SIAM J. MATRIX ANAL. APPL.
Vol. 16, No. 4, pp. 1223–1240, October 1995

ON THE SYMMETRIC AND UNSYMMETRIC SOLUTION SET OF INTERVAL SYSTEMS*

GÖTZ ALEFELD† AND GÜNTER MAYER‡

Abstract. We consider the solution set S of real linear systems $Ax = b$ with the $n \times n$ coefficient matrix A varying between a lower bound $\underline{A}$ and an upper bound $\overline{A}$, and with b similarly varying between $\underline{b}$, $\overline{b}$. First we list some properties on the shape of S if all matrices A are nonsingular. Then we restrict A to be nonsingular and symmetric deriving a complete description for the boundary of the corresponding symmetric solution set S_{sym} in the 2×2 case. Finally we derive a new criterion for the feasibility of the Cholesky method with which bounds for S_{sym} can be found.

Key words. linear interval equations, unsymmetric solution set, enclosures for the solution set of linear interval systems, symmetric linear systems, symmetric solution set, interval Cholesky method, criteria of feasibility for the interval Cholesky method

AMS subject classifications. 65F05, 65G10

1. Introduction. In [2] we introduced the interval Cholesky method in order to find an interval enclosure $[x]^C$ of the *symmetric solution set*

$$S_{\text{sym}} := \{x \in \boldsymbol{R}^n \mid Ax = b,\ A = A^T \in [A],\ b \in [b]\}, \tag{1.1}$$

where $[A] = [A]^T$ is a given $n \times n$ matrix with real compact intervals as entries, and where $[b]$ is a given vector with n real compact intervals as components. We showed that $[x]^C$ need not enclose the *solution set*

$$S := \{x \in \boldsymbol{R}^n \mid Ax = b,\ A \in [A],\ b \in [b]\} \supseteq S_{\text{sym}}, \tag{1.2}$$

where in this definition the symmetry of A is dropped. This phenomenon is not astonishing, since, in general, S_{sym} differs from S as was shown in [2] by a simple example.

In this paper (§4) we want to intensify our study on the symmetric solution set S_{sym}. To this end, in §3 we repeat some characteristic properties of S. Parts of them are stated and proved in [4]. We will prove them again in a much shorter way than in [4] following the lines in [8]. We then turn over to properties of S_{sym}. For 2×2 matrices S_{sym} can be represented in each orthant O as the intersection of S, O, and two sets of which the boundary is formed by conic sections. Thus, one deduces at once that in the general $n \times n$ case, the boundary ∂S_{sym} can be curvilinear in contrast to ∂S, which is shown in [4] to be the surface of a polytope.

In the second part of our paper (§5) we prove new criteria for the feasibility of the interval Cholesky method. Assuming the midpoint matrix $\check{A}$ of $[A]$ to be symmetric and positive definite we will show, for example, that the method results in an enclosing interval $[x]^C$ if the spectral radius of $\frac{1}{2}|\check{A}^C|d([A])$ is less than 1, where $d([A]) \in \boldsymbol{R}^{n\times n}$ denotes the diameter of $[A]$ and where $|\check{A}^C|$ is a matrix which is defined later.

We mention that symmetric interval systems have also been considered by Jansson [5]. In his paper the symmetric solution set is enclosed by an iterative process.

---

* Received by the editors May 16, 1994; accepted for publication (in revised form) by N. Higham November 8, 1994.

† Institut für Angewandte Mathematik, Universität Karlsruhe, D-76128 Karlsruhe, Germany (goetz.alefeld@mathematik.uni-karlsruhe.de).

‡ Fachbereich Mathematik, Universität Rostock, D-18051 Rostock, Germany (guenter.mayer@mathematik.uni-rostock.de).

2. Preliminaries. We start this section with some notations that we use throughout the paper.

By $\boldsymbol{R}^n$, $\boldsymbol{R}^{m\times n}$, $\boldsymbol{IR}$, $\boldsymbol{IR}^n$, $\boldsymbol{IR}^{m\times n}$, we denote the set of real vectors with n components, the set of real $m \times n$ matrices, the set of intervals, the set of interval vectors with n components, and the set of $m \times n$ interval matrices, respectively. By "interval" we always mean a real compact interval. Interval vectors and interval matrices are vectors and matrices, respectively, with interval entries. We write intervals in brackets with the exception of degenerate intervals (so-called *point intervals*) which we identify with the element being contained, and we proceed similarly with interval vectors and interval matrices. Examples are the ith column $e^{(i)}$ of the $n \times n$ identity matrix I and the null matrix O. As usual, we identify $\boldsymbol{R}^{n\times 1}$ and $\boldsymbol{IR}^{n\times 1}$ with $\boldsymbol{R}^n$ and $\boldsymbol{IR}^n$, respectively. We use the notation $[a] = [\underline{a}, \overline{a}] \in \boldsymbol{IR}$ simultaneously without further reference and, in an analogous way, we write $[x] = [\underline{x}, \overline{x}] = ([x]_i) \in \boldsymbol{IR}^n$ and $[A] = [\underline{A}, \overline{A}] = ([a]_{ij}) \in \boldsymbol{IR}^{n\times n}$. For $[a]$, $[b] \in \boldsymbol{IR}$ we define

$$\begin{aligned}
\check{a} &:= (\underline{a} + \overline{a})/2 && \text{midpoint,}\\
|[a]| &:= \max\{|\underline{a}|, |\overline{a}|\} && \text{absolute value,}\\
d([a]) &:= \overline{a} - \underline{a} && \text{diameter,}\\
q([a],[b]) &:= \max\{|\underline{a} - \underline{b}|\,,\, |\overline{a} - \overline{b}|\} && \text{distance,}\\
\beta([a],[b]) &:= |[a]| + q([a],[b]). &&
\end{aligned} \tag{2.1}$$

For interval vectors and interval matrices, these quantities are defined entrywise, i.e., they are real vectors and matrices, respectively. In particular, $|x| = (|x_i|) \in \boldsymbol{R}^n$ for point vectors x. We equip $\boldsymbol{R}^n$ and also $\boldsymbol{R}^{n\times n}$ with the natural partial ordering $\le$. In addition we write $x < y$ or, equivalently, $y > x$ for vectors $x = (x_i)$, $y = (y_i) \in \boldsymbol{R}^n$ if $x_i < y_i$ for $i = 1, \ldots, n$. With the definition

$$\langle [a] \rangle := \begin{cases} 0 & \text{if } 0 \in [a] \in \boldsymbol{IR},\\ \min\{|\underline{a}|, |\overline{a}|\} & \text{otherwise,} \end{cases}$$

we construct the *comparison matrix* $\langle [A] \rangle := (c_{ij}) \in \boldsymbol{R}^{n\times n}$ of $[A]$ by setting

$$c_{ij} := \begin{cases} \langle [a]_{ij} \rangle & \text{if } i = j,\\ -|[a]_{ij}| & \text{if } i \ne j. \end{cases}$$

We call $[A] \in \boldsymbol{IR}^{n\times n}$ *regular* if no matrix $\tilde{A} \in [A]$ is singular, and we write $\rho(A)$ for the spectral radius of $A \in \boldsymbol{R}^{n\times n}$. Intervals $[a]$ are named *zero symmetric* if $\underline{a} = -\overline{a}$. For interval vectors and interval matrices zero-symmetry is defined entrywise.

We close this section by noting equivalent formulations of nonempty intersections of intervals and by recalling two properties of the function β above, which are proved in [6, Lemma 1.7.5, p. 28].

LEMMA 2.1. *Let $[a]$, $[b] \in \boldsymbol{IR}$. Then the following properties are equivalent.*

(a) $[a] \cap [b] \ne \emptyset$.

(b) $\underline{a} \le \overline{b}$ *and* $\overline{a} \ge \underline{b}$.

(c) $|\check{a} - \check{b}| \le \frac{1}{2} d([a]) + \frac{1}{2} d([b])$.

LEMMA 2.2. *With β from (2.1) the following properties hold.*

(a) *If $[a]_i$, $[b]_i \in \boldsymbol{IR}$, $[a]_i \subseteq [b]_i$ for $i = 1, \ldots, n$, then*

$$\beta\left([a]_1 \cdot \ldots \cdot [a]_n\,,\ [b]_1 \cdot \ldots \cdot [b]_n\right) \le \beta\left([a]_1, [b]_1\right) \cdot \ldots \cdot \beta([a]_n, [b]_n)\,.$$

(b) *If* $[a],\ [b] \in \boldsymbol{IR},\ [a] \subseteq [b]$ *and* $\langle [a] \rangle > q([a],[b])$, *then*

$$\beta\left([a]^{-1},[b]^{-1}\right) \le \left(\langle [a] \rangle - q([a],[b])\right)^{-1},$$

where $[c]^{-1} := \{\, c^{-1} \mid c \in [c] \,\}$ *for* $[c] \in \boldsymbol{IR},\ 0 \notin [c]$.

