The ratio between the Toeplitz and the unstructured condition number

Siegfried M. Rump and H. Sekigawa

Abstract. Recently it was shown that the ratio between the normwise Toeplitz structured condition number of a linear system and the general unstructured condition number has a finite lower bound. However, the bound was not explicit, and nothing was known about the quality of the bound. In this note we derive an explicit lower bound only depending on the dimension n, and we show that this bound is almost sharp for all n.

Mathematics Subject Classification (2000). 15A12, 26D05.

Keywords. Structured condition number, Toeplitz matrix, Mahler measure, polynomial norms.

1. Notation and problem formulation

For a system of linear equations $Ax = b$ with $A \in \mathbb{R}^{n\times n}$, $x, b \in \mathbb{R}^n$, the condition number characterizes the sensitivity of the solution x with respect to infinitely small perturbations of the matrix A. For $\varepsilon > 0$, denote

$$M_\varepsilon := M_\varepsilon(A) := \{\Delta A \in \mathbb{R}^{n\times n} : \|\Delta A\| \le \varepsilon \|A\|\}, \tag{1.1}$$

where throughout the paper $\|\cdot\|$ denotes the spectral norm for matrices and for vectors. Denote by $P_\varepsilon := P_\varepsilon(A, x)$ the set of all vectors $\Delta x \in \mathbb{R}^n$ for which there exists $\Delta A \in M_\varepsilon$ with $(A + \Delta A)(x + \Delta x) = Ax$. Then the (unstructured) normwise condition number is defined by

$$\kappa(A, x) := \lim_{\varepsilon\to 0} \sup_{\Delta x \in P_\varepsilon} \frac{\|\Delta x\|}{\varepsilon \|x\|}. \tag{1.2}$$

It is well known that $\kappa(A, x) = \|A^{-1}\|\,\|A\|$, such that the (unstructured) condition number does not depend on x.

If the matrix A has some structure, it seems reasonable to restrict the set M_ε to matrices with similar structure. For $a = (a_{-(n-1)}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{n-1})$, the $n \times n$ Toeplitz matrix $T_n(a)$ is of the form

$$T := T_n(a) := \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{-1} & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_{-(n-1)} & \cdots & a_{-1} & a_0 \end{pmatrix} \in \mathbb{R}^{n\times n}. \tag{1.3}$$

For given (nonsingular) Toeplitz matrix T, restricting M_ε to Toeplitz matrices changes (1.2) into the Toeplitz condition number $\kappa^{\mathrm{Toep}}(T, x)$ [10], [13], [2, Section 13.3]. Since the set of perturbations ΔA is restricted, it follows $\kappa^{\text{Toep}}(T,x) \le \kappa(T,x) = \|T^{-1}\|\,\|T\|$. Note that in contrast to the general condition number, the Toeplitz condition number depends on x. However, there exists always a worst case x such that both condition numbers coincide [13, Theorem 4.1]:

$$\sup_{x\neq 0}\left\{\kappa^{\text{Toep}}(T,x)\right\} = \|T^{-1}\|\,\|T\| \ .$$

In [13, Theorem 10.2] it was shown that $\kappa^{\text{Toep}}(T,x) \ge 2^{-1/2}\sqrt{\kappa(T,x)}$ (see also [2, Theorem 13.14]), hence the ratio $\kappa^{\text{Toep}}/\kappa$ is bounded below by $[2\|T^{-1}\|\,\|T\|]^{-1/2}$. The question arises, how small can the Toeplitz condition number actually be compared to the general condition number?

In a recent survey paper on Toeplitz and Hankel matrices [4], Böttcher and Rost note "One expects that $\kappa^{\text{Toep}}(T,x)$ is in general significantly smaller than $\kappa(T,x)$, but, curiously up to now no convincing example in this direction is known." Furthermore, Böttcher and Rost continue to note that, as proved in [3] (submitted in 2002 but appeared in 2005), it seems rather hopeless to find examples numerically (see also [2, Theorem 13.20]):

Theorem 1.1 (Böttcher, Grudsky, 2002). *Let $x_0, x_1, \dots, x_{n-1} \in \mathbb{C}$ be independent random variables whose real and imaginary parts are subject to the standard normal distribution and put $x = (x_j)_{j=0}^{n-1}$. There are universal constants $\delta \in (0,\infty)$ and $n_0 \in \mathbb{N}$ such that*

$$\text{Probability}\left(\frac{\kappa^{\text{Toep}}(T_n(a),x)}{\kappa(T_n(a),x)} \ge \frac{\delta}{n^{3/2}}\right) > \frac{99}{100}$$

for all finitely supported sequences a and all $n \ge n_0$.

Notice that generically $\kappa(T_n(a),x)$ remains bounded or increases exponentially fast as n goes to infinity. Since in the case of exponential growth the factor $\delta/n^{3/2}$ is harmless, it follows that with high probability that $\kappa^{\text{Toep}}(T_n(a),x)$ increases exponentially fast together with $\kappa(T_n(a),x)$.

In [13] the first author showed a lower bound on the ratio $\kappa^{\text{Toep}}/\kappa$ which surprisingly depends only on the solution x, not on A (see also [2, Theorem 13.16]). However, despite some examples of small dimension (inspired by Heinig [9]) no general examples could be derived.

In this note we

1. derive a general lower bound on $\kappa^{\text{Toep}}(T,x)/\kappa(T,x)$ only depending on the dimension n, and
2. show that this lower bound is almost sharp for all n.

The solution of both problems is based on the minimization of the smallest singular value of a class of Toeplitz matrices (2.2) and its surprising connection to a lower bound on the coefficients of the product of two polynomials. We will prove in Corollary 2.11 that

$$\frac{2n}{\Delta^{n-1}} \ge \inf\{\frac{\kappa^{\text{Toep}}(A,x)}{\kappa(A,x)} : \ A \in \mathbb{R}^{n\times n} \text{ Toeplitz}, \ 0 \neq x \in \mathbb{R}^n\} > \frac{\sqrt{2}}{n\Delta^{n-1}},$$

where $\Delta = 3.209912\ldots$.

We denote by $\sigma_{\min}(A)$ the smallest singular value of the matrix A, and by J the permutation matrix ("flip matrix") mapping $(1,\dots,n)$ into $(n,\dots,1)$.

2. Main results

Let a linear system $Ax = b$ with Toeplitz matrix A be given. The defining equation $(A+\Delta A)(x+\Delta x) = Ax$ with $\|\Delta A\| \le \varepsilon\|A\|$ implies

$$\Delta x = -A^{-1}\Delta A x + \mathcal{O}(\varepsilon). \tag{2.1}$$

For Toeplitz perturbations, we have $\Delta A = T(\Delta a)$ with $\Delta a \in \mathbb{R}^{2n-1}$ according to (1.3), and using ideas from [10] a computation shows [13]

$$\Delta A x = J\Psi_x \Delta a \quad \text{with} \quad \Psi_x := \begin{pmatrix} x_1 & x_2 & \dots & x_n & & & \\ & x_1 & x_2 & \dots & x_n & & \\ & & & \dots & & & \\ & & & x_1 & x_2 & \dots & x_n \end{pmatrix} \in \mathbb{R}^{n\times(2n-1)}. \tag{2.2}$$

In [13, Lemma 6.3] it was shown that the spectral matrix norm of ΔA and Euclidean norm of Δa are related by

$$\frac{1}{\sqrt{n}}\|\Delta A\| \le \|\Delta a\| \le \sqrt{2}\|\Delta A\|. \tag{2.3}$$

Combining this with the definition of $\kappa^{\mathrm{Toep}}(A,x)$ and (2.1) yields [13, Theorem 6.5]

$$\kappa^{\mathrm{Toep}}(A,x) = \gamma\frac{\|A^{-1}J\Psi_x\|\,\|A\|}{\|x\|} \quad \text{with} \quad \frac{1}{\sqrt{n}} \le \gamma \le \sqrt{2}, \tag{2.4}$$

so that $\|A^{-1}J\Psi_x\| \ge \|A^{-1}\|\sigma_{\min}(\Psi_x)$ implies [13, Corollary 6.6]

$$\frac{\kappa^{\mathrm{Toep}}(A,x)}{\kappa(A,x)} \ge \frac{1}{\sqrt{n}}\frac{\|A^{-1}J\Psi_x\|}{\|A^{-1}\|\,\|x\|} \ge \frac{1}{\sqrt{n}}\frac{\sigma_{\min}(\Psi_x)}{\|x\|}. \tag{2.5}$$

Surprisingly, this lower bound depends only on the solution x. That means, a given solution x implies a lower bound for $\kappa^{\mathrm{Toep}}(A,x)/\kappa(A,x)$ for any Toeplitz matrix A.

We will show that the lower bound in (2.5) is achievable up to a small factor. For this we first construct for given x a Toeplitz matrix A with ratio $\kappa^{\mathrm{Toep}}/\kappa$ near $\sigma_{\min}(\Psi_x)/\|x\|$.

