:

$\texttt{if}\ (\tilde{x}_k \geq \max\{0, v_k \underset{\sim}{x}_k, \hat{x}_k\})\ \texttt{then}\ \overline{x}_k = \tilde{x}_k$
$\texttt{else}\ \overline{x}_k = \min\{v_k \tilde{x}_k, -\hat{x}_k\}$
$\texttt{if}\ (\underset{\sim}{x}_k \leq \min\{0, v_k \tilde{x}_k, -\hat{x}_k\})\ \texttt{then}\ \underline{x}_k = \underset{\sim}{x}_k$
$\texttt{else}\ \underline{x}_k = \max\{v_k \underset{\sim}{x}_k, \hat{x}_k\}$

Then, we will prove the following result:

- Σ is empty iff for some $k \in [1..n]$ we have* $\underline{x}_k > \overline{x}_k$.
- If Σ is not empty, then for all $k \in [1..n]$, the kth component of $\Box\Sigma$ is $[\underline{x}_k, \overline{x}_k]$.

We will split the proof into two parts. In the first part (Section 4), we assume that some solution points exist with $x_k \geq 0$ and some others with $x_k \leq 0$. We compute the least upper bound of $|x_k|$ with $(x \in \Sigma) \wedge (x_k \geq 0)$, and with $(x \in \Sigma) \wedge (x_k \leq 0)$. We show that the former equals to $\tilde{x}_k$ and the latter to $-\underset{\sim}{x}_k$ (Corollaries 4.1 and 4.2). Then, $[\underset{\sim}{x}_k, \tilde{x}_k]$ is the projection of $\Box\Sigma$ on the kth coordinate. These bounds are valid under our assumption that x_k can be positive and negative in Σ. This can be checked, thanks to Proposition 6.1 and Corollary 6.1 in Section 6.

In the second part (Section 5), we assume that the kth coordinates of solution points are either all positive or all negative. Therefore, we need to compute the lower bound of $|x_k|$. With $x_k \geq 0$, this bound can be either $v_k \underset{\sim}{x}_k$ or $\hat{x}_k$, and with $x_k \leq 0$, it can be either $v_k \tilde{x}_k$ or $-\hat{x}_k$ (see Proposition 5.4 and Corollary 5.3).

3. Preliminary Results

It is well-known that the united solution set of an interval linear system $\mathbf{A}x = \mathbf{b}$ can be equivalently written as a set of scalar linear inequalities, with respect to the absolute vector $|x|$. This observation originates from Oettli & Prager [14]. It has been extended to AE-solution sets by Shary [16]. If we specialize his result to our case (a right-quantified system with mid $(\mathbf{A}) = I$), it can be written in our notation:

PROPOSITION 3.1.

$$x \in \Sigma \iff |x - b| \leq \Delta|x| + \delta.$$

It is worth stressing that δ is not necessarily positive, contrary to the Oettli-Prager formula.

By fixing the sign of each component of x, the absolute values can be dropped. Let us split $\mathbb{R}^n$ alongside each axis into a set of 2^n subspaces, called *orthants*, where all the components have constant sign.

---

* $\underline{x}_k > \overline{x}_k$ occurs iff both $\underline{x}_k$ and $\overline{x}_k$ are obtained by the "else" statements.

DEFINITION 3.1. Let Q be a diagonal matrix such that $\forall i$, $(1 \leq i \leq n)$ $|Q_{ii}| = 1$. The set $\{x \mid Qx \geq 0\}$ is called an **orthant** of $\mathbb{R}^n$.

To lighten notations, we will not dissociate an orthant from its characteristic matrix Q. Hence, we will talk about an *orthant* Q (Q being a matrix), and sometimes write $x \in Q$ instead of $Qx \geq 0$ to emphasize the membership nature of such condition.

Moreover, notice that $x \in Q \iff Qx = |x|$. Therefore, as a convenient convention, we shall write $|x|$ instead of Qx, as soon as x is known to be in Q.

PROPOSITION 3.2. *Let Q be an orthant*

$$x \in \Sigma \cap Q \iff \begin{cases} (I - \Delta Q)x \leq b + \delta, & (a) \\ (I + \Delta Q)x \geq b - \delta, & (b) \\ Qx \geq 0. & (c) \end{cases} \tag{3.1}$$

Proof.

$$x \in \Sigma \cap Q \iff \begin{cases} |x - b| \leq \Delta|x| + \delta, \\ Qx \geq 0 \end{cases} \iff \begin{cases} x - \Delta|x| \leq b + \delta, \\ -x - \Delta|x| \leq -b + \delta, \\ Qx \geq 0 \end{cases}$$

which is equivalent to (3.1). □

Thanks to Proposition 3.2, we fall back into the classical situation of optimizing bounds of variables constrained by linear inequalities, for which linear programming techniques are well suited [1], [12]. However, the number of orthants grows exponentially with the dimension of the problem, so that calls to the linear programming solver become prohibitive.

In the *classical* case (the case with only existential quantifiers), Hansen [7] and Bliek [2] found that a combinatorial sweep of the orthants was useless with preconditioned systems, and they both give an explicit formula for $\Box\Sigma$. The next proposition is just a reformulation of Proposition 3.2 with inequalities gathered in a different way. The Hansen-Bliek method is based upon this technical result, and so is ours.

PROPOSITION 3.3. *Assume $A = I$, and let Q be an orthant*

$$x \in \Sigma \cap Q \iff \begin{cases} (I - \Delta)Qx \leq Qb + \delta, & (a) \\ (I + \Delta)Qx \geq Qb - \delta, & (b) \\ Qx \geq 0. & (c) \end{cases} \tag{3.2}$$

Proof. Let $i \in [1..n]$. If $Q_{ii} = 1$ then (3.1.a) implies $[(I - \Delta Q)x]_i \leq [b + \delta]_i$, i.e. $x_i - \sum_{j=1}^{n} (\Delta_{ij} Q_{jj}) x_j \leq b_i + \delta_i$.

This can be rewritten $Q_{ii}x_i - \sum_{j=1}^{n} \Delta_{ij}(Q_{jj}x_j) \leq Q_{ii}b_i + \delta_i$, or, in a more compact way, $[(I - \Delta)Qx]_i \leq [Qb + \delta]_i$.

If $Q_{ii} = -1$, then (3.1.b) implies $[(I + \Delta Q)x]_i \geq [b - \delta]_i$, i.e., $x_i + \sum_{j=1}^{n}(\Delta_{ij}Q_{jj})x_j \geq b_i - \delta_i$. As before, we get $Q_{ii}x_i - \sum_{j=1}^{n} \Delta_{ij}(Q_{jj}x_j) \leq Q_{ii}b_i + \delta_i$, and, again, $[(I - \Delta)Qx]_i \leq [Qb + \delta]_i$. Since the latter relation is true for any i, (3.2.a) holds.

Repeating the same implications while swapping roles of (3.1.a) and (3.1.b), one obtains also (3.2.b). Furthermore, each inequality of (3.1.a) and (3.1.b) is used once and only once under an equivalent form, in the system gathering (3.2.a) and (3.2.b). Therefore, conjunction of (3.2.a), (3.2.b) and (3.2.c) is also a characterization of $\Sigma \cap Q$. □

4. Upper Bounds of $|x_k|$

Remember that k is fixed. We denote by $\Sigma^{(k+)}$ the set $\{x \in \mathbb{R}^n \mid x \in \Sigma \wedge x_k \geq 0\}$. Similarly, $\Sigma^{(k-)}$ stands for $\{x \in \mathbb{R}^n \mid x \in \Sigma \wedge x_k \leq 0\}$. In this section, we give an expression for the upper bound of $|x_k|$ with $x \in \Sigma^{(k+)}$. The upper bound of $|x_k|$ with $x \in \Sigma^{(k-)}$ will immediately follow by symmetry.

The cornerstone of Hansen and Bliek's approach is the capability to break the combinatorial sweep we highlighted in the previous section, by locating a specific orthant on which the global maximum is reached.

We proceed in two steps: First, we arbitrarily fix an orthant Q, and formally identify the point that maximizes $(Qx)_k$ *locally*, i.e., for $x \in \Sigma \cap Q$. There is only one possible expression (noted x_Q) for this maximum. As all the orthants share this expression for their maximum, then, it is easy to isolate an orthant (noted $Q^{(k+)}$) that has an overall maximum.

Although the results here are exactly the same as in the classical case, the proofs proposed are different because we have to take into account quantifiers of **b**.*

For a given orthant Q, define $x_Q := QM(Qb + \delta)$ so that $(I - \Delta)Qx_Q = (Qb + \delta)$. The next lemma shows that $|x_Q|$ maximizes $|x|$ in $\Sigma \cap Q$.

LEMMA 4.1 (Local maximum). *The following three conditions are equivalent:*

(i) $\Sigma \cap Q \neq \emptyset$,

(ii) $x_Q \in \Sigma \cap Q$,

(iii) $Qx_Q \geq 0 \wedge \Delta Qx_Q \geq -\delta$.

Furthermore, if one of these conditions holds, we have:

(iv) $|x_Q| = \max_{x \in \Sigma \cap Q} |x|$.

