SUPPORT THEORY FOR PRECONDITIONING*

ERIK G. BOMAN AND BRUCE HENDRICKSON†

Abstract. We present *support theory*, a set of techniques for bounding extreme eigenvalues and condition numbers for matrix pencils. Our intended application of support theory is to enable proving condition number bounds for preconditioners for symmetric, positive definite systems. One key feature sets our approach apart from most other works: We use *support numbers* instead of generalized eigenvalues. Although closely related, we believe support numbers are more convenient to work with algebraically.

This paper provides the theoretical foundation of support theory and describes a set of analytical tools and techniques. For example, we present a new theorem for bounding support numbers (generalized eigenvalues) where the matrices have a known factorization (not necessarily square or triangular). This result generalizes earlier results based on graph theory. We demonstrate the utility of this approach by a simple example: block Jacobi preconditioning on a model problem. Also, our analysis of a new class of preconditioners, *maximum-weight basis preconditioners*, in [7] is based on results contained in this paper.

1. Introduction. The solution of linear systems of equations is at the heart of many computations in science, engineering and other disciplines. Iterative methods are often the most efficient means to solve such systems. In many cases, the matrix describing the system is symmetric, positive definite, in which case the preconditioned conjugate gradients method is the algorithm of choice. The cost of using an iterative method like preconditioned conjugate gradients is the cost of a single iteration (involving the operation of the matrix and of the preconditioner on a vector), multiplied by the number of iterations. Preconditioning is important to keep the number of iterations small. For (preconditioned) conjugate gradients or Chebyshev iteration, the number of iterations is known to be bounded by a constant times the square root of the condition number (after preconditioning). This analysis is based on Chebyshev polynomials and represents a worst-case scenario so in practice the number of iterations may be much smaller, for instance, when the eigenvalues are clustered. Still, the spectral condition number is a useful indicator of the quality of a preconditioner.

The dual goals of finding a preconditioner that is both of good quality and inexpensive to compute and apply often conflict, and the design of effective preconditioners continues to be a very active area of research. Many of the best preconditioners are specialized to individual problems. Some general–purpose preconditioning techniques include variants of incomplete factorizations, approximate inverses, algebraic multilevel methods or domain decomposition. None of these approaches is a panacea, and preconditioning remains as much an art as a science. One of the biggest problems with preconditioning is that convergence analysis is generally limited to simple model problems. For problems with irregular numerical or topological structure, condition number bounds are generally difficult to obtain.

Much work has been done in the field of bounding eigenvalues and condition numbers. In this paper we introduce *support theory* as a mathematical framework to analyze condition numbers of preconditioned systems. Our focus will be on symmet-

*This paper is dedicated to the memory of Fred Howes. This work was funded by the Applied Mathematical Sciences program, U.S. Department of Energy, Office of Energy Research and performed at Sandia National Labs, a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the U.S. DOE under contract number DE-AC-94AL85000. Submitted for journal publication in June 2001; revised October 2002.

†Discrete Algorithms and Math Dept., Sandia National Labs, Albuquerque, NM 87185-1111. {eboman,bah}@cs.sandia.gov

ric positive definite (spd) and symmetric positive semidefinite (spsd) systems. We provide a set of tools, with which one can bound *support numbers* (to be defined in the next section). Support numbers are closely related to generalized eigenvalues. Several authors have earlier derived eigenvalue bound techniques for certain families of preconditioners, in particular incomplete factorizations; see, for example, work by Axelsson [3, 1], Beauwens [4, 5], Magolu [22, 21], and Notay [23, 24]. Although some of the basic tools in the present paper have implicitly been used earlier by others, we believe that our main support theory results (§ 4) are new and different. Also, these results apply to all spsd matrices, not just M–matrices.

Many of our *support theory* techniques can be viewed as an algebraic generalization of recent work on a little–known technique called *support–graph* preconditioning; hence the name. Several core ideas in support–graph theory can be traced back to Beauwens [5], and were rediscovered by Vaidya who used them to study spanning tree preconditioners [30]. The techniques were extended and applied to multilevel methods by Gremban [13], Gremban et al. [14], Reif [26], and Bern et al. [6]. The resulting methods have been applied to the analysis of incomplete Cholesky factorization by Guattery [15] and by Bern et al. [6] and to multilevel diagonal scaling [6]. Unfortunately, support–graph theory is fairly limited in its applicability. It only applies to symmetric, positive semidefinite, diagonally dominant M-matrices (a subset of Stieltjes matrices) and, in some cases, to all spsd diagonally dominant matrices. In contrast, our algebraic support theory applies to all symmetric positive semidefinite matrices. Furthermore, as we discuss in §9, support–graph theory is a special case of our methodology.

In this paper we present a collection of propositions and theorems, some of which are quite elementary and correspond to well known facts in linear algebra. We show that the *support number* used in our analysis is the largest generalized eigenvalue in a certain subspace. More specifically, support numbers are well-defined under rank deficiency and in that sense more robust than generalized eigenvalues. The support number definition is often easier to work with than that of eigenvalues. Our hope is that by reformulating results in terms of support numbers and gathering them into a single paper, this will become a useful resource for future work. This paper forms the foundation for several forthcoming papers by the present authors and collaborators.

In §2 we review the concept of support number and describe how it can be used to bound condition numbers. In §3 we provide a collection of fundamental algebraic properties of support numbers. This is followed in §4 with our most important set of tools and techniques for analyzing preconditioners. In §5 we expand our tool kit to address diagonal matrices (preconditioners). A few basic results about Schur complements are stated in §6. We then present some fairly specialized techniques for analyzing Hadamard products and negative semidefinite matrices in §7 and §8, respectively. We discuss the relationship between this paper and previous work on support–graph theory in §9. In §10 we demonstrate how our support tools can be used to analyze a simple, well–known preconditioner, namely block Jacobi preconditioning. In §11 we propose a generalization of support numbers that may be useful for analyzing nonsymmetric or indefinite systems.

2. Support Theory Definitions and Concepts. The main goal of the *support theory* in this paper is to provide techniques to bound the generalized eigenvalues and condition number for a matrix pencil (A, B). Think of B as being a preconditioner for A. We only study real matrices in this paper, but most of the results carry over to the complex case (substitute Hermitian for symmetric). If both A and B are spd, then the convergence of many preconditioned iterative methods (and specifically, preconditioned conjugate gradients) depends on the condition number of the preconditioned operator $B^{-1/2}AB^{-1/2}$. We define the generalized (spectral) condition number by

$$\kappa(A,B) \equiv \kappa(B^{-1/2}AB^{-1/2}) = \frac{\lambda_{\max}(B^{-1/2}AB^{-1/2})}{\lambda_{\min}(B^{-1/2}AB^{-1/2})} = \frac{\lambda_{\max}(A,B)}{\lambda_{\min}(A,B)},$$

where $\lambda(A)$ denotes an eigenvalue of A while $\lambda(A,B)$ denotes a generalized eigenvalue for (A,B).

The central concept in support theory is the *support number* of a matrix pair (A,B), sometimes simply called the *support*. We remark that the definition we use is slightly different from the one in [6] and [13], but only when A or B is indefinite.

DEFINITION 2.1. *The support number of* (A,B) *is defined by*

$$\sigma(A,B) = \min\left\{t \in \mathbb{R} \,|\, x^T(\tau B - A)x \geq 0 \textit{ for all } x \textit{ and for all } \tau \geq t\right\}$$

For some pencils (A,B), there is no such t and we define the support number $\sigma(A,B)$ to be ∞. Similarly, if $\tau B - A$ is positive semidefinite (psd) for all τ we define the support number to be $-\infty$. (This cannot happen if B is psd.) The definition above does not require A and B to be symmetric. However, symmetric matrices will be the main focus of this paper. We remark that by choosing $B = I$, the techniques in this paper can be used to bound the largest eigenvalue and spectral condition number of A. For symmetric matrices, the support is closely related to a generalized eigenvalue. Axelsson [1, Corollary 2.1] showed the following result:

LEMMA 2.2. *Suppose A is spsd and B is spd. For any τ such that $\lambda_{\min}(\tau B - A) \geq 0$ we have*

$$\lambda_{\max}(B^{-1}A) \leq \tau.$$

In other words, an upper bound on the support number $\sigma(A,B)$ is also a bound on the generalized eigenvalue $\lambda_{\max}(A,B) \equiv \max\{\lambda \,|\, Ax = \lambda Bx, x \neq 0\}$. (More general versions of this lemma can be found as Theorem 3.16 and Theorem 10.1 in [2].) Next, we elaborate on this important result and include the case where B is spsd and may be singular. The theorem below is an extension of Gremban's Support Lemma [13, Lemma 4.4] and similar lemmas in [6].

THEOREM 2.3. *Let A and B be symmetric matrices.*

1. *If B is spd, then $\sigma(A,B) = \lambda_{\max}(A,B)$.*
2. *If B is spsd and $Null(B) \subseteq Null(A)$, then*

$$\sigma(A,B) = \max\left\{\lambda \,|\, Ax = \lambda Bx,\ Bx \neq 0\right\},$$

or equivalently,

$$\sigma(A,B) = \lambda_{\max}(Z^T AZ, Z^T BZ),$$

where Z is such that the columns of Z span the range of B.

3. *If B is not spsd, then $\sigma(A,B)$ is infinite.*

Proof. The first part follows from the variational characterization

$$\lambda_{\max}(A,B) = \tau \max_{x \neq 0} \frac{x^T Ax}{x^T(\tau B)x},$$

where B is assumed to be spd. For any τ such that $x^T(\tau B - A)x \geq 0$ the condition above implies that $\lambda_{\max}(A, B) \leq \tau$. Equality holds when τ is the largest generalized eigenvalue and x is the corresponding eigenvector. To show the second part, use the same argument but restrict x to the space where $Bx \neq 0$. For the third part, let x be a vector such that $x^T Bx < 0$. Then $x^T(\tau B - A)x < 0$ for any sufficiently large τ, so the support is unbounded (infinite). ∎

The support number can therefore be interpreted as an extension of generalized eigenvalues that is robust under rank–deficiency. When both matrices are spd, then the (generalized) condition number is the ratio of the largest to smallest generalized eigenvalues.

