SIAM J. Matrix Anal. Appl.
Vol. 25, No. 1, pp. 31–56

STRUCTURED PERTURBATIONS PART II: COMPONENTWISE DISTANCES*

SIEGFRIED M. RUMP†

Abstract. In the second part of this paper we study condition numbers with respect to componentwise perturbations in the input data for linear systems and for matrix inversion, and the distance to the nearest singular matrix. The structures under investigation are linear structures, namely symmetric, persymmetric, skewsymmetric, symmetric Toeplitz, general Toeplitz, circulant, Hankel, and persymmetric Hankel structures. We give various formulas and estimations for the condition numbers. For all structures mentioned except circulant structures we give explicit examples of linear systems $A_\varepsilon x = b$ with parameterized matrix A_ε such that the unstructured componentwise condition number is $\mathcal{O}(\varepsilon^{-1})$ and the structured componentwise condition number is $\mathcal{O}(1)$. This is true for the important case of componentwise relative perturbations in the matrix and in the right-hand side. We also prove corresponding estimations for circulant structures. Moreover, bounds for the condition number of matrix inversion are given. Finally, we give for all structures mentioned above explicit examples of parameterized (structured) matrices A_ε such that the (componentwise) condition number of matrix inversion is $\mathcal{O}(\varepsilon^{-1})$, but the componentwise distance to the nearest singular matrix is $\mathcal{O}(1)$. This is true for componentwise relative perturbations. It shows that, unlike the normwise case, there is no reciprocal proportionality between the componentwise condition number and the distance to the nearest singular matrix.

Key words. componentwise structured perturbations, condition number, distance to singularity

AMS subject classifications. 15A12, 65F35

PII. S0895479802405744

1. Motivation. In the first part of this paper we investigated structured perturbations with respect to normwise distances. There is some drawback to that. For example, many matrices arising from some discretization are sparse. When using a normwise distance, system zeros may be altered by a perturbation into nonzero elements, which usually does not correspond to the underlying model.

System zeros can be modeled in the context of normwise distances such as, for example, symmetric tridiagonal or tridiagonal Toeplitz matrices (see section 7 of Part I of this paper). If components differ much in size, there is the problem that normwise distances alter small components relatively more often than larger components.

To overcome this difficulty a common approach is to use componentwise distances. Consider a linear system $Ax = b$. For some (structured) weight matrix E, structured perturbations $A + \Delta A$ with $|\Delta A| \le \varepsilon|E|$ are considered, where absolute value and comparison are to be understood componentwise. This offers much freedom. For example, for E being the matrix of all 1's, the inequality above is equivalent to $||\Delta A||_m \le \varepsilon$, where $||A||_m := \max |A_{ij}|$, so that there is a finite ratio between this (structured) componentwise condition number and the (structured) normwise condition number (as considered in Part I). That means, in a way, the componentwise approach includes the normwise.

But componentwise perturbations offer much more freedom. For example, for Toeplitz perturbations one need not change the structure when dealing with banded or triangular Toeplitz matrices; in fact, for the common case of componentwise relative

*Received by the editors April 17, 2002; accepted for publication (in revised form) by N. J. Higham January 13, 2003; published electronically May 15, 2003.
http://www.siam.org/journals/simax/25-1/40574.html

†Technical University of Hamburg-Harburg, Schwarzenbergstr. 95, 21071 Hamburg, Germany (rump@tu-harburg.de).

perturbations such a structure is preserved per se. Also, the sensitivity with respect to one, or a couple of, components fits easily into the concept. Therefore, there has been quite some interest in componentwise perturbations in recent years; cf. [13, 14, 2, 10, 20, 22].

However, this much freedom implies drastic consequences for the ratio between the structured and the unstructured condition numbers. It has been mentioned as an advantage in the application of componentwise perturbations that certain components may be excluded from perturbations by setting the corresponding weights to zero. But zero weights *change the structure* of the perturbations and they lower the degree of freedom of perturbations.

One of the most common weights is $E = |A|$, corresponding to componentwise relative perturbations in the matrix A. In this case, zero components of the matrix shrink the space of admissible perturbations. Consider, for example, symmetric Toeplitz perturbations. Then for a specific $n \times n$ symmetric Toeplitz matrix with two nonzero components in the first row, normwise distances allow n degrees of freedom, whereas for componentwise distances *the specific matrix* reduces the degrees of freedom to 2. On the other hand, if this is the given data and if the zeros in the matrix are intrinsic to the model, then there is no more freedom for the perturbation of the input data.

As a consequence there are examples where the structured condition number is near 1, whereas the unstructured condition number can be arbitrarily large. Surprisingly, this is even the case for symmetric linear systems and the case of componentwise relative perturbations of the matrix *and* the right-hand side. This fact does not create much hope that algorithms can be found at all that are stable with respect to componentwise (relative) perturbations. We add more comments about that in the last section.

2. Introduction and notation. Let nonsingular $A \in M_n(\mathbb{R})$ and $x, b \in \mathbb{R}^n$, $x \neq 0$, be given with $Ax = b$. The componentwise condition number of this linear system with respect to a weight matrix $E \in M_n(\mathbb{R})$ and a weight vector $f \in \mathbb{R}^n$ is defined by

$$\begin{aligned}\mathrm{cond}_{E,f}(A,x) := \lim_{\varepsilon\to 0}\sup\Big\{ \frac{\|\Delta x\|_\infty}{\varepsilon\|x\|_\infty} : \ & (A+\Delta A)(x+\Delta x) = b + \Delta b, \Delta A \in M_n(\mathbb{R}), \\ & \Delta b \in \mathbb{R}^n, |\Delta A| \le \varepsilon|E|, \ |\Delta b| \le \varepsilon|f| \Big\}. \end{aligned} \tag{2.1}$$

Note that the weights E, f may have negative entries, but only $|E|, |f|$ is used. Other definitions assume nonnegative weights beforehand. Usually this does not cause problems. In the following, however, we will use skewsymmetric E as well; therefore we choose the definition as in (2.1).

We use the same symbol $\|\cdot\|_\infty$ for the vector maximum norm and the matrix row sum norm. Here and throughout the paper *we use absolute value and comparison of vectors and matrices always componentwise*. For example, $|\Delta A| \le \varepsilon|E|$ is equivalent to $|\Delta A_{ij}| \le \varepsilon|E_{ij}|$ for all i, j. It is well known [16, Theorem 7.4] that

$$\mathrm{cond}_{E,f}(A,x) = \frac{\| \, |A^{-1}| \, |E| \, |x| + |A^{-1}| \, |f| \, \|_\infty}{\|x\|_\infty}. \tag{2.2}$$

This generalizes the Skeel condition number [24]

$$\mathrm{cond}_A(A,x) = \frac{\| \, |A^{-1}| \, |A| \, |x| \, \|_\infty}{\|x\|_\infty}, \tag{2.3}$$

where $E = A$ depicts componentwise relative perturbations in A, and the omitted f indicates that the right-hand side is unchanged. As usual we use $\|\cdot\|_\infty$ in case of componentwise perturbations, whereas the spectral norm $\|\cdot\|_2$ is used in case of normwise perturbations (see Part I of this paper).

For specific right-hand sides the normwise and componentwise condition number can be arbitrarily far apart. For instance, for the well-known example by Kahan [18]

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} \varepsilon \\ -1 \\ 1 \end{pmatrix} \tag{2.4}$$

one computes for normwise and componentwise relative perturbations in the matrix and the right-hand side

$$\kappa_{|A|,|Ax|}(A,x) = 1.4\varepsilon^{-1} \quad \text{but} \quad \mathrm{cond}_{|A|,|Ax|}(A,x) = 2.5.$$

In case of linear systems with special matrices such as Toeplitz or band matrices, algorithms are known that are faster than a general linear system solver. For such a special solver only structured perturbations are possible, for example, Toeplitz or band. Therefore one may ask whether the sensitivity of the solution changes when restricting perturbations to structured perturbations.

Let $M_n^{\mathrm{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ denote a set of matrices of a certain structure. In this paper we will focus on linear structures, namely

$$\mathrm{struct} \in \{\mathrm{sym}, \mathrm{persym}, \mathrm{skewsym}, \mathrm{symToep}, \mathrm{Toep}, \mathrm{circ}, \mathrm{Hankel}, \mathrm{persymHankel}\}, \tag{2.5}$$

that is, symmetric, persymmetric, skewsymmetric, symmetric Toeplitz, general Toeplitz, circulant, Hankel, and persymmetric Hankel matrices. We define, similarly to (2.1), the *structured componentwise condition number* by

$$\begin{aligned}\mathrm{cond}_{E,f}^{\mathrm{struct}}(A,x) := \lim_{\varepsilon\to 0}\sup\Big\{\frac{\|\Delta x\|_\infty}{\varepsilon\|x\|_\infty} :\ & (A+\Delta A)(x+\Delta x) = b+\Delta b, \Delta A \in M_n^{\mathrm{struct}}(\mathbb{R}), \\ & \Delta b \in \mathbb{R}^n, |\Delta A| \le \varepsilon|E|, |\Delta b| \le \varepsilon|f|\Big\}.\end{aligned} \tag{2.6}$$

We mention that for $A \in M_n^{\mathrm{struct}}(\mathbb{R})$ and all structures under investigation $\Delta A \in M_n^{\mathrm{struct}}(\mathbb{R})$ is equivalent to $A + \Delta A \in M_n^{\mathrm{struct}}(\mathbb{R})$. Therefore it suffices to assume $\Delta A \in M_n^{\mathrm{struct}}(\mathbb{R})$ in (2.6).

For a specialized solver, for example, for symmetric Toeplitz A, only symmetric Toeplitz perturbations of A are possible because only the first row of A is input to the algorithm. A considerable factor between the structured condition number (2.6) and the general, unstructured condition number (2.1) may shed light on the stability of an algorithm.

Indeed, there may be huge factors between (2.1) and (2.6). Let, for example,

$$A = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & \ddots \\ & & \ddots & \ddots \end{pmatrix} \in M_n(\mathbb{R}) \quad \text{and} \quad x = (1,-1,1,-1,\ldots)^T \in \mathbb{R}^n.$$

Then for $n = 200$ we have

$$\mathrm{cond}_A(A,x) = 2.02\cdot 10^4, \quad \mathrm{cond}_A^{\mathrm{sym}}(A,x) = 1.02\cdot 10^4, \quad \text{and} \quad \mathrm{cond}_A^{\mathrm{symToep}}(A,x) = 1.$$

In this example no perturbations in the right-hand side are allowed. This does not conform with Wilkinson's classical ansatz of error analysis, where he relates the computational error to a perturbation in the input data. It may seem artificial to allow variations of some input data, namely the input matrix A, and not of others such as the right-hand side b. However, this depends on the point of view. The right-hand side may be given exactly where perturbations do not make sense, for example, when solving $Ax = e_1$ for computing the first column of A^{-1}, where e_1 denotes the first column of the identity matrix. For this problem, and also in case the problem is solved through $Ax = e_1$, the input of the problem is only A.

A numerical algorithm solves a nearby problem, and for the judgment of stability of an algorithm disregarding perturbations in some input data seems inadequate. However, in case of banded A, for example, the zeros outside the band are *not input* to a band solver, and therefore perturbations in those should *not* be taken into account. Among others, these are motivations for looking at componentwise perturbations.

However, we will see that things change significantly in the rugged world of componentwise perturbations compared to the smooth world of normwise perturbations. We are especially interested in the estimation of $\mathrm{cond}^{\mathrm{struct}}/\mathrm{cond}$, a question also posed in [13]. Even for handsome structures such as symmetric matrices there are examples where $\mathrm{cond}_{A,b}(A,x)$ is arbitrarily large compared to $\mathrm{cond}_{A,b}^{\mathrm{sym}}(A,x) \sim 1$. Note that this is true for the important case of componentwise relative perturbations in the matrix *and* in the right-hand side.

We will give similar examples for all structures in (2.5) except circulant matrices. In the latter case we give almost sharp estimations for $\mathrm{cond}^{\mathrm{struct}}/\mathrm{cond}$. As we will see, for circulant structures the ratio may only become small for ill-conditioned matrices. The worst case is about $\mathrm{cond}^{\mathrm{struct}} \sim \sqrt{\mathrm{cond}}$.

The examples mentioned above are valid for a specific solution x and corresponding right-hand side. The worst case structured condition number for componentwise relative perturbations, the supremum over all x, however, will be shown to be not far away from the corresponding unstructured condition number for all structures in (2.5).