* For instance, as contrary to the existentially quantified case, we cannot pick some matrix $\tilde{A} \in \mathbf{A}$ and some vector $\tilde{b} \in \mathbf{b}$ and check that $\tilde{A}x = \tilde{b}$ to prove the membership of x in Σ.

Proof. Obviously (ii) implies (i). Let us prove (i) implies (ii). First, x_Q obviously satisfies (3.2.a). Now, by (i) there exists $x \in \Sigma \cap Q$. So by (3.2.a) $(I-\Delta)Qx \leq Qb+\delta$. Multiplying the latter inequality by the positive matrix M gives $Qx \leq Qx_Q$. By (3.2.c) $0 \leq Qx$ and therefore x_Q also satisfies (3.2.c). Also $Qx \leq Qx_Q$ gives $(I + \Delta)Qx \leq (I + \Delta)Qx_Q$ because $I + \Delta \geq 0$. Therefore as x satisfies (3.2.b) (i.e. $Qb - \delta \leq (I + \Delta)Qx$), x_Q also satisfies (3.2.b) which completes the proof of (i) implies (ii). Let us now prove that (ii) is equivalent to (iii). As x_Q satisfies (3.2.a) by construction, and as $Qx_Q \geq 0$ is present in both (ii) (through (3.2.c)) and (iii), we just have to prove that x_Q satisfies (3.2.b) (i.e. $(I + \Delta)Qx_Q \geq Qb - \delta$)) if and only if $\Delta Qx_Q \geq -\delta$. Just subtract the equality $(I - \Delta)Qx_Q = (Qb + \delta)$ (which defines x_Q) to $(I + \Delta)Qx_Q \geq Qb - \delta$ to obtain the equivalent inequality $2\Delta Qx_Q \geq -2\delta$. Therefore $\Delta Qx_Q \geq -\delta$ is equivalent to (3.2.b). To prove that (i) $\Longrightarrow$ (iv), let again x be in $\Sigma \cap Q$. We saw that (i) implies $Qx \leq Qx_Q$, i.e. $|x| \leq Qx_Q$. But $x_Q \in Q$, since (i) implies (ii). Therefore $|x_Q| = \max_{x\in\Sigma \cap Q} |x|$. □

We now compute the maximum of $|x_k|$ (or simply x_k) for x ranging over $\Sigma^{(k+)}$. Define Q^* and $Q^{(k+)}$ as the following diagonal matrices:*

$$Q^*_{ii} := \text{sign}(b_i); \qquad Q^{(k+)}_{ii} := \begin{cases} 1 & \text{if } i = k, \\ \text{sign}(b_i) & \text{otherwise.} \end{cases}$$

PROPOSITION 4.1 (Global maximum).

$$\Sigma^{(k+)} \neq \emptyset \implies \begin{cases} \Sigma \cap Q^{(k+)} \neq \emptyset, \\ (x_{Q^{(k+)}})_k = \max\{x_k \mid x \in \Sigma^{(k+)}\}. \end{cases}$$

Proof. Consider any $x \in \Sigma^{(k+)}$ and denote its orthant by Q. Therefore $Q_{kk} = 1$ and $Qb \leq Q^{(k+)}b$. Therefore, as M is positive, we have

$$M(Qb + \delta) \leq M(Q^{(k+)}b + \delta).$$

By definition of x_Q and $x_{Q^{(k+)}}$ this means $Qx_Q \leq Q^{(k+)}x_{Q^{(k+)}}$. Now, using ($i$) $\Longrightarrow$ (iv) in Lemma 4.1, we have $Qx \leq Qx_Q$. Therefore we have $x_k \leq (x_{Q^{(k+)}})_k$ (because $Q_{kk} = Q^{(k+)}_{kk} = 1$).

This inequality holding for any $x \in \Sigma^{(k+)}$, we have

$$(x_{Q^{(k+)}})_k \geq \max\{x_k | x \in \Sigma^{(k+)}\}.$$

Now, by the part "(i) implies (iii)" of Lemma 4.1, $\Sigma \cap Q \neq \emptyset$ implies both $0 \leq Qx_Q$ and $-\delta \leq \Delta Qx_Q$. This gives both $0 \leq Q^{(k+)}x_{Q^{(k+)}}$ and $-\delta \leq \Delta Q^{(k+)}x_{Q^{(k+)}}$ (because

* We define sign(0) := 1 (but −1 would make no difference).

$Qx_Q \leq Q^{(k+)}x_{Q^{(k+)}}$ and $\Delta \geq 0$). Applying the part "(*iii*) implies (*ii*)" of Lemma 4.1, we have $x_{Q^{(k+)}} \in \Sigma \cap Q^{(k+)} \subseteq \Sigma^{(k+)}$. Therefore $\Sigma \cap Q^{(k+)} \neq \emptyset$ and

$$(x_{Q^{(k+)}})_k \leq \max\{x_k \mid x \in \Sigma^{(k+)}\},$$

which completes the proof. □

Remark 4.1. By the contrapositive of Proposition 4.1, if $(x_{Q^{(k+)}})_k < 0$ then $\Sigma^{(k+)}$ has to be empty.

Proposition 4.1 shows that computing the maximum of x_k in $\Sigma^{(k+)}$ only requires the computation of the kth component of $Q^{(k+)}x_{Q^{(k+)}}$. Therefore, one does not need to compute all the components of such vectors, and the next corollary shows that these n reals can be computed directly from a single vector x^*, as in the classical Hansen-Bliek algorithm.

COROLLARY 4.1. *Let $\tilde{x}$ be defined from x^* (see Section 2).*

$$(x_{Q^{(k+)}})_k = \tilde{x}_k, \qquad (i)$$
$$\Sigma^{(k+)} \neq \emptyset \implies \tilde{x}_k = \max\{x_k \mid x \in \Sigma^{(k+)}\}. \qquad (ii)$$

Proof. Thanks to Proposition 4.1, we just have to prove (*i*).

Clearly, $Q^*b = |b|$ and $x^* = Q^*x_{Q^*} = M(Q^*b + \delta)$. Now, if $\text{sign}(b_k) = 1$, we have $Q^{(k+)}_{kk} = Q^*_{kk}$, i.e., $Q^{(k+)} = Q^*$, and therefore $(x_{Q^{(k+)}})_k = x^*_k$. Otherwise, $Q^{(k+)} = Q^* + 2e_ke_k^T$ (e_k is the kth column of I) and $|x_{Q^{(k+)}}| = M(Q^{(k+)}b + \delta) = M((Q^* + 2e_ke_k^T)b + \delta)$ so $(x_{Q^{(k+)}})_k = x^*_k + 2m_{kk}b_k$. Finally, $(x_{Q^{(k+)}})_k = \tilde{x}_k$. □

As a direct consequence, we can compute the lower bound of x_k, with $x \in \Sigma^{(k-)}$.

COROLLARY 4.2.

$$\Sigma^{(k-)} \neq \emptyset \implies \underset{\sim}{x}_k = \min\{x_k \mid x \in \Sigma^{(k-)}\}.$$

Proof. Just notice that $\mathbf{A}x = \mathbf{b}$ is equivalent to $\mathbf{A}(-x) = (-\mathbf{b})$ and apply Corollary 4.1 to this new system. □

5. Lower Bound of $|x_k|$

The same idea is applied for the minimum of $|x_k|$ as in the previous section, except that the situation gets much worse since a local minimum can have up to $2n$ different expressions, instead of one. Each of these expressions will be called a (formal) *candidate*.

In the classical case, there is still a single candidate for minimization so the situation becomes now really different with quantifiers. This means that our contribution lies, for the most part, in this section.

We shall prove first in Section 5.1 that the local minimum of $|x_k|$ (the minimum in a given orthant Q) is necessarily reached at a point that belongs to a specific line called the "support line" (of Q). Next, in Section 5.2, we show that only $2n$ points of this line can be the *candidates* for miminization. We give an explicit formula for each of them, dealing with two distinct situations (the *inner candidates*, and the *border candidates*). Further, in Section 5.3 we show that the minimizer is precisely the one that maximizes the kth component (in absolute value) among all the candidates. This finally provides the local minimum.

In Section 5.4 we prove, as in the case of the upper bound, that the global minimum (i.e., the minimum of all the orthant minima) is met on a formally identified orthant, which is still by the way $Q^{(k+)}$.

5.1. The Support Line

The support line of an orthant is defined as the intersection of the $(n-1)$ hyperplanes obtained by dropping the kth row in the system (3.2.a).

To drop some rows (or columns) of a matrix X, we resort to an enhanced system of indices appropriate for entries and submatrices at the same time. If X is a matrix, as usual, X_{ij} represents an entry if both i and j are integers. But i may also have special values such as ":", "$\neq k$", or "$< k$" to denote respectively all the rows, all the rows but k, and the $(k-1)$ first rows. Likewise, j may take these special values to denote a subset of columns. In this way, $X_{\neq k,:}$ becomes the matrix X with the kth row dropped, and all the columns.

The same system of indices is applied to vectors. We shall need further the following property of submatrices:

LEMMA 5.1.