PROPOSITION 2.4. *When A and B are both spd, the generalized condition number $\kappa(A, B)$ satisfies $\kappa(A, B) = \sigma(A, B)\sigma(B, A)$.*

Proof. By Theorem 2.3, $\sigma(A, B) = \lambda_{\max}(A, B)$, and thus $\sigma(B, A) = 1/\lambda_{\min}(A, B)$. ∎

The condition number is unbounded (infinite) if either A or B is rank deficient, but $\sigma(A, B)\sigma(B, A)$ may still be finite, and can therefore be viewed as a more robust generalization of the condition number. In practice one should be cautious about using a singular matrix as a preconditioner.

Our technique to bound the support of (A, B) is to break the matrices up into pieces which are in some sense *simpler*. In the sections that follow, *simple* can mean different things, for example, sparse and of low rank. We will rely heavily upon the following splitting principle, a slight variation of Lemma 4.7 in [13].

PROPOSITION 2.5 (Splitting). *Split A and B into $A = A_1 + A_2 + \ldots + A_q$ and $B = B_1 + B_2 + \ldots + B_q$. If all B_i are psd, then $\sigma(A, B) \leq \max_i \sigma(A_i, B_i)$.*

The key to proving good support bounds is to find good splittings of A and B. (We remark that “multisplitting” might be a more appropriate term since the matrices can be split into several parts.) In our framework, each B_i must be positive semidefinite while there is no restriction of the definiteness of A_i. However, in practice we usually employ splittings where all the A_i are also spsd.

An important observation for using support theory is that one may use different splittings of A and B when proving bounds on $\sigma(A, B)$ and $\sigma(B, A)$. Different splittings may give quite different bounds on the condition number, so identifying good splittings is crucial.

In some applications, there is a natural splitting of the form $A = \Sigma_i A_i$. For example, in finite element analysis, A could correspond to the global mass or stiffness matrix while each A_i corresponds to an element matrix. Analysis by splitting into element matrices is a technique used by several authors and goes at least back to the early 1970s. Irons and Treharne [19] described the splitting theorem in the context of finite elements as “a familiar but undervalued theorem” and advocated that it should be taught in finite element courses. More recently, Wathen [32, 20] used the splitting property to prove upper and lower eigenvalue bounds for element-by-element preconditioners. Similar splittings are also used in domain decomposition [28]. We do not discuss finite elements any further here because it is outside the scope of the present paper.

3. Fundamental properties of support numbers. We state some fundamental properties of the support number and skip the simplest proofs.

PROPOSITION 3.1. *When A is psd and $\alpha \neq 0$, then $\sigma(\alpha A, A) = \alpha$.*

PROPOSITION 3.2. *Let B be psd and $\alpha > 0$. Then $\sigma(\alpha A, B) = \alpha\sigma(A, B)$ and $\sigma(A, \alpha B) = \alpha^{-1}\sigma(A, B)$.*

Proposition 3.3. *If B is psd, then*

$$\sigma(A+C,B) \leq \sigma(A,B)+\sigma(C,B).$$

Proposition 3.4. *If B and C are psd, then*

$$\sigma(A,B+C) \leq \frac{\sigma(A,B)\sigma(A,C)}{\sigma(A,B)+\sigma(A,C)} \leq \frac{1}{2}\max\left\{\sigma(A,B),\sigma(A,C)\right\}.$$

Proof. Using Propositions 3.2 and 2.5, we have that $\sigma(A,B+C) = \sigma(\frac{1}{2}A+\frac{1}{2}A,B+C) \leq \frac{1}{2}\max\{\sigma(A,B),\sigma(A,C)\}$, which proves the weaker bound. The stronger bound is derived similarly by a splitting $A = \alpha A + (1-\alpha)A$ for α such that $\alpha\sigma(A,B) = (1-\alpha)\sigma(A,C)$. □

Proposition 3.5. *If B and C are psd, then*

$$\sigma(A,B) \leq \sigma(A+C,B).$$

When A and $B-C$ are also psd, then

$$\sigma(A,B) \leq \sigma(A,B-C).$$

The triangle inequality holds for support numbers.

Proposition 3.6. *Suppose that B and C are psd. Then*

$$\sigma(A,C) \leq \sigma(A,B)\sigma(B,C)$$

Note that none of the lemmas so far requires symmetry. The support number essentially ignores the nonsymmetric part of the matrices, as shown below.

Proposition 3.7. *Suppose that B is psd. Then $\sigma(A,B) = \sigma(A^T,B) = \sigma(A,B^T)$, and hence*

$$\sigma(A,B) = \sigma(Sym(A),Sym(B)),$$

where $Sym(X) \equiv \frac{1}{2}(X+X^T)$ denotes the symmetric part of X.

Proof. The result follows from Definition 2.1 and the fact that $x^TAx = x^TA^Tx$ for any square (not necessarily symmetric) matrix. □

Corollary 3.8. *Suppose that A, B and C are spsd. Then*

$$\sigma(AC,B) = \sigma(CA,B) \text{ and } \sigma(A,BC) = \sigma(A,CB)$$

.

Proof. By using Proposition 3.7 and the symmetry of A and C, we have that $\sigma(AC,B) = \sigma((AC)^T,B) = \sigma(C^TA^T,B) = \sigma(CA,B)$. Similarly for the second part. □

We will use a well–known eigenvalue result, see for example Corollary 3.14 in [2].

Lemma 3.9. *Let A and B be spsd matrices of the same order. Then*

$$\lambda_{\max}(AB) \leq \lambda_{\max}(A)\lambda_{\max}(B)$$

Using this Lemma and Theorem 2.3, we get the following results for symmetric matrices.

PROPOSITION 3.10. *When A, B and C are all spsd, then*

$$\sigma(AC, B) \leq \lambda_{\max}(C)\sigma(A, B).$$

Proof. Suppose B is nonsingular. Then $\sigma(AC, B) = \lambda_{\max}(B^{-1}AC) \leq \lambda_{\max}(B^{-1}A)\lambda_{\max}(C) \leq \lambda_{\max}(C)\sigma(A, B)$. If B is singular, the same argument holds in a subspace (the range of B). □

The next proposition extends lemmas that were used by Gremban [13] and by Bern et al. [6] to partially factor a matrix and preconditioner while maintaining a bound on the support number.

PROPOSITION 3.11. *Let $B \in \mathbb{R}^{n \times n}$ be spsd. Then for any $G \in \mathbb{R}^{n \times p}$,*

$$\sigma(G^T AG, G^T BG) \leq \sigma(A, B),$$

and if $Null(G^T) \subseteq Null(A)$ and $Null(G^T) \subseteq Null(B)$ then

$$\sigma(G^T AG, G^T BG) = \sigma(A, B).$$

Proof. Let $\tau = \sigma(A, B)$. Then $x^T(\tau B - A)x \geq 0 \quad \forall x$. For any $y \in \mathbb{R}^p$, let $x = Gy$. Then $y^T G^T(\tau B - A)Gy \geq 0$, and it follows that $\sigma(G^T AG, G^T BG) \leq \tau$. This proves the first part of the proposition. For the second part, note that $Null(G^T) = Range(G)^\perp$. Any vector $x \in \mathbb{R}^n$ can be split into two parts, $x = \hat{x} + \tilde{x}$, where $\hat{x} \in Range(G)$ and $\tilde{x} \in Null(G^T)$. Suppose $Null(G^T) \subseteq Null(A)$ and $Null(G^T) \subseteq Null(B)$. It follows that $x^T(\tau B - A)x = \hat{x}^T(\tau B - A)\hat{x}$, and since $\hat{x} \in Range(G)$ there exists y such that $\hat{x} = Gy$. □

One can multiply both arguments with an spd matrix without changing the support.

PROPOSITION 3.12. *Suppose that A and B are spsd and C is spd. Then*

$$\sigma(A, B) = \sigma(AC, BC) = \sigma(CA, CB)$$

Proof. Let L be a Cholesky factor of C, i.e., $C = LL^T$. Then $\sigma(AC, BC) = \sigma(L^T AL, L^T BL) = \sigma(A, B)$, where the last equality relies on Proposition 3.11. □

The corollary below follows by choosing first $C = A^{-1}$, then $C = B^{-1}$.

COROLLARY 3.13. *Suppose that A and B are spd. Then $\sigma(A, B) = \sigma(B^{-1}, A^{-1})$.*

The next result is a slight generalization of Lemma 3.3 in [6], which was used to prove a bound on modified incomplete Cholesky preconditioners.

PROPOSITION 3.14. *When A and B are psd then*

$$\sigma(A, B) \leq \frac{1}{1 - \sigma(A - B, A)}.$$

Proof. Let $\tau' = \sigma(A - B, A)$. Observe that $\tau' \leq 1$ because $A - (A - B) = B$ is psd. Also, $\tau' A - (A - B)$ is psd by Definition 2.1. We have that $\tau' A - (A - B) = B - (1 - \tau')A$, hence $\frac{1}{1-\tau'}B - A$ is also psd because $\tau' \leq 1$. Consequently, $\sigma(A, B) \leq \frac{1}{1-\tau'}$. □

The following proposition may be useful when A and B are spsd but not diagonally dominant since there are more efficient algorithms for solving diagonally dominant systems. By choosing C to be diagonal with sufficiently large positive elements, $A+C$ and $B+C$ can be made diagonally dominant.