Moreover, bounds for the condition number of matrix inversion are given. Finally, we give for all structures in (2.5) explicit examples of parameterized (structured) matrices A_ε such that the condition number of matrix inversion is $\mathcal{O}(\varepsilon^{-1})$, but the componentwise distance to the nearest singular matrix is $\mathcal{O}(1)$. This is again true for the important case of componentwise relative perturbations. It shows that, unlike in the normwise case, there is no reciprocal proportionality between the componentwise condition number and distance to the nearest singular matrix. Recall that for normwise perturbations the structured condition number is equal to the reciprocal of the (structured) distance to the nearest singular matrix (Part I, Theorem 12.1).

We will use the following notation:

$M_n(\mathbb{R})$	set of real $n \times n$ matrices
$M_n^{\mathrm{struct}}(\mathbb{R})$	set of structured real $n \times n$ matrices
$\lVert \cdot \rVert_\infty$	infinity or row sum norm
E	some (weight) matrix, $E \in M_n(\mathbb{R})$
f	some (weight) vector, $f \in \mathbb{R}^n$
I, I_n	identity matrix (with n rows and columns)
e	vector of all 1's, $e \in \mathbb{R}^n$

$(\mathbf{1})$	matrix of all 1's, $(\mathbf{1}) = ee^T \in M_n(\mathbb{R})$
S	signature matrix, i.e., $\lvert S\rvert = I$ or $S = \operatorname{diag}(\pm 1, \dots, \pm 1)$
J, J_n	permutation matrix mapping $(1, \dots, n)^T$ into $(n, \dots, 1)^T$
$\sigma_{\min}(A)$	smallest singular value of A
$\lambda_{\min}(A)$	smallest eigenvalue of symmetric A

In this paper we treat explicitly the important (linear) structures in (2.5). However, we also derive formulas for general linear structures similar to those derived in [13]. We mention that this includes structures in the right-hand side by treating an augmented linear system of dimension $n+1$. Such structures appear, for example, in the Yule–Walker problem [11, section 4.7.2].

3. Componentwise perturbations. Throughout this paper let nonsingular $A \in M_n(\mathbb{R})$ be given together with $0 \neq x \in \mathbb{R}^n$ and weights $E \in M_n(\mathbb{R})$, $f \in \mathbb{R}^n$. Denote $b := Ax$.

The standard proof [16, Theorem 7.4] of (2.2) uses that $(A+\Delta A)(x+\Delta x) = b+\Delta b$ and $Ax = b$ imply

$$\Delta x = A^{-1}(-\Delta A x + \Delta b) + \mathcal{O}(\varepsilon^2). \tag{3.1}$$

This is true independent of ΔA, structured or not. It follows that

$$\operatorname{cond}^{\text{struct}}_{E,f}(A,x) = \sup\left\{ \frac{\|A^{-1}\Delta A x + A^{-1}\Delta b\|_\infty}{\|x\|_\infty} :\ \Delta A \in M_n^{\text{struct}}(\mathbb{R}), \Delta b \in \mathbb{R}^n, \right.$$
$$\left. |\Delta A| \le |E|, |\Delta b| \le |f| \right\}. \tag{3.2}$$

This is again true for all structures including the unstructured case $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$. For the estimation of $\|\Delta x\|_\infty$, in case of structured perturbations of ΔA we use the ansatz as in [13] (see also Part I of this paper). All structures in (2.5) are linear structures. That means for given "struct" every matrix ΔA in $M_n^{\text{struct}}(\mathbb{R})$ depends linearly on some k parameters $\Delta p \in \mathbb{R}^k$. The number of parameters k depends on the structure; see Table 6.1 in Part I of this paper. Denote the vector of stacked columns of ΔA by $\operatorname{vec}(\Delta A) \in \mathbb{R}^{n^2}$. Then there is a bijective correspondence

$$\operatorname{vec}(\Delta A) = \Phi^{\text{struct}} \cdot \Delta p \tag{3.3}$$

between $\operatorname{vec}(\Delta A)$ and the parameters Δp by some matrix $\Phi^{\text{struct}} \in M_{n^2,k}(\mathbb{R})$. Note that Φ^{struct} is fixed for every structure and given size $n \in \mathbb{N}$. Also note that Φ^{struct} contains for all structures in (2.5) exactly one nonzero entry in each row.

In case of structured componentwise perturbations it seems natural to assume $E \in M_n^{\text{struct}}(\mathbb{R})$. This implies existence of $p_E \in \mathbb{R}^k$ with $\operatorname{vec}(E) = \Phi^{\text{struct}} \cdot p_E$, and because Φ^{struct} contains exactly one nonzero entry per row we have the nice equivalence

$$\Delta A \in M_n^{\text{struct}}(\mathbb{R}) \quad \text{and} \quad |\Delta A| \le |E| \Leftrightarrow \operatorname{vec}(\Delta A) = \Phi^{\text{struct}} \cdot \Delta p \quad \text{and} \quad |\Delta p| \le |p_E|. \tag{3.4}$$

That means the set of all $|\Delta p| \le |p_E|$ maps one-to-one to the set of ΔA allowed in (3.2). In that respect structured componentwise perturbations are easier to handle than structured normwise perturbations. Finally, observe $\Delta A x = (x^T \otimes I)\Phi^{\text{struct}}\Delta p$ for $\otimes$ denoting the Kronecker product, and with the abbreviation

$$\Psi_x^{\text{struct}} := (x^T \otimes I)\Phi^{\text{struct}} \tag{3.5}$$

we obtain the following formula for $\mathrm{cond}^{\mathrm{struct}}_{E,f}(A,x)$, which was also observed in [13].

THEOREM 3.1. *For nonsingular* $A \in M_n(\mathbb{R})$, $0 \neq x \in \mathbb{R}^n$, $E \in M_n^{\mathrm{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$, *and* $f \in \mathbb{R}^n$ *such that* $\mathrm{vec}(E) = \Phi^{\mathrm{struct}} p_E$ *for* $p_E \in \mathbb{R}^k$ *we have*

$$\mathrm{cond}^{\mathrm{struct}}_{E,f}(A,x) = \frac{\|\, |A^{-1}\Psi_x|\,|p_E| + |A^{-1}|\,|f|\,\|_\infty}{\|x\|_\infty}. \tag{3.6}$$

We note that Theorem 3.1 contains (2.2) for unstructured perturbations. In that case it is just $\Phi = I_{n^2}$ and $\Psi_x = x^T \otimes I_n$. Then $\mathrm{vec}(E) = p_E$ implies $|A^{-1}(x^T \otimes I)|\,|p_E| = |x^T \otimes A^{-1}|\,|p_E| = (|x^T| \otimes |A^{-1}|)|p_E| = |A^{-1}|\,|E|\,|x|$ using [17, Lemmas 4.2.10 and 4.3.1].

In Part I of this paper, we concluded in Corollary 6.6 of the corresponding Theorem 6.5 that a lower bound on the ratio of the structured and unstructured normwise condition number only depends on the solution vector x and not on the matrix A. This is not possible in the componentwise case. In the normwise case the factor $\|E\|$ cancelled, whereas in the componentwise case p_E may consist of components large in absolute value corresponding to columns of $|A^{-1}\Psi_x|$ being small in absolute value. This does indeed happen, as we will see in the explicit examples in the next sections.

For the structures in (2.5) the matrix Ψ_x is large but sparse. A mere count of operations shows that $A^{-1}\Psi_x$ requires not more than n^3 multiplications and additions for all structures in (2.5). Moreover, frequently it is not the exact value but rather an approximation of (3.6) that is sufficient. For that purpose efficient methods requiring some $\mathcal{O}(n^2)$ flops are available; see, for example, [12, 15].

To simplify and focus the discussion we observe that

$$\begin{aligned}
&A \in M_n^{\mathrm{sym}}(\mathbb{R}) \Leftrightarrow JA \in M_n^{\mathrm{persym}}(\mathbb{R}) \Leftrightarrow AJ \in M_n^{\mathrm{persym}}(\mathbb{R}),\\
&A \in M_n^{\mathrm{symToep}}(\mathbb{R}) \Leftrightarrow JA \in M_n^{\mathrm{persymHankel}}(\mathbb{R}) \Leftrightarrow AJ \in M_n^{\mathrm{persymHankel}}(\mathbb{R}),\\
&A \in M_n^{\mathrm{Toep}}(\mathbb{R}) \Leftrightarrow JA \in M_n^{\mathrm{Hankel}}(\mathbb{R}) \Leftrightarrow AJ \in M_n^{\mathrm{Hankel}}(\mathbb{R}).
\end{aligned} \tag{3.7}$$

By rewriting (3.1) into

$$\Delta x = (JA)^{-1}(-J\Delta A x + J\Delta b) + \mathcal{O}(\varepsilon^2) \quad \text{and} \quad J\Delta x = (AJ)^{-1}(-\Delta AJ\cdot Jx + \Delta b) + \mathcal{O}(\varepsilon^2)$$

and observing $|\Delta A| \le |E| \;\Leftrightarrow\; |J\Delta A| \le |JE| \;\Leftrightarrow\; |\Delta AJ| \le |EJ|$ and $|\Delta b| \le |f| \;\Leftrightarrow\; |J\Delta b| \le |Jf|$ we obtain the following.

THEOREM 3.2. *For nonsingular* $A \in M_n(\mathbb{R})$ *and* $0 \neq x \in \mathbb{R}^n$ *there holds*

$$\begin{aligned}
\mathrm{cond}^{\mathrm{sym}}_{E,f}(A,x) &= \mathrm{cond}^{\mathrm{persym}}_{JE,Jf}(JA,x) = \mathrm{cond}^{\mathrm{persym}}_{EJ,f}(AJ,Jx),\\
\mathrm{cond}^{\mathrm{symToep}}_{E,f}(A,x) &= \mathrm{cond}^{\mathrm{persymHankel}}_{JE,Jf}(JA,x) = \mathrm{cond}^{\mathrm{persymHankel}}_{EJ,f}(AJ,Jx),\\
\mathrm{cond}^{\mathrm{Toep}}_{E,f}(A,x) &= \mathrm{cond}^{\mathrm{Hankel}}_{JE,Jf}(JA,x) = \mathrm{cond}^{\mathrm{Hankel}}_{EJ,f}(AJ,Jx).
\end{aligned}$$

Therefore we will focus our discussion on symmetric, symmetric Toeplitz, and Hankel matrices, and the results will, mutatis mutandis, be valid for persymmetric, persymmetric Hankel, and general Toeplitz matrices, respectively.

4. Condition number for general x.

For the case of unstructured componentwise relative perturbations $E = A$ and $f = b$ it does not make much difference whether perturbations in the right-hand side are allowed or not. Indeed, (2.2) and $|b| = |Ax| \le |A|\,|x|$ imply

$$\mathrm{cond}_A(A, x) \le \mathrm{cond}_{A,b}(A, x) \le 2\mathrm{cond}_A(A, x). \tag{4.1}$$

A similar estimation is valid for the condition number under normwise perturbations (see Part I, equation (4.1)). Since the normwise condition number $\kappa_E(A, x) = \|A^{-1}\|_2\|E\|_2$ without perturbations in the right-hand side does not depend on x, condition is an inherent property of the matrix—at least for unstructured *normwise* perturbations. This is no longer the case for *componentwise* perturbations. Consider

$$A := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1+\varepsilon \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{4.2}$$

Then (2.2) implies

$$\begin{array}{rlll} \mathrm{cond}_A(A, x) = 4\varepsilon^{-1} + \mathcal{O}(1) & \text{but} & \mathrm{cond}_A(A, y) = 1, & \text{and} \\ \mathrm{cond}_{A,b}(A, x) = 8\varepsilon^{-1} + \mathcal{O}(1) & \text{but} & \mathrm{cond}_{A,b}(A, y) = 2. & \end{array} \tag{4.3}$$

So condition subject to componentwise perturbations is no longer an intrinsic matrix property but depends on the solution x (and therefore on the right-hand side). Note that the norms of rows and columns of A in (4.2) are of similar size.

There are similar examples for structured perturbations. For instance, the same data (4.2) yield for symmetric perturbations

$$\mathrm{cond}^{\mathrm{sym}}_{A,b}(A, x) = 6\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}^{\mathrm{sym}}_{A,b}(A, y) = 2,$$

and there are similar examples for the other structures in (2.5).

We may ask what is the worst case condition number for all x. We first observe the following.