Let fixed but arbitrary $x \in \mathbb{R}^n$ be given. For simplicity assume $\|x\| = 1$. First we will show that for $\delta > 0$ there exists a Toeplitz matrix A with $\|A^{-1}J\Psi_x\| < \|A^{-1}\|\sigma_{\min}(\Psi_x) + \delta$.

Denote by $y \in \mathbb{R}^n$, $\|y\| = 1$, a left singular vector of Ψ_x to $\sigma_{\min}(\Psi_x)$, so that $\|y^T\Psi_x\| = \sigma_{\min}(\Psi_x)$. By $\Psi_x\Psi_x^T = J\Psi_x\Psi_x^T J^T$ we may assume either $y = Jy$ or $y = -Jy$. Define by

$$L(p_1,\dots,p_n) := \begin{pmatrix} p_1 & & & \\ p_2 & \ddots & & \\ \vdots & \ddots & \ddots & \\ p_n & \dots & p_2 & p_1 \end{pmatrix} \in \mathbb{R}^{n\times n},$$

a lower triangular Toeplitz matrix depending on $p \in \mathbb{R}^n$. Define

$$B := L(y_1, y_2, \dots, y_n) \quad \text{and} \quad C := L(0, y_n, y_{n-1}, \dots, y_2)$$

and

$$R_\varepsilon := (B+\varepsilon I)(B+\varepsilon I)^T - CC^T. \tag{2.6}$$

If R_ε is invertible, then the Gohberg-Semencul formula ([8], see also [4, Th. 3.3]) implies that $A_\varepsilon := R_\varepsilon^{-1}$ is a (symmetric) Toeplitz matrix. Furthermore, a direct computation using $y = \pm Jy$ yields

$$R_0 = yy^T \tag{2.7}$$

which implies $\det(R_0) = 0$ for $n \geq 2$. The determinant of R_ε is a monic polynomial of degree $2n$ in ε, thus R_ε is nonsingular for all $0 \neq \varepsilon < \varepsilon_0$ for small enough ε_0. Hence there is a constant α, independent of ε, with

$$\|R_\varepsilon \Psi_x\| \leq \|yy^T \Psi_x\| + \alpha\varepsilon = \sigma_{\min}(\Psi_x) + \alpha\varepsilon,$$

the latter equality because $\sigma^2_{\min}(\Psi_x)$ is the only nonzero eigenvalue of $yy^T\Psi_x(yy^T\Psi_x)^T$. Since $A_\varepsilon = R_\varepsilon^{-1}$ is a Toeplitz matrix and $y = \pm Jy$, (2.4) implies the following result, which is trivially also true for $n = 1$.

Theorem 2.1. *Let $0 \neq x \in \mathrm{IR}^n$ be given. Then for all $\delta > 0$ there exists a Toeplitz matrix $A \in \mathrm{IR}^{n\times n}$ with*

$$\|A^{-1}\|\sigma_{\min}(\Psi_x) \leq \|A^{-1}J\Psi_x\| < \|A^{-1}\|\sigma_{\min}(\Psi_x) + \delta$$

and

$$\kappa^{\mathrm{Toep}}(A, x) = \gamma \cdot \kappa(A, x)\frac{\sigma_{\min}(\Psi_x)}{\|x\|} + \delta' \quad \textit{for} \quad \frac{1}{\sqrt{n}} \leq \gamma \leq \sqrt{2} \quad \textit{and} \quad 0 \leq \delta' < \sqrt{2}\delta.$$

For $x \neq 0$, the matrix Ψ_x has full rank because otherwise each $n \times n$ submatrix of Ψ_x would be singular, taking the leftmost submatrix in Ψ_x would imply $x_1 = 0$, the second leftmost would imply $x_2 = 0$ and so forth. Thus

$$\mu_n := \min_{0\neq x\in\mathrm{IR}} \frac{\sigma_{\min}(\Psi_x)}{\|x\|} = \min_{\|x\|=1} \sigma_{\min}(\Psi_x) > 0 \tag{2.8}$$

for all n, and Theorem 2.1 yields

Corollary 2.2. *For all n,*

$$\sqrt{2}\mu_n \geq \inf\{\frac{\kappa^{\mathrm{Toep}}(A, x)}{\kappa(A, x)} : \ A \in \mathrm{IR}^{n\times n} \text{ Toeplitz}, \ 0 \neq x \in \mathrm{IR}^n\} \geq \frac{1}{\sqrt{n}}\mu_n.$$

In the remaining of the paper we will estimate μ_n to characterize the infimum of $\kappa^{\mathrm{Toep}}/\kappa$. The matrix Ψ_x is also known as "polynomial matrix" [1]. Identifying a vector $x = (x_1, \ldots, x_n) \in \mathrm{IR}^n$ with the polynomial $x(t) := \sum_{\nu=0}^{n-1} x_{n-\nu}t^\nu \in \mathrm{IR}[t]$, a little computation yields

$$z = y^T\Psi_x \quad \Leftrightarrow \quad z(t) = y(t)x(t), \tag{2.9}$$

and therefore

$$y^T\Psi_x = x^T\Psi_y. \tag{2.10}$$

This, of course, can also be verified by direct computation. We define the norm $\|x(t)\|$ of a polynomial by the norm $\|x\|$ of its coefficient vector. Since $\|y^T\Psi_x\| = \sigma_{\min}(\Psi_x)$, we can characterize μ_n by

$$\mu_n = \min\{\|PQ\| : \ P, Q \in \mathrm{IR}[t], \ \deg(P) = \deg(Q) = n-1, \ \|P\| = \|Q\| = 1\}. \tag{2.11}$$

To give lower and upper bounds for μ_n, we first describe some related results for polynomials. The supremum norm $\|P\|_E$ of a complex univariate polynomial P on a compact set $E \subset \mathbb{C}$ is defined as

$$\|P\|_E := \sup_{z\in E} |P(z)|. \tag{2.12}$$

In [11] Kneser gave the exact lower bound for the supremum norm on the interval $[-1, 1]$ of the product of two polynomials.

[1]Many thanks to Ludwig Elsner, Bielefeld, for pointing to this connection.

Theorem 2.3 (Kneser, 1934). *Suppose that $PQ = R$, where P, Q and R are complex polynomials of degree m, $n-m$ and n, respectively. Then for all m and n*

$$\|P\|_{[-1,1]}\|Q\|_{[-1,1]} \leq K_{m,n}\|R\|_{[-1,1]},$$

where

$$K_{m,n} := 2^{n-1}\prod_{k=1}^{m}\left(1+\cos\frac{(2k-1)\pi}{2n}\right)\prod_{k=1}^{n-m}\left(1+\cos\frac{(2k-1)\pi}{2n}\right).$$

This bound is exactly attained by the Chebyshev polynomial of degree n.

To estimate μ_n, we need similar results for the unit disk D. Boyd's result in [5, 6] gives a sharp inequality for this case. To describe Boyd's results, we define the Mahler measure. For a complex polynomial F in k variables the Mahler measure of F is defined as

$$M(F) := \exp\left(\int_0^1 \cdots \int_0^1 \log|F(e^{2\pi\sqrt{-1}t_1},\ldots,e^{2\pi\sqrt{-1}t_k})|dt_1\cdots dt_k\right).$$

Theorem 2.4 (Boyd, 1992/94). *Let R be a polynomial of degree n with complex coefficients and suppose that $PQ = R$. Then for the norm $\|\cdot\|_D$ as in (2.12) on the unit disc D*

$$\|P\|_D\|Q\|_D \leq \delta^n\|R\|_D,$$

where $\delta = M(1+x+y-xy) = 1.7916228\ldots$. The constant is best possible.

As written in section 3 of [5], the constant δ can be expressed in terms of Clausen's integral

$$Cl_2(\theta) = -\int_0^\theta \log\left(2\sin\frac{t}{2}\right)dt = \sum_{k=1}^{\infty}\frac{\sin k\theta}{k^2},$$

or, in terms of $I(\theta)$, where

$$I(\theta) = \int_0^\theta \log\left(2\cos\frac{t}{2}\right)dt = Cl_2(\pi-\theta).$$

Using Catalan's constant $G = I(\pi/2) = Cl_2(\pi/2) \approx 0.9160$, we can write $\delta = e^{2G/\pi}$.

Theorem 2.4 implies a lower bound for μ_n. To obtain an upper bound for μ_n, we estimate the supremum norms of the following polynomials. Define $F_n(t)$ as $t^{2n}+(-1)^n$. Let $\widehat{P}_n(t)$ be the monic polynomial of degree n with the zeros of $F_n(t)$ in the right half plane, and $\widehat{Q}_n(t)$ be the monic polynomial of degree n with the zeros of $F_n(t)$ in the left half plane. It follows $\widehat{P}_n\widehat{Q}_n = F_n$ and $\widehat{Q}_n(t) = (-1)^n\widehat{P}_n(-t)$.

