Finding Norms of Hadamard Multipliers

C. C. Cowen,*

Purdue University, West Lafayette, Indiana
P. A. Ferguson,

Purdue University, West Lafayette, Indiana
D. K. Jackman,

Wooster College, Wooster, Ohio
E. A. Sexauer,

University of Richmond, Richmond, Virginia
C. Vogt,

Purdue University, West Lafayette, Indiana

and

H. J. Woolf

Purdue University, West Lafayette, Indiana

Submitted by Roger A. Horn

ABSTRACT

The norm of a matrix B as a Hadamard multiplier is the norm of the map $X \mapsto X \bullet B$, where $\bullet$ is the Hadamard or entrywise product of matrices. Watson proposed an algorithm for finding lower bounds for the Hadamard multiplier norm of a matrix. It is shown how Watson's algorithm can be used to give upper bounds as well which, in many cases, yield the Hadamard multiplier norm to any desired accuracy. A sharp form of Wittstock's decomposition theorem is proved for the special case of Hadamard multiplication.

---

*Supported in part by a grant from the National Science Foundation including support from the REU program for all but the first author.

655 Avenue of the Americas, New York, NY 10010 0024- 3795/93/$6.00

1. INTRODUCTION

By the *Hadamard product*, also called the Schur product, we mean the entry–wise product of matrices: if A and B are $m \times n$ matrices, their Hadamard product, $A \bullet B$ is the $m \times n$ matrix whose entries are $a_{jk}b_{jk}$. In this paper, we study the operator on the set $\mathcal{M}_n$ of $n \times n$ real matrices given by $X \mapsto X \bullet B$, for a fixed B. For a matrix X in $\mathcal{M}_n$, $\|X\|_{\mathrm{sp}}$ will denote its spectral norm (i.e., the largest singular value of X) and the norm K_B of the Hadamard multiplier is its norm as an operator on $\mathcal{M}_n$ with the spectral norm, that is,

$$K_B = \max\{\|X \bullet B\|_{\mathrm{sp}} : \|X\|_{\mathrm{sp}} \leq 1\}$$

Recently, Watson [15] proposed an algorithm for finding, among other things, a lower bound for K_B. Using a factorization result of [3] based on the factorization theorem of Haagerup [6, 12], we show how to use Watson's algorithm to give an upper bound for K_B as well. In many interesting cases, for example the triangular truncation matrices studied in [3] or McEachin's matrices studied in [11, 5], the upper and lower bounds agree. (Sample results from the algorithm coded in Matlab [10] are included.) Hadamard multiplication is an example of a completely bounded map [12]; we give a sharp form of Wittstock's decomposition theorem for completely bounded maps in the special case of Hadamard multiplication by a Hermitian matrix. This result could, in principle, also be used to find K_B when B is Hermitian, but the resulting optimization problem does not seem likely to lead to a rapidly converging algorithm. However, if B is Hermitian with rank 2, this formulation does lead to a relatively simple one variable optimization problem that can be solved explicitly. More information about the Hadamard product and its properties can be found in [7].

The authors would like to thank the referees for several helpful suggestions that led to improvements and simplifications of our presentation.

2. WATSON'S ALGORITHM

Throughout, $\langle \cdot, \cdot \rangle$ will denote the standard Euclidean inner product on $\mathbf{R}^n$ and by the norm of a vector, we will mean the Euclidean norm. If P is in $\mathcal{M}_n$, we will denote its (Hilbert space) adjoint, that is, its transpose, by P^*. Much of what is presented here can be immediately generalized to the complex case, but the algorithm is designed for the real case.

First of all, we want to reinterpret Watson's algorithm which he motivated by use of subgradients in the context of convex optimization. We think of Hadamard multiplication by B as a linear transformation on the vector space $\mathcal{M}_n$. The dual of $\mathcal{M}_n$ as a vector space is isomorphic to $\mathcal{M}_n$ where, for G in $\mathcal{M}_n$, the linear functional Λ_G is defined by

$$\Lambda_G(X) = \mathrm{trace}(G^* X)$$

Regarding $\mathcal{M}_n$ as normed space with the spectral norm, its dual space is normed by the *trace norm*

$$\|\Lambda_G\| = \|G\|_{\mathrm{tr}} = \max\{\mathrm{trace}(V^* G) : \|V\|_{\mathrm{sp}} = 1\}$$

that is, the trace norm of G is the sum of the singular values of G.

For B, G, and X in $\mathcal{M}_n$, the identity

$$\Lambda_G(B \bullet X) = \mathrm{trace}(G^*(B \bullet X)) = \mathrm{trace}((B \bullet G)^* X) = \Lambda_{B \bullet G}(X)$$

shows that the dual of Hadamard multiplication by B is again Hadamard multiplication by B. Since the norm of a transformation and its dual are the same, we have

$$K_B = \max\{\|G \bullet B\|_{\mathrm{tr}} : \|G\|_{\mathrm{tr}} \leq 1\}$$

Our interpretation of Watson's algorithm is that it finds successively better estimates for the norm and the dual norm of Hadamard multiplication by B. Notice that since the unitaries are the extreme points of the unit ball of $\mathcal{M}_n$ with the spectral norm and the rank one matrices of norm 1 are the extreme points of the unit ball of $\mathcal{M}_n$ with the trace norm that the Hadamard multiplier norm is achieved at such matrices. Moreover, using singular value decomposition, given a matrix H in $\mathcal{M}_n$, it is easy to find a unitary matrix U so that $\|H\|_{\mathrm{tr}} = \mathrm{trace}(H^* U)$ and a rank one matrix G with $\|G\|_{\mathrm{tr}} = 1$ so that $\|H\|_{\mathrm{sp}} = \mathrm{trace}(G^* H)$. Specifically, if $H = VDW^*$ is the singular value decomposition of H where V and W are unitary and D is the diagonal matrix whose diagonal entries are the singular values of H arranged in non–increasing order, then letting $U = VW^*$,

$$\begin{aligned} \|H\|_{\mathrm{tr}} &= \mathrm{trace}(D) = \mathrm{trace}(W^*(WDV^*)V) \\ &= \mathrm{trace}((WDV^*)(VW^*)) = \mathrm{trace}(H^* U) \end{aligned} \tag{1}$$