3. The solution set S. In this section we recall some properties of the solution set S defined in (1.2). To this end, we always assume that a fixed regular interval matrix $[A] \in \boldsymbol{IR}^{n\times n}$ and a fixed interval vector $[b] \in \boldsymbol{IR}^n$ are given. Then the elements of S can be characterized in two equivalent ways.

THEOREM 3.1. *The following three properties are equivalent.*

(a) $x \in S$;

(b) $|\check{A}x - \check{b}| \le \frac{1}{2}d([A])|x| + \frac{1}{2}d([b])$;

(c) $[A]x \cap [b] \ne \emptyset$.

The equivalence (a) $\Leftrightarrow$ (b) is known as Oettli–Prager criterion [7], the equivalence (a) $\Leftrightarrow$ (c) is due to Beeck [3]. We will omit the proof.

To derive some more properties on S we decompose $\boldsymbol{R}^n$ into its closed orthants $O_k,\ k = 1,\dots,2^n$, which are uniquely determined by the signs $s_{k_j} \in \{-1,+1\},\ j = 1,\dots n$, of the components of their interior points. Hence, if O denotes some orthant, fixed by the signs $s_1,\dots,s_n$, then $x = (x_i) \in O$ fulfills

$$x_j \begin{cases} \ge 0 & \text{if } s_j = 1, \\ \le 0 & \text{if } s_j = -1. \end{cases} \tag{3.1}$$

For $[A],\ [b]$ as above, and for $i,j = 1,\dots,n$, let

$$c_{ij} := \begin{cases} \underline{a}_{ij} & \text{if } s_j = 1, \\ \overline{a}_{ij} & \text{if } s_j = -1, \end{cases} \tag{3.2}$$

and

$$d_{ij} := \begin{cases} \overline{a}_{ij} & \text{if } s_j = 1, \\ \underline{a}_{ij} & \text{if } s_j = -1. \end{cases} \tag{3.3}$$

Denote by $\underline{H}_i,\ \overline{H}_i$, the half spaces

$$\left.\begin{aligned} \underline{H}_i &:= \left\{ y \in \boldsymbol{R}^n \,\middle|\, \sum_{j=1}^n c_{ij}y_j \le \overline{b}_i \right\} \\ \overline{H}_i &:= \left\{ y \in \boldsymbol{R}^n \,\middle|\, \sum_{j=1}^n d_{ij}y_j \ge \underline{b}_i \right\} \end{aligned}\right\} \quad i = 1,\dots,n. \tag{3.4}$$

Note that $\underline{H}_i,\ \overline{H}_i$ depend on the choice of the orthant O. By means of these half spaces we can represent $S \cap O$ in the following way (cf. also [8, Cor. 1.2]).

THEOREM 3.2. *Let* $[A] \in \boldsymbol{IR}^{n\times n}$ *be regular and let* O *denote any orthant of* $\boldsymbol{R}^n$. *Then*

$$S \cap O = \bigcap_{i=1}^n (\underline{H}_i \cap \overline{H}_i) \cap O. \tag{3.5}$$

In particular, if $S \cap O$ is nonempty, it is convex, compact, connected, and a polytope.

S is compact, connected, but not necessarily convex. It is the union of finitely many convex polytopes.

Proof. Let $[a] \in \boldsymbol{IR}$, $\xi \in \boldsymbol{R}$. Then

$$\xi \cdot [a] = \begin{cases} [\xi \underline{a}, \xi \overline{a}] & \text{if } \xi \geq 0, \\ [\xi \overline{a}, \xi \underline{a}] & \text{if } \xi < 0. \end{cases}$$

Hence (3.5) follows from Lemma 2.1(a), (b), from Theorem 3.1(a), (c), and from the definition of $\underline{H}_i$, $\overline{H}_i$.

Since O, $\underline{H}_i$, $\overline{H}_i$ are convex, the same holds for $S \cap O$ because of (3.5). This in turn shows that $S \cap O$ is connected. The compactness and the connectivity of S follows from the same property of $[A] \times [b]$ and from the continuity of the function

$$g : \begin{cases} [A] \times [b] & \rightarrow & \boldsymbol{R}^n, \\ (A, b) & \mapsto & A^{-1} b, \end{cases}$$

the range of which is S. Now S being compact the same holds for $S \cap O$ since O is closed. The remaining property of S follows trivially from

$$S = \bigcup_{j=1}^{2^n} (S \cap O_j)$$

and from (3.5), where O_j, $j = 1, \ldots, 2^n$, denote the orthants of $\boldsymbol{R}^n$ numbered arbitrarily. $\square$

That S can be nonconvex is seen by the following example.

Example 3.3. Let $[A] = \begin{pmatrix} 1 & 0 \\ [-1, 1] & 1 \end{pmatrix}$, $[b] = \begin{pmatrix} [-1, 1] \\ 0 \end{pmatrix}$. Then S is given by $S = \{(x, y) \mid |y| \leq |x| \leq 1\}$ as illustrated in Fig. 1.

THEOREM 3.4. *Let $[A]$ be a point matrix. Then S is a parallelepiped; in particular, S is convex.*

Proof. Let $[A] = [A, A]$, and denote the columns of A^{-1} by $c^1, \ldots, c^n$. Then

$$S = \left\{ A^{-1} \underline{b} + \sum_{j=1}^{n} t_j c^j \mid 0 \leq t_j \leq d([b]_j),\ j = 1, \ldots, n \right\}.$$

This proves the theorem. $\square$

We remark that a necessary and sufficient criterion for the convexity of S can be found in [9].

4. On the symmetric solution set S_{sym}. We now turn over to the symmetric solution set S_{sym} defined in (1.1). We again assume $[A] \in \boldsymbol{IR}^{n \times n}$ to be regular, and, in addition, to fulfill

$$[A] = [A]^T,$$

which is equivalent to $\underline{A} = \underline{A}^T$ and $\overline{A} = \overline{A}^T$.

We first prove two simple properties of S_{sym}.

