Non-Asymptotic Preconditioned Krylov, Accretive GMRES, Power-Iterated Nyström, and Deflated Lanczos Envelopes

Lluis Eriksson
Independent Researcher, Östersund, Sweden
editor@airr.science

November 29, 2026

Abstract

We establish four self-contained, explicit, non-asymptotic quantitative theorems unifying preconditioned Krylov linear solvers, non-normal GMRES, randomized Nyström low-rank approximation, and symmetric Lanczos eigensolvers. First, for symmetric positive definite (SPD) systems $Ax = f$ paired with a perturbed SPD preconditioner $\widetilde{M} = M + \Delta M$ of relative Loewner error $\delta := \|M^{-1/2}\Delta M M^{-1/2}\|_2 < 1$, we prove the sharp, attainable two-sided condition-number sandwich $\max\big(1, \frac{1-\delta}{1+\delta}\kappa\big) \le \widetilde{\kappa} \le \frac{1+\delta}{1-\delta}\kappa$, a finite-step two-sided outlier-deflated Chebyshev conjugate-gradient (CG) error ceiling valid across arbitrary upper and lower spectral outliers (including degenerate bulk spectra $a = b$), and its explicit perturbed-preconditioner transfer (Theorem 2.1). Second, for strictly accretive non-normal matrices $A = H + iS$ ($H \succ 0$, $S = S^*$), we establish the exact Hermitian-inverse identity $H(A^{-1}) = H^{-1/2}(I + B^2)^{-1}H^{-1/2}$ with $B := H^{-1/2}SH^{-1/2}$ ($\sigma := \|B\|_2$), deriving the sharp k-step GMRES contraction $\|r_k\|_2 \le (1 - \alpha\gamma)^{k/2}\|r_0\|_2$ that uniformly improves upon Elman's bound (strictly whenever $\ker(H - \alpha I) \cap \ker(A^*A - \|A\|_2^2 I) = \{0\}$, in particular for every non-scalar SPD matrix), alongside a strictly contractive imaginary-axis Chebyshev ceiling $\Phi_k(\sigma) < 1$ for $H^{-1/2}$-preconditioned GMRES (Theorem 3.1). Third, for q-step power-iterated Nyström approximations $\widehat{A}_q^{\mathrm{Nys}}$ of PSD matrices, we construct an exact $(n-k)\times(n-k)$ reduced-residual operator R_q unitarily equivalent on its nonzero spectrum to the dominating residual, yielding a two-regime saturating spectral ceiling that remains finite as the subspace tangent $\|G\|_2 \to \infty$ and an exact Gaussian spectral-moment trace expectation $\mathbb{E}[\mathrm{tr}(A - \widehat{A}_q^{\mathrm{Nys}})] \le \mathrm{tr}(\Lambda_2) + (p-1)^{-1}\,\mathrm{tr}(\Lambda_2^{2q+1})\,\mathrm{tr}(\Lambda_1^{-2q})$ (Theorem 4.1). Fourth, for symmetric Lanczos top Ritz pairs $(\theta_1^{(m)}, y_1^{(m)})$, we prove a saturating s-cluster deflated Kaniel–Paige–Saad envelope bounded by $\lambda_1 - \lambda_n$ even as $c_1 \to 0^+$, paired with exact attainable Bhatia–Davis lower and Temple–Kato upper residual-to-angle sandwiches (Theorem 5.1). All four theorems, exact attainments, and boundary regimes are deterministically verified in an MIT-licensed Python test suite.

1 Introduction and Scope

Iterative Krylov methods and randomized sketching algorithms are the foundational computational primitives of large-scale numerical linear algebra [1, 2, 5]. Classical convergence analyses, however, frequently suffer from four practical non-asymptotic deficiencies: (i) standard conjugate-gradient (CG) Chebyshev envelopes deteriorate dramatically in the presence of even a handful of extreme outliers and rarely quantify the effect of an inexact or perturbed preconditioner $\widetilde{M} = M + \Delta M$; (ii) Elman's classical non-normal GMRES contraction $\sqrt{1 - \alpha^2/\|A\|_2^2}$ loses a full square in condition number even on SPD matrices, while naive odd-degree Chebyshev polynomials on shifted imaginary segments $1 + i[-\sigma, \sigma]$ exceed unity for large σ (already at $k = 1$ whenever $\sigma > 1$); (iii) deterministic randomized Nyström spectral ceilings of the form $\lambda_{k+1}(1+\eta^{2q}\|G\|_2^2)$ diverge to $+\infty$ as the sketching tangent $\|G\|_2$ grows, and classical gap-ratio trace bounds ignore the spectral decay of both head and tail eigenvalues; and (iv) classical Kaniel–Paige–Saad (KPS) Ritz-value ceilings diverge as the starting-vector overlap $c_1 \to 0^+$ and stall on tight top eigenvalue clusters.

In this work we resolve all four limitations in a unified, self-contained framework of four explicit, non-asymptotic theorems (Sections 2–5) whose proofs require only elementary matrix congruences, Chebyshev extremal identities, orthogonal-projection factorizations, and inverse-Wishart expectations. Every constant is explicit, every boundary and degenerate regime is treated without gaps, and every theorem is verified computationally in Section 6 via a deterministic companion repository (https://github.com/lluiseriksson/krylov-nystrom-preconditioning-geometry, tag `v1.0.0`).

Notation. Throughout, $\mathbb{S}^n$, $\mathbb{S}^n_+$, and $\mathbb{S}^n_{++}$ denote the real symmetric, symmetric positive semidefinite (PSD), and symmetric positive definite (SPD) $n \times n$ matrices, with Loewner partial order $X \preceq Y$ iff $Y - X \in \mathbb{S}^n_+$. For $A \in \mathbb{C}^{n\times n}$, A^* denotes the conjugate transpose, $H(A) := \frac{1}{2}(A + A^*)$ its Hermitian part, and $\|A\|_2$ its spectral norm. For $X \in \mathbb{R}^{p\times q}$, X^+ is the Moore–Penrose pseudoinverse, $P_X := X(X^\top X)^+ X^\top$ is the orthogonal projector onto range(X), and $\kappa_2(W) := \lambda_{\max}(W)/\lambda_{\min}(W)$ for $W \in \mathbb{S}^n_{++}$. We write T_d for the degree-d Chebyshev polynomial of the first kind ($T_0(x) = 1$, $T_1(x) = x$, $T_{d+1}(x) = 2xT_d(x) - T_{d-1}(x)$), adopting standard conventions $\prod_\emptyset = 1$ and $0^0 = 1$.

2 Loewner Preconditioner Sandwiches and Deflated Chebyshev–CG

Theorem 2.1 (Preconditioner Loewner Sandwich and Two-Sided Outlier-Deflated PCG)**.** *Let $A, M \in \mathbb{S}^n_{++}$ ($n \geq 2$), let $\widetilde{M} = M + \Delta M \in \mathbb{S}^n_{++}$ satisfy $\delta := \|M^{-1/2}\Delta M M^{-1/2}\|_2 \in [0, 1)$, and define $W := M^{-1/2}AM^{-1/2}$ and $\widetilde{W} := \widetilde{M}^{-1/2}A\widetilde{M}^{-1/2}$ with eigenvalues $\lambda_1(W) \geq \cdots \geq \lambda_n(W) > 0$, $\lambda_1(\widetilde{W}) \geq \cdots \geq \lambda_n(\widetilde{W}) > 0$, and condition numbers $\kappa := \kappa_2(W)$ and $\widetilde{\kappa} := \kappa_2(\widetilde{W})$.*

(a) ***Sharp Two-Sided Spectral and Condition-Number Sandwich.*** *For every $i \in \{1, \ldots, n\}$,*

$$\frac{\lambda_i(W)}{1+\delta} \;\leq\; \lambda_i(\widetilde{W}) \;\leq\; \frac{\lambda_i(W)}{1-\delta}, \qquad \max\!\left(1,\, \frac{1-\delta}{1+\delta}\,\kappa\right) \;\leq\; \widetilde{\kappa} \;\leq\; \frac{1+\delta}{1-\delta}\,\kappa. \tag{1}$$

For every $\delta \in [0, 1)$, the upper condition-number bound is exactly attained for every $\kappa \geq 1$, and the lower bound is exactly attained whenever $\kappa \geq \frac{1+\delta}{1-\delta}$.