PROPOSITION 3.15. *Suppose A and B are psd. Then for any psd C and $\alpha > 0$ such that $\alpha\sigma(A+C, B+\alpha C) \leq 1$, then*

$$\sigma(A, B) \leq \sigma(A+C, B+\alpha C).$$

Proof. For any $\alpha > 0$ there exists a τ such that $\tau(B+\alpha C) - (A+C)$ is spsd. Consequently, $\tau B - A$ is spsd when $(1-\tau\alpha)C$ is spsd. By assumption, $\tau\alpha \leq 1$, so the desired result follows. ☐

When A and B have block diagonal structure, the support number can be computed by looking at the blocks independently and taking the maximum. This is a special case of splitting where equality holds.

PROPOSITION 3.16. *Suppose B is psd and A, B are of the form*

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22} \end{pmatrix}.$$

Then $\sigma(A, B) = \max\{\sigma(A_{11}, B_{11}), \sigma(A_{22}, B_{22})\}$.

In some situations it is helpful to obtain a support bound by expanding the matrices into a higher dimension. The following proposition explains how.

PROPOSITION 3.17. *Let A_{11}, B_{11} denote principal submatrices of A and B, respectively. Then $\sigma(A_{11}, B_{11}) \leq \sigma(A, B)$.*

Proof. Let $\tau = \sigma(A, B)$. Then $\tau B - A$ is positive semidefinite. Any principal submatrix of $\tau B - A$ is also positive semidefinite; in particular, $\tau B_{11} - A_{11}$. ☐

4. Main support results. This section contains our main results. Recall from Lemma 2.5 that we want to break A and B into sums of *simple* pieces. A key kind of simplicity that we will exploit is to have the pieces be of low rank. We can exploit the fact that symmetric rank–1 and rank–2 matrices have spectra that are simple to express.

LEMMA 4.1. *Let $A = uu^T$. Then all eigenvalues of A are zero except $\lambda_1(A) = u^Tu$. Furthermore, if B is invertible (nonsingular) then all generalized eigenvalues of (A, B) are zero except $\lambda = u^TB^{-1}u$.*

LEMMA 4.2. *Let $A = uv^T + vu^T$. Then all the eigenvalues of A are zero except $\lambda_{1,2}(A) = \pm\|u\|_2\|v\|_2 + u^Tv$.*

Lemma 4.1 gives us a formula for the support for a symmetric rank–1 matrix A.

PROPOSITION 4.3. *Let $A = uu^T$ and let B be spd. Then*

$$\sigma(A, B) = u^TB^{-1}u.$$

Proof. From Theorem 2.3 we have that $\sigma(A, B) = \lambda_{\max}(A, B)$. By Lemma 4.1, all the eigenvalues $\lambda(A, B)$ are zero except one, which is $u^TB^{-1}u$. Since B is spd, $u^TB^{-1}u > 0$ for any u, so $\lambda_{\max}(A, B) = u^TB^{-1}u$. ☐

Next we show a more general result that includes the case where B is semidefinite and does not have full rank.

THEOREM 4.4 (Rank–1 Support Theorem). *Suppose $u \in \mathbb{R}^n$ is in the range of $V \in \mathbb{R}^{n\times k}$. Then*

$$\sigma(uu^T, VV^T) = \min_w w^Tw \text{ subject to } Vw = u.$$

Proof. Let w be a vector that satisfies $Vw = u$. By applying Proposition 3.11, we get

$$\sigma(uu^T, VV^T) = \sigma(Vww^TV^T, VV^T) \leq \sigma(ww^T, I) = w^Tw$$

Next we prove that there exists a w such that equality holds. The smallest norm solution to $Vw = u$ is given by $w = V^+u$, where V^+ is the Moore–Penrose pseudoinverse of V [12, p. 243]. We have that $\sigma(uu^T, VV^T) = \lambda_{\max}(V^+uu^T(V^+)^T) = \|V^+u\|_2^2$. ∎

We remark that any w that satisfies $Vw = u$ gives an upper bound on $\sigma(uu^T, VV^T)$. Further observe that when V has full column rank, then there is a unique w such that $Vw = u$. The theorem above can also be restated in terms of the pseudoinverse, that is, $\sigma(uu^T, VV^T) = \|V^+u\|_2^2$.

Note that all spsd matrices can be constructed as a sum of symmetric outer products like those in the theorem. For instance, the Cholesky decomposition (in outer-product form) provides such a splitting. However, there are many alternatives and the Cholesky decomposition may not be the best choice for proving bounds or building preconditioners.

In the special case where each column of U and V have only two nonzero entries and these entries have the same magnitude, this proposition reduces to the Congestion–Dilation Lemma discussed in §9.1. The Congestion–Dilation Lemma is based on a specific graph interpretation that we will examine in §9, and is the cornerstone of support–graph theory [13, 6]. In support–graph theory, the vector u with its two nonzeros in locations i and j represents an edge between vertices i and j, and the set of columns of V correspond to a path (a sequence of edges) between the same vertices. Unfortunately, only a very limited class of matrices can be represented as sums of outer products of these specialized vectors. Specifically, as discussed in §9, if the two values are of the opposite sign, then all symmetric, diagonally dominant, positive semidefinite, M-matrices can be generated. And if values of the same sign are included then the class grows to be all symmetric, diagonally dominant, positive semidefinite matrices. Support–graph theory is limited to these classes of matrices. But with a general u, the much more important class of symmetric, positive semidefinite matrices can be addressed.

We next state the higher–rank generalization of Theorem 4.4.

THEOREM 4.5 (Symmetric Product Support). *Suppose $U \in \mathbb{R}^{n\times k}$ is in the range of $V \in \mathbb{R}^{n\times p}$. Then*

$$\sigma(UU^T, VV^T) = \min_W \|W\|_2^2 \text{ subject to } VW = U.$$

Proof. Let W satisfy $VW = U$. Then

$$\begin{aligned}\sigma(UU^T, VV^T) &= \sigma(VWW^TV^T, VV^T) \leq \sigma(WW^T, I) \\ &= \lambda_{\max}(WW^T) = \|W\|_2^2.\end{aligned}$$

As in the proof of Theorem 4.4, one can show that equality is achieved for $W = V^+U$. ∎

We will often use this Theorem as a tool to obtain an upper bound on $\sigma(UU^T, VV^T)$. Note that any W for which $VW = U$ provides an upper bound on the support number. One special case of interest is when the columns of U are a subset of the columns of V (or vice versa).

COROLLARY 4.6. *Suppose the columns of U are a subset of the columns of V. Then $\sigma(UU^T, VV^T) \leq 1$.*

The result above follows by letting W be an appropriate subset of the identity matrix, so $\|W\|_2^2 \leq 1$. Alternatively, it is easy to show that $VV^T - UU^T$ is spsd, which also gives a bound of one for the support number.

The following theorem is a slight generalization of Theorem 4.5.

THEOREM 4.7. *Suppose $U \in \mathbb{R}^{n\times k}$ is in the range of $V \in \mathbb{R}^{n\times p}$ and let $D \in \mathbb{R}^{k\times k}$ be symmetric. Then*

$$\sigma(UDU^T, VV^T) \leq \lambda_{\max}(WDW^T) \leq \lambda_{\max}(D)\|W\|_2^2$$

for all W such that $VW = U$.

Proof. Let W satisfy $VW = U$. Then

$$\sigma(UDU^T, VV^T) = \sigma(VWDW^TV^T, VV^T) \leq \sigma(WDW^T, I) = \lambda_{\max}(WDW^T),$$

which proves the first part. The second part follows from $\lambda_{\max}(WDW^T) = \lambda_{\max}(DW^TW) \leq \lambda_{\max}(D)\lambda_{\max}(W^TW) = \lambda_{\max}(D)\|W\|_2^2$. ☐

Recall that the support number may be negative.

COROLLARY 4.8. *Suppose $U \in \mathbb{R}^{n\times k}$ is in the range of $V \in \mathbb{R}^{n\times p}$ and let D be a block diagonal matrix in $\mathbb{R}^{k\times k}$, where the blocks are either of the type ± 1 or $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Then $\sigma(UDU^T, VV^T) \leq \|W\|_2^2$ for all W such that $VW = U$.*

Proof. The eigenvalues of D can only take on two different values: 1 or -1. Hence $\lambda_{\max}(D) \leq 1$, and the result follows from Theorem 4.7. ☐

We remark that any symmetric matrix (possibly indefinite) has a decomposition of the type UDU^T where U is square and lower triangular and D is as described in the corollary above. However, this may not be the best way to apply the corollary.

Further note that $\|W\|_2^2$ may be expensive to compute. Nonetheless, as is well known, the 2–norm can be bounded by easy-to-compute quantities.

LEMMA 4.9. *For any matrix W, we have that*

(i) $\|W\|_2^2 \leq \|W\|_1\|W\|_\infty = (\max_j \sum_i |W_{ij}|)\left(\max_i \sum_j |W_{ij}|\right)$,

(ii) $\|W\|_2^2 \leq \|W\|_F^2 = \sum_{i,j} W_{ij}^2$.

Most of the preceding set of results have involved symmetric outer products to construct low rank matrices. We now extend the Rank–1 Support Theorem to the rank-2 case.

THEOREM 4.10. *Suppose $u, v \in \mathbb{R}^n$ are in the range of $Y \in \mathbb{R}^{n\times k}$. Then*

$$\sigma(uv^T + vu^T, YY^T) \leq \|w\|_2\|\hat{w}\|_2 + w^T\hat{w}$$

for any w and $\hat{w}$ such that $Yw = u$ and $Y\hat{w} = v$.

Proof.

$$\begin{aligned}
\sigma(uv^T + vu^T, YY^T) &= \sigma(Y(w\hat{w}^T + \hat{w}w^T)Y^T, YY^T) \\
&\leq \sigma(w\hat{w}^T + \hat{w}w^T, I) \\
&= \lambda_{\max}(w\hat{w}^T + \hat{w}w^T) \\
&= \|w\|_2\|\hat{w}\|_2 + w^T\hat{w} \quad \text{by Lemma 4.2}
\end{aligned}$$

☐

Corollary 4.11. *Suppose $u, v \in \mathbb{R}^n$ are in the range of $Y \in \mathbb{R}^{n\times k}$. Then*

$$\sigma(uv^T + vu^T, YY^T) \le 2\|w\|_2\|\hat{w}\|_2 \le \|w\|_2^2 + \|\hat{w}\|_2^2$$

for all $w, \hat{w}$ such that $Yw = u$, $Y\hat{w} = v$.

