Low-rank approximation of non-Hermitian operators from spectral entangling strength

Yuxuan Zhang

18 November 2026 Version 1.0

Abstract

We prove a dimension-independent operator-norm approximation theorem for arbitrary bipartite operators. Let $J_\alpha(X)$ be the largest $\ell_{2\alpha}$ norm of the output Schmidt coefficients when X acts on normalized product inputs with arbitrary local ancillas. For $0 < \alpha < 1/3$ and every integer $D \geq 0$, there exists an operator Y of operator Schmidt rank at most D such that

$$\|X - Y\| \leq J_\alpha(X)(D+1)^{-(1/(2\alpha)-3/2)}.$$

No Hermiticity or locality is assumed. The proof optimizes over purified local input states, uses their first variations to control two column square sums, and truncates the resulting tensor decomposition. The power of D is optimal, as follows from the rank-one obstruction underlying the known failure at $\alpha \geq 1/3$. This establishes the positive low-order range of the non-Hermitian approximation question discussed by Kim et al. It is an existence theorem; efficient computation of the optimizing inputs is not established.

1 The approximation question

An operator may generate little entanglement on product inputs while remaining difficult to approximate uniformly on all input states. Operator Schmidt truncation naturally controls the Hilbert–Schmidt norm. Converting a normalized Hilbert–Schmidt error into an operator-norm error can cost a factor depending on dimension. To obtain a dimension-independent guarantee, one needs information that controls more than the Schmidt spectrum at the maximally entangled local inputs.

Kim et al. [1] introduced an ancilla-assisted spectral entangling strength and asked whether its low-order versions control operator Schmidt approximation. Their published Note added reports a positive result for Hermitian operators at $\alpha < 1/2$, and a non-Hermitian rank-one obstruction at $\alpha \geq 1/3$. Their discussion leaves analogous guarantees for general non-Hermitian operators as a further question. We give such a guarantee throughout $0 < \alpha < 1/3$, with the optimal power-law exponent.

The distinction from neighboring results concerns both the hypothesis and the error norm. Compression theorems for local dynamics [3] impose structure on the generator. Results from operator entanglement scaling [4] control selected correlation functions using a particular vectorization of the operator. Here the hypothesis is only a bound on an ancilla-assisted strength, and the conclusion is worst-case operator-norm approximation across one bipartition.

Prior-work status. The published Note added of Ref. [1] attributes its Hermitian result and non-Hermitian obstruction to Kim, Kusunoki, and Kuwahara, listed there as unpublished reference 93 [2]. A complete manuscript of that follow-up was not available in the public sources checked on

18 November 2026. Its possible overlap with the positive-range theorem below cannot therefore be excluded. No claim of first-publication priority is made. The negative-range obstruction is treated as prior work.

2 Statement and input representation

Let A, B be finite-dimensional complex Hilbert spaces and let $X \in \mathcal{L}(A \otimes B)$. The operator Schmidt rank $\mathrm{OSR}(X)$ is the least number of terms in a decomposition $X = \sum_s A_s \otimes B_s$. Throughout, $\|\cdot\|$ without a subscript denotes the operator norm. For $p > 0$, define

$$j_p(X) = \sup_{a,b} \left(\sum_s \lambda_s\big((X \otimes I)|a\rangle_{AA'}|b\rangle_{BB'}\big)^p \right)^{1/p}. \tag{1}$$

The inputs are unit vectors, the local ancillas have arbitrary finite dimensions, and the λ_s are the unnormalized Schmidt coefficients across $AA' : BB'$. Thus $j_p = J_{p/2}$; J_α is denoted $\bar{J}_\alpha$ in Definition 2 of Ref. [1]. For $p < 1$, the expression in parentheses defines a sequence quasinorm, not a norm.