(b) ***Two-Sided Outlier-Deflated Chebyshev–CG Ceiling.*** *Suppose the spectrum of W partitions into $m \geq 0$ upper outliers $\lambda_1 \geq \cdots \geq \lambda_m > b$, a bulk subset inside $[a, b]$ with $0 < a \leq b$, and $r \geq 0$ lower outliers $0 < \mu_1 \leq \cdots \leq \mu_r < a$ (with $m + r < n$). Let $\rho_{\text{bulk}} := \frac{\sqrt{b/a}-1}{\sqrt{b/a}+1} \in [0, 1)$. Then for every integer $k \geq m + r$ and any right-hand side $f \in \mathbb{R}^n$ and initial guess $x_0 \in \mathbb{R}^n$, the exact-arithmetic preconditioned CG error $e_k := x_k - A^{-1}f$ on (A, M) satisfies*

$$\|e_k\|_A \;\leq\; \left(\prod_{j=1}^{m}\left(1 - \frac{a}{\lambda_j}\right)\right)\left(\prod_{i=1}^{r}\left(\frac{b}{\mu_i} - 1\right)\right)\frac{2\,\rho_{\text{bulk}}^{k-m-r}}{1 + \rho_{\text{bulk}}^{2(k-m-r)}}\,\|e_0\|_A. \tag{2}$$

In particular, when $a = b$ ($\rho_{\text{bulk}} = 0$), (2) holds at $k = m + r$ (with $0^0 = 1$) and gives exact termination $\|e_k\|_A = 0$ for all $k \geq m + r + 1$.

(c) ***Perturbed-Preconditioner Bulk Transfer.*** *If, in addition to (b), $\lambda_m > \frac{1+\delta}{1-\delta}b$ (when $m \geq 1$) and $\mu_r < \frac{1-\delta}{1+\delta}a$ (when $r \geq 1$), then with $\widetilde{a} := \frac{a}{1+\delta}$, $\widetilde{b} := \frac{b}{1-\delta}$, and $\widetilde{\rho}_{\text{bulk}} := \frac{\sqrt{\frac{1+\delta}{1-\delta}\frac{b}{a}}-1}{\sqrt{\frac{1+\delta}{1-\delta}\frac{b}{a}}+1}$, the preconditioned CG error $\widetilde{e}_k := \widetilde{x}_k - A^{-1}f$ on $(A, \widetilde{M})$ starting from the same initial iterate $\widetilde{x}_0 = x_0$ (so $\widetilde{e}_0 = e_0$) satisfies for every $k \geq m + r$:*

$$\|\widetilde{e}_k\|_A \;\leq\; \left(\prod_{j=1}^{m}\left(1 - \frac{1-\delta}{1+\delta}\,\frac{a}{\lambda_j}\right)\right)\left(\prod_{i=1}^{r}\left(\frac{1+\delta}{1-\delta}\,\frac{b}{\mu_i} - 1\right)\right)\frac{2\,\widetilde{\rho}_{\text{bulk}}^{k-m-r}}{1 + \widetilde{\rho}_{\text{bulk}}^{2(k-m-r)}}\,\|e_0\|_A. \tag{3}$$

Proof. (a) By definition of $\delta = \|M^{-1/2}\Delta M M^{-1/2}\|_2 < 1$, we have $-\delta I \preceq M^{-1/2}\Delta M M^{-1/2} \preceq \delta I$. Congruence by $M^{1/2}$ yields $(1-\delta)M \preceq \widetilde{M} \preceq (1+\delta)M$. Because matrix inversion is operator-antitone on $\mathbb{S}^n_{++}$, $(1+\delta)^{-1}M^{-1} \preceq \widetilde{M}^{-1} \preceq (1-\delta)^{-1}M^{-1}$. Congruence by $A^{1/2}$ gives

$$(1+\delta)^{-1}A^{1/2}M^{-1}A^{1/2} \;\preceq\; A^{1/2}\widetilde{M}^{-1}A^{1/2} \;\preceq\; (1-\delta)^{-1}A^{1/2}M^{-1}A^{1/2}.$$

Now $W = M^{-1/2}AM^{-1/2}$ is similar to $M^{-1}A$, which is similar to $A^{1/2}M^{-1}A^{1/2}$, so $\lambda_i(W) = \lambda_i(A^{1/2}M^{-1}A^{1/2})$ for every i, and identically $\lambda_i(\widetilde{W}) = \lambda_i(A^{1/2}\widetilde{M}^{-1}A^{1/2})$. Weyl's monotonicity theorem immediately yields the first inequality of (1). Dividing the $i=1$ inequality by the $i=n$ inequality and noting $\widetilde{\kappa} \geq 1$ gives the condition-number sandwich. For exact attainment for any $n \geq 2$, take $A = \mathrm{diag}(\kappa, 1, \kappa I_{n-2})$ and $M = I_n$ (so $\kappa_2(W) = \kappa$):

- setting $\widetilde{M} = \mathrm{diag}(1-\delta, 1+\delta, (1+\delta)I_{n-2})$ gives eigenvalues $\frac{\kappa}{1-\delta} \geq \frac{\kappa}{1+\delta} \geq \frac{1}{1+\delta}$ for $\widetilde{W}$, hence $\widetilde{\kappa} = \frac{1+\delta}{1-\delta}\kappa$ for every $\kappa \geq 1$;
- setting $\widetilde{M} = \mathrm{diag}(1+\delta, 1-\delta, (1+\delta)I_{n-2})$ gives eigenvalues $\frac{\kappa}{1+\delta}$ (multiplicity $n-1$) and $\frac{1}{1-\delta}$ for $\widetilde{W}$, so whenever $\kappa \geq \frac{1+\delta}{1-\delta}$ we have $\frac{\kappa}{1+\delta} \geq \frac{1}{1-\delta}$ and hence $\widetilde{\kappa} = \frac{1-\delta}{1+\delta}\kappa$.

(b) By the classical Krylov minimax optimality of exact-arithmetic preconditioned CG [1],

$$\|e_k\|_A \;\leq\; \min_{\substack{p\in\mathcal{P}_k \\ p(0)=1}} \max_{\lambda\in\mathrm{spec}(W)} |p(\lambda)|\,\|e_0\|_A.$$

Set $d := k-m-r \geq 0$ and define the outlier-annihilating factor $q(\lambda) := \prod_{j=1}^{m}(1-\lambda/\lambda_j)\prod_{i=1}^{r}(1-\lambda/\mu_i)$, which satisfies $\deg(q) = m+r$, $q(0)=1$, and $q(\lambda_j) = q(\mu_i) = 0$ for all outliers. For every $\lambda \in [a,b]$:

- each upper outlier $\lambda_j > b$ satisfies $0 < 1-b/\lambda_j \leq 1-\lambda/\lambda_j \leq 1-a/\lambda_j < 1$, so $|1-\lambda/\lambda_j| \leq 1-a/\lambda_j$;
- each lower outlier $0 < \mu_i < a$ satisfies $1-\lambda/\mu_i \leq 1-a/\mu_i < 0$, so $|1-\lambda/\mu_i| = \lambda/\mu_i - 1 \leq b/\mu_i - 1$.

If $a < b$, set $p_k(\lambda) := q(\lambda)\,T_d\!\left(\frac{b+a-2\lambda}{b-a}\right)/T_d\!\left(\frac{b+a}{b-a}\right) \in \mathcal{P}_k$. Then $p_k(0)=1$, p_k vanishes at all $m+r$ outliers, and on $[a,b]$ we have $\left|\frac{b+a-2\lambda}{b-a}\right| \leq 1$ so $|T_d(\frac{b+a-2\lambda}{b-a})| \leq 1$. Writing $c := \frac{b+a}{b-a} = \frac{1+\rho_{\mathrm{bulk}}^2}{2\rho_{\mathrm{bulk}}} > 1$, the standard identity $T_d(\frac{t+t^{-1}}{2}) = \frac{t^d+t^{-d}}{2}$ at $t = \rho_{\mathrm{bulk}}^{-1}$ gives $T_d(c)^{-1} = \frac{2\rho_{\mathrm{bulk}}^d}{1+\rho_{\mathrm{bulk}}^{2d}}$, proving (2). If $a = b$ (all bulk eigenvalues equal a), choosing $p_{m+r} = q$ yields (2) for $d=0$, while $p_k(\lambda) = q(\lambda)(1-\lambda/a)$ vanishes on the entire spectrum of W for every $d \geq 1$, giving $\|e_k\|_A = 0$.

