Detecting and Solving Hyperbolic Quadratic Eigenvalue Problems

Guo, Chun-Hua and Higham, Nicholas J. and Tisseur, Françoise

2007

MIMS EPrint: **2007.117**

ISSN 1749-9097

DETECTING AND SOLVING HYPERBOLIC QUADRATIC EIGENVALUE PROBLEMS*

CHUN-HUA GUO[†], NICHOLAS J. HIGHAM[‡], AND FRANÇOISE TISSEUR[‡]

Abstract. Hyperbolic quadratic matrix polynomials $Q(\lambda) = \lambda^2 A + \lambda B + C$ are an important class of Hermitian matrix polynomials with real eigenvalues, among which the overdamped quadratics are those with nonpositive eigenvalues. Neither the definition of overdamped nor any of the standard characterizations provides an efficient way to test if a given Q has this property. We show that a quadratically convergent matrix iteration based on cyclic reduction, previously studied by Guo and Lancaster, provides necessary and sufficient conditions for Q to be overdamped. For weakly overdamped Q the iteration is shown to be generically linearly convergent with constant at worst 1/2, which implies that the convergence of the iteration is reasonably fast in almost all cases of practical interest. We show that the matrix iteration can be implemented in such a way that when overdamping is detected a scalar $\mu < 0$ is provided that lies in the gap between the n largest and n smallest eigenvalues of the $n \times n$ quadratic eigenvalue problem (QEP) $Q(\lambda)x = 0$. Once such a μ is known, the QEP can be solved by linearizing to a definite pencil that can be reduced using already available Cholesky factorizations to a standard Hermitian eigenproblem. By incorporating an initial preprocessing stage that shifts a hyperbolic Q so that it is overdamped, we obtain an efficient algorithm that identifies and solves a hyperbolic or overdamped QEP maintaining symmetry throughout and guaranteeing real computed eigenvalues.

Key words. quadratic eigenvalue problem, hyperbolic, overdamped, weakly overdamped, quadratic matrix polynomial, quadratic matrix equation, solvent, cyclic reduction, doubling algorithm

AMS subject classifications. 15A18, 15A24, 65F15, 65F30

1. Introduction.

The quadratic eigenvalue problem (QEP) is to find scalars λ and nonzero vectors x and y satisfying $Q(\lambda)x = 0$ and $y^*Q(\lambda) = 0$, where

$$Q(\lambda) = \lambda^2 A + \lambda B + C, \quad A, B, C \in \mathbb{C}^{n\times n} \tag{1.1}$$

is a quadratic matrix polynomial. The vectors x and y are right and left eigenvectors corresponding to the eigenvalue λ. The many applications of the QEP, as well as its theory and algorithms for solving it, are surveyed by Tisseur and Meerbergen [27].

Our interest in this work is in Hermitian quadratic matrix polynomials—those with Hermitian A, B, and C—and more specifically those that are *hyperbolic*. Hyperbolic quadratics, and the subclass of overdamped quadratics, are defined as follows. For Hermitian X and Y we write $X > Y$ ($X \geq Y$) if $X - Y$ is positive definite (positive semidefinite).

DEFINITION 1.1. *$Q(\lambda)$ is hyperbolic if A, B, and C are Hermitian, $A > 0$, and*

$$(x^*Bx)^2 > 4(x^*Ax)(x^*Cx) \quad \textit{for all nonzero } x \in \mathbb{C}^n. \tag{1.2}$$

*Revision of August 19, 2008. This work was supported by a Royal Society-Wolfson Research Merit Award to the second author.

[†]Department of Mathematics and Statistics, University of Regina, Regina, SK S4S 0A2, Canada (chguo@math.uregina.ca, http://www.math.uregina.ca/~chguo/). The research of this author was supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada.

[‡]School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (higham@ma.man.ac.uk, http://www.ma.man.ac.uk/~higham/, ftisseur@ma.man.ac.uk, http://www.ma.man.ac.uk/~ftisseur/). The work of both authors was supported by Engineering and Physical Sciences Research Council grant EP/D079403.

Definition 1.2. *$Q(\lambda)$ is overdamped if it is hyperbolic with $B > 0$ and $C \geq 0$.*

Overdamped quadratics arise in overdamped systems in structural mechanics [22, Sec. 7.6].

Any eigenpair of Q satisfies $x^*Q(\lambda)x = 0$ and hence

$$\lambda = \frac{-x^*Bx \pm \sqrt{(x^*Bx)^2 - 4(x^*Ax)(x^*Cx)}}{2\,x^*Ax}. \tag{1.3}$$

Therefore the eigenvalues of a hyperbolic Q are real and those of an overdamped Q are real and nonpositive. Both classes of quadratics have other important spectral properties, which we summarize in section 2.

We have two aims. The first is to devise an efficient and reliable numerical test for hyperbolicity or overdamping of a given Hermitian quadratic. The second aim is to build upon an affirmative test result an efficient algorithm for solving the QEP that exploits hyperbolicity, and in particular that guarantees real computed eigenvalues in floating point arithmetic.

Part of the motivation for testing overdamping concerns the stability of gyroscopic systems. It is known that a gyroscopic system $G(\lambda) = \lambda^2 A_g + \lambda B_g + C_g$ with $A_g, C_g > 0$ and B_g Hermitian indefinite and nonsingular is stable whenever the quadratic $Q_g(\lambda) = \lambda^2 A_g + \lambda |B_g| + C_g$ is overdamped [9]. Here $|B_g|$ is the Hermitian positive definite square root of B_g^2 (i.e., the Hermitian polar factor of the Hermitian matrix B_g) [12].

Checking the hyperbolicity condition (1.2) is a nontrivial task and plausible sufficient conditions for hyperbolicity may be incorrect. For example, it is claimed in [21] that when $A = I$, $B > 0$, and $C \geq 0$, Q is hyperbolic if $B > 2C^{1/2}$. That this claim is false has been shown by Barkwell and Lancaster [1].

Guo and Lancaster [9] propose testing overdamping by using a matrix iteration based on cyclic reduction to compute two solvents (solutions) of the quadratic matrix equation

$$Q(X) = AX^2 + BX + C = 0 \tag{1.4}$$

and then computing an extremal eigenvalue of each solvent. A definiteness test on Q evaluated at the average of the two extremal eigenvalues finally determines whether Q is overdamped. We show that the same iteration can be used to test overdamping in a much more efficient way that does not necessarily require the iteration to be run to convergence, even for a positive test result. Our test is based on a more complete understanding of the behavior of the matrix iteration, developed in section 3.

In section 4 we extend the convergence analysis to weakly overdamped quadratics, for which the strict inequality in (1.2) is replaced by a weak inequality ($\geq$). The key idea is to use an interpretation of the matrix iteration as a doubling algorithm. We show that for weakly overdamped Q with equality in (1.2) for some nonzero x, the iteration is linearly convergent with constant at worst 1/2 in the generic case. A reasonable speed of convergence can therefore be expected in almost all practically important cases.

In section 5 we turn to algorithmic matters. We first show how a hyperbolic Q can be shifted to make it overdamped. Then we specify our test for overdamping, which requires only the building blocks of Cholesky factorization, matrix multiplication, and solution of triangular systems. We then show how after a successful test the eigensystem of an overdamped Q can be efficiently computed in a way that exploits the symmetry and definiteness and guarantees real computed eigenvalues.

Veselić [28] and Higham, Tisseur, and Van Dooren [19] have previously shown that every hyperbolic quadratic can be reformulated as a definite pencil $L(\lambda) = \lambda X + Y$ of twice the dimension, and this connection is explored in detail and in more generality by Higham, Mackey, and Tisseur [16]. However, the algorithm developed here is the first practical procedure for arranging that X or Y is a definite matrix and hence allowing symmetry and definiteness to be fully exploited.

Section 6 concludes the paper with a numerical experiment that provides further insight into the theory and algorithms.

2. Preliminaries.

We first recall the definition of a definite pencil.

DEFINITION 2.1. *A Hermitian pencil $L(\lambda) = \lambda X + Y$ is definite (or equivalently, the matrices X, Y form a definite pair) if $(z^*Xz)^2 + (z^*Yz)^2 > 0$ for all nonzero $z \in \mathbb{C}^n$.*

Definite pairs have the desirable properties that they are simultaneously diagonalizable under congruence and, in the associated eigenproblem $L(\lambda)x = 0$, the eigenvalues are real and semisimple[1].

The next result gives three conditions each equivalent to the condition (1.2) in the definition of hyperbolic quadratic.

THEOREM 2.2. *Let the $n \times n$ quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ be Hermitian with $A > 0$ and let*

$$\gamma = \min_{\|x\|_2=1} [(x^*Bx)^2 - 4(x^*Ax)(x^*Cx)]. \tag{2.1}$$

The following statements are equivalent.

(a) *Q is hyperbolic.*

(b) *$\gamma > 0$.*

(c) *$x^*Q(\lambda)x = 0$ has two distinct real zeros for all nonzero $x \in \mathbb{C}^n$.*

(d) *$Q(\mu) < 0$ for some $\mu \in \mathbb{R}$.*

Proof. (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) is immediate. (c) $\Leftrightarrow$ (d) follows from Markus [25, Lem. 31.15]. $\square$

Hyperbolic quadratics have many interesting properties [25, Sec. 31].

THEOREM 2.3. *Let the $n \times n$ quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ be hyperbolic.*

(a) *The $2n$ eigenvalues of $Q(\lambda)$ are all real and semisimple.*

(b) *There is a gap between the n largest and n smallest eigenvalues, that is, the eigenvalues can be ordered $\lambda_1 \geq \cdots \geq \lambda_n > \lambda_{n+1} \geq \cdots \geq \lambda_{2n}$.*

(c) *$Q(\mu) < 0$ for all $\mu \in (\lambda_{n+1}, \lambda_n)$ and $Q(\mu) > 0$ for all $\mu \in (-\infty, \lambda_{2n}) \cup (\lambda_1, \infty)$.*

(d) *There are n linearly independent eigenvectors associated with the n largest eigenvalues and likewise for the n smallest eigenvalues.*

(e) *The quadratic matrix equation $Q(X) = 0$ in (1.4) has a solvent $S^{(1)}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and a solvent $S^{(2)}$ with eigenvalues $\lambda_{n+1}, \ldots, \lambda_{2n}$. Moreover,*

$$Q(\lambda) = \big(\lambda I - S^{(2)^*}\big) A (\lambda I - S^{(1)}) = \big(\lambda I - S^{(1)^*}\big) A (\lambda I - S^{(2)}).$$

The n largest eigenvalues of a hyperbolic quadratic are called the primary eigenvalues and the n smallest eigenvalues are the secondary eigenvalues. The solvents $S^{(1)}$ and $S^{(2)}$ having as their eigenvalues the primary eigenvalues and the secondary eigenvalues, respectively, are referred to as the *primary* and *secondary solvents*.