Proof. The result follows from Theorem 4.10 and the Cauchy-Schwarz inequality. □

We can extend Theorem 4.10 to the case where U and V are matrices.

Theorem 4.12. *Suppose $U, V \in \mathbb{R}^{n\times p}$ are in the range of $Y \in \mathbb{R}^{n\times k}$. Then*

$$\sigma(UV^T + VU^T, YY^T) \le \lambda_{\max}(W\hat{W}^T + \hat{W}W^T) \le 2\|W\|_2\|\hat{W}\|_2$$

for any W and $\hat{W}$ such that $YW = U$ and $Y\hat{W} = V$.

We omit the proof because it is essentially a combination of the proofs of Theorem 4.10 and of Corollary 4.11.

5. Diagonal support. In §4 we described tools for bounding support numbers when the pieces involved have low rank. Another kind of simple structure we can exploit occurs when one of the matrices is diagonal. Any matrix can be supported by a positive diagonal matrix. We remark that computing the exact support $\sigma(A, B)$ when B is diagonal is not much easier than for a general spd B and requires the computation of an extremal eigenvalue.

Fortunately, we will see that it is easy to obtain a bound. We need the following well–known fact, which is easily derived from Gerschgorin's Theorem.

Lemma 5.1. *If A is symmetric, weakly (strictly) diagonally dominant, and has non-negative diagonal entries, then A is spsd (spd).*

Using the above lemma, one way to bound $\sigma(A, B)$ is to find τ such that $\tau B - A$ is diagonally dominant with positive diagonal entries. Unfortunately, this strategy only works for certain B and further, computing the optimal value of τ may require the solution of a linear program. However, when B is diagonal we can obtain a bound as follows.

Theorem 5.2. *Suppose A is symmetric (not necessarily spd) and B is diagonal with $b_{ii} \ge 0$ for all i. Assume that $W = \{w_{ij}\}$ satisfies $w_{ij} > 0$ and $w_{ij} = 1/w_{ji}$ for all i and j, and that $b_{ii} = 0$ only if $a_{ii} + \sum_{j\ne i} w_{ij}|a_{ij}| \le 0$. Then*

$$\sigma(A, B) \le \max_i \left\{ \frac{a_{ii} + \sum_{j\ne i} w_{ij}|a_{ij}|}{b_{ii}} \right\},\ b_{ii} \ne 0.$$

Proof. We will describe how to find a spsd matrix $\hat{A}$ such that $D \equiv A + \hat{A}$ is diagonal. From Proposition 3.5 it follows that $\sigma(A, B) \le \sigma(A + \hat{A}, B) = \sigma(D, B)$. Let $\hat{A} = \sum_{ij} \hat{A}_{ij}$, where $\hat{A}_{ij}$ is chosen to cancel out the off–diagonal element a_{ij}. Specifically, $\hat{A}_{ij}$ is zero except in rows and columns i and j where it is

$$\begin{pmatrix} |a_{ij}|/w_{ij} & -a_{ij} \\ -a_{ij} & |a_{ij}|w_{ij} \end{pmatrix} = \begin{pmatrix} |a_{ij}|w_{ji} & -a_{ij} \\ -a_{ij} & |a_{ij}|w_{ij} \end{pmatrix}.$$

Consequently, $D = A + \hat{A}$ is diagonal. By simple algebra, $d_{ii} = a_{ii} + \sum_{j\ne i} w_{ij}|a_{ij}|$, and the desired result follows. □

By setting $B = I$, we obtain an interesting eigenvalue bound.

Corollary 5.3. *Let A be a symmetric matrix (not necessarily spd). Then for any positive matrix W such that $w_{ij} = 1/w_{ji}$ for all i and j,*

$$\lambda_{\max}(A) \le \max_i \left\{ a_{ii} + \sum_{j \neq i} w_{ij} |a_{ij}| \right\}.$$

By setting all the w_{ij} values to be 1, we get a different special case.

Corollary 5.4. *Suppose A is symmetric (not necessarily spd), $B \ge 0$ is diagonal, and $b_{ii} = 0$ only if $a_{ii} + \sum_{j \neq i} |a_{ij}| \le 0$. Then*

$$\sigma(A, B) \le \max_i \left\{ \frac{a_{ii} + \sum_{j \neq i} |a_{ij}|}{b_{ii}} \right\}, \; b_{ii} \neq 0.$$

When $B = I$ and all the w_{ij} values are 1, then each of these corollaries reduces to Gerschgorin's well–known bound on the maximal eigenvalue. Furthermore, Theorem 5.2 contains as a special case the scaled Gerschgorin bound obtained by diagonal scaling of A, that is, the Gerschgorin eigenvalue bound for SAS^{-1} where S is diagonal.

How can we choose W to improve the bound? Computing the optimal W is difficult and could even be more expensive that computing $\lambda_{\max}(A)$ directly. Intuitively, we want to choose w_{ij} small when row i has a large (absolute) row sum, i.e., when $a_{ii} + \sum_{k \neq i} |a_{ik}|$ is large. One possible such strategy is to let

$$w_{ij} = \frac{a_{jj} + \sum_{k \neq j} |a_{jk}| - a_0}{a_{ii} + \sum_{k \neq i} |a_{ik}| - a_0},$$

where $a_0 = \min_i a_{ii}$. (Because we subtract a_0, the bound is invariant under shifting of the eigenvalues.) We remark that the proposed bound is often, but not always, better than the Gerschgorin bound. For example, for

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 6 & 3 \\ 1 & 3 & 9 \end{pmatrix},$$

the Gerschgorin bound is 13 but our new bound is 11.7. The largest eigenvalue is 11.3.

An alternative approach is to start out with $w_{ij} \equiv 1$ and then iteratively pick an entry w_{ij} to adjust. Keeping all other coefficients fixed, one can compute a new value for w_{ij} that tightens the eigenvalue bound.

We note that tighter bounds may be obtained by using matrices with nonzeros in more than two rows (columns) to cancel out positive off–diagonals. Such a strategy requires finding cliques in the graph of the matrix. We do not examine this option any further here.

A technique used by several previous authors for preconditioning diagonally dominant matrices, is to first subtract a diagonal matrix such that the remaining part is semidefinite and rank deficient. Then one preconditions the semidefinite part using support theory and adds back the diagonal part. The following lemma is used. (Note that in this and the subsequent lemmas, D is a general spsd matrix, but for current purposes we are interested in the case where D is diagonal.)

LEMMA 5.5. *If A is symmetric and B and D are spsd, then* $\sigma(A+D, B+D) \leq \max\{\sigma(A,B), 1\}$.

Clearly, the diagonal elements are not fully exploited in this approach. Basically, B supports A while D only supports itself. Going to the other extreme, we could let D support both A and D, which yields $\sigma(A+D, B+D) \leq \sigma(A+D, D) \leq \sigma(A, D)+1$. This method is also unsatisfactory because B is not utilized at all. A better approach is to let parts of D support A and parts of it support itself. From this idea we obtain the following result.

PROPOSITION 5.6. *If A is symmetric and B and D are spsd, then*

$$\sigma(A+D, B+D) \leq \frac{1+\sigma(A,D)}{1+\sigma(A,D)/\sigma(A,B)}.$$

Proof. We use the splitting $B + D = (B + \alpha D) + (1 - \alpha)D$, and by applying Proposition 2.5 and Proposition 3.4 we find that

$$\begin{aligned}\sigma(A+D, B+D) &\leq \max\{\sigma(A, B+\alpha D), \sigma(D, (1-\alpha)D\} \\ &\leq \max\left\{\frac{\sigma(A,B)+\sigma(A,D)}{\alpha\sigma(A,B)+\sigma(A,D)}, \frac{1}{1-\alpha}\right\}\end{aligned}$$

for any α such that $0 < \alpha < 1$. We want the tightest possible bound, which occurs when the two arguments in max are equal. Hence we solve the following equation for α,

$$(1-\alpha)(\sigma(A,B)+\sigma(A,D)) = \alpha\sigma(A,B)+\sigma(A,D),$$

which has the solution

$$\alpha = \frac{\sigma(A,D)(1+\sigma(A,B)}{\sigma(A,B)(1+\sigma(A,D))}.$$

The desired support bound is $1/(1-\alpha)$, which after some algebra is shown to equal

$$\frac{1}{1-\alpha} = \frac{\sigma(A,B)(1+\sigma(A,D))}{\sigma(A,B)+\sigma(A,D)} = \frac{1+\sigma(A,D)}{1+\sigma(A,D)/\sigma(A,B)}.$$

□

6. Schur complement support. Another special matrix structure that commonly arises in practice is the Schur complement — the remaining portion of a matrix after a subset of rows and columns have been factored (by Gaussian elimination). This section contains tools to address this special matrix structure.

A matrix can be supported in a "higher–dimensional space" using the Schur complement.

PROPOSITION 6.1. *Let A and B be spsd and of the form*

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix},$$

where B_{22} is nonsingular. Then $\sigma(A,B) = \sigma(A_{11}, B_{11} - B_{12}B_{22}^{-1}B_{12}^T)$.

Proof. Let $G^T = \begin{pmatrix} I & -B_{12}B_{22}^{-1} \\ 0 & I \end{pmatrix}$, which is always nonsingular. Let S denote the Schur complement $B_{11} - B_{12}B_{22}^{-1}B_{12}^T$. It is easy to verify that $G^TAG = A$ and

$G^T BG = \begin{pmatrix} S & 0 \\ 0 & B_{22} \end{pmatrix}$. By Proposition 3.11, $\sigma(G^T AG, G^T BG) = \sigma(A, B)$. Since the lower right block of A is zero, the support number is determined by the upper left blocks of the block diagonal matrix pencil $(G^T AG, G^T BG)$ and we have that $\sigma(G^T AG, G^T BG) = \sigma(A_{11}, S)$. $\square$

A useful special case of the preceding result is as follows.