LEMMA 4.1. *For nonsingular $A \in M_n(\mathbb{R})$ and $M^{\mathrm{struct}}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$ we have*

$$\sup_{x\neq 0} \mathrm{cond}^{\mathrm{struct}}_{E,f}(A, x) = \sup_{|x|=e} \mathrm{cond}^{\mathrm{struct}}_{E,f}(A, x). \tag{4.4}$$

Proof. In view of (3.2) the supremum over all $0 \ne x \in \mathbb{R}^n$ in (4.4) can obviously be replaced by the supremum over all $\|x\|_\infty = 1$, the same as $|x| \le e$ with at least one $|x_i| = 1$. The assertion follows easily. ∎

For unstructured perturbations, formula (2.2) and Lemma 4.1 imply

$$\sup_{x\neq 0} \mathrm{cond}_{E,f}(A, x) = \mathrm{cond}_{E,f}(A, e) = \|\,|A^{-1}|\,|E|e + |A^{-1}|\,|f|\,\|_\infty, \tag{4.5}$$

and for no perturbations in the right-hand side

$$\sup_{x\neq 0} \mathrm{cond}_E(A, x) = \mathrm{cond}_E(A, e) = \|\,|A^{-1}|\,|E|\,\|_\infty.$$

For some structures the supremum for structured perturbations (4.4) is equal to the worst case (4.5) for unstructured perturbations.

THEOREM 4.2. *Let $M^{\mathrm{struct}}_n(\mathbb{R}) \subseteq M_n(\mathbb{R})$ and nonsingular $A \in M^{\mathrm{struct}}_n(\mathbb{R})$, $E \in M^{\mathrm{struct}}_n(\mathbb{R})$, and $f \in \mathbb{R}^n$ be given. If, for every signature matrix S, $B \in M^{\mathrm{struct}}_n(\mathbb{R})$ implies $SBS \in M^{\mathrm{struct}}_n(\mathbb{R})$, then*

$$\sup_{x\neq 0} \mathrm{cond}^{\mathrm{struct}}_{E,f}(A, x) = \sup_{x\neq 0} \mathrm{cond}_{E,f}(A, x) = \|\,|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|\,\|_\infty. \tag{4.6}$$

Proof. Let i denote the row of $|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|$ for which the maximum is achieved in the ∞-norm, and denote by S the signature matrix with $S_{\nu\nu} := sign(A^{-1})_{i\nu}$. Then

$$(A^{-1} \cdot S|E|S \cdot Se)_i = (|A^{-1}|\,|E|\,e)_i,$$

and the result follows by (3.2), choosing $\Delta A := S|E|S \in M_n^{\text{struct}}(\mathbb{R})$ with $|\Delta A| = |E|$ and the obvious choice of Δb. ∎

COROLLARY 4.3. *For* struct $\in$ {sym, persym, skewsym} *and nonsingular* $A \in M_n^{\text{struct}}(\mathbb{R})$, $E \in M_n^{\text{struct}}(\mathbb{R})$, $f \in \mathbb{R}^n$, *it follows that*

$$\begin{aligned}\sup_{x\neq 0} \text{cond}_{E,f}^{\text{struct}}(A,x) &= \sup_{x\neq 0} \text{cond}_{E,f}(A,x) = \text{cond}_{E,f}(A,e)\\ &= \|\,|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|\,\|_\infty\,,\end{aligned}$$

and therefore

$$\sup_{x\neq 0} \text{cond}_{E}^{\text{struct}}(A,x) = \|\,|A^{-1}|\,|E|\,\|_\infty$$

for no perturbations in the right-hand side.

Proof. For struct $\in$ {sym, skewsym} the result follows by Theorem 4.2, and for persymmetric matrices by Theorem 3.2. ∎

For other structures things change if the structure imposes too many restrictions on the choice of the elements. If not, the following theorem gives at least two-sided bounds for the worst case condition number for all x.

THEOREM 4.4. *Let* $M_n^{\text{struct}}(\mathbb{R}) \subseteq M_n(\mathbb{R})$ *be given such that for every individual column there is no dependency between the elements; in other words, for every* $c \in \mathbb{R}^n$ *and every index* $i \in \{1,\dots,n\}$ *there exists* $B \in M_n^{\text{struct}}(\mathbb{R})$ *with the ith column* B_i *of* B *equal to* c. *For such* $M_n^{\text{struct}}(\mathbb{R})$ *and given nonsingular* $A \in M_n^{\text{struct}}(\mathbb{R})$, $E \in M_n^{\text{struct}}(\mathbb{R})$, $f \in \mathbb{R}^n$, *it follows that*

$$n^{-1}\alpha \le \sup_{x\neq 0} \text{cond}_{E,f}^{\text{struct}}(A,x) \le \alpha, \tag{4.7}$$

where

$$\alpha := \text{cond}_{E,f}(A,e) = \sup_{x\neq 0} \text{cond}_{E,f}(A,x) = \|\,|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|\,\|_\infty.$$

Estimation (4.7) *is especially true for* struct $\in$ {circ, Toep, Hankel}.

Proof. Denote by $i \in \{1,\dots,n\}$ an index with

$$\alpha = \sup_{\|x\|_\infty = 1} \text{cond}_{E,f}(A,x) = \|\,|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|\,\|_\infty = (|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|\,)_i.$$

There is $j \in \{1,\dots,n\}$ with $(|A^{-1}|\,|E|\,e)_i \le n(|A^{-1}|\,|E|)_{ij}$. Choose $\Delta A \in M_n^{\text{struct}}(\mathbb{R})$ with $|\Delta A| \le |E|$ such that $\Delta A_{\nu i} = sign((A^{-1})_{i\nu}) \cdot |E_{\nu j}|$. Then $(A^{-1}\Delta A)_{ij} = (|A^{-1}|\,|E|)_{ij}$, and for suitable x with $|x| = e$ and for suitable Δb we obtain

$$\begin{aligned}|(A^{-1}\Delta A x + A^{-1}\Delta b)_i| &\ge (|A^{-1}|\,|E|)_{ij} + (|A^{-1}|\,|f|)_i\\ &\ge n^{-1}(|A^{-1}|\,|E|\,e + |A^{-1}|\,|f|)_i = n^{-1}\alpha.\end{aligned}$$

Now the left inequality in (4.7) follows by (3.2), and the right inequality is obvious. ∎

The assumptions on $M_n^{\text{struct}}(\mathbb{R})$ are also satisfied for struct $\in$ {sym, persym, skewsym}, where we already obtained the sharp result in Corollary 4.3. However, the assumptions are not satisfied for symmetric Toeplitz structures of dimension $n \ge 3$, and therefore also are not for persymmetric Hankel structures. Indeed, we will give general examples of symmetric Toeplitz matrices of dimension $n = 3$ and $n \ge 5$ such that

$$\text{cond}_A(A, e) = \varepsilon^{-1} + \mathcal{O}(1) \quad \text{but} \quad \sup_{x \neq 0} \text{cond}_A^{\text{symToep}}(A, x) = 1 + \mathcal{O}(\varepsilon). \tag{4.8}$$

Note that we use the weight matrix $E = A$ but do not allow perturbations in the right-hand side. Of course, after introducing some small weight f for a perturbation in the right-hand side formula, (4.8) is still valid in weaker form.

Let us explore this example for $n = 3$ in more detail. Consider (for small $\alpha \in \mathbb{R}$)

$$A = \begin{pmatrix} 0 & 1 & \alpha \\ 1 & 0 & 1 \\ \alpha & 1 & 0 \end{pmatrix} \quad \text{such that} \quad A^{-1} = (2\alpha)^{-1} \begin{pmatrix} -1 & \alpha & 1 \\ \alpha & -\alpha^2 & \alpha \\ 1 & \alpha & -1 \end{pmatrix}. \tag{4.9}$$

General $\Delta A \in M_3^{\text{symToep}}(\mathbb{R})$ with $|\Delta A| \le |A|$ is of the form

$$\Delta A = \begin{pmatrix} 0 & a & \alpha b \\ a & 0 & a \\ \alpha b & a & 0 \end{pmatrix} \quad \text{with} \quad |a| \le 1, |b| \le 1.$$

Then

$$A^{-1}\Delta A = \frac{1}{2} \begin{pmatrix} a+b & 0 & a-b \\ \alpha(b-a) & 2a & \alpha(b-a) \\ a-b & 0 & a+b \end{pmatrix} \quad \text{such that} \quad \|A^{-1}\Delta A\|_\infty \le 1 + \mathcal{O}(\alpha).$$

But

$$|A^{-1}|\,|A| = \begin{pmatrix} 1 & \alpha^{-1} & 1 \\ \alpha & 1 & \alpha \\ 1 & \alpha^{-1} & 1 \end{pmatrix} \quad \text{implies} \quad \|\,|A^{-1}|\,|A|\,\|_\infty = \alpha^{-1} + \mathcal{O}(1)$$

such that (3.2) and (2.2) imply

$$\sup_{x \neq 0} \text{cond}_A^{\text{symToep}}(A, x) \le 1 + \mathcal{O}(\alpha) \quad \text{but} \quad \text{cond}_A(A, e) = \alpha^{-1} + \mathcal{O}(1). \tag{4.10}$$

Note that $\text{cond}_A(A, x) \sim \alpha^{-1}$ is true for all $x \in \mathbb{R}^n$ with $|x_2|$ not too small.

The situation as in (4.10) cannot happen if perturbations in the right-hand side are allowed, at least not for the important case of relative perturbations $E = A, f = b$. In that case the worst case condition number is of the order of $\|\,|A^{-1}|\,|A|\,\|_\infty$, as shown by the following theorem.

THEOREM 4.5. *Let arbitrary $M_n^{\text{struct}}(\mathbb{R})$ be given and nonsingular $A \in M_n^{\text{struct}}(\mathbb{R})$. Then for componentwise relative perturbations in the matrix and in the right-hand side, i.e., for $E = A$ and $f = Ax$, we have*

$$n^{-1}\|\,|A^{-1}|\,|A|\,\|_\infty \le \sup_{x \neq 0} \text{cond}_{A,Ax}^{\text{struct}}(A, x) \le 2\|\,|A^{-1}|\,|A|\,\|_\infty.$$

Remark 4.6. Note that the weight $f = Ax$ for the right-hand side depends on x. This problem does not occur in the normwise case because in that case the worst case structured condition number (for all x) is equal to the unstructured condition number for *all* weights E, f (see Part I, Theorem 4.1).

Proof. On the one hand, (3.2) implies

$$\sup_{x\neq 0} \mathrm{cond}^{\mathrm{struct}}_{A,Ax}(A,x) \geq \sup_{|x|=e} \| \, |A^{-1}| \, |Ax| \, \|_\infty \geq n^{-1} \| \, |A^{-1}| \, |A| \, \|_\infty .$$

On the other hand, (3.2) and (2.2) yield

$$\begin{aligned} \sup_{x\neq 0} \mathrm{cond}^{\mathrm{struct}}_{A,Ax}(A,x) &\leq \sup_{x\neq 0} \mathrm{cond}_{A,Ax}(A,x) = \mathrm{cond}_{A,Ae}(A,e) \\ &= \| \, |A^{-1}| \, |A|e + |A^{-1}| \, |Ae| \, \|_\infty \leq 2 \| \, |A^{-1}| \, |A| \, \|_\infty . \qquad \square \end{aligned}$$

We mention that

$$\| \, |A^{-1}| \, |A| \, \|_\infty = \inf_{D_1, D_2} \kappa_\infty(D_1 A D_2) = \varrho(|A^{-1}| \, |A|), \tag{4.11}$$

where ϱ denotes the spectral radius and the infimum is taken over nonsingular diagonal D_i. So this quantity is the infimum ∞-norm condition number with respect to unstructured and normwise perturbations in the matrix. The right equality in (4.11) was proved by Bauer [3] for the case where $|A|$ and $|A^{-1}|$ have positive entries. The proof gives D_1 and D_2 explicitly by using the right Perron vector of $|A^{-1}| \, |A|$. The argument is also valid for general A, as shown by [23].

The question remains of whether at least some of the previous results for the worst case structured condition number (for all x) can be shown for specific x, i.e., specific right-hand side. A worst case scenario in that respect would be if for the natural weights $E = A$ and $f = b$, i.e., componentwise relative perturbations in the matrix and the right-hand side, there exist A, b, and x with $\mathrm{cond}_{A,b}(A,x)$ arbitrarily large, whereas $\mathrm{cond}^{\mathrm{struct}}_{A,b}(A,x) = \mathcal{O}(1)$. Indeed, we will show that for all structures mentioned in (2.5) except circulants there are such general examples.