[1] An eigenvalue of a matrix polynomial $P(\lambda) = \sum_{k=0}^{\ell} \lambda^k P_k$ is semisimple if it appears only in 1×1 Jordan blocks in a Jordan triple for P [7].

Hyperbolicity can also be defined for matrix polynomials P of arbitrary degree [25, Sec. 31]. The notion has recently been extended in [16] by replacing the assumption of a positive definite leading coefficient matrix with $P(\omega) > 0$ for some $\omega \in \mathbb{R} \cup \{\infty\}$.

The next result gives some characterizations of an overdamped quadratic. First, we need a simple lemma.

LEMMA 2.4. *Let $Q(\lambda) = \lambda^2 A + \lambda B + C$ be Hermitian and let $\mu > 0$. Then $Q(-\mu) < 0$ if and only if $B > \mu A + \mu^{-1} C$.*

Proof. Immediate from $Q(-\mu) = \mu^2 A - \mu B + C < 0 \Leftrightarrow \mu A - B + \mu^{-1} C < 0$. ∎

THEOREM 2.5. *Let $Q(\lambda) = \lambda^2 A + \lambda B + C$ be Hermitian with $A > 0$. Then the following statements are equivalent.*

(a) *$Q(\lambda)$ is overdamped.*

(b) *$Q(\lambda)$ is hyperbolic and all its eigenvalues are real and nonpositive.*

(c) *$B > 0$, $C \geq 0$, and $B > \mu A + \mu^{-1} C$ for some $\mu > 0$.*

Proof. (a) $\Leftrightarrow$ (b) is proved in [9, Thm. 5]. (b) $\Rightarrow$ (c): by Theorem 2.3 (c), $Q(\widetilde{\mu}) < 0$ for some $\widetilde{\mu} < 0$; (c) follows on invoking Lemma 2.4. (c) $\Rightarrow$ (a): $B > \mu A + \mu^{-1} C$ with $\mu > 0$ implies $Q(-\mu) < 0$ by Lemma 2.4, which implies Q is hyperbolic by Theorem 2.2 (d) and hence overdamped since $B > 0$ and $C \geq 0$. ∎

It follows from (b) in Theorem 2.5 that if we know an upper bound, say θ, on the largest eigenvalue λ_1 of a hyperbolic quadratic Q then, with $\lambda = \mu + \theta$, the quadratic Q_θ defined by

$$
\begin{aligned}
Q(\lambda) = Q(\mu + \theta) &= \mu^2 A + \mu(B + 2\theta A) + C + \theta B + \theta^2 A \\
&= \mu^2 A_\theta + \mu B_\theta + C_\theta \\
&=: Q_\theta(\mu)
\end{aligned} \tag{2.2}
$$

is overdamped. Thus any hyperbolic quadratic can be transformed into an overdamped quadratic by an appropriate shifting of the eigenvalues. Hence for the purposes of testing hyperbolicity and overdamping it suffices to consider overdamping. We make this restriction in the next two sections and consider in section 5 how to implement the shifting in practice.

3. An iteration for testing overdamping.

Suppose we have a Hermitian quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$, where we assume throughout this section that $A > 0$, $B > 0$, and $C \geq 0$. The challenge is how to test the hyperbolicity (or equivalently, overdamping) condition (1.2) or, equivalently, condition (c) in Theorem 2.5.

The primary and secondary solvents $S^{(1)}$ and $S^{(2)}$ of an overdamped quadratic can be found efficiently by applying an iteration based on cyclic reduction [2], [9]. The iteration is

$$
\begin{aligned}
&S_0 = B, \quad A_0 = A, \quad B_0 = B, \quad C_0 = C, \\
&S_{k+1} = S_k - A_k B_k^{-1} C_k, \\
&A_{k+1} = A_k B_k^{-1} A_k, \\
&B_{k+1} = B_k - A_k B_k^{-1} C_k - C_k B_k^{-1} A_k, \\
&C_{k+1} = C_k B_k^{-1} C_k.
\end{aligned} \tag{3.1}
$$

The next theorem summarizes properties of the iteration proved in [9, Lem. 6, Thm. 7 and proof].

THEOREM 3.1. *Let $Q(\lambda) = \lambda^2 A + \lambda B + C$ be an $n \times n$ overdamped quadratic with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n > \lambda_{n+1} \geq \cdots \geq \lambda_{2n}$. Consider iteration* (3.1) *and any matrix norm $\|\cdot\|$.*

(a) *The iterates satisfy $A_k > 0$, $C_k \geq 0$, $B_k > 0$ for all $k \geq 0$.*

(b) *$\|A_k\|\,\|C_k\|$ converges quadratically to zero with*

$$\limsup_{k\to\infty} \sqrt[2^k]{\|A_k\|\,\|C_k\|} \leq \frac{\lambda_n}{\lambda_{n+1}} < 1.$$

(c) *S_k converges quadratically to a nonsingular matrix $\widehat{S}$ with*

$$\limsup_{k\to\infty} \sqrt[2^k]{\|S_k - \widehat{S}\|} \leq \frac{\lambda_n}{\lambda_{n+1}} < 1. \tag{3.2}$$

(d) *The primary and secondary solvents of $Q(X)$, $S^{(1)}$ and $S^{(2)}$ respectively, are given by*

$$S^{(1)} = -\widehat{S}^{-1}C, \quad S^{(2)} = -A^{-1}\widehat{S}^{*}. \tag{3.3}$$

The next lemma reveals a crucial property of iteration (3.1) for overdamped quadratics. The "only if" part of the result is [9, Lem. 6].

LEMMA 3.2. *Let $\mu > 0$ and assume $A_k > 0$ and $C_k \geq 0$. In* (3.1)*, $B_k > \mu^{2^k} A_k + \mu^{-2^k} C_k$ if and only if $A_{k+1} > 0$, $C_{k+1} \geq 0$, and $B_{k+1} > \mu^{2^{k+1}} A_{k+1} + \mu^{-2^{k+1}} C_{k+1}$.*

Proof. "$\Rightarrow$": we have

$$\begin{aligned} B_{k+1} &= B_k - A_k B_k^{-1} C_k - C_k B_k^{-1} A_k \\ &= B_k - \big(\mu^{2^k} A_k + \mu^{-2^k} C_k\big) B_k^{-1} \big(\mu^{2^k} A_k + \mu^{-2^k} C_k\big) \\ &\qquad + \mu^{2^{k+1}} A_k B_k^{-1} A_k + \mu^{-2^{k+1}} C_k B_k^{-1} C_k \\ &> \mu^{2^{k+1}} A_k B_k^{-1} A_k + \mu^{-2^{k+1}} C_k B_k^{-1} C_k, \end{aligned}$$

where we have used the fact that $X - YX^{-1}Y > Y - YY^{-1}Y = 0$ when $X > Y > 0$. Clearly, $A_{k+1} > 0$ and $C_{k+1} \geq 0$ since $B_k^{-1} > 0$.

"$\Leftarrow$": as in the first part we have

$$B_{k+1} = B_k - F_k B_k^{-1} F_k + F_{k+1}, \tag{3.4}$$

where $F_k = \mu^{2^k} A_k + \mu^{-2^k} C_k$. Now if $B_{k+1} > \mu^{2^{k+1}} A_{k+1} + \mu^{-2^{k+1}} C_{k+1} = F_{k+1}$ then (3.4) gives $B_k - F_k B_k^{-1} F_k > 0$. Note that $B_k - F_k B_k^{-1} F_k$ is the Schur complement of $B_k > 0$ in

$$T = \begin{bmatrix} B_k & F_k \\ F_k & B_k \end{bmatrix}.$$

So we have $T > 0$ and it follows that $B_k - F_k > 0$ (for example, by looking at the (1,1) block of the congruence $\left[\begin{smallmatrix} I & -I \\ 0 & I \end{smallmatrix}\right] T \left[\begin{smallmatrix} I & 0 \\ -I & I \end{smallmatrix}\right]$). Therefore $B_k > F_k = \mu^{2^k} A_k + \mu^{-2^k} C_k$. $\square$

In view of Theorem 2.5 (c), Lemma 3.2 implies that Q is overdamped *if and only if* any one of the quadratics

$$Q_k(\lambda) = \lambda^2 A_k + \lambda B_k + C_k \tag{3.5}$$

generated during the iteration is overdamped, assuming that $A_k > 0$ and $C_k \geq 0$ for all k. Note that the latter assumption holds if $B_k > 0$ for all k.

COROLLARY 3.3. *Let Q be a Hermitian quadratic with $A, B > 0$ and $C \geq 0$. For iteration* (3.1) *and any fixed $m \geq 0$, if $B_k > 0$ for $k = 1\colon m-1$ and*

$$B_m > \mu^{2^m} A_m + \mu^{-2^m} C_m \tag{3.6}$$

for some $\mu > 0$, then $B > \mu A + \mu^{-1} C$ and Q is overdamped.