COROLLARY 6.2. *Suppose A and B are spsd and of the form*

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha A_{11} + VV^T & \beta V \\ \beta V^T & \beta^2 I \end{pmatrix},$$

where $\alpha > 0$, $\beta \neq 0$, and V is any matrix of appropriate dimensions. Then

$$\sigma(A, B) = 1/\alpha.$$

Proof. Proposition 6.1 yields

$$\begin{aligned} \sigma(A, B) &= \sigma(A_{11}, \alpha A_{11} + VV^T - \beta V \beta^{-2} \beta V^T) \\ &= \sigma(A_{11}, \alpha A_{11}) = 1/\alpha. \end{aligned}$$

$\square$

This corollary contains the Clique-Star Lemma from [13, 6] as a special case, where $\alpha = 1/k$, $\beta = 1$, $A_{11} = kI - ee^T$, $V = e$, and e is a vector of all ones. The Clique-Star Lemma was used by Gremban [13] in the analysis of multilevel support–graph preconditioners (see also [6]).

7. Hadamard product support. In this section we restate some known results about eigenvalues and Hadamard products in terms of support numbers. The *Hadamard* product is the elementwise matrix product, that is, if $C = A \circ B$ then $c_{ij} = a_{ij} b_{ij}$ for all i, j. Schur [27] proved several properties of the Hadamard product, including the important results below.

LEMMA 7.1. *If A and C are both spsd, then*

$$\lambda_{\min}(A)\lambda_{\min}(C) \leq \lambda_i(A \circ C) \leq \lambda_{\max}(A)\lambda_{\max}(C) \text{ for all } i.$$

COROLLARY 7.2. *If A and C are both spsd, then $A \circ C$ is also spsd.*

The next proposition follows directly from Schur's results.

PROPOSITION 7.3. *If $A, B,$ and C are spsd, then*

$$\sigma(A \circ C, B \circ C) \leq \sigma(A, B).$$

Proof. Let $\tau = \sigma(A, B)$, so $\tau B - A$ is spsd. By Corollary 7.2, $(\tau B - A) \circ C = \tau(B \circ C) - A \circ C$ is also spsd for any spsd C. $\square$

Restating a variation of Schur's result [25, Lemma 2.1] in support theory notation, we get the proposition below.

PROPOSITION 7.4. *Suppose A is spsd and C is symmetric. Let D_A denote the diagonal matrix with the same diagonal as A. Then*

$$\sigma(A \circ C, D_A) \leq \lambda_{\max}(C).$$

If C is spd, then we also have

$$\sigma(D_A, A \circ C) \leq \frac{1}{\lambda_{\min}(C)}.$$

Fiedler and Markham [11] proved the following result.

PROPOSITION 7.5. *Suppose A is spsd and C is spd. Then*

$$\sigma(A, A \circ C) \leq e^T C^{-1} e,$$

where e is the all ones vector.

This result may be useful in our context when, for example, the preconditioner B has a sparsity pattern that is a subset of the nonzeros of A, so there exists a C such that $A \circ C = B$. As a simple example, consider the case when B (and hence also C) is diagonal. Then $\sigma(A, B) \leq \sum_i (a_{ii}/b_{ii})$. Observe that when $B = I$ this bound reduces to the well-known trace bound, $\lambda_{\max}(A) \leq tr(A) = \sum_i a_{ii}$.

Recently, several extensions to the Fiedler–Markham result (Proposition 7.5) have been developed [25, 17]. These extensions hold when C is either positive definite or *conditionally* positive definite, that is, positive definite in a subspace.

8. Supporting negative semidefinite parts. It is trivial to support a negative semidefinite matrix.

PROPOSITION 8.1. *If A is negative semidefinite, then $\sigma(A, 0) = -\infty$. Furthermore, $\sigma(A, B) \leq 0$ for any positive semidefinite B.*

This proposition gives us two preconditioning strategies when applied to a part A_i of a matrix A. First, any negative semidefinite part of A can be ignored (preconditioned by 0). We remark that a better condition number bound may possibly be obtained by utilizing the negative semidefinite part. Second, we can add any positive semidefinite matrix B_i to a preconditioner B and the support number $\sigma(A, B)$ will not increase. Implicitly, there is a corresponding term $A_i = 0$, so $\sigma(A_i, B_i) \leq 0$ for any psd B_i. It may seem strange to make the preconditioner B more complicated than necessary, but in fact B can often be made "simpler" (for example, sparser) by adding additional psd terms. This strategy is particularly well suited for cancelling out off-diagonal elements that make the preconditioner hard to factor.

Recall that when we split a preconditioner B into parts, $B = \sum_i B_i$, we normally require that all B_i be positive semidefinite. There is one exception to this rule. A matrix B_i may be indefinite or negative definite if it is supported by a set of psd matrices $\sum_{j \in S} B_j$ with support at most one. The combined matrix $B_i + \sum_{j \in S} B_j$ is then psd. In the expression

$$\tau B - A = \tau \sum_{i=1}^{k'} B_i - \sum_{i=1}^{k} A_i,$$

A and B are not necessarily decomposed into the same number of terms, that is, $k' \neq k$ is allowed. Hence some terms in B can be used to support non-psd terms in B. A special case of this technique was used by Bern et al. [6, Sec. 3.2].

9. Laplacian matrices and support graphs. As mentioned in the introduction, several previous authors have analyzed preconditioners using a closely related technique called *support–graph theory*. In this section we review the essentials of support–graph theory and show that they are just a special case (albeit a very useful one) of our basis support results from §4. Specifically, in Theorem 4.4 we showed how to support a rank one matrix uu^T with a larger symmetric matrix VV^T. In support–graph theory the vectors u and the columns of v are generally limited to have two nonzeros each. And the two nonzeros are of equal magnitude. Recall that a basic tool in support theory is to split a general matrix into simpler parts. What classes of matrices can be split into sums of such restricted outer products?

Consider first the case where the two nonzeros in u are of opposite sign, so $u_i = \sqrt{\alpha}$ and so $u_j = -\sqrt{\alpha}$. Then the nonzero portion uu^T (in rows and columns i and j) is just

$$\begin{pmatrix} \alpha & -\alpha \\ -\alpha & \alpha \end{pmatrix}.$$

A positive linear combination of such matrices can produce any matrix that is symmetric, positive semidefinite, diagonally dominant, has non-positive off-diagonal elements and has zero row (and column) sums. We call this class of matrices for *Laplacians* or Laplacian matrices. (Remark that Gremban used the term *generalized Laplacians* [13].) This class of matrices includes many standard discretizations of Laplace's or Poisson's equation and other elliptic equations and so is quite important in practice. By also including u vectors with a single nonzero one can augment the diagonal values, thus allowing matrices with positive row sums. This corresponds to different (e.g., Dirichlet) boundary conditions in the differential equation.

If we also allow the two nonzeros in the u vector to be of equal sign, then the nonzero contribution from uu^T is

$$\begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix}.$$

Any positive linear combination of such matrices is symmetric positive semidefinite and diagonally dominant, but now the off diagonal values are non-negative. Combining all these observations it is easy to show the following.

Proposition 9.1. *A symmetric matrix A with non-negative diagonal entries is diagonally dominant if and only if there exists a decomposition of the form $A = UU^T$, where each column of U has either one nonzero or exactly two nonzero entries and these two entries have the same magnitude. Furthermore, if all off-diagonal entries of A are non-positive, then A is also an M-matrix and any column of U with two nonzeros has entries of opposite signs.*

The columns of U are easy to construct in linear time. Each symmetric pair of off-diagonal nonzeros in A corresponds to a single column of U. Additional columns of U can be added to augment the diagonals. This correspondence between nonzeros of A and simple columns of U can be expressed in terms of graph theory. Specifically, consider the rows of the symmetric matrix A to be vertices of a graph, and for each nonzero off-diagonal a_{ij} add an edge between vertices i and j with weight equal to a_{ij}. Note that in our current notation, each such edge corresponds to a column of U. This relationship between Laplacian matrices, and more generally, diagonally dominant matrices, and graphs is at the heart of support–graph theory.

9.1. Congestion–Dilation results. Key tools in support–graph theory are various forms of what are called Congestion–Dilation Lemmas. Here we show that they follow directly from Theorem 4.4. A *path* between vertices i and j is a series of edges which leads from i to j. Let e^{ij} be a vector corresponding to the edge between i and j

in which all elements are zero, except for $e_i^{ij} = 1$ and $e_j^{ij} = -1$. Define $E^{ij} = e^{ij}(e^{ij})^T$. Consider the set of vectors comprising a path from i to j. By adding or subtracting these vectors as appropriate, all the intermediate values will cancel and the result will be equal to e^{ij}. In this way, a path can be used to support an edge. In particular, as we state more formally below, the support number is equal to the *dilation*, the number of edges in the path. A preconditioner containing a set of such paths can be built which supports any symmetric, diagonally dominant matrix with non-positive off-diagonals. This was Vaidya's key observation and is a principal idea in support–graph theory.

Note that a single edge in the preconditioner might be on many such support paths. In this case, the support number also depends on the number of paths it must support – its *congestion*. These observations are made more rigorous in the following results.

PROPOSITION 9.2 (Edge Congestion–Dilation). *Suppose $A = aE^{1,k+1}$ for some k, and that $B = b\sum_{i=1}^{k} E^{i,i+1}$ where a and b are positive and E^{ij} is as defined above. Then $\sigma(A, B) = (a/b)k$.*

Proof. Applying Theorem 4.4 with $u = \sqrt{a}e^{1,k+1}$ and $V = \sqrt{b}\left(e^{1,2}, e^{2,3}, \ldots, e^{k,k+1}\right)$ we find that $w = V^{+}u = \sqrt{\frac{a}{b}}(1, 1, \ldots, 1)^T$, and hence $||w||_2^2 = (a/b)k$. ∎

This is essentially the version of the Congestion–Dilation Lemma given in [6]. A more general result can be derived for the case in which the coefficients in B are of varying magnitude.