$(\forall k \in [1..n])\ (I-\Delta)_{\neq k,\neq k}$ *is an M-matrix.*

Proof. Two well-known properties of M-matrices underlie the proof:

$$A \text{ is an M-matrix} \iff (\exists u > 0)\, Au > 0, \tag{5.1}$$

$$A \text{ is an M-matrix} \implies (\forall i \neq k)\, A_{ik} \leq 0. \tag{5.2}$$

As $I-\Delta$ is an M-matrix, by (5.1), $\exists u > 0$ such that $(I-\Delta)u > 0$. By (5.2), for such u, $(I-\Delta)_{\neq k,\neq k}(u_{\neq k}) > 0$ and by (5.1) again $(I-\Delta)_{\neq k,\neq k}$ is an M-matrix. □

Let us get now to the heart of the matter.

DEFINITION 5.1. Let Q be an orthant. We call the "support" line and denote by $\mathcal{L}_Q^{(k)}$ the following set:

$$\mathcal{L}_Q^{(k)} := \{x \mid (I-\Delta)_{\neq k,:}Qx = (Qb+\delta)_{\neq k}\}.$$

To prove that the minimum of $|x_k|$ is reached on $\mathcal{L}_Q^{(k)}$ (see Proposition 5.1 below), we first show that every point $\tilde{x}$ of the solution set can be projected on

$\mathcal{L}_Q^{(k)}$ orthogonally to the axis x_k, while remaining in the solution set. Then, we will simply have to apply this property to *any* point that minimizes $|x_k|$.

LEMMA 5.2. *Let Q be an orthant that intersects Σ*

$$(\forall \check{x} \in \Sigma \cap Q)(\exists x \in \Sigma \cap Q \cap \mathcal{L}_Q^{(k)})(|x_k| = |\check{x}_k|).$$

Proof. In this proof, we need to change the kth row of a matrix. Notations can become quickly burdensome, so we assume that $k = 1$ to lighten them. There is no loss of generality since this assumption means only a reordering of variables.

Let $\check{x}$ be in $\Sigma \cap Q$. We shall build an appropriate point x. First, put

$$\Delta' := \begin{pmatrix} 0 \\ (I - \Delta)_{\neq 1,:} \end{pmatrix}.$$

A well-known property of spectral radius states that

$$0 \leq \Delta' \leq \Delta \Longrightarrow \rho(\Delta') \leq \rho(\Delta).$$

Therefore, $\rho(\Delta') < 1$, and $I - \Delta'$ is still an M-matrix. We can consider the unique x such that

$$(I - \Delta')Qx = \begin{pmatrix} |\check{x}_1| \\ (Qb + \delta)_{\neq 1} \end{pmatrix}.$$

In this way, x already verifies $(Qx)_1 = |\check{x}_1|$ and $x \in \mathcal{L}_Q^{(1)}$. We need to prove now that $x \in \Sigma \cap Q$. For that end, we first prove that $|\check{x}| \leq |x|$.

Let us denote by M' the inverse of $(I - \Delta)_{\neq 1, \neq 1}$. By Lemma 5.1, M' is also positive.

We have

$$\begin{cases} (Qx)_1 &= |\check{x}|_1, \\ (Qx)_{\neq 1} &= M'(Qb + \delta)_{\neq 1} + M'\Delta_{\neq 1, 1}|\check{x}|_1. \end{cases}$$

Since $\check{x} \in \Sigma \cap Q$, it satisfies (3.2.a), i.e., $(I - \Delta)|\check{x}| \leq Qb + \delta$. And in particular,

$$(I - \Delta)_{\neq 1,:}|\check{x}| \leq (Qb + \delta)_{\neq 1}.$$

By moving $\check{x}_1$ on the right side, we get $(I-\Delta)_{\neq 1, \neq 1}|\check{x}|_{\neq 1} \leq (Qb+\delta)_{\neq 1}+\Delta_{\neq 1, 1}|\check{x}|_1$.

And by multiplying both sides by M': $|\check{x}|_{\neq 1} \leq M'(Qb + \delta)_{\neq 1} + M'\Delta_{\neq 1, 1}|\check{x}|_1$.

We recognize $(Qx)_{\neq 1}$ on the right side, so that $|\check{x}|_{\neq 1} \leq (Qx)_{\neq 1}$ and finally $|\check{x}| \leq Qx$. We have seen that $x \in \mathcal{L}_Q^{(1)}$, so what is left to be proven is $x \in \Sigma \cap Q$. To that end, we use the characterization given by system (3.2).

$\rightarrow$ x satisfies (3.2.a). First, $[(I - \Delta)|x|]_1 = |\check{x}|_1 - \sum_{i=1}^{n} \Delta_{1i}|x_i|$.

But $|\check{x}|_1 - \sum_{i=1}^{n} \Delta_{1i}|x_i| \leq |\check{x}|_1 - \sum_{i=1}^{n} \Delta_{1i}|\check{x}|_i = [(I - \Delta)|\check{x}|]_1 \leq [Qb + \delta]_1$.

Second, by definition of x, $[(I-\Delta)|x|]_{\neq 1} = (Qb+\delta)_{\neq 1}$. By gathering the two previous relations we obtain the desired inequalities $(I-\Delta)|x| \leq Qb+\delta$.

→ x satisfies (3.2.b). Indeed, $\check{x} \in \Sigma \cap Q$ implies that $(I+\Delta)|\check{x}| \geq Qb-\delta$, which in turn means that $(I+\Delta)|x| \geq Qb-\delta$.

→ x satisfies (3.2.c) because $Qx \geq |\check{x}| \geq 0$. □

PROPOSITION 5.1. *Let Q be an orthant that intersects Σ*

$$(\exists w \in \Sigma \cap Q \cap \mathcal{L}_Q^{(k)})(\forall x \in \Sigma \cap Q)(|w_k| \leq |x_k|).$$

Proof. Since $\Sigma \cap Q$ is a compact set, the minimal value of $|x_k|$, with $x \in \Sigma \cap Q$, is reached at a point $\check{x}$ of the set. By applying Lemma 5.2 to $\check{x}$, we get a point w of the line $\mathcal{L}_Q^{(k)}$ that also minimizes the kth component (in absolute value) within $\Sigma \cap Q$. □

At last, here is a key property of $\mathcal{L}_Q^{(k)}$ related to its orientation, that we will refer to in Section 5.3. Roughly speaking, it states that if we follow the trajectory of $\mathcal{L}_Q^{(k)}$ in the orthant Q, then all the components increase simultaneously (in absolute value) as $|x_k|$ increases.

LEMMA 5.3. *Let Q be an orthant, $x \in \mathcal{L}_Q^{(k)}$ and $x' \in \mathcal{L}_Q^{(k)}$.*

$$(Qx)_k \leq (Qx')_k \iff \forall i \neq k\ (Qx)_i \leq (Qx')_i.$$

Proof.

$$\begin{aligned} x \in \mathcal{L}_Q^{(k)} &\iff (I-\Delta)_{\neq k,:}Qx = (Qb+\delta)_{\neq k} \\ &\iff (I-\Delta)_{\neq k,\neq k}(Qx)_{\neq k} = (Qb+\delta)_{\neq k} + (Qx)_k\Delta_{k,:} \\ &\iff \forall i \neq k,\ \ (Qx)_i = (M'(Qb+\delta)_{\neq k})_i + (Qx)_k M'\Delta_{ki}, \end{aligned}$$

where M' denotes the inverse of $(I-\Delta)_{\neq k,\neq k}$, which is positive by Lemma 5.1. The last relation holds also by substituting x' for x. Now, just remember that $\Delta_{ki} \geq 0$. □

5.2. The Candidate Points

Remember that we are looking for the minimum of $|x_k|$ in the simplex defined by system (3.2). A well-known property of linear programming states that a point optimizing a linear criterion always lies at a vertex [3].

So far, we know that the point of $\Sigma^{(k+)} \cap Q$ minimizing $|x_k|$ belongs to the line $\mathcal{L}_Q^{(k)}$. This means that it is found at the intersection of $\mathcal{L}_Q^{(k)}$ and one of the $(2n+1)$ remaining hyperplanes of $\Sigma^{(k+)} \cap Q$ (among those given by system (3.2)).

Yet, we can get rid of the hyperplane whose equation is $(I-\Delta)_{k,:}Qx = (Qb+\delta)_k$, since the intersection of the latter and $\mathcal{L}_Q^{(k)}$ is precisely x_Q, the point maximizing $|x_k|$ (see Section 4).

The $2n$ other points are called *candidates*. The point $y_Q^{(k,i)}$ obtained by considering the hyperplane defined by the ith row of (3.2.b) is called an *inner candidate*, whereas the point $z_Q^{(k,i)}$ obtained by the ith row of (3.2.c) (i.e., $(Qx)_i = 0$) is called a *border candidate*.