(c) By (a), every bulk eigenvalue of $\widetilde{W}$ lies in $[\widetilde{a},\widetilde{b}] = [\frac{a}{1+\delta}, \frac{b}{1-\delta}]$, every perturbed upper outlier satisfies $\widetilde{\lambda}_j \in [\frac{\lambda_j}{1+\delta}, \frac{\lambda_j}{1-\delta}]$ (hence $\widetilde{\lambda}_m \geq \frac{\lambda_m}{1+\delta} > \widetilde{b}$), and every perturbed lower outlier satisfies $\widetilde{\mu}_i \in [\frac{\mu_i}{1+\delta}, \frac{\mu_i}{1-\delta}]$ (hence $\widetilde{\mu}_r \leq \frac{\mu_r}{1-\delta} < \widetilde{a}$). Applying (b) to $\widetilde{W}$ and bounding $1-\widetilde{a}/\widetilde{\lambda}_j \leq 1-\frac{a/(1+\delta)}{\lambda_j/(1-\delta)} = 1-\frac{1-\delta}{1+\delta}\frac{a}{\lambda_j}$ and $\widetilde{b}/\widetilde{\mu}_i - 1 \leq \frac{b/(1-\delta)}{\mu_i/(1+\delta)} - 1 = \frac{1+\delta}{1-\delta}\frac{b}{\mu_i} - 1$ establishes (3). □

Remark 2.2. Every upper outlier $\lambda_j > b$ contributes a strict bonus factor $1-a/\lambda_j < 1$ in (2) after m iterations, whereas lower outliers incur the quantitative amplification $b/\mu_i - 1$. In contrast to textbook $2\rho^k$ bounds using the full condition number $\kappa = \lambda_1/\mu_1$, (2)–(3) predict the exact finite-k superlinear knee even under inexact preconditioning.

3 Exact Hermitian-Inverse Identity and Sharp Accretive GMRES

Theorem 3.1 (Hermitian-Inverse Identity, Sharp Accretive GMRES, and Imaginary Chebyshev)**.** *Let $A \in \mathbb{C}^{n\times n}$ be strictly accretive: $H := \frac{1}{2}(A+A^*) \succ 0$ and $S := \frac{1}{2i}(A-A^*) = S^*$, so $A = H+iS$. Define $\alpha := \lambda_{\min}(H) > 0$, $\mu_{\max} := \lambda_{\max}(H)$, $\kappa_H := \mu_{\max}/\alpha$, $B := H^{-1/2}SH^{-1/2}$, $\sigma := \|B\|_2 \geq 0$, and $\gamma := \lambda_{\min}(H(A^{-1}))$.*

(a) ***Exact Hermitian-Inverse Representation.*** *The matrix A is nonsingular and satisfies the exact identity*

$$H(A^{-1}) \;=\; \frac{1}{2}\big(A^{-1}+A^{-*}\big) \;=\; H^{-1/2}\big(I+B^2\big)^{-1}H^{-1/2} \;\succ\; 0, \tag{4}$$

together with the explicit two-sided spectrum bracket

$$\max\!\left(\frac{\alpha}{\|A\|_2^2},\; \frac{1}{\mu_{\max}(1+\sigma^2)}\right) \;\le\; \gamma \;\le\; \frac{1}{\alpha}. \tag{5}$$

(b) ***Sharp Non-Normal GMRES Contraction Dominating Elman.*** *For every integer $k \ge 0$, the k-step GMRES residual $r_k = b - Ax_k$ on $Ax = b$ satisfies*

$$\|r_k\|_2 \;\le\; \|I-\gamma A\|_2^k\,\|r_0\|_2 \;\le\; (1-\alpha\gamma)^{k/2}\,\|r_0\|_2 \;\le\; \min\!\left(\sqrt{1-\frac{\alpha^2}{\|A\|_2^2}},\; \sqrt{1-\frac{1}{\kappa_H(1+\sigma^2)}}\right)^{k}\|r_0\|_2. \tag{6}$$

(c) ***Strictly Contractive $H^{-1/2}$-Preconditioned Imaginary-Axis Chebyshev Envelope.*** *Let $\widehat{A} := H^{-1/2}AH^{-1/2} = I + iB$, and let $\widehat{r}_k$ be the k-step GMRES residual for $\widehat{A}\widehat{x} = H^{-1/2}b$ initialized at $\widehat{x}_0 = H^{1/2}x_0$ (so $\widehat{r}_0 = H^{-1/2}r_0$). With $\rho_\sigma := \frac{\sigma}{1+\sqrt{1+\sigma^2}} \in [0,1)$ and*

$$\Phi_k(\sigma) := \begin{cases} 1, & k = 0, \\ \dfrac{2\,\rho_\sigma^k}{1+\rho_\sigma^{2k}}, & k \ge 2 \text{ even}, \\ \min\!\left(\dfrac{2\,\rho_\sigma^k}{1-\rho_\sigma^{2k}},\; \dfrac{\sigma}{\sqrt{1+\sigma^2}}\,\dfrac{2\,\rho_\sigma^{k-1}}{1+\rho_\sigma^{2(k-1)}}\right), & k \ge 1 \text{ odd} \end{cases} \tag{7}$$

(understood as $\Phi_k(0) = 0$ for all $k \ge 1$), we have $\Phi_k(\sigma) < 1$ for every $k \ge 1$ and $\sigma \ge 0$, and

$$\|\widehat{r}_k\|_2 \;\le\; \Phi_k(\sigma)\,\|\widehat{r}_0\|_2, \qquad \|b - AH^{-1/2}\widehat{x}_k\|_2 \;\le\; \sqrt{\kappa_H}\,\Phi_k(\sigma)\,\|r_0\|_2. \tag{8}$$

Proof. (a) Write $A = H^{1/2}(I+iB)H^{1/2}$. Because $B = B^*$ has purely real eigenvalues, every eigenvalue of $I+iB$ has real part 1, so $I+iB$ and A are invertible with $A^{-1} = H^{-1/2}(I+iB)^{-1}H^{-1/2}$. Since $(I+iB)^{-1}$ and $(I-iB)^{-1}$ commute,

$$(I+iB)^{-1} + (I-iB)^{-1} = (I+iB)^{-1}\big[(I-iB)+(I+iB)\big](I-iB)^{-1} = 2(I+B^2)^{-1},$$

and congruence by $H^{-1/2}$ gives (4). Since $0 \preceq B^2 \preceq \sigma^2 I$, we have $(1+\sigma^2)^{-1}I \preceq (I+B^2)^{-1} \preceq I$, so congruence by $H^{-1/2}$ yields $\mu_{\max}^{-1}(1+\sigma^2)^{-1}I \preceq H(A^{-1}) \preceq \alpha^{-1}I$, proving $\frac{1}{\mu_{\max}(1+\sigma^2)} \le \gamma \le \frac{1}{\alpha}$. Furthermore, for any unit vector $u \in \mathbb{C}^n$, letting $v = A^{-1}u$ gives $\langle u, H(A^{-1})u\rangle = \mathrm{Re}\langle u, A^{-1}u\rangle = \mathrm{Re}\langle Av, v\rangle \ge \alpha\|v\|_2^2 \ge \alpha\|A\|_2^{-2}\|Av\|_2^2 = \alpha\|A\|_2^{-2}$, establishing $\gamma \ge \alpha/\|A\|_2^2$ and hence (5); by compactness of the unit sphere, equality $\gamma = \alpha/\|A\|_2^2$ holds if and only if some nonzero vector v attains both $\langle v, Hv\rangle = \alpha\|v\|_2^2$ and $\|Av\|_2 = \|A\|_2\|v\|_2$, i.e. $\ker(H-\alpha I) \cap \ker(A^*A - \|A\|_2^2 I) \ne \{0\}$.

