Interval Computations
No 1, 1994

Suboptimal Enclosures for the Interval Buneman Algorithm for Arbitrary Block Dimension

Hartmut Schwandt

The interval arithmetic Buneman algorithm is a "fast solver" for a class of block tridiagonal systems with interval coefficients. In [11], we have introduced a modification for arbitrary block dimension. In the present paper, we discuss enclosure properties depending on the block dimension of the coefficient matrix in selected cases for which optimal enclosures cannot be expected.

Субоптимальные включения в интервальном алгоритме Бунемана при произвольной размерности блоков

Х. Швандт

Алгоритм Бунемана, использующий интервальную арифметику, представляет собой быстрый решатель для определенного класса блочных трехдиагональных систем с интервальными коэффициентами. В [11] нами предложена модификация этого алгоритма, допускающая произвольную размерность блока. В настоящей работе рассматриваются включающие свойства алгоритма в зависимости от блочной размерности матрицы коэффициентов для некоторых случаев, в которых невозможно получение оптимальных включающих множеств.

1 Introduction

The interval arithmetic analogue of one of the "classical" fast Poisson solves, the Buneman algorithm [4, 5], is dedicated to the treatment of systems $(\mathbf{M}, \mathbf{b})$ of linear form with a block tridiagonal matrix $\mathbf{M} = (-\mathrm{S}, \mathbf{A}, -\mathrm{T})$, where S, T are real $p \times p$ matrices, $p \in \mathbb{N}$, where $\mathbf{A}$ is a $p \times p$ matrix with real compact intervals as coefficients and where $\mathbf{b}$ is a vector with N interval components, $N = pq$, $q \in \mathbb{N}$. The specific form of $\mathbf{M}$ results from necessary commutativity conditions, which are satisfied in the main applications of *IBU* as a "linear" solver in Newton-like interval methods for nonlinear systems of equations resulting from difference methods for almost linear partial BVP [8, 9]. In contrast to the solution of linear point (i.e. noninterval) systems, an interval method like the interval Buneman algorithm *IBU* yields an interval vector including the set of solutions $SOL(\mathbf{M}, \mathbf{b}) := \{x \in \mathbb{R}^N | \mathrm{Mx} = \mathrm{b}, \mathrm{M} \in \mathbf{M}, \mathrm{b} \in \mathbf{b}\} \subseteq \mathbf{x} := IBU(\mathbf{M}, \mathbf{b})$. As, in general, $SOL(\mathbf{M}, \mathbf{b})$ is not an interval vector, we can at most expect $\mathbf{x}$ to be an optimal enclosure, i.e. the tightest interval vector including this set [1, 3]. In a previous paper [11], we have introduced a modification *IBUD* for arbitrary block dimension. The original Buneman algorithm is defined for a block dimension of the form $q = 2^{r+1} - 1$ [4, 5]. A more flexible choice of q is, however, often desirable if $IBU(D)$ is applied as part of other algorithms. As an example we mention a domain decomposition method [12] where *IBUD* is needed for the treatment of systems on rectangular subdomains whose number and size just depend on the restriction on q. Suboptimal enclosures which are discussed in the present paper are admissible if only a reduced precision of the results is required or if their difference to an optimal enclosure can be neglected relative to the machine precision. Newton-like methods in which $IBU(D)$ is used as a "linear" solver can be typical applications for the latter case. In analogy to the corresponding noninterval algorithm which is applicable to point systems $\mathrm{Mx} = \mathrm{b}$, $\mathrm{M} = (-\mathrm{S}, \mathrm{A}, -\mathrm{T})$ of arbitrary block dimension and arbitrary block size [13], the interval method *IBUD* can be carried out for interval systems for any value of q, but optimal enclosures can be guaranteed under suitable conditions only for particular values for q while the size of the blocks can be arbitrary. In the present paper, we modify *IBUD* for an efficient treatment of a nonoptimal block dimension and we develop an estimate for the additional width by which it can be decided whether acceptable enclosures can be obtained. The theoretical results are illustrated by numerical examples.

2 Notation and basic results

For the introduction of interval arithmetic and interval methods, we refer to [1], for example. We denote real numbers by $a, \dots, z$ or by Greek letters, real point vectors and matrices by $\mathrm{a}, \dots, \mathrm{z}$ and $\mathrm{A}, \dots, \mathrm{Z}$, resp., intervals by $A, \dots, Z$, real interval vectors and matrices by $\mathbf{a}, \dots, \mathbf{z}$ and $\mathbf{A}, \dots, \mathbf{Z}$, resp. We consider real, compact intervals: $A = [\mathrm{i}(A), \mathrm{s}(A)]$, $\mathrm{i}(A) := \min\{a | a \in A\}$, $\mathrm{s}(A) := \max\{a | a \in A\}$. For interval vectors and matrices we use the notation $\mathbf{a} = (A_i)_{i=1}^N = [\mathrm{i}(\mathbf{a}), \mathrm{s}(\mathbf{a})] = \big([\mathrm{i}(A_i), \mathrm{s}(A_i)]\big)_{i=1}^N$ and $\mathbf{A} = (A_{i,j})_{i,j=1}^N = [\mathrm{i}(\mathbf{A}), \mathrm{s}(\mathbf{A})] = \big([\mathrm{i}(A_{i,j}), \mathrm{s}(A_{i,j})]\big)_{i,j=1}^N$, where the bounds are defined componentwise. We use the componentwise ordering: $\mathrm{A} \leq \mathrm{B} \Leftrightarrow \forall i, j \in \{1, \dots, N\} : a_{i,j} \leq b_{i,j}$. $\mathrm{I}(\mathbb{R})$, $\mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big)$, $\mathrm{M}_{NN}(\mathbb{R})$, $\mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big)$ denote the sets of real compact intervals, N-dimensional interval vectors, $N \times N$ point and interval matrices, resp. The absolute value, midpoint, width of an interval are defined by $|A| = \max\{|\mathrm{i}(A)|, |\mathrm{s}(A)|\}$, $\mathrm{m}(A) = \big(\mathrm{i}(A) + \mathrm{s}(A)\big)/2$, $\mathrm{d}(A) = \big(\mathrm{s}(A) - \mathrm{i}(A)\big)/2$.

We further define $[x] := \min\{i \in \mathbb{N} | i \leq x\}$ for arbitrary $x \in \mathbb{R}^+$.

A real point matrix A is called an M matrix if $a_{i,j} \leq 0$ for $i \neq j$, if A^{-1} exists and if $\mathrm{A}^{-1} \geq \mathrm{O}$ (see [14], e.g.). An interval matrix $\mathbf{A}$ is called an interval M matrix if all $\mathrm{A} \in \mathbf{A}$ are M matrices. By $\rho(\mathrm{A})$ and $\mathrm{spec}(\mathrm{A})$, we denote the spectral radius and the spectrum of a quadratic real matrix A.

Lemma 2.1. *If* $\mathrm{A}, \mathrm{B} \geq \mathrm{O}$, *if all eigenvectors of* A *are eigenvectors of* B *and if* A *has an eigenvector* $\mathrm{c} > 0$, *then* $\|\mathrm{A}\|_{\mathrm{c,c}} \leq \rho(\mathrm{A})$, $\|\mathrm{B}\|_{\mathrm{c,c}} \leq \rho(\mathrm{B})$ *in the matrix norm corresponding to the monotone vector norm* $\|\mathrm{x}\|_{\mathrm{c}} := \max\limits_{1 \leq i \leq N} \left\{ \dfrac{|x_i|}{c_i} \right\}$.

Proof. The relation $\mathrm{Ac} = \lambda \mathrm{c}$ with $\lambda \in \mathrm{spec}(\mathrm{A})$ is equivalent to $\lambda = \frac{1}{c_i} \sum\limits_{j=1}^{N} a_{i,j} c_j > 0 \;\forall\, 1 \leq i \leq N$. For arbitrary $\mathrm{x} \in \mathbb{R}^N$ this implies

$$\begin{aligned}\|\mathrm{Ax}\|_{\mathrm{c}} = \max_{1\le i\le N}\left\{\frac{1}{c_i}\left|\sum_{j=1}^{N} a_{i,j}x_j\right|\right\} &\le \max_{1\le i\le N}\left\{\frac{1}{c_i}\sum_{j=1}^{N} a_{i,j}c_j\frac{|x_j|}{c_j}\right\}\\ &\le \lambda\|\mathrm{x}\|_{\mathrm{c}} \le \rho(\mathrm{A})\|\mathrm{x}\|_{\mathrm{c}}.\end{aligned}$$

c is also an eigenvector of B, therefore there exists $\mu \in \mathrm{spec}(\mathrm{B})$ such that $\mathrm{Bc} = \mu\mathrm{c}$, hence $\mu > 0$ and $\|\mathrm{Bx}\|_{\mathrm{c}} \le \mu\|\mathrm{x}\|_{\mathrm{c}} \le \rho(\mathrm{B})\|\mathrm{x}\|_{\mathrm{c}}$ as above. □

Corollary 2.2. *Under the conditions of Lemma 2.1, the following assertions hold:*
(a) $\mathrm{A} \ge \mathrm{O}$ *irreducible* $\Rightarrow \|\mathrm{A}\|_{\mathrm{c,c}} = \rho(\mathrm{A})$;
(b) $\mathrm{Bc} = \rho(\mathrm{B})\mathrm{c} \Rightarrow \|\mathrm{B}\|_{\mathrm{c,c}} = \rho(\mathrm{B})$.

Proof. If $\mathrm{A} \ge \mathrm{O}$ is irreducible, the theorem of Perron-Frobenius implies the existence of a unique eigenvector $\mathrm{c} > 0$ to the eigenvalue $\rho(\mathrm{A})$ of A ([14], Th. 2.1). The second assertion is obvious from the proof of Lemma 2.1. □

For elementary rules of interval analysis we refer to [1], for example. For convenience we mention some of these rules which are particularly significant for the present paper.

$$\begin{aligned}&A, B, C, D \in \mathrm{I}(\mathbb{R}), \quad A \subseteq B, C \subseteq D, \quad \bullet \in \{+,-,*,/\}\\ &\Rightarrow A \bullet B \subseteq C \bullet D;\end{aligned} \tag{1}$$

$$\begin{aligned}&X = -X \Rightarrow X = \frac{1}{2}\mathrm{d}(X)[-1,1];\\ &\mathbf{x} = -\mathbf{x} \Rightarrow \mathrm{A}\mathbf{x} = -\mathrm{A}\mathbf{x} = |\mathrm{A}|\mathbf{x};\end{aligned} \tag{2}$$

$$\mathrm{A}, \mathrm{B} \ge \mathrm{O} \Rightarrow \forall \mathbf{x} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big): \quad \mathrm{A}(\mathrm{B}\mathbf{x}) = (\mathrm{AB})\mathbf{x}; \tag{3}$$

$$\mathrm{i}(\mathbf{A}), \mathrm{i}(\mathbf{B}) \ge \mathrm{O} \Rightarrow \forall \mathbf{x} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big): \ (\mathbf{A} + \mathbf{B})\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}; \tag{4}$$

$$\mathrm{i}(A) \ge 0 \Rightarrow \mathrm{d}(A) \le \mathrm{s}(A); \tag{5}$$

$$\mathrm{A} \in \mathrm{M}_{NN}(\mathbb{R}) \Rightarrow \forall \mathbf{x} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big): \quad \mathrm{d}(\mathrm{A}\mathbf{x}) = |\mathrm{A}|\mathrm{d}(\mathbf{x}); \tag{6}$$

$$0 \in A \Rightarrow \forall B \in \mathrm{I}(\mathbb{R}): \quad 0 \in AB; \tag{7}$$

$$\begin{aligned}&\mathrm{i}(A), \mathrm{i}(B) \ge 0 \Rightarrow \mathrm{i}(AB) = \mathrm{i}(A)\mathrm{i}(B) \ge 0;\\ &\mathrm{s}(A) \le 0 \le \mathrm{i}(B) \Rightarrow \mathrm{s}(AB) = \mathrm{s}(A)\mathrm{i}(B) \le 0.\end{aligned} \tag{8}$$

The following elementary properties are relevant for the use of any interval method *LES* for the treatment of systems of linear form $(\mathbf{A}, \mathbf{y})$.

$$\forall \mathbf{A} \in \mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big),\ \mathbf{y},\mathbf{z} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big):\quad \mathit{LES}(\mathbf{A},\mathbf{y}+\mathbf{z}) \subseteq \mathit{LES}(\mathbf{A},\mathbf{y}) + \mathit{LES}(\mathbf{A},\mathbf{z}); \tag{9}$$

$$\forall \mathrm{A} \in \mathrm{M}_{NN}(\mathbb{R}),\ \mathbf{y},\mathbf{z} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big):\quad \mathit{LES}(\mathrm{A},\mathbf{y}+\mathbf{z}) = \mathit{LES}(\mathrm{A},\mathbf{y}) + \mathit{LES}(\mathrm{A},\mathbf{z}); \tag{10}$$

$$\forall \mathbf{A} \in \mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big),\ \mathbf{y} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big), \alpha \in \mathbb{R}:\quad \mathit{LES}(\mathbf{A},\alpha\mathbf{y}) = \alpha\mathit{LES}(\mathbf{A},\mathbf{y}); \tag{11}$$

$$\forall \mathbf{A},\mathbf{B} \in \mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big),\ \mathbf{y},\mathbf{z} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big):\quad \mathbf{A} \subseteq \mathbf{B},\ \mathbf{y} \subseteq \mathbf{z} \Rightarrow \mathit{LES}(\mathbf{A},\mathbf{y}) \subseteq \mathit{LES}(\mathbf{B},\mathbf{z}); \tag{12}$$

$$\mathbf{A}_n \to \mathbf{A},\ \mathbf{y}_n \to \mathbf{y}\ (n \to \infty) \Rightarrow \quad \mathit{LES}(\mathbf{A}_n,\mathbf{y}_n) \to \mathit{LES}(\mathbf{A},\mathbf{y})\ (n \to \infty). \tag{13}$$

The tightest interval vector which still includes the set of solutions $\mathit{SOL}(\mathbf{A},\mathbf{y})$ is called an *optimal enclosure* of $\mathit{SOL}(\mathbf{A},\mathbf{y})$. The following cases are typical for optimal enclosures:

If $\mathbf{A} \in \mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big)$ is an interval M matrix, then

$$\begin{array}{lll} \text{a)} & \mathbf{A} \equiv \mathrm{A} \in \mathrm{M}_{NN}(\mathbb{R}) \Rightarrow \forall \mathbf{y} \in \mathrm{V}_N\big(\mathrm{I}(\mathbb{R})\big): & \\ & & \mathit{LES}(\mathrm{A},\mathbf{y}) = \big[\mathrm{A}^{-1}\mathrm{i}(\mathbf{y}), \mathrm{A}^{-1}\mathrm{s}(\mathbf{y})\big] = \mathrm{A}^{-1}\mathbf{y} \\ \text{b)} & 0 \in \mathbf{y} \quad \Rightarrow & \mathit{LES}(\mathbf{A},\mathbf{y}) = \big[\big(\mathrm{i}(\mathbf{A})\big)^{-1}\mathrm{i}(\mathbf{y}), \big(\mathrm{i}(\mathbf{A})\big)^{-1}\mathrm{s}(\mathbf{y})\big]; \\ & & \mathrm{o} \in \mathit{LES}(\mathbf{A},\mathbf{y}) \\ \text{c)} & 0 \le \mathrm{i}(\mathbf{y}) \quad \Rightarrow & \mathit{LES}(\mathbf{A},\mathbf{y}) = \big[\big(\mathrm{s}(\mathbf{A})\big)^{-1}\mathrm{i}(\mathbf{y}), \big(\mathrm{i}(\mathbf{A})\big)^{-1}\mathrm{s}(\mathbf{y})\big]; \\ & & \mathrm{i}\big(\mathit{LES}(\mathbf{A},\mathbf{y})\big) \ge \mathrm{o} \\ \text{d)} & \mathrm{s}(\mathbf{y}) \le 0 \quad \Rightarrow & \mathit{LES}(\mathbf{A},\mathbf{y}) = \big[\big(\mathrm{i}(\mathbf{A})\big)^{-1}\mathrm{i}(\mathbf{y}), \big(\mathrm{s}(\mathbf{A})\big)^{-1}\mathrm{s}(\mathbf{y})\big]; \\ & & \mathrm{s}\big(\mathit{LES}(\mathbf{A},\mathbf{y})\big) \le \mathrm{o}. \end{array} \tag{2.14}$$

3 The algorithm

The subsequent discussion is based on an algorithm for *IBUD* which has been derived in [11]. For convenience, we repeat its formulation and we briefly recall the main properties. We assume:

> *A system of linear form* $(\mathbf{M},\mathbf{b})$, *where* $\mathbf{b} \in V_N\big(\mathrm{I}(\mathbb{R})\big)$, $\mathbf{M} = (-\mathrm{S},\mathbf{A},-\mathrm{T}) \in \mathrm{M}_{NN}\big(\mathrm{I}(\mathbb{R})\big)$ *block tridiagonal with* q *block rows,* $\mathrm{S},\mathrm{T} \in \mathrm{M}_{pp}(\mathbb{R})$, $\mathbf{A} \in \mathrm{M}_{pp}\big(\mathrm{I}(\mathbb{R})\big)$, $N = pq$, $p,q \in \mathbb{N}$. (3.1)

The system under consideration is one with a block tridiagonal interval matrix with, except for the first and last, identical block rows. In [11], we followed principally the ideas of the Buneman algorithm. Starting from a point system $\mathrm{Mx} = \mathrm{y}$ with $\mathrm{M} \in \mathbf{M}$, $\mathrm{b} \in \mathbf{b}$, block cyclic reduction first yields reduced systems $\mathrm{M}^{(r)}\mathrm{x}^r = \mathrm{b}^r$, $\mathrm{M}^{(r)} \in \mathrm{M}_{pj_r,pj_r}(\mathbb{R})$, $\mathrm{x}^r, \mathrm{b}^r \in \mathrm{V}_{pj_r}(\mathbb{R})$, $0 \le r \le r_q$, $r_q = [\log_2(q)] + 1$, where

$$\mathrm{M}^{(r)} = \begin{pmatrix} \mathrm{A}^{(r)} & -\mathrm{T}^{(r)} & & & \\ -\mathrm{S}^{(r)} & \mathrm{A}^{(r)} & -\mathrm{T}^{(r)} & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathrm{S}^{(r)} & \mathrm{A}^{(r)} & -\mathrm{T}^{(r)} \\ & & & -\mathrm{S}^{(r)} & \mathrm{B}^{(r)}\left(\mathrm{C}^{(r)}\right)^{-1} \end{pmatrix},$$

$$\mathrm{x}^r = \begin{pmatrix} \mathrm{x}_{2^r} \\ \mathrm{x}_{2^{r+1}} \\ \vdots \\ \mathrm{x}_{j_r-2^r} \\ \mathrm{x}_{j_r} \end{pmatrix}, \qquad \mathrm{b}^r = \begin{pmatrix} \mathrm{b}^r_{2^r} \\ \mathrm{b}^r_{2^{r+1}} \\ \vdots \\ \mathrm{b}^r_{j_r-2^r} \\ \mathrm{b}^r_{j_r} \end{pmatrix}. \tag{2}$$

In the sequel, superscripts on real numbers, intervals, real and interval vectors and matrices indicate reduction or solution steps of (block) cyclic reduction algorithms. Parentheses are set to avoid any confusion with powers, except for vectors: $\lambda_i^{(r)}$, y_j^r, $X^{(r)}$, $\mathbf{A}^{(r)}$, e.g.