DEFINITION 5.2 (Inner candidates). Let Q be an orthant, $i \in [1..n]$, and

$$\Lambda_i := \begin{pmatrix} (I-\Delta)_{<k,:} \\ (I+\Delta)_{i,:} \\ (I-\Delta)_{>k,:} \end{pmatrix}.$$

If Λ_i is regular, we call **inner candidate** and denote by $y_Q^{(k,i)}$ the point such that

$$Qy_Q^{(k,i)} := \Lambda_i^{-1} \begin{pmatrix} (Qb+\delta)_{<k} \\ (Qb-\delta)_i \\ (Qb+\delta)_{>k} \end{pmatrix}. \tag{5.3}$$

Hence, $y_Q^{(k,i)}$ lies at the intersection of $\mathcal{L}_Q^{(k)}$ and the hyperplane

$$\{x \mid (I+\Delta)_{i,:}Qx = (Qb-\delta)_i\}.$$

We define similarly the *border candidate*.

DEFINITION 5.3 (Border candidates). Let Q be an orthant, $i \in [1..n]$, and

$$\Gamma_i := \begin{pmatrix} (I-\Delta)_{<k,:} \\ I_{i,:} \\ (I-\Delta)_{>k,:} \end{pmatrix}.$$

If Γ_i is regular, we call **border candidate** and denote by $z_Q^{(k,i)}$ the point such that

$$Qz_Q^{(k,i)} := \Gamma_i^{-1} \begin{pmatrix} (Qb+\delta)_{<k} \\ 0 \\ (Qb+\delta)_{>k} \end{pmatrix}. \tag{5.4}$$

Again, $z_Q^{(k,i)}$ lies at the intersection of $\mathcal{L}_Q^{(k)}$ and the hyperplane $\{x \mid (Qx)_i = 0\}$.

Of course, if a hyperplane is parallel to $\mathcal{L}_Q^{(k)}$, the corresponding candidate does not exist. This does not change the fact that the minimum is still met at some other candidate point. Notice that the candidates $y_Q^{(k,k)}$ and $z_Q^{(k,k)}$ always exist since Λ_k and Γ_k are regular. It means that in the "worst" case, the minimum is still found among them.

We will show in appendix the following property (see Proposition 8.1):

$$\forall i \neq k \quad \Lambda_i \text{ is regular} \iff \Gamma_i \text{ is regular} \iff m_{ik} \neq 0.$$

5.2.1. *I e Ca did e*

No , he e io a o o: Fo a o ha Q, ca e de e i e he k h co o e of $Q^{(k,i)}_{Q}$? The e o o i io o ide a a e .

PROPOSITION 5.2. *Le Q be a h . We h e:*

$$(Q^{(k,k)}_{Q})_k = v_k \left[m_{kk}(Qb - \delta)_k - \sum_{j \neq k} m_{kj}(Qb + \delta)_j \right] \quad i\ h \quad v_k = \frac{1}{2m_{kk} - 1}.$$

A d f $i \neq k$, if $m_{ik} \neq 0$ he

$$(Q^{(k,i)}_{Q})_k = \alpha_{ki}(Qb)_i + \sum_{j \neq k}(m_{kj} - m_{ij}\alpha_{ki})(Qb + \delta)_j \quad i\ h \quad \alpha_{ki} = m_{kk} / m_{ik}.$$

P f. Fo he a e of c a i , i he fo o i g, D_i i a d fo $e_k e_i^T$, i.e., fo a ze o fi ed a i e ce fo $(D_i)_{ki} := 1$.

We ha e $\Lambda_i = (I - D_k)(I - \Delta) + D_i(I + \Delta) = (I - D_k - D_i)(I - \Delta) + 2D_i$.

A o iced befo e, Λ_k i eg a . A d if $i \neq k, m_{ik} \neq 0$ i ie b P o o i io 8.1, ha Λ_i i eg a . Th , e ca i e $\Lambda_i^{-1}\Lambda_i = I$, i.e., $\Lambda_i^{-1}\Lambda_i M = M$.

The $\Lambda_i^{-1}[(I - D_k - D_i)(I - \Delta) + 2D_i]M = M$. A d i ce $(I - \Delta)M = I$, e ge $\Lambda_i^{-1}[(I - D_k - D_i) + 2D_i M] = M$. Le e he bo λ fo he e ie of Λ^{-1}. Wi h ch e a io , o e ca ea i co e he k h o of Λ_i^{-1}.

Wi h $i = k$, e ge :

- $\lambda_{kk}(2m_{kk} - 1) = m_{kk} \Longrightarrow \lambda_{kk} = (m_{kk} / (2m_{kk} - 1)) = m_{kk}v_k$.
- $2m_{kj}\lambda_{kk} + \lambda_{kj} = m_{kj} \Longrightarrow \lambda_{kj} = m_{kj}(1 - 2\lambda_{kk}) = m_{kj}(1 - 2m_{kk}v_k) = -m_{kj}v_k$.

So ha ojec io of (5.3) o he k h coo di a e gi e :

$$\begin{aligned}(Q^{(k,k)}_{Q})_k &= \lambda_{kk}(Qb - \delta)_k + \sum_{j \neq k} \lambda_{kj}(Qb + \delta)_j \\ &= v_k \left[m_{kk}(Qb - \delta)_k - \sum_{j \neq k} m_{kj}(Qb + \delta)_j \right].\end{aligned}$$

Wi h $i \neq k$, e ge :

- $2\lambda_{kk}m_{ik} = m_{kk} \Longrightarrow \lambda_{kk} = 1 / 2(m_{kk} / m_{ik}) = 1 / 2\alpha_{ki}$.
- $2m_{ij}\lambda_{kk} + \lambda_{kj} = m_{kj} \Longrightarrow \lambda_{kj} = m_{kj} - (m_{ij}m_{kk}) / m_{ik} = m_{kj} - m_{ij}\alpha_{ki}$.
- $(2m_{ii} - 1)\lambda_{kk} + \lambda_{ki} = m_{ki} \Longrightarrow \lambda_{ki} = m_{ki} - (m_{ii}m_{kk}) / m_{ik} + 1 / 2(m_{kk} / m_{ik}) = m_{ki} - m_{ii}\alpha_{ki} + 1 / 2\alpha_{ki}$.

A d ea o i g a befo e ead o:

$$\begin{aligned}(Qy_Q^{(k,i)})_k &= \lambda_{kk}(Qb-\delta)_i + \lambda_{ki}(Qb+\delta)_i + \sum_{j\neq k, j\neq i}^{n} \lambda_{kj}(Qb+\delta)_j \\ &= (\lambda_{kk}+\lambda_{ki})(Qb)_i + (\lambda_{ki}-\lambda_{kk})\delta_i + \sum_{j\neq k, j\neq i}^{n} \lambda_{kj}(Qb+\delta)_j \\ &= \alpha_{ki}(Qb)_i + \sum_{j\neq k}^{n}(m_{kj}-m_{ij}\alpha_{ki})(Qb+\delta)_j. \qquad \square\end{aligned}$$

Remark 5.1. In particular we have:

$$(Qy_Q^{(k,i)})_k = (\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki})(Qb)_i + \sum_{\substack{j\neq k\\ j\neq i}}^{n} (m_{kj} - m_{ij}\alpha_{ki})(Qb)_j + \varphi(\delta),$$

where $\varphi(\delta)$ is an expression that does not depend on Q.

In the next corollary, we reintroduce x_Q, as defined in Section 4.

COROLLARY 5.1. *Let Q be an orthant. Let S be a diagonal matrix filled with* 1 *except for $S_{kk} := -1$ (this means that (SQ) is the symmetric orthant of Q with respect to $(x_k = 0)$). We have:*

$$(Qy_Q^{(k,k)})_k = -v_k\big((SQ)x_{(SQ)}\big)_k.$$

Proof. Simply compare the expression obtained in Proposition 5.2 with the definition of $x_{(SQ)}$ (Section 4). $\square$

COROLLARY 5.2. *Let Q be an orthant, $x \in \mathcal{L}_Q^{(k)}$ and $i \in [1..n]$ such that $m_{ik} \neq 0$.*

$$(Qx)_k \geq (Qy_Q^{(k,i)})_k \quad \Longrightarrow \quad (I+\Delta)_{i,:}Qx \geq (Qb-\delta)_i.$$

Proof. By Lemma 5.3, $(Qx)_k \geq (Qy_Q^{(k,i)})_k$ implies that $Qx \geq Qy_Q^{(k,i)}$. Combining this with $I+\Delta \geq 0$ and $(I+\Delta)_{i,:}Qy_Q^{(k,i)} = (Qb-\delta)_i$ gives the desired result. $\square$

5.2.2. *Border Candidates*

Explicit formulae for border candidates are obtained in the same way as inner candidates, and with simpler calculus. We will be more concise.