Similarly, letting $G = xy^*$ where x and y are the first columns of V and W respectively,

$$\|H\|_{\mathrm{sp}} = \mathrm{trace}(x^* VDW^* y) = \mathrm{trace}(yx^* VDW^*) = \mathrm{trace}(G^* H) \tag{2}$$

With this in mind, we can easily describe Watson's algorithm.

Watson's Algorithm

Choose a unitary matrix U_0.

For $k = 1$ to K_{stop}

$s_k = \|B\bullet U_{k-1}\|_{\text{sp}}$
Choose G_k rank one with $\|G_k\|_{\text{tr}} = 1$
and $\text{trace}(G_k^*(B\bullet U_{k-1})) = s_k$
$t_k = \|B\bullet G_k\|_{\text{tr}}$
Choose U_k unitary with $\text{trace}((B\bullet G_k)^* U_k) = t_k$

End

Since

$$s_k = \|B\bullet U_{k-1}\|_{\text{sp}} \leq K_B \|U_{k-1}\|_{\text{sp}} = K_B$$

and

$$t_k = \|B\bullet G_k\|_{\text{tr}} \leq K_B \|G_k\|_{\text{tr}} = K_B$$

the sequences of estimates s_k and t_k are bounded. Moreover, for each k,

$$\begin{aligned} s_{k+1} &= \|B\bullet U_k\|_{\text{sp}} \\ &= \max\{\text{trace}(G^*(B\bullet U_k)) : \text{rank}(G) = 1 \ \text{ and } \ \|G\|_{\text{tr}} = 1\} \\ &\geq \text{trace}(G_k^*(B\bullet U_k)) = \text{trace}((B\bullet G_k)^* U_k) = t_k \end{aligned}$$

and

$$\begin{aligned} t_k &= \|B\bullet G_k\|_{\text{tr}} = \max\{\text{trace}((B\bullet G_k)^* U) : U \text{ is unitary}\} \\ &\geq \text{trace}((B\bullet G_k)^* U_{k-1}) = \text{trace}(G_k^*(B\bullet U_{k-1})) = s_k \end{aligned}$$

so that

$$s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq K_B$$

It follows that the estimates s_k and t_k converge to a lower bound for K_B and that, by possibly choosing subsequences, we can find a unitary matrix U_∞ and a rank one matrix G_∞ with $\|G_\infty\|_{\text{tr}} = 1$ so that

$$\lim s_k = \lim t_k = \|(B\bullet G_\infty)\|_{\text{tr}} = \|(B\bullet U_\infty)\|_{\text{sp}} \leq K_B$$

Naturally, we hope that the limits are actually equal to K_B and this does happen for some interesting matrices B. However, it does not always happen; indeed, taking $U_0 = I$ just gives the limit equal to the largest absolute value diagonal entry, not always the same as K_B. We need a good upper estimate of K_B in order to decide if a particular lower estimate is giving acceptable results.

For P an $n \times n$ matrix with columns $P_1, P_2, \cdots, P_n$, let

$$c(P) = \max\{\|P_1\|, \|P_2\|, \cdots, \|P_n\|\}$$

Haagerup [6] (or see [12, pp. 110–116], [1], [13], or [9]) showed that if B is an $n \times n$ matrix then the norm of B as a Hadamard multiplier is

$$K_B = \min\{c(S)c(R) : S^*R = B\}$$

We will need the following extension of Haagerup's theorem from [3].

Theorem A. *Let B be a non–zero $n \times n$ matrix. If S and R are $n \times n$ matrices and $S^*R = B$ is a factorization of B with $c(S) = c(R) = \sqrt{K_B}$, there is a unitary matrix U and unit vectors x and y so that if X and Y are the diagonal matrices with diagonals x and y, then*

$$K_B = \langle B \bullet Uy, x\rangle = \mathit{trace}((xy^*)^* B \bullet U) = \|B \bullet U\|_{\mathrm{sp}}$$

$RYU^* = SX$, and the columns satisfy $\|S_j\| = c(S)$ and $\|R_k\| = c(R)$ whenever the components x_j and y_k are non-zero. Conversely, if S and R are $n \times n$ matrices satisfying $B = S^*R$ and x and y are unit vectors such that $\|S_j\| = \|R_k\| = c(S) = c(R)$ whenever x_j and y_k are non-zero and U is a unitary matrix so that $RYU^* = SX$, then $K_B = c(R)c(S) = \|B \bullet U\|_{\mathrm{sp}}$.

We want to use the convergents coming from Watson's algorithm to find a factorization of B. Since every factorization of B gives an upper bound for K_B and Watson's algorithm gives a lower bound, we hope to obtain accurate estimates for the Hadamard multiplier norm.