An arbitrary block dimension q can be handled due to the introduction of the matrices $\mathrm{B}^{(r)}$, $\mathrm{C}^{(r)}$ in the last block equation [13]. The instability of this block cyclic reduction procedure is avoided in the Buneman algorithm and its variants by replacing the b_j^r by auxiliary vectors and by a factorization of the matrices $\mathrm{A}^{(r)}$, $\mathrm{B}^{(r)}$, $\mathrm{C}^{(r)}$. For $0 \le r \le r_q$, we obtain

$$\begin{aligned} \mathrm{A}^{(r)} &= \prod_{i=1}^{2^r} \left(\mathrm{A} - \alpha_i^{(r)}\mathrm{U}\right), & \alpha_i^{(r)} &= 2\cos\left(\tfrac{2i-1}{2^{r+1}}\pi\right), \\ \mathrm{B}^{(r)} &= \prod_{i=1}^{k_r} \left(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\right), & \lambda_i^{(r)} &= 2\cos\left(\tfrac{i}{k_r+1}\pi\right), \\ \mathrm{C}^{(r)} &= \prod_{i=1}^{l_r} \left(\mathrm{A} - \mu_i^{(r)}\mathrm{U}\right), & \mu_i^{(r)} &= 2\cos\left(\tfrac{i}{l_r+1}\pi\right) \end{aligned} \tag{3.3}$$

where $k_0 = 1$, $l_0 = 0$, and

$$k_{r+1} = \begin{cases} k_r + 2^r \\ k_r + 2^{r+1}. \end{cases} \qquad l_{r+1} = \begin{cases} l_r & \text{if } q_r \text{ even,} \\ k_r & \text{if } q_r \text{ odd.} \end{cases} \tag{3.4}$$

Starting from $\mathrm{M} \in \mathbf{M}$, $\mathrm{b} \in \mathbf{b}$, the interval arithmetic evaluation [1] of the resulting real point system, mainly by the repeated application of (2.1) and (2.12), leads to the following algorithm (3.5) from [11] to which we refer in the sequel. In (3.5) $L\tilde{E}S$ denotes a partial algorithm for the treatment of partial systems with the matrices $\mathbf{A}^{(r)}$. $L\tilde{E}S$ will be defined in detail after (3.5).

Modified *IBU* for arbitrary block dimension *(IBUD)* (3.5)

$q_0 = j_0 = q$, $r_q = [\log_2(q)] + 1$; $r_s := \min\{r | 0 \le r \le r_q, q_r \text{ even}\}$
$\mathrm{S}^{(0)} = \mathrm{S}$; $\mathrm{T}^{(0)} = \mathrm{T}$; $\mathrm{U}^{(0)} = \mathrm{U} = \sqrt{\mathrm{ST}}$; $\mathbf{x}_0 = \mathbf{x}_{q+1} = \mathrm{o}$;

reduction phase

```
for $j := 1$ to $q$ do     $\mathbf{p}_j^0 = \mathrm{o}$; $\mathbf{q}_j^0 = \mathbf{b}_j$;
for $r := 0$ to $r_s - 1$ do
    $\mathrm{S}^{(r+1)} = \left(\mathrm{S}^{(r)}\right)^2$; $\mathrm{T}^{(r+1)} = \left(\mathrm{T}^{(r)}\right)^2$; $\mathrm{U}^{(r+1)} = \mathrm{S}^{(r)}\mathrm{T}^{(r)}$;
    $q_{r+1} = q_r \operatorname{div} 2$; $j_{r+1} = j_r - 2^r$; $k_{r+1} = k_r + 2^{r+1}$; $l_{r+1} = k_r$;
    for $j := 2^{r+1}$ step $2^{r+1}$ to $j_{r+1}$ do
        $\mathbf{p}_j^{r+1} = \mathbf{p}_j^r + L\tilde{E}S\left(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{p}_{j-2^r}^r + \mathrm{T}^{(r)}\mathbf{p}_{j+2^r}^r + \mathbf{q}_j^r\right)$;          (3.6a)
        $\mathbf{q}_j^{r+1} = \mathrm{S}^{(r)}\mathbf{q}_{j-2^r}^r + \mathrm{T}^{(r)}\mathbf{q}_{j+2^r}^r + 2\mathrm{U}^{(r+1)}\mathbf{p}_j^{r+1}$
for $r := r_s$ to $r_q - 1$ do
    $\mathrm{S}^{(r+1)} = \left(\mathrm{S}^{(r)}\right)^2$; $\mathrm{T}^{(r+1)} = \left(\mathrm{T}^{(r)}\right)^2$; $\mathrm{U}^{(r+1)} = \mathrm{S}^{(r)}\mathrm{T}^{(r)}$; $q_{r+1} = q_r \operatorname{div} 2$;
    if $q_r$ even then
        $j_{r+1} = j_r$; $k_{r+1} = k_r + 2^r$; $l_{r+1} = l_r$;
        $\mathbf{p}_{j_{r+1}}^{r+1} = \mathbf{p}_{j_{r+1}}^r + L\tilde{E}S\left(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\left(\mathrm{S}^{(r)}\mathbf{p}_{j_{r+1}-2^r}^r + \mathbf{q}_{j_{r+1}}^r\right)\right)$;          (3.6b)
        $\mathbf{q}_{j_{r+1}}^{r+1} = \mathrm{S}^{(r)}\mathbf{q}_{j_{r+1}-2^r}^r + \mathrm{U}^{(r+1)}\mathbf{p}_{j_{r+1}}^{r+1}$
    else
        $j_{r+1} = j_r - 2^r$; $k_{r+1} = k_r + 2^{r+1}$; $l_{r+1} = k_r$;
        $\mathbf{p}^{r+1}_{j_{r+1}} = \mathbf{p}^{r}_{j_{r+1}} + \widetilde{LES}\Big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{p}^{r}_{j_{r+1}-2^r} + \mathrm{T}^{(r)}\mathbf{p}^{r}_{j_{r+1}+2^r} + \mathbf{q}^{r}_{j_{r+1}}\Big)\,;$   (3.6c)
        $\mathbf{q}^{r+1}_{j_{r+1}} = \mathrm{S}^{(r)}\mathbf{q}^{r}_{j_{r+1}-2^r} + \mathrm{U}^{(r+1)}\mathbf{p}^{r+1}_{j_{r+1}}$
            $+ \widetilde{LES}\Big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{A}^{(r)}\Big(\mathrm{T}^{(r)}\mathbf{q}^{r}_{j_{r+1}+2^r} + \mathrm{U}^{(r+1)}\mathbf{p}^{r+1}_{j_{r+1}}\Big)\Big);$
    for $j := 2^{r+1}$ step $2^{r+1}$ to $j_{r+1} - 2^{r+1}$ do
        $\mathbf{p}^{r+1}_{j} = \mathbf{p}^{r}_{j} + \widetilde{LES}\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{p}^{r}_{j-2^r} + \mathrm{T}^{(r)}\mathbf{p}^{r}_{j+2^r} + \mathbf{q}^{r}_{j}\big)\,;$
        $\mathbf{q}^{r+1}_{j} = \mathrm{S}^{(r)}\mathbf{q}^{r}_{j-2^r} + \mathrm{T}^{(r)}\mathbf{q}^{r}_{j+2^r} + 2\mathrm{U}^{(r+1)}\mathbf{p}^{r+1}_{j}$
```

solution phase

```
if $r_s = r_q$ then $\mathbf{x}_{2^{r_q}} = \mathbf{p}^{r_q}_{2^{r_q}} + \widetilde{LES}\big(\mathbf{A}^{(r_q)}, \mathbf{q}^{r_q}_{2^{r_q}}\big)$
    else $\mathbf{x}_{2^{r_q}} = \mathbf{p}^{r_q}_{2^{r_q}} + \widetilde{LES}\big(\mathbf{B}^{(r_q)}, \mathbf{C}^{(r_q)}\mathbf{q}^{r_q}_{2^{r_q}}\big)\,;$
for $r := r_q - 1$ step $-1$ to $\min\{r_s, r_q - 1\} + 1$ do
    for $j := 2^r$ step $2^{r+1}$ to $j_r - 2^r$ do
        $\mathbf{x}_j = \mathbf{p}^{r}_{j} + \widetilde{LES}\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{x}_{j-2^r} + \mathrm{T}^{(r)}\mathbf{x}_{j+2^r} + \mathbf{q}^{r}_{j}\big)\,;$
    if $q_r$ odd then
        $\mathbf{x}_{j_r} = \mathbf{p}^{r}_{j_r} + \widetilde{LES}\Big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\big(\mathrm{S}^{(r)}\mathbf{x}_{j_r-2^r} + \mathbf{q}^{r}_{j_r}\big)\Big);$
for $r := \min\{r_s, r_q - 1\}$ step $-1$ to $0$ do
    for $j := 2^r$ step $2^{r+1}$ to $j_r$ do
        $\mathbf{x}_j = \mathbf{p}^{r}_{j} + \widetilde{LES}\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{x}_{j-2^r} + \mathrm{T}^{(r)}\mathbf{x}_{j+2^r} + \mathbf{q}^{r}_{j}\big)\,.$
```

For the solution of the partial systems, we define the abstract method $\widetilde{LES}$ as follows. Instead of $\mathbf{A}^{(r)}$, $\mathbf{B}^{(r)}$, $\mathbf{C}^{(r)}$, we use interval extensions [1] $\mathbf{A} - \alpha\mathrm{U}$ of the matrices $\mathrm{A} - \alpha\mathrm{U}$. *LES* denotes any interval method which is applicable to coefficient matrices $\mathrm{A} - \alpha\mathrm{U}$. Typical examples are the interval Gauss algorithm *IGA* [1], interval arithmetic cyclic reduction *ICR* for tridiagonal systems [10], the interval Cholesky algorithm [2] or *IBU*(*D*) itself in discretizations of three-dimensional problems.

```
    $\mathbf{z} = \widetilde{LES}\big(\mathbf{A}^{(r)}, \mathbf{y}\big)$ :   (3.7)
        $\mathbf{z}_0 = \mathbf{y}$
        for $i := 1$ to $2^r$ do $\mathbf{z}_i = LES\big(\mathbf{A} - \alpha^{(r)}_i\mathrm{U}, \mathbf{z}_{i-1}\big)$
        $\mathbf{z} := \mathbf{z}_{2^r}$
```

```
$\mathbf{z} = L\tilde{E}S\left(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{y}\right)$ :                                (3.8)
    $\mathbf{z}_0 = \mathbf{y}$
    for $i := 1$ to $l_r$ do
        $\mathbf{z}_i = \mathbf{z}_{i-1} + \left(\lambda_i^{(r)} - \mu_i^{(r)}\right)\mathrm{U}LES\left(\mathbf{A} - \lambda_i^{(r)}\mathrm{U}, \mathbf{z}_{i-1}\right)$
    for $i := l_r + 1$ to $k_r$ do $\mathbf{z}_i = LES\left(\mathbf{A} - \lambda_i^{(r)}\mathrm{U}, \mathbf{z}_{i-1}\right)$
    $\mathbf{z} := \mathbf{z}_{k_r}$.
```

As already mentioned in [11], nonoptimal enclosures are caused by step (3.6c). Therefore, we have determined in that paper all values of the block dimension q for which step (3.6c) is never used. In the present context, our goal is to consider nonoptimal values of q. We use the following algorithm for (3.6c) which will be justified in the next section:

```
$\mathbf{z} = L\tilde{E}S\left(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{A}^{(r)}\mathbf{y}\right)$ :
    $\mathbf{z}_0 = \mathbf{y}$
    for $i := 1$ to $k_r - 1$ do
        $\mathbf{z}_i = \mathbf{z}_{i-1} + \left(\lambda_i^{(r)} - \gamma_{j(i)}^{(r)}\right)\mathrm{U}LES\left(\mathbf{A} - \lambda_i^{(r)}\mathrm{U}, \mathbf{z}_{i-1}\right)$          (3.9)
    $\mathbf{z}_{k_r} = \mathbf{z}_{k_r-1} + \left(\lambda_{k_r}^{(r)} - \gamma_{j(k_r)}^{(r)}\right)\mathrm{U}LES\left(\mathbf{A} - \lambda_{k_r}^{(r)}\mathrm{U}, \mathbf{z}_{k_r-1}\right)$
    $\mathbf{z} := \mathbf{z}_{k_r}$
```

where

$$\left(\gamma_i^{(r)}\right)_{i=1}^{k_r} := \text{merge}\left\{\left(\mu_i^{(r)}\right)_{i=1}^{l_r}, \left(\alpha_i^{(r)}\right)_{i=1}^{2^r} | \gamma_{i+1}^{(r)} \le \gamma_i^{(r)}\right\}, \gamma_1^{(r)} = \alpha_1^{(r)} \tag{3.10}$$

i.e. the $\gamma_i^{(r)}$ are ordered by size. In the sequel, we use the ordering

$$j(i) = i + 1 \quad (1 \le i \le k_r - 1); \quad j(k_r) = 1 \tag{3.11}$$

which will be discussed in the next section under the aspect of controllable enclosures. Principally, any ordering is admissible. Note that we have shown in [11] for (3.8) that

$$\forall r \in \{1, \ldots, r_q\} \quad \forall i \in \{1, \ldots, l_r\} : \quad \lambda_i^{(r)} \ge \mu_i^{(r)}. \tag{3.12}$$

We summarize the main properties of *IBUD* from [11] in the following theorem. α is defined as one of the roots $\alpha_i^{(r)}$, $\lambda_i^{(r)}$, $\mu_i^{(r)}$ of (3.3).

Theorem 3.1. *Assume (3.1). Under the conditions*

(a) $\forall \mathrm{A} \in \mathbf{A}$: $\mathrm{AS} = \mathrm{SA}$, $\mathrm{AT} = \mathrm{TA}$, *and* $\mathrm{ST} = \mathrm{TS}$

(b) $\mathrm{U} = \sqrt{\mathrm{ST}} \geq \mathrm{O}$, *i.e.* $\mathrm{U}^2 = \mathrm{ST}$, *and* U^{-1} *exists*

(c) LES is an interval method for systems with the coefficient structure of $(\mathbf{A} - \alpha\mathrm{U}, \mathbf{y})$

(d) LES is applicable to interval M *matrices and satisfies (2.9)–(2.13)*

(e) LES yields optimal enclosures for $\mathit{SOL}(\mathbf{A} - \alpha\mathrm{U}, \mathbf{y})$ *according to (2.14,* a–d*)*

(f) $\mathbf{M}$ *and* $\mathbf{A} + \alpha_1^{r_q}\mathrm{U}$, $\mathbf{A} - \alpha_1^{r_q}\mathrm{U}$ *are interval* M *matrices,*

the following assertions hold:

1) IBUD can be applied to $(\mathbf{M}, \mathbf{b})$ *and satisfies (2.9)–(2.13)*

2) $\mathit{SOL}(\mathbf{M}, \mathbf{b}) \subseteq \mathit{LES}(\mathbf{M}, \mathbf{b})$

3) $\forall r \in \{0, \ldots, r_s - 1\} : \mathrm{A}^{(r)} = \mathrm{B}^{(r)} \left(\mathrm{C}^{(r)}\right)^{-1}$*; IBU and IBUD coincide in these steps.*

If, in addition,

(g) $q = 2^n(2^m + 1) - 1$, $n, m \in \mathbb{N}$,

then

4) IBUD yields optimal enclosures of (3.1) according to (2.14, a–d*).*

The assertions 1 and 2 also hold for interval H matrices [11]. Note that for $\mathbf{M} = (-\mathrm{S}, \mathbf{A}, -\mathrm{T})$ condition (f) implies $\mathrm{S}, \mathrm{T} \geq \mathrm{O}$, $\mathrm{U} \geq \mathrm{O}$. (b) is satisfied, for example, for $\mathrm{S} = \mathrm{T}$, T invertible, or $\mathrm{S} = \alpha\mathrm{I}$, $\mathrm{T} = \beta\mathrm{I}$, $\alpha, \beta > 0$ and (a) for $\mathbf{A} = \mathrm{A} + D\mathrm{I}$. Typical problems result from standard discretizations of elliptic BVP with difference methods. As simple characteristical examples (see [9]), we mention discretizations of Dirichlet problems $Lu = f$ on a rectangle Ω, $u = g$ on $\partial\Omega$, with central difference quotients on a grid with constant mesh size:

(i) the nine point formula for $Lu = u_{xx} + u_{yy}$ leading to $\mathrm{A} = (-4, 20, -4)$, $\mathrm{S} = \mathrm{T} = (1, 4, 1)$,

(ii) a five point formula for $Lu = \big(a(x)u_x\big)_x + bu_{yy}$, $a, b > 0$, leading to $\mathrm{A} = (-\gamma_{i-1}, 2\beta + \gamma_{i-1} + \gamma_i, -\gamma_i)$, $\mathrm{S} = \mathrm{T} = \beta\mathrm{I}$,

(iii) a five point formula for $Lu = a(x)u_{xx} + bu_{yy} + c(x)u_x + du_y + e(x)u$, $a, b > 0$, $e \geq 0$, under appropriate conditions on the mesh size leading to $\mathrm{A} = (-\gamma_i, \varepsilon_i + \alpha + \beta, -\delta_i)$, $\mathrm{S} = \alpha\mathrm{I}$, $\mathrm{T} = \beta\mathrm{I}$.

Interval matrices $\mathbf{M}$, i.e. $\mathbf{A} = \mathrm{A} + D\mathrm{I}$ result, for example, from the application of $IBU(D)$ in Newton-like methods where D is due to the interval arithmetic treatment of the nonlinearity of problems like $\mathrm{f(x)} = \mathrm{Mx} + \varphi(\mathrm{x}) = \mathrm{o}$ with a diagonal function φ, where $\varphi' \geq 0$ on Ω (see [9]).