Intuitively, we can think of the scalars μ^{2^m} and μ^{-2^m} in (3.6) as trying to balance A_m and C_m. This suggests that (3.6) could be replaced by $B_m > \widetilde{A}_m + \widetilde{C}_m$ if the iteration is scaled so that $\widetilde{A}_m$ and $\widetilde{C}_m$ are balanced. Normwise balancing is included in the following scaled version of (3.1), introduced in [9]; it generates iterates $\widetilde{A}_k$, B_k (unchanged from (3.1)), and $\widetilde{C}_k$ according to

$$\begin{aligned}
&\alpha_0 = \sqrt{\|C\|/\|A\|}, \\
&\widetilde{A}_0 = \alpha_0 A, \quad B_0 = B, \quad \widetilde{C}_0 = \alpha_0^{-1} C, \\
&A_{k+1} = \widetilde{A}_k B_k^{-1} \widetilde{A}_k, \\
&B_{k+1} = B_k - \widetilde{A}_k B_k^{-1} \widetilde{C}_k - \widetilde{C}_k B_k^{-1} \widetilde{A}_k, \\
&C_{k+1} = \widetilde{C}_k B_k^{-1} \widetilde{C}_k, \\
&\alpha_{k+1} = \sqrt{\|C_{k+1}\|/\|A_{k+1}\|}, \\
&\widetilde{A}_{k+1} = \alpha_{k+1} A_{k+1}, \quad \widetilde{C}_{k+1} = \alpha_{k+1}^{-1} C_{k+1}.
\end{aligned} \tag{3.7}$$

Here we have assumed that $C \neq 0$ (the overdamping condition holds for the trivial case $C = 0$ by (1.2)); thus $\alpha_k > 0$ for each $k \geq 0$. The scaling procedure ensures that $\|\widetilde{A}_k\| = \|\widetilde{C}_k\|$ and $\|\widetilde{A}_k\| \, \|\widetilde{C}_k\| = \|A_k\| \, \|C_k\|$.

The next result describes the behavior of the scaled iteration.

THEOREM 3.4. *A Hermitian quadratic Q with $A, B > 0$ and $0 \neq C \geq 0$ is overdamped if and only if in* (3.7)

$$B_k > 0 \text{ for all } k, \quad \lim_{k\to\infty} \widetilde{A}_k = 0, \quad \lim_{k\to\infty} \widetilde{C}_k = 0, \quad \lim_{k\to\infty} B_k > 0, \tag{3.8}$$

and in this case

$$\limsup_{k\to\infty} \sqrt[2^k]{\|\widetilde{A}_k\|} = \limsup_{k\to\infty} \sqrt[2^k]{\|\widetilde{C}_k\|} \leq \left(\frac{\lambda_n}{\lambda_{n+1}}\right)^{1/2}, \tag{3.9}$$

$$\limsup_{k\to\infty} \sqrt[2^k]{\|B_k - \widehat{B}\|} \leq \frac{\lambda_n}{\lambda_{n+1}} \tag{3.10}$$

with $\widehat{B} = A(S^{(1)} - S^{(2)})$.

Proof. Assume that the conditions in (3.8) hold. Then $B_m > \widetilde{A}_m + \widetilde{C}_m$ for some $m \geq 0$. It is easy to see that the iterates $\widetilde{A}_k$ and $\widetilde{C}_k$ defined in (3.7) are related to A_k and C_k in (3.1) by

$$\widetilde{A}_k = \alpha_0^{2^k} \alpha_1^{2^{k-1}} \dots \alpha_{k-1}^2 \alpha_k A_k, \quad \widetilde{C}_k = \alpha_0^{-2^k} \alpha_1^{-2^{k-1}} \dots \alpha_{k-1}^{-2} \alpha_k^{-1} C_k, \quad k \geq 0.$$

So $B_m > \widetilde{A}_m + \widetilde{C}_m$ implies $B_m > \mu^{2^m} A_m + \mu^{-2^m} C_m$ with $\mu = \alpha_0 \alpha_1^{2^{-1}} \alpha_2^{2^{-2}} \dots \alpha_m^{2^{-m}}$, which implies Q is overdamped by Corollary 3.3.

Now assume the QEP is overdamped. Then, from Theorem 3.1 (a), $B_k > 0$ for each $k \ge 0$, while, since $\|\widetilde{A}_k\| = \|\widetilde{C}_k\| = (\|A_k\|\|C_k\|)^{1/2}$, Theorem 3.1 (b) implies $\lim \widetilde{A}_k = \lim \widetilde{C}_k = 0$ and that (3.9) holds. To show the convergence of B_k, we note that from (3.1), $B_{k+1} = B_k - (S_k - S_{k+1}) - (S_k - S_{k+1})^*$, which implies

$$B_k = B_0 - (S_0 - S_k) - (S_0 - S_k)^* = -B + S_k + S_k^*.$$

In view of (3.2), (3.3) and $B_k > 0$, (3.10) holds with $\widehat{B} = -B + \widehat{S} + \widehat{S}^* = A(S^{(1)} - S^{(2)}) \ge 0$. Since the sequence $\{\|B_k^{-1}\|\}$ is known to be bounded (see the proof of [9, Thm. 7]), we have $\widehat{B} > 0$. $\square$

The next result confirms that μ can be removed from (3.6) for the scaled iteration. It follows readily from Theorem 3.4 and its proof.

Corollary 3.5. *A Hermitian quadratic Q with $A, B > 0$ and $0 \ne C \ge 0$ is overdamped if and only if for some $m \ge 0$, $B_k > 0$ for $k = 1\colon m-1$ in* (3.7) *and $B_m > \widetilde{A}_m + \widetilde{C}_m$.*

The corollary is important for two reasons. First, it provides a basis for an elegant, practical test for overdamping, as definiteness of a matrix is easily tested numerically. Second, in the case of an affirmative test result a μ with $Q(-\mu) < 0$ can be identified, and such a μ is very useful when we go on to solve the QEP, as we will show in section 5.

From a numerical point of view it is preferable to work with the original data as much as possible. The following variant of Corollary 3.5 tests the overdamping condition using the original quadratic Q and will be the basis of the algorithm in section 5. It follows readily from Corollary 3.3 and Theorem 3.4 and its proof.

Corollary 3.6. *A Hermitian quadratic Q with $A, B > 0$ and $0 \ne C \ge 0$ is overdamped if and only if for some $m \ge 0$, $B_k > 0$ for $k = 1\colon m-1$ in* (3.7) *and $Q(-\mu_m) < 0$, where $\mu_m = \alpha_0 \alpha_1^{2^{-1}} \alpha_2^{2^{-2}} \dots \alpha_m^{2^{-m}} > 0$ and the α_k are defined in* (3.7).

Usually, only a few iterations of the cyclic reduction algorithm (3.7) will be necessary. To illustrate, we consider a quadratic $Q(\lambda)$ of dimension $n = 100$ defined by

$$A = I,\ B = \beta \begin{bmatrix} 20 & -10 & & & \\ -10 & 30 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 30 & -10 \\ & & & -10 & 20 \end{bmatrix},\ C = \begin{bmatrix} 15 & -5 & & \\ -5 & 15 & \ddots & \\ & \ddots & \ddots & -5 \\ & & -5 & 15 \end{bmatrix}, \tag{3.11}$$

where $\beta > 0$ is a real parameter. This example, which comes from a damped mass-spring system, is used in [13] with $\beta = 1$. We use the 1-norm in (3.7). Tables 3.1 and 3.2 report the number of iterations required to demonstrate that Q is overdamped, through verification of the conditions in Corollary 3.6, or that it is not overdamped, through generation of a non-positive definite iterate B_m. Note that when Q is "strongly" overdamped and when Q is far from being overdamped, the overdamping condition is shown to hold or not after just a few iterations.

4. Convergence analysis for weakly overdamped quadratics. For the example at the end of section 3 and some $\beta_0 \in (0.5196152422, 0.5196152423)$, the inequality (1.2) holds as a weak inequality with equality attained for some nonzero x. We have seen that the overdamping test requires a very small number of iterations

TABLE 3.1
Number of iterations m to verify that the quadratic defined by (3.11) *is overdamped.*

β	1	0.62	0.61	0.53	0.52	0.5197	0.519616	0.51961525	0.5196152423
m	0	0	1	1	2	3	5	8	12

TABLE 3.2
Number of iterations m to verify that the quadratic defined by (3.11) *is* not *overdamped.*

β	0.36	0.47	0.50	0.51	0.5196	0.519615	0.51961524	0.5196152422
m	1	2	3	4	8	11	15	17

when β is not close to β_0. When $\beta \approx \beta_0$, the number of iterations increases, but is still under 20 in our experiments. The purpose of this section is to explain this behavior by showing that the convergence of iteration (3.1) is reasonably fast even when the QEP is weakly overdamped in the sense defined as follows.

DEFINITION 4.1. *$Q(\lambda)$ is weakly hyperbolic if A, B, and C are Hermitian, $A > 0$, and*

$$\gamma = \min_{\|x\|_2=1} [(x^*Bx)^2 - 4(x^*Ax)(x^*Cx)] \geq 0. \tag{4.1}$$

DEFINITION 4.2. *$Q(\lambda)$ is weakly overdamped if it is weakly hyperbolic with $B > 0$ and $C \geq 0$.*

The eigenvalues of a weakly hyperbolic Q are real and those of a weakly overdamped Q are real and nonpositive. The following result collects further properties of a weakly overdamped quadratic [25, §31].

THEOREM 4.3. *Let $Q(\lambda) = \lambda^2 A + \lambda B + C$ be a weakly overdamped $n \times n$ quadratic.*

(a) *If $\gamma = 0$ in* (4.1) *then $Q(\lambda)$ has $2n$ real eigenvalues that can be ordered $\lambda_1 \geq \cdots \geq \lambda_n = \lambda_{n+1} \geq \cdots \geq \lambda_{2n}$. The partial multiplicities*[2] *of λ_n are at most 2, and the eigenvalues other than λ_n are semisimple.*

(b) $Q(\lambda_n) \leq 0$.

(c) *The quadratic matrix equation $Q(X) = 0$ in* (1.4) *has a solvent $S^{(1)}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and a solvent $S^{(2)}$ with eigenvalues $\lambda_{n+1}, \ldots, \lambda_{2n}$.*

In the overdamped case considered in the previous section, convergence results for the iteration (3.1) are established using matrix identities obtained from the cyclic reduction method. Those identities do not contain enough information about (3.1) to allow a proof of convergence for weakly overdamped quadratics with $\gamma = 0$, for which $\lambda_{n+1} = \lambda_n$. We now study this critical case and thereby obtain a better understanding of the convergence of the iteration for overdamped QEPs with $\lambda_n \approx \lambda_{n+1}$. The next lemma shows that (3.1) remains well defined in the critical case, which is the setting for the rest of this section.

LEMMA 4.4. *For a weakly overdamped quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ with $\gamma = 0$ in* (4.1)*, there is a positive real number μ such that for the iteration* (3.1)*,*

$$A_k > 0, \quad C_k \geq 0, \quad B_k \geq \mu^{2^k} A_k + \mu^{-2^k} C_k \tag{4.2}$$

for all $k \geq 0$.