Lemma 2.5. *For the norm $\|\cdot\|_D$ as in (2.12) on the unit disc D, the following inequalities hold true.*

$$e^{\frac{\pi}{8n}}\delta^n > \|\widehat{P}_n\|_D = (-1)^n\widehat{P}_n(-1) = \|\widehat{Q}_n\|_D = \widehat{Q}_n(1) > \delta^n.$$

Remark 2.6. When n is even, $2K_{n/2,n} = \widehat{Q}_n(1)^2$, where $K_{p,q}$ is the constant in Theorem 2.3.

Combining Theorem 2.4 and Lemma 2.5, where the proof of the latter is deferred to the appendix, with (2.11), we obtain an upper and a lower bound for μ_n. Before we state our final result, we prove that we may assume without loss of generality that polynomials P and Q minimizing μ_n as in (2.11) must both have all their roots on the unit circle. This is also useful to identify such polynomials P and Q numerically for small n. In fact, the following Theorem 2.7 shows more, namely that for fixed (normed) Q there is a (normed) polynomial P with only roots on the unit circle and minimizing $\|PQ\|$.

Theorem 2.7. *For two nonzero real univariate polynomials P and Q with $\|P\| = \|Q\| = 1$, there exists a real univariate polynomial P' such that $\deg(P') = \deg(P)$, $\|P'\| = 1$, all zeros of P' lie on the unit circle and $\|P'Q\| \le \|PQ\|$.*

The proof of Theorem 2.7 is rather involved, and thus deferred to the appendix. An immediate consequence is the following corollary.

Corollary 2.8.

$$\begin{aligned} \mu_n \quad = \quad \min\{\|PQ\| : \ & P, Q \in \mathbb{R}[t],\ \deg(P) = \deg(Q) = n-1,\ \|P\| = \|Q\| = 1, \\ & \textit{and } P,\ Q \textit{ have all zeros on the unit circle}\}. \end{aligned}$$

Now we can prove the following upper and lower bounds for μ_n.

Theorem 2.9.

$$\frac{\sqrt{2}(n+1)}{\Delta^n} > \mu_{n+1} \ge \frac{2}{\sqrt{2n+1}\Delta^n},$$

where $\Delta := e^{4G/\pi}$ for Catalan's constant G. It is $\Delta = \delta^2$, where δ is the constant in Theorem 2.4. Note that $\Delta = 3.209912\ldots$.

Remark 2.10. Using Proposition 2.12 at the end of this section, we can improve the upper bound to

$$\frac{C\sqrt{n+1}}{\Delta^n},$$

where C is a constant independent of n.

Proof. Let F be a complex polynomial $\sum_{\nu=0}^{n} a_\nu t^\nu$. Then, the following inequalities among norms of F hold.

$$\sqrt{n+1}\|F\| \ge |F|_1 \ge \|F\|_D \ge \|F\|. \tag{2.13}$$

Here, $|F|_1$ is defined as $\sum_{\nu=0}^{n} |a_\nu|$. Real polynomials P and Q minimizing μ_n have all their roots on the unit circle. For this case the right-most inequality in (2.13) improves into

$$\|F\|_D \ge \sqrt{2}\|F\| \tag{2.14}$$

which follows from a much more general result[2] in [14], see also [15, (7.71)]. From Theorem 2.4, for *real* polynomials P and Q of degree n, we have

$$\frac{\|PQ\|_D}{\|P\|_D\|Q\|_D} \ge \frac{1}{\delta^{2n}} = \frac{1}{\Delta^n}.$$

Therefore, for polynomials P and Q with $\|P\| = \|Q\| = 1$, the inequalities

$$\|PQ\| \ge \frac{2\|PQ\|_D}{\sqrt{2n+1}\|P\|_D\|Q\|_D} \ge \frac{2}{\sqrt{2n+1}\Delta^n}$$

follow from (2.13) and (2.14). This proves the lower bound for μ_{n+1}.

Let $\widehat{P}_n$ and $\widehat{Q}_n$ be as in Lemma 2.5. An upper bound for $\|\widehat{P}_n\widehat{Q}_n\|/(\|\widehat{P}_n\|\|\widehat{Q}_n\|)$ is an upper bound for μ_{n+1}. Since $\|\widehat{P}_n\widehat{Q}_n\| = \sqrt{2}$, the inequalities

$$\frac{\|\widehat{P}_n\widehat{Q}_n\|}{\|\widehat{P}_n\|\|\widehat{Q}_n\|} \le \frac{\sqrt{2}}{(\|\widehat{P}_n\|_D/\sqrt{n+1})(\|\widehat{Q}_n\|_D/\sqrt{n+1})} < \frac{\sqrt{2}(n+1)}{\Delta^n}$$

follow from (2.13) and Lemma 2.5. ∎

---

[2] Thanks to P. Batra, Hamburg, for pointing to this reference.

Inserting this into Corollary 2.2 characterizes the asymptotic behavior of the worst ratio between the unstructured and structured condition number for Toeplitz matrices.

Corollary 2.11. *For all n,*

$$\frac{2n}{\Delta^{n-1}} > \inf\{\frac{\kappa^{\mathrm{Toep}}(A,x)}{\kappa(A,x)} : A \in \mathrm{IR}^{n\times n} \text{ Toeplitz}, \ 0 \neq x \in \mathrm{IR}^n\} > \frac{\sqrt{2}}{n\Delta^{n-1}}, \tag{2.15}$$

where $\Delta = 3.209912\ldots$ is the constant in Theorem 2.9.

We can improve the upper bound using the following Proposition, the proof of which is given in the Appendix.

Proposition 2.12.

$$\lim_{n\to\infty} \frac{||\widehat{P}_n||n^{1/4}}{||\widehat{P}_n||_D} = \lim_{n\to\infty} \frac{||\widehat{Q}_n||n^{1/4}}{||\widehat{Q}_n||_D} = \frac{1}{\sqrt{2}}.$$

By similar arguments of the proof for Theorem 2.9, we obtain the following improved upper bound.

Corollary 2.13. *There exists a constant $C > 0$ such that for all n,*

$$\frac{C\sqrt{n}}{\Delta^{n-1}} > \inf\{\frac{\kappa^{\mathrm{Toep}}(A,x)}{\kappa(A,x)} : A \in \mathrm{IR}^{n\times n} \text{ Toeplitz}, \ 0 \neq x \in \mathrm{IR}^n\},$$

where $\Delta = 3.209912\ldots$ is the constant in Theorem 2.9.

3. Approximation of μ_n

Next we show how to approximate $\Psi_x \in \mathrm{IR}^{n\times(2n-1)}$ minimizing $\sigma_{\min}(\Psi_x)$. Using Ψ_x, a Toeplitz matrix with small ratio $\kappa^{\mathrm{Toep}}/\kappa$ can be constructed following the discussion preceding Theorem 2.1. For given unit vector $x \in \mathrm{IR}^n$ and $x(t) := \sum_{\nu=0}^{n-1} x_{n-\nu}t^\nu$ define Ψ_x as in (2.2), and let $y \in \mathrm{IR}^n$ be a unit left singular vector to $\sigma_{\min}(\Psi_x)$ of Ψ_x. With $y(t) := \sum_{\nu=0}^{n-1} y_{n-\nu}t^\nu$ as in the discussion following Corollary 2.2 we have

$$\|x\| = \|y\| = \|x(t)\| = \|y(t)\| = 1 \quad \text{and} \quad \|x(t)y(t)\| = \|y^T\Psi_x\| = \sigma_{\min}(\Psi_x).$$

For fixed $x(t)$, the polynomial $y(t)$ minimizes $\|x(t)y(t)\|$ subject to $\|y(t)\| = 1$. Now (2.10) implies $\|x^T\Psi_y\| = \sigma_{\min}(\Psi_x)$ and therefore $\sigma_{\min}(\Psi_y) \le \sigma_{\min}(\Psi_x)$. Iterating the process, that is replacing x by y and computing y as a left singular vector to $\sigma_{\min}(\Psi_x)$, generates a monotonically decreasing and therefore convergent sequence. Practical experience suggests that for generic starting vector x this sequence converges mostly to the same limit, presumably μ_n. In any case this limit is an upper bound for μ_n. Table 1 displays this limit for some values of n.

To ensure that the limit is not a local but the global minimum $\min_x \sigma_{\min}(\Psi_x)$, a verified global optimization method was used [12] for computing rigorous lower and upper bounds for μ_n. Such methods take all procedural, approximation and rounding errors into account and are, provided the computer system works to its specifications, rigorous (see, for example, [7]). For given n and using (2.11) this means $2n$ variables. This was possible up to $n = 5$ with reasonable effort. The right-most column in Table 1 displays the computed bounds for μ_n. For larger values of n, the number of variables was significantly reduced using Theorem 2.7. Since minimizers P, Q have only roots on the unit circle it follows $P(z) = \pm z^n P(1/z)$ and similarly for Q, i.e. the coefficient vectors are (skew-)symmetric to reflection. Using this allows the computation of rigorous bounds for μ_n up to $n = 8$ with moderate effort.[3] Computational evidence supports the following conjecture.