Suppose the G's and U's are chosen successively as in the outline of Watson's algorithm above where $U_k = V_k W_k^*$ and $G_k = x_k y_k^*$ using the singular value decomposition $B \bullet (x_k y_k^*) = V_k D_k W_k^*$ as in equations (1) and (2) and $s_k = \|B \bullet U_{k-1}\|_{\mathrm{sp}}$ and $t_k = \|B \bullet G_k\|_{\mathrm{tr}}$. If s, t, x, y, D, V, and W are the limits (perhaps of subsequences) of s_k, t_k, x_k, y_k, D_k, V_k, W_k respectively, then $s = t$, $\|x\| = \|y\| = 1$, V and W are unitary and satisfy

$$B \bullet (xy^*) = VDW^*$$

and

$$\begin{aligned} s &= \|B \bullet (xy^*)\|_{\mathrm{tr}} = \mathrm{trace}\left((VW^*)^* B \bullet (xy^*)\right) \\ &= \mathrm{trace}\left((xy^*)^* B \bullet (VW^*)\right) = \|B \bullet (VW^*)\|_{\mathrm{sp}} \end{aligned}$$

Now let X and Y be the diagonal matrices with diagonals x and y so that $B \bullet (xy^*) = XBY$. If X and Y are invertible, letting C be the non–negative diagonal matrix with $C^2 = D$, then

$$XBY = B\bullet(xy^*) = VDW^* = VC^2W^*$$

implies

$$B = (X^{-1}VC)(CW^*Y^{-1})$$

Taking

$$S = CV^*X^{-1} \quad \text{and} \quad R = CW^*Y^{-1} \tag{3}$$

and $U = VW^*$, we have $B = S^*R$,

$$RYU^* = (CW^*Y^{-1})Y(VW^*)^* = CV^* = (CV^*X^{-1})X = SX$$

and $\|B\bullet U\|_{\mathrm{sp}} = s$.

While these conditions are not quite enough for the converse statement of Theorem A, this factorization is highly suggestive. Moreover, it indicates the need for investigating the conditions under which X and Y are not invertible. The last section of the paper gives some results of using the authors' Matlab code implementing Watson's algorithm to give a lower bound for K_B and using the factorization suggested above to give an upper bound. For the triangular truncation matrices and McEachin's matrices, the resulting X and Y are invertible and the computed values of the upper and lower bounds differ by less than 2×10^{-14} for sizes up to 50×50; in other words, for these matrices, the algorithm works to machine accuracy. (In addition, checking the extra hypotheses in the converse statement from Theorem A shows that the computed matrices satisfy (up to machine accuracy) these hypotheses, so the theory indicates they should give K_B.)

It is easy to see that if B is a matrix for which x and y are the standard unit basis vectors e_i and e_j, then $K_B = \|B\bullet(xy^*)\|_{\mathrm{tr}} = |b_{ij}|$, so the Hadamard multiplier norm of B is given by one of its entries. More generally, if $\widehat{B}$ is the submatrix of B whose rows and columns are associated with the non–zero components of x and y, then constructing the vectors $\widehat{x}$ and $\widehat{y}$ by omitting the zero components, we have

$$K_B = \|B\bullet(xy^*)\|_{\mathrm{tr}} = \|\widehat{B}\bullet(\widehat{x}\widehat{y}^*)\|_{\mathrm{tr}} \leq K_{\widehat{B}} \leq K_B$$

That is, the non–invertible X and Y are associated with submatrices of B that have the same Hadamard multiplier norm.

We will say a matrix B is *Hadamard irreducible* if there are no proper submatrices of B with the same Hadamard multiplier norm as B. In our experience, a large difference between the computed upper and lower bounds for K_B is due to the presence of a submatrix with a large multiplier norm (as compared with K_B), that is, with matrices that are reducible in this sense. Indeed, in such cases, the sequences x_k and y_k converge to vectors whose non–zero components are associated with just such a submatrix. Further analysis of this situation suggests a multistep procedure for finding an optimal Haagerup factorization in this case as well. Our analysis, below, of the case in which a proper submatrix has the same Hadamard multiplier norm as B leads to Theorem 2.1 which is the formal justification of the multistep procedure. The application of Theorem 2.1 in finding norms is illustrated in section 4.

Suppose B is an $n \times n$ matrix with block form

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

where B_{11} is invertible, Hadamard irreducible, and $K_{B_{11}} = K_B$. Let $B = S^*R$ be a factorization of B as in Theorem A. By multiplying S and R on the left by an appropriate unitary, we may assume without loss of generality that S is upper triangular. In block form, this is

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} S_{11}^* & 0 \\ S_{12}^* & S_{22}^* \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}$$

In particular, we have $B_{11} = S_{11}^* R_{11}$. Since

$$c(S_{11})c(R_{11}) \leq c(S)c(R) = K_B = K_{B_{11}} \leq c(S_{11})c(R_{11})$$

we see that in fact $c(S_{11}) = c(R_{11}) = \sqrt{K_B}$. But B_{11} is Hadamard irreducible, so each column of S_{11} and R_{11} has the same norm and the inequality

$$K_B = K_{B_{11}} = c(R_{11})^2 \leq c(R_{11})^2 + c(R_{21})^2 = c(R)^2 = K_B$$

shows that $R_{21} = 0$.

The invertibility of B_{11} implies that S_{11} and R_{11} are invertible also and it follows that $S_{12} = (R_{11}^*)^{-1} B_{21}^*$ and $R_{12} = (S_{11}^*)^{-1} B_{12}$. Since

$$B_{22} = S_{12}^* R_{12} + S_{22}^* R_{22} = B_{21} R_{11}^{-1} (S_{11}^*)^{-1} B_{12} + S_{22}^* R_{22}$$

Substituting, we find

$$S_{22}^* R_{22} = B_{22} - B_{21} B_{11}^{-1} B_{12}$$

which we will denote by A. (In particular, $K_A \leq K_B$ but we will improve this estimate.)