Proof. We have $\lambda_n \leq \lambda_1 \leq 0$. If $\lambda_n = 0$ then, from Theorem 4.3, $C = Q(\lambda_n) \leq 0$. Since $C \geq 0$ we must have $C = 0$. However, $\gamma > 0$ for the trivial case $C = 0$.

[2] The partial multiplicities of an eigenvalue of Q are the sizes of the Jordan blocks in which it appears in a Jordan triple for Q [7].

Therefore, $\lambda_n < 0$ since $\gamma = 0$. It then follows from $Q(\lambda_n) \le 0$ that $B \ge \mu A + \mu^{-1}C$ for $\mu = -\lambda_n > 0$. The inequalities in (4.2) are then proved inductively using the technique from the proof of the first part of Lemma 3.2. □

Lin and Xu [24] recently showed that Meini's iterations based on cyclic reduction for the matrix equation $X + A^*X^{-1}A = Q$ [26] can also be derived from a structure-preserving doubling algorithm. Following their approach we show that the iteration (3.1) is related to a doubling algorithm and we use this observation to study the convergence of (3.1) for weakly overdamped quadratics. The rate of convergence will be shown to be at least linear with constant 1/2 in the generic case, which is the case where $\lambda_n = \lambda_{n+1}$ is a multiple eigenvalue with partial multiplicities all equal to 2 (that is, λ_n occurs only in 2×2 Jordan blocks). This rate and constant are to be expected in view of the results of Guo in [8].

LEMMA 4.5. *Let* $X = \begin{bmatrix} A & 0 \\ H & -I \end{bmatrix}$ *and* $Y = \begin{bmatrix} G & I \\ C & 0 \end{bmatrix}$ *be block* 2×2 *matrices with* $n \times n$ *blocks. When* $H + G$ *is nonsingular there exist* $2n \times 2n$ *matrices* $\widetilde{X}$ *and* $\widetilde{Y}$ *such that* (a) $\widetilde{X}Y = \widetilde{Y}X$ *and* (b) $\widetilde{X}X$, $\widetilde{Y}Y$ *have the same zero and identity blocks as* X *and* Y, *respectively.*

Proof. Applying block row permutations and block Gaussian elimination to $\begin{bmatrix} X \\ Y \end{bmatrix}$ yields $P\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} U \\ 0 \end{bmatrix}$, where $U = \begin{bmatrix} G & I \\ G+H & 0 \end{bmatrix}$ and

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & 0 & I & 0 \\ 0 & I & I & 0 \\ \hline I & -A(G+H)^{-1} & -A(G+H)^{-1} & 0 \\ 0 & C(G+H)^{-1} & C(G+H)^{-1} & -I \end{array}\right].$$

Since $[\,P_{21} \quad P_{22}\,]\begin{bmatrix} X \\ Y \end{bmatrix} = 0$, the required equality $\widetilde{Y}X = \widetilde{X}Y$ is satisfied with $\widetilde{X} := -P_{22}$ and $\widetilde{Y} := P_{21}$. Furthermore,

$$\widetilde{X}X = \begin{bmatrix} A(G+H)^{-1}A & 0 \\ H - C(G+H)^{-1}A & -I \end{bmatrix}, \qquad \widetilde{Y}Y = \begin{bmatrix} G - A(G+H)^{-1}C & I \\ C(G+H)^{-1}C & 0 \end{bmatrix}. \qquad □$$

Lemma 4.5 and its proof suggest the following recurrence

$$X_{k+1} = \widetilde{X}_k X_k, \quad Y_{k+1} = \widetilde{Y}_k Y_k, \qquad k \ge 0, \tag{4.3}$$

with

$$X_k = \begin{bmatrix} A_k & 0 \\ H_k & -I \end{bmatrix}, \quad Y_k = \begin{bmatrix} G_k & I \\ C_k & 0 \end{bmatrix} \tag{4.4}$$

and

$$\widetilde{X}_k = \begin{bmatrix} A_k(G_k+H_k)^{-1} & 0 \\ -C_k(G_k+H_k)^{-1} & I \end{bmatrix}, \quad \widetilde{Y}_k = \begin{bmatrix} I & -A_k(G_k+H_k)^{-1} \\ 0 & C_k(G_k+H_k)^{-1} \end{bmatrix},$$

which leads to

$$\begin{aligned} A_{k+1} &= A_k(G_k+H_k)^{-1}A_k, \\ G_{k+1} &= G_k - A_k(G_k+H_k)^{-1}C_k, \\ H_{k+1} &= H_k - C_k(G_k+H_k)^{-1}A_k, \\ C_{k+1} &= C_k(G_k+H_k)^{-1}C_k. \end{aligned} \tag{4.5}$$

With

$$A_0 = A, \quad C_0 = C, \quad G_0 = 0, \quad H_0 = B \tag{4.6}$$

the iteration (3.1) is recovered from (4.5) by letting $B_k = G_k + H_k$ and $S_k = H_k^*$. By Lemma 4.4, $B_k > 0$ for all $k \geq 0$. Therefore with the starting matrices (4.6), iteration (4.5) is well defined. Note that X_k in (4.4) is nonsingular for all $k \geq 0$ and, from (4.3) and the property that $\widetilde{X}_k Y_k = \widetilde{Y}_k X_k$,

$$X_{k+1}^{-1} Y_{k+1} = (\widetilde{X}_k X_k)^{-1} \widetilde{Y}_k Y_k = X_k^{-1} \widetilde{X}_k^{-1} \widetilde{Y}_k Y_k = X_k^{-1} Y_k X_k^{-1} Y_k = (X_k^{-1} Y_k)^2. \tag{4.7}$$

It follows from (4.7) that for all $k \geq 0$,

$$X_k^{-1} Y_k = \left(X_0^{-1} Y_0\right)^{2^k}. \tag{4.8}$$

The identity (4.8) is what we need to prove the convergence of (4.5) with (4.6) and hence the convergence of (3.1).

The next result describes the convergence behavior in the generic case.

THEOREM 4.6. *Let $Q(\lambda)$ be weakly overdamped with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n = \lambda_{n+1} \geq \cdots \geq \lambda_{2n}$ and assume that the partial multiplicities of λ_n are all equal to* 2*. Let $S^{(1)}$ and $S^{(2)}$ be the primary and secondary solvents of $Q(X) = 0$, respectively, and assume that λ_n is a semisimple eigenvalue of $S^{(1)}$ and $S^{(2)}$. Then the iterates G_k, H_k, A_k, and C_k defined by* (4.5) *and* (4.6) *satisfy*

$$\limsup_{k\to\infty} \sqrt[k]{\|G_k - AS^{(1)}\|} \leq \frac{1}{2}, \quad \limsup_{k\to\infty} \sqrt[k]{\|H_k + AS^{(2)}\|} \leq \frac{1}{2},$$
$$\limsup_{k\to\infty} \sqrt[k]{\|A_k\| \|C_k\|} \leq \frac{1}{4}.$$

Proof. We start by making the change of variables (or scaling) $\lambda = \mu\theta$, where $\theta = |\lambda_n| > 0$ (see the proof of Lemma 4.4) so that $\mu_n = \mu_{n+1} = -1$ and define $\widehat{Q}(\mu) = \mu^2 \widehat{A} + \mu \widehat{B} + \widehat{C}$ with $(\widehat{A}, \widehat{B}, \widehat{C}) = (\theta A, B, \theta^{-1} C)$. For this triple denote the iterates of (4.5) by $\widehat{A}_k$, $\widehat{G}_k$, $\widehat{H}_k$, $\widehat{C}_k$. It is easy to see that for all $k \geq 0$, $\widehat{G}_k = G_k$, $\widehat{H}_k = H_k$, $\widehat{A}_k = \theta^{2^k} A_k$, and $\widehat{C}_k = \theta^{-2^k} C_k$, so that $\|A_k\| \|C_k\| = \|\widehat{A}_k\| \|\widehat{C}_k\|$. The primary and secondary solvents of $\widehat{A} S^2 + \widehat{B} S + \widehat{C} = 0$ are $\widehat{S}^{(1)} = \theta^{-1} S^{(1)}$ and $\widehat{S}^{(2)} = \theta^{-1} S^{(2)}$, respectively. Note that $\widehat{A} \widehat{S}^{(i)} = A S^{(i)}$, $i = 1, 2$. To avoid notational clutter, we omit the hats on matrices in the rest of the proof.

We now consider the iterations for the block 2×2 matrices X_k and Y_k in (4.4). With $A_0 = A$, $C_0 = C$, $G_0 = 0$, and $H_0 = B$, the pencil

$$\mu X_0 + Y_0 = \mu \begin{bmatrix} A & 0 \\ B & -I_n \end{bmatrix} + \begin{bmatrix} 0 & I_n \\ C & 0 \end{bmatrix} \tag{4.9}$$

is a linearization of $Q(\mu)$ [7]. Hence $-X_0^{-1} Y_0$ and $Q(\mu)$ have the same eigenvalues, with the same partial multiplicities. Suppose there are r 2×2 Jordan blocks associated with eigenvalues equal to $\mu_n = -1$, where $r \geq 1$ by assumption. Rearranging the Jordan canonical form of $X_0^{-1} Y_0$ appropriately yields

$$V^{-1}(X_0^{-1} Y_0) V = \begin{bmatrix} D_2 \oplus I_r & 0 \oplus I_r \\ 0 & D_1 \oplus I_r \end{bmatrix} =: D_V, \tag{4.10}$$

$$W^{-1}(X_0^{-1} Y_0) W = \begin{bmatrix} D_2 \oplus I_r & 0 \\ 0 \oplus I_r & D_1 \oplus I_r \end{bmatrix} =: D_W, \tag{4.11}$$

where V and W are nonsingular, D_1 and D_2 are $(n-r)\times(n-r)$ diagonal matrices containing the (semisimple) eigenvalues less than 1 and greater than 1 in modulus, respectively, and $M \oplus N$ denotes $\left[\begin{smallmatrix} M & 0 \\ 0 & N \end{smallmatrix}\right]$. Now partition V and W as block 2×2 matrices with $n\times n$ blocks:

$$V = \begin{bmatrix} V_1 & V_3 \\ V_2 & V_4 \end{bmatrix}, \qquad W = \begin{bmatrix} W_1 & W_3 \\ W_2 & W_4 \end{bmatrix},$$

and note that from (4.10)–(4.11),

$$X_0^{-1}Y_0 \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}(D_2 \oplus I_r), \qquad X_0^{-1}Y_0 \begin{bmatrix} W_3 \\ W_4 \end{bmatrix} = \begin{bmatrix} W_3 \\ W_4 \end{bmatrix}(D_1 \oplus I_r). \tag{4.12}$$