[3]Thanks to Kyoko Makino for performing the verified global optimization using the COSY-package [1].

n	approximate μ_n	rigorous bounds of μ_n	$\widehat{\mu}_n = \frac{\|\widehat{P}_n\widehat{Q}_n\|}{\|\widehat{P}_n\|\|\widehat{Q}_n\|}$	$\widehat{\mu}_n/\mu_n$
2	0.70710678118655	[0.70710678118 , 0.70710678119]	0.707107	1.0000
3	0.33333333333333	[0.33333333333 , 0.33333333334]	0.353553	1.0607
4	0.13201959446019	[0.13201959446 , 0.13201959447]	0.141421	1.0712
5	0.04836936580270	[0.04836936580 , 0.04836936581]	0.051777	1.0705
6	0.01702151213258	[0.01702151213 , 0.01702151214]	0.018183	1.0682
7	0.00584679996238	[0.00584679996 , 0.00584679997]	0.006234	1.0662
8	0.00197621751074	[0.00197621751 , 0.00197621752]	0.002104	1.0647

TABLE 1

Conjecture 3.1. *There are polynomials $P, Q \in \mathbb{R}[t]$ with $degP = degQ = n-1$ and $\|P\| = \|Q\| = 1$ with $\mu_n = \|PQ\|$ such that all coefficients of P are positive, $Q(t) = P(-t)$ and all roots of P and Q lie on the unit circle. The roots $a_\nu \pm ib_\nu$ of P have all positive real parts a_ν, and the roots of Q are $-a_\nu \pm ib_\nu$.*

Finally, the values $\widehat{\mu}_n = \frac{\|\widehat{P}_n\widehat{Q}_n\|}{\|\widehat{P}_n\|\|\widehat{Q}_n\|}$ for the polynomials $\widehat{P}_n, \widehat{Q}_n$ as in Lemma 2.5 and the ratio $\widehat{\mu}_n/\mu_n$ is displayed as well. It seems that $\widehat{P}_n, \widehat{Q}_n$ are not far from the optimum. This is supported by Proposition 2.12.

4. Appendix

Proof of Lemma 2.5. Since we can write

$$\widehat{Q}_n(t) = \begin{cases} (t+1)\prod_{k=1}^{\frac{n-1}{2}}\left(t^2 + 2t\cos\frac{k\pi}{n} + 1\right), & \text{if } n \text{ is odd,} \\ \prod_{k=1}^{\frac{n}{2}}\left(t^2 + 2t\cos\frac{(2k-1)\pi}{2n} + 1\right), & \text{if } n \text{ is even,} \end{cases} \tag{4.1}$$

and $\cos\frac{k\pi}{n}$, $\cos\frac{(2k-1)\pi}{2n} > 0$ for the values of k in question, we have $\|\widehat{Q}_n\|_D = \widehat{Q}_n(1)$. From the definition of $\widehat{Q}_n$, we have $\|\widehat{Q}_n\|_D = \|\widehat{P}_n\|_D$ and $\widehat{Q}_n(1) = (-1)^n\widehat{P}_n(-1)$.

First we prove the inequalities in Lemma 2.5 when n is odd. From (4.1), we have

$$\widehat{Q}_n(1) = 2^{\frac{n+1}{2}}\prod_{k=1}^{\frac{n-1}{2}}\left(1 + \cos\frac{k\pi}{n}\right).$$

Therefore,

$$\log\widehat{Q}_n(1) = \frac{(n+1)\log 2}{2} + \sum_{k=1}^{\frac{n-1}{2}}\log\left(1 + \cos\frac{k\pi}{n}\right).$$

Let a and b be real numbers with $a < b$. For a real function f such that $f'' \le 0$ on the interval $[a, b]$, we have

$$(b-a)f\left(\frac{a+b}{2}\right) \ge \int_a^b f(x)dx \ge (b-a)\frac{f(a)+f(b)}{2}. \tag{4.2}$$

Applying (4.2) to $f = \log(1+\cos\pi x)$ on intervals $[0, \frac{1}{2n}]$, $[\frac{2k-1}{2n}, \frac{2k+1}{2n}]$ ($k = 1, 2, \ldots, \frac{n-1}{2}$) for an upper estimation, and on intervals $[\frac{k}{n}, \frac{k+1}{n}]$ ($k = 0, 1, \ldots, \frac{n-3}{2}$), $[\frac{n-1}{2n}, \frac{1}{2}]$ for a lower estimation, we have

$$\begin{aligned}\frac{\log \widehat{Q}_n(1)}{n} - \frac{(n+1)\log 2}{2n} + \frac{1}{2n}\log\left(1+\cos\frac{\pi}{4n}\right) &\geq \int_0^{\frac{1}{2}} \log(1+\cos\pi x)dx \\ &\geq \frac{\log \widehat{Q}_n(1)}{n} - \frac{\log 2}{2} - \frac{1}{4n}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right).\end{aligned}$$

Since $1+\cos\pi x = 2\cos^2\frac{\pi x}{2}$, it follows

$$\int_0^{\frac{1}{2}} \log(1+\cos\pi x)dx = \int_0^{\frac{1}{2}} \log\left(2\cos^2\frac{\pi x}{2}\right)dx = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \log\left(2\cos\frac{t}{2}\right)dt - \frac{\log 2}{2} = \log\delta - \frac{\log 2}{2}. \qquad (4.3)$$

From (4.3), we have

$$\frac{\log \widehat{Q}_n(1)}{n} - \frac{\log 2}{2n} + \frac{1}{2n}\log\left(1+\cos\frac{\pi}{4n}\right) \geq \log\delta \geq \frac{\log \widehat{Q}_n(1)}{n} - \frac{1}{4n}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right).$$

Therefore, the following inequalities hold.

$$n\log\delta + \frac{1}{4}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right) \geq \log \widehat{Q}_n(1) \geq n\log\delta + \frac{\log 2}{2} - \frac{1}{2}\log\left(1+\cos\frac{\pi}{4n}\right). \qquad (4.4)$$

Since

$$\log\left(1+\cos\frac{(n-1)\pi}{2n}\right) = \log\left(1+\sin\frac{\pi}{2n}\right) < \log\left(1+\frac{\pi}{2n}\right) < \frac{\pi}{2n},$$

we can estimate the left-hand side of (4.4) by

$$n\log\delta + \frac{1}{4}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right) < n\log\delta + \frac{\pi}{8n}.$$

An estimation for the right-hand side of (4.4) is as follows. Since $\log 2 > \log\left(1+\cos\frac{\pi}{4n}\right)$, we have

$$n\log\delta + \frac{\log 2}{2} - \frac{1}{2}\log\left(1+\cos\frac{\pi}{4n}\right) > n\log\delta,$$

and therefore

$$n\log\delta + \frac{\pi}{8n} > \log \widehat{Q}_n(1) > n\log\delta$$

proves Lemma 2.5 for odd n. When n is even, we have

$$\log \widehat{Q}_n(1) = \frac{n\log 2}{2} + \sum_{k=1}^{\frac{n}{2}} \log\left(1+\cos\frac{(2k-1)\pi}{2n}\right).$$

Applying (4.2) to $f = \log(1+\cos\pi x)$ on intervals $[\frac{k}{n}, \frac{k+1}{n}]$ ($k = 0, 1, \ldots, \frac{n}{2}-1$), $[\frac{n-1}{2n}, \frac{1}{2}]$ for an upper estimation, and on intervals $[0, \frac{1}{2n}]$, $[\frac{2k-1}{2n}, \frac{2k+1}{2n}]$ ($k = 1, 2, \ldots, \frac{n}{2}-1$), $[\frac{n-1}{2n}, \frac{1}{2}]$ for a lower estimation, we have

$$\begin{aligned}\frac{\log \widehat{Q}_n(1)}{n} - \frac{\log 2}{2} &\geq \int_0^{\frac{1}{2}} \log(1+\cos\pi x)\,dx \\ &\geq \frac{\log \widehat{Q}_n(1)}{n} - \frac{\log 2}{2} - \frac{1}{2n}\log\left(1+\cos\frac{\pi}{2n}\right) + \frac{\log 2}{2n} - \frac{1}{4n}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right).\end{aligned}$$

Therefore, the inequalities

$$n\log\delta + \frac{1}{2}\log\left(1+\cos\frac{\pi}{2n}\right) - \frac{\log 2}{2} + \frac{1}{4}\log\left(1+\cos\frac{(n-1)\pi}{2n}\right) \geq \log \widehat{Q}_n(1) \geq n\log\delta$$

hold, and from similar arguments for odd n, the inequalities

$$n\log\delta + \frac{\pi}{8n} > \log \widehat{Q}_n(1) > n\log\delta$$

prove Lemma 2.5 for even n. ■

To prove Theorem 2.7, we need the following lemmas, corollaries and algorithm.