We may take the point of view that B_{11} and its factorization are given. In the circumstances we are considering, this also determines S_{12} and R_{12}, but it does not determine the factorization of A. Since K_B is known, we see that if u_j is the j^{th} column of S_{12} and v_j is the j^{th} column of S_{22}, then $\|u_j\|^2 + \|v_j\|^2 \le K_B$ so that $\|v_j\|^2 \le K_B - \|u_j\|^2$. We obtain similar inequalities for the columns of R_{22}. Thus, we want to factor A as $S_{22}^* R_{22}$ where the columns of S_{22} and R_{22} satisfy these inequalities. For each column of S_{22}, let $\phi_j = (K_B - \|u_j\|^2)^{-1/2}$ and let ψ_j be defined similarly for each column of R_{22}. If Φ and Ψ are the diagonal matrices whose diagonal entries are the ϕ_j and the ψ_j respectively, we see that A has a factorization $S_{22}^* R_{22}$ satisfying the inequalities for the column norms if and only if the matrix $Z = \Phi S_{22}^* R_{22} \Psi = \Phi A \Psi$ has a factorization $Z = \widehat{S}_{22}^* \widehat{R}_{22}$ with $c(\widehat{S}_{22}) \le 1$ and $c(\widehat{R}_{22}) \le 1$, that is $K_Z \le 1$. We summarize this analysis in the following theorem.

Theorem 2.1. *Suppose B is an $n \times n$ matrix with block form*

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

where B_{11} is invertible, Hadamard irreducible, and $K_{B_{11}} = K_B$. If $B_{11} = S_{11}^* R_{11}$ is a factorization with $c(S_{11})^2 = c(R_{11})^2 = K_B$, then B has a factorization of the form $B = S^* R$ where $K_B = c(S)^2 = c(R)^2$ and

$$S = \begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{pmatrix} \quad \textit{and} \quad R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix}$$

if and only if $K_Z \le 1$ where $Z = \Phi(B_{22} - B_{21} B_{11}^{-1} B_{12})\Psi$ as in the preceding discussion.

Corollary 2.2. *Suppose B is a non–zero $n \times n$ matrix such that $b_{11} \ge |b_{ij}|$ for $1 \le i,j \le n$ and $b_{11} > |b_{1j}|$ and $b_{11} > |b_{i1}|$ for $1 < i,j \le n$. Then $K_B = b_{11}$ if and only if $K_Z \le 1$ where*

$$Z_{ij} = \frac{(b_{11} b_{ij} - b_{i1} b_{1j})}{\sqrt{b_{11}^2 - b_{i1}^2}\sqrt{b_{11}^2 - b_{1j}^2}} \quad \textit{for } i,j = 2, \cdots, n$$

Although we have been unable to prove that Watson's algorithm converges to K_B for most starting unitaries and that the procedure above finds a factorization that shows the lower bound from Watson's algorithm is exact, the algorithm seems to be very robust. Of course, it is easy to produce starting unitaries that are stationary points for the algorithm: for example, the identity is a stationary point for the algorithm that gives the largest diagonal entry as the lower estimate. However, we have been unable to find a matrix and non–stationary starting unitary for which Watson's algorithm converges to a local maximum less than K_B.

3. WITTSTOCK'S DECOMPOSITION THEOREM

Recall that a positive semidefinite matrix B has $K_B = \max\{b_{jj} : j = 1, 2, \cdots, n\}$. First, we show that a Hermitian matrix can be split in such a way that its multiplier norm can be easily obtained from the pieces. This result is a sharp form, for our special case, of Wittstock's decomposition theorem [12, page 107].

THEOREM 3.1. *Let B be a Hermitian matrix. If P_1 and Q_1 are positive semidefinite matrices so that $B = P_1 - Q_1$, then $K_B \leq K_{P_1+Q_1}$. Moreover, there are positive semidefinite matrices P and Q so that $B = P - Q$ and $K_B = K_{P+Q}$ with rank(P) being the number of positive eigenvalues of B and rank(Q) being the number of negative eigenvalues.*

The theorem says we must be concerned with the ways in which the matrix B can be split as $B = P - Q$ where P and Q are positive semidefinite and rank(P) is the number of positive eigenvalues of B and rank(Q) is the number of negative eigenvalues of B. One such splitting is given by the spectral theorem: let P_0 be the restriction of B to the subspace spanned by the eigenvectors corresponding to the positive eigenvalues of B and let Q_0 be the restriction of $-B$ to the subspace spanned by the eigenvectors corresponding to the negative eigenvalues of B. We will call this splitting the *spectral splitting*. That this is not the only, or always the optimal, splitting of B is what gives this theorem interest.

Proof. (of Theorem 3.1.) Assume first that B is an invertible Hermitian matrix. It follows from the Theorem and Corollary 3 of [2, pages 183 and 195] that

$$K_B = \min\left\{K_Z : \begin{pmatrix} Z & B \\ B & Z \end{pmatrix} \geq 0\right\}$$

(See [5] and [9] for discussions of this characterization.)