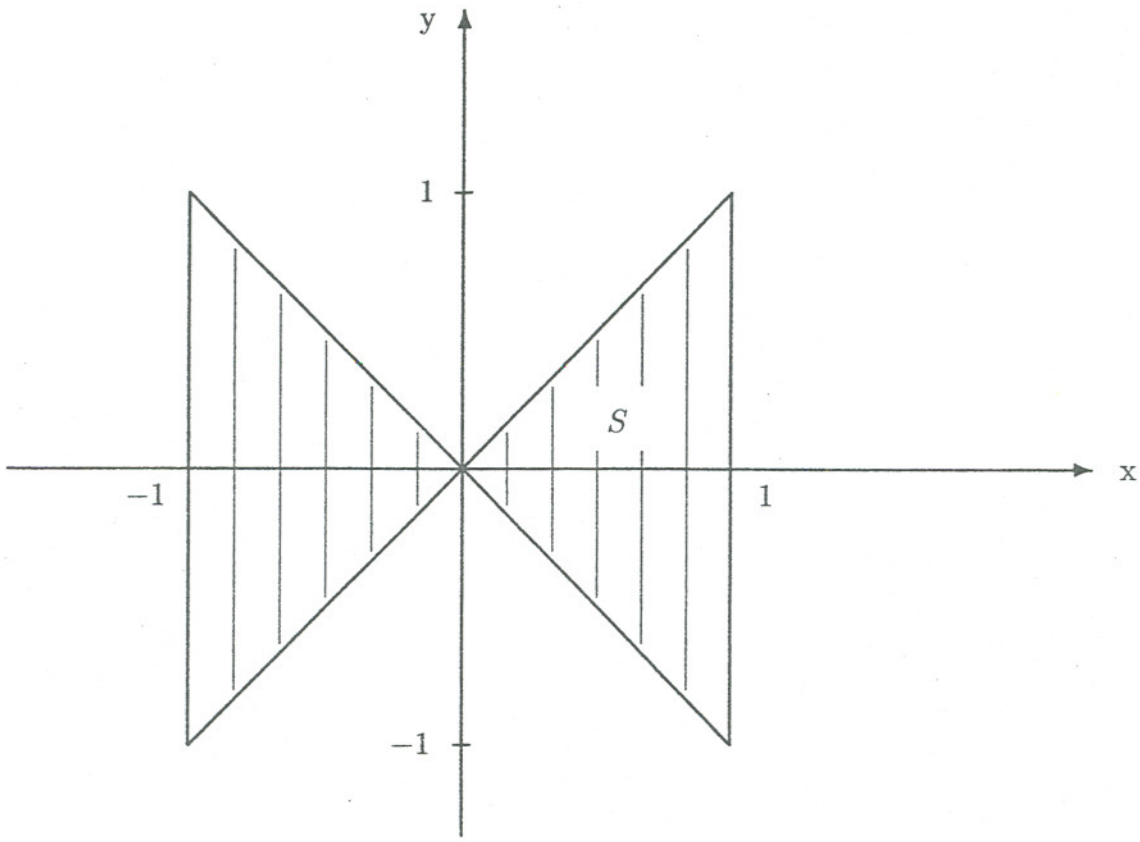

FIG. 1. *The shape of the solution set S in Example 3.3.*

THEOREM 4.1. *Let $[A] = [A]^T \in \boldsymbol{IR}^{n\times n}$ be regular. Then S_{sym} is compact and connected.*

Proof. Define $[A]_{\text{sym}} := \{A \in [A] |\ A = A^T\}$. Then

$$f : \begin{cases} [A]_{\text{sym}} \times [b] & \rightarrow \quad \boldsymbol{R}^n, \\ (A, b) & \mapsto \quad A^{-1}b \end{cases} \tag{4.1}$$

is continuous. Let $\{A_k\}$ be an infinite sequence from $[A]_{\text{sym}}$. Since the $(1, 1)$-entries of A_k are all contained in the compact set $[a]_{11}$, there is a subsequence $\{A_k^{(1)}\}$ of $\{A_k\}$ such that its $(1, 1)$-entries are convergent. By the same reason one can choose a subsequence $\{A_k^{(2)}\}$ of $\{A_k^{(1)}\}$ such that the $(1, 2)$-entries are convergent. It is obvious that the $(1, 1)$-entries of $\{A_k^{(2)}\}$ keep this property. Repeating the arguments by running through the indices (i, j), $1 \le i \le j \le n$ and taking into account the symmetry of A_k shows that there is a convergent subsequence of $\{A_k\}$, which proves $[A]_{\text{sym}}$ to be compact. Therefore, $[A]_{\text{sym}} \times [b]$ is compact, and the same holds for the range S_{sym} of f.

If $A_1,\ A_2 \in [A]_{\text{sym}}$ then the line segment $A_1 + t(A_2 - A_1) \in [A]_{\text{sym}},\ 0 \le t \le 1$. Hence $[A]_{\text{sym}}$ is connected and also $[A]_{sym} \times [b]$. Using the continuous function f from (4.1) once more shows S_{sym} to be connected. ◻

We next investigate S_{sym} in the 2×2 case more carefully. To this end, as in §3, we fix an orthant O given by the signs $s_1, \ldots, s_n$ of the components of its interior points. We define $\underline{H}_i$, $\overline{H}_i$ as in (3.2)–(3.4) and $e,\ f \in \boldsymbol{R}^n$ by

$$e_i := \begin{cases} \underline{b}_i & \text{if } s_i = 1, \\ \overline{b}_i & \text{if } s_i = -1, \end{cases}$$

$$f_i := \begin{cases} \overline{b}_i & \text{if } s_i = 1, \\ \underline{b}_i & \text{if } s_i = -1. \end{cases}$$

For $n = 2$ we use the sets

$$C^- := \left\{ y \in \boldsymbol{R}^2 \mid \underline{a}_{11} y_1^2 - \overline{a}_{22} y_2^2 - f_1 y_1 + e_2 y_2 \le 0 \right\}, \tag{4.2}$$

$$C^+ := \left\{ y \in \boldsymbol{R}^2 \mid \overline{a}_{11} y_1^2 - \underline{a}_{22} y_2^2 - e_1 y_1 + f_2 y_2 \ge 0 \right\}. \tag{4.3}$$

Obviously, each of these two sets has a conic section as boundary provided that $\underline{a}_{11}^2 + \overline{a}_{22}^2 \neq 0$ for C^- and, similarly, $\overline{a}_{11}^2 + \underline{a}_{22}^2 \neq 0$ for C^+. As for the hyperplanes $\underline{H}_i$, $\overline{H}_i$ in §3 we point out that C^-, C^+ depend on the choice of the orthant O. However, the type of the conic section is independent of O if one does not distinguish between hyperbolas and pairs of intersecting straight lines, and if one considers a single point as an ellipse. If each symmetric matrix from $[A]$ is positive definite then $\underline{a}_{ii} > 0$, $i = 1, 2$, hence the boundary of C^- and C^+ is formed by hyperbolas in the above-mentioned generalized sense.

We now describe S_{sym} in the 2×2 case by means of S, C^-, and C^+ .

THEOREM 4.2. *Let $[A] = [A]^T \in \boldsymbol{IR}^{2\times 2}$ be regular and let O denote any orthant of $\boldsymbol{R}^2$. Then*

$$S_{\text{sym}} \cap O = S \cap O \cap C^- \cap C^+ . \tag{4.4}$$

In particular, if $S_{\text{sym}} \cap O$ is nonempty, it is compact, but not necessarily convex.

Proof. The compactness follows from Theorem 4.1. The nonconvexity is shown by Example 4.4. It remains to prove (4.4).

$\subseteq$: Let $x \in S_{\text{sym}} \cap O$. Then $x \in S \cap O$, and there exists a *symmetric* matrix $A \in [A]$ and a vector $b \in [b]$ such that $Ax = b$. With $[t] := [a]_{12} = [a]_{21}$ and $t := a_{12} = a_{21}$ we get

$$a_{11} x_1 + t x_2 = b_1 \ , \tag{4.5}$$

$$t x_1 + a_{22} x_2 = b_2 \ . \tag{4.6}$$

Multiplying (4.5) by x_1 and (4.6) by x_2 and substituting $t x_1 x_2$ we obtain

$$a_{11} x_1^2 - a_{22} x_2^2 = b_1 x_1 - b_2 x_2 \ .$$

Thus

$$x^T \begin{pmatrix} [a]_{11} & 0 \\ 0 & -[a]_{22} \end{pmatrix} x \cap x^T \begin{pmatrix} [b]_1 \\ -[b]_2 \end{pmatrix} \neq \emptyset \, , \tag{4.7}$$

whence, by Lemma 2.1, we get equivalently

$$\underline{a}_{11} x_1^2 - \overline{a}_{22} x_2^2 \le f_1 x_1 - e_2 x_2 \ ,$$
$$\overline{a}_{11} x_1^2 - \underline{a}_{22} x_2^2 \ge e_1 x_1 - f_2 x_2 \ .$$

This means $x \in C^-$ and $x \in C^+$, respectively. Therefore, $S_{\text{sym}} \cap O \subseteq S \cap O \cap C^- \cap C^+$.

$\supseteq$: Let

$$x \in S \cap O \cap C^- \cap C^+ . \tag{4.8}$$

Since $x \in S$, there are $A \in [A]$, $b \in [b]$ such that

$$Ax = b \tag{4.9}$$

holds. We are going to show that $A \in [A]$ in (4.9) can be chosen to be symmetric when changing $b \in [b]$ appropriately. To simplify the notation we use

$$t_1 := a_{12} \in [a]_{12} \quad \text{and} \quad t_2 := a_{21} \in [a]_{21} = [a]_{12} =: [t]$$

for the two off-diagonal entries of A in (4.9).

If $t_1 = t_2$ then $x \in S_{\text{sym}} \cap O$. Therefore, assume $t_1 \neq t_2$, say

$$t_1 < t_2 . \tag{4.10}$$

If $x_1 = 0$ then A can be replaced in (4.9) by the symmetric matrix

$$A_{\text{sym}} := \begin{pmatrix} a_{11} & t_1 \\ t_1 & a_{22} \end{pmatrix}$$

thus showing $x \in S_{\text{sym}} \cap O$. Analogously one proceeds for $x_2 = 0$.