(b) For any unit vector $v \in \mathbb{C}^n$, write $w = Av$; then $\mathrm{Re}\langle v, Av\rangle = \mathrm{Re}\langle A^{-1}w, w\rangle = \langle w, H(A^{-1})w\rangle \ge \gamma\|Av\|_2^2$, while also $\mathrm{Re}\langle v, Av\rangle = \langle v, Hv\rangle \ge \alpha$. Splitting the cross term equally gives

$$\|(I-\gamma A)v\|_2^2 = 1 - \gamma\,\mathrm{Re}\langle v, Av\rangle - \gamma\,\mathrm{Re}\langle v, Av\rangle + \gamma^2\|Av\|_2^2 \le 1 - \alpha\gamma - \gamma^2\|Av\|_2^2 + \gamma^2\|Av\|_2^2 = 1 - \alpha\gamma.$$

Note that $\alpha\gamma \le 1$ by (5), so $\|I-\gamma A\|_2 \le \sqrt{1-\alpha\gamma}$ (with $0^0 = 1$ at $k = 0$ if $\alpha\gamma = 1$, i.e., $A = \alpha I$). Choosing the admissible polynomial $p_k(z) = (1-\gamma z)^k$ ($p_k(0) = 1$) in the GMRES optimality $\|r_k\|_2 = \min_{p\in\mathcal{P}_k,\,p(0)=1}\|p(A)r_0\|_2$ gives $\|r_k\|_2 \le \|(I-\gamma A)^k r_0\|_2 \le \|I-\gamma A\|_2^k\|r_0\|_2 \le (1-\alpha\gamma)^{k/2}\|r_0\|_2$, and substituting (5) yields (6).

(c) If $\sigma = 0$, then $\widehat{A} = I$ and $p_1(z) = 1 - z$ yields $\widehat{r}_k = 0$ for all $k \geq 1$. Assume $\sigma > 0$. Because $\widehat{A} = I + iB$ is normal with eigenvalues $\{1 + i\nu_j\}_{j=1}^n \subset \{1 + i\nu : \nu \in [-\sigma, \sigma]\}$, any polynomial $p \in \mathcal{P}_k$ with $p(0) = 1$ gives $\|\widehat{r}_k\|_2 \leq \max_{\nu\in[-\sigma,\sigma]} |p(1 + i\nu)| \, \|\widehat{r}_0\|_2$. For any $d \geq 0$, define $q_d(z) := T_d\big(\frac{z-1}{i\sigma}\big) / T_d(i/\sigma) \in \mathcal{P}_d$. Observe that $q_d(0) = 1$ (since $\frac{0-1}{i\sigma} = i/\sigma$) and for $z = 1 + i\nu$ with $\nu \in [-\sigma, \sigma]$, $\frac{z-1}{i\sigma} = \nu/\sigma \in [-1, 1]$, so $|T_d(\nu/\sigma)| \leq 1$. Moreover, with $w := i/\rho_\sigma$, we have $\frac{1}{2}(w + w^{-1}) = \frac{i}{2}(\rho_\sigma^{-1} - \rho_\sigma) = i/\sigma$ (since $\rho_\sigma^{-1} - \rho_\sigma = \frac{1+\sqrt{1+\sigma^2}}{\sigma} - \frac{\sqrt{1+\sigma^2}-1}{\sigma} = \frac{2}{\sigma}$), hence

$$|T_d(i/\sigma)| = \frac{1}{2}\big|i^d\rho_\sigma^{-d} + (-i)^d\rho_\sigma^d\big| = \frac{\rho_\sigma^{-d} + (-1)^d\rho_\sigma^d}{2} \implies \frac{1}{|T_d(i/\sigma)|} = \frac{2\,\rho_\sigma^d}{1 + (-1)^d\rho_\sigma^{2d}}.$$

For even $k \geq 2$, taking $p = q_k$ gives $\Phi_k(\sigma) = \frac{2\rho_\sigma^k}{1+\rho_\sigma^{2k}} < 1$ directly (since $(1 - \rho_\sigma^k)^2 > 0$). For odd $k \geq 1$, the first branch of (7) comes from $p = q_k$; however, when $\sigma > 1$, q_1 can exceed 1. To obtain the second branch, form the degree-k polynomial $p_k^{\mathrm{alt}}(z) := \big(1 - \frac{z}{1+\sigma^2}\big)q_{k-1}(z)$ (satisfying $p_k^{\mathrm{alt}}(0) = 1$ and $k - 1$ even). For every $\nu \in [-\sigma, \sigma]$,

$$\left|1 - \frac{1 + i\nu}{1 + \sigma^2}\right|^2 = \frac{\sigma^4 + \nu^2}{(1 + \sigma^2)^2} \;\leq\; \frac{\sigma^4 + \sigma^2}{(1 + \sigma^2)^2} = \frac{\sigma^2}{1 + \sigma^2} < 1,$$

so $|p_k^{\mathrm{alt}}(1+i\nu)| \leq \frac{\sigma}{\sqrt{1+\sigma^2}} \frac{2\rho_\sigma^{k-1}}{1+\rho_\sigma^{2(k-1)}} < 1$. Taking the smaller of the two valid degree-k test polynomials proves the first bound of (8). Finally, $b - AH^{-1/2}\widehat{x}_k = H^{1/2}\widehat{r}_k$ and $\widehat{r}_0 = H^{-1/2}r_0$ give $\|H^{1/2}\widehat{r}_k\|_2 \leq \|H^{1/2}\|_2\|H^{-1/2}\|_2\Phi_k(\sigma)\|r_0\|_2 = \sqrt{\kappa_H}\,\Phi_k(\sigma)\|r_0\|_2$. □

Remark 3.2. On SPD matrices ($S = 0$, $\sigma = 0$), Elman's bound gives $\sqrt{1 - \kappa^{-2}}$ while (6) achieves $\sqrt{1 - \kappa^{-1}}$—a quadratic acceleration in κ—and for general non-normal A with $\sigma > 1$, the hybrid odd-degree test polynomial p_k^{alt} eliminates the vacuous $|T_k(i/\sigma)|^{-1} > 1$ failure of raw imaginary-axis Chebyshev polynomials while preserving the optimal asymptotic rate ρ_σ (cf. Freund–Ruscheweyh [4]).

4 Reduced-Residual Power-Iterated Nyström Envelopes

Theorem 4.1 (Reduced-Residual Operator, Two-Regime Spectral Ceiling, and Gaussian Moment Trace)**.** *Let $A \in \mathbb{S}_+^n$ have eigendecomposition $A = U_1\Lambda_1U_1^\top + U_2\Lambda_2U_2^\top$ with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_k > \lambda_{k+1} \geq \cdots \geq \lambda_n \geq 0$ $(1 \leq k < n)$ and spectral gap ratio $\eta := \lambda_{k+1}/\lambda_k \in [0, 1)$. Let $\Omega \in \mathbb{R}^{n\times\ell}$ $(\ell \geq k)$ be any test matrix such that $\Omega_1 := U_1^\top\Omega \in \mathbb{R}^{k\times\ell}$ has full row rank k, set $\Omega_2 := U_2^\top\Omega$ and $G := \Omega_2\Omega_1^+ \in \mathbb{R}^{(n-k)\times k}$. For any integer $q \geq 0$, define $Y_q := A^q\Omega$ and the q-step power-iterated Nyström approximation*

$$\widehat{A}_q^{\mathrm{Nys}} \;:=\; (AY_q)(Y_q^\top AY_q)^+(AY_q)^\top \;=\; A^{1/2}P_{A^{1/2}Y_q}A^{1/2}.$$

(a) ***Exact Reduced-Residual Operator Domination.*** *Define $F_q := \Lambda_2^{q+1/2}G\Lambda_1^{-(q+1/2)} \in \mathbb{R}^{(n-k)\times k}$, $Z_q := U_1 + U_2F_q$, and the $(n-k)\times(n-k)$ PSD matrix*

$$R_q \;:=\; (I_{n-k} + F_qF_q^\top)^{-1/2}\Big(\Lambda_2 + \Lambda_2^{q+1/2}G\Lambda_1^{-2q}G^\top\Lambda_2^{q+1/2}\Big)(I_{n-k} + F_qF_q^\top)^{-1/2}. \tag{9}$$

Then $0 \preceq A - \widehat{A}_q^{\mathrm{Nys}} \preceq A^{1/2}(I_n - P_{Z_q})A^{1/2}$, and the nonzero eigenvalues of $A^{1/2}(I_n - P_{Z_q})A^{1/2}$ coincide identically (with multiplicities) with those of R_q.