Suppose P_1 and Q_1 are positive semidefinite matrices so that $B = P_1 - Q_1$. Then

$$\begin{pmatrix} P_1 + Q_1 & P_1 - Q_1 \\ P_1 - Q_1 & P_1 + Q_1 \end{pmatrix} = \begin{pmatrix} P_1 + Q_1 & B \\ B & P_1 + Q_1 \end{pmatrix}$$

is also positive semidefinite so, by the characterization above,

$$K_B \leq K_{P_1+Q_1}$$

Let A be a matrix that attains the minimum above, that is, $K_A = K_B$ and

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix} \geq 0$$

The matrix $C = A^{-1/2}BA^{-1/2}$ is an invertible Hermitian matrix with the same inertia as B. If we let $C = C_+ - C_-$ be the spectral splitting of C as a difference of positive semidefinite matrices, then $P = A^{1/2}C_+A^{1/2}$ and $Q = A^{1/2}C_-A^{1/2}$ are positive semidefinite matrices with $P - Q = B$ and $\text{rank}(P)$ and $\text{rank}(Q)$ being the number of positive and negative eigenvalues of B. Since

$$\begin{pmatrix} I & A^{-1/2}BA^{-1/2} \\ A^{-1/2}BA^{-1/2} & I \end{pmatrix}$$
$$= \begin{pmatrix} A^{-1/2} & 0 \\ 0 & A^{-1/2} \end{pmatrix} \begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} A^{-1/2} & 0 \\ 0 & A^{-1/2} \end{pmatrix} \geq 0$$

and $C_+C_- = 0$, it follows that $C_+ + C_- \leq I$ which implies $P + Q \leq A$. Thus, $K_B \leq K_{P+Q} \leq K_A = K_B$ and P and Q satisfy the conclusion of the theorem.

Now if B is not invertible, we can find a sequence of invertible Hermitian matrices converging to B and apply the results above to each of the matrices in the sequence. Since the Hadamard multiplier norm is continuous, we can find convergent subsequences to obtain positive semidefinite matrices P and Q so that $B = P - Q$, $K_B = K_{P+Q}$, and $\text{rank}(P)$ and $\text{rank}(Q)$ are, respectively, at least as large as the number of positive and negative eigenvalues. Let N be the orthogonal projection onto the range of B. Now $B = NBN = NPN - NQN$ and NPN and NQN are positive semidefinte matrices whose ranks are, respectively, the number of positive and negative eigenvalues of B. Since $NPN + NQN \leq P + Q$, the largest diagonal entry of $NPN + NQN$ is no more than the largest diagonal entry of $P + Q$ and we see that $K_B = K_{NPN+NQN}$. Thus, $B = NPN - NQN$ is a splitting of the sort needed for the conclusion. ■

As noted above, the spectral splitting of B is not the only splitting. To find other splittings, we consider the quadratic form $\langle Bv, v\rangle$, and let $\mathcal{P} = \{v : \langle Bv, v\rangle > 0\}$ and $\mathcal{N} = \{v : \langle Bv, v\rangle < 0\}$. In the invertible case, if $B = P - Q$ is a splitting and $Qx = 0$, then $\langle Bx, x\rangle = \langle Px, x\rangle \geq 0$. Since $\text{rank}(A) = n = \text{rank}(P) + \text{rank}(Q)$ and $A = P + Q$, only the zero vector is in the range of P and the range of Q. This means that only the zero vector is in the kernel of P and the kernel of Q. That is, if both $Qx = 0$ and $Px = 0$ then $x = 0$, so the non–zero x with $Qx = 0$ are in $\mathcal{P}$. In particular, the range of Q, which is spanned by the eigenvectors of Q, is a subspace whose orthogonal complement lies in $\mathcal{P}$. Similarly, the range of P, which is spanned by the eigenvectors of P, is a subspace whose orthogonal complement lies in $\mathcal{N}$. The theorem below uses this observation to characterize the splittings in a special case.

THEOREM 3.2. *Suppose B is an $n \times n$ Hermitian matrix with $n - 1$ positive, one negative eigenvalue and suppose u is an eigenvector of B for* *which $B = P_0 - uu^*$ is the spectral splitting of B. For w orthogonal to u, if $\alpha = \sqrt{1 + w^* P_0 w}$, then $Q = (\alpha u + P_0 w)(\alpha u + P_0 w)^*$ is a positive semidefinite matrix with rank one such that $P = B + Q$ is a positive semidefinite matrix with rank $n - 1$. Conversely, if $B = P - Q$ where P is a positive semidefinite matrix with rank $n - 1$ and Q is a positive semidefinite matrix with rank one, then there is a vector w orthogonal to u such that $Q = (\alpha u + P_0 w)(\alpha u + P_0 w)^*$ for $\alpha = \sqrt{1 + w^* P_0 w}$.*

Proof. Every positive semidefinite matrix Q with rank one has the form

$$Q = (\alpha u + v)(\alpha u + v)^*$$

for some $\alpha \geq 0$ and some vector v orthogonal to u. If α were 0, then u would be a vector orthogonal to the range of Q that is not in $\mathcal{P}$, so the remarks above show $\alpha > 0$ for the desired splittings of B as $P - Q$. Since P_0 has $n - 1$ positive eigenvalues, for every vector v orthogonal to u, there is a vector w orthogonal to u so that $P_0 w = v$. Now if the scalar β and the vector x orthogonal to u are such that $\beta u + x$ is orthogonal to the range of Q, then, recalling that $P_0 u = 0$, we have

$$0 = (\alpha u + P_0 w)^* (\beta u + x) = \alpha\beta\|u\|^2 + w^* P_0 x$$

so $\beta\|u\|^2 = -w^* P_0 x/\alpha$. Thus the condition that $\beta u + x$ be in $\mathcal{P}$ is

$$0 < (\beta u + x)^* (P_0 - uu^*)(\beta u + x) = x^* P_0 x - |\beta|^2 \|u\|^4$$

or

$$x^* P_0 x - \frac{1}{\alpha^2} x^* P_0 w w^* P_0 x > 0$$

For this to be true for every x orthogonal to u means that $P_0 - P_0 w w^* P_0/\alpha^2$ is positive definite as an operator on range(P_0). This is equivalent to $\alpha^2 I - (\sqrt{P_0} w)(\sqrt{P_0} w)^* > 0$ which is the same as $\|\sqrt{P_0} w\| < \alpha$.