Lemma 4.1. *Let F and G be nonzero complex univariate polynomials, and ζ be a fixed complex number on the unit circle. Define $\nu : \mathbb{R} \to \mathbb{R}$ by*

$$\nu(r) := \frac{\|(t - r\zeta)F\|}{\|(t - r\zeta)G\|}.$$

Then, the following statements hold.

1. *$\nu(r)$ has a minimum at either $r = 1$ or -1.*
2. *If $\nu(1)$ is not a minimum, then $\nu(r) > \nu(0)$ for any $r > 0$.*

Proof. Since $\nu(r)$ is nonnegative, it is sufficient to prove that $N(r) = \nu(r)^2$ has the above properties. For $P(t) = \sum_{k=0}^{n} a_k t^k$, we have

$$\|(t - r\zeta)P\|^2 = \|P\|^2 r^2 - \left(\zeta \sum_{k=1}^{n} \overline{a}_{k-1} a_k + \overline{\zeta} \sum_{k=1}^{n} a_{k-1} \overline{a}_k \right) r + \|P\|^2. \tag{4.5}$$

Therefore, we can write

$$\begin{aligned} \|(t - r\zeta)F\|^2 &= f_1 r^2 + f_2 r + f_1, \\ \|(t - r\zeta)G\|^2 &= g_1 r^2 + g_2 r + g_1, \end{aligned}$$

where $f_1 = \|F\|^2$, $g_1 = \|G\|^2$, and f_2, g_2 are real numbers. Therefore, we have

$$N'(r) = \frac{(f_1 g_2 - f_2 g_1)(r^2 - 1)}{\|(t - r\zeta)G\|^2}.$$

If $f_1 g_2 - f_2 g_1 = 0$, then $N(r)$ is constant and the statements are clear.

If $f_1 g_2 - f_2 g_1 > 0$, then $N(r)$ tends to $f_1/g_1 = N(0)$, as r tends to $\pm\infty$. Furthermore, $N(r)$ is monotonically increasing on $(-\infty, -1]$, monotonically decreasing on $[-1, 1]$ and monotonically increasing on $[1, \infty)$. Therefore, $N(r)$ has a minimum at $r = 1$.

If $f_1 g_2 - f_2 g_1 < 0$, then similar arguments hold. We have $N(r) > N(0)$ for any $r > 0$ and $N(-1)$ is a minimum. ■

The following corollary immediately follows from Lemma 4.1.

Corollary 4.2. *Let F and G be nonzero complex univariate polynomials in t, and α be a nonzero complex number. Put $\zeta = \alpha/|\alpha|$. (That is, $|\zeta| = 1$.) Then, the following inequality holds.*

$$\frac{\|(t - \alpha)F\|}{\|(t - \alpha)G\|} \geq \min \left\{ \frac{\|(t - \zeta)F\|}{\|(t - \zeta)G\|}, \frac{\|tF\|}{\|tG\|} \right\}.$$

When polynomials F and G are real, the following lemma holds.

Lemma 4.3. *Let F and G be nonzero real univariate polynomials in t, and r be a fixed nonzero real number. Define $\nu : \mathbb{C} \to \mathbb{R}$ by*

$$\nu(\zeta) := \frac{\|(t - r\zeta)F\|}{\|(t - r\zeta)G\|}.$$

We consider $\nu(\zeta)$ a function on the unit circle in $\mathbb{C}$. Then, $\nu(\zeta)$ has a minimum at $\zeta = -1$ or 1.

Proof. Since $\nu(\zeta)$ is nonnegative, it is sufficient to prove that $N(\zeta) = \nu(\zeta)^2$ has a minimum at $\zeta = -1$ or 1. From Equation (4.5) and considering F, $G \in \mathbb{R}[t]$, we have

$$\begin{aligned} \|(t - r\zeta)F\|^2 &= f_1(\zeta + \overline{\zeta}) + f_2, \\ \|(t - r\zeta)G\|^2 &= g_1(\zeta + \overline{\zeta}) + g_2, \end{aligned}$$

where $f_2 = (r^2 + 1)\|F\|^2$, $g_2 = (r^2 + 1)\|G\|^2$ and f_1, $g_1 \in \mathbb{R}$. Put $s = \zeta + \overline{\zeta}$ $(\in \mathbb{R})$. We can write $N(\zeta)$ as $\widetilde{N}(s)$, which is a function of s $(-2 \leq s \leq 2)$. Then we have

$$\widetilde{N}'(s) = \frac{f_1 g_2 - f_2 g_1}{(g_1 s + g_2)^2},$$

That is, $\widetilde{N}'(s)$ is monotonic on $[-2, 2]$. Therefore, it has a minimum at $s = -2$ or 2, which corresponds to $\zeta = -1$ or 1, respectively. ∎

Combining Lemmas 4.1 and 4.3, we can easily see that the following corollary holds.

Corollary 4.4. *Let F and G be nonzero real univariate polynomials, and α be a complex number. Then, the following inequality holds.*

$$\frac{\|(t - \alpha)F\|}{\|(t - \alpha)G\|} \geq \min\left\{ \frac{\|(t - 1)F\|}{\|(t - 1)G\|}, \frac{\|(t + 1)F\|}{\|(t + 1)G\|} \right\}.$$

Finally, we describe the following algorithm.

Algorithm 4.5. Given a real polynomial $P(t) = (t - \alpha)(t - \overline{\alpha})P_0(t)$, where $\alpha \in \mathbb{C}, \notin \mathbb{R}$, this algorithm constructs $F \in \mathbb{R}[t]$ satisfying the following conditions.

1. The degree of F is two.
2. Both zeros of F lie on the unit circle.
3. F satisfies the following inequality.

$$\frac{\|PQ\|}{\|P\|} \geq \frac{\|FP_0Q\|}{\|FP_0\|}.$$

Step 1.: Put $\zeta = \alpha/|\alpha|$.

If

$$\frac{\|(t - \alpha)(t - \overline{\alpha})P_0Q\|}{\|(t - \alpha)(t - \overline{\alpha})P_0\|} \geq \frac{\|(t - \zeta)(t - \overline{\alpha})P_0Q\|}{\|(t - \zeta)(t - \overline{\alpha})P_0\|},$$

then go to Step 2.

Otherwise, go to Step 3.

Step 2.:

Step 2.1.: If

$$\frac{\|(t - \zeta)(t - \overline{\alpha})P_0Q\|}{\|(t - \zeta)(t - \overline{\alpha})P_0\|} \geq \frac{\|(t - \zeta)(t - \overline{\zeta})P_0Q\|}{\|(t - \zeta)(t - \overline{\zeta})P_0\|},$$

then terminate with the output $(t - \zeta)(t - \overline{\zeta})$.

Otherwise, go to Step 2.2.

Step 2.2.: If

$$\frac{\|t(t - \zeta)P_0Q\|}{\|t(t - \zeta)P_0\|} \geq \frac{\|t(t - 1)P_0Q\|}{\|t(t - 1)P_0\|},$$

then put $b_1 = 1$. Otherwise put $b_1 = -1$.

If

$$\frac{\|t(t - b_1)P_0Q\|}{\|t(t - b_1)P_0\|} \geq \frac{\|(t - 1)(t - b_1)P_0Q\|}{\|(t - 1)(t - b_1)P_0\|},$$

then put $b_2 = 1$. Otherwise put $b_2 = -1$.

Terminate with the output $(t - b_1)(t - b_2)$.

Step 3.: If

$$\frac{\|t(t-\overline{\alpha})P_0Q\|}{\|t(t-\overline{\alpha})P_0\|} \geq \frac{\|t(t-1)P_0Q\|}{\|t(t-1)P_0\|},$$

then put $b_3 = 1$. Otherwise, put $b_3 = -1$.

If

$$\frac{\|t(t-b_3)P_0Q\|}{\|t(t-b_3)P_0\|} \geq \frac{\|(t-1)(t-b_3)P_0Q\|}{\|(t-1)(t-b_3)P_0\|},$$

then put $b_4 = 1$. Otherwise put $b_4 = -1$.

Terminate with the output $(t - b_3)(t - b_4)$.

The validity of the algorithm is as follows. In Step 2.1, if the inequality does not hold, then we have

$$\frac{\|(t-\zeta)(t-\overline{\alpha})P_0Q\|}{\|(t-\zeta)(t-\overline{\alpha})P_0\|} \geq \frac{\|t(t-\zeta)P_0Q\|}{\|t(t-\zeta)P_0\|}$$

from Corollary 4.2.