Let now $x_1 \neq 0$ and $x_2 \neq 0$. We first consider the case $x_1 > 0,\ x_2 > 0$, which, by (4.8), means that O is the first quadrant of $\boldsymbol{R}^2$. Our proof is based on the equivalence of (4.9) with

$$t_1 = \frac{b_1 - a_{11}x_1}{x_2} \in [t], \qquad t_2 = \frac{b_2 - a_{22}x_2}{x_1} \in [t] . \tag{4.11}$$

Assume $x \notin S_{\text{sym}} \cap O$. This means that $b \in [b]$ and $A \in [A]$ from (4.9) *cannot* be replaced such that (4.9) is satisfied for some symmetric matrix $A_{\text{sym}} \in [A]$ and some suitably modified vector $b \in [b]$. Taking into account (4.10) we consequently obtain

$$\underline{t} \le t_1 \le t_{\max} := \frac{\overline{b}_1 - \underline{a}_{11}x_1}{x_2} < t_{\min} := \frac{\underline{b}_2 - \overline{a}_{22}x_2}{x_1} \le t_2 \le \overline{t}, \tag{4.12}$$

whence

$$\overline{b}_1 x_1 - \underline{a}_{11} x_1^2 < \underline{b}_2 x_2 - \overline{a}_{22} x_2^2 .$$

Since we supposed O to be the first quadrant this implies $x \notin C^-$, which contradicts (4.8).

Replacing (4.10) by $t_1 > t_2$ and assuming $x \notin S_{\text{sym}} \cap O$ yields

$$\overline{t} \ge t_1 \ge t_{\min} := \frac{\underline{b}_1 - \overline{a}_{11}x_1}{x_2} > t_{\max} := \frac{\overline{b}_2 - \underline{a}_{22}x_2}{x_1} \ge t_2 \ge \underline{t}$$

from which we get the contradiction $x \notin C^+$. Therefore,

$$S \cap O \cap C^- \cap C^+ \subseteq S_{\text{sym}} \cap O \tag{4.13}$$

holds if O is the first quadrant O_1 .

Let now $x \in O \neq O_1$, $x_1 \neq 0$, $x_2 \neq 0$, $s_1 := \mathrm{sign}(x_1)$, $s_2 := \mathrm{sign}(x_2)$, $D_x := \mathrm{diag}(s_1, s_2) \in \boldsymbol{R}^{2\times 2}$. Then (4.9) is equivalent to

$$\hat{A}\hat{x} = \hat{b} \tag{4.14}$$

with $\hat{A} := D_x A D_x \in D_x[A]D_x =: [\hat{A}]$, $\hat{x} := D_x x \in O_1$, $\hat{b} := D_x b \in D_x[b] =: [\hat{b}]$. Let S, S_{sym}, C^-, C^+, e_i, f_i be associated with the given quantities $[A]$, $[b]$, and O, and let $\hat{S}$, $\hat{S}_{\mathrm{sym}}$, $\hat{C}^-$, $\hat{C}^+$, $\hat{e}_i$, $\hat{f}_i$ be the corresponding quantities associated with $[\hat{A}]$, $[\hat{b}]$, and O_1 . Since

$$s_1 f_1 = \left\{ \begin{array}{rll} \overline{b}_1 & \text{if} \quad s_1 = & 1 \\ -\underline{b}_1 & \text{if} \quad s_1 = & -1 \end{array} \right\} = \max\{s_1[b]_1\} = \hat{f}_1$$

and

$$s_2 e_2 = \left\{ \begin{array}{rll} \underline{b}_2 & \text{if} \quad s_2 = & 1 \\ -\overline{b}_2 & \text{if} \quad s_2 = & -1 \end{array} \right\} = \min\{s_2[b]_2\} = \hat{e}_2,$$

we get from $y \in C^-$ the inequality

$$\begin{aligned} 0 &\geq (s_1\underline{a}_{11}s_1)(s_1y_1)^2 - (s_2\overline{a}_{22}s_2)(s_2y_2)^2 - (s_1f_1)(s_1y_1) + (s_2e_2)(s_2y_2) \\ &= \hat{\underline{a}}_{11}\hat{y}_1^2 - \hat{\overline{a}}_{22}\hat{y}_2^2 - \hat{f}_1\hat{y}_1 + \hat{e}_2\hat{y}_2, \end{aligned}$$

where $\hat{y} := D_x y$. Hence $y \in C^-$ implies $\hat{y} \in \hat{C}^-$, and analogously $y \in C^+$ yields $\hat{y} \in \hat{C}^+$. Therefore, $x \in S \cap O \cap C^- \cap C^+$ results in $\hat{x} \in \hat{S} \cap O_1 \cap \hat{C}^- \cap \hat{C}^+$ whence

$$\hat{x} \in \hat{S}_{\mathrm{sym}} \cap O_1 \tag{4.15}$$

as we have proved above. Since (4.15) implies $\hat{A}_{\mathrm{sym}}\hat{x} = \hat{b}$ for some symmetric matrix $\hat{A}_{\mathrm{sym}} \in [\hat{A}]$ and some right-hand side $\hat{b} \in [\hat{b}]$, it yields $x \in S_{\mathrm{sym}} \cap O$ via (4.14). □

The generalization of Theorem 4.2 for the case $n > 2$ is not straightforward since the elimination process performed in the proof does not seem to work in this case.

Since $x \in C^- \cap C^+$ is equivalent to (4.7), we obtain immediately the subsequent corollary from Theorem 3.1(a), (c) and from Theorem 4.2.

COROLLARY 4.3. *For regular matrices* $[A] = [A]^T \in \boldsymbol{IR}^{2\times 2}$ *and* $[b] \in \boldsymbol{IR}^2$ *the following properties are equivalent.*

(a) $x \in S_{\mathrm{sym}}$.

(b) $[A]x \cap [b] \neq \emptyset$ *(i. e.,* $x \in S$*) and*

$$x^T \begin{pmatrix} [a]_{11} & 0 \\ 0 & -[a]_{22} \end{pmatrix} x \cap x^T \begin{pmatrix} [b]_1 \\ -[b]_2 \end{pmatrix} \neq \emptyset .$$

Note that in contrast to Theorem 4.2 no orthant enters explicitly in Corollary 4.3. Therefore, it can be viewed as an analogue of Theorem 3.1.

We now illustrate Theorem 4.2 by two examples. In particular we show that S_{sym} can be nonconvex in the orthants and that its boundary can be curvilinear.

Example 4.4. Let

$$[A] := \begin{pmatrix} 5 & [-4, 0] \\ [-4, 0] & 5 \end{pmatrix} \text{ and } [b] := \begin{pmatrix} 9 \\ 0 \end{pmatrix}.$$

With

$$A_{t_1,t_2} := \begin{pmatrix} 5 & t_1 \\ t_2 & 5 \end{pmatrix}, \; t_1, t_2 \in [-4, 0],$$

we get

$$A_{t_1,t_2}^{-1} = \frac{1}{25 - t_1 t_2} \begin{pmatrix} 5 & -t_1 \\ -t_2 & 5 \end{pmatrix} \geq 0$$

and

$$A_{t_1,t_2}^{-1} \cdot \begin{pmatrix} 9 \\ 0 \end{pmatrix} = \frac{1}{25 - t_1 t_2} \begin{pmatrix} 45 \\ -9t_2 \end{pmatrix}, \quad t_1,\, t_2 \in [-4, 0] \,. \tag{4.16}$$

Hence S and S_{sym} are completely contained in the first quadrant O_1. With the notations of §§3 and 4 we obtain

$$\begin{aligned} \underline{H}_1 &= \{y \in \boldsymbol{R}^2 \mid 5y_1 - 4y_2 \leq 9\}\,, & \overline{H}_1 &= \{y \in \boldsymbol{R}^2 \mid 5y_1 \geq 9\}\,, \\ \underline{H}_2 &= \{y \in \boldsymbol{R}^2 \mid -4y_1 + 5y_2 \leq 0\}\,, & \overline{H}_2 &= \{y \in \boldsymbol{R}^2 \mid 5y_2 \geq 0\}\,, \end{aligned}$$

hence $S = \underline{H}_1 \cap \overline{H}_1 \cap \underline{H}_2 \cap \overline{H}_2 \cap O_1$ is the triangle with the vertices $(1.8, 0)$, $(1.8, 1.44)$, and $(5, 4)$. To describe S_{sym} we list the sets

$$\begin{aligned} C^- &= \left\{y \in \boldsymbol{R}^2 \mid 5y_1^2 - 5y_2^2 - 9y_1 \leq 0\right\}, \\ C^+ &= \left\{y \in \boldsymbol{R}^2 \mid 5y_1^2 - 5y_2^2 - 9y_1 \geq 0\right\}. \end{aligned}$$

Then $K := C^- \cap C^+$ is the hyperbola

$$K: \quad \left(y_1 - \frac{9}{10}\right)^2 - y_2^2 = \frac{81}{100}\,.$$

By (4.16) or by Theorem 4.2 one can see that S_{sym} is that part of the right branch of K which lies between the points $(1.8, 0)$ and $(5, 4)$. The sets S and S_{sym} are illustrated in Fig. 2.