PROPOSITION 5.3. *Let Q be an orthant, $i \in [1..n]$. If $m_{ik} \neq 0$, then we have:*

$$(Qz_Q^{(k,i)})_k = (Qy_Q^{(k,i)})_k - \alpha_{ki}(Qb)_i.$$

Proof. We have $\Gamma_i = (I - D_k)(I - \Delta) + D_i$, so that $\Gamma_i^{-1}((I - D_k) + D_i M) = M$. Let us use γ as the symbol for the entries of Γ_i^{-1}. We deduce that $\gamma_{kk} = m_{kk} / m_{ik} = \alpha_{ki}$. And $\forall j \neq k$, $\gamma_{kk} m_{ij} + \gamma_{kj} = m_{kj}$, i.e., $\gamma_{kj} = m_{kj} - (m_{kk} m_{ij}) / m_{ik}$. Finally we have $(Qz_Q^{(k,i)})_k = \sum_{j \neq k} (m_{kj} - m_{ij}\alpha_{ki})(Qb + \delta)_j = (Qy_Q^{(k,i)})_k - \alpha_{ki}(Qb)_i$. □

5.3. Local Minima

Now that we have well characterized inner and border candidates, it will be more convenient to consider a unique set of (possibly $2n$) candidates. Define:

$$\forall i \in [1..n], \;\; m_{ik} \neq 0, \qquad \begin{cases} c_Q^{(k,i)} & := y_Q^{(k,i)}, \\ c_Q^{(k,n+i)} & := z_Q^{(k,i)}. \end{cases}$$

As noticed before, $y_Q^{(k,k)}$ and $z_Q^{(k,k)}$ always belong to this set.

We define also an appropriate set of indices:

$$\mathcal{I} := \bigcup_{i \in [1..n],\, m_{ik} \neq 0} \{i, (n+i)\}.$$

LEMMA 5.4 (Local minimum). *Let Q be an orthant.*

$$\Sigma \cap Q \neq \emptyset \quad \Longrightarrow \quad \inf_{x \in \Sigma \cap Q} |x_k| = \max_{i \in \mathcal{I}} (Qc_Q^{(k,i)})_k.$$

Proof. Let c be any candidate, and assume it exists $x \in \mathcal{L}_Q^{(k)}$ such that $(Qx)_k < (Qc)_k$. We will prove that $x \notin \Sigma \cap Q$. Then, if $\Sigma \cap Q \neq \emptyset$, $\inf_{x \in \Sigma \cap Q} |x_k|$ is obtained by maximizing $(Qc)_k$, where c ranges over all the candidate points. We prove now $x \notin \Sigma \cap Q$. First, $(Qx)_k < (Qc)_k$ implies $Qx \leq Qc$, by virtue of Lemma 5.3. Now, if c is a border candidate, there exists i such that $(Qc)_i = 0$ (see Definition 5.3). As $Qx \leq Qc$ implies $(Qx)_i \leq (Qc)_i$, then $(Qx)_i \leq 0$. But $(Qx)_i = 0$ would mean that $x = c$,[⋆] contradicting $(Qx)_k < (Qc)_k$. Therefore, $(Qx)_i < 0$ and $x \notin \Sigma \cap Q$. Similarly, if c is an inner candidate, there exists i such that $(I + \Delta)_{i,:} Qc = (Qb - \delta)_i$ (see Definition 5.2). $Qx \leq Qc$, implies $(I + \Delta)_{i,:} Qx \leq (I + \Delta)_{i,:} Qc$, therefore $(I + \Delta)_{i,:} Qx \leq (Qb - \delta)_i$, and the inequality is actually strict since $x \neq c$.[⋆⋆] Again, $x \notin \Sigma \cap Q$. □

5.4. Global Minima

In the previous section, we gave a way to compute the minimum of $|x_k|$ among all the candidates for a fixed orthant. Here, we do an orthogonal operation, i.e., we

⋆ $\mathcal{L}_Q^{(k)}$ is not parallel to $(x_i = 0)$ since $i \in \mathcal{I}$ means that Γ_i is regular.

⋆⋆ $\mathcal{L}_Q^{(k)}$ is not parallel to $((I + \Delta)_{i,:} Qx = (Qb - \delta)_i)$ since $i \in \mathcal{I}$ means that Λ_i is regular.

compute the minimum of a fixed candidate (defined by a given matrix Λ_i or Γ_i)*
among all the orthants.

Define

$$\begin{aligned} \mathcal{O}^+ &:= \{Q \text{ orthant of } \mathbb{R}^n \mid Q_{kk} = +1\}, \\ \mathcal{O}^- &:= \{Q \text{ orthant of } \mathbb{R}^n \mid Q_{kk} = -1\}. \end{aligned}$$

LEMMA 5.5 (Best orthant for candidates).

$$\forall i \in \mathcal{I} \qquad \min_{Q \in \mathcal{O}^+} (Qc_Q^{(k,\,i)})_k = (Q^{(k+)} c_{Q^{(k+)}}^{(k,\,i)})_k$$

More precisely, $\forall i \in [1, n]$ *such that* $i \neq k$ *and* $m_{ik} \neq 0$

$$\text{(i)} \qquad \min_{Q \in \mathcal{O}^+} (Qy_Q^{(k,\,i)})_k = \theta_{ki} \qquad \text{(see Section 2),}$$

$$\text{(ii)} \qquad \min_{Q \in \mathcal{O}^+} (Qz_Q^{(k,\,i)})_k = \theta_{ki} - \alpha_{ki}|b_i|$$

and

$$\text{(iii)} \qquad \min_{Q \in \mathcal{O}^+} (Qy_Q^{(k,\,k)})_k = v_k \underset{\sim}{x}_k,$$

$$\text{(iv)} \qquad \min_{Q \in \mathcal{O}^+} (Qz_Q^{(k,\,k)})_k = 0.$$

Proof. For any candidate $c_Q^{(k,\,i)}$, $Qb \longrightarrow (Qc_Q^{(k,\,i)})_k$ is an affine form (see Remark 5.1 after Proposition 5.2) with, except for $(Qb)_k$, negative coefficients (see Proposition 8.2). Therefore, the minimum in $\mathcal{O}^+$ is reached when all components of $(Qb)_{\neq k}$ are maximized, i.e., when $Q = Q^{(k+)}$. More precisely:

(i) and (ii). By Propositions 5.2 and 5.3, and by the definition of θ_{ki}:

$$\theta_{ki} = (Q^{(k+)} y_{Q^{(k+)}}^{(k,\,i)})_k \qquad \text{and} \qquad \theta_{ki} - \alpha_{ki}|b_i| = (Q^{(k+)} z_{Q^{(k+)}}^{(k,\,i)})_k.$$

$$\begin{aligned} \text{(iii)} \ \min_{Q \in \mathcal{O}^+} Qy_Q^{(k,\,k)} &= \min_{Q \in \mathcal{O}^+} -v_k((SQ)x_{(SQ)})_k && \text{by Corollary 5.1} \\ &= -v_k \max\nolimits_{Q \in \mathcal{O}^+}((SQ)x_{(SQ)})_k && \text{because } v_k \geq 0 \text{ (see [15])} \\ &= -v_k \max\nolimits_{Q \in \mathcal{O}^-}(Qx_Q)_k && \\ &= v_k \underset{\sim}{x}_k && \text{by Corollary 4.2.} \end{aligned}$$

(iv) is obvious. □

We are ready now to put all the things together and provide the global minimum for $|x_k|$ in $\Sigma^{(k+)}$.

PROPOSITION 5.4.

$$\Sigma^{(k+)} \neq \emptyset \Longrightarrow \min_{x \in \Sigma^{(k+)}} x_k = \max\{0, v_k \underset{\sim}{x}_k, \hat{x}_k\}.$$

* We have to point out that Λ_i or Γ_i do not depend on the orthant. In other words, for any orthant Q and Q', $y_Q^{(k,i)}$ exists iff $y_{Q'}^{(k,i)}$ exists. Hence, statement of Lemma 5.5 is coherent.

Proof. Assume $\Sigma^{(k+)} \neq \emptyset$. A familiar "minimax" property states that

$$\max_{i\in\mathcal{I}} \min_{Q\in\mathcal{O}^+} (Qc_Q^{(k,i)})_k \leq \min_{Q\in\mathcal{O}^+} \max_{i\in\mathcal{I}} (Qc_Q^{(k,i)})_k. \tag{5.5}$$

But, thanks to Lemma 5.4:

$$\min_{Q\in\mathcal{O}^+} \max_{i\in\mathcal{I}} (Qc_Q^{(k,i)})_k = \min_{Q\in\mathcal{O}^+} \min_{x\in\Sigma\cap Q} |x_k| = \min_{x\in\Sigma^{(k+)}} |x_k| = \min_{x\in\Sigma^{(k+)}} x_k. \tag{5.6}$$

And thanks to Lemma 5.5:

$$\max_{i\in\mathcal{I}} \min_{Q\in\mathcal{O}^+} (Qc_Q^{(k,i)})_k = \max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k. \tag{5.7}$$

Combining (5.5), (5.6), and (5.7), we get:

$$\max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k \leq \min_{x\in\Sigma^{(k+)}} x_k. \tag{5.8}$$

Now, by Proposition 4.1, $\Sigma^{(k+)} \neq \emptyset \Longrightarrow Q^{(k+)} \neq \emptyset$, and by Lemma 5.4,

$$\max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k = \min_{x\in\Sigma\cap Q^{(k+)}} x_k. \tag{5.9}$$

Combining (5.8) and (5.9), we get:

$$\max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k = \min_{x\in\Sigma^{(k+)}} x_k. \tag{5.10}$$

Thanks to Lemma 5.5 again we have:

$$\max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k = \max\{\max_{i\in\mathcal{I}} \theta_{ki},\ \max_{i\in\mathcal{I}} (\theta_{ki} - \alpha_{ki}|b_i|),\ v_k \underset{\sim}{x}_k,\ 0\}.$$