In Step 2.2, the following inequalities hold from Corollary 4.4.

$$\frac{\|t(t-\zeta)P_0Q\|}{\|t(t-\zeta)P_0\|} \geq \frac{\|t(t-b_1)P_0Q\|}{\|t(t-b_1)P_0\|} \geq \frac{\|(t-b_1)(t-b_2)P_0Q\|}{\|(t-b_1)(t-b_2)P_0\|}.$$

In Step 3, the inequality

$$\frac{\|(t-\alpha)(t-\overline{\alpha})P_0Q\|}{\|(t-\alpha)(t-\overline{\alpha})P_0\|} \geq \frac{\|t(t-\overline{\alpha})P_0Q\|}{\|t(t-\overline{\alpha})P_0\|}$$

holds from Corollary 4.2. Furthermore, the inequalities

$$\frac{\|t(t-\overline{\alpha})P_0Q\|}{\|t(t-\overline{\alpha})P_0\|} \geq \frac{\|t(t-b_3)P_0Q\|}{\|t(t-b_3)P_0\|} \geq \frac{\|(t-b_3)(t-b_4)P_0Q\|}{\|(t-b_3)(t-b_4)P_0\|}$$

hold from Corollary 4.4.

Proof of Theorem 2.7. It is sufficient to show that the following two statements hold.

1. Given $P = (t-a)P_0$, where $a \in \mathbb{R}$, we can construct a real polynomial $R = (t-b)P_0$ ($b = 1$ or -1) satisfying the following inequality.
$$\frac{\|PQ\|}{\|P\|} \geq \frac{\|RQ\|}{\|R\|}.$$
2. Given $P = (t-\alpha)(t-\overline{\alpha})P_0$, where $\alpha \in \mathbb{C}, \notin \mathbb{R}$, we can construct a real polynomial $R = FP_0$ with the inequality
$$\frac{\|PQ\|}{\|P\|} \geq \frac{\|RQ\|}{\|R\|},$$
where F is a univariate real polynomial of degree two with both zeros on the unit circle.

The first statement and the second statement follow from Corollary 4.4 and Algorithm 4.5, respectively. ■

To prove Proposition 2.12, we need some lemmas.

Lemma 4.6. *Let $P(t)$ be a real univariate polynomial of degree n. For an integer $m > n$, the equality*

$$\|P\|^2 = \frac{1}{m}\sum_{k=1}^{m} |P(\omega\zeta^k)|^2$$

holds, where $\omega \in \mathbf{C}$ lies on the unit circle and ζ is a primitive m-th root of unity.

Lemma 4.7. *For arbitrary $\epsilon > 0$, there exists $\theta > 0$ such that the inequality*

$$1 - 2(1-\epsilon)x \geq \frac{1 - \sin x}{1 + \sin x}$$

holds for $\theta \geq x \geq 0$.

Proof. Since

$$\frac{1 - \sin x}{1 + \sin x} = 1 - \frac{2 \sin x}{1 + \sin x},$$

the inequality is equivalent to

$$\frac{\sin x}{1 + \sin x} \geq (1-\epsilon)x. \tag{4.6}$$

As x tends to 0,

$$\frac{\sin x}{x} \to 1, \qquad \frac{1}{1 + \sin x} \to 1,$$

hold. Therefore, for given $\epsilon > 0$, there exists $\theta > 0$ such that the inequality (4.6) holds for $\theta \geq x \geq 0$. ■

Lemma 4.8. *For arbitrary $\epsilon > 0$, there exists $\theta > 0$ such that the inequalities*

$$\exp(-x) \geq 1 - x \geq \exp(-(1+\epsilon)x)$$

hold for $\theta \geq x \geq 0$.

Lemma 4.9 (Jordan's Inequality). *For $\pi/2 \geq x \geq 0$,*

$$x \geq \sin x \geq \frac{2x}{\pi}.$$

Proof of Proposition 2.12. First we prove the proposition when n is odd. It is sufficient to show that for any $\epsilon > 0$, there exists an integer N such that the inequalities

$$\frac{1}{2\sqrt{1-\epsilon}} + \frac{1}{2\sqrt{n}} - \frac{\sqrt{n}}{2}\exp\left(-\frac{\lfloor n^{2/3} \rfloor^2}{n}\right) > \frac{\|\widehat{Q}_n\|^2 \sqrt{n}}{\|\widehat{Q}_n\|_D^2} > \frac{1}{2\sqrt{1+\epsilon}} - \frac{1}{\sqrt{n}} - \frac{\sqrt{n}\exp\left(-(1+\epsilon)\pi n^{1/3}\right)}{2(1+\epsilon)\pi} \tag{4.7}$$

hold for any odd integer $n \geq N$.

Let ζ be $\exp(\pi\sqrt{-1}/n)$. Then we have

$$\|\widehat{Q}_n\|^2 = \frac{1}{2n}\sum_{k=1}^{2n} |\widehat{Q}(\zeta^k)|^2 = \frac{\widehat{Q}_n(1)^2}{2n} + \frac{1}{n}\sum_{k=1}^{(n-1)/2} |\widehat{Q}_n(\zeta^k)|^2.$$

The relation between $|\widehat{Q}_n(\zeta^k)|^2$ and $|\widehat{Q}_n(\zeta^{k-1})|^2$ is as follows.

$$|\widehat{Q}_n(\zeta^k)|^2 = |\widehat{Q}_n(\zeta^{k-1})|^2 \frac{\left|\zeta^{k-1} + \zeta^{-\frac{n+1}{2}}\right|^2}{\left|\zeta^{k-1} + \zeta^{\frac{n-1}{2}}\right|^2} = |\widehat{Q}_n(\zeta^{k-1})|^2 \frac{\left|1 + \zeta^{-\frac{n-1}{2}-k}\right|^2}{\left|1 + \zeta^{\frac{n+1}{2}-k}\right|^2}.$$

Since the equalities

$$|1 + \zeta^j|^2 = (1 + \zeta^j)(1 + \zeta^{-j}) = 2\left(1 + \cos\frac{j\pi}{n}\right),$$

hold for $j \in \mathbb{N}$, we have

$$|\widehat{Q}_n(\zeta^k)|^2 = |\widehat{Q}_n(\zeta^{k-1})|^2 \frac{1 + \cos\left(\frac{\pi}{2} + \frac{(2k-1)\pi}{2n}\right)}{1 + \cos\left(\frac{\pi}{2} - \frac{(2k-1)\pi}{2n}\right)} = |\widehat{Q}_n(\zeta^{k-1})|^2 \frac{1 - \sin\frac{(2k-1)\pi}{2n}}{1 + \sin\frac{(2k-1)\pi}{2n}}. \tag{4.8}$$

First we show the upper bound. Take any $\epsilon > 0$. Then, there exists an integer L such that the above lemma holds for $\theta = L^{-1/3}\pi$. Take any $n \geq L$. Since we have

$$\frac{\pi}{L^{1/3}} \geq \frac{\pi}{n^{1/3}} \geq \frac{(2n^{2/3}-1)\pi}{2n} \geq \frac{(2k-1)\pi}{2n}$$

for $\lfloor n^{2/3} \rfloor \geq k \geq 1$, the following inequalities follow from Lemmas 4.7 and 4.8.

$$\frac{1-\sin\frac{(2k-1)\pi}{2n}}{1+\sin\frac{(2k-1)\pi}{2n}} \leq 1 - \frac{(1-\epsilon)(2k-1)\pi}{n} \leq \exp\left(-\frac{(1-\epsilon)(2k-1)\pi}{n}\right).$$

Therefore, for $\lfloor n^{2/3} \rfloor \geq k \geq 1$ we have

$$\begin{aligned} |\widehat{Q}_n(\zeta^k)|^2 &\leq |\widehat{Q}_n(\zeta^{k-1})|^2 \exp\left(-\frac{(1-\epsilon)(2k-1)\pi}{n}\right) \leq \widehat{Q}_n(1)^2 \prod_{j=1}^{k} \exp\left(-\frac{(1-\epsilon)(2j-1)\pi}{n}\right) \\ &= \widehat{Q}_n(1)^2 \exp\left(-\frac{(1-\epsilon)\pi}{n}\sum_{j=1}^{k}(2j-1)\right) = \widehat{Q}_n(1)^2 \exp\left(-\frac{(1-\epsilon)\pi}{n}k^2\right). \end{aligned}$$

Since the inequality

$$\frac{1-\sin\frac{(2k-1)\pi}{2n}}{1+\sin\frac{(2k-1)\pi}{2n}} \leq 1 - \sin\frac{(2k-1)\pi}{2n}$$

holds for $(n-1)/2 \geq k > \lfloor n^{2/3} \rfloor$, the following inequalities follow from Lemmas 4.8 and 4.9.

$$\frac{1-\sin\frac{(2k-1)\pi}{2n}}{1+\sin\frac{(2k-1)\pi}{2n}} \leq \exp\left(-\sin\frac{(2k-1)\pi}{2n}\right) \leq \exp\left(-\frac{2k-1}{n}\right).$$