4 Suboptimal enclosures

Theorem 3.1 shows that *IBUD* can be carried out for arbitrary q under appropriate conditions while optimal enclosures can be expected only for particular q. The nonoptimality is reflected by an increased interval width, due to a change of sign in one or more of the terms $\lambda_i^{(r)} - \gamma_{j(i)}^{(r)}$ in (3.6c). Based on the ordering (3.11), we prove a qualitative result for the dependence of the additional width for nonoptimal q on the diagonal dominance of $\mathbf{A}$ and we develop an estimate of this width for a restricted set of values for q which are candidates for a moderate increase of the width. In a noninterval system $\mathrm{Mx} = \mathrm{y}$, the problem posed by (3.6c) is reduced to the solution of systems of the form

$$\prod_{i=1}^{k_r} (\mathrm{A} - \lambda_i^{(r)}\mathrm{U})\mathrm{x} = \prod_{i=1}^{l_r} (\mathrm{A} - \mu_i^{(r)}\mathrm{U}) \prod_{i=1}^{2^r} (\mathrm{A} - \alpha_i^{(r)}\mathrm{U})\mathrm{y}. \tag{4.1}$$

From the definition of the k_r and l_r, it follows that

$$(a)\, k_r = l_r + 2^r \quad (b)\, 0 \leq l_r \leq 2^r \quad (c)\, 2^r \leq k_r < 2^{r+1}. \tag{4.2}$$

By merging the $\alpha_i^{(r)}$ and $\mu_i^{(r)}$ appropriately $\big($compare (3.10)$\big)$, (4.1) can be simplified to

$$\prod_{i=1}^{k_r} (\mathrm{A} - \lambda_i^{(r)}\mathrm{U})\mathrm{x} = \prod_{i=1}^{k_r} (\mathrm{A} - \gamma_i^{(r)}\mathrm{U})\mathrm{y}. \tag{4.3}$$

In the interval case, we have to deal with matrices $\mathbf{A} - \alpha \mathrm{U}$. After some manipulations in (4.1) [13] and a subsequent interval extension of (4.3), we get (3.9). This formula ensures that multiplications by $\mathbf{A}$ on the right-hand sides of interval systems are avoided. In a first step, we justify the ordering (3.11). In the sequel, we assume r, $0 \leq r \leq r_q$, to be arbitrary, but fixed.

Lemma 4.1. *In the case of nontrivial use of (3.6c), i.e.* $q \neq 2^n(2^m+1)-1$, $n, m \in \mathbb{N}$, *the following relation holds:*

$$\forall r \in \{1, \ldots, r_q\}\, \forall i \in \{1, \ldots, k_r\} : \quad \lambda_i^{(r)} < \alpha_i^{(r)}.$$

Proof. According to (4.2) we note $k_r + 1 \leq 2^{r+1}$ where equality only holds if q_s is odd for all $s \leq r$, i.e. $k_r = 2^{r+1} - 1 = \sum_{i=0}^{r} 2^i$. Therefore, $k_r + 1 \leq 2^{r+1}$ which is equivalent to $\frac{i}{k_r+1} > \frac{1}{2^{r+1}}$, i.e. the above assertion holds. □

Corollary 4.2. *There does* **not** *exist an ordering such that*

$$\begin{aligned} &\forall i \in \{1, \ldots, k_r\} : \\ &\quad \exists j \in \{1, \ldots, 2^r\} : \lambda_i^{(r)} \geq \alpha_j^{(r)} \quad \textit{or} \quad \exists k \in \{1, \ldots, l_r\} : \lambda_i^{(r)} \geq \mu_k^{(r)}. \end{aligned}$$

Lemma 4.1 and Corollary 4.2 show that optimal enclosures cannot be expected by *IBUD* if (3.6c) and (3.9), resp., are involved. In the sequel, we use the abbreviations

$$l^{(i)} := \frac{i}{l_r + 1}, \quad a^{(i)} := \frac{2i-1}{2^{r+1}}, \quad k^{(i)} := \frac{i}{k_r + 1}.$$

Lemma 4.3. *Each of the intervals* $\big[-2, \lambda_{k_r}^{(r)}\big]$, $\big[\lambda_1^{(r)}, 2\big]$, $\big[\lambda_i^{(r)}, \lambda_{i+1}^{(r)}\big]$, $1 \leq i \leq k_r - 1$, *contains at most one* $\alpha_j^{(r)}$ *and one* $\mu_k^{(r)}$.

Proof. We first note $\frac{1}{k_r+1} \leq \frac{1}{2^r} \leq \frac{1}{l_r+1}$ which implies for all $i \in \{1, \ldots, k_r - 1\}$, $m \in \{1, \ldots, 2^r - 1\}$, $j \in \{1, \ldots, l_r - 1\}$ that $k^{(i+1)} - k^{(i)} \leq a^{(m+1)} - a^{(m)} \leq l^{(j+1)} - l^{(j)}$. On the other hand, (4.2) yields $\frac{1}{2^{r+1}} \leq \frac{1}{k_r+1} \leq \frac{1}{l_r+1}$, hence $\alpha_1^{(r)} \geq \lambda_1^{(r)} \geq \mu_1^{(r)}$. The corresponding relation for the other boundary interval follows from the symmetry of the cosine. □

Lemma 4.4.

$$x, y, z \in \mathbb{R}^+,\ x = y + z \Rightarrow [x] \geq [y] + [z].$$

Proof. Let $f(s)$ be the fractional part of an arbitrary real number, i.e. $s = f(s) + [s]$. The assertion follows from

$$\begin{array}{lll} [x] = [y] + [z], & f(x) = f(y) + f(z) & \text{if } f(y) + f(z) \leq 1 \\ [x] = [y] + [z] + 1, & f(x) = f(y) + f(z) - 1 & \text{if } f(y) + f(z) > 1. \end{array}$$

□

Lemma 4.5. *Define* $\operatorname{card}(k^{(i)})$ *to be the number of* $k^{(i)}$ *(analogously* $l^{(i)}$, $a^{(i)}$*) in a given interval* $[0, t]$, $0 \leq t \leq 1$. *Then*

$$-1 \leq \operatorname{card}(k^{(i)}) - \big(\operatorname{card}(l^{(i)}) + \operatorname{card}(a^{(i)})\big).$$

Proof. According to their definition, the distance $k^{(i+1)} - k^{(i)}$ between two consecutive numbers $k^{(i)}$, $k^{(i+1)}$ (similarly $a^{(i+1)} - a^{(i)}$, $l^{(i+1)} - l^{(i)}$) is always the same. This implies $\operatorname{card}(k^{(i)}) = [tk_r]$ and $\operatorname{card}(l^{(i)}) = [tl_r]$. In the case of the $a^{(i)}$, we have to take account of $a^{(1)} - 0 = 1/2^{r+1}$ instead of $1/2^r$ and we get $[t2^r] \leq \operatorname{card}(a^{(i)}) \leq [t2^r] + 1$. The assertion then follows from Lemma 4.4. □

Corollary 4.6. *Assume an arbitrary subinterval* $I \subseteq [-2, 2]$. *The number of* $\lambda_i^{(r)}$ *included in* I *differs by at most* 1 *from the number of* $\alpha_j^{(r)}$, $\mu_k^{(r)}$ *included in* I.

We summarize the result of the preceding lemmata in

Theorem 4.7. *In the case of nontrivial use of (3.6c), i.e.* $q \neq 2^n(2^m+1)-1$, $n, m \in \mathbb{N}$, *we get*

$$\lambda_{k_r}^{(r)} < \gamma_1^{(r)} \quad \textit{and} \quad \lambda_i^{(r)} \geq \gamma_{i+1}^{(r)}, \quad 1 \leq i \leq k_r - 1$$

for the ordering (3.11) where $\big(\mu_i^{(r)}\big)_{i=1}^{l_r}$, $\big(\alpha_i^{(r)}\big)_{i=1}^{2^r}$ *are merged in* $\big(\gamma_i^{(r)}\big)_{i=1}^{k_r}$ *according to (3.10).*

Proof. According to Lemma 4.1, we have $\lambda_{k_r}^{(r)} \leq \lambda_1^{(r)} < \gamma_1^{(r)} = \alpha_1^{(r)}$ and $\alpha_2^{(r)}, \mu_2^{(r)} \leq \lambda_1^{(r)}$. Therefore, the interval $\big[-2, \lambda_1^{(r)}\big]$ contains $k_r - 1 \quad \gamma_i^{(r)}$, in particular $\gamma_2^{(r)}$, hence $\lambda_1^{(r)} \geq \gamma_2^{(r)}$. Assume now, that $\lambda_j^{(r)} \geq \gamma_{j+1}^{(r)}$ for $1 \leq j \leq i$. According to Lemma 4.5, the interval $\big(\lambda_{i+1}^{(r)}, 2\big]$ contains at most $i+1 \; \gamma_j^{(r)}$, i.e. $\gamma_j^{(r)} \in \big[-2, \lambda_{i+1}^{(r)}\big]$ for $i+2 \leq j \leq k_r$ or $\lambda_{i+1}^{(r)} \geq \gamma_{i+2}^{(r)}$. Finally, there remain $\lambda_{k_r}^{(r)}$ and $\gamma_1^{(r)}$ with $\lambda_{k_r}^{(r)} < \gamma_1^{(r)}$. □

Theorem 4.7 yields a justification for the ordering (3.11) of the $\lambda_i^{(r)}$ and $\gamma_j^{(r)}$ which is only applied in (3.9). This will enable us to develop easily computable estimates for the width in (3.9) (see Lemma 4.9) which will be integrated into an estimate for the overall width of *IBUD* for nonoptimal q. In the noninterval case the terms $\big|\lambda_i^{(r)} - \gamma_{j(i)}^{(r)}\big|$ are minimized in order to minimize the roundoff error [13]. In the interval case, an optimal enclosure requires $\lambda_i^{(r)} \geq \gamma_{j(i)}^{(r)}$ which cannot hold for all i. (3.11) can at least ensure that only one change of the sign in the sequence of the $\lambda_i^{(r)} - \gamma_{j(i)}^{(r)}$, i.e. only one inequality $\lambda_i^{(r)} < \gamma_{j(i)}^{(r)}$ occurs. We do not claim that this strategy is the best one as principally any ordering is admissible. As an example we mention any minimization of the terms $\big|\lambda_i^{(r)} - \gamma_{j(i)}^{(r)}\big|$. But the proof of Lemma 4.9 will illustrate that the complexity of estimates for the width for orderings of this type seems not to be reasonable.

After having defined the final form of (3.9) by Theorem 4.7, we are able to estimate the additional width of the enclosure vector $\mathbf{x} = \mathit{LES}(\mathbf{M}, \mathbf{b})$ in the case of nonoptimal enclosures. In order to achieve this task, we determine an enclosure

$$\mathbf{x} \subseteq \tilde{\mathbf{x}} + \hat{\mathbf{x}} \tag{4.4}$$

where $\tilde{\mathbf{x}} \equiv \mathbf{x}_{opt}$ denotes the optimal enclosure according to Theorem 3.1 and where $\hat{\mathbf{x}}$ denotes an enclosure of the additional width occurring for nonoptimal q. Then we derive an estimate for $\mathrm{d}(\hat{\mathbf{x}})$ depending on $\mathrm{d}(\tilde{\mathbf{x}})$ in a suitable norm.

The larger enclosures are caused by the application of (3.9) in (3.6c). Therefore, we first derive an estimate of the form (4.4) for (3.9). In the sequel, we denote all subvectors leading to $\tilde{\mathbf{x}}$ accordingly by "$\tilde{\ }$" and the corresponding enclosures of the additional width by "$\hat{\ }$". We note some conditions which are needed in the subsequent Lemmata and Theorems:

$$\begin{aligned}
&\textit{(a)–(f) of Theorem 3.1}\\
&\textit{(g)}\quad \forall \mathrm{A} \in \mathbf{A} : \operatorname{spec}(\mathrm{A}) = (\lambda_i)_{i=1}^p,\ \operatorname{spec}(\mathrm{U}) = (\omega_i)_{i=1}^p \subset \mathbb{R}\\
&\qquad \textit{and } \forall i \in \{1, \dots, p\} : \lambda_i \ge 2\omega_i > 0 \qquad\qquad (4.5)\\
&\textit{(h)}\quad \forall \mathrm{A} \in \mathbf{A} : \textit{ all eigenvectors of } \mathrm{A} \textit{ are eigenvectors of } \mathrm{S}, \mathrm{T}, \textit{ and } \mathrm{U}\\
&\textit{(i)}\quad \forall \mathrm{A} \in \mathbf{A} : \mathrm{A} \textit{ is irreducible or } \mathrm{A} \textit{ has an eigenvector } \mathrm{c} > 0.
\end{aligned}$$

The conditions (g) and (h) enable us to treat nonsymmetric matrices like in example (iii) after Theorem 3.1. Obviously, (g) and (h) are satisfied for symmetric positive definite and commuting matrices A, S, T where S = T.

Before we start the discussion of the behaviour of *IBUD* with respect to the occurrence of (3.9), we collect in the following lemma some auxiliary results which are needed in later estimates.

Lemma 4.8. *Assume (3.1)–(3.11), (4.5),* $\mathrm{A} \in \mathbf{A}$*. Define* $\theta \equiv \theta(\lambda, \omega) := \operatorname{arcosh}\big(\lambda/(2\omega)\big)$ *for* $\lambda \ge 2\omega$*,*

$\operatorname{sh}(m, n, \theta) := \frac{\sinh\{m\theta\}}{\sinh\{n\theta\}}$ *for* $\theta > 0$, $\operatorname{sh}(m, n, 0) := 1$ *and* $\lambda_{min} := 1/\rho(\mathrm{A}^{-1})$, $\omega_{max} := \rho(\mathrm{U})$, $s := \rho(\mathrm{S})$, $t := \rho(\mathrm{T}), \theta_i := \theta(\lambda_i, \omega_i)$, $\theta_{min} := \theta(\lambda_{min}, \omega_{max})$.

Then

(a) $(\mathrm{A} - \alpha\mathrm{U})^{-1} \ge \mathrm{O}$ *for all roots* α *in (3.3)*

$$\mathrm{S}^{(r)}, \mathrm{T}^{(r)}, \mathrm{U}^{(r)}, \big(\mathrm{A}^{(r)}\big)^{-1}, \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}, \big(\mathrm{A}-\gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)} \ge \mathrm{O}$$

(b) All eigenvectors of A *are eigenvectors of* $\mathrm{A}^{(r)}$, $\mathrm{B}^{(r)}$, $\mathrm{C}^{(r)}$, $\mathrm{S}^{(r)}$, $\mathrm{T}^{(r)}$, $\mathrm{U}^{(r)}$

(c) $\mathrm{A}^{(r)}$, $\mathrm{B}^{(r)}$, $\mathrm{C}^{(r)}$, $\mathrm{S}^{(r)}$, $\mathrm{T}^{(r)}$, $\mathrm{U}^{(r)}$ *commute*

(d) For $0 \le m < n : \operatorname{sh}(m, n, \theta) \downarrow 0\, (\theta \uparrow \infty)$ *and* $\operatorname{sh}(m, n, \theta) \le 1$

(e) $\big|\lambda_{k_r}^{(r)} - \gamma_1^{(r)}\big| \le 4$

(f) $a_\gamma^{(r)} := \rho\Big(\big(\mathrm{A} - \gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\Big) = \dfrac{1}{(\lambda_{min}/\omega_{max}) - \gamma_1^{(r)}}$

(g)
$$\begin{aligned}
b_{\mathrm{S}}^{(r)} &:= \rho\Big(\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{S}^{(r)}\Big)\\
&= \left(\frac{s}{\omega_{max}}\right)^{2^r}\mathrm{sh}\{l_r+1,k_r+1,\theta_{min}\} \le \left(\frac{s}{\omega_{max}}\right)^{2^r}\\
b_{\mathrm{T}}^{(r)} &:= \rho\Big(\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{T}^{(r)}\Big)\\
&= \left(\frac{t}{\omega_{max}}\right)^{2^r}\mathrm{sh}\{l_r+1,k_r+1,\theta_{min}\} \le \left(\frac{t}{\omega_{max}}\right)^{2^r}\\
b_{\mathrm{TU}}^{(r)} &:= \rho\Big(\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{T}^{(r-1)}\mathrm{U}^{(r-1)}\Big)\\
&= \left(\frac{t}{\omega_{max}}\right)^{2^{r-1}}\mathrm{sh}\{l_r+1,k_r+1,\theta_{min}\} \le \left(\frac{t}{\omega_{max}}\right)^{2^{r-1}}\\
b_{\mathrm{U}}^{(r)} &:= \rho\Big(\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{U}^{(r)}\Big) = \mathrm{sh}\{l_r+1,k_r+1,\theta_{min}\} \le 1
\end{aligned}$$

(h)
$$\begin{aligned}
a_{\mathrm{S}}^{(r)} &:= \rho\Big(\big(\mathrm{A}^{(r)}\big)^{-1}\mathrm{S}^{(r)}\Big) = \left(\frac{s}{\omega_{max}}\right)^{2^r}\frac{1}{2\cosh\{2^r\theta_{min}\}} \le \frac{1}{2}\left(\frac{s}{\omega_{max}}\right)^{2^r}\\
a_{\mathrm{T}}^{(r)} &:= \rho\Big(\big(\mathrm{A}^{(r)}\big)^{-1}\mathrm{T}^{(r)}\Big) = \left(\frac{t}{\omega_{max}}\right)^{2^r}\frac{1}{2\cosh\{2^r\theta_{min}\}} \le \frac{1}{2}\left(\frac{t}{\omega_{max}}\right)^{2^r}\\
a_{\mathrm{U}}^{(r)} &:= \rho\Big(\big(\mathrm{A}^{(r)}\big)^{-1}\mathrm{U}^{(r)}\Big) = \frac{1}{2\cosh\{2^r\theta_{min}\}} \le \frac{1}{2}\\
x^{(r)} &:= \rho\big(\mathrm{X}^{(r)}\big) = \frac{1}{\displaystyle\prod_{i=0}^{r}\big(2\omega^{2^r}\cosh\{2^r\theta_{min}\}\big)} \le \frac{1}{2^{r+1}}
\end{aligned}$$

(i)
$$\begin{aligned}
b_{\gamma\mathrm{A}}^{(r)} &:= \rho\Big(\big(\mathrm{A}-\gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)}\Big)\\
&= b_{\mathrm{A}}^{(r)}\frac{\omega_{max}}{\lambda_{min}-\gamma_1^{(r)}\omega_{max}} \le \frac{1}{\lambda_{min}/\omega_{max}-\gamma_1^{(r)}}
\end{aligned}$$

$$\textit{with } b_{\mathrm{A}}^{(r)} := 1-\mathrm{sh}\{2^r-(l^r+1),2^r+(l^r+1),\theta_{min}\}$$