5. Symmetric, persymmetric, and skewsymmetric matrices.

For the case of no perturbations in the right-hand side it is fairly easy to find parameterized $A = A_\varepsilon$ and x such that $\mathrm{cond}_A(A,x) = \mathcal{O}(\varepsilon^{-1})$ and $\mathrm{cond}^{\mathrm{struct}}_A(A,x) = \mathcal{O}(1)$ for $\mathrm{struct} \in \{\mathrm{sym}, \mathrm{persym}, \mathrm{skewsym}\}$. We found it more difficult to find such examples with perturbations in the right-hand side; in fact, we did not expect there to be any; however, they do exist. We illustrate the first example in more detail. Consider

$$A = A_\varepsilon = \begin{pmatrix} 0 & 1 & 1 & 0 & -1 \\ 1 & 0 & 1 & -1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & \varepsilon & 1 \\ -1 & 0 & 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Then

$$A^{-1} = \varepsilon^{-1} \begin{pmatrix} 1 & -1+\frac{\varepsilon}{2} & 0 & 1 & -1-\frac{\varepsilon}{2} \\ -1+\frac{\varepsilon}{2} & 1-\frac{\varepsilon}{2} & \frac{\varepsilon}{2} & -1 & 1 \\ 0 & \frac{\varepsilon}{2} & 0 & 0 & \frac{\varepsilon}{2} \\ 1 & -1 & 0 & 1 & -1 \\ -1-\frac{\varepsilon}{2} & 1 & \frac{\varepsilon}{2} & -1 & 1+\frac{\varepsilon}{2} \end{pmatrix}.$$

Furthermore,

$$|A^{-1}|\,|A|\,|x| = \varepsilon^{-1}\begin{pmatrix} 8+3\varepsilon \\ 8+5\varepsilon \\ 2\varepsilon \\ 8+\varepsilon \\ 8+5\varepsilon \end{pmatrix} \quad \text{and} \quad |A^{-1}|\,|Ax| = |A^{-1}|\begin{pmatrix} 0 \\ 0 \\ 4 \\ \varepsilon \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \\ 3 \end{pmatrix}. \tag{5.1}$$

Note that $|A^{-1}|\,|Ax|$ is of size $\mathcal{O}(1)$ because the third column of A^{-1}, which meets the component 4 in $|Ax|$, is of size $\mathcal{O}(\varepsilon)$. This is important because this term $|A^{-1}|\,|Ax|$ occurs in both the unstructured condition number (2.2) *and* the structured condition number (cf. Theorem 3.1). Now (2.2) implies

$$\mathrm{cond}_{A,Ax}(A,x) = 8\varepsilon^{-1} + \mathcal{O}(1). \tag{5.2}$$

On the other hand, according to (3.5),

$$\Psi_x^{\mathrm{sym}} = \begin{pmatrix} 1&1&0&1&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&1&0&1&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&1&0&0&0&1&1&0&0&0 \\ 0&0&0&1&0&0&0&1&0&0&0&0&1&1&0 \\ 0&0&0&0&1&0&0&0&1&0&0&0&0&1&1 \end{pmatrix},$$

such that $A^{-1}\Psi_x^{\mathrm{sym}}$ has large elements of size $\mathcal{O}(\varepsilon^{-1})$ in columns 1, 4, 6, 9, 13, and 15, whereas all other columns are comprised of elements of magnitude $\mathcal{O}(1)$. However, the parameter vector p_A such that $\mathrm{vec}(A) = \Phi^{\mathrm{sym}} \cdot p_A$ has zero elements in components 1, 4, 6, 9, 10, 15, a value ε in component 13, and ± 1's otherwise. Therefore

$$|A^{-1}\Psi_x^{\mathrm{sym}}| \cdot |p_A| = \begin{pmatrix} 3 \\ 4 \\ 2 \\ 1 \\ 4 \end{pmatrix} + \mathcal{O}(\varepsilon)$$

such that (5.1) and Theorem 3.1 imply

$$\mathrm{cond}_{A,Ax}^{\mathrm{sym}}(A,x) = 7 + \mathcal{O}(\varepsilon). \tag{5.3}$$

The numbers in (5.2) and (5.3) do not change when replacing A by $A \oplus B$ and prolonging x by k zeros, where $B \in M_k(\mathbb{R})$ denotes any symmetric matrix. Furthermore, Theorem 3.2 implies that the same example applies for persymmetric structures. We proved the following result.

THEOREM 5.1. *For $n \geq 5$, there exist parameterized symmetric $A := A_\varepsilon \in M_n^{\mathrm{sym}}(\mathbb{R})$ and $x \in \mathbb{R}^n$ such that*

$$\mathrm{cond}_{A,Ax}(A,x) = 8\varepsilon^{-1} + \mathcal{O}(1) \quad \textit{and} \quad \mathrm{cond}_{A,Ax}^{\mathrm{sym}}(A,x) = 7 + \mathcal{O}(\varepsilon).$$

For the persymmetric matrix $JA \in M_n^{\mathrm{persym}}(\mathbb{R})$ similar assertions are true.

For the skewsymmetric matrix $A = A_\varepsilon \in M_n^{\mathrm{skewsym}}(\mathbb{R})$ with

$$A := \begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & \varepsilon & 1 \\ 0 & -\varepsilon & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

one computes using (2.2) and Theorem 3.1

$$\mathrm{cond}_{A,Ax}(A,x) = 6\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}_{A,Ax}^{\mathrm{skewsym}}(A,x) = 4 + \mathcal{O}(\varepsilon). \tag{5.4}$$

Skewsymmetric nonsingular matrices are of even dimension. So replacing A by $A \oplus B$ and prolonging x by $2k$ zeros, where $B \in M_k^{\mathrm{skewsym}}(\mathbb{R})$ denotes any skewsymmetric matrix, does not change the numbers in (5.4). We have the following result.

THEOREM 5.2. *For even $n \geq 4$, there exist parameterized skewsymmetric $A := A_\varepsilon \in M_n^{\mathrm{skewsym}}(\mathbb{R})$ and $x \in \mathbb{R}^n$ such that*

$$\mathrm{cond}_{A,Ax}(A,x) = 6\varepsilon^{-1} + \mathcal{O}(1) \quad \textit{and} \quad \mathrm{cond}_{A,Ax}^{\mathrm{skewsym}}(A,x) = 4 + \mathcal{O}(\varepsilon).$$

This result shows a major difference between normwise and componentwise perturbations. In Part I, Theorem 5.3 we proved that for struct $\in \{\mathrm{sym}, \mathrm{persym}, \mathrm{skewsym}\}$ and for *all* x the structured normwise condition number is *equal* to the unstructured normwise condition number. For componentwise perturbations and *specific* x the condition numbers can be arbitrarily far apart, although Corollary 4.3 shows that in the worst case they are identical. We had similar results for normwise perturbations for the other structures in (2.5) (Part I, Theorems 8.4, 9.2, and 10.2). However, the worst case was essentially $\kappa^{\mathrm{struct}} \sim \kappa^{1/2}$; i.e., a big ratio $\kappa/\kappa^{\mathrm{struct}}$ was only possible for ill-conditioned matrices. For componentwise perturbations, $\mathrm{cond}_{A,Ax}^{\mathrm{struct}} \sim 1$ is possible compared to arbitrarily large $\mathrm{cond}_{A,Ax}$—always for the important case of componentwise relative perturbations in the matrix and the right-hand side.

6. Toeplitz and Hankel matrices. Symmetric Toeplitz matrices depend only on n parameters. That makes it less difficult to find examples in the spirit of the previous section. Consider

$$A = A_\varepsilon := \mathrm{Toeplitz}(0, 0, 1+\varepsilon, -1, 1) \quad \text{and} \quad x = (1, 1, 0, 1, 1)^T$$

such that A is the symmetric Toeplitz matrix with first row $[0, 0, 1+\varepsilon, -1, 1]$. Then

$$\mathrm{cond}_{A,Ax}(A,x) = 4\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}_{A,Ax}^{\mathrm{symToep}}(A,x) = 5 + \mathcal{O}(\varepsilon).$$

In case of Toeplitz structures it is a little more subtle to find general $n \times n$ examples. The structure does not permit us to use a simple direct sum as in the previous section. We found the following examples. For even order greater than or equal to 6 consider

$$\begin{aligned} &A = \mathrm{Toeplitz}(0, 0, z, -1, 1, z, 1, \varepsilon) \in M_{6+2k}^{\mathrm{symToep}}(\mathbb{R}) \quad \text{and} \\ &x = (1, 0, z, -1, -1, z, 0, 1) \in \mathbb{R}^{6+2k}, \end{aligned}$$

where $z \in \mathbb{R}^k$ denotes a vector of $k \geq 0$ zeros. Then (2.2) and Theorem 3.1 yield

$$\mathrm{cond}_{A,Ax}(A,x) = 4\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}_{A,Ax}^{\mathrm{symToep}}(A,x) = 6 + \mathcal{O}(\varepsilon)$$

for all $k \geq 0$. For odd order greater than or equal to 7 consider

$$\begin{aligned} A &= \mathrm{Toeplitz}(\varepsilon, -\varepsilon, 0, 0, 0, 0, z, 1, z) \in M_{7+2k}^{\mathrm{symToep}}(\mathbb{R}) \quad \text{and} \\ x &= (z, 0, 1, 1, 0, 1, 1, 0, z) \in \mathbb{R}^{7+2k}, \end{aligned}$$

where $z \in \mathbb{R}^k$ denotes again a vector of k zeros. Then

$$\begin{array}{rcllll} \mathrm{cond}_{A,Ax}(A,x) & = & 4\varepsilon^{-1} + \mathcal{O}(1) & \text{for} & n = 7, & \\ \mathrm{cond}_{A,Ax}(A,x) & = & 4\varepsilon^{-2} + \mathcal{O}(\varepsilon^{-1}) & \text{for} & \text{odd } n \geq 9 & \text{but} \\ \mathrm{cond}^{\mathrm{symToep}}_{A,Ax}(A,x) & = & 5 + \mathcal{O}(\varepsilon) & \text{for} & \text{odd } n \geq 7. & \end{array}$$

For persymmetric Hankel structures we use Theorem 3.2 and, summarizing, we have the following result.

THEOREM 6.1. *For $n \geq 5$, there exist parameterized symmetric Toeplitz matrices $A := A_\varepsilon \in M_n^{\mathrm{symToep}}(\mathbb{R})$ and $x \in \mathbb{R}^n$ such that*

$$\mathrm{cond}_{A,Ax}(A,x) \geq 4\varepsilon^{-1} + \mathcal{O}(1) \quad \textit{and} \quad \mathrm{cond}^{\mathrm{symToep}}_{A,Ax}(A,x) \leq 6 + \mathcal{O}(\varepsilon).$$

For the persymmetric Hankel matrix $JA \in M_n^{\mathrm{persymHankel}}(\mathbb{R})$ similar assertions are true.

For Hankel structures consider

$$\begin{array}{rcl} A & = & \mathrm{Hankel}([0,\varepsilon,-1+\varepsilon,-1,0],[0,1,1,0,0]) \in M_5^{\mathrm{Hankel}}(\mathbb{R}) \quad \text{and} \\ x & = & (1,1,0,1,1)^T \in \mathbb{R}^5, \end{array}$$

where $\mathrm{Hankel}(c,r)$ denotes the Hankel matrix with first column c and last row r. Then (2.2) and Theorem 3.1 give

$$\mathrm{cond}_{A,Ax}(A,x) = 8\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}^{\mathrm{Hankel}}_{A,Ax}(A,x) = 8 + \mathcal{O}(\varepsilon). \tag{6.1}$$

For general even $n \geq 6$ consider

$$\begin{array}{rcl} A & = & \mathrm{Hankel}([\varepsilon,1,z,1,-1,z,0,0],[0,0,z,-1,1,z,1,0]) \in M_{6+2k}^{\mathrm{Hankel}}(\mathbb{R}), \\ x & = & (1,0,z,-1,-1,z,0,1)^T \in \mathbb{R}^{6+2k}, \end{array}$$

where z denotes a vector of $k \geq 0$ zeros. Then

$$\mathrm{cond}_{A,Ax}(A,x) = 8\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad \mathrm{cond}^{\mathrm{Hankel}}_{A,Ax}(A,x) = 7 + \mathcal{O}(\varepsilon)$$

for all even $n \geq 6$. For general odd $n \geq 7$ define

$$\begin{array}{rcl} A & = & \mathrm{Hankel}([\varepsilon,z,0,-1,-1,0,z],[z,0,1-\varepsilon,1,0,0,z]) \in M_{5+2k}^{\mathrm{Hankel}}(\mathbb{R}) \quad \text{and} \\ x & = & (1,1,z,0,z,1,1)^T \in \mathbb{R}^{5+2k}, \end{array}$$

where z denotes a vector of $k \geq 1$ zeros. Then (6.1) is valid as well. Using Theorem 3.2 for general Toeplitz structures we have the following result.