By Theorem 4.3 and our assumption on $S^{(1)}$ and $S^{(2)}$ there exist nonsingular U_1 and U_2 such that

$$-S^{(1)} = U_1(D_1 \oplus I_r)U_1^{-1}, \qquad -S^{(2)} = U_2(D_2 \oplus I_r)U_2^{-1}. \tag{4.13}$$

Since $S^{(i)}$, $i = 1, 2$, is a solution of $Q(X) = 0$, from (4.9) we obtain

$$X_0^{-1}Y_0 \begin{bmatrix} I_n \\ -AS^{(i)} \end{bmatrix} = \begin{bmatrix} I_n \\ -AS^{(i)} \end{bmatrix}(-S^{(i)}), \quad i = 1, 2.$$

On comparing with the invariant subspaces in (4.12) and using (4.13) we deduce that

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} U_2 \\ -AS^{(2)}U_2 \end{bmatrix} Z_1, \qquad \begin{bmatrix} W_3 \\ W_4 \end{bmatrix} = \begin{bmatrix} U_1 \\ -AS^{(1)}U_1 \end{bmatrix} Z_2,$$

with Z_1 and Z_2 nonsingular, where we have also used the fact that there are exactly r eigenvectors of $X_0^{-1}Y_0$ corresponding to the eigenvalue 1. Hence V_1 and W_3 are nonsingular and

$$-AS^{(2)} = V_2V_1^{-1}, \quad -AS^{(1)} = W_4W_3^{-1}. \tag{4.14}$$

By (4.8)–(4.11) we have $V^{-1}(X_k^{-1}Y_k)V = D_V^{2^k}$ and $W^{-1}(X_k^{-1}Y_k)W = D_W^{2^k}$, so that

$$Y_kV = X_kVD_V^{2^k}, \quad Y_kW = X_kWD_W^{2^k}. \tag{4.15}$$

On equating blocks using (4.4) this yields

$$G_kV_1 + V_2 = A_kV_1(D_2^{2^k} \oplus I_r), \tag{4.16}$$
$$G_kV_3 + V_4 = A_kV_1(0 \oplus 2^kI_r) + A_kV_3(D_1^{2^k} \oplus I_r), \tag{4.17}$$
$$C_kV_1 = (H_kV_1 - V_2)(D_2^{2^k} \oplus I_r), \tag{4.18}$$
$$C_kV_3 = (H_kV_1 - V_2)(0 \oplus 2^kI_r) + (H_kV_3 - V_4)(D_1^{2^k} \oplus I_r) \tag{4.19}$$

and

$$G_kW_1 + W_2 = A_kW_1(D_2^{2^k} \oplus I_r) + A_kW_3(0 \oplus 2^kI_r), \tag{4.20}$$
$$G_kW_3 + W_4 = A_kW_3(D_1^{2^k} \oplus I_r), \tag{4.21}$$
$$C_kW_1 = (H_kW_1 - W_2)(D_2^{2^k} \oplus I_r) + (H_kW_3 - W_4)(0 \oplus 2^kI_r), \tag{4.22}$$
$$C_kW_3 = (H_kW_3 - W_4)(D_1^{2^k} \oplus I_r). \tag{4.23}$$

By (4.22) and (4.23) we have

$$C_k(W_3 - W_1(0 \oplus 2^{-k} I_r)) = (H_k W_3 - W_4)(D_1^{2^k} \oplus 0) - (H_k W_1 - W_2)(0 \oplus 2^{-k} I_r). \tag{4.24}$$

By (4.18) we have

$$H_k = V_2 V_1^{-1} + C_k V_1 (D_2^{-2^k} \oplus I_r) V_1^{-1}. \tag{4.25}$$

Inserting (4.25) in (4.24), we obtain

$$\begin{aligned} &C_k\Big(W_3 - W_1(0 \oplus 2^{-k} I_r) - V_1(D_2^{-2^k} \oplus I_r)V_1^{-1}\big(W_3(D_1^{2^k} \oplus 0) - W_1(0 \oplus 2^{-k} I_r)\big)\Big) \\ &\qquad = (V_2 V_1^{-1} W_3 - W_4)(D_1^{2^k} \oplus 0) - (V_2 V_1^{-1} W_1 - W_2)(0 \oplus 2^{-k} I_r), \end{aligned}$$

from which it follows, since D_1 and D_2 are diagonal with diagonal elements of magnitude less than 1 and greater than 1, respectively, that

$$C_k = O(2^{-k}); \tag{4.26}$$

the latter notation means that $\|C_k\| = O(2^{-k})$. It then follows from (4.25) and (4.14) that

$$H_k + AS^{(2)} = H_k - V_2 V_1^{-1} = O(2^{-k}). \tag{4.27}$$

By (4.20) and (4.21),

$$G_k W_3 + W_4 - (G_k W_1 + W_2)(0 \oplus 2^{-k} I_r) = A_k(W_3(D_1^{2^k} \oplus 0) - W_1(0 \oplus 2^{-k} I_r)). \tag{4.28}$$

By (4.16),

$$A_k = (G_k V_1 + V_2)(D_2^{-2^k} \oplus I_r) V_1^{-1}. \tag{4.29}$$

Inserting (4.29) in (4.28) we obtain

$$G_k W_3 + W_4 - (G_k W_1 + W_2)(0 \oplus 2^{-k} I_r) = (G_k V_1 + V_2) M_k$$

with $M_k = O(2^{-k})$. Thus

$$-G_k\big(W_3 - W_1(0 \oplus 2^{-k} I_r) - V_1 M_k\big) = W_4 - W_2(0 \oplus 2^{-k} I_r) - V_2 M_k.$$

It follows from (4.14) that

$$G_k - AS^{(1)} = G_k + W_4 W_3^{-1} = O(2^{-k}). \tag{4.30}$$

Postmultiplying (4.16) by $D_2^{-2^k} \oplus 0$ gives

$$(G_k V_1 + V_2)(D_2^{-2^k} \oplus 0) = A_k V_1 (I_r \oplus 0), \tag{4.31}$$

while postmultiplying (4.17) by $0 \oplus 2^{-k} I_r$ gives

$$(G_k V_3 + V_4)(0 \oplus 2^{-k} I_r) = A_k V_1 (0 \oplus I_r) + A_k V_3 (0 \oplus 2^{-k} I_r). \tag{4.32}$$

Adding (4.31) and (4.32) we get

$$A_k(V_1 + V_3(0 \oplus 2^{-k} I_r)) = (G_k V_1 + V_2)(D_2^{-2^k} \oplus 0) + (G_k V_3 + V_4)(0 \oplus 2^{-k} I_r).$$

It follows that

$$A_k = O(2^{-k}), \tag{4.33}$$

since $\{G_k\}$ has been shown to be bounded. Equations (4.26), (4.27), (4.30), and (4.33) yield the required convergence results. $\square$

For S_k and B_k in iteration (3.1) we obtain the following convergence result.

COROLLARY 4.7. *Under the conditions of Theorem* 4.6*, the iterates S_k and B_k in* (3.1) *satisfy*

$$\limsup_{k\to\infty} \sqrt[k]{\|S_k - \widehat{S}\|} \le \frac{1}{2}, \qquad \limsup_{k\to\infty} \sqrt[k]{\|B_k - \widehat{B}\|} \le \frac{1}{2},$$

where $\widehat{S} = -{S^{(2)}}^* A$ is nonsingular and $\widehat{B} = A(S^{(1)} - S^{(2)}) \ge 0$ is singular.

Proof. The convergence results follow from Theorem 4.6 by noting $B_k = H_k + G_k$ and $S_k = H_k^*$. By (4.27) and (4.30), $\widehat{B} = A(S^{(1)} - S^{(2)})$. We have $\widehat{B} \ge 0$ since $B_k > 0$ for each k, by Lemma 4.4. We now show that $\widehat{B}$ is singular. Using (4.9) it is easy to check that

$$(-X_0^{-1}Y_0)\begin{bmatrix} I & I \\ -AS^{(1)} & -AS^{(2)} \end{bmatrix} = \begin{bmatrix} I & I \\ -AS^{(1)} & -AS^{(2)} \end{bmatrix}(S^{(1)} \oplus S^{(2)}), \tag{4.34}$$

and $S^{(1)} \oplus S^{(2)}$ is diagonalizable. Now $-X_0^{-1}Y_0$ is not diagonalizable, by assumption, since it has at least one eigenvalue of partial multiplicity 2. Thus (4.34) can only hold if $\left[\begin{smallmatrix} I & I \\ -AS^{(1)} & -AS^{(2)} \end{smallmatrix}\right]$ is singular. Thus the Schur complement $\widehat{B} = A(S^{(1)} - S^{(2)})$ is singular. $\square$

In the generic case for a weakly overdamped Q with $\gamma = 0$, in which all the partial multiplicities of λ_n are 2, Q is in some sense irreducible or coupled. The next example shows that this condition is necessary for the conclusions in Theorem 4.6 and Corollary 4.7 (and at the same time answers an open question from [9, Sec. 4]). Consider

$$Q(\lambda) = \lambda^2 A + \lambda B + C = \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

It is easy to see that $\gamma = 0$, so $Q(\lambda)$ is weakly overdamped with eigenvalues $\{0, -1, -1, -2\}$ with $\lambda_2 = \lambda_3 = -1$ semisimple. In (3.1) and (4.5), (4.6),

$$\lim_{k\to\infty} A_k = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \lim_{k\to\infty} B_k = I_2, \quad \lim_{k\to\infty} C_k = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$
$$\lim_{k\to\infty} G_k = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \lim_{k\to\infty} H_k = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Neither A_k nor C_k converges to zero. We also note that the convergence is quadratic for B_k, G_k, H_k. Moreover, B_k converges to a nonsingular matrix. This does not come as a surprise, since $Q(\lambda)$ can be decomposed into the direct sum of two scalar quadratics,

$$Q_1(\lambda) = \lambda^2 + 3\lambda + 2, \qquad Q_2(\lambda) = \lambda^2 + \lambda.$$

It is readily seen that Q_1 is overdamped with eigenvalues $-1, -2$ and that Q_2 is overdamped with eigenvalues $0, -1$. Thus, the convergence of B_k to a positive definite matrix is guaranteed by Theorem 3.4 applied to each component of the direct sum.