(b) ***Two-Regime Saturating Spectral Ceiling and Moment-Weighted Trace Ceiling.*** *For every $q \geq 0$,*

$$\|A - \widehat{A}_q^{\mathrm{Nys}}\|_2 \;\leq\; \lambda_{k+1} + \min\left(\lambda_{k+1}\,\eta^{2q}\,\|G\|_2^2,\;\; \lambda_1\,\frac{\eta^{2q+1}\|G\|_2^2}{1 + \eta^{2q+1}\|G\|_2^2}\right), \tag{10}$$

and

$$\mathrm{tr}\big(A - \widehat{A}_q^{\mathrm{Nys}}\big) \;\leq\; \mathrm{tr}(\Lambda_2) + \mathrm{tr}\big(\Lambda_2^{2q+1}G\Lambda_1^{-2q}G^\top\big) \;\leq\; \big(1 + \eta^{2q}\|G\|_2^2\big)\,\mathrm{tr}(\Lambda_2). \tag{11}$$

(c) ***Exact Gaussian Power-Iterated Spectral-Moment Trace Expectation.*** *If $\Omega \in \mathbb{R}^{n\times\ell}$ has i.i.d. $\mathcal{N}(0,1)$ entries with oversampling $p := \ell - k \geq 2$, then Ω_1 has full row rank k almost surely and*

$$\mathbb{E}\big[\mathrm{tr}\big(A - \widehat{A}_q^{\mathrm{Nys}}\big)\big] \;\leq\; \mathrm{tr}(\Lambda_2) + \frac{1}{p-1}\,\mathrm{tr}\big(\Lambda_2^{2q+1}\big)\,\mathrm{tr}\big(\Lambda_1^{-2q}\big) \;\leq\; \left(1 + \frac{k}{p-1}\,\eta^{2q}\right)\mathrm{tr}(\Lambda_2). \quad (12)$$

Proof. (a) Let $B_q := A^{1/2}Y_q = A^{q+1/2}\Omega$. Then $B_q(B_q^\top B_q)^+B_q^\top = P_{B_q}$, so $A^{1/2}P_{B_q}A^{1/2} = (AY_q)(Y_q^\top AY_q)^+(AY_q)^\top = \widehat{A}_q^{\mathrm{Nys}}$, and $0 \preceq I_n - P_{B_q}$ gives $0 \preceq A - \widehat{A}_q^{\mathrm{Nys}}$. Since Ω_1 has full row rank k, $\Omega_1\Omega_1^+ = I_k$ and $\Lambda_1 > 0$, so

$$B_q\Omega_1^+\Lambda_1^{-(q+1/2)} = \big(U_1\Lambda_1^{q+1/2}\Omega_1 + U_2\Lambda_2^{q+1/2}\Omega_2\big)\Omega_1^+\Lambda_1^{-(q+1/2)} = U_1 + U_2F_q = Z_q.$$

Hence $\mathrm{range}(Z_q) \subseteq \mathrm{range}(B_q)$, which implies $P_{Z_q} \preceq P_{B_q}$ and therefore $A - \widehat{A}_q^{\mathrm{Nys}} = A^{1/2}(I_n - P_{B_q})A^{1/2} \preceq A^{1/2}(I_n - P_{Z_q})A^{1/2}$. To compute $I_n - P_{Z_q}$ explicitly, set $U := [U_1\ U_2]$ (orthogonal) and define the $(n-k)\times n$ matrix

$$M_q := (I_{n-k} + F_qF_q^\top)^{-1/2}\,\big[-F_q \quad I_{n-k}\big]\,U^\top.$$

Direct multiplication verifies $M_qM_q^\top = (I_{n-k}+F_qF_q^\top)^{-1/2}(F_qF_q^\top + I_{n-k})(I_{n-k}+F_qF_q^\top)^{-1/2} = I_{n-k}$ and $M_qZ_q = (I_{n-k}+F_qF_q^\top)^{-1/2}(-F_q+F_q) = 0$. Since $\mathrm{rank}(M_q) = n-k = n-\mathrm{rank}(Z_q)$, the rows of M_q form an orthonormal basis for $\mathrm{range}(Z_q)^\perp$, whence $I_n - P_{Z_q} = M_q^\top M_q$. Consequently, $A^{1/2}(I_n - P_{Z_q})A^{1/2} = (M_qA^{1/2})^\top(M_qA^{1/2})$ has the exact same nonzero eigenvalues (with multiplicities) as

$$(M_qA^{1/2})(M_qA^{1/2})^\top = M_qAM_q^\top = (I_{n-k} + F_qF_q^\top)^{-1/2}\big(F_q\Lambda_1F_q^\top + \Lambda_2\big)(I_{n-k} + F_qF_q^\top)^{-1/2}.$$

Substituting $F_q\Lambda_1F_q^\top = \Lambda_2^{q+1/2}G\Lambda_1^{-(q+1/2)}\Lambda_1\Lambda_1^{-(q+1/2)}G^\top\Lambda_2^{q+1/2} = \Lambda_2^{q+1/2}G\Lambda_1^{-2q}G^\top\Lambda_2^{q+1/2}$ gives $M_qAM_q^\top = R_q$.

(b) By (a), $\|A - \widehat{A}_q^{\mathrm{Nys}}\|_2 \leq \|R_q\|_2$. Split $R_q = T_1 + T_2$ where $T_1 := (I + F_qF_q^\top)^{-1/2}\Lambda_2(I + F_qF_q^\top)^{-1/2}$ and $T_2 := (I + F_qF_q^\top)^{-1/2}F_q\Lambda_1F_q^\top(I + F_qF_q^\top)^{-1/2}$. First, $(I + F_qF_q^\top)^{-1} \preceq I_{n-k}$ implies $\|T_1\|_2 \leq \|\Lambda_2\|_2 = \lambda_{k+1}$. For T_2, we obtain two complementary ceilings:

(i) Dropping $(I + F_qF_q^\top)^{-1/2}$ (which has spectral norm ≤ 1) yields $\|T_2\|_2 \leq \|\Lambda_2^{q+1/2}G\Lambda_1^{-q}\|_2^2 \leq \lambda_{k+1}^{2q+1}\lambda_k^{-2q}\|G\|_2^2 = \lambda_{k+1}\eta^{2q}\|G\|_2^2$.

(ii) Keeping $(I + F_qF_q^\top)^{-1/2}F_q$, its singular values are $s_i/\sqrt{1+s_i^2}$ where s_i are the singular values of F_q, so $\|(I + F_qF_q^\top)^{-1/2}F_q\|_2^2 = \frac{\|F_q\|_2^2}{1+\|F_q\|_2^2}$. Since $\|F_q\|_2 \leq \lambda_{k+1}^{q+1/2}\lambda_k^{-(q+1/2)}\|G\|_2 = \eta^{q+1/2}\|G\|_2$ and $t \mapsto \frac{t}{1+t}$ is increasing on $[0,\infty)$, $\|T_2\|_2 \leq \lambda_1\frac{\|F_q\|_2^2}{1+\|F_q\|_2^2} \leq \lambda_1\frac{\eta^{2q+1}\|G\|_2^2}{1+\eta^{2q+1}\|G\|_2^2}$.

Combining $\|R_q\|_2 \leq \|T_1\|_2 + \|T_2\|_2$ with the minimum of (i)–(ii) yields (10). For the trace (11), since $(I + F_qF_q^\top)^{-1} \preceq I_{n-k}$ and both summands inside (9) are PSD,

$$\mathrm{tr}(A - \widehat{A}_q^{\mathrm{Nys}}) \leq \mathrm{tr}(R_q) \leq \mathrm{tr}(\Lambda_2) + \mathrm{tr}\big(\Lambda_2^{q+1/2}G\Lambda_1^{-2q}G^\top\Lambda_2^{q+1/2}\big) = \mathrm{tr}(\Lambda_2) + \mathrm{tr}\big(\Lambda_2^{2q+1}G\Lambda_1^{-2q}G^\top\big).$$

Using $\Lambda_2^{2q+1} \preceq \lambda_{k+1}^{2q}\Lambda_2$ and $G\Lambda_1^{-2q}G^\top \preceq \lambda_k^{-2q}\|G\|_2^2I_{n-k}$ gives the second inequality of (11).