Theorem 1. *For $0 < p < 2/3$, every $X \in \mathcal{L}(A \otimes B)$ and every integer $D \geq 0$ admit an operator Y with $\mathrm{OSR}(Y) \leq D$ and*

$$\|X - Y\| \leq j_p(X)(D+1)^{-\kappa_p}, \qquad \kappa_p = \frac{1}{p} - \frac{3}{2}. \tag{2}$$

The bound is independent of $\dim A$ and $\dim B$. At $p = 2/3$, the case $D = 0$ remains valid: $\|X\| \leq j_{2/3}(X)$.

For $D = 0$ the approximant is zero. If $j_p(X) > \epsilon > 0$, choosing

$$D = \left\lceil \left(\frac{j_p(X)}{\epsilon} \right)^{1/\kappa_p} \right\rceil - 1 \tag{3}$$

suffices for error at most ϵ; one may instead take $Y = X$ when its exact rank is smaller. If $\epsilon \geq j_p(X)$, take $Y = 0$. In terms of $\alpha = p/2$, the accuracy exponent is $2\alpha/(1-3\alpha)$. The approximant is an arbitrary complex operator: positivity, unitarity, or a channel interpretation are not part of the assertion.

Fix bases and write $\mathrm{vec}(M) = \sum_{a,i} M_{ai}|a\rangle|i\rangle$. A canonical purification of a density matrix ρ is $\mathrm{vec}(\sqrt{\rho})$. Every local pure state with reduced density ρ is related to this purification by an isometry on the ancillary support. That isometry commutes with X and preserves Schmidt coefficients. Consequently ancillas of dimensions $d_A = \dim A$ and $d_B = \dim B$ suffice, and

$$j_p(X)^p = \max_{\rho,\sigma} F(\rho,\sigma), \qquad F(\rho,\sigma) = \sum_s \lambda_s(\rho,\sigma)^p, \tag{4}$$

where ρ, σ range over density matrices and the coefficients are those of

$$w(\rho,\sigma) = (X \otimes I)\big[\mathrm{vec}(\sqrt{\rho}) \otimes \mathrm{vec}(\sqrt{\sigma})\big]. \tag{5}$$

Continuity and finite-dimensional compactness give the maximum. We do not assume joint concavity of F.

For $0 < \delta < 1$, maximize F on the restricted compact domain

$$\rho \geq \frac{\delta}{d_A} I_A, \qquad \sigma \geq \frac{\delta}{d_B} I_B, \qquad \mathrm{Tr}\,\rho = \mathrm{Tr}\,\sigma = 1. \tag{6}$$

Choose a maximizer and put $F_\delta = F(\rho, \sigma)$. The case $X = 0$ is immediate, so assume $X \neq 0$. Then $F_\delta > 0$. Take a Schmidt decomposition

$$w = \sum_{s=1}^{r} \lambda_s \operatorname{vec}(L_s) \otimes \operatorname{vec}(R_s), \qquad \lambda_1 \geq \cdots \geq \lambda_r > 0, \tag{7}$$

with both vector families orthonormal. Define

$$A_s = L_s \rho^{-1/2}, \qquad B_s = R_s \sigma^{-1/2}. \tag{8}$$

Undoing the two invertible input filters in (5) gives the exact identity

$$X = \sum_{s=1}^{r} \lambda_s A_s \otimes B_s. \tag{9}$$

In particular, if the coefficient matrix of w has matrix SVD $U\Lambda V^\dagger$, the right Schmidt ket in (7) is the vector $\overline{V_s}$. This convention matters for complex X.

3 First variations control both column sums

Set

$$K_A = \sum_s \lambda_s^p A_s^\dagger A_s, \qquad K_B = \sum_s \lambda_s^p B_s^\dagger B_s. \tag{10}$$

These operators act on the physical input spaces. The following derivative identity is the main use of the input optimization.

Lemma 2. *At strictly positive inputs, for every Hermitian variation Δ of ρ,*

$$d_\rho F[\Delta] = \frac{p}{2} \operatorname{Tr}(K_A \Delta). \tag{11}$$

The analogous identity holds for σ and K_B.