5. Algorithm for the detection and numerical solution. Let $Q(\lambda) = \lambda^2 A + \lambda B + C$ be Hermitian with $A > 0$. We develop in this section an efficient algorithm that checks if Q is hyperbolic and, if it is, computes some or all of the eigenvalues and associated eigenvectors, exploiting the symmetry and hyperbolicity and thereby preserving the spectral properties.

Our algorithm consists of three steps:

1. *Preprocessing*: this step forms $Q_\theta(\lambda) \equiv Q(\lambda + \theta) = \lambda^2 A_\theta + \lambda B_\theta + C_\theta$ with θ such that $B_\theta > 0$ and $C_\theta \geq 0$, or concludes that Q is not hyperbolic and terminates the algorithm.
2. *Overdamping test*: this step checks the overdamping condition for Q_θ. If Q_θ is overdamped, a $\mu \in \mathbb{R}$ such that $Q_\theta(\mu) = Q(\mu + \theta) < 0$ is also computed; otherwise the algorithm terminates.
3. *Solution*: the quadratic Q_θ is converted into a definite pencil $\lambda X + Y \in \mathbb{C}^{2n \times 2n}$ with $X > 0$ or $Y > 0$. The eigenvalues and eigenvectors of $Q(\lambda)$ are then obtained from the eigendecomposition of a $2n \times 2n$ Hermitian matrix obtained by transforming $\lambda X + Y$ and exploiting the definiteness of X or Y and the block structure of X and Y.

We now detail each of these three steps and compare the cost and stability of our solution process with that of three alternative ways of solving the QEP: the QZ algorithm applied to a linearization of $Q(\lambda)$, the J-orthogonal Jacobi algorithm [28] also applied to a linearization of $Q(\lambda)$, and the method of computing the eigenpairs of the primary and secondary solvents obtained via the cyclic reduction method [9].

At different stages our algorithm needs to test the (semi)definiteness of a matrix. This is done by attempting a Cholesky factorization, with complete pivoting in the case of semidefinitiness: completion of the factorization means the matrix is (semi)definite. This is a numerically stable test, as shown in [10].

5.1. Preprocessing step. The preprocessing step aims to eliminate by simple tests quadratics that are not hyperbolic and to produce, if possible, a shifted quadratic $Q_\theta(\lambda) = Q(\lambda + \theta)$ (with $\theta = 0$ is possible) for which the necessary condition

$$B > 0, \quad C \geq 0 \tag{5.1}$$

for overdamping is satisfied.

If B is singular then by (1.2) Q cannot be hyperbolic. Assume now that B is nonsingular but not positive definite or C is not positive semidefinite. Since $A > 0$ then for $\theta > 0$ large enough, the matrices

$$B_\theta = B + 2\theta A, \quad C_\theta = C + \theta B + \theta^2 A$$

defining the shifted quadratic $Q_\theta(\lambda) = Q(\lambda + \theta)$ with $A_\theta = A$ (see (2.2)) satisfy (5.1). To avoid numerical instability in the formation of B_θ and C_θ (due to the possibly large variation in $\|A\|$, $\|B\|$, and $\|C\|$) we would ideally like to choose θ close to

$$\theta_{\mathrm{opt}} = \inf\{\theta \in \mathbb{R} : \; B + 2\theta A > 0, \; C + \theta B + \theta^2 A \geq 0\}.$$

Rather than solving this optimization problem we choose θ to be an upper bound on the modulus of λ_1, the right-most eigenvalue of Q. With such a shift, all the eigenvalues of Q_θ lie in the left half plane. When Q is hyperbolic, Q_θ is also hyperbolic with real and nonpositive eigenvalues. Thus $B_\theta > 0$ and $C_\theta \geq 0$ by Theorem 2.5. Therefore if $B_\theta \not> 0$ or $C_\theta \not\geq 0$ we can conclude that Q is not hyperbolic. If $B_\theta > 0$ and $C_\theta \geq 0$ we proceed to step 2.

TABLE 5.1
Operation count for the preprocessing step. Matrices are assumed real and of dimension n.

Operations		Cost (flops)
Cholesky factorization of B and C to check definiteness.		$2n^3/3$ or less
Computation of θ when B and/or C not positive definite:		
Cholesky factorization of A.		$n^3/3$
$\|A^{-1}\|$ (1-norm estimation [11, Sec. 15.3], typically 4 solves).		$4n^2$
Form $B_\theta = B + 2\theta A$, $C_\theta = C + \theta B + \theta^2 A$.		$6n^2$
Cholesky factorizations of B_θ and C_θ.		$2n^3/3$ or less
	Total	$5n^3/3$ or less

To construct the shift θ we use the following strategy. Let

$$a = \|A\|, \quad b = \|B\|, \quad c = \|C\|,$$

where $\|\cdot\|$ is any consistent matrix norm. Then from [18, Lem. 3.1 and Lem. 4.1], for every eigenvalue λ of Q we have

$$|\lambda| \le \frac{1}{2}\|A^{-1}\|\left(b + \sqrt{b^2 + 4c/\|A^{-1}\|}\right) =: \sigma_1, \tag{5.2}$$

$$|\lambda| \le \bigl(1 + \|A^{-1}\|\bigr)\max(c^{1/2}, b) =: \sigma_2. \tag{5.3}$$

We take the 1-norm and set $\sigma = \min(\sigma_1, \sigma_2)$. Since σ must greatly overestimate $|\lambda_1|$ when $|\lambda_n| \gg |\lambda_1|$ we carry on one step further and form the shifted quadratic $Q_{-\sigma/2}(\lambda) = Q(\lambda - \sigma/2)$ for which (5.2)–(5.3) give two new bounds τ_1 and τ_2 (and A is unchanged so $\|A^{-1}\|$ can be re-used). We then take $\theta = \min(\sigma, \tau - \frac{1}{2}\sigma)$, where $\tau = \min(\tau_1, \tau_2)$.

As shown by Theorem 3.4, the speed of convergence of iteration (3.7) for overdamped Q depends on the ratio λ_n/λ_{n+1}. An unnecessarily large shift of the spectrum to the left can make this ratio very close to 1, potentially causing slow convergence of the iteration. However, we showed in section 4 that for the generic case of weakly overdamped Q with $\lambda_n = \lambda_{n+1}$ the convergence is at least linear with constant 1/2, so convergence of the iteration cannot be unduly delayed by a conservative choice of shift.

Table 5.1 details the computations and their cost. (Costs of all the operations used here are summarized in [12, App. C].) Preprocessing requires at most $\frac{5}{3}n^3$ flops.

5.2. Overdamping test. The following algorithm is based on Corollary 3.6. It runs the scaled iteration (3.7) until either a non-positive definite B_k or a negative definite $Q(\mu_k)$ is detected, signalling that Q is not overdamped or is overdamped, respectively. The algorithm terminates on one of these conditions or because of possible non-convergence of the iteration for a non-overdamped Q. It is intended to be applied to Q_θ from the preprocessing step.

ALGORITHM 5.1 (overdamping test). *This algorithm tests whether a quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ with $A, B > 0$ and $0 \ne C \ge 0$ is overdamped and, if it is, computes $\mu < 0$ such that $Q(\mu) < 0$. Input parameters are the maximum number of iterations $k_{\max}$ and a convergence tolerance $\epsilon > 0$.*

Set $A_0 = A$, $B_0 = B$, $C_0 = C$.
Set $\alpha_0 = \|C_0\|_1/\|A_0\|_1$, $\mu_0 = -\alpha_0^{1/2}$, $k = 0$.
if $Q(\mu_0) < 0$, $Q(\lambda)$ is (hyperbolic and hence) overdamped, $\mu = \mu_0$, quit, end

4 while $k < k_{\max}$
5 $\quad B_{k+1} = B_k - A_k B_k^{-1} C_k - C_k B_k^{-1} A_k$
6 $\quad$ if $\|B_{k+1} - B_k\|_1 / \|B_{k+1}\|_1 \le \epsilon$, goto line 15, end
7 $\quad$ if $B_{k+1} \not> 0$, Q is not overdamped, quit, end
8 $\quad A_{k+1} = \alpha_k A_k B_k^{-1} A_k$
9 $\quad C_{k+1} = \alpha_k^{-1} C_k B_k^{-1} C_k$
10 $\quad \alpha_{k+1} = \|C_{k+1}\|_1 / \|A_{k+1}\|_1$
11 $\quad \mu_{k+1} = \mu_k \alpha_{k+1}^{1/2^{k+2}}$
12 $\quad$ if $Q(\mu_{k+1}) < 0$, Q is overdamped, $\mu = \mu_{k+1}$, quit, end
13 $\quad k = k + 1$
14 end
15 Q is not overdamped. % See the discussion below.

Note that the crucial definiteness test on line 12 of Algorithm 5.1 is carried out on Q and not on Q_k in (3.5). Hence a positive test can be interpreted irrespective of rounding errors in the iteration: the only errors are in forming $Q(\mu_{k+1})$ and in computing its Cholesky factor. For a non-overdamped Q, it is possible that $B_k > 0$ for all k (see the example at the end of section 4). However, if convergence of the B_k is detected on line 6 then Q is declared not overdamped because by this point an overdamped Q would have been detected, while if $k_{\max}$ is large enough (say $k_{\max} = 20$) and this iteration limit is reached then Q can reasonably be declared not overdamped in view of the fast (quadratic) convergence of (3.7) for an overdamped Q.

The implementation details of Algorithm 5.1 and the cost per iteration are described in Table 5.2. The total cost for m iterations is $\frac{1}{3}n^3$ flops for $m = 0$ and roughly $\frac{20}{3}mn^3$ flops for $m \ge 1$.