(j) $$w^{(r)} := \sum_{k=1}^{2^r}\{s^{2^r-1+k}t^{2^r-k}+s^{2^r-k}t^{2^r-1+k}\} \le 2^{r+1}\max\{s,t\}^{2^{r+1}-1}$$

(k) $a_{\gamma}^{(r)}, a_{\mathrm{S}}^{(r)}, a_{\mathrm{T}}^{(r)}, a_{\mathrm{U}}^{(r)}, b_{\gamma\mathrm{A}}^{(r)}, b_{\mathrm{S}}^{(r)}, b_{\mathrm{T}}^{(r)}, b_{\mathrm{U}}^{(r)}, b_{\mathrm{TU}}^{(r)}, x^{(r)} \downarrow 0\ (\lambda\uparrow\infty).$

Proof. We first note that each $\mathrm{A} \in \mathbf{A}$ is an M matrix as $\mathbf{M}$ is an interval M matrix. Therefore, $\mathrm{A}^{-1} \geq \mathrm{O}$ and with (4.5, g) $\lambda_{min} = \min\{\lambda | \lambda \in \operatorname{spec}(\mathrm{A})\} = 1/\rho(\mathrm{A}^{-1})$. In (a) $\mathrm{S}^{(r)}, \mathrm{T}^{(r)}, \mathrm{U}^{(r)} \geq \mathrm{O}$ follows from the definition of the corresponding matrices. (4.5, f) and the definition of the roots in (3.3) imply $\mathrm{A} - \alpha_1^{(r)}\mathrm{U} \leq \mathrm{A} - \alpha\mathrm{U} \leq \mathrm{A} + \alpha_1^{(r)}\mathrm{U}$. Then $\mathrm{A} - \alpha\mathrm{U}$ is also an M matrix and $\mathrm{O} \leq \big(\mathrm{A} + \alpha_1^{(r)}\mathrm{U}\big)^{-1} \leq \big(\mathrm{A} - \alpha\mathrm{U}\big)^{-1} \leq \big(\mathrm{A} - \alpha_1^{(r)}\mathrm{U}\big)^{-1}$. This also implies $\big(\mathrm{A}^{(r)}\big)^{-1} \geq \mathrm{O}$. We further get

$$\begin{aligned}
&\big(\mathrm{A} - \gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)} = \prod_{i=1}^{k_r-1}\big(\mathrm{A} - \gamma_{i+1}^{(r)}\mathrm{U}\big)^{-1}\prod_{i=1}^{k_r}\big(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\big)^{-1} \\
&= \prod_{i=1}^{k_r-1}\big\{\mathrm{I} + \big(\lambda_i^{(r)} - \gamma_{i+1}^{(r)}\big)\mathrm{U}\big(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\big)^{-1}\big\}\big(\mathrm{A} - \lambda_{k_r}^{(r)}\mathrm{U}\big)^{-1} \geq \mathrm{O}
\end{aligned} \tag{4.6}$$

and, with (3.12),

$$\begin{aligned}
\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)} &= \prod_{i=1}^{l_r}\big(\mathrm{A} - \mu_i^{(r)}\mathrm{U}\big)\prod_{i=1}^{k_r}\big(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\big)^{-1} \\
&= \prod_{i=1}^{l_r}\big\{\mathrm{I} + \big(\lambda_i^{(r)} - \mu_i^{(r)}\big)\mathrm{U}\big(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\big)^{-1}\big\}\prod_{i=l_r+1}^{k_r}\big(\mathrm{A} - \lambda_i^{(r)}\mathrm{U}\big)^{-1} \geq \mathrm{O}.
\end{aligned} \tag{4.7}$$

Parts (b), (c) follow trivially from the definition of the above matrices. For the proof of (d), we recall that with the substitution $\theta := \operatorname{arcosh}(x)$, $x := \lambda/(2\omega)$, $x \geq 1$, we can rewrite sh with Chebyshev polynomials of the second kind

$$\sinh(m\theta) = \frac{\sinh\{m\theta\}}{\sinh\{\theta\}} = \prod_{i=1}^{m-1}\big(2x - \lambda_i^{(m)}\big) = \mathrm{U}_{m-1}(x), \quad \lambda_i^{(m)} = 2\cos\frac{i\pi}{m}$$

as

$$\operatorname{sh}(m,n,\theta) = \frac{\sinh\{m\theta\}}{\sinh\{n\theta\}} = \frac{\prod_{i=1}^{m-1}\big(2x - \lambda_i^{(m)}\big)}{\prod_{i=1}^{n-1}\big(2x - \lambda_i^{(n)}\big)} = \frac{\mathrm{U}_{m-1}(x)}{\mathrm{U}_{n-1}(x)}.$$

Note that $\lambda_i^{(m)} < 2$ and $\lambda \geq 2\omega$ imply $2x > \lambda_i^{(m)}$. Differentiation with respect to x yields

$$\frac{d}{dx}\left(\frac{\mathrm{U}_{m-1}(x)}{\mathrm{U}_{n-1}(x)}\right) = 2\frac{\prod\limits_{i=1}^{m-1}\left(2x-\lambda_i^{(m)}\right)}{\prod\limits_{i=1}^{n-1}\left(2x-\lambda_i^{(n)}\right)}\left(\sum_{i=1}^{m-1}\frac{1}{\left(2x-\lambda_i^{(m)}\right)}-\sum_{i=1}^{n-1}\frac{1}{\left(2x-\lambda_i^{(n)}\right)}\right) < 0.$$

Because of $\theta \uparrow \infty \Leftrightarrow x \uparrow \infty$, the monotone convergence to 0 follows from $m < n$, in particular from $\lambda_i^{(m)} < \lambda_i^{(n)}$. (e) follows from the definition (3.3) of $\gamma_1^{(r)}$, $\lambda_{k_r}^{(r)}$. As A and U have the same eigenvectors, $\left(\mathrm{A}-\gamma_1^{(r)}\mathrm{U}\right)^{-1}\mathrm{U}$ has the eigenvalues $\omega_i/\left(\lambda_i-\gamma_1^{(r)}\omega_i\right)$, where $\lambda_i-\gamma_1^{(r)}\omega_i > \lambda_i - 2\omega_i \geq 0$. Because of $\left(\mathrm{A}-\gamma_1^{(r)}\mathrm{U}\right)^{-1}\mathrm{U} \geq \mathrm{O}$, the theorem of Perron-Frobenius ([14], Th. 2.1) shows that $a_\gamma^{(r)}$ is ab eigenvalue, where $\lambda_i = \lambda_{min} = 1/\rho(\mathrm{A}^{-1})$, $\omega_i = \omega_{max} = \rho(\mathrm{U})$, hence (f) holds.

According to the factorization (3.3), $\mathrm{A}^{(r)}$, $\mathrm{B}^{(r)}$, $\mathrm{C}^{(r)}$ have the nonnegative eigenvalues $\omega_i^{2^r}\cosh\{2^r\theta_i\}$, $\omega_i^{k_r}\sinh\{(k_r+1)\theta_i\}/\sinh\{\theta_i\}$, $\omega_i^{l_r}\sinh\{(l_r+1)\theta_i\}/\sinh\{\theta_i\}$ $(1 \leq i \leq p)$, resp., ([5, 13]), whose maximum is reached with (d) for $\theta_{min} = \theta(\lambda_{min}, \omega_{max})$. With part (b), we get

$$\rho\Big(\left(\mathrm{B}^{(r)}\right)^{-1}\mathrm{C}^{(r)}\mathrm{S}^{(r)}\Big) = \rho\Big(\left(\mathrm{B}^{(r)}\right)^{-1}\mathrm{C}^{(r)}\Big)\rho\left(\mathrm{S}^{(r)}\right) = s^{2^r}\omega_{max}^{l_r-k_r}\mathrm{sh}\{l_r+1, k_r+1, \theta_{min}\}.$$

The other relations in (g) can be shown similarly. (h) follows from $\cosh\{2^r\theta_i\} \geq 1$. According to (a) and to the ([14], 2.1), $\rho\Big(\big(\mathrm{A} -$

$\gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)}\Big)$ is an eigenvalue, hence

$$\begin{aligned}
&\rho\Big(\big(\mathrm{A}-\gamma_1^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)}\Big)\\
&=\prod_{i=1}^{k_r-1}\left\{1+\big(\lambda_i^{(r)}-\gamma_{i+1}^{(r)}\big)\frac{\omega_{max}}{\lambda_{min}-\lambda_i^{(r)}\omega_{max}}\right\}\frac{\omega_{max}}{\lambda_{min}-\lambda_{k_r}^{(r)}\omega_{max}}\\
&=\frac{\omega_{max}}{\lambda_{min}-\gamma_1^{(r)}\omega_{max}}\prod_{i=1}^{k_r}\frac{\lambda_{min}-\gamma_i^{(r)}\omega_{max}}{\lambda_{min}-\lambda_i^{(r)}\omega_{max}} \qquad (4.8)\\
&=\big(1-\mathrm{sh}\big\{2^r-(l^r+1),2^r+(l^r+1),\theta_{min}\big\}\big)\frac{\omega_{max}}{\lambda_{min}-\gamma_1^{(r)}\omega_{max}}\\
&\le\frac{1}{(\lambda_{min}/\omega_{max})-\gamma_1^{(r)}}
\end{aligned}$$

(compare (4.6)). Assertion (j) is trivial. The monotonicity in $\lambda\equiv\lambda_{min}$ follows from (d). For the matrix in (4.8), it can be directly deduced from the expression after the first equality. □

After these preparations, we start with the estimation of the additional width caused by one occurrence of (3.9).

Lemma 4.9. *Assume (3.1)–(3.11) and (4.5, a–f). Let* $\tilde{\mathbf{z}}$ *be the vector in (3.9) corresponding to an optimal enclosure for the global system* $(\mathbf{M},\mathbf{b})$ *in one of the cases (a)* $\mathbf{M}\equiv\mathrm{M}$ *and arbitrary* $\mathbf{b}$*, (b)* $0\in\mathbf{b}$*, (c)* $\mathrm{i}(\mathbf{b})\ge\mathrm{o}$*, or (d)* $\mathrm{s}(\mathbf{b})\le\mathrm{o}$*. Then the vector* $\mathbf{z}$ *computed by (3.9) in IBUD satisfies*

(a) $$\mathbf{z}=\tilde{\mathbf{z}}+\big|\lambda_{k_r}^{(r)}-\gamma_1^{(r)}\big|\big(\mathrm{A}-\gamma_i^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}_d^{(r)}\big)^{-1}\mathrm{C}_d^{(r)}\mathrm{A}_d^{(r)}\mathrm{d}(\mathbf{y})[-1,1]$$

(b) $$\mathbf{z}=\tilde{\mathbf{z}}+\big|\lambda_{k_r}^{(r)}-\gamma_1^{(r)}\big|\big(\mathrm{i}(\mathbf{A})-\gamma_i^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}_{inf}^{(r)}\big)^{-1}\mathrm{C}_{inf}^{(r)}\mathrm{A}_{inf}^{(r)}\mathrm{d}(\mathbf{y})[-1,1]$$

(c) $$\begin{aligned}\mathbf{z}=\tilde{\mathbf{z}}+\big|\lambda_{k_r}^{(r)}-\gamma_1^{(r)}\big|\Big(&\big(\mathrm{i}(\mathbf{A})-\gamma_i^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}_{inf}^{(r)}\big)^{-1}\mathrm{C}_{inf}^{(r)}\mathrm{A}_{inf}^{(r)}\mathrm{s}(\mathbf{y})\\&-\big(\mathrm{s}(\mathbf{A})-\gamma_i^{(r)}\mathrm{U}\big)^{-1}\mathrm{U}\big(\mathrm{B}_{sup}^{(r)}\big)^{-1}\mathrm{C}_{sup}^{(r)}\mathrm{A}_{sup}^{(r)}\mathrm{i}(\mathbf{y})\Big)[-1,1]\end{aligned}$$

(d) *like (c) with* $\mathrm{i}(\mathbf{A})$ *and* $\mathrm{s}(\mathbf{A})$ *exchanged*

where $\mathrm{A}_{inf}^{(r)}:=\prod_{i=1}^{2^r}\big(\mathrm{i}(\mathbf{A})-\alpha_i^{(r)}\mathrm{U}\big)$, $\mathrm{A}_{sup}^{(r)}:=\prod_{i=1}^{2^r}\big(\mathrm{s}(\mathbf{A})-\alpha_i^{(r)}\mathrm{U}\big)$, $\mathrm{A}_d^{(r)}:=\prod_{i=1}^{2^r}\big(\mathrm{A}-\alpha_i^{(r)}\mathrm{U}\big)$*, and similarly defined* $\mathrm{B}_d^{(r)}$, $\mathrm{C}_d^{(r)}$, $\mathrm{B}_{inf}^{(r)}$, $\mathrm{C}_{inf}^{(r)}$, $\mathrm{B}_{sup}^{(r)}$, $\mathrm{C}_{sup}^{(r)}$.

Proof. (3.9) is characterized by the treatment of a series of systems of the form

$$\mathbf{z} = \mathbf{y} + \alpha \mathrm{U} \mathit{LES}(\mathbf{B}, \mathbf{y}) \tag{4.9}$$

with, according to (3.10) and (4.5, f), interval M matrices $\mathbf{B}$ and $\mathrm{U} \geq \mathrm{O}$. The applicability of *LES* follows from (4.5, c, d). *LES* yields optimal enclosures according to (4.5, e), i.e. (2.14).

Assume one of cases (a)–(d) for the main system $(\mathbf{M}, \mathbf{b})$. With $\mathrm{S}^{(r)}, \mathrm{T}^{(r)}, \mathrm{U}^{(r)} \geq \mathrm{O}$ and (4.5, e), it can be shown by induction that in any partial system $\mathit{LES}(\mathbf{A} - \alpha\mathrm{U}, \mathbf{w})$ in *IBUD* before the first occurrence of (3.6c)/(3.9) (including the latter) the same case applies. For (a) this is obvious. For (b) this follows with (2.7) from the relations $\mathrm{o} \in \mathbf{b} \Rightarrow \mathrm{o} \in \mathbf{w} \Rightarrow \mathrm{o} \in \mathit{LES}(\mathbf{A} - \alpha\mathrm{U}, \mathbf{w}) \Rightarrow \mathrm{o} \in \mathbf{y}$ in (3.9). (c), (d) are treated similarly with (2.8).

If (a) holds, we obtain for $\alpha \geq 0$ an optimal enclosure $\mathbf{z} = \tilde{\mathbf{z}}$ and with the abbreviation $\mathbf{B} = \mathbf{A} - \alpha\mathrm{U}$ and $\mathrm{B} = \mathrm{A} - \alpha\mathrm{U}$, resp.,

$$\begin{array}{rcl} \mathrm{i}(\tilde{\mathbf{z}}) & = & \mathrm{i}(\mathbf{y}) + \alpha\mathrm{U}\mathrm{B}^{-1}\mathrm{i}(\mathbf{y}) = (\mathrm{I} + \alpha\mathrm{U}\mathrm{B}^{-1})\mathrm{i}(\mathbf{y}) \\ \mathrm{s}(\tilde{\mathbf{z}}) & = & \mathrm{s}(\mathbf{y}) + \alpha\mathrm{U}\mathrm{B}^{-1}\mathrm{s}(\mathbf{y}) = (\mathrm{I} + \alpha\mathrm{U}\mathrm{B}^{-1})\mathrm{s}(\mathbf{y}) \\ \tilde{\mathbf{z}} & = & (\mathrm{I} + \alpha\mathrm{U}\mathrm{B}^{-1})\mathbf{y} \\ \mathrm{d}(\tilde{\mathbf{z}}) & = & (\mathrm{I} + \alpha\mathrm{U}\mathrm{B}^{-1})\mathrm{d}(\mathbf{y}). \end{array} \tag{4.10}$$

For $\alpha < 0$ an optimal enclosure still has the form (4.10), but *IBUD* yields

$$\begin{array}{rcl} \mathrm{i}(\mathbf{z}) & = & \mathrm{i}(\mathbf{y}) - |\alpha|\mathrm{U}\mathrm{B}^{-1}\mathrm{s}(\mathbf{y}) \\ \mathrm{s}(\mathbf{z}) & = & \mathrm{s}(\mathbf{y}) - |\alpha|\mathrm{U}\mathrm{B}^{-1}\mathrm{i}(\mathbf{y}) \\ \mathbf{z} & = & \mathbf{y} - |\alpha|\mathrm{U}\mathrm{B}^{-1}\mathbf{y} \\ \mathrm{d}(\mathbf{z}) & = & (\mathrm{I} + |\alpha|\mathrm{U}\mathrm{B}^{-1})\mathrm{d}(\mathbf{y}). \end{array} \tag{4.11}$$

The comparison of (4.10) and (4.11) yields

$$\begin{array}{rcl} \mathrm{i}(\mathbf{z}) & = & \mathrm{i}(\tilde{\mathbf{z}}) - |\alpha|\mathrm{U}\mathrm{B}^{-1}\mathrm{d}(\mathbf{y}) \\ \mathrm{s}(\mathbf{z}) & = & \mathrm{s}(\tilde{\mathbf{z}}) + |\alpha|\mathrm{U}\mathrm{B}^{-1}\mathrm{d}(\mathbf{y}) \\ \mathrm{d}(\mathbf{z}) & = & \mathrm{d}(\tilde{\mathbf{z}}) + 2|\alpha|\mathrm{U}\mathrm{B}^{-1}\mathrm{d}(\mathbf{y}). \end{array} \tag{4.12}$$