PROPOSITION 9.3 (Path Congestion–Dilation). *Suppose $A = aE^{1,k+1}$ for some k, and that $B = \sum_{i=1}^{k} b_i E^{i,i+1}$ where a, $b_i > 0$ and E^{ij} is as defined above. Then*

$$\sigma(A, B) = \sum_{i=1}^{k} \frac{a}{b_i}$$

Proof. From Theorem 4.4 with $u = \sqrt{a}e^{1,k+1}$ and $V = \left(\sqrt{b_1}e^{1,2}, \sqrt{b_2}e^{2,3}, \ldots, \sqrt{b_k}e^{k,k+1}\right)$ we find that $w = (\sqrt{\frac{a}{b_1}}, \ldots, \sqrt{\frac{a}{b_k}})^T$, and the result follows. ∎

This proposition says that the support is bounded by the sum of the edge congestions along a path. The Path Congestion–Dilation Proposition is not new; variations of this result have been stated by Gremban [13, Lemma 4.6] and by Guattery [15]. This proposition was also (implicitly) used by Guattery, Leighton, and Miller [16] in their *path resistance method* to bound the Fiedler eigenvalue of Laplacians.

The preceding propositions only consider the support for a single edge by a single path. More interesting is the case for set of edges being supported by a set of paths, that is, we have a graph embedding. The set of edges will correspond to a matrix A and the set of paths to a preconditioner B, where both A and B are Laplacians. Represent A and B by graphs G_A and G_B respectively, and each edge $e \in G_A$ is mapped to a path in G_B that connects the endpoints of e. (Note that a path may be a single edge.) One strategy is to use the Splitting Proposition and break A into a sum of edges and B into a sum of paths, and apply Proposition 9.3 to each of these pairs.

PROPOSITION 9.4 (Basic Graph Congestion–Dilation). *Given Laplacian matrices A and B, choose a mapping of the edges in the graph G_A onto paths in G_B. For each $e \in E(G_A)$, let $path(e)$ denote the corresponding path in G_B, and let $c(f)$ denote the number of supporting paths an edge f participates in, where $f \in E(G_B)$. Then*

$$\sigma(A, B) \leq \max_{e \in E(G_A)} \sum_{f \in path(e)} \frac{a_e\, c(f)}{b_f}$$

This result is a slight extension of the "worst congestion times worst dilation" bound used in [13, 6]. With our Symmetric Product Theorem (Theorem 4.5), we can show the following stronger result, which to the best of our knowledge is new.

THEOREM 9.5 (Graph Congestion–Dilation). *Given Laplacian matrices A and B, choose a mapping of the edges in the graph G_A onto paths in G_B. For each $e \in E(G_A)$, let path(e) denote the corresponding path in G_B. Then*

$$\sigma(A,B) \leq \left(\max_{e \in E(G_A)} \sum_{f \in path(e)} \sqrt{\frac{a_e}{b_f}} \right) \left(\max_{f \in E(G_B)} \sum_{e | f \in path(e)} \sqrt{\frac{a_e}{b_f}} \right),$$

and also

$$\sigma(A,B) \leq \sum_{e \in E(G_A)} \sum_{f \in path(e)} \frac{a_e}{b_f} = \sum_{f \in E(G_B)} \sum_{e | f \in path(e)} \frac{a_e}{b_f}.$$

Proof. Let U, V have the structure described in Proposition 9.1 and satisfy $UU^T = A$ and $VV^T = B$. Let $w_{ef} = \sqrt{a_e}/\sqrt{b_f}$, where $e \in E(G_A)$ and $f \in E(G_B)$, if f belongs to *path*(e). It is straightforward to verify that for appropriately chosen signs (the signs do not affect the norms of W), $W = \{\pm w_{ef}\}$ satisfies $VW = U$. By Theorem 4.5 and Lemma 4.9, $\sigma(A,B) \leq \|W\|_2^2 \leq \|W\|_1 \|W\|_\infty$ and also $\sigma(A,B) \leq \|W\|_2^2 \leq \|W\|_F^2$. □

In the unweighted case (a_e, b_f, and w_{ef} are 0 or 1), the first bound has a simple interpretation: The first term, $\max_e \sum_f w_{ef}$, is the maximum number of support paths that include any particular edge – that is, the maximum congestion. The second term, $\max_f \sum_e w_{ef}$, is the length of the longest path, or the maximum dilation. Thus the support number is bounded by the product of the maximum congestion and the maximum dilation. In the weighted case, the square roots in the definition of w_{ef} are significant and our result is different from previously used bounds.

The second bound, based on the Frobenius norm, shows that the support number is bounded by the sum of all congestions, or equivalently, the sum of all dilations in the graph embeddding. This bound is tighter than the bound in Proposition 9.4 We remark that in the weighted case, the two bounds given in the theorem above are incomparable.

Theorem 9.5 assumes that each edge in G_A is supported by a unique path in G_B. More generally we can support an edge by a (finite) set of paths. This corresponds to a fractional mapping where each edge weight may be split up into several parts and mapped to different paths in G_B. It is straight-forward to extend the theorem above to fractional mappings.

9.2. Vaidya's preconditioners and maximum–weight basis preconditioners. Vaidya [30] used the above graph interpretation to construct preconditioners for Laplacian matrices based on *maximum–weight spanning trees* (MST). (Recall that a spanning tree is a tree that spans all vertices of a given graph, and that the weight of a tree is the sum of the weights of the edges in the tree. There are efficient algorithms to find spanning trees of maximum or minimum weight.) One advantage of using a tree is that the corresponding matrix can be factored in linear time with no fill. It is easy to show that the edges of a spanning tree constitute a basis for a graph and hence also for a Laplacian.

Vaidya stated a version of the following result, which was later described in detail in [6].

Proposition 9.6. *Let A be a Laplacian matrix with a corresponding graph G_A. Let T be a maximum–weight spanning tree for G_A, and let A_T be the Laplacian matrix of T. Then $\sigma(A_T, A) \leq 1$ and $\sigma(A, A_T) \leq mn$, where n is the number of vertices and m is the number of edges in G_A.*

Vaidya claimed but did not prove that his techniques could be extended to all diagonally dominant matrices (that is, graphs with both positive and negative edge weights) using *maximum–weight bases.* This claim is finally proved in a recent paper with Chen and Toledo [7] using techniques from the present paper.

The gist of maximum–weight basis preconditioners is as follows: Let $A = UU^T$. We seek a preconditioner of the form $B = VV^T$, where the columns of V are a subset of the columns of U that span $Range(U)$. The challenge is to choose V such that $\sigma(A, B)$ is small and at the same time B is easy to solve for (factor). The columns of U can be used to define a structure called a *matroid.* It is easy to find a maximal independent set in a matroid, and such a set is a basis that can be chosen as our V. There are many different bases (maximal independent sets), so ideally we would like to find an optimal basis, i.e., one that minimizes $\sigma(UU^T, VV^T)$. However, we conjecture that this problem is NP-hard, so a heuristic or approximation algorithm must be used instead. Vaidya suggested choosing a set of vectors that are linearly independent and have the largest possible norm, that is, a maximum–weight basis, which is a maximal independent set in a *weighted* matroid. For Laplacian matrices (symmetric, diagonally dominant M-matrices with zero row sums), a basis is simply a spanning tree so the maximum-weight-basis problem reduces to the well-studied problem of finding a MST. We refer to [7] for details.

Vaidya's maximum–weight spanning tree preconditioner is simple and not very effective in practice [10]. However, Vaidya also proposed a more complicated preconditioner that contains more edges than just the tree and yields a better condition number. This *augmented* maximum–weight spanning tree preconditioner has been described in [6] and was recently implemented by Chen and Toledo [10]. They concluded that the preconditioner is very effective on discretizations of 2d elliptic partial differential equations, but less so in 3d. Vaidya [30] and Reif [26] proposed to use the augmented MST preconditioners recursively, yielding a solver that runs in near linear time for several classes of matrices; a fairly remarkable result. This recursive algorithm has not yet been studied empirically.

10. Example: Block Jacobi. In this section, we show how support theory can be used to analyse the well-known block Jacobi preconditioner for a model problem. The analysis is purely algebraic. We reproduce known bounds in a different and perhaps simpler way.

10.1. The 1-d model problem. We start with the 1-dimensional (higher dimensions will be considered later) Laplace equation with Dirichlet boundary conditions,

$$-u_{xx} = f(x), \quad x \in \Omega = [0, 1].$$

Suppose that Ω has been uniformly discretized using n points, and let $h = 1/n$. We need to solve a system $Au = f$, where A is a tridiagonal matrix with all twos on the diagonal and -1 on the sub- and super-diagonals, and u and f are discretizations of $u(x)$ and $f(x)$, respectively.

We wish to analyze the block Jacobi method, which corresponds to a simple domain decomposition method without overlap. Let B be the block Jacobi operator for a certain decomposition of A. Note that we do not assume that the blocks have the same sizes, or in other words, the subdomains may vary in size. Let q denote the number of subdomains, or equivalently, the number of diagonal blocks in B.

Consider the following example, where $n = 7$ and $q = 3$.

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & & & & \\ & & & 2 & -1 & & \\ & & & -1 & 2 & & \\ & & & & & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix}.$$

We now bound the eigenvalues of the preconditioned operator $B^{-1/2}AB^{-1/2}$ using support theory. Recall (see Definition 2.1) that the support number $\sigma(A, B)$ is roughly given by $\sigma(A, B) = \min\{t \,|\, tB - A \text{ is psd }\}$, and that $\kappa(B^{-1}A) \le \sigma(A, B)\sigma(B, A)$ (Proposition 2.4). It is easy to bound $\sigma(A, B)$, so the harder to prove bound is $\sigma(B, A)$.