Guo and Lancaster's test for overdamping is based on iteration (3.1), scaled as in (3.7). For the computation of $\widehat{S}$, $\frac{19}{3}\ell n^3$ flops are required where ℓ is the number of iterations for convergence of (3.1). An extra $5n^3$ flops is needed to form the two solvents $S^{(1)}$ and $S^{(2)}$ (which are nonsymmetric in general) via (3.3). Then the smallest eigenvalue λ_n of $S^{(1)}$ and the largest eigenvalue λ_{n+1} of $S^{(2)}$ need to be computed and the definiteness of $Q((\lambda_n + \lambda_{n+1})/2)$ tested. The total cost is $(\frac{19}{3}\ell + \frac{16}{3})n^3$ flops plus the cost of finding λ_n and λ_{n+1}. Since $m \le \ell$, Algorithm 5.1 is clearly the more efficient, possibly significantly so.

We mention two alternative ways to test hyperbolicity. Both are based on the fact that a Hermitian Q with $A > 0$ is hyperbolic if and only if a certain $2n \times 2n$ pair $(\mathcal{A}, \mathcal{B})$ is definite [19, Thm. 3.6]. The first approach is to apply the J-orthogonal Jacobi algorithm of Veselić [28] to $(\mathcal{A}, \mathcal{B})$, since the algorithm breaks down when applied to an indefinite pair. Drawbacks of this approach are that the algorithm uses hyperbolic transformations and so is potentially unstable, and that it must be run to completion to check whether the problem is overdamped, though of course upon completion it has computed the eigenvalues. It requires an initial $\frac{11}{3}n^3$ flops followed by $12sn^3$ flops, where s is the number of sweeps performed. The second approach is to apply to $(\mathcal{A}, \mathcal{B})$ an algorithm of Crawford and Moon [4] for detecting definiteness of Hermitian matrix pairs. Although only linearly convergent, this algorithm usually terminates within 30 iterations with a message of "definite", "indefinite", or "fail" (denoting failure of the algorithm to make a determination). The number 30 here is for difficult problems, for which our algorithm may also need 20 iterations. For easy problems, the Crawford–Moon algorithm needs about 3 iterations, while our algorithm needs 0 or 1 iterations. Since the Crawford–Moon algorithm requires one Cholesky factorization per iteration, here of a $2n \times 2n$ matrix, it needs $\frac{8}{3}n^3$ flops per

Table 5.2
Operation count per complete iteration of Algorithm 5.1. *Matrices are assumed real and of dimension* n.

Operations		Cost (flops)
Cholesky factorization of $B_k = L_k L_k^T$ available from previous step		
Form $V_k = L_k^{-1} A_k$		n^3
Form $W_k = L_k^{-1} C_k$		n^3
Compute $A_k B_k^{-1} C_k = V_k^T W_k$		$2n^3$
Cholesky of B_{k+1}		$n^3/3$
Compute $A_k B_k^{-1} A_k = V_k^T V_k$		n^3
Compute $C_k B_k^{-1} C_k = W_k^T W_k$		n^3
Cholesky of $-Q(\mu_{k+1})$		$n^3/3$
	Total	$20n^3/3$

iteration, and this can be reduced to $\frac{1}{3}n^3$ flops per iteration by working directly with the $n \times n$ quadratic Q through the use of a congruence transformation, as given in the proof of [19, Thm. 3.6] for example. Since our algorithm needs $\frac{20}{3}n^3$ flops per iteration, it is often more efficient than the Crawford–Moon algorithm applied to the pair $(\mathcal{A}, \mathcal{B})$, and is often less efficient than the Crawford–Moon algorithm working on Q via the congruence. However, the Crawford–Moon algorithm with or without the congruence is numerically unreliable, as we now explain.

We use a MATLAB translation of the Fortran code PDFIND from [3] and also modify it so that it exploits the congruence to work only with the quadratic Q. For the quadratic (3.11), we found that for $\beta \in (0.5196152422, 0.5196152423)$ (which is a small interval in which Q changes from being not overdamped to overdamped—see Tables 3.1 and 3.2), both codes often return with a "fail" message when Algorithm 5.1 correctly diagnoses (non-) overdamping. We then considered a scaling of the problem $A \leftarrow \alpha^2 A$, $B \leftarrow \alpha B$ with $\alpha > 0$, which has no effect on the overdamping or on Algorithm 5.1. However, as α decreases, PDFIND becomes more unreliable, due to the increasing ill conditioning of the congruence transformation with decreasing α. To be more specific, we take $\alpha = 10^{-7}$. First, consider $\beta = 0.5157\colon 0.0001\colon 0.5197$. For $\beta = 0.5197$ our algorithm detects overdamping in 3 iterations, and for other values our algorithm detects non-overdamping in at most 8 iterations. PDFIND, using the congruence, incorrectly detects non-overdamping for $\beta = 0.5197$ and fails for $\beta = 0.5157$, 0.5177–0.5181, 0.5188, 0.5189, 0.5194. Next, we take $\beta = 0.51965\colon 0.00001\colon 0.51971$. Our algorithm detects overdamping in at most 5 iterations. PDFIND without the congruence incorrectly detects non-overdamping for 0.51965, 0.51967, 0.51968, and fails for 0.51966, 0.51969, 0.51970. The conclusion is that PDFIND is numerically unreliable whether the congruence is used or not, and when it gives the wrong answer there is no warning. The poor performance of PDFIND when working with the pair is due to the fact that the ill conditioned congruence transformation is implicitly present in the equivalence between Q being hyperbolic and $(\mathcal{A}, \mathcal{B})$ being definite.

For our algorithm, instability could potentially arise if B_k is ill-conditioned. However, we know from [2, p. 40, line 10] that B_k is well-conditioned if $B_0 = B$ is well-conditioned (which is verifiable right from the beginning) and if λ_n/λ_{n+1} is not too close to 1. When λ_n/λ_{n+1} is extremely close to 1, B_k is known to be ill-conditioned for large k. However, B_k appears in our algorithm only in terms like $A_k B_k^{-1} C_k$, and A_k and C_k converge to 0, so the ill-conditioning of B_k only has a limited effect on our algorithm; indeed, instability has not been observed in any of our tests.

Table 5.3

Operation count for the eigenvalue computation, with reference to (5.4)

Operations	Cost (flops)
Cholesky factorizations of $A = L_A L_A^T$ and $-C = L_C L_C^T$ already available.	
Form $R = -(L_A^{-1} L_C)^T$.	$n^3/3$
Form $G = L_A^{-1} B L_A^{-T}$.	$3n^3/2$
Tridiagonalization of $\left[\begin{smallmatrix} -G & R \\ -R^T & 0 \end{smallmatrix}\right]$.	$< 4(2n)^3/3$
Eigenvalues via (e.g.) QR iteration.	$O(n^2)$
Total	$\approx 13n^3$

5.3. Solving hyperbolic QEPs via definite linearizations. Recall that the scalar μ computed by Algorithm 5.1 applied to Q_θ is such that $Q(\mu+\theta) = Q_\theta(\mu) < 0$. Hence with $\omega = \mu + \theta$ we have

$$\begin{aligned} \widetilde{Q}(t) = Q(t+\omega) &= t^2 A + t(B + 2\omega A) + C + \omega B + \omega^2 A \\ &= t^2 \widetilde{A} + t\widetilde{B} + \widetilde{C}, \end{aligned}$$

with $\widetilde{C} = Q(\omega) < 0$ and $\widetilde{A} = A > 0$. The pencils

$$L_1(\lambda) = \lambda \begin{bmatrix} \widetilde{A} & 0 \\ 0 & -\widetilde{C} \end{bmatrix} + \begin{bmatrix} \widetilde{B} & \widetilde{C} \\ \widetilde{C} & 0 \end{bmatrix}, \quad L_2(\lambda) = \lambda \begin{bmatrix} 0 & \widetilde{A} \\ \widetilde{A} & \widetilde{B} \end{bmatrix} + \begin{bmatrix} -\widetilde{A} & 0 \\ 0 & \widetilde{C} \end{bmatrix}$$

are both Hermitian definite linearizations of $\widetilde{Q}$ with positive definite leading coefficient of L_1 and negative definite trailing coefficient of L_2. They share the same eigenvalues as $\widetilde{Q}$ and the eigenvectors of $\widetilde{Q}$ are easy to recover from those of L_1 or L_2. The sensitivity and stability of these linearizations have recently been studied in [14], [15], [17]. It is shown therein that the scaling of Fan, Lin and Van Dooren [6] should be applied to $\widetilde{Q}$ before linearizing. The choice between L_1 and L_2 should be guided by the fact that in terms of conditioning and backward error they favour large and small eigenvalues, respectively. However, if $\widetilde{C}$ or $\widetilde{A}$ is well conditioned and $\|\widetilde{B}\|/(\|\widetilde{A}\|\,\|\widetilde{C}\|)^{1/2}$ is not much bigger than 1 then L_1 or L_2, respectively, can safely be used to stably obtain all the eigenpairs. For more details on conditioning and backward error for L_1 and L_2 see [14], [15], [17].

Using Cholesky factorizations $\widetilde{A} = L_A L_A^T$ and $-\widetilde{C} = L_C L_C^T$, the definite generalized eigenvalue problem $L_1(\lambda)z = 0$ or $L_2(\lambda)z = 0$ is transformed to a Hermitian (or real symmetric) standard eigenvalue problem [5]. For example, $L_1(\lambda)$ reduces to

$$\lambda I + \begin{bmatrix} L_A^{-1}\widetilde{B}L_A^{-T} & -L_A^{-1}L_C \\ -L_C^T L_A^{-T} & 0 \end{bmatrix}. \tag{5.4}$$

As Table 5.3 explains, this phase requires about $13n^3$ flops, giving a grand total of $(\frac{20}{3}m + 13)n^3$ flops.

Guo and Lancaster's solution algorithm has a total cost of $(\frac{19}{3}\ell + 25)n^3$ flops, assuming the eigenvalues of $S^{(1)}$ and $S^{(2)}$ (which are the eigenvalues of Q) are computed by the QR algorithm. In practice this is significantly more than the cost of our algorithm given that $m \le \ell$ is usually small.

The most common way of solving the QEP is to apply the QZ algorithm or a Krylov method to a linearization L of Q. The QZ algorithm applied to the $2n \times 2n$ L costs $240n^3$ flops for the computation of the eigenvalues.

Our algorithm has two important advantages over that of Guo and Lancaster and QZ applied to a linearization, besides its more favorable operation count. First, it work entirely with symmetric matrices, which reduces the storage requirement. Second, it guarantees to produce real eigenvalues in floating point arithmetic; the other two approaches cannot do so because they invoke the QZ algorithm and the nonsymmetric QR algorithm.