THEOREM 6.2. *For $n \geq 5$, there exist parameterized Hankel matrices $A := A_\varepsilon \in M_n^{\mathrm{Hankel}}(\mathbb{R})$ and $x \in \mathbb{R}^n$ such that*

$$\mathrm{cond}_{A,Ax}(A,x) = 8\varepsilon^{-1} + \mathcal{O}(1) \quad \textit{and} \quad \mathrm{cond}^{\mathrm{Hankel}}_{A,Ax}(A,x) \leq 8 + \mathcal{O}(\varepsilon).$$

For the general Toeplitz matrix $JA \in M^{\mathrm{Toep}}(\mathbb{R})$ similar assertions are true.

In summary, for all of the structures struct $\in \{\mathrm{sym}, \mathrm{persym}, \mathrm{skewsym}, \mathrm{symToep}, \mathrm{Toep}, \mathrm{Hankel}, \mathrm{persymHankel}\}$ there are general $n \times n$ examples, $n \geq 5$, such that the unstructured condition number is arbitrarily large, whereas the structured condition number is $\mathcal{O}(1)$. Note that this includes perturbations in the right-hand side. The only exception of the structures in (2.5) to this statement are circulant structures, as we will see in the next section.

For no perturbations in the right-hand side things are even worse. Consider

$$\begin{array}{llllll} A & := & A_\varepsilon & = & \mathrm{Toeplitz}(\varepsilon, v, 1, v, 0) & \text{for odd } n \geq 3 \quad \text{and} \\ A & := & A_\varepsilon & = & \mathrm{Toeplitz}(\varepsilon, w, 1, w, 0, 0) & \text{for even } n \geq 6, \end{array} \tag{6.2}$$

where $v \in \mathbb{R}^{k-1}$ and $w \in \mathbb{R}^{k-2}$ denote zero vectors for $k := \lfloor n/2 \rfloor$. We will show that for these matrices a linear system $Ax = b$ is *always well conditioned* with respect to componentwise symmetric Toeplitz perturbations, that is, *for all* x. On the other hand, the linear system is *ill conditioned* for generic x with respect to componentwise general perturbations. We illustrate the proof for $n = 5$. Let $x \in \mathbb{R}^5$ with $\|x\|_\infty = 1$ be given. According to (3.5) and (3.6) we calculate

$$A^{-1} = \frac{1}{2} \begin{pmatrix} \varepsilon^{-1} & 0 & 1 & 0 & -\varepsilon^{-1} \\ 0 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 \\ -\varepsilon^{-1} & 0 & 1 & 0 & \varepsilon^{-1} \end{pmatrix} + \mathcal{O}(\varepsilon) \quad \text{and}$$

$$\Psi_x^{\mathrm{symToep}} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ x_2 & x_1 + x_3 & x_4 & x_5 & 0 \\ x_3 & x_2 + x_4 & x_1 + x_5 & 0 & 0 \\ x_4 & x_3 + x_5 & x_2 & x_1 & 0 \\ x_5 & x_4 & x_3 & x_2 & x_1 \end{pmatrix},$$

and from this

$$A^{-1}\Psi_x = \frac{1}{2}\varepsilon^{-1} \begin{pmatrix} x_1 - x_5 & x_2 - x_4 & 0 & x_4 - x_2 & x_5 - x_1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ x_5 - x_1 & x_4 - x_2 & 0 & x_2 - x_4 & x_1 - x_5 \end{pmatrix} + \mathcal{O}(1).$$

By construction (3.3) we have $\mathrm{vec}(A) = \Phi^{\mathrm{symToep}} \cdot p_A$ with $p_A = (\varepsilon, 0, 1, 0, 0)^T$. The only element of p_A of size 1 meets the zero column in $A^{-1}\Psi_x$, so $|A^{-1}\Psi_x|\,|p_A| = \mathcal{O}(1)$, and by (3.6)

$$\mathrm{cond}_A^{\mathrm{symToep}}(A, x) = \mathcal{O}(1) \quad \text{for all } 0 \neq x \in \mathbb{R}^n.$$

By (3.6) this remains true without the assumption $\|x\|_\infty = 1$. On the other hand,

$$|A^{-1}|\,|A| = \begin{pmatrix} 1 & 0 & \varepsilon^{-1} & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & \varepsilon^{-1} & 0 & 1 \end{pmatrix} + \mathcal{O}(\mathrm{eps}).$$

So (2.2) implies

$$\mathrm{cond}_A(A, x) \geq \varepsilon^{-1}|x_3|.$$

The computation above extends to all matrices in (6.2) and we have the following result.

THEOREM 6.3. *Let $\varepsilon > 0$ and a matrix $A := A_\varepsilon$ according to* (6.2) *be given. Then the following are true:*

(i) *For all $0 \neq x \in \mathbb{R}^n$ we have $\mathrm{cond}_A^{\mathrm{symToep}}(A, x) = \mathcal{O}(1)$.*

(ii) *Let $x \in \mathbb{R}^n$ be given and denote $\alpha := |x_{k+1}|$ for odd n and $\alpha := \max(|x_k|, |x_{k+1}|)$ for even n, where $k = \lfloor n/2 \rfloor$. Then*

$$\mathrm{cond}_A(A, x) \geq \varepsilon^{-1}\alpha.$$

The example is possible because the symmetric Toeplitz structure imposes severe restrictions on the possible perturbations so that the assumptions of Theorem 4.4 are not satisfied.

7. Circulant matrices. Better estimations of the ratio $\mathrm{cond}^{\mathrm{circ}}/\mathrm{cond}$ of the componentwise condition numbers are possible because circulant matrices commute (because they are diagonalized by the Fourier matrix; cf. [7, 11]). This implies for $A, \Delta A \in M_n^{\mathrm{circ}}(\mathbb{R})$ that $A^{-1}\Delta A = \Delta A \cdot A^{-1}$ and therefore

$$\Delta x = -\Delta A \cdot A^{-1}x + A^{-1}\Delta b + \mathcal{O}(\varepsilon^2).$$

This implies the following nice characterization.

THEOREM 7.1. *Let nonsingular $A \in M_n^{\mathrm{circ}}(\mathbb{R})$, $x \in \mathbb{R}^n$, $E \in M_n^{\mathrm{circ}}(\mathbb{R})$, $f \in \mathbb{R}^n$ be given. Then*

$$\mathrm{cond}_{E,f}^{\mathrm{circ}}(A, x) = \frac{\|\, |E|\,|A^{-1}x| + |A^{-1}|\,|f|\,\|_\infty}{\|x\|_\infty}.$$

To estimate the ratio $\mathrm{cond}^{\mathrm{circ}}/\mathrm{cond}$ we first show that

$$\|\, |A|\,|x|\,\|_\infty \geq n^{-1}\|A\|_\infty \|x\|_2 \quad \text{for} \quad A \in M_n^{\mathrm{circ}}(\mathbb{R}). \tag{7.1}$$

For $A \in M_n^{\mathrm{circ}}(\mathbb{R})$ we have $\sum_i |A_{ij}| = \|A\|_1$ for any j and $\|A\|_1 = \|A\|_\infty$, and therefore

$$\begin{aligned} \|\, |A|\,|x|\,\|_\infty &\geq n^{-1}\sum_i (|A|\,|x|)_i = n^{-1}\sum_i\sum_j |A_{ij}|\,|x_j| = n^{-1}\sum_j \|A\|_1 |x_j| \\ &= n^{-1}\|A\|_1\|x\|_1 \geq n^{-1}\|A\|_\infty\|x\|_2. \end{aligned}$$

This implies

$$\begin{aligned} \|\, |A|\,|A^{-1}x|\,\|_\infty &\geq n^{-1}\|A\|_\infty\|A^{-1}x\|_2 \quad \text{and} \\ \|\, |A^{-1}|\,|Ax|\,\|_\infty &\geq n^{-1}\|A^{-1}\|_\infty\|Ax\|_2. \end{aligned}$$

For $x \in \mathbb{R}^n$ we have

$$\|x\|_2^2 = x^TA^{-1}Ax \leq \|x^TA^{-1}\|_2\|Ax\|_2 = \|A^{-1}x\|_2\|Ax\|_2$$

using $\|C^Tx\|_2 = \|Cx\|_2$ for $C \in M_n^{\mathrm{circ}}(\mathbb{R})$; see Part I, Lemma 8.2. Putting things together we obtain for relative perturbations $E = A$ and $f = Ax$

$$\begin{aligned} \mathrm{cond}_{A,Ax}^{\mathrm{circ}}(A, x) &= \|\, |A|\,|A^{-1}x| + |A^{-1}|\,|Ax|\,\|_\infty/\|x\|_\infty \\ &\geq n^{-1}\max(\|A\|_\infty\|A^{-1}x\|_2, \|A^{-1}\|_\infty\|Ax\|_2)/\|x\|_\infty \\ &\geq n^{-1}\sqrt{\|A\|_\infty\|A^{-1}\|_\infty\|A^{-1}x\|_2\|Ax\|_2}/\|x\|_\infty \\ &\geq n^{-1}\sqrt{\|A\|_\infty\|A^{-1}\|_\infty}\frac{\|x\|_2}{\|x\|_\infty}. \end{aligned} \tag{7.2}$$

With this and (4.1) and (2.3) we also obtain a lower bound on the ratio $\mathrm{cond}^{\mathrm{circ}}_{A,Ax}/\mathrm{cond}_{A,Ax}$ by

$$
\begin{aligned}
\mathrm{cond}^{\mathrm{circ}}_{A,Ax}(A,x) &\geq n^{-1}\sqrt{\|\,|A^{-1}|\,|A|\,\|_\infty}\,\frac{\|x\|_2}{\|x\|_\infty} \\
&\geq n^{-1}\sqrt{\frac{\|\,|A^{-1}|\,|A|x|\|_\infty}{\|x\|_\infty}}\cdot\frac{\|x\|_2}{\|x\|_\infty} \\
&\geq n^{-1}\sqrt{\frac{1}{2}\mathrm{cond}_{A,Ax}(A,x)}\cdot\frac{\|x\|_2}{\|x\|_\infty} \\
&\geq 2^{-1/2}n^{-1}\sqrt{\mathrm{cond}_{A,Ax}(A,x)}.
\end{aligned}
\tag{7.3}
$$

We have the following result.

THEOREM 7.2. *Let a nonsingular circulant $A \in M^{\mathrm{circ}}_n(\mathbb{R})$ and $0 \neq x \in \mathbb{R}^n$ be given. Then*

$$
\begin{aligned}
\mathrm{cond}^{\mathrm{circ}}_{A,Ax}(A,x) &\geq n^{-1}\sqrt{\|A\|_\infty\|A^{-1}\|_\infty}\cdot\frac{\|x\|_2}{\|x\|_\infty} \\
&\geq 2^{-1/2}n^{-1}\sqrt{\mathrm{cond}_{A,Ax}(A,x)}\cdot\frac{\|x\|_2}{\|x\|_\infty}.
\end{aligned}
$$

We think that the factor n^{-1} in both lower bounds of Theorem 7.2 can be replaced by the factor $n^{-1/2}$. If this is true, it is easy to find examples verifying that the overestimation in either case is bounded by a small constant factor.

The ratio $\mathrm{cond}^{\mathrm{circ}}_{A,Ax}/\mathrm{cond}_{A,Ax}$ can only become large for ill-conditioned linear systems. The question remains of whether this changes if we forbid perturbations in the right-hand side. This is indeed the case, and it is very simple to find examples. Just take a vector $r \in \mathbb{R}^n$ with uniformly distributed random first $n-1$ components in $[-1,1]$ and set $r_n := -\sum_{i=1}^{n-1} r_i + \varepsilon$. Then define $A = [\mathrm{circ}(r)]^{-1}$ and $x = e$. Obviously $\mathrm{circ}(r)e = \varepsilon e$, so A has an eigenvalue ε^{-1} to the eigenvector e; one can see that most likely $A > 0$, so Theorem 7.1 implies $\mathrm{cond}^{\mathrm{circ}}_A(A,e) = \|\,|A|\,|A^{-1}e|\,\|_\infty = 1$. On the other hand, (2.3) implies $\mathrm{cond}_A(A,e) = \|\,|A^{-1}|\,|A|e|\|_\infty = \|\,|A^{-1}|\,|A|\,\|_\infty$, and extensive numerical experience shows that it is likely that $\mathrm{cond}_A(A,x) \sim \varepsilon^{-1}$. An explicit example is a matrix A constructed as above with $r_1 = \cdots = r_{n-1} = 1$. Then

$$
\mathrm{cond}_A(A,e) = (2n-2)\varepsilon^{-1} + \mathcal{O}(1) \quad\text{and}\quad \mathrm{cond}^{\mathrm{circ}}_A(A,e) = 1
$$

for $n \geq 2$.