LEMMA 10.1. *Let A be the discrete Laplace operator as defined above, and let B be a block diagonal approximation for A formed by dropping some of the off-diagonal entries. Then $\sigma(A, B) \le 2$.*

Proof. We observe that $2B - A$ is diagonally dominant with positive diagonal and hence positive semidefinite (by Lemma 5.1). Thus, $\sigma(A, B) \le 2$ because $t = 2$ in Definition 2.1 ensures that $tB - A$ is psd. ▯

In order to bound $\sigma(B, A)$ we will use the Symmetric Product Support Theorem (4.5). We factorize $A = VV^T$ and $B = UU^T$, where

$$V = \begin{pmatrix} 1 & 1 & & & & & \\ & -1 & 1 & & & & \\ & & -1 & 1 & & & \\ & & & \ddots & \ddots & & \\ & & & & -1 & 1 & \\ & & & & & -1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 1 & & & & & & & & \\ & -1 & 1 & & & & & & & \\ & & -1 & 1 & & & & & & \\ & & & & 1 & 1 & & & & \\ & & & & & -1 & 1 & & & \\ & & & & & & & 1 & 1 & \\ & & & & & & & & -1 & 1 \end{pmatrix}.$$

We seek a matrix W such that $VW = U$. One such W is

$$W = \begin{pmatrix} 1 & & & 1 & 1 & & & & & \\ & 1 & & -1 & -1 & & & & & \\ & & 1 & -1 & -1 & & & & & \\ & & & & -1 & & & & & \\ & & & & & 1 & 1 & & & \\ & & & & & & 1 & 1 & 1 & \\ & & & & & & 1 & 1 & & 1 \end{pmatrix}.$$

Since U and V have many columns in common, W is constructed to have mostly columns with only one nonzero element, which is 1. A few columns of W will have all nonzero entries either above or below a certain index. These columns correspond to the boundary of a block.

By inspection, the largest row sum in W for the example above is 3 and the largest column sum is 4, so $\|W\|_2^2 \le \|W\|_1 \|W\|_\infty \le 12$. Columns in W with more than one nonzero correspond to boundaries between subdomains. The corresponding columns in U have to be "supported" from the external boundary, $\partial\Omega$. In general, we can show that W has at most $q+1$ nonzeros in each row (one nonzero for each boundary of $q/2$ subdomains), and at most $(n+1)/2$ nonzeros in each column. Since each nonzero in W is ± 1, it follows that $\sigma(B, A) \le (q+1)(n+1)/2$.

LEMMA 10.2. *Let A be the discrete Laplace operator as defined above, and let B be a block diagonal approximation for A with $q > 1$ blocks formed by dropping some of the off-diagonal entries. Then $\sigma(B, A) \le (q+1)(n+1)/2$.*

Another way to obtain this result is to use the Congestion-Dilation Lemma (9.4) for graphs. In our case, we need support paths from the boundary nodes to each interior node that is on the boundary of a subdomain. Consequently, the dilation is $O(n)$ while the congestion is $O(q)$, which also gives the support bound $O(nq)$. (The factor $1/2$ comes from routing half the support paths from each boundary.)

By combining the two bounds on the support numbers, we get the following bound on the condition number:

THEOREM 10.3. *Let A be the discrete Laplace operator as defined above, and let B be a block diagonal approximation for A with $q > 1$ blocks formed by dropping some of the off-diagonal entries. Then the condition number κ satisfies $\kappa(B^{-1}A) \le (q+1)(n+1)$.*

A more detailed analysis in [9] showed that the condition number is bounded by $qn+q+1$. Our bound agrees with that bound up to a lower order term, and is simpler to derive. Since the Chang–Schultz bound is known to be tight [9], our bound is also tight asymptotically.

For the special case where uniform blocks are used, let $H = hn/q$ such that H denotes the subdomain size. This gives us the well-known result from domain decomposition that the condition number is bounded by $O(1/(hH))$.

10.2. Higher dimensions. We will show that the following result holds for block Jacobi preconditioning in higher dimensions than one.

THEOREM 10.4. *Consider a regular $n_1 \times n_2 \times \ldots \times n_d$ grid in d dimensions. Let A be the finite difference discretization of the Laplace equation. Suppose the domain is partitioned into subdomains, possibly in an unstructured fashion. Let B be the block Jacobi preconditioner corresponding to this partitioning (domain decomposition). Then*

$$\kappa(B^{-1}A) = O(\max_{1 \le i \le d} n_i q_i),$$

where q_i is the maximum number of subdomains along any line in the ith dimension.

Proof. Split $A = A_1 + A_2 + \ldots + A_d$ and similarly, $B = B_1 + \ldots + B_d$, where A_i corresponds to the Laplace finite difference operator along the lines in the ith dimension. Similarly, let B_i correspond to the block Jacobi approximation in the i'th dimension. By the Splitting Lemma (2.5), we have that

$$\sigma(B, A) \le \max_i \{\sigma(B_i, A_i)\}$$

Consider the algebraic equations along one line of gridpoints. Such a subset of equations corresponds precisely to the 1-d problem we analyzed in the previous section. Hence, $\sigma(B_i, A_i) = O(n_i q_i)$, and it follows that

$$\sigma(B, A) = O(\max_i n_i q_i).$$

The desired condition number bound follows by noting that $\sigma(A, B) \leq 2$ as in the 1-d case. $\square$

For a regular grid on the unit cube with $n^{1/d}$ grid points in each dimension and a uniform partitioning ($H = 1/q$) we obtain the expected bound $\sigma(B, A) = O(1/(hH))$.

10.3. Block Jacobi summary. We have rederived known bounds for block Jacobi using support theory. While a traditional analysis is based on calculating the eigenvectors (eigenfunctions) of the Laplacian, the support theory analysis is purely algebraic and does not require analytic expressions for the eigenvectors. Our analysis is a bit similar to the one in [9], but simpler in several ways. One advantage of our analysis is that it is easy to analyze nonuniform (irregular) decompositions of a domain. In this example, we only examined the Laplace equation on a structured grid. Our analysis tools also apply to more complicated equations and unstructured grids, though it is harder to obtain any general (a priori) bound.

11. Extensions to general matrices. Support theory was developed with symmetric, positive definite (spd) systems in mind. Nevertheless, much of the theory developed in the preceeding sections can be extended to general (including indefinite and nonsymmetric) matrices through a small change in the definition of support number.

DEFINITION 11.1. *For real matrices A and B with same number of colums, the* generalized *support number of (A, B) is defined by*

$$\hat{\sigma}(A, B) = \min\left\{t \,|\, x^T(\tau^2 B^T B - A^T A)x \geq 0 \text{ for all } x \text{ and for all } \tau \geq t\right\}$$

Note that generalized support numbers cannot be negative.

Since both $B^T B$ and $A^T A$ are symmetric, positive semidefinite (psd), all of the techniques introduced in the previous sections can be used to analyze $\hat{\sigma}(A, B)$. When $B^T B$ has full rank, then by Theorem 2.3 $\hat{\sigma}(A, B) = \sqrt{\lambda_{\max}(A^T A, B^T B)} = \mu_{\max}(A, B)$, where $\mu_{\max}(A, B)$ is the largest generalized singular value of the matrix pencil (A, B). (For a brief description of generalized singular values, see §8.7.3 of [12]. We use μ to denote singular values since the letter σ is reserved for support numbers in this paper.)

The spectral condition number $\kappa_2(C)$ is defined as $\kappa_2(C) = \|C\|_2\|C^{-1}\|_2 = \mu_{\max}(C)/\mu_{\min}(C)$, where again μ denotes a singular value. For nonsingular B, it follows that $\kappa_2(B^{-1}A) \leq \hat{\sigma}(A, B)\hat{\sigma}(B, A)$. When A and B are singular but share the same nullspace, then $\hat{\sigma}(A, B)\hat{\sigma}(B, A)$ bounds the effective condition number of the pencil (A, B) outside the nullspace. In short, the generalized support number can be used to bound the condition number in much the same way as the standard support number.

The quadratic form in the definition of generalized support can be factored in a useful manner. Specifically, for $A, B \in \mathbb{R}^{n \times m}$,

$$x^T(\tau^2 B^T B - A^T A)x = x^T(\tau B^T - A^T)(\tau B + A)x.$$

If A and B have different sizes, one can pad the smaller matrix with zeros. When A and B are both symmetric, this factorization reveals a close relationship between the generalized support number and the (standard) support number. Since the product of two symmetric psd matrices is also psd, the quadratic form on the right will be nonnegative if both the matrix terms are psd. These terms have the form used in standard support numbers, which leads to the following.

PROPOSITION 11.2. *If A and B are symmetric, then $\bar{\sigma}(A,B) \leq \max\{\sigma(A,B), \sigma(-A,B)\}$. Equality holds when B is spsd.*

If B is not psd, then $\sigma(A,B)$ is infinite and the bound becomes useless. In the case where both A and B are spsd a further reduction is possible. In this case, $\sigma(-A,B)$ is nonpositive, so Proposition 11.2 reduces to the following.

COROLLARY 11.3. *When A and B are both spsd, then $\hat{\sigma}(A,B) = \sigma(A,B)$.*

Thus, generalized support numbers are strict generalizations of the support numbers we defined in §2. Note however, that there is a discrepancy in definitions if either A or B is not psd. For example, if A is symmetric but negative definite, then the standard support number σ will be negative and corresponds to the largest (rightmost) generalized eigenvalue of (A,B). In contrast, the generalized support number $\hat{\sigma}$ is always nonnegative and corresponds to the largest magnitude of a generalized eigenvalue of (A,B).

Some of the propositions presented in this paper hold for generalized support numbers as well as the standard support number, but not all. In particular, the Splitting Lemma (Lemma 2.5) needs to be modified, as shown below.