6. Numerical experiment. We describe an experiment that illustrates the behavior of our algorithm for testing overdamping. More extensive testing of this algorithm, and of the preprocessing and solve procedures described in section 5, will be presented in a future publication. Our experiments were performed in MATLAB 7.4 (R2007a), for which the unit roundoff is $u = 2^{-53} \approx 1.1 \times 10^{-16}$. We took $k_{\max} = 30$ and $\epsilon = u$ in Algorithm 5.1.

We first describe a useful technique for generating symmetric quadratic matrix polynomials with prescribed eigenvalues and eigenvectors and positive definite coefficient matrices.

Let (λ_k, v_k), $k = 1\colon 2n$ be a set of given real eigenpairs such that with

$$\begin{aligned} \Lambda &= \mathrm{diag}(\lambda_1, \dots, \lambda_{2n}) =: \Lambda_1 \oplus \Lambda_2, \qquad && \Lambda_1, \Lambda_2 \in \mathbb{R}^{n\times n}, \\ V &:= [\, v_1, \dots, v_{2n}\,] =: [\, V_1 \quad V_2\,], && V_1, V_2 \in \mathbb{R}^{n\times n}, \end{aligned}$$

V_1 and V_2 are nonsingular and

$$V_1 V_1^T = V_2 V_2^T, \qquad V_1\Lambda_1 V_1^T - V_2 \Lambda_2 V_2^T =: \Gamma \text{ is nonsingular.} \tag{6.1}$$

Then the symmetric quadratic polynomial defined by the matrices

$$A = \Gamma^{-1}, \quad B = -A(V_1\Lambda_1^2 V_1^T - V_2\Lambda_2^2 V_2^T)A, \tag{6.2a}$$

$$C = -A(V_1\Lambda_1^3 V_1^T - V_2\Lambda_2^3 V_2^T)A + B\Gamma B, \tag{6.2b}$$

has eigenpairs (λ_k, v_k), $k = 1\colon 2n$ (see [23] for example). We now show how to generate a potentially overdamped quadratic.

LEMMA 6.1. *Assume that* $0 > \lambda_1 \ge \cdots \ge \lambda_n > \lambda_{n+1} \ge \cdots \ge \lambda_{2n}$. *Then* Γ *is nonsingular and the matrices generated by* (6.2) *satisfy* $A > 0$, $B > 0$, *and* $C > 0$.

Proof. It follows from Weyl's theorem [20, p. 181] that $\Gamma > 0$ and hence that $A > 0$. All matrices V_2 that satisfy the first constraint in (6.1) can be written $V_1 U$ for some orthogonal U. Hence we can write

$$B = -AV_1(\Lambda_1^2 - U\Lambda_2^2 U^T)V_1^T A = -AV_1(H_1^2 - H_2^2)V_1^T A,$$

where $H_1 = \Lambda_1$ and $H_2 = U\Lambda_2 U^T$, and again Weyl's theorem guarantees that $B > 0$.

It is known that (V, Λ, PV^T) where $P = \mathrm{diag}(I_n, -I_n)$ forms a self adjoint triple for $Q(\lambda)$ [7, Sec. 10.2]. Since Q has no zero eigenvalues, C is nonsingular and a formula for its inverse is easily obtained from the resolvent form of $Q(\lambda)$: for $\lambda \ne \lambda_i$,

$$Q(\lambda)^{-1} = V(\lambda I_{2n} - \Lambda)^{-1} PV^T.$$

Setting $\lambda = 0$ in the above expression gives

$$C^{-1} = -V\Lambda^{-1}PV^T = -V_1(H_1^{-1} - H_2^{-1})V_1^T$$

and once again Weyl's theorem guarantees that C^{-1}, and therefore also C, is positive definite. $\square$

We use the following eigenvalue distributions:

Table 6.1

Minimum, average, and maximum number of iterations performed by Algorithm 5.1 *and percentage of overdamped problems, for each* n *and matrix type.*

n	type 1		type 2		type 3	
5	0.0, 2.4, 6.0	100%	0.0, 0.8, 3.0	100%	0.0, 2.4, 5.0	25%
10	0.0, 3.6, 10.0	100%	0.0, 0.5, 3.0	100%	2.0, 2.7, 4.0	5%
50	0.0, 4.2, 11.0	100%	0.0, 2.1, 4.0	100%	2.0, 2.1, 3.0	0%
100	3.0, 6.2, 10.0	100%	0.0, 2.6, 4.0	100%	2.0, 2.0, 2.0	0%
250	2.0, 6.0, 11.0	100%	2.0, 3.0, 4.0	100%	2.0, 2.0, 2.0	0%
500	3.0, 7.5, 11.0	100%	2.0, 3.0, 4.0	100%	2.0, 2.0, 2.0	0%

type 1: λ_k, $k = 1{:}\,2n$, is uniformly distributed in $[-100, -1]$.

type 2: λ_k is uniformly distributed in $[-100, -6]$ for $k = n+1{:}\,2n$ and $[-5, -1]$ for $k = 1{:}\,n$.

type 3: λ_k is uniformly distributed in $[-100, 20]$. B and C are then shifted as in (2.2) with $\theta = 1.1\lambda_1$ to ensure that the eigenvalues are all negative.

We took $V_1 = U_1$ and $V_2 = V_1 U_2$, where U_1 and U_2 are random orthogonal matrices from the Haar distribution [11, Sec. 28.3]. For types 1 and 2, A, B, and C are all positive definite by construction; for type 3 nothing can be said about the definiteness of A, B, and C. Table 6.1 shows the minimum, average, and maximum number of iterations for Algorithm 5.1 over 20 quadratics for each of several values of n, along with the percentage of Q found to be overdamped for each n and matrix type. In all cases where Q was deemed overdamped, the computed μ was verified to lie in $(\lambda_{n+1}, \lambda_n)$.

We make several observations.

- For all three eigenvalue distributions, Algorithm 5.1 is quick to terminate, especially for types 2 and 3, with only very occasional need for more than 10 iterations. The gap between λ_n and λ_{n+1} is larger for type 2 than type 1, which explains the greater number of iterations for type 1.
- With V_1 orthogonal the coefficients matrices A, B, and C are well conditioned, with 2-norm condition numbers of order 10^2. If instead we take V_1 a random matrix with 2-norm condition number 10^4 (computed in MATLAB as `gallery('randsvd',n,1e4,...)`, the condition numbers of A, B, and C are of order 10^8 and the number of iterations of the algorithm increases, though only slightly: the maximum number of iterations over all tests is 13 and the largest average over all n rises to 7.8, 3.1 and 3.2 for types 1, 2, and 3, respectively.
- After detecting overdamping an average of 6–9 more iterations are needed for convergence of the block cyclic iteration. Recall that the algorithm of Guo and Lancaster [9] needs to iterate to convergence in order to show overdamping.

Acknowledgements. This work was started while the first author visited MIMS in the School of Mathematics at the University of Manchester in 2005; he thanks the School for its hospitality. We thank Qiang Ye for helpful comments concerning the algorithm of Crawford and Moon.

REFERENCES

[1] L. Barkwell and P. Lancaster, *Overdamped and gyroscopic vibrating systems*, Trans. AME: J. Applied Mechanics, 59 (1992), pp. 176–181.

[2] D. A. Bini, L. Gemignani, and B. Meini, *Computations with infinite Toeplitz matrices and polynomials*, Linear Algebra Appl., 343–344 (2002), pp. 21–61.

[3] C. R. Crawford, *ALGORITHM 646 PDFIND: A routine to find a positive definite linear combination of two real symmetric matrices*, ACM Trans. Math. Software, 12 (1986), pp. 278–282.

[4] C. R. Crawford and Y. S. Moon, *Finding a positive definite linear combination of two Hermitian matrices*, Linear Algebra Appl., 51 (1983), pp. 37–48.

[5] P. I. Davies, N. J. Higham, and F. Tisseur, *Analysis of the Cholesky method with iterative refinement for solving the symmetric definite generalized eigenproblem*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 472–493.

[6] H.-Y. Fan, W.-W. Lin, and P. Van Dooren, *Normwise scaling of second order polynomial matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 252–256.

[7] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.

[8] C.-H. Guo, *Convergence rate of an iterative method for a nonlinear matrix equation*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 295–302.

[9] C.-H. Guo and P. Lancaster, *Algorithms for hyperbolic quadratic eigenvalue problems*, Math. Comp., 74 (2005), pp. 1777–1791.

[10] N. J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl., 103 (1988), pp. 103–118.

[11] ———, *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second ed., 2002.

[12] ———, *Functions of Matrices: Theory and Computation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[13] N. J. Higham and H.-M. Kim, *Numerical analysis of a quadratic matrix equation*, IMA J. Numer. Anal., 20 (2000), pp. 499–519.

[14] N. J. Higham, R.-C. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1218–1241.

[15] N. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 1005–1028.

[16] ———, *Definite matrix polynomials and their linearization by definite pencils*, MIMS EPrint 2007.97, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Apr. 2008.

[17] N. J. Higham, D. S. Mackey, F. Tisseur, and S. D. Garvey, *Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problems*, Internat. J. Numer. Methods Eng., 73 (2008), pp. 344–360.

[18] N. J. Higham and F. Tisseur, *Bounds for eigenvalues of matrix polynomials*, Linear Algebra Appl., 358 (2003), pp. 5–22.

[19] N. J. Higham, F. Tisseur, and P. M. Van Dooren, *Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems*, Linear Algebra Appl., 351–352 (2002), pp. 455–474.

[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[21] D. J. Inman and A. N. Andry, Jr., *Some results on the nature of eigenvalues of discrete damped linear systems*, J. Appl. Mech., 47 (1980), pp. 927–930.

[22] P. Lancaster, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966. Reprinted by Dover, New York, 2002.

[23] P. Lancaster, *Inverse spectral problems for semisimple damped vibrating systems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 279–301.

[24] W.-W. Lin and S.-F. Xu, *Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 26–39.

[25] A. S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, American Mathematical Society, Providence, RI, USA, 1988.

[26] B. Meini, *Efficient computation of the extreme solutions of $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$*, Math. Comp., 71 (2002), pp. 1189–1204.

[27] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.

[28] K. Veselić, *A Jacobi eigenreduction algorithm for definite matrix pairs*, Numer. Math., 64 (1993), pp. 241–269.