Proof. Write $S = d\sqrt{\rho}[\Delta]$. The derivative of the matrix square root is Hermitian and satisfies

$$\sqrt{\rho}\, S + S\sqrt{\rho} = \Delta. \tag{12}$$

Keep the exact decomposition (9) fixed at the base input while varying ρ. Then

$$dw = \sum_s \lambda_s \operatorname{vec}(A_s S) \otimes \operatorname{vec}(R_s).$$

Differentiating the sum of powers of the nonzero Schmidt coefficients yields

$$\begin{aligned} dF &= p \operatorname{Re} \sum_s \lambda_s^{p-1} \langle \operatorname{vec}(L_s) \otimes \operatorname{vec}(R_s), dw \rangle \\ &= p \operatorname{Re} \sum_s \lambda_s^p \operatorname{Tr}(L_s^\dagger A_s S) \\ &= p \operatorname{Re} \operatorname{Tr}(\sqrt{\rho} K_A S) \\ &= \frac{p}{2} \operatorname{Tr}[K_A (S\sqrt{\rho} + \sqrt{\rho} S)]. \end{aligned} \tag{13}$$

Equation (12) proves the claim.

For completeness, this differentiation is valid even when the realignment of X has deficient rank. On the open domain $\rho, \sigma > 0$, the coefficient matrix of (5) is obtained from the fixed realignment of X by invertible left and right multipliers. Its rank is therefore constant. Its positive singular values stay separated from zero locally, while the remaining singular values stay identically zero. The trace spectral derivative used in (13) is thus taken on a fixed-rank smooth family. It neither differentiates t^p through a newly appearing singular value at zero nor requires simple positive singular values; repeated positive values have the same spectral derivative in any orthonormal basis of the repeated subspace. □

Orthonormality in (7) also gives

$$\operatorname{Tr}(K_A\rho) = \operatorname{Tr}(K_B\sigma) = F_\delta. \tag{14}$$

For any density matrix τ on A, the segment from ρ to $\rho_\tau = \delta I_A/d_A + (1-\delta)\tau$ lies in the feasible domain (6). At the chosen maximum its one-sided derivative is nonpositive. Lemma 2 and (14) give

$$(1-\delta)\operatorname{Tr}(K_A\tau) \le F_\delta - \frac{\delta}{d_A}\operatorname{Tr} K_A \le F_\delta.$$

Taking the supremum over τ, and then repeating the argument on B, yields

$$K_A \le \frac{F_\delta}{1-\delta} I_A, \qquad K_B \le \frac{F_\delta}{1-\delta} I_B. \tag{15}$$

The floor in (6) permits singular unconstrained maximizing inputs: all derivatives were computed at positive densities, using feasible directions. No assertion that the maximizing point is interior is needed.

4 A two-column estimate and truncation

Lemma 3. *If finite operator families satisfy $\sum_s C_s^\dagger C_s \le I_A$ and $\sum_s E_s^\dagger E_s \le I_B$, then for arbitrary scalars c_s,*

$$\left\| \sum_s c_s C_s \otimes E_s \right\| \le \left(\sum_s |c_s|^2 \right)^{1/2}. \tag{16}$$

Proof. For any unit vector $\psi \in A \otimes B$, vector Cauchy–Schwarz gives

$$\left\| \sum_s c_s (C_s \otimes E_s)\psi \right\|^2 \le \left(\sum_s |c_s|^2 \right) \langle\psi| \sum_s C_s^\dagger C_s \otimes E_s^\dagger E_s |\psi\rangle.$$

The positive operator in the last expression is at most

$$\left(\sum_s C_s^\dagger C_s \right) \otimes \left(\sum_t E_t^\dagger E_t \right) \le I_{AB}.$$