For (b), we replace B by $\mathrm{i}(\mathbf{B})$ in (4.10)–(4.12), for (c) (similarly for (d)), we get

$$\begin{array}{rcl} \mathrm{i}(\tilde{\mathbf{z}}) & = & \mathrm{i}(\mathbf{y}) - |\alpha|\mathrm{U}\,\mathrm{i}(\mathbf{B})^{-1}\mathrm{s}(\mathbf{y}) \\ \mathrm{s}(\tilde{\mathbf{z}}) & = & \mathrm{s}(\mathbf{y}) - |\alpha|\mathrm{U}\,\mathrm{s}(\mathbf{B})^{-1}\mathrm{i}(\mathbf{y}) \\ \mathrm{d}(\tilde{\mathbf{z}}) & = & \mathrm{d}(\mathbf{y}) + |\alpha|\mathrm{U}\,\mathrm{d}\big(\mathit{LES}(\mathbf{B}, \mathbf{y})\big), \end{array} \tag{4.13}$$

$$\begin{aligned}
\mathrm{i}(\mathbf{z}) &= \mathrm{i}(\tilde{\mathbf{z}}) - |\alpha| \mathrm{U}\, \mathrm{d}\big(\mathit{LES}(\mathbf{B}, \mathbf{y})\big) \\
\mathrm{s}(\mathbf{z}) &= \mathrm{s}(\tilde{\mathbf{z}}) + |\alpha| \mathrm{U}\, \mathrm{d}\big(\mathit{LES}(\mathbf{B}, \mathbf{y})\big) \\
\mathrm{d}(\mathbf{z}) &= \mathrm{d}(\tilde{\mathbf{z}}) + |\alpha| \mathrm{U}\, \mathrm{d}\big(\mathit{LES}(\mathbf{B}, \mathbf{y})\big).
\end{aligned} \tag{4.14}$$

The above relations can be summarized in

$$\begin{aligned}
&\mathbf{z} = \tilde{\mathbf{z}} + |\alpha| \mathrm{U}\, \mathrm{d}\big(\mathit{LES}(\mathbf{B}, \mathbf{y})\big)[-1, 1] && \text{for (c), (d)} \\
&\mathbf{z} = \tilde{\mathbf{z}} + |\alpha| \mathrm{U}\, \mathrm{i}(\mathbf{B})^{-1} \mathrm{d}(\mathbf{y})[-1, 1] && \text{for (b)} \\
&\mathbf{z} = \tilde{\mathbf{z}} + |\alpha| \mathrm{U} \mathrm{B}^{-1} \mathrm{d}(\mathbf{y})[-1, 1] && \text{for (a).}
\end{aligned} \tag{4.15}$$

We first consider (a). For $\alpha = \lambda_i^{(r)} - \gamma_{i+1}^{(r)} \geq 0$, $\mathbf{z} = \mathbf{z}_i$, $\mathbf{y} = \mathbf{z}_{i-1}$, and the interval M matrix $\mathbf{B} = \mathbf{A} - \lambda_i^{(r)} \mathrm{U} \equiv \mathrm{A} - \lambda_i^{(r)} \mathrm{U} \equiv \mathrm{B}$ (4.10) yields

$$\begin{aligned}
\tilde{\mathbf{z}}_i &= \tilde{\mathbf{z}}_{i-1} + \big(\lambda_i^{(r)} - \gamma_{i+1}^{(r)}\big) \mathrm{U} \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} \tilde{\mathbf{z}}_{i-1} \\
&= \Big\{ \mathrm{I} + \big(\lambda_i^{(r)} - \gamma_{i+1}^{(r)}\big) \mathrm{U} \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} \Big\} \tilde{\mathbf{z}}_{i-1} \\
&= \Big(\big(\mathrm{A} - \gamma_{i+1}^{(r)} \mathrm{U}\big) \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} \Big) \tilde{\mathbf{z}}_{i-1} \\
&= \prod_{j=1}^{i} \Big(\big(\mathrm{A} - \gamma_{i+1}^{(r)} \mathrm{U}\big) \big(\mathrm{A} - \lambda_j^{(r)} \mathrm{U}\big)^{-1} \Big) \mathbf{y}
\end{aligned} \tag{4.16}$$

$$\mathrm{d}(\tilde{\mathbf{z}}_i) = \prod_{j=1}^{i} \Big(\big(\mathrm{A} - \gamma_{j+1}^{(r)} \mathrm{U}\big) \big(\mathrm{A} - \lambda_j^{(r)} \mathrm{U}\big)^{-1} \Big) \mathrm{d}(\mathbf{y})$$

for $1 \leq i \leq k_r$ because of (2.4), (2.6) and

$$\mathrm{I} + \big(\lambda_i^{(r)} - \gamma_{i+1}^{(r)}\big) \mathrm{U} \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} = \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} \big(\mathrm{A} - \gamma_{i+1}^{(r)} \mathrm{U}\big) \geq \mathrm{O} \tag{4.17}$$

with optimal enclosures $\mathbf{z}_i = \tilde{\mathbf{z}}_i$ for $1 \leq i \leq k_r - 1$. With $\alpha = \lambda_{k_r}^{(r)} - \gamma_1^{(r)} < 0$, $\mathbf{z} = \mathbf{z}_{k_r}$, $\mathbf{y} = \mathbf{z}_{k_r - 1}$, we further get

$$\mathbf{z}_{k_r} = \tilde{\mathbf{z}}_{k_r} + \big|\lambda_{k_r}^{(r)} - \gamma_1^{(r)}\big| \mathrm{U} \big(\mathrm{A} - \lambda_{k_r}^{(r)} \mathrm{U}\big)^{-1} \mathrm{d}(\mathbf{z}_{k_r - 1})[-1, 1]. \tag{4.18}$$

With (4.16), (4.18), (2.2), (2.3), and $\mathbf{z}_{k_r - 1} = \tilde{\mathbf{z}}_{k_r - 1}$, the computed enclosure $\mathbf{z} = \mathbf{z}_{k_r}$ satisfies

$$\mathbf{z} = \tilde{\mathbf{z}} + \Big(\big|\lambda_{k_r}^{(r)} - \gamma_1^{(r)}\big| \mathrm{U} \prod_{i=2}^{k_r} \big(\mathrm{A} - \gamma_i^{(r)} \mathrm{U}\big) \prod_{i=1}^{k_r} \big(\mathrm{A} - \lambda_i^{(r)} \mathrm{U}\big)^{-1} \Big) \mathrm{d}(\mathbf{y})[-1, 1] \tag{4.19}$$

with $\gamma^{(r)}_{k_r+1} := \gamma^{(r)}_1$. With (3.3), this equivalent to the assertion for case (a). The other two cases are treated similarly using (4.15). For case (b), we note with (2.7) $\mathrm{o} \in \tilde{\mathbf{z}}_{i-1} \Rightarrow \mathrm{o} \in \tilde{\mathbf{z}}_i$ in (4.16). For (c), the repeated application of (2.14, c) yields $\mathrm{o} \leq \mathrm{i}(\tilde{\mathbf{z}}_{i-1}) \Rightarrow \mathrm{o} \leq \mathrm{i}(\tilde{\mathbf{z}}_i)$ with (2.4), (2.8) and

$$\begin{aligned}
\tilde{\mathbf{z}}_i &= \tilde{\mathbf{z}}_{i-1} + \big(\lambda^{(r)}_i - \gamma^{(r)}_{i+1}\big)\mathrm{U}\big[\big(\mathrm{s}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\mathrm{i}(\tilde{\mathbf{z}}_{i-1}),\\
&\qquad\qquad \big(\mathrm{i}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\mathrm{s}(\tilde{\mathbf{z}}_{i-1})\big]\\
&= \Big[\Big\{\mathrm{I} + \big(\lambda^{(r)}_i - \gamma^{(r)}_{i+1}\big)\mathrm{U}\big(\mathrm{s}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\Big\}\mathrm{i}(\tilde{\mathbf{z}}_{i-1}),\\
&\qquad \Big\{\mathrm{I} + \big(\lambda^{(r)}_i - \gamma^{(r)}_{i+1}\big)\mathrm{U}\big(\mathrm{i}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\Big\}\mathrm{s}(\tilde{\mathbf{z}}_{i-1})\Big]\\
&= \Big[\big(\mathrm{s}(\mathbf{A}) - \gamma^{(r)}_{i+1}\mathrm{U}\big)\big(\mathrm{s}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\mathrm{i}(\tilde{\mathbf{z}}_{i-1}),\\
&\qquad \big(\mathrm{i}(\mathbf{A}) - \gamma^{(r)}_{i+1}\mathrm{U}\big)\big(\mathrm{i}(\mathbf{A}) - \lambda^{(r)}_i\mathrm{U}\big)^{-1}\mathrm{s}(\tilde{\mathbf{z}}_{i-1})\Big]\\
&= \Big[\prod_{j=1}^{i}\Big(\big(\mathrm{s}(\mathbf{A}) - \gamma^{(r)}_{j+1}\mathrm{U}\big)\big(\mathrm{s}(\mathbf{A}) - \lambda^{(r)}_j\mathrm{U}\big)^{-1}\Big)\mathrm{i}(\mathbf{y})\\
&\qquad \prod_{j=1}^{i}\Big(\big(\mathrm{i}(\mathbf{A}) - \gamma^{(r)}_{j+1}\mathrm{U}\big)\big(\mathrm{i}(\mathbf{A}) - \lambda^{(r)}_j\mathrm{U}\big)^{-1}\Big)\mathrm{s}(\mathbf{y})\Big].
\end{aligned} \tag{4.20}$$

For (d), we apply (2.14, d) exchanging $\mathrm{i}(\mathbf{A})$ and $\mathrm{s}(\mathbf{A})$ and using $\mathrm{s}(\tilde{\mathbf{z}}_{i-1}) \leq \mathrm{o} \Rightarrow \mathrm{s}(\tilde{\mathbf{z}}_i) \leq \mathrm{o}$ and (2.8). □

Orderings with more than one inequality $\lambda^{(r)}_i < \gamma^{(r)_{j(i)}}$ are possible, but they require a nested application of (4.18) also for $i \neq k_r$. In this case, the assertions of Lemma 4.9 become extremely complex. Lemma 4.9 indicates the additional width in one application of (3.9) with the ordering (3.11). The following theorem describes the propagation of this width across *IBUD*. We define

$$r_t := \min\{r | r_s + 1 \leq r \leq r_q, r \text{ odd}\}. \tag{4.21}$$

Theorem 4.10. *Assume (4.5), (3.1)–(3.11) and (a)* $\mathbf{M} \equiv \mathrm{M}$, *i.e.* $\mathbf{A} \equiv \mathrm{A}$, $\mathbf{b}$ *arbitrary or (b)* $\mathrm{o} \in \mathbf{b}$ *or (c)* $\mathrm{o} \leq \mathrm{i}(\mathbf{b})$ *or (d)* $\mathrm{s}(\mathbf{b}) \leq \mathrm{o}$. *Then*

1) $\exists c(q) \geq 0$*:*

$$\max_{1\leq j\leq q}\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_{\mathrm{c}}\} \leq c(q)\begin{cases}\max\limits_{1\leq j\leq q}\{\|\mathrm{d}(\tilde{\mathbf{x}}_j)\|_{\mathrm{c}}\} & a),\ b)\\ \max\limits_{1\leq j\leq q}\{\|\,|\mathrm{s}(\tilde{\mathbf{x}}_j)|\,\|_{\mathrm{c}}\} & c),\ d)\end{cases}$$

where $\|\mathrm{x}\|_{\mathrm{c}} := \max\limits_{1\le i\le p}\left\{\dfrac{|\mathrm{x}_i|}{c_i}\right\}$, $\mathrm{Ac} = \rho(\mathrm{A})\mathrm{c}$ *for (a) and* $\mathrm{i}(\mathbf{A})\mathrm{c} = \rho\big(\mathrm{i}(\mathbf{A})\big)\mathrm{c}$ *otherwise*

2) $c(q) \downarrow 0\,(d \uparrow \infty)$ *for* $\mathbf{A}$ *replaced by* $\mathbf{A} + d\mathrm{I}$.

If, in addition,

(j) $q = 2^n\{2^m(2^l+1)+1\} - 1$, $n \ge 0$, $m, l > 0$

then

3) $r_s = n$, $r_t = n+m$, $r_q = n+m+l$

4) *the assertions 1), 2) hold for*

$$\begin{aligned} c(q) \;:=\; & 2\max\left\{1, \max\left\{\prod_{r=i}^{r_q-1}\left(a_{\mathrm{S}}^{(r)} + a_{\mathrm{T}}^{(r)}\right)\middle|\, 0 \le i \le r_q - 1\right\}\right\} \\ & * \prod_{r=r_t+2}^{r_q}\left(1 + b_{\mathrm{U}}^{(r)}\right)\left|\lambda_{k_{r_t}}^{(r_t)} - \gamma_1^{(r_t)}\right|\rho\Big(\big(\mathrm{A} - \gamma_1^{(r_t)}\mathrm{U}\big)^{-1}\mathrm{U}\Big) \\ & * \left\{b_{\mathrm{U}}^{(r_t+1)} b_{\mathrm{S}}^{(r_t)}\left(x^{(r_t-1)}w^{(r_t-1)}\right) + b_{\mathrm{T}}^{(r_t)} b_{\mathrm{S}}^{(r_t-1)}\left(x^{(r_t-2)}w^{(r_t-2)}\right)\right. \\ & \quad \left. + \, b_{\mathrm{TU}}^{(r_t+1)} b_{\mathrm{A}}^{(r_t)}\left(1 + a_{\mathrm{U}}^{(r_t)}\right) \prod_{r=r_s+1}^{r_t-1}\left(1 + b_{\mathrm{U}}^{(r)}\right) a_{\mathrm{T}}^{(r_s)}\right\} \end{aligned}$$

where all constants are defined by Lemma 4.8 (A *replaced by* $\mathrm{i}(\mathbf{A})$ *for (b), (c), (d)).*

If, in addition,

(k) $a_{\mathrm{S}}^{(r)} + a_{\mathrm{T}}^{(r)} \le 1 \quad \forall i \in \{0, \ldots, r_q - 1\}$ *and* $\rho(\mathrm{S}) \le \rho(\mathrm{T})$

then

5) $\max\limits_{1\le j\le q}\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_{\mathrm{c}}\} = \|\mathrm{d}(\tilde{\mathbf{x}}_{2^{r_q}})\|_{\mathrm{c}}$.

Proof. In the following proof, we use (2.9)–(2.13) without mentioning it explicitly. Assertion 3) obviously follows from the definition of q in (j) as $q = 2^{n+m+l} + 2^{n+m} + 2^n - 1$. The first deviation from an optimal enclosure occurs for $r = r_t + 1$, i.e. the first step in the reduction phase after a change from an even r to an odd r: The width increases while treating the linear system in the computation of $\mathbf{q}^{r+1}_{j_{r+1}}$ if q_r is odd. More precisely, for $r = r_t$ and $j = j_{r_t+1}$, we note with (3.6c)

$$\begin{aligned}
\mathbf{q}^{r+1}_{j} &= \mathrm{U}^{(r+1)}\mathbf{p}^{r+1}_{j} + \mathrm{S}^{(r)}\mathbf{q}^{r}_{j-2^r} \\
&\quad + L\tilde{E}S\big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{A}^{(r)}\big\{\mathrm{U}^{(r+1)}\mathbf{p}^{r+1}_{j} + \mathrm{T}^{(r)}\mathbf{q}^{r}_{j+2^r}\big\}\big) \qquad (22)\\
&= \tilde{\mathbf{q}}^{r+1}_{j} + \hat{\mathbf{q}}^{r+1}_{j}
\end{aligned}$$

$$\begin{aligned}
\tilde{\mathbf{q}}^{r+1}_{j} &= \mathrm{U}^{(r+1)}\tilde{\mathbf{p}}^{r+1}_{j} + \mathrm{S}^{(r)}\tilde{\mathbf{q}}^{r}_{j-2^r} \\
&\quad + \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)}\big\{\mathrm{U}^{(r+1)}\tilde{\mathbf{p}}^{r+1}_{j} + \mathrm{T}^{(r)}\tilde{\mathbf{q}}^{r}_{j+2^r}\big\}
\end{aligned} \qquad (23)$$

$$\begin{aligned}
\hat{\mathbf{q}}^{r+1}_{j} &:= \big|\lambda^{(r)}_{k_r} - \gamma^{(r)}_1\big|\mathrm{U}L\tilde{E}S\Big(\big(\mathbf{A} - \gamma^{(r)}_1\mathrm{U}\big)\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{A}^{(r)}\big(\mathrm{U}^{(r+1)}\mathrm{d}(\tilde{\mathbf{p}}^{r+1}_{j}) \\
&\qquad\qquad + \mathrm{T}^{(r)}\mathrm{d}(\tilde{\mathbf{q}}^{r}_{j+2^r})\big)[-1,1]\Big) \\
&= \big|\lambda^{(r)}_{k_r} - \gamma^{(r)}_1\big|\mathrm{U}\big(\mathrm{A} - \gamma^{(r)}_1\mathrm{U}\big)^{-1}\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{A}^{(r)}\big(\mathrm{U}^{(r+1)}\mathrm{d}(\tilde{\mathbf{p}}^{r+1}_{j}) \\
&\qquad\qquad + \mathrm{T}^{(r)}\mathrm{d}(\tilde{\mathbf{q}}^{r}_{j+2^r})\big)[-1,1].
\end{aligned}$$

For the case (a) which we consider first in order to facilitate the discussion. For the same reason, we further assume even q_r for $r \geq r_t + 1$ which is equivalent to (j) and to the property that (3.6c) is used only once in the whole algorithm (3.5) for *IBUD*. (j) can be expressed by

$$q_r \text{ is} \begin{cases} \text{even} & \text{for } r_s \leq r \leq r_t - 1 \text{ and } r_t + 1 \leq r \leq r_q \\ \text{odd} & \text{for } r = r_t \text{ and } 0 \leq r \leq r_s - 1. \end{cases} \qquad (4.24)$$