Now suppose w, α, and Q are as in the statement of the theorem. Since

$$\alpha^2 = 1 + w^* P_0 w > w^* P_0 w = \|\sqrt{P_0} w\|^2$$

all vectors z that are orthogonal to the range of Q are in $\mathcal{P}$ so $z^* P z = z^* (B + Q) z = z^* B z > 0$ which implies the rank of P is at least $n - 1$. But

$$\begin{aligned}
&P\left(-\frac{\alpha}{\|u\|^2} u + w\right) \\
&\quad= \left(P_0 - uu^* + (\alpha u + P_0 w)(\alpha u + P_0 w)^*\right)\left(-\frac{\alpha}{\|u\|^2} u + w\right) \\
&\quad= P_0 w + \alpha u + (\alpha u + P_0 w)(-\alpha^2 + w^* P_0 w) \\
&\quad= (\alpha u + P_0 w)(1 - \alpha^2 + w^* P_0 w) = 0
\end{aligned}$$

so P has rank exactly $n-1$.

To prove the converse, suppose $B = P - Q$ is a splitting as in the theorem and $Q = (\alpha u + P_0 w)(\alpha u + P_0 w)^*$ for some $\alpha > 0$ and w orthogonal to u. Suppose $\beta u + x$ is a non–zero vector in the kernel of P where x is orthogonal to u. We have

$$\begin{aligned} P(\beta u + x) &= \left(P_0 - uu^* + (\alpha u + P_0 w)(\alpha u + P_0 w)^*\right)(\beta u + x) \\ &= P_0 x - \beta\|u\|^2 u + (\alpha u + P_0 w)(\alpha\beta\|u\|^2 + w^* P_0 x) \\ &= \left((\alpha^2 - 1)\beta\|u\|^2 + \alpha w^* P_0 x\right) u + P_0\left(x + \alpha\beta\|u\|^2 w + (w^* P_0 x) w\right) \end{aligned}$$

For this to be 0, since u is orthogonal to the range of P_0, we must have

$$(\alpha^2 - 1)\beta\|u\|^2 + \alpha w^* P_0 x = 0$$

so

$$w^* P_0 x = -(\alpha^2 - 1)\beta\|u\|^2/\alpha$$

Moreover, since P_0 is invertible on the set of vectors orthogonal to u, we must have

$$\begin{aligned} 0 &= x + \alpha\beta\|u\|^2 w + (w^* P_0 x) w \\ &= x + \left(\alpha\beta\|u\|^2 - (\alpha^2 - 1)\beta\|u\|^2 \frac{1}{\alpha}\right) w = x + \frac{\beta\|u\|^2}{\alpha} w \end{aligned}$$

Using this in the above equality, we see

$$-(\alpha^2 - 1)\frac{\beta\|u\|^2}{\alpha} = w^* P_0 x = -\frac{\beta\|u\|^2}{\alpha} w^* P_0 w$$

If β were 0, this would mean $x = 0$ and $\beta u + x = 0$, so β is non–zero and

$$\alpha^2 - 1 = w^* P_0 w$$

as in the conclusion. ■

We can apply these ideas to get a straightforward minimization problem for the rank two case. Suppose u and v are linearly independent vectors in $\mathbf{R}^n$ and $B = vv^* - uu^*$ is a Hermitian matrix. (Note that all Hermitian matrices with rank 2 for which neither B nor $-B$ is positive semidefinite are of this form.) We need to write $B = P - Q$ where P and Q are positive semidefinite rank one matrices with $P + Q$ having its largest diagonal entry as small as possible. So

$$P = (\alpha u + \beta v)(\alpha u + \beta v)^* \quad \text{and} \quad Q = (\gamma u + \delta v)(\gamma u + \delta v)^*$$

for some scalars $\alpha, \beta, \gamma, \delta$. The independence of u and v implies uu^*, uv^*, vu^*, and vv^* are linearly independent matrices so equating $vv^* - uu^*$ and

$P - Q$ gives $\alpha^2 - \gamma^2 = 1$, $\alpha\beta - \gamma\delta = 0$, and $\beta^2 - \delta^2 = -1$. Without loss of generality, this means

$$P = (\alpha u + \beta v)(\alpha u + \beta v)^* \quad \text{and} \quad Q = (\beta u + \alpha v)(\beta u + \alpha v)^*$$

where $\alpha^2 - \beta^2 = 1$ and $\alpha > 0$. Now

$$P + Q = (\alpha^2 + \beta^2)(vv^* + uu^*) + 2\alpha\beta(vu^* + uv^*)$$

Setting $t = 2\alpha\beta$, we see $\alpha^2 + \beta^2 = \sqrt{1 + t^2}$. Thus

$$K_B = \min_t \max\left(\operatorname{diag}\left(\sqrt{1+t^2}(vv^* + uu^*) + t(vu^* + uv^*)\right)\right)$$

If u_j and v_j are the j^{th} components of u and v respectively, letting $f_j(t) = \sqrt{1+t^2}(v_j^2 + u_j^2) + 2tv_ju_j$ gives

$$K_B = \min_t \max_j f_j(t)$$

Since each of the f_j is a convex function,

$$K_B = f_{j_0}(t_0)$$

where either f_{j_0} has a minimum at t_0 and $f_{j_0}(t_0) \geq f_j(t_0)$ for $j = 1, \cdots, n$ or there are $i_0 \neq j_0$ so that $f_{i_0}(t_0) = f_{j_0}(t_0) \geq f_j(t_0)$ for $j = 1, \cdots, n$. That is, to find K_B, we need only check the minima of the f_j and the points where the graphs cross.