PROPOSITION 11.4. *For splittings $A = A_1 + A_2$ and $B = B_1 + B_2$ where $B_1^T B_2$ is psd (possibly zero),*

$$\hat{\sigma}(A_1 + A_2, B_1 + B_2) \leq \max\left\{\hat{\sigma}(A_1,B_1), \hat{\sigma}(A_2,B_2), \sqrt{\max\{0, \sigma(A_1^T A_2, B_1^T B_2)\}}\right\}.$$

Proof. We have that $A^T A = (A_1+A_2)^T(A_1+A_2) = A_1^T A_1 + A_1^T A_2 + A_2^T A_1 + A_2^T A_2$, and similarly for $B^T B$. Hence

$$\begin{aligned} &x^T(\tau^2 B^T B - A^T A)x \\ &\quad \leq x^T\left((\tau^2 B_1^T B_1 - A_1^T A_1) + (\tau^2 B_2^T B_2 - A_2^T A_2) + (2\tau^2(B_1^T B_2 - A_1^T A_2)\right)x. \end{aligned}$$

Now choose τ by the right hand side bound in the Proposition. Since each of the three terms in the quadratic form above is then nonnegative, the total quadratic form must also be nonnegative. The desired result follows from Definitions 2.1 and 11.1. ☐

In the special case when $A_1^T A_2$ and $B_1^T B_2$ are both zero, this reduces to the standard splitting property.

Finally, it is possible that the standard support number may provide an indication about convergence even for non-spd systems. An analysis by Starke [29] shows that the residual of the GMRES method can be bounded by a simple function of the support number (although he did not use that terminology). We have not tried to determine which approach gives better bounds.

12. Summary and future work. All the results in this paper that hold for real symmetric matrices generalize to complex Hermitian matrices. This feature complements the work of Howle and Vavasis [18], who considered complex symmetric matrices. It is more difficult to go from symmetric to nonsymmetric systems. A major difficulty is that the correspondence between the support number and the largest generalized eigenvalue (Theorem 2.3) breaks down. In Section 11 we proposed to use the *generalized* support number, which is closely related to the generalized singular values, to bound the condition number in the non-spd case. The convergence analysis for iterative methods for nonsymmetric problems is quite complicated and further work is needed.

In the symmetric case, the Chebyshev (semi–)iterative method [31, 33] can benefit from support analysis because good bounds on the extreme eigenvalues are required.

We remark that Chebyshev iteration has the same worst–case complexity as conjugate gradients but requires no inner products. This may give Chebyshev iteration an advantage for large–scale problems on parallel computers. Also note that in general, the convergence of iterative methods depends not only on the extreme eigenvalues but on the distribution of all the eigenvalues. The support theory presented here only bounds the extreme eigenvalues. It is more difficult to obtain bounds for interior eigenvalues. See [1] for some such results.

The present paper extends the existing support–graph theory [6] from spsd, diagonally dominant M-matrices to a much wider class of matrices, namely all spsd matrices. Our framework is purely algebraic and no longer relies on graph theory (though graphs may still be useful in an analysis). The work presented here has enabled us to generalize Vaidya's preconditioners to all spd diagonally dominant matrices [7]. Using vectors with two nonzeros but possibly different magnitudes, we conjecture that the max-weight–basis preconditioners can be extended to all H–matrices.

The authors believe that the tools presented in the present paper are well suited both to analyze existing preconditioners and to develop new types of preconditioners. Promising candidates to analyze include incomplete factorizations and algebraic multilevel methods. The earlier support–graph theory has already been successfully applied to a multilevel preconditioner by Gremban [13], and to incomplete factorization preconditioners by Guattery [15] and Bern et al. [6]. However, the results are restricted to fairly specific problem instances and matrix classes. We predict that the techniques presented in the present paper can be used to extend some of these methods and results to all symmetric positive definite matrices. In a forthcoming paper [8], we will show how to apply the support theory presented here to elliptic finite element systems.

The support preconditioners we and others have developed all rely on using the Rank–1 Support Theorem (Theorem 4.4) or the Symmetric Product Support Theorem (Theorem 4.5) where columns of U and V correspond to edges in a graph (that is, they have only two nonzeros and these have the same magnitude). An open question is whether efficient preconditioners can be constructed that employ column vectors with three or more nonzeros. Although the theory in the present paper can handle this situation, a major obstacle in practice is that the resulting preconditioner may be difficult to solve for (i.e., factorize).

Acknowledgments. We wish to thank John Gilbert for introducing us to support-graph preconditioners, Roger Horn for pointing us to the results in §7, and David Keyes for suggesting the example in §10. We also thank Michele Benzi, Gene Golub, Steve Guattery, and Sivan Toledo for helpful comments and suggestions.

REFERENCES

[1] O. Axelsson. Bounds of eigenvalues of preconditioned matrices. *SIAM J. Matrix Analysis*, 13(3):847–862, 1992.

[2] O. Axelsson. *Iterative Solution Methods.* Cambridge University Press, 1994.

[3] O. Axelsson and V. A. Barker. *Finite Element Solution of Boundary Value Problems. Theory and Computation.* Academic Press, Orlando, FL, 1984.

[4] R. Beauwens. Upper eigenvalue bounds for pencils of matrices. *Lin. Alg. Appl.*, 62:87–104, 1984.

[5] R. Beauwens. Lower eigenvalue bounds for pencils of matrices. *Lin. Alg. Appl.*, 85:101–119, 1987.

[6] M. Bern, J. R. Gilbert, B. Hendrickson, N. Nguyen, and S. Toledo. Support-graph preconditioners. Technical report, School of Computer Science, Tel-Aviv University, 2001. Submitted

for publication. A shorter version appeared in *Proceedings of the 5th Copper Mountain Conference On Iterative Methods*, 1998.

[7] E. G. Boman, D. Chen, B. Hendrickson, and S. Toledo. Maximum-weight-basis preconditioners. *Numerical Linear Algebra with Appl.*, to appear, 2002.

[8] E. G. Boman and B. Hendrickson. Provably good preconditioners for finite element solution of elliptic problems. Manuscript in preparation. A related talk was given at Preconditioning 2001, Tahoe City, California, April 2001.

[9] M. Y. Chang and M. H. Schultz. Bounds on block diagonal preconditioning. *Parallel Algorithms and Applications*, 1:141–164, 1993.

[10] D. Chen and S. Toledo. Implementation and evaluation of Vaidya's preconditioners. Technical report, University of Tel-Aviv, Israel, 2001. Available at `http://www.math.tau.ac.il/~stoledo/pubs.html`.

[11] M. Fiedler and T. L. Markham. An observation on the Hadamard product of Hermitian matrices. *Lin. Alg. Appl.*, 215:179–182, 1995.

[12] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, USA, second edition, 1989.

[13] K. Gremban. *Combinatorial Preconditioners for Sparse, Symmetric, Diagonally Dominant Linear Systems*. PhD thesis, School of Computer Science, Carnegie-Mellon University, 1996. Available as Tech. Report CMU-CS-96-123.

[14] K. Gremban, G.L. Miller, and M. Zagha. Performance evaluation of a new parallel preconditioner. In *Proc. 9th International Parallel Processing Symposium*, 1995. A longer version is available as Tech. Report CMU-CS-94-205, Carnegie Mellon University.

[15] S. Guattery. Graph embedding techniques for bounding condition numbers of incomplete factor preconditioners. Technical Report 97-47, ICASE, NASA Langley Research Center, 1997.

[16] S. Guattery, F. T. Leighton, and G. Miller. Path resistance method for bounding λ_2 of a Laplacian. In *Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 201–210, 1997.

[17] Roger A. Horn. Personal communication, 2001.

[18] V. Howle and S. Vavasis. An iterative method for solving complex-symmetric systems arising in electric power modeling. Technical report, Cornell University, 2000.

[19] B. M. Irons and G. Treharne. A bound theorem in eigenvalues and its practical applications. In *Proc. Third Conference on Matrix Methods in Structural Mechanics*. Wright–Patterson AFB, Ohio, 1971.

[20] H-C. Lee and A. J. Wathen. On element-by-element preconditioning for general elliptic problems. *Comp. Meth. Appl. Mech. Eng.*, 92:215–229, 1991.

[21] M. M. Magolu. Analytical bounds for block approximate factorizations. *Lin. Alg. Appl.*, 179:33–57, 1993.

[22] M. M. Magolu and Y. Notay. On the conditioning analysis of block approximate factorization methods. *Lin. Alg. Appl.*, 154-156:583–599, 1991.

[23] Y. Notay. Conditioning analysis of modified block incomplete factorizations. *Lin. Alg. Appl.*, 154-156:711–722, 1991.

[24] Y. Notay. Upper eigenvalue bounds and related modified incomplete factorizations. In *Iterative Methods in Linear Algebra*, pages 551–562. North–Holland, Amsterdam, 1992.

[25] R. Reams. Hadamard inverses, square roots and products of almost semidefinite matrices. *Lin. Alg. Appl.*, 288:35–43, 1999.

[26] J.H. Reif. Efficient approximate solution of sparse linear systems. *Computers and Mathematics with Applications*, 36(9):37–58, 1998. See also errata in volume 38, page 141, 1999.

[27] I. Schur. Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. *J. Reine Angew. Math.*, 140:1–28, 1911.

[28] B. Smith, P. E. Bjørstad, and W. Gropp. *Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, 1996.

[29] G. Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Num. Math.*, 78:103–117, 1997.

[30] P. M. Vaidya. Solving linear equations with symmetric diagonally dominant matrices by constructing good preconditioners. Unpublished manuscript. A talk based on this manuscript was presented at the IMA Workshop on Graph Theory and Sparse Matrix Computations, Minneapolis, October 1991.

[31] R. S. Varga. *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1962.

[32] A. J. Wathen. An analysis of some element-by-element preconditioning techniques. *Comp. Meth. Appl. Mech. Eng.*, 74:271–287, 1989.

[33] D. M. Young. *Iterative Solution of Large Linear Systems*. Academic Press, New York, NY, USA, 1971.