THEOREM 7.3. *Given $n \geq 2$, there exists $A \in M^{\mathrm{circ}}_n(\mathbb{R})$ with*

$$
\mathrm{cond}_A(A,e) \geq \mathcal{O}(\varepsilon^{-1}) \quad\textit{and}\quad \mathrm{cond}^{\mathrm{circ}}_A(A,e) = 1.
$$

The results show that with respect to normwise and componentwise perturbations circulants behave similarly (Part I, Theorems 8.1, 8.4, and equation (8.1)). Besides normality, a reason for that is that circulants commute.

8. Inversion of structured matrices. Similar to the structured componentwise condition number for linear systems, the structured componentwise condition number for matrix inversion is defined for $A \in M^{\mathrm{struct}}_n(\mathbb{R})$ and given weight matrix $E \in M^{\mathrm{struct}}_n(\mathbb{R})$ by

$$
\mu^{\mathrm{struct}}_E(A) := \lim_{\varepsilon\to 0}\sup\left\{\frac{\|(A+\Delta A)^{-1} - A^{-1}\|_\infty}{\varepsilon\|A^{-1}\|_\infty};\quad \Delta A \in M^{\mathrm{struct}}_n(E),\ |\Delta A| \leq \varepsilon|E|\right\}.
\tag{8.1}
$$

The unstructured condition number $\mu_E(A)$, that is, for $M_n^{\text{struct}}(\mathbb{R}) = M_n(\mathbb{R})$, satisfies the following bounds:

$$n^{-1}\alpha \le \mu_E(A) \le \alpha \quad \text{for} \quad \alpha := \frac{\|\,|A^{-1}|\,|E|\,|A^{-1}|\,\|_\infty}{\|A^{-1}\|_\infty}. \tag{8.2}$$

This follows by the well-known ansatz (see, for example, [16, proof of Theorem 6.4])

$$(A+\Delta A)^{-1} - A^{-1} = -A^{-1}\Delta A A^{-1} + \mathcal{O}(\|\Delta A\|^2). \tag{8.3}$$

From this the right inequality in (8.2) is obvious. Denoting the ith row and jth column of A^{-1} by $A^{-1}_{i,:}$ and $A^{-1}_{:,j}$, respectively, we have

(8.4)
$$(A^{-1}\Delta A A^{-1})_{ij} = (|A^{-1}|\,|E|\,|A^{-1}|)_{ij} \quad \text{for} \quad \Delta A := \operatorname{diag}(\mathit{sign}(A^{-1}_{i,:}))|E|\operatorname{diag}(\mathit{sign}(A^{-1}_{:,j})),$$

which implies the left inequality of (8.2).

In case of normwise perturbations the condition numbers for matrix inversion and for an arbitrary linear system with the same matrix (for no perturbations in the right-hand side) are both equal to $\|A^{-1}E\|_2$. In case of componentwise perturbations the condition number depends on the solution (see (4.2) and (4.3)). We may ask whether there is a relation between $\mu_E(A)$ and the supremum of $\operatorname{cond}_E(A,x)$ over all x.

Definition 8.1. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{R})$ and $E \in M_n^{\text{struct}}(\mathbb{R})$ be given. Then*

$$\operatorname{cond}_E^{\text{struct}}(A) := \sup_{x\neq 0} \operatorname{cond}_E^{\text{struct}}(A,x).$$

In Corollary 4.3 we saw

$$\operatorname{cond}_E(A) = \operatorname{cond}_E^{\text{struct}}(A) = \|\,|A^{-1}|E\|_\infty \quad \text{for struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}. \tag{8.5}$$

Obviously (8.2) implies

$$\mu_E(A) \le \operatorname{cond}_E(A).$$

However, this inequality may be arbitrarily weak. Consider

$$A = A_\varepsilon = \begin{pmatrix} \varepsilon & 1 & 0 \\ 1 & \varepsilon & 1 \\ 0 & 1 & \varepsilon \end{pmatrix} \quad \text{with} \quad \frac{\|\,|A^{-1}|\,|A|\,|A^{-1}|\,\|_\infty}{\|A^{-1}\|_\infty} = 3, \tag{8.6}$$
$$\text{but} \quad \operatorname{cond}_A(A,e) = \|\,|A^{-1}|\,|A|\,\|_\infty = \varepsilon^{-1}.$$

Note that this is for componentwise relative perturbations, i.e., $E = A$. Denote $b := Ae$. Then (8.6) implies that the linear system $Ax = b$ is ill conditioned for small ε, but matrix inversion of A is well conditioned for every $\varepsilon > 0$. This might lead to the apparent contradiction that solving the linear system by $x = A^{-1}b$ removes the ill-conditionedness. This is of course not the case. In our example we have

$$A^{-1} = (2\varepsilon)^{-1}\begin{pmatrix} 1 & \varepsilon & -1 \\ \varepsilon & 0 & \varepsilon \\ -1 & \varepsilon & 1 \end{pmatrix} + \mathcal{O}(1),$$

so an $\mathcal{O}(1)$ change in A^{-1} is a small perturbation. However, $b = Ae = (1+\varepsilon, 2+\varepsilon, 1+\varepsilon)^T$, so an $\mathcal{O}(1)$ change in A^{-1} causes an $\mathcal{O}(1)$ perturbation in $x = A^{-1}b = e$, which is of the order of 100% change.

The condition number $\mu_E(A)$ depends on diagonal scaling of A (and E). We may ask for the optimal condition number with respect to two-sided diagonal scaling. For this we obtain the following result.

THEOREM 8.2. *Let nonsingular $A \in M_n(\mathbb{R})$ and $E \in M_n(\mathbb{R})$ be given. Denote*

$$\mu_E^{\text{opt}}(A) := \inf_{D_1, D_2} \mu_{D_1 E D_2}(D_1 A D_2), \tag{8.7}$$

where the infimum is taken over nonsingular diagonal matrices. Define

$$r := \min_{i,j} \frac{(|A^{-1}|\,|E|\,|A^{-1}|)_{ij}}{|A^{-1}|_{ij}}, \tag{8.8}$$

where $\alpha/0 := \infty$ for $\alpha \ge 0$. Then

$$n^{-1} r \le \mu_E^{\text{opt}}(A) \le r. \tag{8.9}$$

Proof. Let i, j be indices realizing the minimum in the definition (8.8) of r and let $D^{(\nu)} := \operatorname{diag}(\varepsilon, \ldots, \varepsilon, 1, \varepsilon, \ldots, \varepsilon)$ with the 1 at the νth position. Defining $D_1^{-1} := D^{(j)}$ and $D_2^{-1} := D^{(i)}$ we obtain by (8.2)

$$\begin{aligned}\mu_E^{\text{opt}}(A) \le \mu_{D_1 E D_2}(D_1 A D_2) &\le \frac{\|D^{(i)}|A^{-1}|\,|E|\,|A^{-1}|D^{(j)}\|_\infty}{\|D^{(i)} A^{-1} D^{(j)}\|_\infty} \\ &= \frac{(|A^{-1}|\,|E|\,|A^{-1}|)_{ij}}{|A^{-1}|_{ij}} + \beta\varepsilon = r + \beta\varepsilon\end{aligned}$$

for a constant β not depending on ε. This proves the right inequality in (8.9). Denote $C := |A^{-1}|\,|E|\,|A^{-1}|$ and let $\|A^{-1}\|_\infty = \sum_\nu |A^{-1}|_{i\nu}$ and $\|C\|_\infty = \sum_\nu C_{j\nu}$. Then by (8.2) and the definition (8.8) of r

$$n\mu_E(A) \ge \frac{\|C\|_\infty}{\|A^{-1}\|_\infty} = \frac{\sum\limits_\nu C_{j\nu}}{\sum\limits_\nu |A^{-1}|_{i\nu}} \ge \frac{\sum\limits_\nu C_{i\nu}}{\sum\limits_\nu |A^{-1}|_{i\nu}} \ge \frac{r\sum\limits_\nu |A^{-1}|_{i\nu}}{\sum\limits_\nu |A^{-1}|_{i\nu}} = r. \qquad \square$$

We note that one may measure the componentwise relative perturbation of $(A+\Delta A)^{-1}$ versus A^{-1} subject to componentwise perturbations of A. Then (cf. [4, 16])

$$\begin{aligned}\tilde{\mu}_E(A) \quad &:= \quad \lim_{\varepsilon\to 0} \sup\left\{ \frac{|(A+\Delta A)^{-1} - A^{-1}|_{ij}}{\varepsilon |A^{-1}|_{ij}} : \quad \Delta A \in M_n(\mathbb{R}),\ |\Delta A| \le \varepsilon |E| \right\} \\ &= \quad \max_{ij} \frac{(|A^{-1}|\,|E|\,|A^{-1}|)_{ij}}{|A^{-1}|_{ij}}.\end{aligned}$$

The structured componentwise condition number for the inverse can be bounded by adapting the approach for linear systems. Let $A, \Delta A \in M_n^{\text{struct}}(\mathbb{R})$, $\operatorname{vec}(\Delta A) = \Phi^{\text{struct}} p_{\Delta A}$, and $\operatorname{vec}(E) = \Phi^{\text{struct}} p_E$. Then $|\Delta A| \le |E|$ is equivalent to $|p_{\Delta A}| \le |p_E|$. In view of (8.5) we note that (see [17, Lemma 4.3.1])

$$\operatorname{vec}(A^{-1}\Delta A A^{-1}) = (A^{-T} \otimes A^{-1})\operatorname{vec}(\Delta A) = (A^{-T} \otimes A^{-1})\Phi^{\text{struct}} p_{\Delta A}. \tag{8.10}$$

This implies the following.

THEOREM 8.3. *Let nonsingular $A \in M_n^{\text{struct}}(\mathbb{R})$ and $E \in M_n^{\text{struct}}(\mathbb{R})$ be given. Let $B \in M_n(\mathbb{R})$ with*

$$\text{vec}(B) = |(A^{-T} \otimes A^{-1})\Phi^{\text{struct}}|\,|p_E| \tag{8.11}$$

and denote

$$\alpha := \frac{\|B\|_\infty}{\|A^{-1}\|_\infty}.$$

Then

$$n^{-1}\alpha \le \mu_E^{\text{struct}}(A) \le \alpha. \tag{8.12}$$

Remark 8.4. The result includes (8.2) because, in the unstructured case, $\Phi = I_{n^2}$ and $\text{vec}(B) = |A^{-T} \otimes A^{-1}|\,|p_E| = (|A^{-T}| \otimes |A^{-1}|)|p_E| = \text{vec}(|A^{-1}|\,|E|\,|A^{-1}|)$ by [17, Lemma 4.3.1].

Proof. Let $dA \in M_n^{\text{struct}}(\mathbb{R})$ such that $\|A^{-1}dAA^{-1}\|_\infty = \sup\{\|A^{-1}\Delta AA^{-1}\| : |\Delta A| \le |E|\}$, and denote $\text{vec}(dA) = \Phi^{\text{struct}}p_{dA}$. Then $|p_{dA}| \le |p_E|$ implies $|\text{vec}(A^{-1}dAA^{-1})| = |(A^{-T} \otimes A^{-1})\Phi^{\text{struct}} \cdot p_{dA}| \le \text{vec}(B)$, and the right inequality in (8.12) follows by (8.10), (8.3), and the definition (8.1). On the other hand, let the index k, $1 \le k \le n^2$, be such that $\max_{\mu,\nu}|B_{\mu\nu}| = (\text{vec}(B))_k$. Denote $C := (A^{-T} \otimes A^{-1})\Phi^{\text{struct}}$ and set diagonal $D \in M_{n^2}(\mathbb{R})$ with $D_{\nu\nu} := sign(C_{k\nu})$. Furthermore, define $p_{\Delta A} := D|p_E|$ and let $\Delta A \in M_n^{\text{struct}}(\mathbb{R})$ with $\text{vec}(\Delta A) = \Phi^{\text{struct}}p_{\Delta A}$. Then

$$\beta := (\text{vec}(A^{-1}\Delta AA^{-1}))_k = ((A^{-T} \otimes A^{-1})\Phi^{\text{struct}}p_{\Delta A})_k = (|C|\,|p_E|)_k = (\text{vec}(B))_k$$

and

$$\mu_E^{\text{struct}}(A) \ge \frac{\|A^{-1}\Delta AA^{-1}\|_\infty}{\|A^{-1}\|} \ge \frac{\beta}{\|A^{-1}\|} \ge n^{-1}\frac{\|B\|_\infty}{\|A^{-1}\|}. \qquad \square$$

The question remains of whether, as for normwise perturbations, there is a relation between the reciprocal of the matrix condition number and the componentwise distance to the nearest singular matrix. This question will be treated in the next section.