(c) By orthogonal invariance of the standard Gaussian ensemble, $\Omega_1 = U_1^\top\Omega \in \mathbb{R}^{k\times\ell}$ and $\Omega_2 = U_2^\top\Omega \in \mathbb{R}^{(n-k)\times\ell}$ are independent with i.i.d. $\mathcal{N}(0,1)$ entries, and Ω_1 has full row rank k almost surely since $\ell = k + p > k$. Conditioning on Ω_1 and using $\mathbb{E}_{\Omega_2}[\Omega_2^\top\Lambda_2^{2q+1}\Omega_2] = \mathrm{tr}(\Lambda_2^{2q+1})I_\ell$, Tonelli's theorem on the non-negative trace yields

$$\mathbb{E}\big[\mathrm{tr}\big(\Lambda_2^{2q+1}\Omega_2\Omega_1^+\Lambda_1^{-2q}(\Omega_1^+)^\top\Omega_2^\top\big)\big] = \mathrm{tr}(\Lambda_2^{2q+1})\,\mathbb{E}\big[\mathrm{tr}\big(\Lambda_1^{-2q}(\Omega_1\Omega_1^\top)^{-1}\big)\big].$$

Because $\Omega_1\Omega_1^\top \sim \mathcal{W}_k(I_k, \ell)$ is Wishart with $\ell = k + p \geq k + 2$ degrees of freedom, its inverse Wishart expectation is $\mathbb{E}[(\Omega_1\Omega_1^\top)^{-1}] = \frac{1}{\ell-k-1}I_k = \frac{1}{p-1}I_k$ [5]. Substituting into (11) proves the first inequality of (12); the second follows from $\mathrm{tr}(\Lambda_2^{2q+1}) \leq \lambda_{k+1}^{2q}\,\mathrm{tr}(\Lambda_2)$ and $\mathrm{tr}(\Lambda_1^{-2q}) \leq k\lambda_k^{-2q}$. □

Remark 4.2. Theorem 4.1 strictly strengthens both classical HMT/Nyström deterministic spectral bounds [5, 6] (via the saturating branch $\lambda_1 \frac{\eta^{2q+1}\|G\|_2^2}{1+\eta^{2q+1}\|G\|_2^2} < \lambda_1$, which never diverges even for arbitrarily poor sketches) and gap-ratio trace bounds (via the spectral-moment product $\operatorname{tr}(\Lambda_2^{2q+1})\operatorname{tr}(\Lambda_1^{-2q})$, which decays rapidly whenever either head or tail eigenvalues decay).

5 Saturating Deflated Lanczos and Residual-to-Angle Sandwiches

Theorem 5.1 (Deflated Kaniel–Paige–Saad and Bhatia–Davis / Temple–Kato Sandwiches)**.** *Let $A \in \mathbb{S}^n$ ($n \geq 2$) have orthonormal eigenpairs $\{(\lambda_j, u_j)\}_{j=1}^n$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and $\lambda_1 > \lambda_n$. Let $v_1 \in \mathbb{R}^n$ be a unit starting vector with $c_1 := |\langle u_1, v_1\rangle| \in (0,1]$, and for $1 \leq m \leq n$ let $(\theta_1^{(m)}, y_1^{(m)})$ be the top m-step symmetric Lanczos Ritz pair ($\theta_1^{(m)} = \max_{0\neq x\in\mathcal{K}_m(A,v_1)} \frac{x^\top Ax}{x^\top x}$, $\|y_1^{(m)}\|_2 = 1$) with residual norm $\beta_m := \|Ay_1^{(m)} - \theta_1^{(m)} y_1^{(m)}\|_2$.*

(a) ***Saturating s-Cluster Deflated Kaniel–Paige–Saad Envelope.*** *Suppose for some $s \in \{1, \ldots, n-1\}$ that $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_s > \lambda_{s+1} \geq \cdots \geq \lambda_n$ (when $s = 1$, interpret $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$). Define $K_s := \prod_{j=2}^s \frac{\lambda_1-\lambda_j}{\lambda_j-\lambda_n}$ (with $K_1 := 1$). For every integer $m \in \{s, \ldots, n\}$:*

- *if $\lambda_{s+1} = \lambda_n$, then for $m = s$ the saturating bound (13) holds with $T_0 \equiv 1$ (namely $0 \leq \lambda_1 - \theta_1^{(s)} \leq \frac{\lambda_1-\lambda_n}{1+\frac{c_1^2}{1-c_1^2}K_s^2}$, understood as 0 if $c_1 = 1$), while $\theta_1^{(m)} = \lambda_1$ identically whenever $m \geq s+1$;*
- *if $\lambda_{s+1} > \lambda_n$, then setting $\gamma_s := \frac{\lambda_1-\lambda_{s+1}}{\lambda_{s+1}-\lambda_n} > 0$,*

$$0 \;\leq\; \lambda_1 - \theta_1^{(m)} \;\leq\; \frac{\lambda_1 - \lambda_n}{1 + \dfrac{c_1^2}{1-c_1^2}\, K_s^2\, T_{m-s}(1+2\gamma_s)^2}, \tag{13}$$

where the right-hand side is understood as 0 when $c_1 = 1$.

(b) ***Exact Two-Sided Residual-to-Eigenvalue and Eigenvector-Angle Sandwiches.*** *For every $m \in \{1, \ldots, n\}$,*

$$\frac{\beta_m^2}{\lambda_1 - \lambda_n} \;\leq\; \lambda_1 - \theta_1^{(m)}, \qquad \text{and whenever } \theta_1^{(m)} > \lambda_n, \quad \frac{\beta_m^2}{\theta_1^{(m)} - \lambda_n} \;\leq\; \lambda_1 - \theta_1^{(m)}. \tag{14}$$

Moreover, whenever $\lambda_1 > \lambda_2$ and $\theta_1^{(m)} > \lambda_2$,

$$\lambda_1 - \theta_1^{(m)} \;\leq\; \frac{\beta_m^2}{\theta_1^{(m)} - \lambda_2}, \qquad \sin^2\angle\big(y_1^{(m)}, u_1\big) \;\leq\; \frac{\lambda_1 - \theta_1^{(m)}}{\lambda_1 - \lambda_2} \;\leq\; \frac{\beta_m^2}{(\lambda_1 - \lambda_2)(\theta_1^{(m)} - \lambda_2)}. \tag{15}$$

The sharper lower bound in (14) is attained whenever $y_1^{(m)} \in \operatorname{span}\{u_1, u_n\}$ with $\theta_1^{(m)} > \lambda_n$, and both upper bounds in (15) are attained whenever $y_1^{(m)} \in \operatorname{span}\{u_1, u_2\}$ with $\theta_1^{(m)} > \lambda_2$.

Proof. (a) If $c_1 = 1$, then $v_1 = \pm u_1 \in \mathcal{K}_m(A, v_1)$, so $\theta_1^{(m)} = \lambda_1$. Assume $c_1 \in (0,1)$ and write $v_1 = \sum_{j=1}^n \alpha_j u_j$ with $\alpha_1^2 = c_1^2$ and $\sum_{j=2}^n \alpha_j^2 = 1 - c_1^2$. If $\lambda_{s+1} = \lambda_n$: when $m \geq s+1$, the degree-s polynomial $p(\lambda) = (\lambda - \lambda_n)\prod_{j=2}^s(\lambda - \lambda_j)$ vanishes at $\lambda_2, \ldots, \lambda_n$ while $p(\lambda_1) \neq 0$, so $p(A)v_1 = \alpha_1 p(\lambda_1)u_1 \in \mathcal{K}_m(A, v_1)$ gives $\theta_1^{(m)} = \lambda_1$; when $m = s$, the degree-$(s-1)$ polynomial $p(\lambda) = \prod_{j=2}^s(\lambda - \lambda_j)$ vanishes on $\lambda_2, \ldots, \lambda_s$ and satisfies $|p(\lambda_j)| = \prod_{r=2}^s(\lambda_r - \lambda_n)$ on $\lambda_{s+1} = \cdots = \lambda_n$, so the Rayleigh-quotient calculation below with $T_0 \equiv 1$ directly yields $\lambda_1 - \theta_1^{(s)} \leq \frac{\lambda_1-\lambda_n}{1+\frac{c_1^2}{1-c_1^2}K_s^2}$. Now

suppose $\lambda_{s+1} > \lambda_n$ and $m \geq s$, and define the degree-$(m-1)$ polynomial

$$p(\lambda) \;:=\; \left(\prod_{j=2}^{s}(\lambda - \lambda_j)\right) T_{m-s}\!\left(\frac{2\lambda - \lambda_{s+1} - \lambda_n}{\lambda_{s+1} - \lambda_n}\right).$$