Hence, for $(n-1)/2 \geq k > \lfloor n^{2/3} \rfloor$, we have

$$\begin{aligned} |\widehat{Q}_n(\zeta^k)|^2 &\leq |\widehat{Q}_n(\zeta^{k-1})|^2 \exp\left(-\frac{2k-1}{n}\right) \leq \widehat{Q}_n(1)^2 \prod_{j=1}^{k} \exp\left(-\frac{2j-1}{n}\right) \\ &= \widehat{Q}_n(1)^2 \exp\left(-\frac{1}{n}\sum_{j=1}^{k}(2j-1)\right) = \widehat{Q}_n(1)^2 \exp\left(-\frac{k^2}{n}\right). \end{aligned}$$

Therefore, the following inequality holds.

$$\frac{\widehat{Q}_n(1)^2}{2n} + \frac{\widehat{Q}_n(1)^2}{n} \sum_{k=1}^{\lfloor n^{2/3} \rfloor} \exp\left(-\frac{(1-\epsilon)\pi}{n}k^2\right) + \frac{\widehat{Q}_n(1)^2}{n} \sum_{k=\lfloor n^{2/3} \rfloor+1}^{(n-1)/2} \exp\left(-\frac{k^2}{n}\right) \geq \|\widehat{Q}_n\|^2.$$

Here,

$$\sum_{k=1}^{\lfloor n^{2/3} \rfloor} \exp\left(-\frac{(1-\epsilon)\pi}{n}k^2\right) < \int_0^\infty \exp\left(-\frac{(1-\epsilon)\pi}{n}x^2\right)dx = \frac{1}{2}\sqrt{\frac{n}{1-\epsilon}}$$

holds since

$$\int_0^\infty \exp(-cx^2)dx = \frac{1}{\sqrt{c}}\int_0^\infty \exp(-x^2)dx = \frac{1}{2}\sqrt{\frac{\pi}{c}}$$

holds for $c > 0$. Then we have

$$\begin{aligned} \sum_{k=\lfloor n^{2/3} \rfloor+1}^{(n-1)/2} \exp\left(-\frac{k^2}{n}\right) &< \int_{\lfloor n^{2/3} \rfloor}^\infty \exp\left(-\frac{x^2}{n}\right)dx < \int_{\lfloor n^{2/3} \rfloor}^\infty x\exp\left(-\frac{x^2}{n}\right)dx \\ &= \left[-\frac{n}{2}\exp\left(-\frac{x^2}{n}\right)\right]_{\lfloor n^{2/3} \rfloor}^\infty = -\frac{n}{2}\exp\left(-\frac{\lfloor n^{2/3} \rfloor^2}{n}\right). \end{aligned}$$

Hence, the following inequality holds.

$$\widehat{Q}(1)^2\left(\frac{1}{2n}+\frac{1}{2\sqrt{(1-\epsilon)n}}-\frac{1}{2}\exp\left(-\frac{\lfloor n^{2/3}\rfloor^2}{n}\right)\right)>\|\widehat{Q}_n\|^2.$$

Therefore, we obtain the upper bound. That is, the inequality

$$\frac{1}{2\sqrt{n}}+\frac{1}{2\sqrt{1-\epsilon}}-\frac{\sqrt{n}}{2}\exp\left(-\frac{\lfloor n^{2/3}\rfloor^2}{n}\right)>\frac{\|\widehat{Q}_n\|^2\sqrt{n}}{\widehat{Q}_n(1)^2} \tag{4.9}$$

holds for $n\geq L$.

Next, we show the lower bound. From (4.8) we have

$$|\widehat{Q}_n(\zeta^j)|^2>\widehat{Q}_n(1)^2\prod_{k=1}^{j}\left(1-\sin\frac{(2k-1)\pi}{2n}\right)^2.$$

Take any $\epsilon>0$. Then, there exists an integer M such that the above lemma holds for $\theta=M^{-1/3}\pi$. Take any $n\geq M$. Since for $\lfloor n^{2/3}\rfloor\geq k\geq 1$ we have

$$\frac{\pi}{M^{1/3}}\geq\frac{\pi}{n^{1/3}}\geq\frac{(2n^{2/3}-1)\pi}{2n}\geq\frac{(2k-1)\pi}{2n},$$

the following inequalities follow from Lemma 4.8.

$$1-\sin\frac{(2k-1)\pi}{2n}\geq 1-\frac{(2k-1)\pi}{2n}\geq\exp\left(\frac{-(1+\epsilon)(2k-1)\pi}{2n}\right).$$

Hence, for $\lfloor n^{2/3}\rfloor\geq k\geq 1$ we have

$$|\widehat{Q}_n(\zeta^k)|^2>\widehat{Q}_n(1)^2\exp\left(\sum_{j=1}^{k}\frac{-(1+\epsilon)(2k-1)\pi}{n}\right)=\widehat{Q}_n(1)^2\exp\left(\frac{-(1+\epsilon)\pi j^2}{n}\right).$$

Therefore, the following inequalities hold.

$$\begin{aligned}\|\widehat{Q}_n\|^2 &> \frac{1}{n}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}|\widehat{Q}_n(\zeta^k)|^2>\frac{1}{n}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}\left(\widehat{Q}_n(1)^2\exp\left(\frac{-(1+\epsilon)\pi k^2}{n}\right)\right)\\ &= \frac{\widehat{Q}_n(1)^2}{n}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}\exp\left(\frac{-(1+\epsilon)\pi k^2}{n}\right).\end{aligned}$$

The following estimation holds.

$$\sum_{k=1}^{\lfloor n^{2/3}\rfloor}\exp\left(\frac{-(1+\epsilon)\pi k^2}{n}\right)>\int_1^{\lfloor n^{2/3}\rfloor+1}\exp\left(\frac{-(1+\epsilon)\pi x^2}{n}\right)dx.$$

For $a>0$ and $c\geq 1$ we have

$$\begin{aligned}\int_1^a\exp(-cx^2)dx &= \int_0^\infty\exp(-cx^2)dx-\int_0^1\exp(-cx^2)dx-\int_a^\infty\exp(-cx^2)dx\\ &> \frac{\sqrt{\pi}}{2\sqrt{c}}-1-\int_a^\infty x\exp(-cx^2)dx\end{aligned}$$

and

$$\int_a^\infty x\exp(-cx^2)dx=\left[-\frac{\exp(-cx^2)}{2c}\right]_a^\infty=\frac{\exp(-ca^2)}{2c}.$$

Hence, the inequalities

$$\begin{aligned}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}\exp\left(\frac{-(1+\epsilon)\pi k^2}{n}\right) &> \frac{\sqrt{n}}{2\sqrt{1+\epsilon}}-1-\frac{n\exp\left(-\frac{(1+\epsilon)\pi}{n}(\lfloor n^{2/3}\rfloor+1)^2\right)}{2(1+\epsilon)\pi}\\ &> \frac{\sqrt{n}}{2\sqrt{1+\epsilon}}-1-\frac{n\exp\left(-(1+\epsilon)\pi n^{1/3}\right)}{2(1+\epsilon)\pi}\end{aligned}$$

hold. Therefore, we have

$$\begin{aligned}\|\widehat{Q}_n\|^2 &> \frac{1}{n}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}|\widehat{Q}_n(\zeta^k)|^2 > \frac{\widehat{Q}_n(1)^2}{n}\sum_{k=1}^{\lfloor n^{2/3}\rfloor}\exp\left(\frac{-(1+\epsilon)\pi j^2}{n}\right)\\ &> \frac{\widehat{Q}_n(1)^2}{n}\int_1^{\lfloor n^{2/3}\rfloor+1}\exp\left(\frac{-(1+\epsilon)\pi x^2}{n}\right)dx\\ &> \widehat{Q}_n(1)^2\left(\frac{1}{2\sqrt{n(1+\epsilon)}}-\frac{1}{n}-\frac{\exp\left(-(1+\epsilon)\pi n^{1/3}\right)}{2(1+\epsilon)\pi}\right).\end{aligned}$$

Then, we obtain the lower bound. That is, the inequality

$$\frac{\|\widehat{Q}_n\|^2\sqrt{n}}{\widehat{Q}_n(1)^2} > \frac{1}{2\sqrt{1+\epsilon}}-\frac{1}{\sqrt{n}}-\frac{\sqrt{n}\exp\left(-(1+\epsilon)\pi n^{1/3}\right)}{2(1+\epsilon)\pi} \tag{4.10}$$

holds for $n \geq M$. Combining (4.9) and (4.10), we have the statement (4.7) for $N = \max\{L, M\}$ when n is odd.