9. Distance to singularity. For normwise and unstructured perturbations the condition number is equal to the reciprocal of the distance to the nearest singular matrix. Moreover, we showed in Part I, Theorem 12.1 that this is also true for structured (normwise) perturbations, that is,

$$\delta_E^{\text{struct}}(A) = \kappa_E(A)^{-1},$$

which is true for all our structures (2.5) under investigation. In the limit, a matrix has condition number ∞ iff it is singular, that is, the distance to singularity is 0.

The question arises of whether a similar result can be proved for componentwise perturbations, unstructured or structured. The componentwise (structured) distance to the nearest singular matrix is defined by

$$\text{d}_E^{\text{struct}}(A) := \min\{\alpha : \Delta A \in M_n^{\text{struct}}(\mathbb{R}),\ |\Delta A| \le \alpha|E|,\ A + \Delta A \text{ singular}\}. \tag{9.1}$$

For normwise perturbations, the distance to singularity $\delta_E^{\text{struct}}(A)$ as well as the condition number $\kappa_E(A)$ depend on row and column diagonal scaling of the matrix. This is no longer true for componentwise perturbations. The unstructured distance to singularity $\mathrm{d}_E(A)$ as well as the structured distance is independent of row and column diagonal scaling (as long as, of course, the scaled matrix remains in the structure). That is, for positive diagonal D_1, D_2,

$$\mathrm{d}_{D_1ED_2}(D_1AD_2) = \mathrm{d}_E(A),$$

and for $A, E, D_1AD_2, D_1ED_2 \in M_n^{\mathrm{struct}}(\mathbb{R})$,

$$\mathrm{d}_{D_1ED_2}^{\mathrm{struct}}(D_1ED_2) = \mathrm{d}_E^{\mathrm{struct}}(A).$$

This is simply because $|\Delta A| \le \alpha|E| \;\Leftrightarrow\; |D_1\Delta AD_2| \le \alpha D_1|E|D_2$ and $\det(A+\Delta A) = 0 \;\Leftrightarrow\; \det(D_1AD_2 + D_1\Delta AD_2) = 0$. Furthermore, $A + \tilde{E} = A(I + A^{-1}\tilde{E})$ is singular iff -1 is an eigenvalue of $A^{-1}\tilde{E}$ so that definition (9.1) implies

(9.2)
$$\mathrm{d}_E^{\mathrm{struct}}(A) = \Big[\max\{|\lambda| :\; \tilde{E} \in M_n^{\mathrm{struct}}(\mathbb{R}),\, |\tilde{E}| \le |E|,\, \lambda \text{ real eigenvalue of } A^{-1}\tilde{E}\}\Big]^{-1}.$$

Note that the maximum is taken only over real eigenvalues of $A^{-1}\tilde{E}$. For unstructured perturbations the linearity of the determinant in each matrix element implies that the matrices $\tilde{E}$ can be restricted to the boundary $|\tilde{E}| = |E|$, i.e., finitely many matrices:

$$\mathrm{d}_E(A) = \Big[\max\{|\lambda| :\; \tilde{E} \in M_n(\mathbb{R}),\, |\tilde{E}| = |E|,\, \lambda \text{ real eigenvalue of } A^{-1}\tilde{E}\}\Big]^{-1}. \tag{9.3}$$

This is not true for structured perturbations; that is, the maximum may only be achieved for some $|\tilde{E}| \ne |E|$. An example for symmetric structures was given in [20].

With respect to the condition number things are even more involved. In the normwise case, we have $\kappa_E^{\mathrm{struct}}(A) = \|A^{-1}\|\,\|E\|$ (see Part I, Theorem 11.1) for all structures in (2.5), and it is the same condition number for matrix inversion as for linear systems when taking the supremum over all x. This is no longer true in the componentwise case. Here the condition numbers for matrix inversion and a linear system with the same matrix may be arbitrarily far apart (cf. the example in (8.6)). So if there is a relation at all between distance to singularity and the reciprocal of a condition number we first have to discuss which is the "right" condition number to choose.

Let us first consider the condition number $\mu_E(A)$ of the matrix inverse as defined in the previous section. Consider

$$A = A_\varepsilon = \begin{pmatrix} -\varepsilon & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad E = A.$$

By (9.3) we calculate

$$\mathrm{d}_A(A) = \frac{1}{4}\varepsilon^{1/2} + \mathcal{O}(1).$$

On the other hand (8.2) yields

$$\mu_A(A) \le \frac{\|\,|A^{-1}|\,|A|\,|A^{-1}|\,\|_\infty}{\|A^{-1}\|_\infty} \le 8. \tag{9.4}$$

Note that this is true for the most common case $E = A$ of componentwise relative perturbations. The same example applies to structured perturbations. The perturbation ΔA with $|\Delta A| = \mathrm{d}_A(A)|A|$ and $\det(A + \Delta A) = 0$ is a symmetric matrix. That means

$$\mathrm{d}_A^{\mathrm{sym}}(A) = \frac{1}{4}\varepsilon^{1/2} + \mathcal{O}(1) \quad \text{and} \quad \mu_A^{\mathrm{sym}}(A) \le \mu_A(A) \le 8.$$

So the condition number of the matrix inverse does not seem appropriate for our anticipated results.

To proceed let us first consider unstructured componentwise perturbations. Then, by Corollary 4.3, the worst case condition number for all x is $\mathrm{cond}_E(A) = \sup_{x \ne 0} \mathrm{cond}_E(A, x) = \mathrm{cond}_E(A, e) = \|\, |A^{-1}|\,|E|\,\|_\infty$. We choose no perturbations in the right-hand side because we are interested in the *matrix property* of distance to singularity. By column diagonal scaling, $\|\, |A^{-1}|\,|E|\,\|_\infty$ may become arbitrarily large. Therefore we choose optimal diagonal scaling for which [3, 9, 23]

$$\inf_{D_1, D_2} \mathrm{cond}_{D_1 E D_2}(D_1 A D_2) = \varrho(|A^{-1}|\,|E|), \tag{9.5}$$

the infimum taken over nonsingular diagonal D_ν, ϱ denoting the spectral radius. Note that $\varrho(|A^{-1}|\,|E|)$ is also equal to the minimum *normwise* condition number $\kappa_{E,\infty}(A)$ with respect to the ∞-norm achievable by diagonal scaling. For this minimum condition number we could indeed show an inverse proportionality to $\mathrm{d}_E(A)$ as by [21, Proposition 5.1]

$$\frac{1}{\varrho(|A^{-1}|\,|E|)} \le \mathrm{d}_E(A) \le \frac{(3 + 2\sqrt{2})n}{\varrho(|A^{-1}|\,|E|)}. \tag{9.6}$$

The left inequality is an equality for large classes of matrices, e.g., M-matrices. Moreover, explicit $n \times n$ examples, $n \ge 1$, were given [21] with $\mathrm{d}_A(A) = n\varrho(|A^{-1}|\,|A|)^{-1}$, so there is not much room for improvement in (9.6).

The question remains of whether a similar result is possible in case of structured componentwise perturbations. Unfortunately, for all structures in (2.5) the answer is no. Following we give a sequence of examples showing that. The first example for the symmetric case will be treated in more detail; the rest follow similarly. All examples will be given for the important case of componentwise relative perturbations of the matrix entries.

Let

$$A = A_\varepsilon = \begin{pmatrix} \varepsilon & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & \varepsilon \\ 1 & 1 & \varepsilon & 0 \end{pmatrix} \in M_4^{\mathrm{sym}}(\mathbb{R})$$

be given for $\varepsilon > 0$. A general symmetric perturbation of A subject to componentwise relative perturbations is

$$\tilde{A} = \begin{pmatrix} \varepsilon(1+\delta_1) & 0 & 1+\delta_2 & 1+\delta_3 \\ 0 & 0 & 1+\delta_4 & 1+\delta_5 \\ 1+\delta_2 & 1+\delta_4 & 0 & \varepsilon(1+\delta_6) \\ 1+\delta_3 & 1+\delta_5 & \varepsilon(1+\delta_6) & 0 \end{pmatrix} \in M_4^{\mathrm{sym}}(\mathbb{R}).$$

Then $\mathrm{d}_A^{\mathrm{sym}}(A)$ is the smallest α such that $|\delta_\nu| \le \alpha$ and $\det \tilde{A} = 0$. With Maple [25] we calculate

$$\begin{aligned} \det \tilde{A} &= c_0 + c_2\varepsilon^2 \quad \text{with} \\ c_0 &= ((1+\delta_2)(1+\delta_5) - (1+\delta_3)(1+\delta_4))^2, \\ c_2 &= 2(1+\delta_1)(1+\delta_4)(1+\delta_5)(1+\delta_6). \end{aligned}$$

In order to move $\det \tilde{A}$ into zero, the second summand $c_2\varepsilon^2$ must be zero or negative. This implies $\mathrm{d}_A^{\mathrm{sym}}(A) \ge 1$ and therefore, of course

$$\mathrm{d}_A^{\mathrm{sym}}(A) = 1$$

because $\tilde{A} \equiv 0$ for $\delta_\nu \equiv -1$. On the other hand,

$$\mathrm{cond}_A(A) = \mathrm{cond}_A^{\mathrm{sym}}(A) = \| \, |A^{-1}| \, |A| \, \|_\infty = 4\varepsilon^{-1} + \mathcal{O}(1).$$

Moreover, (9.5) implies

$$\inf_D \mathrm{cond}_{DAD}^{\mathrm{sym}}(DAD) \ge \varrho(|A^{-1}| \, |A|) = 2.8\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \quad d_A(A) = \varrho(|A^{-1}| \, |A|)^{-1}$$

so that there are arbitrarily ill-conditioned, though optimally scaled, symmetric matrices with $\mathrm{d}_A^{\mathrm{sym}}(A) = 1$. In other words, no relative perturbation less than 100% may move A into the manifold of (symmetric) singular matrices. The example above is extendable to higher dimensions by choosing $A \oplus I$. By Theorem 3.2 the example extends also to persymmetric structures.

For the skewsymmetric case consider

$$A = A_\varepsilon = \begin{pmatrix} 0 & 0 & -1 & 1-\varepsilon & 0 & 0 \\ 0 & 0 & -\varepsilon & 0 & 1+\varepsilon & 1 \\ 1 & \varepsilon & 0 & -1 & 0 & 0 \\ -1+\varepsilon & 0 & 1 & 0 & 0 & 0 \\ 0 & -1-\varepsilon & 0 & 0 & 0 & -\varepsilon \\ 0 & -1 & 0 & 0 & \varepsilon & 0 \end{pmatrix} \in M_6^{\mathrm{skewsym}}(\mathbb{R}).$$

Here and in the following examples we define $\tilde{A}$ (as for the symmetric case) by multiplying the components of A by $1+\delta_\nu$ with rowwise numbering of the δ_ν. Then

$$\det \tilde{A} = \varepsilon^4(1-\varepsilon)^2(1+\delta_2)^2(1+\delta_3)^2(1+\delta_7)^2,$$

implying

$$\mathrm{d}_A^{\mathrm{skewsym}}(A) = 1.$$

On the other hand,

$$\mathrm{cond}_A(A) = 6\varepsilon^{-2} + \mathcal{O}(\varepsilon^{-1}) \quad \text{and} \quad \mathrm{cond}_A^{\mathrm{skewsym}}(A) = 2\varepsilon^{-2} + \mathcal{O}(\varepsilon^{-1}),$$

whereas

$$\begin{aligned} &\inf_D \mathrm{cond}_{DAD}^{\mathrm{skewsym}}(DAD) \ge \varrho(|A^{-1}| \, |A|) = 6\varepsilon^{-3/2} + \mathcal{O}(\varepsilon^{1/2}) \quad \text{and} \\ &\mathrm{d}_A(A) = \varrho(|A^{-1}| \, |A|)^{-1} \end{aligned}$$

such that A is truly ill conditioned for arbitrary diagonal scaling with respect to relative componentwise skewsymmetric perturbations. Now (8.5) and (9.5) imply for $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$

$$\inf_D \text{cond}^{\text{struct}}_{DED}(DAD) = \inf_D \text{cond}_{DED}(DAD) = \varrho(|A^{-1}|\,|E|),$$

the infimum taken over positive diagonal D. So we have the following result.