Our next example shows that parts of a parabola, of a circle, and straight lines can also form the boundary of S_{sym}.

Example 4.5. Let

$$[A] := \begin{pmatrix} 1 & [1,2] \\ [1,2] & [-1,0] \end{pmatrix}, \quad [b] := \begin{pmatrix} 4 \\ [1,2] \end{pmatrix}, \quad A_{\alpha,\beta,\gamma} := \begin{pmatrix} 1 & \alpha \\ \beta & \gamma \end{pmatrix} \in [A]$$

with α, $\beta \in [1,2]$, $\gamma \in [-1,0]$. Since $\det A_{\alpha,\beta,\gamma} = \gamma - \alpha\beta \leq -1$, the interval matrix $[A]$ is regular with

$$A_{\alpha,\beta,\gamma}^{-1} = \frac{1}{\det A_{\alpha,\beta,\gamma}} \begin{pmatrix} \gamma & -\alpha \\ -\beta & 1 \end{pmatrix}.$$

With $b_1 \geq 2b_2 \geq 2$ we get $A_{\alpha,\beta,\gamma}^{-1} \cdot b \geq 0$ for any choice $A_{\alpha,\beta,\gamma} \in [A]$, $b \in [b]$. Hence S and S_{sym} are completely contained in the first quadrant O_1. Using the notation above we obtain for O_1 the following sets:

$$\begin{aligned} \underline{H}_1 &= \{y \in \boldsymbol{R}^2 \mid y_1 + y_2 \leq 4\}, \overline{H}_1 = \{y \in \boldsymbol{R}^2 \mid y_1 + 2y_2 \geq 4\}, \\ \underline{H}_2 &= \{y \in \boldsymbol{R}^2 \mid y_1 - y_2 \leq 2\}, \overline{H}_2 = \{y \in \boldsymbol{R}^2 \mid 2y_1 \geq 1\}\,, \end{aligned}$$

$$\begin{aligned} C^- &= \left\{y \in \boldsymbol{R}^2 \mid y_1^2 - 4y_1 + y_2 \leq 0\right\} = \left\{y \in \boldsymbol{R}^2 \mid y_2 \leq 4 - (y_1 - 2)^2\right\}, \\ C^+ &= \left\{y \in \boldsymbol{R}^2 \mid y_1^2 + y_2^2 - 4y_1 + 2y_2 \geq 0\right\} = \left\{y \in \boldsymbol{R}^2 \mid (y_1 - 2)^2 + (y_2 + 1)^2 \geq 5\right\}. \end{aligned}$$

The set $S = \underline{H}_1 \cap \overline{H}_1 \cap \underline{H}_2 \cap \overline{H}_2 \cap O_1$ is the convex hull of the points $(\frac{1}{2}, \frac{7}{4})$, $(\frac{1}{2}, \frac{7}{2})$, $(\frac{8}{3}, \frac{2}{3})$ and $(3, 1)$. The boundary of $S_{\text{sym}} = S \cap O_1 \cap C^- \cap C^+$ is formed by the following four curves.

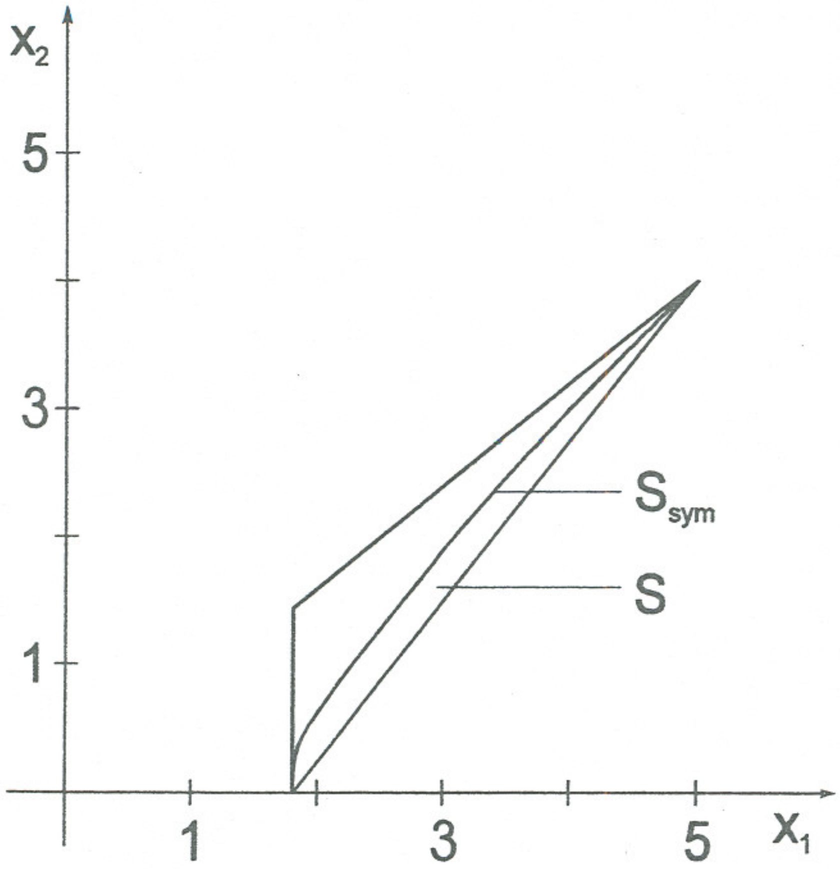

FIG. 2. *The shape of the solution sets S and S_{sym} in Example* 4.4.

(i) The straight line between $(\frac{1}{2}, \frac{7}{4})$ and $(\frac{8}{5}, \frac{6}{5})$.
(ii) The straight line between $(1, 3)$ and $(3, 1)$.
(iii) The part of the parabola $y_2 = 4 - (y_1 - 2)^2$ between $(\frac{1}{2}, \frac{7}{4})$ and $(1, 3)$.
(iv) The part of the circle $(y_1 - 2)^2 + (y_2 + 1)^2 = 5$ between $(\frac{8}{5}, \frac{6}{5})$ and $(3, 1)$.

The situation is illustrated in Fig. 3.

5. Computing enclosures for S_{sym}. As was shown in [2], S_{sym} can be enclosed by the vector $[x]^C$, which results from the following interval version of the well-known Cholesky method, for which we assume $[A] = [A]^T \in \boldsymbol{IR}^{n \times n}$, and $[b] \in \boldsymbol{IR}^n$.

Step 1. "LL^T decomposition"
for $j := 1$ **to** n **do**

$$[l]_{jj} := \left([a]_{jj} - \sum_{k=1}^{j-1} [l]_{jk}^2 \right)^{\frac{1}{2}} ;$$

for $i := j + 1$ **to** n **do**

$$[l]_{ij} := \left([a]_{ij} - \sum_{k=1}^{j-1} [l]_{ik} [l]_{jk} \right) / [l]_{jj} ;$$