According to the definition of r_s and r_t, $\mathbf{p}^{r}_{j_r}$, $\mathbf{q}^{r}_{j_r}$ are computed with (3.6a) for $0 \leq r \leq r_s$, with (3.6b) for $r_s + 1 \leq r \leq r_t$, for $r = r_t + 1$ with (3.6c) and for $r_t + 2 \leq r \leq r_q$ with (3.6b) or (3.6c) in the general case, under restriction (j) only with (3.6b). In the sequel, we extensively use the commutativity of the matrices $\mathrm{A}^{(r)}$, $\mathrm{B}^{(r)}$, $\mathrm{S}^{(r)}$, $\mathrm{T}^{(r)}$, $\mathrm{U}^{(r)}$, and of all matrices derived from them, which results from the commutativity of A, S, and T, hence also of U, without mentioning it explicitly. With (j), we only admit even q_r for $r \geq r_t + 1$. In this case $j_r = j_{r+1}$, hence $j_r - 2^{r+1} + 2^r < j_r = j_{r+1}$, i.e. the additional width introduced in (3.6c) propagates only to $\mathbf{p}^{r}_{j_r}$ and $\mathbf{q}^{r}_{j_r}$ for $r \geq r_t + 2$. We then get

for $r = r_t + 2$, $j = j_r$ (25)

$$\begin{aligned}
\mathbf{p}_j^r &= \mathbf{p}_j^{r-1} + L\tilde{E}S\big(\mathbf{B}^{(r-1)}, \mathbf{C}^{(r-1)}\{\mathrm{S}^{(r-1)}\mathbf{p}_{j-2^{r-1}}^{r-1} + \mathbf{q}_j^{r-1}\}\big) \\
&= \tilde{\mathbf{p}}_j^{r-1} + L\tilde{E}S\big(\mathbf{B}^{(r-1)}, \mathbf{C}^{(r-1)}\{\mathrm{S}^{(r-1)}\tilde{\mathbf{p}}_{j-2^{r-1}}^{r-1} + (\tilde{\mathbf{q}}_j^{r-1} + \hat{\mathbf{q}}_j^{r-1})\}\big) \\
&\subseteq \tilde{\mathbf{p}}_j^r + \hat{\mathbf{p}}_j^r \\
\tilde{\mathbf{p}}_j^r &:= \tilde{\mathbf{p}}_j^{r-1} + L\tilde{E}S\big(\mathbf{B}^{(r-1)}, \mathbf{C}^{(r-1)}\{\mathrm{S}^{(r-1)}\tilde{\mathbf{p}}_{j-2^{r-1}}^{r-1} + \tilde{\mathbf{q}}_j^{r-1}\}\big) \\
\hat{\mathbf{p}}_j^r &:= L\tilde{E}S\big(\mathbf{B}^{(r-1)}, \mathbf{C}^{(r-1)}\hat{\mathbf{q}}_j^{r-1}\big) = \big(\mathrm{B}^{(r-1)}\big)^{-1}\mathbf{C}^{(r-1)}\hat{\mathbf{q}}_j^{r-1}
\end{aligned}$$

and

$$\begin{aligned}
&\text{for } r > r_t + 2,\ j = j_r \\
\hat{\mathbf{p}}_j^r &:= \hat{\mathbf{p}}_j^{r-1} + \big(\mathrm{B}^{(r-1)}\big)^{-1}\mathbf{C}^{(r-1)}\hat{\mathbf{q}}_j^{r-1}
\end{aligned} \tag{26}$$

as well as

$$\begin{aligned}
&\text{for } r \geq r_t + 2,\ j = j_r \\
\mathbf{q}_j^r &= \mathrm{S}^{(r-1)}\mathbf{q}_{j-2^{r-1}}^{r-1} + \mathrm{U}^{(r)}\mathbf{p}_j^r \\
&\subseteq \mathrm{S}^{(r-1)}\tilde{\mathbf{q}}_{j-2^{r-1}}^{r-1} + \mathrm{U}^{(r)}(\tilde{\mathbf{p}}_j^r + \hat{\mathbf{p}}_j^r) \subseteq \tilde{\mathbf{q}}_j^r + \hat{\mathbf{q}}_j^r \\
\tilde{\mathbf{q}}_j^r &= \mathrm{S}^{(r-1)}\tilde{\mathbf{q}}_{j-2^{r-1}}^{r-1} + \mathrm{U}^{(r)}\tilde{\mathbf{p}}_j^r \\
\hat{\mathbf{q}}_j^r &:= \mathrm{U}^{(r)}\hat{\mathbf{p}}_j^r.
\end{aligned} \tag{27}$$

Note that $\hat{\mathbf{p}}_{j_r}^r$, $\hat{\mathbf{q}}_{j_r}^r$ are symmetric interval vectors, i.e. (2.14, b) applies for each application of LES in $L\tilde{E}S$ and the resulting vectors are again symmetric.

The solution phase is started with ($r_t < r_q$, otherwise q, i.e. the enclosure, would be optimal)

$$\begin{aligned}
&r = r_q,\ j = j_r = 2^{r_q} \\
\mathbf{x}_j &= \mathbf{p}_j^r + L\tilde{E}S\big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\mathbf{q}_j^r\big) \\
&\subseteq (\tilde{\mathbf{p}}_j^r + \hat{\mathbf{p}}_j^r) + L\tilde{E}S\big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}(\tilde{\mathbf{q}}_j^r + \hat{\mathbf{q}}_j^r)\big) \subseteq \tilde{\mathbf{x}}_j + \hat{\mathbf{x}}_j \\
\tilde{\mathbf{x}}_j &= \tilde{\mathbf{p}}_j^r + L\tilde{E}S\big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\tilde{\mathbf{q}}_j^r\big) \\
\hat{\mathbf{x}}_j &= \hat{\mathbf{p}}_j^r + L\tilde{E}S\big(\mathbf{B}^{(r)}, \mathbf{C}^{(r)}\hat{\mathbf{q}}_j^r\big) = \hat{\mathbf{p}}_j^r + \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\hat{\mathbf{q}}_j^r.
\end{aligned} \tag{28}$$

In the steps $r = r_q - 1$ to $r = 0$, the additional width propagates to all $\mathbf{x}_j$. To continue the enclosure of this error, we first note that, according to (4.24), $j_r = j_{r+1}$, i.e. $j_{r+1} - 2^{r+1} < j_{r+1}$ for $r \geq r_t + 1$. Therefore, $\mathbf{p}_{j_r}^r$, $\mathbf{q}_{j_r}^r$

are not optimal for $r \geq r_t + 1$ and optimal for $r \leq r_t$. In the solution phase, only for $r = r_q$ they influence the computation of $\mathbf{x}$ under the assumption (4.24), as $\mathbf{x}_{j_r} = \mathbf{x}_{j_{r_q}} = \mathbf{x}_{2^{r_q}}$ has been computed already in the step $r = r_q$. With $\mathbf{x}_0 := \mathrm{o}$, we then obtain

$$\begin{aligned}
&r = r_q - 1(-1)0,\ j := 2^r(2^{r+1})j_r - 2^r \qquad (29)\\
\mathbf{x}_j &= \mathbf{p}_j^r + L\tilde{\acute{E}}S\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\mathbf{x}_{j-2^r} + \mathrm{T}^{(r)}\mathbf{x}_{j+2^r} + \mathbf{q}_j^r\big)\\
&\subseteq \tilde{\mathbf{p}}_j^r + L\tilde{\acute{E}}S\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}(\tilde{\mathbf{x}}_{j-2^r} + \hat{\mathbf{x}}_{j-2^r}) + \mathrm{T}^{(r)}(\tilde{\mathbf{x}}_{j+2^r} + \hat{\mathbf{x}}_{j+2^r}) + \tilde{\mathbf{q}}_j^r\big)\\
&\subseteq \tilde{\mathbf{x}}_j + \hat{\mathbf{x}}_j\\
\tilde{\mathbf{x}}_j &= \tilde{\mathbf{p}}_j^r + L\tilde{\acute{E}}S\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\tilde{\mathbf{x}}_{j-2^r} + \mathrm{T}^{(r)}\tilde{\mathbf{x}}_{j+2^r} + \tilde{\mathbf{q}}_j^r\big)\\
\hat{\mathbf{x}}_j &= L\tilde{\acute{E}}S\big(\mathbf{A}^{(r)}, \mathrm{S}^{(r)}\hat{\mathbf{x}}_{j-2^r} + \mathrm{T}^{(r)}\hat{\mathbf{x}}_{j+2^r}\big) = \big(\mathrm{A}^{(r)}\big)^{-1}\big(\mathrm{S}^{(r)}\hat{\mathbf{x}}_{j-2^r} + \mathrm{T}^{(r)}\hat{\mathbf{x}}_{j+2^r}\big).
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
&r = r_q - 1(-1)\min\{r_s, r_q - 1\} + 1,\ j = j_r \qquad (30)\\
\mathbf{x}_j &= \mathbf{p}_j^r + L\tilde{\acute{E}}S\big(\mathbf{B}^{(r)}, \mathrm{C}^{(r)}\big\{\mathrm{S}^{(r)}\mathbf{x}_{j-2^r} + \mathbf{q}_j^r\big\}\big) \subseteq \tilde{\mathbf{x}}_j + \hat{\mathbf{x}}_j\\
\tilde{\mathbf{x}}_j &= \tilde{\mathbf{p}}_j^r + L\tilde{\acute{E}}S\big(\mathbf{B}^{(r)}, \mathrm{C}^{(r)}\big\{\mathrm{S}^{(r)}\tilde{\mathbf{x}}_{j-2^r} + \tilde{\mathbf{q}}_j^r\big\}\big)\\
\hat{\mathbf{x}}_j &:= L\tilde{\acute{E}}S\big(\mathbf{B}^{(r)}, \mathrm{C}^{(r)}\mathrm{S}^{(r)}\hat{\mathbf{x}}_{j-2^r}\big) = \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{S}^{(r)}\hat{\mathbf{x}}_{j-2^r}
\end{aligned}$$

with odd q_r $\big($see (3.5)$\big)$. We now repeatedly insert (4.25) in (4.26), (4.27) and vice versa in order to finally obtain from (4.28):

$$\hat{x}_{2^{r_q}} = \prod_{r=r_t+2}^{r_q} \Big(\mathrm{I} + \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{U}^{(r)}\Big)\big(\mathrm{B}^{(r_t+1)}\big)^{-1}\mathrm{C}^{(r_t+1)}\hat{\mathbf{q}}_{j_{r_t+1}}^{r_t+1}. \qquad (31)$$

Before inserting (4.23) into (4.31), we first estimate the term $\mathrm{U}^{(r+1)}\mathrm{d}\big(\tilde{\mathbf{p}}_{j_r}^{(r+1)}\big) + \mathrm{T}^{(r)}\mathrm{d}\big(\tilde{\mathbf{q}}_{j_r+2^r}^r\big)$ for $r = r_t$ in (4.23). In case (a), the optimal enclosure is given by $\tilde{\mathbf{x}} = \lfloor \mathrm{M}^{-1}\mathrm{i}(\mathbf{b}), \mathrm{M}^{-1}\mathrm{s}(\mathbf{b})\rfloor$, i.e. $\tilde{\mathrm{x}} = \mathrm{M}^{-1}\tilde{\mathrm{b}}$ with $\tilde{\mathrm{b}} := \mathrm{d}(\mathbf{b})$, $\tilde{\mathrm{x}} := \mathrm{d}(\tilde{\mathbf{x}})$. we then formally apply the interval algorithm (3.5) to the point system $\mathrm{M}\tilde{\mathrm{x}} = \tilde{\mathrm{b}}$ and we get auxiliary vectors $\tilde{\mathrm{p}}_j^r := \mathrm{d}(\tilde{\mathbf{p}}_j^r)$, $\tilde{\mathrm{q}}_j^r := \mathrm{d}(\tilde{\mathbf{q}}_j^r)$. For convenience, we define the abbreviations:

$$\tilde{\mathrm{w}}^{s,j} := \sum_{k=1}^{2^s} \big\{\mathrm{S}^{2^s-1+k}\mathrm{T}^{2^s-k}\mathrm{d}\big(\tilde{\mathbf{x}}_{j-(2k-1)}\big) + \mathrm{S}^{2^s-k}\mathrm{T}^{2^s-1+k}\mathrm{d}\big(\tilde{\mathbf{x}}_{j+(2k-1)}\big)\big\}$$

$$\begin{aligned}
\mathrm{M}^{(r)} &= \prod_{i=r_s+1}^{r} \Big(\mathrm{I} + \big(\mathrm{B}^{(i)}\big)^{-1}\mathrm{C}^{(i)}\mathrm{U}^{(i)}\Big)\big(\mathrm{A}^{(r_s)}\big)^{-1}\mathrm{T}^{(r_s)} \\
\mathrm{X}^{(r)} &= \prod_{i=0}^{r} \big(\mathrm{A}^{(i)}\big)^{-1} \\
\tilde{\mathrm{z}}^{r,j} &:= \big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{X}^{(r-1)}\mathrm{S}^{(r)}\tilde{\mathrm{w}}^{r-1,j-2^r}.
\end{aligned} \tag{32}$$

Using the relations

$$\begin{aligned}
\mathrm{B}^{(r+1)}\big(\mathrm{C}^{(r+1)}\big)^{-1} &= \mathrm{A}^{(r)}\mathrm{B}^{(r)}\big(\mathrm{C}^{(r)}\big)^{-1} - \mathrm{U}^{(r+1)} \\
\mathrm{B}^{(r)}\big(\mathrm{C}^{(r)}\big)^{-1} &= \mathrm{A}^{(r)} \quad (0 \le r \le r_s)
\end{aligned} \tag{4.33}$$

the equations (3.2) and $\tilde{\mathrm{y}}_j^r = \mathrm{A}^{(r)}\tilde{\mathrm{p}}_j^r + \tilde{\mathrm{q}}_j^r \; (j \neq j_r)$ and $\tilde{\mathrm{y}}_j^r = \mathrm{B}^{(r)}\big(\mathrm{C}^{(r)}\big)^{-1}\tilde{\mathrm{p}}_j^r + \tilde{\mathrm{q}}_j^r \; (j = j_r)$, we can show by induction that

for $0 \le r \le r_s - 1$, $j := 2^{r+1}(2^{r+1})j_{r+1}$ (34)

$$\begin{aligned}
\mathrm{d}\big(\tilde{\mathbf{p}}_j^{r+1}\big) &= \mathrm{d}\big(\tilde{\mathbf{x}}_j\big) - \mathrm{X}^{(r)}\tilde{\mathrm{w}}^{r,j} \\
\mathrm{d}\big(\tilde{\mathbf{q}}_j^{r+1}\big) &= -\big\{\mathrm{S}^{(r+1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j-2^{r+1}}\big) + \mathrm{T}^{(r+1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j+2^{r+1}}\big)\big\} + \mathrm{A}^{(r+1)}\mathrm{X}^{(r)}\tilde{\mathrm{w}}^{r,j}
\end{aligned}$$

for $r_s \le r \le r_t - 1$, $j := j_{r+1}$ (35)

$$\begin{aligned}
\mathrm{d}\big(\tilde{\mathbf{p}}_j^{r+1}\big) &= \mathrm{d}\big(\tilde{\mathbf{x}}_j\big) - \mathrm{M}^{(r)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big) - \tilde{\mathrm{z}}^{r,j} \\
\mathrm{d}\big(\tilde{\mathbf{q}}_j^{r+1}\big) &= -\big\{\mathrm{S}^{(r+1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j-2^{r+1}}\big) + \mathrm{U}^{(r+1)}\mathrm{M}^{(r)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big)\big\} \\
&\quad + \mathrm{B}^{(r+1)}\big(\mathrm{C}^{(r+1)}\big)^{-1}\tilde{\mathrm{z}}^{r,j}
\end{aligned}$$

and finally

for $r = r_t$, $j := j_{r+1}$ (36)

$$\begin{aligned}
\mathrm{d}\big(\tilde{\mathbf{p}}_j^{r+1}\big) &:= \mathrm{d}\big(\tilde{\mathbf{x}}_j\big) - \big(\mathrm{A}^{(r)}\big)^{-1}\Big(\mathrm{T}^{(r)}\mathrm{M}^{(r-1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big) \\
&\qquad + \mathrm{S}^{(r)}\mathrm{X}^{(r-1)}\tilde{\mathrm{w}}^{r-1,j-2^r} + \mathrm{T}^{(r)}\tilde{\mathrm{z}}^{r-1,j+2^r}\Big).
\end{aligned}$$

With (4.35) and (4.36), we get for $r = r_t + 1$, $j := j_{r+1}$

$$\begin{aligned}
&\mathrm{U}^{(r_t+1)}\mathrm{d}\big(\tilde{\mathbf{p}}_{j_{r_t}}^{r_t+1}\big) + \mathrm{T}^{(r_t)}\mathrm{d}\big(\tilde{\mathbf{q}}_{j_{r_t}+2^{r_t}}^{r_t}\big) \\
&\quad = \mathrm{U}^{(r_t+1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_t}}\big) - \mathrm{U}^{(r_t+1)}\big(\mathrm{A}^{(r_t)}\big)^{-1} \\
&\qquad * \Big(\mathrm{T}^{(r_t)}\mathrm{M}^{(r_t-1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big) + \mathrm{S}^{(r_t)}\mathrm{X}^{(r_t-1)}\tilde{\mathrm{w}}^{r_t-1,j_{r_t}-2^{r_t}} + \mathrm{T}^{(r_t)}\tilde{\mathrm{z}}^{r_t-1,j_{r_t}+2^{r_t}}\Big)
\end{aligned} \tag{37}$$

$$
\begin{aligned}
&\quad - \left\{ \mathrm{S}^{(r_t)}\mathrm{T}^{(r_t)}\mathrm{d}\big(\tilde{\mathbf{x}}_j\big) + \mathrm{T}^{(r_t)}\mathrm{U}^{(r_t)}\mathrm{M}^{(r_t-1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big)\right\} \\
&\quad + \mathrm{B}^{(r_t)}\big(\mathrm{C}^{(r_t)}\big)^{-1}\mathrm{T}^{(r_t)}\tilde{\mathrm{z}}^{r_t-1,j_{r_t}+2^{r_t}} \\
= &\; -\mathrm{U}^{(r_t+1)}\mathrm{S}^{(r_t)}\mathrm{X}^{(r_t)}\tilde{\mathrm{w}}^{r_t-1,j_{r_t}-2^{r_t}} \\
&\quad - \Big(\mathrm{I} + \big(\mathrm{A}^{(r_t)}\big)^{-1}\mathrm{U}^{(r_t)}\Big)\mathrm{T}^{(r_t)}\mathrm{U}^{(r_t)}\mathrm{M}^{(r_t-1)}\mathrm{d}\big(\tilde{\mathbf{x}}_{j_{r_s}+2^{r_s}}\big) \\
&\quad + \big(\mathrm{A}^{(r_t)}\big)^{-1}\mathrm{B}^{(r_t+1)}\big(\mathrm{C}^{(r_t+1)}\big)^{-1}\mathrm{T}^{(r_t)}\tilde{\mathrm{z}}^{r_t-1,j+2^{r_t}}.
\end{aligned}
$$