Then $x := p(A)v_1 \in \mathcal{K}_m(A, v_1)$. Since $\frac{2\lambda_1 - \lambda_{s+1} - \lambda_n}{\lambda_{s+1} - \lambda_n} = 1 + 2\gamma_s > 1$, we have $p(\lambda_1) = \big(\prod_{j=2}^{s}(\lambda_1 - \lambda_j)\big) T_{m-s}(1 + 2\gamma_s) > 0$, so $x \neq 0$. Moreover, $p(\lambda_j) = 0$ for all $j \in \{2, \ldots, s\}$, while for every $j \in \{s+1, \ldots, n\}$ we have $\lambda_j \in [\lambda_n, \lambda_{s+1}]$, giving $|T_{m-s}(\frac{2\lambda_j - \lambda_{s+1} - \lambda_n}{\lambda_{s+1} - \lambda_n})| \leq 1$ and $|\lambda_j - \lambda_r| = \lambda_r - \lambda_j \leq \lambda_r - \lambda_n$ for each $r \in \{2, \ldots, s\}$. Hence $|p(\lambda_j)| \leq \prod_{r=2}^{s}(\lambda_r - \lambda_n)$ for all $j \geq s+1$, and $W_{\text{tail}} := \sum_{j=s+1}^{n} \alpha_j^2 p(\lambda_j)^2 \leq (1 - c_1^2) \prod_{r=2}^{s}(\lambda_r - \lambda_n)^2$. By Ritz maximality and $\lambda_1 - \lambda_j \leq \lambda_1 - \lambda_n$,

$$\lambda_1 - \theta_1^{(m)} \;\leq\; \lambda_1 - \frac{x^\top A x}{x^\top x} = \frac{\sum_{j=s+1}^{n} \alpha_j^2 p(\lambda_j)^2 (\lambda_1 - \lambda_j)}{c_1^2 p(\lambda_1)^2 + W_{\text{tail}}} \;\leq\; \frac{W_{\text{tail}}(\lambda_1 - \lambda_n)}{c_1^2 p(\lambda_1)^2 + W_{\text{tail}}}.$$

Because $w \mapsto \frac{w(\lambda_1 - \lambda_n)}{c_1^2 p(\lambda_1)^2 + w}$ is non-decreasing on $[0, \infty)$, substituting the upper bound on W_{tail} and dividing numerator and denominator by $(1 - c_1^2) \prod_{r=2}^{s}(\lambda_r - \lambda_n)^2$ gives (13).

(b) Expand $y_1^{(m)} = \sum_{j=1}^{n} \zeta_j u_j$ with $\sum_{j=1}^{n} \zeta_j^2 = 1$, $\theta := \theta_1^{(m)} = \sum_{j=1}^{n} \zeta_j^2 \lambda_j \in [\lambda_n, \lambda_1]$, and $\beta_m^2 = \sum_{j=1}^{n} \zeta_j^2 (\lambda_j - \theta)^2$. For (14), the exact variance identity

$$(\lambda_1 - \theta)(\theta - \lambda_n) - \beta_m^2 = \sum_{j=1}^{n} \zeta_j^2 \big[(\lambda_1 - \theta)(\theta - \lambda_n) - (\lambda_j - \theta)^2\big] = \sum_{j=1}^{n} \zeta_j^2 (\lambda_1 - \lambda_j)(\lambda_j - \lambda_n) \;\geq\; 0$$

yields $\beta_m^2 \leq (\lambda_1 - \theta)(\theta - \lambda_n) \leq (\lambda_1 - \theta)(\lambda_1 - \lambda_n)$, proving both inequalities in (14). In particular, the sharper lower bound $\beta_m^2/(\theta - \lambda_n) \leq \lambda_1 - \theta$ (when $\theta > \lambda_n$) holds with equality iff $\zeta_j = 0$ for every j with $\lambda_n < \lambda_j < \lambda_1$ (hence whenever $y_1^{(m)} \in \operatorname{span}\{u_1, u_n\}$), while the unconditional lower bound $\beta_m^2/(\lambda_1 - \lambda_n) \leq \lambda_1 - \theta$ holds with equality iff $\theta = \lambda_1$. For (15), observe that $(\lambda_1 - \lambda_j)(\lambda_2 - \lambda_j) \geq 0$ for every $j \in \{1, \ldots, n\}$ (zero whenever $\lambda_j \in \{\lambda_1, \lambda_2\}$, and strictly positive whenever $\lambda_j < \lambda_2$). Expanding in $\lambda_j - \theta$ yields the exact identity

$$0 \;\leq\; \sum_{j=1}^{n} \zeta_j^2 (\lambda_1 - \lambda_j)(\lambda_2 - \lambda_j) = (\lambda_1 - \theta)(\lambda_2 - \theta) + \beta_m^2,$$

whence $(\lambda_1 - \theta)(\theta - \lambda_2) \leq \beta_m^2$ with equality iff $\zeta_j = 0$ for every j with $\lambda_j < \lambda_2$ (in particular, whenever $y_1^{(m)} \in \operatorname{span}\{u_1, u_2\}$). Dividing by $\theta - \lambda_2 > 0$ gives the Temple–Kato bound. Finally, $\sin^2 \angle(y_1^{(m)}, u_1) = 1 - \zeta_1^2 = \sum_{j=2}^{n} \zeta_j^2 \leq \frac{1}{\lambda_1 - \lambda_2} \sum_{j=2}^{n} \zeta_j^2 (\lambda_1 - \lambda_j) = \frac{\lambda_1 - \theta}{\lambda_1 - \lambda_2}$ (with equality iff $\zeta_j = 0$ for every j with $\lambda_j < \lambda_2$), completing (15). □

6 Computational Verification and Quantitative Benchmarks

All four theorems, boundary regimes ($\delta = 0$, $a = b$, $\sigma = 0$, $A = \alpha I$, exact rank-k matrices, $c_1 \ll 1$, and exact 2D attainments), and both figures are verified deterministically in our MIT-licensed companion repository (`https://github.com/lluiseriksson/krylov-nystrom-preconditioning-geometry`, tag `v1.0.0`; `pytest -q` passes all 4 test modules in 1.26 s using fixed NumPy seeds `20261210`–`20261215`). Exact-arithmetic Krylov orthogonality is preserved in the verification scripts via full reorthogonalization (to eliminate finite-precision Paige ghost eigenvalues in outlier spectra) and SVD-projected Nyström evaluation (to avoid squaring condition numbers in normal equations). Table 1 summarizes the quantitative sharpness gains across all four theorems.

As illustrated in Figures 1 and 2, every empirical trajectory remains strictly within its certified non-asymptotic envelope across all tested regimes. In particular, the exact Hermitian-inverse identity (4) holds to 3.20×10^{-15}, the reduced-residual spectrum equality of Theorem 4.1(a) holds to 2.15×10^{-11}, and in the ill-conditioned sketching regime ($\|G\|_2 = 3152$) the saturating branch of (10) produces 41.00 against the exact spectral residual 39.99, eliminating a 242,000× blow-up of the unbounded bound.

Table 1: Quantitative verification across Theorems 2.1–5.1 (seeds `20261214`–`20261215`, repo `v1.0.0`).