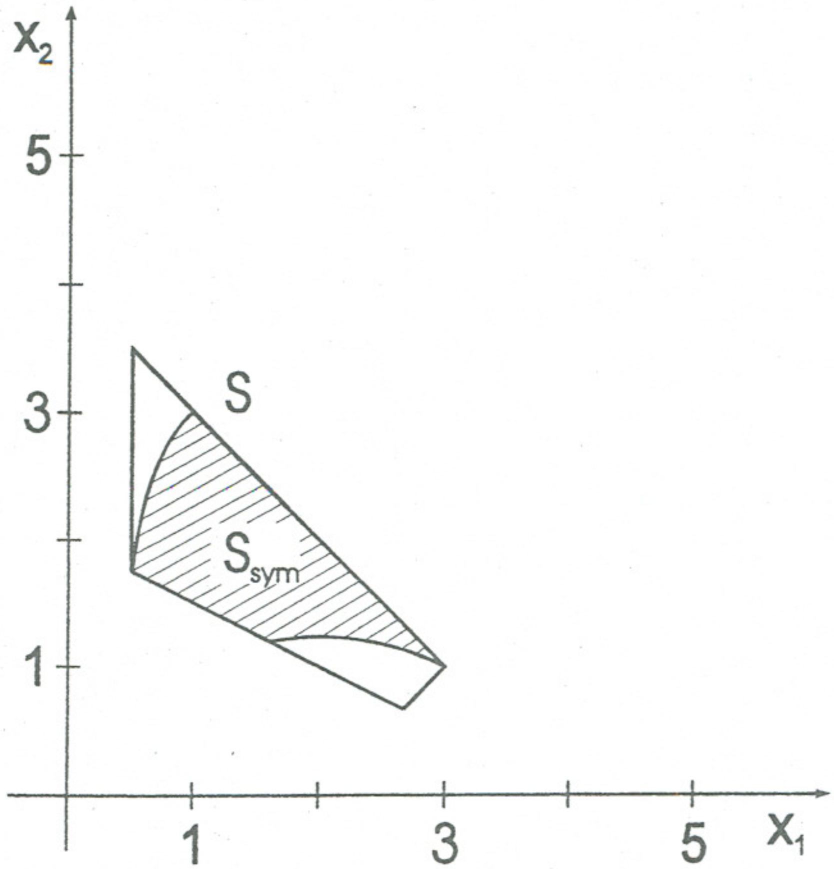

FIG. 3. *The shape of the solution sets S and S_{sym} in Example* 4.5.

Step 2. Forward substitution
for $i := 1$ **to** n **do**

$$[y]_i := \left([b]_i - \sum_{j=1}^{i-1} [l]_{ij}[y]_j \right) / [l]_{ii} \ ;$$

Step 3. Backward substitution
for $i := n$ **downto** 1 **do**

$$[x]_i^C := \left([y]_i - \sum_{j=i+1}^{n} [l]_{ji}[x]_j^C \right) / [l]_{ii} \ ;$$

$\text{ICh}([A],[b]) := [x]^C$.

Here,

$$[a]^2 := \left\{ a^2 \mid a \in [a] \right\} \tag{5.1}$$

and

$$[a]^{1/2} := \sqrt{[a]} := \{ \sqrt{a} \mid a \in [a] \}$$

for intervals $[a]$.

In contrast to the classical, i.e., noninterval Cholesky method, it is an open question when the interval Cholesky method is feasible. In [2] several criteria are given that guarantee the existence of $[x]^C$. We add here two new ones as well as a nonexistence criterion, which we formulate first.

THEOREM 5.1. *If $[A] = [A]^T \in \boldsymbol{IR}^{n\times n}$ contains at least one symmetric matrix A which is not positive definite, then $[x]^C$ does not exist.*

Proof. We first recall that a real symmetric matrix has an LL^T-decomposition with positive diagonal entries l_{ii} if and only if this matrix is positive definite (see [11]). L can be computed by the Cholesky method. Assume now that $A = A^T \in [A] = [A]^T$ is not positive definite. Then the Cholesky method will break down. This is the case if and only if for some index j either l_{jj} cannot be computed because of

$$a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2 < 0$$

(see Step 1) or y_i cannot be computed because of $l_{ii} = 0$ (see Step 2). By the inclusion monotonicity of the interval arithmetic, either $[l]_{jj}$ does not exist, or $0 \in [l]_{ii}$ and the interval Cholesky method will break down. □

Example 4.5 illustrates Theorem 5.1: Since

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in [A] = \begin{pmatrix} 1 & [1,2] \\ [1,2] & [-1,0] \end{pmatrix}$$

is not positive definitive, $[x]^C$ does not exist for $[A]$. Note, however, that the interval Gaussian algorithm is feasible for this interval matrix.

Before formulating our new feasibility criterion we need some preparations.

By Theorem 3.4 in [2] we have for $[y]$ from Step 2 in the interval version of the Cholesky method

$$[y] = [D^n]\left([L^{n-1}]\left([D^{n-1}]\left(\ldots\left([L^2]\left([D^2]\left([L^1]\left([D^1][b]\right)\right)\right)\right)\ldots\right)\right)\right)$$

and

$$[x]^C = [D^1]\left([L^1]^T\left([D^2]\left(\ldots\left([L^{n-2}]^T\left([D^{n-1}]\left([L^{n-1}]^T\left([D^n][y]\right)\right)\right)\right)\ldots\right)\right)\right), \tag{5.2}$$

where the diagonal matrices $[D^s]$ and the lower triangular matrices $[L^s]$ are defined for $s = 1, \ldots, n-1$ by

$$[d^s]_{ij} := \begin{cases} 1 & \text{if } i = j \neq s, \\ 1/[l]_{ss} & \text{if } i = j = s, \\ 0 & \text{otherwise}, \end{cases}$$

$$[l^s]_{ij} := \begin{cases} 1 & \text{if } i = j, \\ -[l]_{is} & \text{if } i > j = s, \\ 0 & \text{otherwise}, \end{cases}$$

with $[l]_{ij}$ from the Cholesky method. (Note that $[l]_{ij}$ is computed in the jth step of the "LL^T-decomposition"). By (5.2) it is easy to see that the mapping

$$f : \begin{cases} \boldsymbol{IR}^n & \to \quad \boldsymbol{IR}^n, \\ [b] & \mapsto \quad ICh([A],[b]) \end{cases} \tag{5.3}$$

is a *sublinear* one in the sense of [6, p. 98], i.e.,

(i) $\quad [b] \subseteq [c] \;\Rightarrow\; f([b]) \subseteq f([c])$,

(ii) $\quad \alpha \in \boldsymbol{R} \;\Rightarrow\; f(\alpha[b]) = \alpha f([b])$,

(iii) $\quad f([b] + [c]) \;\subseteq\; f([b]) + f([c])$ for $[b],\ [c] \in \boldsymbol{IR}^n$.

An easy computation yields

$$| [D^n] | \cdot | [L^{n-1}] | \cdot | [D^{n-1}] | \cdot \ldots \cdot | [L^2] | \cdot | [D^2] | \cdot | [L^1] | \cdot | [D^1] | = \langle [L] \rangle^{-1}$$

again with $[L] = ([l]_{ij})$ from the Cholesky method. Hence, for the particular "right-hand side" $[\hat{b}] := [-1,1]e$, where $e = (1,\ldots,1)^T$, one gets

$$\mathrm{ICh}([A],[\hat{b}]) = [x]^C = \langle [L]^T \rangle^{-1} \left(\langle [L] \rangle^{-1} [\hat{b}] \right) = \left(\langle [L]^T \rangle^{-1} \langle [L] \rangle^{-1} \right) [\hat{b}] \, .$$

With the abbreviation

$$| [A]^C | := \langle [L]^T \rangle^{-1} \langle [L] \rangle^{-1}, \tag{5.4}$$

one therefore obtains for any $[b] \subseteq [\hat{b}]$ the inclusion

$$\mathrm{ICh}([A],[b]) \subseteq | [A]^C | \, [\hat{b}] \, .$$

Thus, $| [A]^C |$ can be thought of as a *measure* for the width of the enclosure $\mathrm{ICh}([A],[b])$ of S_{sym} that does not depend on the right-hand side $[b]$ as long as $[b]$ is contained in $[\hat{b}]$. The condition $[b] \subseteq [\hat{b}]$ can be considered as a sort of normalization. If it no longer holds, replace $[\hat{b}]$ by $t[\hat{b}]$ with $t > 0$ as small as possible such that $[b] \subseteq t[\hat{b}]$ is valid. Then

$$\mathrm{ICh}([A],[b]) \subseteq t | [A]^C | \, [\hat{b}],$$

hence $t| [A]^C |$ is a corresponding measure.

By (5.2) we also get

$$\left| \left(\mathrm{ICh}\left([A], [-e^{(1)}, e^{(1)}] \right), \ldots, \mathrm{ICh}\left([A], [-e^{(n)}, e^{(n)}] \right) \right) \right| = | [A]^C | \, ,$$

hence $| [A]^C |$ is the *absolute value* of the sublinear mapping f in the sense of [6, p. 100]. By an elementary rule of the diameter d (cf. [1]) one proves at once the property

$$d(f([b])) \geq | [A]^C | \, d([b])$$

of f which is then called *normal* in [6, p. 102].

We next recall an equivalent definition of Step 1 in the interval Cholesky method.