Next we prove the statement when n is even. Let ζ_{4n} be $\exp(\pi\sqrt{-1}/2n)$. Note that ζ_{4n}^2 is a primitive $2n$-th root of unity. Then, we have

$$\|\widehat{Q}_n\|^2 = \frac{1}{2n}\sum_{k=1}^{2n}|\widehat{Q}(\zeta_{4n}^{2k-1})|^2 = \frac{1}{n}\sum_{k=1}^{n/2}|\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2.$$

The relation between $|\widehat{Q}_n(\zeta_{4n}^{2k+1})|^2$ and $|\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2$ is as follows.

$$|\widehat{Q}_n(\zeta_{4n}^{2k+1})|^2 = |\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2\frac{|\zeta_{4n}^{2k-1}+\zeta_{4n}^{-n-1}|^2}{|\zeta_{4n}^{2k-1}+\zeta_{4n}^{n-1}|^2} = |\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2\frac{|1+\zeta_{4n}^{-n-2k}|^2}{|1+\zeta_{4n}^{n-2k}|^2}.$$

Since

$$|1+\zeta_{4n}^j|^2 = (1+\zeta_{4n}^j)(1+\zeta_{4n}^{-j}) = 2\left(1+\cos\frac{j\pi}{2n}\right),$$

we have

$$|\widehat{Q}_n(\zeta_{4n}^{2k+1})|^2 = |\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2\frac{1+\cos(\frac{\pi}{2}+\frac{k\pi}{n})}{1+\cos(\frac{\pi}{2}-\frac{k\pi}{n})} = |\widehat{Q}_n(\zeta_{4n}^{2k-1})|^2\frac{1-\sin\frac{k\pi}{n}}{1+\sin\frac{k\pi}{n}}.$$

From similar arguments for odd n, given $\epsilon > 0$ there exists an integer N such that the following inequalities hold for any even integer $n \geq N$.

$$\frac{1}{2\sqrt{1-\epsilon}}-\frac{\sqrt{n}}{2}\exp\left(-\frac{\lfloor n^{2/3}\rfloor^2}{n}\right) > \frac{\|\widehat{Q}_n\|^2\sqrt{n}}{|\widehat{Q}_n(\zeta_{4n})|^2} > \frac{1}{2\sqrt{1+\epsilon}}-\frac{1}{\sqrt{n}}-\frac{\sqrt{n}\exp\left(-(1+\epsilon)\pi n^{1/3}\right)}{2(1+\epsilon)\pi}.$$

That is, we have

$$\frac{\|\widehat{Q}_n\|^2\sqrt{n}}{|\widehat{Q}_n(\zeta_{4n})|^2} \to \frac{1}{2} \tag{4.11}$$

as n tends to infinity.

According to the following Lemma, we have

$$\lim_{n\to\infty} \frac{|\widehat{Q}_n(\zeta_{4n})|^2}{\|\widehat{Q}_n\|_D^2} = 1,$$

and combining with (4.11), we have the statement. ∎

Lemma 4.10.

$$\lim_{n\to\infty} \frac{|\widehat{Q}_n(\zeta_{4n})|^2}{\|\widehat{Q}_n\|_D^2} = 1.$$

Proof. Since the following inequalities

$$\widehat{Q}_n(1) = \prod_{k=-n/2+1}^{n/2} (1+\zeta_{4n}^{2k-1}), \qquad \widehat{Q}_n(\zeta_{4n}) = \prod_{k=-n/2+1}^{n/2} (\zeta_{4n}+\zeta_{4n}^{2k-1})$$

hold, we have

$$\begin{aligned}
\widehat{Q}_n(1)^2 &= \prod_{k=-n/2+1}^{n/2} (1+\zeta_{4n}^{2k-1})^2 = \prod_{k=1}^{n/2} (1+\zeta_{4n}^{2k-1})^2(1+\zeta_{4n}^{-2k+1})^2 = \prod_{k=1}^{n/2}\left(2+2\cos\frac{(2k-1)\pi}{2n}\right)^2, \\
|\widehat{Q}_n(\zeta_{4n})|^2 &= \prod_{k=-n/2+1}^{n/2} |1+\zeta_{4n}^{2k-2}|^2 = \prod_{k=1}^{n/2} |1+\zeta_{4n}^{2k-2}|^2 \cdot |1+\zeta_{4n}^{-2k}|^2 \\
&= \prod_{k=1}^{n/2}\left(2\cos\frac{\pi}{2n}+2\cos\frac{(2k-1)\pi}{2n}\right)^2.
\end{aligned}$$

Therefore, the following inequalities hold.

$$\begin{aligned}
\frac{|\widehat{Q}_n(\zeta_{4n})|^2}{\widehat{Q}_n(1)^2} &= \prod_{k=1}^{n/2} \frac{\left(\cos\frac{\pi}{2n}+\cos\frac{(2k-1)\pi}{2n}\right)^2}{\left(1+\cos\frac{(2k-1)\pi}{2n}\right)^2} \geq \prod_{k=1}^{n/2} \frac{\left(1-\frac{\pi^2}{8n^2}+\cos\frac{(2k-1)\pi}{2n}\right)^2}{\left(1+\cos\frac{(2k-1)\pi}{2n}\right)^2} \\
&= \prod_{k=1}^{n/2}\left(1-\frac{\pi^2}{8n^2\left(1+\cos\frac{(2k-1)\pi}{2n}\right)^2}\right)^2 > \left(1-\frac{\pi^2}{8n^2}\right)^n.
\end{aligned}$$

That is, we have

$$1 \geq \frac{|\widehat{Q}_n(\zeta_{4n})|^2}{\widehat{Q}_n(1)^2} > \left(1-\frac{\pi^2}{8n^2}\right)^n.$$

Therefore, we have

$$\frac{|\widehat{Q}_n(\zeta_{4n})|^2}{\widehat{Q}_n(1)^2} \to 1$$

as n tends to infinity. ∎

Acknowledgement

The first author wishes to thank Prashant Batra, Ludwig Elsner and Arnold Schönhage for their valuable hints and fruitful discussion. The authors thank Kyoko Makino for computing the rigorous lower and upper bounds in Table 1. Furthermore our thanks to two unknown referees for their constructive comments.

References

[1] M. Berz. From Taylor series to Taylor models. In *Nonlinear problems in accelerator physics, AIP Conference proceedings*, number CP405, pages 1–27, 1997.

[2] A. Böttcher and S. Grudsky. *Spectral properties of banded Toeplitz matrices*. SIAM, Philadelphia, 2005.

[3] A. Böttcher and S. Grudsky. Structured condition numbers of large Toeplitz matrices are rarly better than usual condition numbers. *Numerical Linear Algebra and its Applications*, 12:95–102, 2005.

[4] A. Böttcher and K. Rost. Topics in the numerical linear algebra of Toeplitz and Hankel matrices. *Mitt. Ges. Angew. Math. Mech.*, 27(2):174–188, 2004.

[5] D.W. Boyd. Two sharp inequalities for the norm of a factor of a polynomial. *Mathematika*, 39:341–349, 1992.

[6] D.W. Boyd. Sharp inequalities for the product of polynomials. *Bull. London Math. Soc.*, 26:449–454, 1994.

[7] S. Dallwig, A. Neumaier, and H. Schichl. GLOPT - a program for constrained global optimization. In I. M. Bomze et al., editor, *Developments in global optimization*, pages 19–36. Kluwer Academic Publishers, 1997.

[8] I. Gohberg and A.A. Semencul. The inversion of finite Toeplitz matrices and their continual analogues. *Matem. Issled*, 7:201–223, 1972.

[9] G. Heinig. private communication.

[10] D.J. Higham and N.J. Higham. Backward error and condition of structured linear systems. *SIAM J. Matrix Anal. Appl.*, 13(1):162–175, 1992.

[11] H. Kneser. Das Maximum des Produkts zweier Polynome. In *Sitz. Preuss. Akad. Wiss., Phys.-Math. Kl.*, pages 426–431, 1934.

[12] A. Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica*, 13:271–369, 2004.

[13] S.M. Rump. Structured perturbations Part I: Normwise distances. *SIAM J. Matrix Anal. Appl. (SIMAX)*, 25(1):1–30, 2003.

[14] E.B. Saff and T. Sheil-Small. Coefficient and integral mean estimates for albgebraic and trigonometric polynomials with restricted zeros. *J. London Math. Soc.*, 9:16–22, 1974.

[15] T. Sheil-Small. *Complex polynomials*, volume 75 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2002.

Siegfried M. Rump
Institute for Reliable Computing
Hamburg University of Technology
Schwarzenbergstraße 95
Hamburg 21071
Germany

and

Waseda University
Faculty of Science and Engineering
3–4–1 Okubo, Shinjuku-ku
Tokyo 169–8555
Japan

e-mail: `rump@tu-harburg.de`

H. Sekigawa
NTT Communication Science Laboratories
Nippon Telegraph and Telephone Corporation
3–1 Morinosato-Wakamiya, Atsugi-shi
Kanagawa, 243-0198
Japan

e-mail: `sekigawa@theory.brl.ntt.co.jp`