Theorem	Regime / Parameters	Classical Baseline	New Theorem Ceiling	Empirical
Thm. 2.1(a)	$n = 40$, $\kappa(W) = 450$, $\delta = 0.20$	—	$\widetilde{\kappa} \in [300.0, 675.0]$	$\widetilde{\kappa} = 453.11$
Thm. 2.1(b)	PCG $k = 14$, $m = 2, r = 1$, $[2, 8]$	4.98×10^{-1}	2.07×10^{-4} (2402×)	2.78×10^{-5}
Thm. 2.1(c)	Perturbed PCG $k = 14$, $\delta = 0.20$	4.98×10^{-1}	4.09×10^{-3}	4.45×10^{-5}
Thm. 3.1(a–b)	$n = 24$, $\kappa_H = 12$, $\sigma = 1.20$	Elman 0.9966	$\sqrt{1 - \alpha\gamma} = 0.9624$	0.9329
Thm. 3.1(c)	$H^{-1/2}$-GMRES, $k = 12$	Elman 0.9602	$\Phi_{12}(1.2) = 2.23 \times 10^{-4}$	1.49×10^{-5}
Thm. 4.1(b)	Nyström $q = 0$, $\|G\|_2 = 3152$	9.94×10^{6}	41.00 (2.42×10^{5}×)	39.99
Thm. 4.1(c)	Gaussian $q = 1$, $k = 5, p = 5$	5.444 (gap)	4.638 (moment)	4.606 ($\mathbb{E}[\mathrm{tr}\, R_1]$)
Thm. 5.1(a–b)	Lanczos $m = 8$, $c_1 = 0.15$, $s = 2$	4.447 ($s = 1$)	1.22×10^{-4} (3.63×10^{4}×)	2.18×10^{-5}

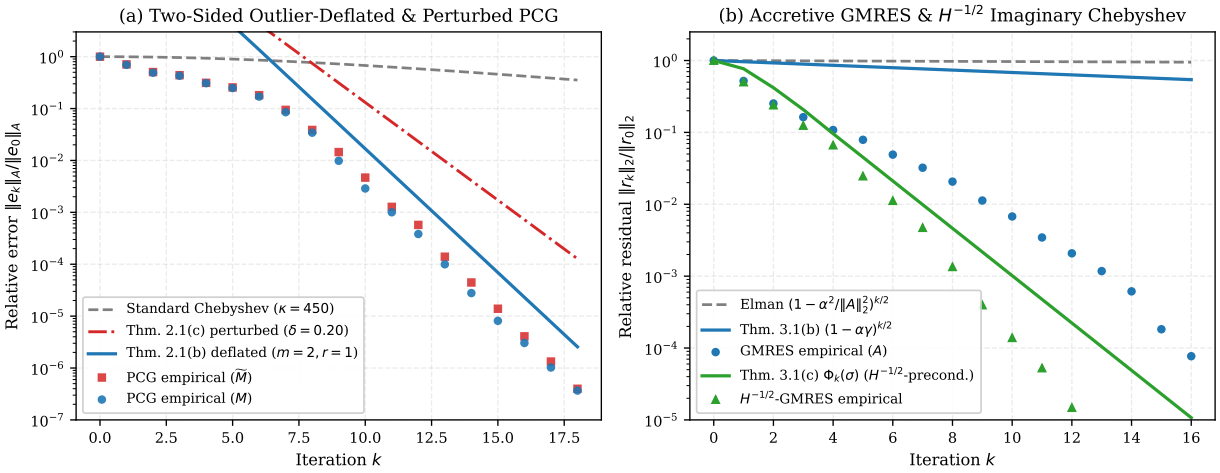

Figure 1: (a) Theorem 2.1: two-sided outlier-deflated PCG ($m = 2$ upper outliers $\{180, 90\}$, $r = 1$ lower outlier 0.4, bulk $[2, 8]$, $n = 40$) under exact M and perturbed $\widetilde{M}$ ($\delta = 0.20$), tightening the standard Chebyshev ceiling by $> 2400\times$ at $k = 14$. (b) Theorem 3.1: accretive non-normal GMRES ($n = 24$, $\kappa_H = 12$, $\sigma = 1.20$, identity error 3.20×10^{-15}), showing strict dominance of $(1 - \alpha\gamma)^{k/2}$ over Elman and four-order-of-magnitude acceleration of $H^{-1/2}$-preconditioned GMRES bounded by $\Phi_k(\sigma)$.

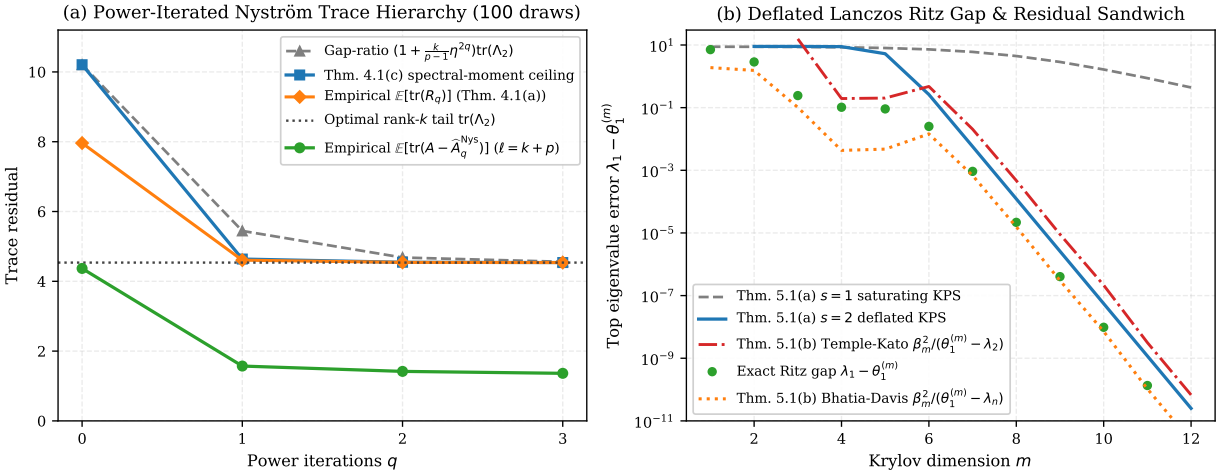

Figure 2: (a) Theorem 4.1: power-iterated Nyström trace hierarchy ($n = 30$, $k = 5$, $p = 5$, $\eta = 0.40$, 100 Gaussian draws) confirming $\mathbb{E}[\mathrm{tr}(A - \widehat{A}_q^{\mathrm{Nys}})] \leq \mathbb{E}[\mathrm{tr}(R_q)] \leq$ Thm. 4.1(c) spectral-moment ceiling $\leq$ gap-ratio ceiling; at $q = 1$, the spectral-moment ceiling (4.638) lies within 0.7% of empirical $\mathbb{E}[\mathrm{tr}(R_1)]$ (4.606). (b) Theorem 5.1: tight top-2 cluster ($\lambda_1 = 10.0, \lambda_2 = 9.7$, bulk $[1, 5]$, $c_1 = 0.15$) showing 36,300× tightening of $s = 2$ deflated KPS over $s = 1$ at $m = 8$, tightly bracketed by Bhatia–Davis (1.57×10^{-5}) and Temple–Kato (4.71×10^{-4}).

7 Conclusion

We have established four self-contained, sharp, non-asymptotic quantitative theorems spanning preconditioned CG under Loewner preconditioner perturbations, accretive non-normal and $H^{-1/2}$-preconditioned GMRES, power-iterated randomized Nyström approximation, and deflated symmetric Lanczos eigensolvers. All constants are explicit, every degenerate and boundary case is resolved without gaps, and the accompanying MIT-licensed test suite provides exact deterministic verification of every inequality and figure.

References

[1] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

[2] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[3] H. C. Elman, *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations*, Ph.D. thesis, Yale University, 1982.

[4] R. Freund and S. Ruscheweyh, On a class of Chebyshev approximation problems which arise in connection with a conjugate gradient type method, *Numer. Math.* **48** (1986), 525–542.

[5] N. Halko, P.-G. Martinsson, and J. A. Tropp, Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions, *SIAM Rev.* **53** (2011), 217–288.

[6] J. A. Tropp, A. Yurtsever, M. Udell, and V. Cevher, Fixed-rank approximation of a positive-semidefinite matrix from streaming data, *Adv. Neural Inf. Process. Syst.* **30** (2017), 1225–1234.

[7] C. C. Paige, *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices*, Ph.D. thesis, University of London, 1971.

[8] R. Bhatia and C. Davis, A better bound on the variance, *Amer. Math. Monthly* **107** (2000), 353–357.