DEFINITION 5.2. ([2]) *Let either* $[A] = ([a]_{11}) \in \boldsymbol{IR}^{1\times 1}$ *or*

$$[A] = \begin{pmatrix} [a]_{11} & [c]^T \\ [c] & [A'] \end{pmatrix} = [A]^T \in \boldsymbol{IR}^{n\times n}, \; n > 1, \; [c] \in \boldsymbol{IR}^{n-1},$$
$$[A'] \in \boldsymbol{IR}^{(n-1)\times(n-1)}.$$

(a) $\Sigma_{[A]} := [A'] - (1/[a]_{11}) \, [c][c]^T \in \boldsymbol{IR}^{(n-1)\times(n-1)}$ *is termed the* Schur complement *(of the* $(1,1)$ *entry* $[a]_{11}$*) provided* $n > 1$ *and* $0 \notin [a]_{11}$*. In the product* $[c][c]^T$ *we assume that* $[c]_i[c]_i$ *is evaluated as* $[c]_i^2$ (see (5.1)). $\Sigma_{[A]}$ *is not defined if* $n = 1$ *or if* $0 \in [a]_{11}$.

(b) *We call the pair* $([L],[L]^T)$ *the* Cholesky decomposition *of* $[A]$ *if* $0 < \underline{a}_{11}$ *and if either* $n = 1$ *and* $[L] = (\sqrt{[a]_{11}})$ *or*

$$[L] = \begin{pmatrix} \sqrt{[a]_{11}} & 0 \\ \dfrac{[c]}{\sqrt{[a]_{11}}} & [L'] \end{pmatrix}, \tag{5.5}$$

where $([L'], ([L'])^T)$ *is the Cholesky decomposition of* $\Sigma_{[A]}$ *provided that it exists.*

As was shown in [2] the matrix $[L]$ of the Cholesky method and that of the Definition 5.2(b) are identical.

The proof of our main result, Theorem 5.4, is heavily based on the following lemma.

LEMMA 5.3. *Let the Cholesky decomposition* $([L], [L]^T)$ *of* $[A] = [A]^T \in \boldsymbol{IR}^{n\times n}$ *exist, and let* $[B] = [B]^T \supseteq [A]$ *be such that for a suitable* $u > 0$ *we have*

$$q([A], [B])u < \langle [L] \rangle \langle [L]^T \rangle u \,. \tag{5.6}$$

Then the Cholesky method is feasible for $[B]$.

Proof by induction. The proof proceeds similarly as for Lemma 4.5.14 in [6].

Let $n = 1$. Then (5.6) implies $u > 0$. Again (5.6) together with $0 < \underline{a}_{11}$ yields

$$(\underline{a}_{11} - \underline{b}_{11})u \le q([A], [B])u < \langle [a]_{11} \rangle u = \underline{a}_{11} u \,,$$

hence

$$0 < \underline{b}_{11} u$$

follows. This shows $0 < \underline{b}_{11} = \langle [b]_{11} \rangle$ which proves the existence of $\mathrm{ICh}([B], [b])$ for $n = 1$.

Assume now that the statement is true for some dimension $n \ge 1$, and let (5.6) hold for

$$[A] = \begin{pmatrix} [a]_{11} & [c]^T \\ [c] & [A'] \end{pmatrix} \subseteq [B] = \begin{pmatrix} [b]_{11} & [d]^T \\ [d] & [B'] \end{pmatrix} \in \boldsymbol{IR}^{(n+1)\times(n+1)} \,. \tag{5.7}$$

We first show $\underline{b}_{11} > 0$. With

$$q([A], [B]) = (q_{ij}) = \begin{pmatrix} q_{11} & r^T \\ r & Q' \end{pmatrix} \tag{5.8}$$

we get from (5.6)

$$\sum_{j=1}^{n+1} q_{1j} u_j < \langle [a]_{11} \rangle \, u_1 - \sum_{j=2}^{n+1} |\, [a]_{1j} \,|\, u_j \,,$$

hence

$$\underline{a}_{11} - q_{11} = \langle [a]_{11} \rangle - q_{11} > \left\{ \sum_{j=2}^{n+1} (q_{1j} + |\, [a]_{1j} \,|)\, u_j \right\} / u_1 \ge 0 \,.$$

Together with (5.7) this implies $0 < \underline{b}_{11} = \langle [b]_{11} \rangle$, whence the Schur complement $\Sigma_{[B]} \supseteq \Sigma_{[A]}$ exists.

By our assumptions, the Schur complement $\Sigma_{[A]}$ has a Cholesky decomposition $([L'], [L']^T)$. If we can show that

$$q(\Sigma_{[A]}\,,\, \Sigma_{[B]})u' < \langle [L'] \rangle \langle [L']^T \rangle u' \tag{5.9}$$

holds for some vector $u' > 0$ then $\Sigma_{[B]}$ has a Cholesky decomposition, say $([\hat{L}'], [\hat{L}']^T)$, by the hypothesis of our induction, and with

$$[\hat{L}] := \begin{pmatrix} \sqrt{[b]_{11}} & O \\ \dfrac{[d]}{\sqrt{[b]_{11}}} & [\hat{L}'] \end{pmatrix},$$

we obtain the Cholesky decomposition $([\hat{L}], [\hat{L}]^T)$ of $[B]$.

To prove (5.9) we apply β from (2.1) componentwise, and use the notation from (5.8) as well as that of Lemma 2.2. We then get

(5.10)
$$\begin{aligned} q(\Sigma_{[A]}\,,\, \Sigma_{[B]}) &= q\left([A'] - [c][c]^T[a]_{11}^{-1}, [B'] - [d][d]^T[b]_{11}^{-1}\right) \\ &\le Q' + q([c][c]^T[a]_{11}^{-1}, [d][d]^T[b]_{11}^{-1}) \\ &= Q' - |\,[c][c]^T[a]_{11}^{-1}\,| + \beta([c][c]^T[a]_{11}^{-1}, [d][d]^T[b]_{11}^{-1}) \\ &= Q' - |\,[c]\,|\,|\,[c]^T\,|\,\langle [a]_{11}\rangle^{-1} + \beta([c][c]^T[a]_{11}^{-1}, [d][d]^T[b]_{11}^{-1}) \\ &\le Q' - |\,[c]\,|\,|\,[c]^T\,|\,\langle [a]_{11}\rangle^{-1} + \beta([c],[d]) \cdot \beta([c]^T,[d]^T) \cdot \beta([a]_{11}^{-1}, [b]_{11}^{-1}) \\ &= Q' - |\,[c]\,|\,|\,[c]^T\,|\,\langle [a]_{11}\rangle^{-1} + (|\,[c]\,| + r)(|\,[c]\,| + r)^T \beta([a]_{11}^{-1}, [b]_{11}^{-1})\,. \end{aligned}$$

We now want to apply Lemma 2.2 (b) on the last factor in (5.10). To this end we must show

$$\langle [a]_{11} \rangle \;>\; q([a]_{11}, [b]_{11}) = q_{11}\,. \tag{5.11}$$

Therefore, we set $u = \begin{pmatrix} u_1 \\ u' \end{pmatrix}$ in (5.6). With (5.5) and with the notation (5.8), we then obtain

$$\begin{pmatrix} q_{11} & r^T \\ r & Q' \end{pmatrix}\begin{pmatrix} u_1 \\ u' \end{pmatrix} < \begin{pmatrix} \sqrt{\langle [a]_{11}\rangle} & O \\ -\dfrac{|\,[c]\,|}{\sqrt{\langle [a]_{11}\rangle}} & \langle [L'] \rangle \end{pmatrix}\begin{pmatrix} \sqrt{\langle [a]_{11}\rangle} & -\dfrac{|\,[c]^T\,|}{\sqrt{\langle [a]_{11}\rangle}} \\ O & \langle [L']^T \rangle \end{pmatrix}\begin{pmatrix} u_1 \\ u' \end{pmatrix},$$

whence

$$q_{11}u_1 + r^T u' < \langle [a]_{11} \rangle u_1 - |\,[c]^T\,| u' \tag{5.12}$$

and

$$r u_1 + Q'u' < -|\,[c]\,| u_1 + |\,[c]\,|\,|\,[c]^T\,|\,\langle [a]_{11}\rangle^{-1} u' + \langle [L'] \rangle \langle [L']^T \rangle u'\,. \tag{5.13}$$

Since $u_1 > 0$, the inequality (5.12) implies (5.11), and Lemma 2.2(b) and (5.10) yield

$$q(\Sigma_{[A]}\,,\, \Sigma_{[B]}) \le Q' - |\,[c]\,|\,|\,[c]^T\,|\,\langle [a]_{11}\rangle^{-1} + (|\,[c]\,| + r)(|\,[c]\,| + r)^T (\langle [a]_{11}\rangle - q_{11})^{-1}\,.$$

Together with (5.11), (5.12), (5.13), this implies

$$\begin{aligned} q(\Sigma_{[A]}\,,\, \Sigma_{[B]})u' &\le Q'u' - |\,[c]\,|\,|\,[c]^T\,|\langle [a]_{11}\rangle^{-1}u' \\ &\quad + (|\,[c]\,| + r)(\langle [a]_{11}\rangle - q_{11})^{-1}(|\,[c]\,| + r)^T u' \\ &< -r u_1 - |\,[c]\,| u_1 + \langle [L'] \rangle\ \langle [L']^T \rangle u' \\ &\quad + (|\,[c]\,| + r)(\langle [a]_{11}\rangle - q_{11})^{-1}(\langle [a]_{11}\rangle u_1 - q_{11}u_1) \\ &= \langle [L'] \rangle \langle [L']^T \rangle u'. \end{aligned}$$

This proves (5.9) and terminates the induction. □

We are now ready to prove our main result.