Where $\mathcal{I} := \{i \in [1..n], i \neq k \mid m_{ik} \neq 0\}$. And since $\forall i \in \mathcal{I}$, $\theta_{ki} - \alpha_{ki}|b_i| \leq \theta_{ki}$ (because $\alpha_{ki} \geq 0$), we get:

$$\max_{i\in\mathcal{I}} (Q^{(k+)} c_{Q^{(k+)}}^{(k,i)})_k = \max\{0,\ v_k \underset{\sim}{x}_k,\ \hat{x}_k\}. \tag{5.11}$$

Finally, combining (5.10) and (5.11) leads to:

$$\max\{0,\ v_k \underset{\sim}{x}_k,\ \hat{x}_k\} = \min_{x\in\Sigma^{(k+)}} x_k. \qquad \square$$

COROLLARY 5.3.

$$\Sigma^{(k-)} \neq \emptyset \implies \max_{x\in\Sigma^{(k-)}} x_k = \min\{0,\ v_k \tilde{x}_k,\ -\hat{x}_k\}.$$

Proof. Consider $\mathbf{A}(-x) = -\mathbf{b}$ and apply Proposition 5.4. $\square$

We finish this section by a simple trick. One can notice that 0 does not appear as a candidate in the "else" statements of the scheme given in Section 2. Its removal is justified by the following fact: as soon as 0 belongs to $[\underline{x}_k, \overline{x}_k]$ (the kth component of $\square\Sigma$), we have

$$(\square\Sigma) \cap \Sigma^{(k+)} \neq \emptyset \qquad \text{and} \qquad (\square\Sigma) \cap \Sigma^{(k-)} \neq \emptyset,$$

so that $[\underline{x}_k, \overline{x}_k] = [\underset{\sim}{x}_k, \tilde{x}_k]$. Consequently, 0 can be discarded in any other case.

6. Existence of Solutions

So far, we have proven that the bounds of $\square\Sigma^{(k+)}$ we compute are correct. But we have always assumed that $\Sigma^{(k+)} \neq \emptyset$. In this section, we show that ($\tilde{x}_k \geq \max\{0, v_k \underset{\sim}{x}_k, \hat{x}_k\}$) is a sufficient condition for solutions to exist in $\Sigma^{(k+)}$, providing that either this condition or its negative counterpart is also verified for all the other components.

LEMMA 6.1. $\forall k \in [1..n]$,

If $b_k \geq 0$ **then** $\tilde{x}_k < \max\{0, v_k \underset{\sim}{x}_k\} \implies \underset{\sim}{x}_k > 0$
else $\underset{\sim}{x}_k > \min\{0, v_k \tilde{x}_k\} \implies \tilde{x}_k < 0$

Proof. First, from the expressions of $\tilde{x}_k$ and $\underset{\sim}{x}_k$ follows $\underset{\sim}{x}_k = -\tilde{x}_k + 2m_{kk}b_k$, whatever the sign of b_k is.

Assume $b_k \geq 0$. If $\tilde{x}_k < 0$ then $-\tilde{x}_k > 0$ so $\underset{\sim}{x}_k > 0$. If $\tilde{x}_k < v_k \underset{\sim}{x}_k$ then $-\tilde{x}_k > -v_k \underset{\sim}{x}_k$ so $\underset{\sim}{x}_k > -v_k \underset{\sim}{x}_k + 2m_{kk}b_k$. It follows that $(1 + v_k)\underset{\sim}{x}_k > 2m_{kk}b_k$, which again means that $\underset{\sim}{x}_k > 0$ ($m_{kk} \geq 0$). Hence $\tilde{x}_k < \max\{0, v_k \underset{\sim}{x}_k\} \implies \underset{\sim}{x}_k > 0$.

Assume now $b_k < 0$. If $\underset{\sim}{x}_k > 0$ then $-\underset{\sim}{x}_k < 0$ so $\tilde{x}_k < 0$. If $\underset{\sim}{x}_k > v_k \tilde{x}_k$ then $-\tilde{x}_k + 2m_{kk}b_k > v_k \tilde{x}_k$. It follows that $(1 + v_k)\tilde{x}_k < 2m_{kk}b_k$, which again means that $\tilde{x}_k < 0$. Hence $\underset{\sim}{x}_k > \max\{0, v_k \tilde{x}_k\} \implies \tilde{x}_k < 0$. □

Before giving a sufficient condition for the non-emptiness of $\Sigma^{(k+)}$, we have to introduce an intermediate result on the non-emptiness of $\Sigma \cap Q^*$ (definition of Q^* is given in Section 4).

LEMMA 6.2.

$$\Sigma \cap Q^* \neq \emptyset \iff (\forall k \in [1..n]) \quad \tilde{x}_k \geq \max\{0, v_k \underset{\sim}{x}_k\}$$

or

$$\underset{\sim}{x}_k \leq \min\{0, v_k \tilde{x}_k\}.$$

Proof. For the forward implication, assume $\Sigma \cap Q^* \neq \emptyset$. If $Q^*_{kk} = 1$, then $\Sigma^{(k+)}$ is non empty. It follows by Proposition 5.4 that both 0 and $v_k \underset{\sim}{x}_k$ are lower than $\min\{x_k, x \in \Sigma^{(k+)}\}$, and by Corollary 4.1 that $\max\{x_k, x \in \Sigma^{(k+)}\}$ is $\tilde{x}_k$. Hence, $\tilde{x}_k \geq \max\{0, v_k \underset{\sim}{x}_k\}$. If $Q^*_{kk} = -1$, a similar argument leads to $\underset{\sim}{x}_k \leq \min\{0, v_k \tilde{x}_k\}$.

For the backward implication, we prove that $x_{Q^*} \in \Sigma$. First, x_{Q^*} satisfies (3.2.a) by definition. Now, let k be in $[1..n]$, and remind that x^* stands for $Q^* x_{Q^*}$ (see Section 4). If $b_k \geq 0$ then $x^*_k = \tilde{x}_k$. On the one hand, $\tilde{x}_k \geq 0$ (because $\tilde{x}_k < 0$ would imply $\underset{\sim}{x}_k > 0$ by Lemma 6.1), so $x^*_k \geq 0$; on the other hand, $\tilde{x}_k \geq v_k \underset{\sim}{x}_k$ (for the same reason). By Lemma 5.5, $v_k \underset{\sim}{x}_k = (Q^* y^{(k,k)}_{Q^*})_k$ (since $Q^* = Q^{(k+)}$), so we have $x^*_k \geq (Q^* y^{(k,k)}_{Q^*})_k$ which implies $(I + \Delta)_{k,:} x^* \geq (Q^* b - \delta)_k$, by Corollary 5.2. If $b_k \leq 0$, $x^*_k = -\underset{\sim}{x}_k$. Then $\underset{\sim}{x}_k \leq 0$ implies $x^*_k \geq 0$ and $\underset{\sim}{x}_k \leq v_k \tilde{x}_k$ implies $-\underset{\sim}{x}_k \geq (Q^* y^{(k,k)}_{Q^*})_k$. We conclude in the same way. We have proven that $x^*_k \geq 0$ and $(I + \Delta)_{k,:} x^* \geq (Q^* b - \delta)_k$ for any $k \in [1..n]$. Hence, x_{Q^*} satisfies (3.2.b) and (3.2.c). Finally $x_{Q^*} \in \Sigma$, and $\Sigma \cap Q^* \neq \emptyset$. □

We are now in position to provide a sufficient condition for $\Sigma^{(k+)}$ to be non empty.

PROPOSITION 6.1.

$$\Sigma^{(k+)} \neq \emptyset \iff \begin{cases} \tilde{x}_k \geq \max\{0,\ v_k \underset{\sim}{x}_k,\ \hat{x}_k\}, \\ (\forall i \neq k) \quad \tilde{x}_i \geq \max\{0,\ v_i \underset{\sim}{x}_i\} \quad \textbf{or} \quad \underset{\sim}{x}_i \leq \min\{0,\ v_i \tilde{x}_i\}. \end{cases}$$

Proof. The forward implication is easy to obtain by adapting the first part of the previous proposition proof. For the backward implication, we shall prove that $x_{Q^{(k+)}}$ satisfies system (3.2). For readability, put $t := Q^{(k+)} x_{Q^{(k+)}}$ and notice that $x_{Q^*} \in \Sigma \cap Q^*$ by Lemma 6.2. First, $x_{Q^{(k+)}}$ satisfies (3.2.a) by definition. Next, $\forall i \neq k$, $t_i = x^*_i$ and $x^*_i \geq 0$ since $x_{Q^*} \in \Sigma \cap Q^*$. Further, $t_k = \tilde{x}_k$ by Corollary 4.1 and $\tilde{x}_k \geq 0$ by hypothesis. Hence $t \geq 0$, i.e., $x_{Q^{(k+)}}$ satisfies (3.2.c). It remains to proof that $x_{Q^{(k+)}}$ satisfies (3.2.b):