THEOREM 9.1. *Let* $\text{struct} \in \{\text{sym}, \text{persym}, \text{skewsym}\}$. *Then for every* $\varepsilon > 0$ *there exists* $A := A_\varepsilon \in M_n^{\text{struct}}(\mathbb{R})$ *with*

$$\inf_D \text{cond}^{\text{struct}}_{DAD}(DAD) > \varepsilon^{-1} \quad \textit{and} \quad \mathrm{d}^{\text{struct}}_A(A) = 1.$$

For the symmetric Toeplitz case consider

$$A = A_\varepsilon = \text{Toeplitz}(0, 1, 1, -\varepsilon) \in M_4^{\text{symToep}}. \tag{9.7}$$

Then defining $\tilde{A}$ as before yields

$$\begin{aligned} \det \tilde{A} &= c_0 + c_1\varepsilon + c_2\varepsilon^2 \quad \text{with} \\ c_0 &= (2+\delta_1+\delta_2)^2(\delta_2-\delta_1)^2, \\ c_1 &= 2(1+\delta_1)(1+\delta_3)((1+\delta_1)^2+(1+\delta_2)^2), \quad \text{and} \\ c_2 &= (1+\delta_1)^2(1+\delta_3)^2. \end{aligned}$$

For $|\delta_\nu| < 1$, c_0 is nonnegative, whereas both c_1 and c_2 are positive. Therefore

$$\mathrm{d}^{\text{symToep}}_A(A) = 1.$$

On the other hand, for $x = (1, -1, 1, -1)^T$,

$$\begin{aligned} &\text{cond}_A(A) = 2\varepsilon^{-1} + \mathcal{O}(1) \quad \text{and} \\ &\sup_{x\neq 0} \text{cond}^{\text{symToep}}_A(A) \geq \text{cond}^{\text{symToep}}_A(A, x) = 2\varepsilon^{-1} + \mathcal{O}(1) \end{aligned}$$

such that A is truly ill conditioned subject to relative componentwise symmetric Toeplitz perturbations. For Toeplitz structures, diagonal scaling is, in general, not possible. For completeness we note

$$\varrho(|A^{-1}|\,|A|) = 2\varepsilon^{-1/2} + \mathcal{O}(1) = [\mathrm{d}_A(A)]^{-1}.$$

The same example applies, according to Theorem 3.2, to persymmetric Hankel structures.

For the general Toeplitz case consider

$$A = A_\varepsilon = \text{Toeplitz}([0, 1, -1, 0], [0, 1, -1, -\varepsilon]) \in M_4^{\text{Toep}}(\mathbb{R}),$$

where $\text{Toeplitz}(c, r)$ denotes the (general) Toeplitz matrix with first column c and first row r. Defining $\tilde{A}$ as before yields

$$\begin{aligned} \det \tilde{A} &= c_0 + c_1\varepsilon \quad \text{with} \\ c_0 &= ((1+\delta_1)(1+\delta_4) - (1+\delta_1)(1+\delta_5))^2 \quad \text{and} \\ c_1 &= (1+\delta_3)(1+\delta_4)^3 + (1+\delta_1)(1+\delta_3)(1+\delta_5)^2. \end{aligned}$$

For $|\delta_\nu| < 1$ the determinant is positive, so

$$\mathrm{d}_A^{\mathrm{Toep}}(A) = 1.$$

On the other hand,

$$\mathrm{cond}_A(A,e) = 4\varepsilon^{-1} + \mathcal{O}(1) = \mathrm{cond}_A^{\mathrm{Toep}}(A,e),$$

so A is truly ill conditioned subject to relative componentwise general Toeplitz perturbations. By Theorem 3.2 this also covers the Hankel case. For completeness we note

$$\varrho(|A^{-1}|\,|A|) = 2\sqrt{2}\varepsilon^{-1/2} + \mathcal{O}(1) = [\mathrm{d}_A(A)]^{-1}.$$

Finally, define for the circulant case

$$A = A_\varepsilon = \mathrm{circ}(1,\varepsilon,1,0) \in M_4^{\mathrm{circ}}(\mathbb{R}).$$

For $\tilde{A}$ defined as before we get

$$\begin{array}{rcl} \det\tilde{A} & = & \alpha\beta \quad \text{with} \\ \alpha & = & (2+\delta_1+\delta_3)^2 - \varepsilon^2(1+\delta_2)^2 \quad \text{and} \\ \beta & = & (\delta_1-\delta_3)^2 + \varepsilon^2(1+\delta_2)^2. \end{array}$$

For small ε both factors are nonzero for $|\delta_\nu| < 1$, so

$$\mathrm{d}_A^{\mathrm{circ}}(A) = 1.$$

On the other hand, for $x = (1,1,1,-1)^T$,

$$\begin{array}{l} \mathrm{cond}_A(A,x) = 2\varepsilon^{-1} + \mathcal{O}(1) = \mathrm{cond}_A^{\mathrm{circ}}(A,x) \quad \text{and} \\ \varrho(|A^{-1}|\,|A|) = 2\varepsilon^{-1} + \mathcal{O}(1) = [\mathrm{d}_A(A)]^{-1}. \end{array}$$

Summarizing, we have the following result.

THEOREM 9.2. *Let* struct $\in \{\mathrm{symToep}, \mathrm{Toep}, \mathrm{circ}, \mathrm{Hankel}, \mathrm{persymHankel}\}$. *Then for every* $\varepsilon > 0$ *there exists* $A := A_\varepsilon \in M_n^{\mathrm{struct}}(\mathbb{R})$ *and* $x \in \mathbb{R}^n$ *with* $|x| = e$ *such that*

$$\mathrm{cond}_A^{\mathrm{struct}}(A,x) > \varepsilon^{-1} \quad \textit{and} \quad \mathrm{d}_A^{\mathrm{struct}}(A) = 1.$$

10. Conclusion. Summarizing, depending on the perturbation in use, we face severe differences in the sensitivity of the solution of a linear system. An extreme example is symmetric Toeplitz perturbations. In that case, Theorem 6.3 implies that for the matrices defined in (6.2) the solution $A^{-1}b$ is well conditioned subject to structured componentwise perturbations in the matrix for *all* right-hand sides b. However, for unstructured componentwise perturbations it is ill conditioned for *generic* right-hand side b. This is true when perturbations are restricted to the matrix.

We saw similar examples with a perfectly well-conditioned linear system with respect to componentwise structured perturbations in the matrix *and* the right-hand side, but being arbitrarily ill conditioned with respect to componentwise general (unstructured) perturbations. We presented such examples for all perturbations under investigation except circulants, for which almost sharp estimations for the ratio between the structured and unstructured condition numbers were derived.

So far it seems that componentwise perturbations may produce some quite unexpected and unwanted effects. One reason, as mentioned in the first section, is that zero weights produce certain substructures of the given structure. In particular, the degrees of freedom may be significantly reduced. Then a problem may become well conditioned because not much room is left to produce "bad" perturbations.

This may lead to the conclusion that it is rather unlikely we will find algorithms for the problems and structures under investigation in this paper that are stable with respect to componentwise perturbations. One might even conclude that this seems to be an intrinsic property of componentwise perturbations.

Fortunately, this seems not to be the case. There are other structures for which very fast and accurate algorithms have been developed for the solution of linear systems or matrix inversion and also for other problems such as LU-decomposition and the computation of singular values. For example, those problems can be solved with small componentwise relative backward error for Vandermonde-like or Cauchy matrices [16, section 22], [5, 8, 6]. This is especially remarkable because Vandermonde and Cauchy matrices are reputed for being persistently ill conditioned (with respect to unstructured perturbations; see [3] in Part I).

This is of course a question of exploiting the data, or of developing the "right" algorithms, but also is sometimes facilitated by choosing a clever set of input data. Consider, for example, the problem of matrix inversion, LU-decomposition, or computation of singular values for weakly diagonally dominant M-matrices. Small perturbations in the diagonal elements can cause arbitrarily large perturbations in the result. However, another choice of input data changes the situation [19, 1]: The mentioned problems are well conditioned with respect to the off-diagonal elements and the row sums as input data.

The problem with stability with respect to componentwise (relative) perturbations, structured or not, is that in the course of a computation one single subtraction producing some cancellation may ruin the result in the componentwise backward sense. The backward error of the result of the subtraction is small with respect to uncorrelated perturbations of the operands. However, perturbations are correlated if the operands are the result of previous computations. A typical example can be seen when solving (2.4) with Gaussian elimination.

It seems more and more difficult to design structured solvers for linear systems over the structures in (2.5) being stable with respect to structured componentwise perturbations. Are there such algorithms?

A candidate might be circulant matrices because of their rich algebraic properties. In fact, a normwise stable algorithm already exists [26]. Moreover, in contrast to the other perturbations under investigation, the worst case unstructured componentwise condition number in this case is at most about the square of the structured condition number (for perturbations in the matrix and the right-hand side; see Theorem 7.2).

Finally, there does not seem to be much relation between the distance to singularity and the reciprocal of a condition number in case of componentwise structured perturbations. This is the case for the matrix inverse condition number $\mu_E(A)$ as well as for $\mathrm{cond}_E^{\mathrm{struct}}(A)$, the supremum of $\mathrm{cond}_E^{\mathrm{struct}}(A, x)$ for all x. But maybe an appropriate structured componentwise condition number for that purpose is still to be defined.

Acknowledgments. I wish to thank the students of the summer course on structured perturbations held during my sabbatical at Waseda University, Tokyo. Also my thanks to the two anonymous referees for their thorough reading and most constructive and useful comments.

REFERENCES

[1] A. S. Alfa, J. Xue, and Q. Ye, *Entrywise perturbation theory for diagonally dominant M-matrices with applications*, Numer. Math., 90 (2002), pp. 401–414.
[2] S. G. Bartels and D. J. Higham, *The structured sensitivity of Vandermonde-like systems*, Numer. Math., 62 (1992), pp. 17–33.
[3] F. L. Bauer, *Optimally scaled matrices*, Numer. Math., 5 (1963), pp. 73–87.
[4] F. L. Bauer, *Genauigkeitsfragen bei der Lösung linearer Gleichungssysteme*, Z. Angew. Math. Mech., 46 (1966), pp. 409–421.
[5] A. Björck, *Component-wise perturbation analysis and error bounds for linear least squares solutions*, BIT, 31 (1991), pp. 238–244.
[6] T. Boros, T. Kailath, and V. Olshevsky, *The fast parallel Björck-Pereyra-type algorithm for parallel solution of Cauchy linear equations*, Linear Algebra Appl., 302/303 (1999), pp. 265–293.
[7] P. J. Davis, *Circulant Matrices*, John Wiley, New York, 1979.
[8] J. Demmel, *Accurate singular value decompositions of structured matrices*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 562–580.
[9] J. W. Demmel, *The componentwise distance to the nearest singular matrix*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 10–19.
[10] I. Gohberg and I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 688–704.
[11] G. H. Golub and C. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[12] W. Hager, *Condition estimates*, SIAM J. Sci. Stat. Comput., 5 (1984), pp. 311–316.
[13] D. J. Higham and N. J. Higham, *Backward error and condition of structured linear systems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 162–175.
[14] D. J. Higham and N. J. Higham, *Componentwise perturbation theory for linear systems with multiple right-hand sides*, Linear Algebra Appl., 174 (1992), pp. 111–129.
[15] N. J. Higham, *Experience with a matrix norm estimator*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 804–809.
[16] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
[17] R. A. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
[18] W. M. Kahan, *Numerical linear algebra*, Canad. Math. Bull., 9 (1966), pp. 757–801.
[19] C. O'Cinneide, *Relative-error bounds for the LU decomposition via the GTH algorithm*, Numer. Math., 73 (1996), pp. 507–519.
[20] S. M. Rump, *Structured perturbations and symmetric matrices*, Linear Algebra Appl., 278 (1998), pp. 121–132.
[21] S. M. Rump, *Ill-conditioned matrices are componentwise near to singularity*, SIAM Rev., 41 (1999), pp. 102–112.
[22] S. M. Rump, *Ill-conditionedness need not be componentwise near to ill-posedness for least squares problems*, BIT, 39 (1999), pp. 143–151.
[23] S. M. Rump, *Optimal scaling for p-norms and componentwise distance to singularity*, IMA J. Numer. Anal., 23 (2003), pp. 1–9.
[24] R. Skeel, *Scaling for numerical stability in Gaussian elimination*, J. ACM, 26 (1979), pp. 494–526.
[25] Maple V, *Release* 7.0*, Reference Manual*, Waterloo Maple, Inc., Waterloo, Ontario, Canada, 2001.
[26] P. Y. Yalamov, *On the almost strong stability of the circular deconvolution algorithm*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 358–363.