THEOREM 5.4. *Let* $[A]$, $[B] \in \boldsymbol{IR}^{n\times n}$, $[A] = [A]^T$, $[B] = [B]^T$, *and suppose that* $\mathrm{ICh}([A],[b])$ *exists. If*

$$\rho(|\ [A]^C|\ q([A],[B])) < 1, \tag{5.14}$$

then the Cholesky method is feasible for $[B]$.

Proof. Let $Q := q([A],[B])$, $[C] := [A] + [-Q,Q]$. Then $[B] \subseteq [C]$, and $\mathrm{ICh}([B],[b])$ exists if $\mathrm{ICh}([C],[b])$ does. By (5.14) the inverse of $I - |\,[A]^C\,|\,Q$ exists and can be represented as Neumann series

$$(I - |\ [A]^C|Q)^{-1} = \sum_{k=0}^{\infty} (|\,[A]^C\,|Q)^k \geq 0\ .$$

With any $v \in \boldsymbol{R}^n$ satisfying $v > 0$ define

$$u := \left(I - |\,[A]^C\,|Q\right)^{-1}\ |\,[A]^C\,|v\ . \tag{5.15}$$

Since $|\,[A]^C\,| \geq 0$ and $\left(I - |\,[A]^C\,|Q\right)^{-1} \geq 0$ are regular each of their rows contains at least one positive entry. Therefore $|\,[A]^C\,|v > 0$ and $u > 0$. Now (5.15) yields

$$|\ [A]^C|Qu = u - |\,[A]^C\,|v,$$

whence

$$\begin{aligned} Qu &= \langle [L]\rangle\langle [L]^T\rangle u - v \\ &< \langle L\rangle\langle L\rangle^T u, \end{aligned}$$

with $([L],[L]^T)$ being the Cholesky decomposition of $[A]$. Hence, Lemma 5.3 guarantees the feasibility of the Cholesky method for $[C]$ and therefore also for $[B]$. □

We illustrate Theorem 5.4 by a simple example.

Example 5.5. Let

$$[B] := \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & [0,2] \\ 2 & [0,2] & 4 \end{pmatrix}.$$

Then $\langle [B]\rangle \cdot (1,1,1)^T = 0$, hence $\langle [B]\rangle$ is singular. In particular, $\langle [B]\rangle$ is not an M–matrix (which requires $\langle [B]\rangle^{-1} \geq 0$; cf. [2]), whence, by definition, $[B]$ is not an H–matrix. Therefore, Theorem 4.2 in [2] does not apply. Consider now

$$[A] := \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 1 \\ 2 & 1 & 4 \end{pmatrix} \subseteq [B]\ .$$

Since $\langle [A]\rangle$ is irreducibly diagonally dominant, the interval Cholesky method is feasible for $[A]$ by Corollary 4.3 (ii) in [2], for example. A simple computation yields

$$[L] = \begin{pmatrix} 2 & 0 & 0 \\ 1 & \sqrt{3} & 0 \\ 1 & 0 & \sqrt{3} \end{pmatrix}, \qquad \langle [L]\rangle^{-1} = \frac{\sqrt{3}}{6}\begin{pmatrix} \sqrt{3} & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

and

$$|\,[A]^C| = \langle [L]^T\rangle^{-1}\langle [L]\rangle^{-1} = \frac{1}{12}\begin{pmatrix} 5 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 4 \end{pmatrix}.$$

From

$$q([A],[B]) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

we get the matrix

$$|\,[A]^C|q([A],[B]) = \begin{pmatrix} 0 & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \end{pmatrix},$$

which has the eigenvalues $-\frac{1}{3},\ 0,\ \frac{1}{3}$. Therefore, Theorem 5.4 applies. The elements $[\hat{l}]_{ij}$ that result from the interval Cholesky method for $[B]$ are given by

$$[\hat{L}] = \begin{pmatrix} 2 & 0 & 0 \\ 1 & \sqrt{3} & 0 \\ 1 & [-1,1]/\sqrt{3} & [\sqrt{8},3]/\sqrt{3} \end{pmatrix}.$$

Our example also illustrates the following corollary.

COROLLARY 5.6. *Let the midpoint matrix $\check{A}$ of $[A] = [A]^T \in \boldsymbol{IR}^{n\times n}$ be positive definite, and assume that*

$$\rho\left(\frac{1}{2}|\check{A}^C|\ d([A])\right) < 1\ .$$

Then the interval Cholesky method is feasible for $[A]$.

Proof. Because of $[A] = [A]^T$, the matrix $\check{A}$ is symmetric. Since it is positive definite by assumption, the interval Cholesky method is feasible for $\check{A}$ when viewed as a point matrix. Taking into account $q(\check{A},[A]) = \frac{1}{2}d([A])$, the assertion is a direct consequence of Theorem 5.4. □

6. Concluding remarks. We stress the fact that the main purpose of this paper is to give criteria for the feasibility of the interval Cholesky method. If this feasibility is guaranteed—for example, this is the case if one of the criteria presented in this paper or in [2] holds—the question arises immediately how close the symmetric solution set S_{sym} is included. Especially, what is the relation between the results of applying the Gaussian algorithm (or some other method) and the interval Cholesky method, respectively? In [2] it was shown by simple examples that generally no comparison is possible. The examples from [2] can be generalized to arbitrary large dimensions $n > 2$ without any difficulties. Hence up to now it is not clear under which conditions on the given interval matrix the interval Cholesky method is superior to the interval Gaussian algorithm or vice versa. The investigation of this question and/or some statistics about the width of the bounds for systems of larger dimension will be part of further research.

We also mention that for a given real system a very careful analysis of the floating-point Cholesky decomposition was performed in [10]. If the matrix as well as the right-hand side are afflicted with tolerances then bounds are computed for the set of all solutions for data within tolerances.

Acknowledgments. The authors are grateful to two anonymous referees for a series of comments and remarks that improved the paper considerably.

REFERENCES

[1] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.

[2] G. Alefeld and G. Mayer, *The Cholesky method for interval data*, Linear Algebra Appl., 194 (1993), pp. 161–182.

[3] H. Beeck, *Über Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten*, Computing, 10 (1972), pp. 231–244.

[4] D. J. Hartfiel, *Concerning the solution set of $Ax = b$ where $P \leq A \leq Q$ and $p \leq b \leq q$*, Numer. Math., 35 (1980), pp. 355–359.

[5] C. Jansson, *Interval linear systems with symmetric matrices, skew-symmetric matrices and dependencies in the right-hand side*, Computing, 46 (1991), pp. 265–274.

[6] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.

[7] W. Oettli and W. Prager, *Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides*, Numer. Math., 6 (1964), pp. 405–409.

[8] J. Rohn, *Interval linear systems*, Freiburger Intervall-Berichte, 84/7 (1984), pp. 33–58.

[9] J. Rohn, *On nonconvexity of the solution set of a system of linear interval equations*, BIT, 30 (1989), pp. 161–165.

[10] S. M. Rump, *Inclusion of the solution of large linear systems with positive definite symmetric M-matrix*, in Computer Arithmetic and Enclosure Methods, L. Atanassova and J. Herzberger, eds., Elsevier, Amsterdam, 1992, pp. 339–347.

[11] G. Strang, *Linear Algebra and Its Applications*, Third ed., Harcourt Brace Jovanovich, San Diego, 1988.