- Consider $i \neq k$.
  Assume first $m_{ik} \neq 0$. Then, the inner candidate $y^{(k,i)}_{Q^{(k+)}}$ exists (see Proposition 8.1). We have $t_k \geq \hat{x}_k$ by hypothesis, and $\hat{x}_k \geq \theta_{ki} = (Q^{(k+)} y^{(k,i)}_{Q^{(k+)}})_k$ by definition. Then, by Corollary 5.2, we obtain $(I + \Delta)_{i,:} t \geq (Q^{(k+)} b - \delta)_i$.
  Assume now $m_{ik} = 0$. We can notice that the expressions for $(Q^{(k+)} y^{(i,i)}_{Q^{(k+)}})_i$ and $(Q^* y^{(i,i)}_{Q^*})_i$ (see Proposition 5.2) differ only by the term involving respectively $Q^{(k+)}_{kk}$ and Q^*_{kk}. But this term has a factor m_{ik}, so it is null in both cases. Therefore, these expressions coincide. Next, $t_i = x^*_i$ and $x^*_i \geq (Q^* y^{(i,i)}_{Q^*})_i$ since $x_{Q^*} \in \Sigma \cap Q^*$. So $t_i \geq (Q^{(k+)} y^{(i,i)}_{Q^{(k+)}})_i$, and we can apply Corollary 5.2. We find again $(I + \Delta)_{i,:} t \geq (Q^{(k+)} b - \delta)_i$.
- We have $t_k \geq v_k \underset{\sim}{x}_k$ by hypothesis, and $v_k \underset{\sim}{x}_k = (Q^{(k+)} y^{(k,k)}_{Q^{(k+)}})$ by Lemma 5.5. Hence, still by Corollary 5.2, we find $(I + \Delta)_{k,:} t \geq (Q^{(k+)} b - \delta)_k$. □

The negative counterpart of the previous proposition is:

Table 1.

k	numerical result	formal result	Gauss-Seidel	Krawczyk
1	$[4.71053, 11.8576]$	$[\theta_{13}, \tilde{x}_1]$	$[-2.96871, 11.8576]$	$[-3.8576, 11.8576]$
2	$[-9.84177, -6.09412]$	$[\underset{\sim}{x}_2, -\theta_{25}]$	$[-9.84177, -5.10549]$	$[-9.84177, -4.15830]$
3	$[-1.36076, 4.27215]$	$[\underset{\sim}{x}_3, \tilde{x}_3]$	$[-2.04993, 4.27215]$	$[-2.27215, 4.27215]$
4	$[8.09474, 15.81013]$	$[\theta_{43}, \tilde{x}_4]$	$[2.51899, 15.81013]$	$[1.18987, 15.81013]$
5	$[-6.7943, -2.53322]$	$[\underset{\sim}{x}_5, v_5\tilde{x}_5]$	$[-6.7943, -1.71375]$	$[-6.7943, -1.2057]$

COROLLARY 6.1.

$$\Sigma^{(k-)} \neq \emptyset \iff \begin{cases} \underset{\sim}{x}_k \leq \min\{0, v_k\tilde{x}_k, -\hat{x}_k\}, \\ (\forall i \neq k) \quad \tilde{x}_i \geq \max\{0, v_i\underset{\sim}{x}_i\} \quad \textbf{or} \quad \underset{\sim}{x}_i \leq \min\{0, v_i\tilde{x}_i\}. \end{cases}$$

7. Application

Let us first check the validity of our algorithm on one example, and compare the result to what other methods yield.

Take $\mathbf{A}x = \mathbf{b}$, with $\mathbf{A} = I \pm \Delta$ and:

$$\Delta = \begin{pmatrix} .1 & .1 & .1 & .1 & .1 \\ .1 & .2 & .1 & .1 & .1 \\ .2 & .3 & .1 & .2 & .2 \\ .1 & .4 & .1 & .1 & .1 \\ .1 & .5 & .1 & .1 & .1 \end{pmatrix},$$

$$\mathbf{b} = ([1,7]\ [-10,-4]\ [-6,8]\ [8,9]\ [-10,2])^T.$$

The vector of quantifiers for **b** being:

$$(\exists, \forall, \forall, \forall, \forall)^T.$$

In Table 1, for $k = 1..5$, we show the result obtained for the kth component of $\square\Sigma$. In the second column, the bounds are given. In the third column, we write which candidates form these bounds.

So far that we know, the other options for computing an outer estimation of AE-solution sets are the generalized Gauss-Seidel iteration [6], [16], [18], or a Krawczyk-like iteration for AE-solution sets [16], [19]. Results obtained via the latters appear in the two next columns.

We also determined $\square\Sigma$ with the 2^5 simplexes derived from Proposition 3.2, thanks to a `Maple` program. As expected, each bound obtained via our method coincides with the corresponding one given by `Maple`.

We see in the "formal result" column of the table that any candidate can be the "good" one, so it seems unlikely that a simpler formula than ours could be found for $\square\Sigma$.

Now, there is a motivation of our work drawn on practical matters.

Recently, Goldsztejn [4] provided a new branch-and-prune algorithm for paving the solution space of non-linear systems under parameter uncertainty. It combines a test for inner boxes and a special filtering procedure, that takes into account quantifiers associated with parameters. In the last stage of this procedure, there is a need to approximate outwardly a right-quantified solution set, as sharp as possible. It would require too much room to give insight into the reasons why a right-quantified solution set arises. The interested reader may refer to [4] for details.

8. Discussion

We have presented a new algorithm that gives exact bounds for the interval hull of right-quantified solution sets, along with a step-by-step proof.

It is arguable whether the class of right-quantified solution sets is not restrictive. It is so, indeed, from an academic standpoint. However, by tackling right-quantified solution sets instead of united solution sets, the overall number of possibilities for any bound has gone up from 2 to $n + 1$. The difficulty has increased significantly as it is, and AE-solution sets with arbitrary quantifiers on the matrix coefficients are much more inextricable! In any case, we do not believe that the Hansen-Bliek method can be adapted to *all* the AE-solution sets, merely because it strongly relies on the positivity of $\Delta = \text{rad}\,(\mathbf{A})$.

Furthermore, from a practical standpoint, right-quantified solution sets have turned out to be the cornerstone of an algorithm for efficiently narrowing domains in the framework of non-linear square systems with quantified parameters.

Improvements on our work would comprise evaluating the influence of our method on the performances of Goldsztejn's algorithm, returning the hull of each connected components of the right-quantified solution set, or dealing with some specific quantified matrices, e.g., with universally quantifiers arranged column wise.

Appendix

In all this appendix, we fix $i \neq k$. Define the matrix Δ' as follows:

$$\Delta' := (I - D_k + D_i)\Delta \qquad \text{with } D_k = e_k e_k^T \text{ and } D_i = e_k e_i^T.$$

Δ' is a copy of Δ except that the ith row is duplicated in the kth row. Now, as $(I - \Delta)$ is an M-matrix, there exists $u > 0$ such that $(I - \Delta)u > 0$. Since each row of $I - \Delta'$ is also a row of $I - \Delta$, we have $(I - \Delta')u > 0$ and therefore $I - \Delta'$ is still an M-matrix. Now, put

$$M' := (I - \Delta')^{-1}.$$

LEMMA 8.1.

$$m'_{kk} = m'_{ik} + 1.$$

Proof. We have $(\; - \Delta')M' = \;$, so $M' = \; + \Delta' M'$ and $M' - \; = \Delta' M'$. Now $\Delta'_{k,:} = \Delta'_{i,:} \Longrightarrow (M' - \;)_{k,:} = (M' - \;)_{i,:}$, and, in particular, $m'_{kk} = m'_{ik} + 1$. □

LEMMA 8.2.

$$\begin{aligned} m'_{kk} &= \frac{m_{kk}}{m_{kk} - m_{ik}}, \\ m'_{ki} &= m_{ki} - m_{kk}\frac{1 + m_{ki} - m_{ii}}{m_{kk} - m_{ik}}, \\ \forall j \neq k,\ j \neq i,\ m'_{kj} &= m_{kj} - m_{kk}\frac{m_{kj} - m_{ij}}{m_{kk} - m_{ik}}. \end{aligned}$$

Proof. We have:

$$\begin{aligned} M'(\; - \Delta') = \; &\iff M'\big(\; - (\; - D_k + D_i)\Delta\big) = \\ &\iff M'\big(\; - \Delta + (D_k - D_i)\Delta\big)M = M \\ &\iff M'\big((\; - \Delta) - (D_k - D_k\Delta) + (D_i - D_i\Delta) + (D_k - D_i)\big)M \\ &\quad = M \\ &\iff M'\big(\; - D_k + D_i + (D_k - D_i)M\big) = M. \end{aligned}$$

For instance, with $k = 1$, the latter relation can be rewritten:

$$M' \begin{pmatrix} (m_{11} - m_{i1}) & (m_{12} - m_{i2}) & \cdots & (1 + m_{1i} - m_{ii}) & \cdots & (m_{1n} - m_{in}) \\ 0_{n-1} & & & {}_{n-1} & & \end{pmatrix} = M.$$