We insert (4.32), (4.37) in (4.23) and then (4.23) in (4.31). As $\hat{x}_{2^{r_q}}$ is a symmetric interval vector, it is sufficient to estimate its width. We use (2.2) and the nonnegativity of all matrices whose spectral radius has been estimated in Lemma 4.8. According to this Lemma, all these spectral radii $\rho(.)$ are positive eigenvalues of the respective matrices and they are all functions of $\lambda_{min}(\mathrm{A}) = 1/\rho(\mathrm{A}^{-1})$ and $\omega_{max}(\mathrm{U}) = \rho(\mathrm{U})$. Then (4.5, i) shows the existence of a vector $\mathrm{c} > 0$, which is an eigenvector for all matrices in Lemma 4.8 (f)–(j). From Lemma 2.1, we deduce $\|.\|_{\mathrm{c,c}} \le \rho(.)$ in the **same** matrix norm for all these matrices. This results in the inequality

$$
\|\mathrm{d}(\hat{x}_{2^{r_q}})\|_{\mathrm{c}} \le c_4 \max_{1\le j\le q}\left\{\|\mathrm{d}(\tilde{\mathrm{x}}_j)\|_{\mathrm{c}}\right\}
$$

where

$$
\begin{aligned}
c_4 &= 2\textstyle\prod_{r=r_t+2}^{r_q}\Big(1+\big\|\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{U}^{(r)}\big\|_{\mathrm{c,c}}\Big)\big|\lambda_{k_{r_t}}^{(r_t)} - \gamma_1^{(r_t)}\big| \\
&\quad *\Big(\big\|\big(\mathrm{A}-\gamma_1^{(r_t)}\mathrm{U}\big)^{-1}\mathrm{U}\big\|_{\mathrm{c,c}}(c_1+c_2)+c_3\Big) \\
c_1 &:= \big\|\big(\mathrm{B}^{(r_t+1)}\big)^{-1}\mathrm{C}^{(r_t+1)}\mathrm{U}^{(r_t+1)}\big\|_{\mathrm{c,c}}\big\|\big(\mathrm{B}^{(r_t)}\big)^{-1}\mathrm{C}^{(r_t)}\mathrm{S}^{(r_t)}\big\|_{\mathrm{c,c}} \\
&\quad *\big(\big\|\mathrm{X}^{(r_t-1)}\big\|_{\mathrm{c,c}}v^{(r_t-1)}\big) \\
c_2 &:= \big\|\big(\mathrm{B}^{(r_t)}\big)^{-1}\mathrm{C}^{(r_t)}\mathrm{T}^{(r_t)}\big\|_{\mathrm{c,c}}\big\|\big(\mathrm{B}^{(r_t-1)}\big)^{-1}\mathrm{C}^{(r_t-1)}\mathrm{S}^{(r_t-1)}\big\|_{\mathrm{c,c}} \\
&\quad *\big(\big\|\mathrm{X}^{(r_t-2)}\big\|_{\mathrm{c,c}}v^{(r_t-2)}\big) \\
c_3 &:= \big\|\big(\mathrm{B}^{(r_t+1)}\big)^{-1}\mathrm{C}^{(r_t+1)}\mathrm{T}^{(r_t)}\mathrm{U}^{(r_t)}\big\|_{\mathrm{c,c}} \\
&\quad *\big\|\big(\mathrm{B}^{(r_t)}\big)^{-1}\mathrm{C}^{(r_t)}\mathrm{A}^{(r_t)}\big(\mathrm{A}-\gamma_1^{(r_t)}\mathrm{U}\big)^{-1}\mathrm{U}\big\|_{\mathrm{c,c}}\Big(1+\big\|\big(\mathrm{A}^{(r_t)}\big)^{-1}\mathrm{U}^{(r_t)}\big\|_{\mathrm{c,c}}\Big) \\
&\quad *\prod_{r=r_s+1}^{r_t-1}\Big(1+\big\|\big(\mathrm{B}^{(r)}\big)^{-1}\mathrm{C}^{(r)}\mathrm{U}^{(r)}\big\|_{\mathrm{c,c}}\Big)\big\|\big(\mathrm{A}^{(r_s)}\big)^{-1}\mathrm{T}^{(r_s)}\big\|_{\mathrm{c,c}} \\
v^{(r)} &:= \sum_{k=1}^{2^r}\left\{\|\mathrm{S}\|_{\mathrm{c,c}}{}^{2^r-1+k}\|\mathrm{T}\|_{\mathrm{c,c}}{}^{2^r-k} + \|\mathrm{S}\|_{\mathrm{c,c}}{}^{2^r-k}\|\mathrm{T}\|_{\mathrm{c,c}}{}^{2^r-1+k}\right\}.
\end{aligned}
\tag{4.38}
$$

Before terminating the proof of assertion 4), i.e. of assertion 1) under restriction (j), we note

$$\text{for } r := r_q - 1(-1)0: \tag{4.39a}$$

$$\begin{aligned}\max_{j=2^r(2^{r+1})j_r-2^r}\left\{\left\|\mathrm{d}(\hat{\mathbf{x}}_j)\right\|_{\mathrm{c}}\right\} &\le \left(\left\|\left(\mathrm{A}^{(r)}\right)^{-1}\mathrm{S}^{(r)}\right\|_{\mathrm{c,c}} + \left\|\left(\mathrm{A}^{(r)}\right)^{-1}\mathrm{T}^{(r)}\right\|_{\mathrm{c,c}}\right) \\ &\qquad * \max_{j=2^{r+1}(2^{r+1})j_r}\left\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_{\mathrm{c}}\right\} \\ &\le \left(a_{\mathrm{S}}^{(r)} + a_{\mathrm{T}}^{(r)}\right)\max_{j=2^{r+1}(2^{r+1})j_r}\left\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_{\mathrm{c}}\right\}\end{aligned}$$

and, again with Lemma 4.8,

$$\text{for } r := r_q - 1(-1)0: \tag{4.39b}$$

$$\begin{aligned}\|\mathrm{d}(\hat{\mathbf{x}}_{j_r})\|_{\mathrm{c}} &\le \begin{cases}\left\|\left(\mathrm{B}^{(r)}\right)^{-1}\mathrm{C}^{(r)}\mathrm{S}^{(r)}\right\|_{\mathrm{c,c}}\|\mathrm{d}(\hat{\mathbf{x}}_{j-2^r})\|_{\mathrm{c}} \\ \qquad \left(r := r_q - 1(-1)\min\{r_s, r_q - 1\} + 1\right) \\ \left\|\left(\mathrm{A}^{(r)}\right)^{-1}\mathrm{S}^{(r)}\right\|_{\mathrm{c,c}}\|\mathrm{d}(\hat{\mathbf{x}}_{j-2^r})\|_{\mathrm{c}} \\ \qquad \left(r := \min\{r_s, r_q - 1\}(-1)0\right)\end{cases} \\ &\le \|\mathrm{d}(\hat{\mathbf{x}}_{j-2^r})\|_{\mathrm{c}}.\end{aligned}$$

Note also that according to our assumption (4.24), q is defined by r_s, r_t, i.e. $c(q) \equiv c(r_s, r_t)$. With

$$c'(q) := c_4 \max\left\{1, \max\left\{\prod_{r=i}^{r_q-1}\left(\left\|\left(\mathrm{A}^{(r)}\right)^{-1}\mathrm{S}^{(r)}\right\|_{\mathrm{c,c}} + \left\|\left(\mathrm{A}^{(r)}\right)^{-1}\mathrm{T}^{(r)}\right\|_{\mathrm{c,c}}\right) \;\middle|\; 0 \le i \le r_q - 1\right\}\right\} \tag{4.40}$$

and using Lemma 4.8, we get $c'(q) \le c(q)$ with $c(q)$ defined in assertion 4). Under restriction (j) the monotone convergence $c(q) \downarrow 0$ $(d \uparrow \infty)$, i.e. assertion 2), immediately follows with Lemma 4.8 using $d \uparrow \infty \Leftrightarrow \theta \uparrow \infty \Leftrightarrow \lambda' := \lambda + d \uparrow \infty$ and the monotonicity in d from the right-hand side in the relation $\left\|\left(\mathrm{B}^{(r_t)}\right)^{-1}\mathrm{C}^{(r_t)}\mathrm{A}^{(r_t)}\left(\mathrm{A} - \gamma_1^{(r_t)}\mathrm{U}\right)^{-1}\mathrm{U}\right\|_{\mathrm{c,c}} \le b_{\mathrm{A}}^{(r_t)}\left\|\left(\mathrm{A} - \gamma_1^{(r_t)}\mathrm{U}\right)^{-1}\mathrm{U}\right\|_{\mathrm{c,c}}$. With (4.38), we have completed the proof of assertion 4). From (4.39) we immediately deduce assertion 5).

If we remove the restriction (j) in order to prove assertions 1), 2) for general q, we are faced with considerably more complicated enclosures. First, (4.23) has to be replaced by

$$\begin{aligned}\hat{\mathbf{q}}_j^{r+1} := {} & \mathrm{U}^{(r+1)}\hat{\mathbf{p}}_j^{r+1} + \left|\lambda_{k_r}^{(r)} - \gamma_1^{(r)}\right|\mathrm{U}\,LES\Big(\left(\mathbf{A} - \lambda_{k_r}^{(r)}\mathrm{U}\right)\mathbf{B}^{(r)}, \\ & \mathbf{C}^{(r)}\mathbf{A}^{(r)}\left\{\mathrm{U}^{(r+1)}\left(\mathrm{d}(\tilde{\mathbf{p}}_j^{r+1}) + \mathrm{d}(\hat{\mathbf{p}}_j^{r+1})\right)\right. \\ & \qquad \left. + \mathrm{T}^{(r)}\left(\mathrm{d}(\tilde{\mathbf{q}}_{j+2^r}^r) + \mathrm{d}(\hat{\mathbf{q}}_{j+2^r}^r)\right)\right\}[-1, 1]\Big)\end{aligned} \tag{4.23'}$$

also for possibly several $r \geq r_t + 1$ and $j = j_r$. In the subsequent steps of both the reduction and the solution phase, additional terms depending on $\hat{\mathbf{p}}^r_{j_r}$, $\hat{\mathbf{q}}^r_{j_r}$ have to be added. The assertions 1), 2), however, remain valid, as the proof is based on the behaviour of the spectral radii listed in Lemma 4.8 for $d \to \infty$. Because of the unpredictable number of case decision, a useful closed expression like (4.38) cannot be derived and in addition we cannot guarantee anymore that the maximal error in the chosen norm occurs for $\hat{x}_{2^{r_q}}$.

The whole proof can now be repeated for case (b) by applying (2.7) and (2.14, b) for *LES*. In case (c), (d), we only get a weaker estimate. We simplify (c), (d) from Lemma 4.9 with (2.5) to

$$\mathbf{z} \subseteq \tilde{\mathbf{z}} + \left|\lambda^{(r)}_{k_r} - \gamma^{(r)}_1\right| \left(\mathrm{i}(\mathbf{A}) - \gamma^{(r)}_i \mathrm{U}\right)^{-1} \mathrm{U} \left(\mathrm{B}^{(r)}_{\mathit{inf}}\right)^{-1} \mathrm{C}^{(r)}_{\mathit{inf}} \mathrm{A}^{(r)}_{\mathit{inf}} |\mathrm{s}(\mathbf{y})| [-1, 1] \quad (4.41)$$

and replace (4.23) and all subsequent estimates appropriately by applying (2.8). □

Condition (k), i.e. assertion 5) is satisfied in many cases (compare Lemma 4.8), as $\cosh(2^r\theta_{min}) \gg 1$ often holds. But for $\cosh(2^r\theta_{min}) \approx 1$, we possibly get $a^{(r)}_{\mathrm{S}} + a^{(r)}_{\mathrm{T}} \geq 1$, taking into account that $\frac{\mathrm{S}}{\omega_{max}} = \sqrt{\frac{S}{t}}$ and $\frac{\mathrm{t}}{\omega_{max}} = \sqrt{\frac{t}{S}}$ for $\mathrm{S} \neq \mathrm{T}$, hence $\frac{\mathrm{t}}{\omega_{max}} + \frac{\mathrm{S}}{\omega_{max}} \geq 1$. For $\mathrm{S} = \mathrm{T}$, Theorem 4.10 can be significantly simplified:

Corollary 4.11. *If, in addition,*

k′) $\mathrm{S} = \mathrm{T}$

then

4′) the assertions 1), 2) hold for

$$\begin{aligned} c(q) \;:=\; & 2 \prod_{r=r_t+2}^{r_q} (1 + b^{(r)}) \left|\lambda^{(r_t)}_{k_{r_t}} - \gamma^{(r_t)}_1\right| \rho\left((\mathrm{A} - \gamma^{(r_t)}_1 \mathrm{U})^{-1} \mathrm{U}\right) \\ & * \Big\{ b^{(r_t+1)} b^{(r_t)} 2^{r_t - 1} \prod_{i=0}^{r_t - 1} a^{(i)} + b^{(r_t)} b^{(r_t - 1)} 2^{r_t - 2} \prod_{i=0}^{r_t - 2} a^{(i)} \\ & + b^{(r_t+1)} b^{(r_t)}_{\mathrm{A}} (1 + a^{(r_t)}) \prod_{r=r_s+1}^{r_t - 1} (1 + b^{(r)}) a^{(r_s)} \Big\} \end{aligned}$$

where $b^{(r)} \equiv b^{(r)}_{\mathrm{U}}$, $a^{(r)} \equiv a^{(r)}_{\mathrm{U}}$

$$5')\ \max_{1\le j\le q}\left\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_{\mathrm{c}}\right\} = \|\mathrm{d}(\tilde{\mathbf{x}}_{2^{r_q}})\|_{\mathrm{c}}$$

Proof. In the special case $\mathrm{S} = \mathrm{T}$, i.e. $\forall r \in \{0,\dots,r_q\} : \mathrm{S}^{(r)} = \mathrm{T}^{(r)} = \mathrm{U}^{(r)}$, $c(q)$ can be simplified with $a^{(r)} \equiv a^{(r)}_{\mathrm{U}}$, $b^{(r)} \equiv b^{(r)}_{\mathrm{U}}$, and $x^{(r)}w^{(r)} = 2^r \prod_{i=0}^{r} a^{(i)}$. We further note (compare Lemma 4.8) $a^{(r)}_{\mathrm{S}} + a^{(r)}_{\mathrm{T}} = 2a^{(r)}_{\mathrm{U}} = \frac{1}{\cosh(2^r\theta_{min})} \le 1$. □

The conditions (4.5, g–i) ensure that estimates for nonsymmetric coefficients matrices like that of the five point discretization of the elliptic operator $a(x)u_{xx} + bu_{yy} + c(x)u_x + du_y$ can be treated. In the symmetric case, the Euclidean norm can be used.

Corollary 4.12. *For symmetric* A, S, T, *the assertions 2), 4) of Theorem 4.10 can be improved:*

$$2')\quad \|\mathrm{d}(\hat{\mathbf{x}})\|_\infty \le \begin{cases} \max\limits_{1\le j\le q}\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_2\} \;\le\; c(q)\max\limits_{1\le j\le q}\{\|\mathrm{d}(\tilde{\mathbf{x}}_j)\|_2\} \\ \qquad\qquad\qquad\qquad\;\le\; c(q)\sqrt{p}\,\|\mathrm{d}(\tilde{\mathbf{x}})\|_\infty & a),\ b) \\ \max\limits_{1\le j\le q}\{\|\mathrm{s}(\hat{\mathbf{x}}_j)\|_2\} \;\le\; c(q)\max\limits_{1\le j\le q}\{\||\mathrm{s}(\tilde{\mathbf{x}}_j)|\|_2\} \\ \qquad\qquad\qquad\qquad\;\le\; c(q)\sqrt{p}\||\mathrm{s}(\tilde{\mathbf{x}})|\|_\infty & c),\ d) \end{cases}$$

$$4')\ \max_{1\le j\le q}\{\|\mathrm{d}(\hat{\mathbf{x}}_j)\|_2\} = \|\mathrm{d}(\hat{\mathbf{x}}_{2^{r_q}})\|_2.$$

Proof. If A, S, and T are symmetric, then all matrices needed in the definition of $c(q)$ commute. This implies their symmetry and $\rho(.) = \|.\|_{2,2}$. □

In the cases (a), (b), we have derived the desired estimate for $\mathrm{d}(\hat{\mathbf{x}})$ depending on $\mathrm{d}(\tilde{\mathbf{x}})$. In the cases (c), (d), we have to accept instead a dependence of $\mathrm{d}(\hat{\mathbf{x}})$ on $|\mathrm{s}(\tilde{\mathbf{x}})|$ which is due to (c), (d) of Lemma 4.9, i.e. basically (2.14, c) and (2.14, d).

Corollary 4.13. *With* $\mathrm{C} = (\delta_{i,j} c_i)_{i,j=1}^{p}$, *we get*

$$\|\mathrm{d}(\hat{\mathbf{x}})\|_\infty \leq \begin{cases} c(q)\|\mathrm{C}\|_\infty \max\limits_{1\leq j\leq q}\{\|\mathrm{d}(\tilde{\mathbf{x}}_j)\|_{\mathrm{c}}\} & \\ \qquad \leq c(q)\|\mathrm{C}\|_\infty \, \|\mathrm{C}^{-1}\|_\infty \, \|\mathrm{d}(\tilde{\mathbf{x}})\|_\infty & \text{a), b)} \\ c(q)\|\mathrm{C}\|_\infty \max\limits_{1\leq j\leq q}\{\||\mathrm{s}(\tilde{\mathbf{x}}_j)|\|_{\mathrm{c}}\} & \\ \qquad \leq c(q)\|\mathrm{C}\|_\infty \| \mathrm{C}^{-1}\|_\infty \, \||\mathrm{s}(\tilde{\mathbf{x}})|\|_\infty & \text{c), d).} \end{cases}$$

The corollaries contain simplifications, mainly estimates in the Euclidean and the ∞-norm which can be significant for computational purposes. The eigenvector c defining the $\|.\|_{\mathrm{c}}$-norm should also be accessible in many cases. The practical significance of the estimates containing $\|\mathrm{C}\|_\infty$, $\|\mathrm{C}^{-1}\|_\infty$ depends on the size of these norms in a specific context.

The assertions of Theorem 4.10 are the key for the application of *IBUD* for values of the block dimension q for which optimal enclosures cannot be expected. We have derived an estimate for the additional width of the enclosure vector $\mathit{IBUD}(\mathbf{M}, \mathbf{b})$ when compared to an optimal enclosure in the "usual" four cases, in which such an enclosure is principally possible. Assertion 2) shows that *IBUD* can yield satisfactory enclosures for any value of q under the condition that the coefficient matrix $\mathbf{M}$ is sufficiently diagonally dominant.