The values of the functions where the graphs cross do not seem to have convenient expressions, but it turns out that the minimum value of f_j is $|b_{jj}|$ and the value of f_k at this point is $(2b_{jk}^2 - b_{kk}b_{jj})/|b_{jj}|$. Since K_B is at least as large as $|b_{jj}|$ for each j, we have another proof of Theorem 9 of [4] for this case.

Corollary 3.3. *Let $B = (b_{ij})$ be a real Hermitian $n \times n$ matrix with rank 2. If $b_{jj} > 0$ and b_{jj} is the largest diagonal entry of B, then the following are equivalent:*

(i) $K_B = b_{jj}$.

(ii) Every 2×2 principal submatrix containing b_{jj} has Hadamard multiplier norm b_{jj}.

(iii) For $1 \leq k \leq n$,

$$b_{jj}^2 + b_{jj}b_{kk} - 2b_{jk}^2 \geq 0$$

4. NUMERICAL RESULTS

A Matlab program `FindKB` that carries out Watson's iterative algorithm, finds a factorization of the type discussed in Section 2 along with the unitaries V and W and vectors x and y, and computes upper and lower bounds for K_B can be obtained (`cowen@math.purdue.edu`) from the first author. The inputs are the matrix B whose Hadamard multiplier norm is to be found, a number of iterations to be done, and a starting unitary matrix. The lower bound for K_B is computed as in the description of the Watson algorithm. The upper bound $c(R)c(S)$ for K_B given by Haagerup's Theorem is computed using Equation (3) to find an S and $R = (S^*)^{-1}B$ to find the corresponding factor R. The algorithm fails if the matrix X used in the computation of S is not invertible.

All the computations reported here were done with Matlab [10] version 4.1 on a Macintosh Quadra 950 and the starting unitary was generated by the program (by QR–factorization of a random matrix) rather than being user specified. This algorithm produces the following results for the triangular truncation matrices, that is, for the $n \times n$ matrices T for which $T_{ij} = 1$ when $i + j \leq n + 1$ and $T_{ij} = 0$ when $i + j > n + 1$. In each case, the lower bound from Watson's algorithm and the upper bound from the resulting factorization differ by less than 2×10^{-14}. The lower bound is given in the table below for K_T for $n = 2, \cdots, 50$.

One of the convenient features of the triangular truncation matrices is that the convergents computed in Watson's algorithm give a factorization of T for which the Haagerup upper bound is essentially equal to the lower bound. We believe this happens because there are no submatrices of T with the same Hadamard multiplier norm. The same is true of McEachin's matrices [11] the norms of some of which were computed in [5].

It might be noted that the stopping criterion written into the algorithm used by the authors is a number, `Kstop`, of iterations. It would be easy to modify the stopping criterion to be, for example, that the difference between two successive iterates of the lower bound be smaller than a given epsilon. Some caution needs to be exercised, however, as the various output variables converge at quite different rates. For example, in computing the Hadamard multiplier norm for the 50×50 triangular truncation matrix, the lower bound estimate is correct to 14 decimal places after `Kstop = 60` iterations, whereas, at that point, the upper bound estimate is only correct to 7 decimal places. To get the upper bound estimate to be correct to 14 decimal places requires about `Kstop = 130` iterations.

Table 1

n	K_T	n	K_T

		26	1.84464503202383
2	1.15470053837925	27	1.85566122168631
3	1.25319726474218	28	1.86629172019688
4	1.32612266648237	29	1.87656283483510
5	1.38423194069267	30	1.88649825965243
6	1.43262139077559	31	1.89611941132500
7	1.47412419160510	32	1.90544571264071
8	1.51048356114476	33	1.91449483313437
9	1.54285069801945	34	1.92328289442527
10	1.57202669921690	35	1.93182464630082
11	1.59859201851078	36	1.94013361841505
12	1.62298098962732	37	1.94822225154975
13	1.64552725432741	38	1.95610201165968
14	1.66649276502696	39	1.96378348934513
15	1.68608702122257	40	1.97127648693326
16	1.70448024411923	41	1.97859009497756
17	1.72181264978498	42	1.98573275968352
18	1.73820113209454	43	1.99271234252354
19	1.75374417916747	44	1.99953617310344
20	1.76852555639746	45	2.00621109617797
21	1.78261711023281	46	2.01274351357634
22	1.79608093350023	47	2.01913942168573
23	1.80897105939639	48	2.02540444504639
24	1.82133480230775	49	2.03154386653308
25	1.83321383040637	50	2.03756265453114

If the matrix X needed in the computation of S is not invertible, a multistep procedure based on Theorem 2.1 is needed. As an example of the techniques needed to handle a more complicated case, consider the following 5×5 matrix. The steps indicated below outline the Matlab program `Multistep` which can also be obtained from the first author.

```
B =
   -1.6688  -0.2358   0.9445   1.1862   0.9198
    1.5766   1.9792  -1.0445   0.6810   0.1804
   -0.9223   0.7425  -2.1884   0.8284   0.3286
    0.2944   0.0148   0.9386   0.2742  -0.0324
   -0.4420   0.0872   0.2947   0.2091  -0.2410
```

After convergence, the lower bound for K_B is 2.2835 but the vectors x and y in the algorithm have zeros in their last two components, so no upper bound is computed. However, following the ideas of Theorem 2.1, we suspect the upper left 3×3 submatrix of B has the same Hadamard multiplier norm. Factoring this submatrix and continuing as in Theorem 2.1, we compute

```
S11 =                                   S12 =
    0.5259  -0.7162  -1.4586               0.5972   0.1564
    1.0427  -1.2669   0.3949              -0.1623   0.2531
   -0.9590  -0.4069   0.0091              -0.3096  -0.3919

R11 =                                   R12 =
    0.2369  -0.7355   1.5046              -0.5261  -0.1716
   -1.4669  -0.8122   0.0187               0.1852   0.2171
    0.2752  -1.0406  -0.1395              -1.3241  -0.8172

PHI =                                   PSI =
    0.7444        0                        2.1357        0
         0   0.6999                             0   0.8061
```

and

```
Z =
    0.3315  -0.0886
   -0.4101  -0.3325
```

Using `FindKB` on this matrix, we find $K_Z = 0.4375$, which is good since Theorem 2.1 requires that $K_Z \leq 1$ to proceed. The corresponding optimal factorization (for the upper bound from the Haagerup theorem) is

```
Sz =                             Rz =
    0.4164  -0.6511                 0.6509   0.4190
   -0.5140  -0.1165                -0.1177   0.5118
```