With such relation, entries of M' can be computed directly. We get for instance $m'_{kk}(m_{kk} - m_{ik}) = m_{kk}$ which means that $m'_{kk} = m_{kk} / (m_{kk} - m_{ik})$. The other entries come similarly, so we skip the details. □

PROPOSITION 8.1 Existence of candidates.

$$\Lambda_i \text{ is regular } \iff \Gamma_i \text{ is regular } \iff m'_{kk} \neq 1 \iff m_{ik} \neq 0.$$

Proof. We shall proof that $Ker(\Lambda_i) = \{0\} \iff m_{ik} \neq 0$.
First, we transform Λ_i in this way:

$$\Lambda_i = (\; - D_k)(\; - \Delta) + D_i(\; + \Delta) = \big(\; - (\; - D_k - D_i)\Delta\big) + D_i - D_k.$$

Define (as in Corollary 5.1) $S := \; - 2D_k$. Then,

$$\begin{aligned} S\Lambda_i &= \big(S - S(\; - D_k - D_i)\Delta\big) + S(D_i - D_k) \\ &= \big(S - (\; - D_k + D_i)\Delta\big) + (D_k - D_i) \\ &\quad \text{(because } S(\; - D_k - D_i) = \; - D_k + D_i \text{ and } S(D_i - D_k) = D_k - D_i) \\ &= \big(\; - 2D_k - (\; - D_k + D_i)\Delta\big) + D_k - D_i \qquad \text{(by definition of } S) \\ &= \big(\; - (\; - D_k + D_i)\Delta\big) - D_k - D_i \\ &= (\; - \Delta') - D_k - D_i. \end{aligned}$$

$$\begin{aligned}\Lambda_i x = 0 \quad &\iff \quad S\Lambda_i x = 0 \qquad \text{(because } S \text{ is a regular diagonal matrix)}\\ &\iff \quad (I - \Delta')x - (D_k + D_i)x = 0\\ &\iff \quad x = M'(D_k + D_i)x\\ &\iff \quad \forall j \in [1..n] \qquad x_j = m'_{jk}(x_k + x_i).\end{aligned}$$

The latter relation yields $x_k = m'_{kk}(x_k + x_i)$ and $x_i = m'_{ik}(x_k + x_i)$. Hence, $x_i = (m'_{kk} - 1)(x_k + x_i) = x_k - x_k - x_i = -x_i$ i.e., $x_i = 0$. Finally,

$$\Lambda_i x = 0 \iff (x_i = 0 \text{ and } \forall j \neq i,\ x_j = m'_{jk} x_k). \tag{8.1}$$

Clearly $m'_{kk} = 1$ implies that $\forall x_k \in \mathbb{R}$ the vector $(m'_{1k}x_k, \ldots, 0, \ldots, m'_{nk}x_k)$ satisfies the right side of (8.1), so the left side, and then $dim(Ker(\Lambda_i)) > 0$.

Conversely, if $m'_{kk} \neq 1$, one can see that only 0_n satisfies the right side of (8.1).

Therefore, $Ker(\Lambda_i) = \{0\} \iff m'_{kk} \neq 1$. With Lemma 8.2 we have

$$m'_{kk} = 1 \iff \frac{m_{kk}}{m_{kk} - m_{ik}} = 1 \iff m_{ik} = 0.$$

Summing up:

$$\Lambda_i \text{ regular} \iff Ker(\Lambda_i) = \{0\} \iff m'_{kk} \neq 1 \iff m_{ik} \neq 0.$$

The proof is similar for Γ_i. We have

$$\Gamma_i = (I - D_k)(I - \Delta) + D_i = I - (I - D_k)\Delta + D_i - D_k.$$

If we multiply it by the regular matrix $S = I - 2D_k + D_i$ we fall back again into

$$S\Gamma_i = (I - \Delta') - D_k - D_i,$$

using the fact that $S(I - D_k) = I - D_k + D_i$ and $S(D_i - D_k) = D_k - D_i$. □

PROPOSITION 8.2. *If $m_{ik} \neq 0$, Then*

(*i*) $\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki} \leq 0$,

(*ii*) $\forall j \neq k,\ j \neq i,\ m_{kj} - m_{ij}\alpha_{ki} \leq 0$.

Proof. (*i*). We have

$$\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki} = \frac{m_{kk}}{m_{ik}} + m_{ki} - \frac{m_{ii}m_{kk}}{m_{ik}} = m_{ki} + m_{kk}\frac{1 - m_{ii}}{m_{ik}}.$$

Let us calculate $(1 - m'_{kk})(\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki})$. We find

$$\left(1 - \frac{m_{kk}}{m_{kk} - m_{ik}}\right)\left(m_{ki} + m_{kk}\frac{1 - m_{ii}}{m_{ik}}\right),$$

and, by developing,

$$m_{ki} + m_{kk}\left(\frac{1 - m_{ii}}{m_{ik}} - \frac{m_{ki} + m_{kk}(1 - m_{ii})\,/\,m_{ik}}{m_{kk} - m_{ik}}\right),$$

i.e.,

$$m_{ki} - m_{kk}\left(\frac{m_{ki} + (1 - m_{ii})}{m_{kk} - m_{ik}}\right),$$

which is m'_{ki}. By Proposition 8.1, $m_{ik} \neq 0$ means also $m'_{kk} \neq 1$. By virtue of Lemma 8.1, $m'_{kk} \geq 1$; hence, $m'_{kk} > 1$. Then $\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki} = m'_{ki} / (1 - m'_{kk})$. And as $m'_{ki} \geq 0$ and $m'_{kk} > 1$, we obtain $\alpha_{ki} + m_{ki} - m_{ii}\alpha_{ki} \leq 0$. Similarly, if we calculate $(1 - m'_{kk})(m_{kj} - m_{ij}\alpha_{ki})$, we get m'_{kj} and conclude in the same way. □

References

1. Aberth, O.: The Solution of Linear Interval Equations by a Linear Programming Method, *Linear Algebra and Its Applications* **259** (1997), pp. 271–279.
2. Bliek, C.: *Computer Methods for Design Automation*, PhD Thesis, Massachusetts Institute of Technology, 1992.
3. Dantzig, G. B.: *Linear Programming and Extensions*, Princeton University Press, 1963.
4. Goldsztejn, A.: A Branch and Prune Algorithm for the Approximation of Non-Linear AE-Solution Sets, in: *ACM SAC*, 2006, pp. 1650–1654.
5. Goldsztejn, A.: A Right-Preconditioning Process for the Formal-Algebraic Approach to Inner and Outer Estimation of AE-Solution Sets, *Reliable Computing* **11** (6) (2005), pp. 443–478.
6. Goldsztejn, A.: *Définition et Applications des Extensions des Fonctions Réelles aux Intervalles Généralisés*, PhD Thesis, Université de Nice-Sophia Antipolis, 2005.
7. Hansen, E. R.: Bounding the Solution of Interval Linear Equations, *SIAM J. Numer. Anal.* **29** (5) (1992), pp. 1493–1503.
8. Heindl, G., Kreinovich, V., and Lakeyev, A. V.: Solving Linear Interval Systems Is NP-Hard Even If We Exclude Overflow and Underflow, *Reliable Computing* **4** (4) (1998), pp. 377–381.
9. Markov, S., Popova, E., and Ullrich, Ch.: On the Solution of Linear Algebraic Equations Involving Interval Coefficients, in: Margenov, S. and Vassilevski, P. (eds), Iterative Methods in Linear Algebra, II, *IMACS Series in Computational and Applied Mathematics* **3** (1996), pp. 216–225.
10. Moore, R.: *Interval Analysis*, Prentice Hall, 1966.
11. Neumaier, A.: A Simple Derivation of the Hansen-Bliek-Rohn-Ning-Kearfott Enclosure for Linear Interval Equations, *Reliable Computing* **5** (2) (1999), pp. 131–136.
12. Neumaier, A.: *Interval Methods for Systems of Equations*, Cambridge University Press, 1990.
13. Ning, S. and Kearfott, R. B.: A Comparison of Some Methods for Solving Linear Interval Equations, *SIAM J. Numer. Anal.* **34** (1) (1997), pp. 1289–1305.
14. Oettli, W.: On the Solution Set of a Linear System with Inaccurate Coefficients, *SIAM J. Numer. Anal.* **2** (1) (1965), pp. 115–118.
15. Rohn, J.: Cheap and Tight Bounds: The Recent Result by E. Hansen Can Be Made More Efficient, *Interval Computations* **1** (4) (1993), pp. 13–21.
16. Shary, S. P.: A New Technique in Systems Analysis Under Interval Uncertainty and Ambiguity, *Reliable Computing* **8** (5) (2002), pp. 321–418.
17. Shary, S. P.: Algebraic Solutions to Interval Linear Equations and Their Application, in: *IMACS—GAMM International Symposium on Numerical Methods and Error Bounds*, 1996.
18. Shary, S. P.: Interval Gauss-Seidel Method for Generalized Solution Sets to Interval Linear Systems, *Reliable Computing* **7** (2) (2001), pp. 141–155.
19. Shary, S. P.: Outer Estimation of Generalized Solution Sets to Interval Linear Systems, *Reliable Computing* **5** (3) (1999), pp. 323–335.