For a selected set of values of q, i.e. those which cause only one occurrence of (3.6c)/(3.9), the width can be explicitly estimated by a rather ugly, but very easily computable constant $c(q)$. It is principally possible to derive an estimate for general q. But on the one hand, such an estimate would have to be computed, due to the case decisions which are necessary for general q, by an algorithm rather than by a simple constant like $c(q)$ from assertion 5). On the other hand, more general q than those defined by condition (j) are characterized by more than one application of (3.6c). Numerical results will show that for these values of q acceptable enclosures can be observed only for a very (in practice usually unrealistically) strong diagonal dominance of $\mathbf{M}$. Unfortunately, the constants defining $c(q)$ are not equally monotone in r: while $a_\gamma^{(r)}$, $a_{\mathrm{T}}^{(r)}$, $a_{\mathrm{U}}^{(r)}$, $b_{\gamma\mathrm{A}}^{(r)}$, $b_{\mathrm{S}}^{(r)}$, $b_{\mathrm{T}}^{(r)}$, $b_{\mathrm{U}}^{(r)}$, $b_{\mathrm{TU}}^{(r)}$, $x^{(r)}$ decrease with increasing r, $\bigl|\lambda_{k_{r_t}}^{(r_t)} - \gamma_1^{(r_t)}\bigr|$, $\rho\bigl((\mathrm{A} - \gamma_1^{(r_t)}\mathrm{U})^{-1}\mathrm{U}\bigr)$ or the products involving $1 + a_{\mathrm{U}}^{(r)}$, $1 + b_{\mathrm{U}}^{(r)}$ may increase. Therefore, it is not possible to predict suitable q from monotonicity arguments on r.

Assertion 4) is only valid for the q defined by (j) because for more general q, $\mathbf{x}_j$ in (4.30) can depend also from $\hat{\mathbf{p}}_j^{(r)}$ and $\hat{\mathbf{q}}_j^{(r)}$. This dependence has to be mainly considered in (4.37).

5 Numerical results

The numerical examples have been computed on a workstation IBM RS 6000/560. We have used a simulation of an interval arithmetic which follows the principles described in [1]. All bounds are computed by the near-IEEE floating point arithmetic on the IBM workstation. All rounding errors are included in the computed interval vector. Lower and upper bounds are multiplied by factors $1 - 2^{-53}$ and $1 + 2^{-52}$. Underflow results are avoided by substituting them by one of the constants 0 or $\pm pmin$, where $pmin = 2^{-1022}$ denotes the smallest positive normalized machine number. This simulation represents a compromise between precision and speed. The usually slight overestimates caused by this simulation and indirectly by the underlying floating point arithmetic can be reduced by using an accurate arithmetic like that proposed in [6, 7] and related publications. This has to be paid for by a reduction of speed by roughly at least one order of magnitude. In the following examples the quality of enclosures is indicated by the following relative width

$$rd(x) := \max_{1 \le i \le N} \left\{ \frac{\mathrm{d}(\hat{X}_i)}{\mathrm{d}(\tilde{X}_i)} \right\} \tag{5.1}$$

where $\hat{\mathbf{x}}$ denotes an estimate for the difference of $\mathbf{x} = \mathit{IBUD}(\mathbf{M}, \mathbf{b})$ to the optimal enclosure $\tilde{\mathbf{x}} = \mathbf{x}_{opt}$ in case (a) according to $\mathbf{x} \subseteq \tilde{\mathbf{x}} + \hat{\mathbf{x}}$.

Figure 1 illustrates the dependence of the width on the diagonal dominance of the coefficient matrix for the example $\mathbf{M} \equiv \mathrm{M} = (-\mathrm{I}, \mathrm{A} + d\mathrm{I}, -\mathrm{I})$, $\mathrm{A} = (-1, 4, -1)$, $d = 10^k$, $-7 \le k \le 3$ or $k = 0$, $\mathbf{x}_{opt} = ([1,2])_{i=1}^N$, $\mathbf{b} := [\mathrm{M}\,\mathrm{i}(\mathbf{x}_{opt}), \mathrm{M}\,\mathrm{s}(\mathbf{x}_{opt})]$, $127 \le q \le 146$, $p = 255$.

As predicted, we observe widths which strongly differ depending on q for small d, but which decrease to 0 with increasing diagonal dominance. Table 1 confirms this fact for the same example for some values of d and values $q = 2^n\big(2^m(2^l + 1) + 1\big) - 1$ for which optimal enclosures cannot be expected. This table illustrates the fact that $c(q)$ is a good measure

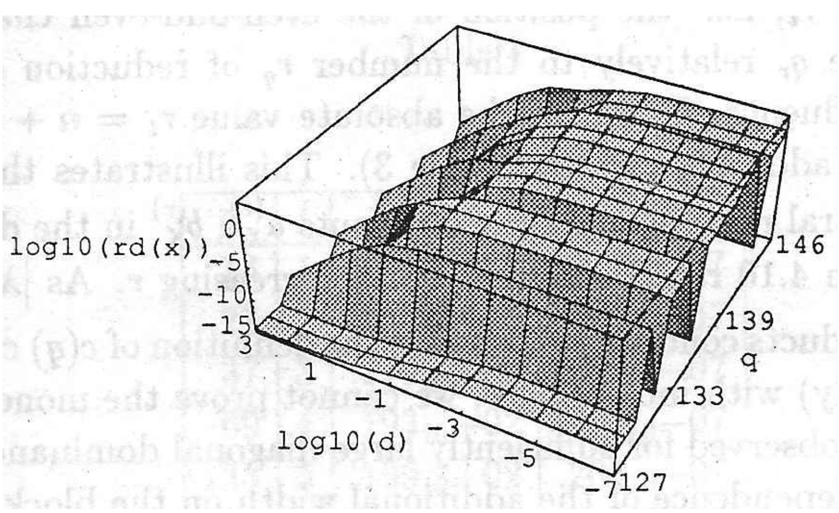

Figure 1

for the relative additional width (5.1). Note that this is not necessarily an upper bound for (5.1) according to its definition in Theorem 4.10 on the basis of maxima over the blocks $1 \leq j \leq q$ in the Euclidean norm. The last column reveals that the computed relative width remains roughly in the order of the relative machine precision while values for $c(q)$ indicate that the theoretical additional width is smaller by orders of magnitude. This is due to unavoidable overestimates of the rounding errors by the simulated interval arithmetic. These overestimates are mainly caused by the fact that in the treatment of the sequences of subsystems with matrices $\mathbf{A}-\alpha\mathrm{U}$, in particular for large r, the absolute values of right-hand sides and intermediate enclosure vectors extremely vary over almost all orders of magnitude while the initial r.h.s. and the final enclosures have a reasonable size. It is typical for all variants of the Buneman algorithm, that, depending on the used floating point arithmetic, underflow frequently occurs already for moderate system sizes. As underflow results are replaced by the next admissible normalized machine number, they can be subject to a significant overestimate of the relative error in subsequent steps.

The Tables 2, 3 show the behaviour of $c(q)$, i.e. the additional width, depending on the constants m, n, l in the definition of q in Theorem 4.10 (j). While $l = r_q - r_t$, i.e. the position of the even-odd-even change in the sequence of the q_r relatively to the number r_q of reduction steps, only has a minor influence (Table 2), the absolute value $r_t = n+m$ strongly determines the additional width (Table 3). This illustrates the fact that almost all spectral radii defining the constants $a_{\mathrm{X}}^{(r)}$, $b_{\mathrm{X}}^{(r)}$ in the definition of $c(q)$ in

q	$d = 0$		$d = 0.1$	
	$c(q)$	$rd(x)$	$c(q)$	$rd(x)$
130	$.363_{10}{+}01$	$.150_{10}{+}01$	$.122_{10}{+}01$	$.672_{10}{+}00$
132	$.103_{10}{+}02$	$.543_{10}{+}01$	$.109_{10}{+}01$	$.975_{10}{+}00$
133	$.199_{10}{+}02$	$.834_{10}{+}01$	$.182_{10}{+}01$	$.141_{10}{+}01$
136	$.218_{10}{+}02$	$.138_{10}{+}02$	$.181_{10}{+}00$	$.242_{10}{+}00$
137	$.423_{10}{+}02$	$.229_{10}{+}02$	$.339_{10}{+}00$	$.365_{10}{+}00$
139	$.789_{10}{+}02$	$.340_{10}{+}02$	$.484_{10}{+}00$	$.464_{10}{+}00$
144	$.397_{10}{+}02$	$.292_{10}{+}02$	$.222_{10}{-}02$	$.367_{10}{-}02$
145	$.780_{10}{+}02$	$.521_{10}{+}02$	$.457_{10}{-}02$	$.561_{10}{-}02$

q	$d = 1$		$d = 10$	
	$c(q)$	$rd(x)$	$c(q)$	$rd(x)$
130	$.769_{10}{-}01$	$.714_{10}{-}01$	$.134_{10}{-}03$	$.134_{10}{-}03$
132	$.501_{10}{-}02$	$.533_{10}{-}02$	$.197_{10}{-}07$	$.197_{10}{-}07$
133	$.637_{10}{-}02$	$.629_{10}{-}02$	$.205_{10}{-}07$	$.205_{10}{-}07$
136	$.520_{10}{-}05$	$.580_{10}{-}05$	$.108_{10}{-}15$	$.355_{10}{-}14$
137	$.665_{10}{-}05$	$.669_{10}{-}05$	$.110_{10}{-}15$	$.355_{10}{-}14$
139	$.689_{10}{-}05$	$.688_{10}{-}05$	$.111_{10}{-}15$	$.355_{10}{-}14$
144	$.220_{10}{-}11$	$.253_{10}{-}11$	$.135_{10}{-}32$	$.333_{10}{-}14$
145	$.285_{10}{-}11$	$.290_{10}{-}11$	$.137_{10}{-}32$	$.333_{10}{-}14$

Table 1

q	l	$d = 0$	$d = 10$
13	1	$.123_{10}{+}02$	$.205_{10}{-}07$
21	2	$.156_{10}{+}02$	$.205_{10}{-}07$
37	3	$.179_{10}{+}02$	$.205_{10}{-}07$
69	4	$.193_{10}{+}02$	$.205_{10}{-}07$
133	5	$.199_{10}{+}02$	$.205_{10}{-}07$
261	6	$.200_{10}{+}02$	$.205_{10}{-}07$
517	7	$.200_{10}{+}02$	$.205_{10}{-}07$

Table 2: $q = 2^{n+m+l} + 2^{n+m} + 2^n - 1 = 4 * 2^l + 5 \quad (n = m = 1)$

q	n	m	$n+m$	$d=0$	$d=0.1$
6	0	1	1	$.216_{10}{+}01$	$.114_{10}{+}01$
12	0	2	2	$.674_{10}{+}01$	$.108_{10}{+}01$
13	1	1	2	$.123_{10}{+}02$	$.181_{10}{+}01$
24	0	3	3	$.156_{10}{+}02$	$.181_{10}{+}00$
25	1	2	3	$.294_{10}{+}02$	$.339_{10}{+}00$
27	2	1	3	$.521_{10}{+}02$	$.484_{10}{+}00$
48	0	4	4	$.313_{10}{+}02$	$.222_{10}{-}02$
49	1	3	4	$.610_{10}{+}02$	$.457_{10}{-}02$
51	2	2	4	$.115_{10}{+}03$	$.697_{10}{-}02$
55	3	1	4	$.201_{10}{+}03$	$.781_{10}{-}02$
96	0	5	5	$.532_{10}{+}02$	$.175_{10}{-}06$
97	1	4	5	$.105_{10}{+}03$	$.382_{10}{-}06$
99	2	3	5	$.204_{10}{+}03$	$.604_{10}{-}06$
103	3	2	5	$.382_{10}{+}03$	$.690_{10}{-}06$
111	4	1	5	$.661_{10}{+}03$	$.694_{10}{-}06$
192	0	6	6	$.567_{10}{+}02$	$.578_{10}{-}15$
193	1	5	6	$.113_{10}{+}03$	$.131_{10}{-}14$
195	2	4	6	$.225_{10}{+}03$	$.211_{10}{-}14$
199	3	3	6	$.439_{10}{+}03$	$.244_{10}{-}14$
207	4	2	6	$.819_{10}{+}03$	$.246_{10}{-}14$
223	5	1	6	$.137_{10}{+}04$	$.246_{10}{-}14$
384	0	7	7	$.199_{10}{+}02$	$.339_{10}{-}32$
385	1	6	7	$.404_{10}{+}02$	$.781_{10}{-}32$
387	2	5	7	$.828_{10}{+}02$	$.127_{10}{-}31$
391	3	4	7	$.171_{10}{+}03$	$.148_{10}{-}31$
399	4	3	7	$.347_{10}{+}03$	$.149_{10}{-}31$
415	5	2	7	$.655_{10}{+}03$	$.149_{10}{-}31$
447	6	1	7	$.103_{10}{+}04$	$.149_{10}{-}31$
768	0	8	8	$.108_{10}{+}01$	$.609_{10}{-}67$
769	1	7	8	$.227_{10}{+}01$	$.142_{10}{-}66$

Table 3: $q = 2^{n+m+1} + 2^{n+m} + 2^n - 1 = 3 * 2^{n+m} + 2^n - 1 \quad (l = 1)$

Theorem 4.10 rapidly decrease with increasing r. As $\left|\lambda_{k_{r_t}}^{(r_t)} - \gamma_1^{(r_t)}\right|$, $a_{\gamma^{\mathrm{A}}}^{(r)}$ and the products concerning $1 + b_{\mathrm{U}}^{(r)}$ in the definition of $c(q)$ can increase (relatively slowly) with increasing r, we cannot prove the monotonicity in r which can be observed for sufficiently large diagonal dominance. Table 4 illustrates the dependence of the additional width on the block size p.

p	$c(q)$
7	$.218_{10}{-}06$
15	$.499_{10}{-}05$
31	$.122_{10}{-}04$
63	$.155_{10}{-}04$
127	$.164_{10}{-}04$
255	$.167_{10}{-}04$
511	$.167_{10}{-}04$

Table 4: $q = 151 \quad p = 2^{k+1} - 1 \quad (2 \le k \le 8)$

q	$k = -3$		$k = -1$	
	$rd(x)$	$c(q)$	$rd(x)$	$c(q)$
130	$.144_{10}{+}00$	$.167_{10}{+}00$	$.148_{10}{+}00$	$.170_{10}{+}00$
132	$.225_{10}{-}01$	$.220_{10}{-}01$	$.264_{10}{-}01$	$.232_{10}{-}01$
133	$.329_{10}{-}01$	$.301_{10}{-}01$	$.340_{10}{-}01$	$.318_{10}{-}01$
136	$.124_{10}{-}03$	$.957_{10}{-}04$	$.133_{10}{-}03$	$.108_{10}{-}03$
137	$.157_{10}{-}03$	$.134_{10}{-}03$	$.168_{10}{-}03$	$.151_{10}{-}03$
139	$.167_{10}{-}03$	$.143_{10}{-}03$	$.179_{10}{-}03$	$.161_{10}{-}03$
144	$.995_{10}{-}09$	$.714_{10}{-}09$	$.115_{10}{-}08$	$.921_{10}{-}09$
145	$.125_{10}{-}08$	$.102_{10}{-}08$	$.144_{10}{-}08$	$.131_{10}{-}08$

Table 5

Table 5 shows similar results for a non-symmetric example where the non-symmetry is determined by the “mesh size” h in accordance with typical discretizations of general elliptic problems which are not given in self-adjoint form. We consider the example

$$\begin{aligned}
\mathbf{M} &= (-\mathrm{S}, \mathbf{A}, -\mathrm{T}),\\
\mathrm{S} &= -\big(b-(h/2)e\big)\mathrm{I},\\
\mathrm{T} &= -\big(b+(h/2)e\big)\mathrm{I},\\
\mathbf{A} &= \Big(-\big(a-(h/2)c\big), 2(a+b), -\big(a+(h/2)c\big)\Big) + [0,1]\mathrm{I} + d\mathrm{I},\\
a &= 4.1, \quad b = 3.1, \quad c = 1.0, \quad e = 1.1, \quad d = 2,\\
\mathbf{x} &= \big([i+j, 2(i+j)]\big)_{i=1}^{p}{}_{j=1}^{q}, \quad p = 255, \quad h = 10^{-k}, \quad k = -3, -1.
\end{aligned}$$

6 Conclusion

The interval Buneman algorithm can be defined, like its noninterval counterpart, for arbitrary values of the block dimension. This increase of flexibility is useful in applications where the system size resulting from the requirement $q = 2^{k+1} - 1$ in the variants of the original Buneman algorithm is not adequate for various reasons. Under appropriate conditions *IBUD* yields optimal enclosures for selected values of the block dimension q. In the present paper, we have extended the range of suitable values by admitting also nonoptimal enclosures. For the special choice $q = 2^n\big(2^m(2^l+1)+1\big)-1$, we can prove suboptimal enclosures already in the presence of a moderate degree of diagonal dominance of the coefficient matrix. The enclosure quality can be controlled by fairly easily computable estimates.

Acknowledgment. I am indebted to the referees for their helpful comments and suggestions.

References

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

[2] Alefeld, G. and Mayer, G. *The Cholesky method for interval data* (in preparation).

[3] Barth, W. and Nuding, E. *Optimale Lösung von Intervallgleichungssystemen.* Computing **12** (1974), pp. 117–125.

[4] Buneman, O. *A compact noniterative Poisson solver.* Institute for Plasma Research Report 294, Stanford University, 1969.

[5] Buzbee, B., Golub, G., and Nielson, C. *On direct methods for solving Poisson's equation.* SIAM J. Num. Anal. **7** (1970), pp. 627–656.

[6] Kulisch, U. and Miranker, W. *Computer arithmetic in theory and practice.* Academic Press, New York, 1981.

[7] Kulisch, U. and Miranker, W. *A new approach to scientific computation.* Academic Press, New York, 1983.

[8] Schwandt, H. *An interval arithmetic approach for the construction of an almost globally convergent method for the solution of the nonlinear Poisson equation on the unit square.* SIAM J. Sci. Stat. Comp. **5** (1984), pp. 427–452.

[9] Schwandt, H. *Interval arithmetic methods for systems of nonlinear equations arising from discretizations of quazilinear elliptic and parabolic partial differential equations.* Appl. Num. Math. **3** (1987), pp. 257–287.

[10] Schwandt, H. *Cyclic reduction for tridiagonal systems of equations with interval coefficients on vector computers.* SIAM J. Num. Anal. **26** (1989), pp. 661–680.

[11] Schwandt, H. *The interval Buneman algorithm for arbitrary block dimension.* To appear in Computing Suppl. **9** (1993).

[12] Schwandt, H. *An interval arithmetic domain decomposition method for a class of elliptic PDE on non-rectangular domains.* To appear in J. Comp. Appl. Math. (1993).

[13] Sweet, R. *A cyclic reduction algorithm for solving block tridiagonal systems of arbitrary dimension.* SIAM J. Num. Anal. **14** (1977), pp. 706–720.

[14] Varga, R. *Matrix iterative analysis.* Prentice Hall, Englewood Cliffs, New Jersey, 1962.

Received: April 1, 1993
Revised version: August 10, 1993

Technische Universität Berlin
Fachbereich Mathematik, MA 6–4
Straße des 17. Juni 136
D–10623 Berlin
Germany
E-mail: `schwandt@math.tu-berlin.de`