We now compute S_{22} and R_{22} from `Sz, Rz, PHI` and `PSI` to use in the factorization of B. In this case, we get

```
S =
   0.5259  -0.7162  -1.4586   0.5972   0.1564
   1.0427  -1.2669   0.3949  -0.1623   0.2531
  -0.9590  -0.4069   0.0091  -0.3096  -0.3919
        0        0        0   0.5593  -0.9303
        0        0        0  -0.6905  -0.1665

R =
   0.2369  -0.7355   1.5046  -0.5261  -0.1716
  -1.4668  -0.8122   0.0187   0.1852   0.2171
```

```
 0.2752  -1.0406  -0.1395  -1.3241  -0.8172
      0        0        0   0.3048   0.5199
      0        0        0  -0.0551   0.6349
```

It can be checked that $B = S^*R$ and that the norms of the columns of S and R are

```
cS =
  1.5111   1.5111   1.5111   1.1262   1.0655

cR =
  1.5111   1.5111   1.5111   1.4698   1.1907
```

so $K_B \leq c(S)c(R) = 2.2835$ which agrees with lower bound found earlier.

Although, as we noted above, we have been unable to prove that the algorithm always works, it is self checking; that is, if the upper and lower bound estimate of K_B are acceptably close, we know the algorithm has worked in this case. In a typical problem, one would apply `FindKB` to a matrix. If the upper and lower estimates of the Hadamard multiplier norm are significantly different, then the non–zero components of x and y indicate the rows and columns, respectively, of the matrix that comprise the important submatrix $\widehat{B}$. Then, `Multistep` can be used to find the resulting factorization and upper bound estimate of $\widehat{B}$. If the lower right corner produced by the algorithm is also Hadamard reducible, the process must be iterated to find the eventual factorization.

Our experience is that Watson's algorithm is quite robust. In addition to the structured triangular truncation and McEachin matrices, we generated 165 random 10×10 matrices using Matlab's `randn(10,10)` command; that is, the entries of the matrices were chosen from a normal distribution with mean 0 and standard deviation 1. In each case, the Watson algorithm converged to what was shown by the subsequent upper bound estimates to be the Hadamard multiplier norm of the matrix. Random starting unitaries were used and 500 iterations of the algorithm were sufficient to give agreement of the upper and lower estimates for K_B of 7 to 14 decimal places in the Hadamard irreducible cases. (Watson's algorithm converges much faster than that to the lower bound estimate, but a large number of iterations are required to get a factorization that gives an acceptably small upper bound.) Of the 165 cases, 90 matrices were Hadamard irreducible, 37 had a 9×9 submatrix with the same Hadamard multiplier norm, 12 had an 8×8 submatrix, 1 each had 7×7, 6×6, and 5×5 submatrices, 5 had 2×2 submatrices, and in 18 cases the Hadamard multiplier norm was the absolute value of the largest entry, i.e. a 1×1 submatrix.

REFERENCES

1 T. Ando, R. A. Horn, and C. R. Johnson, The singular values of a Hadamard product: a basic inequality, *Linear Multilinear Alg.* 21:345–365(1987).

2 T. Ando and K. Okubo, Induced norms of the Schur multiplier operator, *Linear Alg. Appl.* 147:181–199(1991).

3 J. R. Angelos, C. C. Cowen, and S. K. Narayan, Triangular truncation and finding the norm of a Hadamard multiplier, *Linear Alg. Appl.* 170:117–135(1992).

4 C. C. Cowen, K. E. Debro, and P. D. Sepanski, Geometry and the norms of Hadamard multipliers, *Linear Alg. Appl.*, to appear.

5 C. C. Cowen, M. A. Dritschel, and R. C. Penney, Norms of Hadamard multipliers, *SIAM J. Matrix Anal. Appl.* 15:313–320(1994).

6 U. Haagerup, Decompositions of completely bounded maps on operator algebras, preprint.

7 R. A. Horn, The Hadamard product, *Proc. Symposia Appl. Math.* 40:87–169(1990).

8 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

9 R. Mathias, Matrix completions, norms and Hadamard products, *Proc. Amer. Math. Soc.* 117:905–918(1993).

10 The MathWorks, *MATLAB*, Version 4.1, The MathWorks, South Natick, Massachusetts, 1993.

11 R. V. McEachin, A sharp estimate for an operator inequality, *Proc. Amer. Math. Soc.* 115:161–165(1992).

12 V. I. Paulsen, *Completely Bounded Maps and Dilations*, Pitman Research Notes in Mathematics #146, Wiley, New York, 1986.

13 V. I. Paulsen, S. C. Power, and R. R. Smith, Schur products and matrix completions, *J. Functional Analysis* 85:151–178(1989).

14 I. Schur, Bemerkungen zur theorie de beschränkten bilinearformen mit unendlich vielen veränderlichen, *J. reine angew. Math.* 140:1–28(1911).

15 G. A. Watson, Estimating Hadamard multiplier norms, with application to triangular truncation, *Linear Alg. Appl.*, to appear.