Indeed, the difference is a sum of positive tensor products with $s \ne t$. Since ψ can be entangled, the result bounds the full operator norm. □

Let $c_\delta = F_\delta/(1-\delta)$ and set

$$C_s = \lambda_s^{p/2} A_s/\sqrt{c_\delta}, \qquad E_s = \lambda_s^{p/2} B_s/\sqrt{c_\delta}.$$

Equation (15) verifies the two hypotheses of Lemma 3. Applying it to any subset I gives

$$\left\|\sum_{s\in I}\lambda_s A_s\otimes B_s\right\| \le \frac{F_\delta}{1-\delta}\left(\sum_{s\in I}\lambda_s^{2-2p}\right)^{1/2}. \tag{17}$$

This estimate converts the input-dependent Schmidt decomposition into an operator-norm bound without introducing a row square sum or a dimension factor.

Proof of Theorem 1. If $D \ge r$, take $Y = X$. Otherwise put

$$Y_{\delta,D} = \sum_{s\le D}\lambda_s A_s\otimes B_s, \qquad q = 2-2p.$$

Then $\mathrm{OSR}(Y_{\delta,D}) \le D$. For $0 < p < 2/3$ we have $q > p$, and monotonicity of the Schmidt coefficients gives

$$\begin{aligned}\lambda_{D+1}^p &\le \frac{F_\delta}{D+1},\\ \sum_{s>D}\lambda_s^q &\le \lambda_{D+1}^{q-p}\sum_{s>D}\lambda_s^p \le F_\delta^{q/p}(D+1)^{-(q-p)/p}.\end{aligned} \tag{18}$$

Together with (17), this implies

$$\begin{aligned}\|X - Y_{\delta,D}\| &\le \frac{F_\delta^{1/p}}{1-\delta}(D+1)^{-\kappa_p}\\ &\le \frac{j_p(X)}{1-\delta}(D+1)^{-\kappa_p}.\end{aligned} \tag{19}$$

Here $1 + q/(2p) = 1/p$ and $(q-p)/(2p) = 1/p - 3/2$. The calculation includes $D = 0$. At $p = 2/3$ and $D = 0$, use $q = p$ in (17) to obtain the same norm bound with zero exponent.

Fix X, D and let $\delta \downarrow 0$. For $\delta \le 1/2$, the approximating operators in (19) are bounded in operator norm by $\|X\| + 2j_p(X)$. Realignment identifies $\mathrm{OSR}(Y) \le D$ with an ordinary matrix-rank bound, so this set is closed. A subsequence therefore converges in the fixed finite-dimensional space to an operator Y of rank at most D. Continuity of the norm proves (2). The compactness argument is used separately for each finite dimension; only the resulting estimate is uniform in dimension. □

5 Optimality of the power and the threshold

The published Note added of Ref. [1] already reports a non-Hermitian rank-one obstruction at $\alpha \ge 1/3$. We spell out this mechanism to compare the power in Theorem 1; the obstruction itself is not claimed as a new result.

For $A = B = \mathbb{C}^N$, let

$$|\Phi_N\rangle = N^{-1/2}\sum_{i=1}^{N}|i,i\rangle, \qquad T_N = |1,1\rangle\langle\Phi_N|. \tag{20}$$

Proposition 4. *For $0 < p \le 1$, $j_p(T_N) = N^{1/p-3/2}$. For every $0 \le D < N$,*

$$\inf_{\mathrm{OSR}(Y)\le D}\|T_N - Y\| \ge \sqrt{1 - D/N}. \tag{21}$$

Proof. Write the normalized local input coefficient matrices as P, Q, so $\|P\|_F = \|Q\|_F = 1$. The output is $|1,1\rangle$ tensored with an ancillary vector whose coefficient matrix is $N^{-1/2}P^TQ$. This matrix has rank at most N and trace norm at most $N^{-1/2}$, since $\left\|P^TQ\right\|_1 \le \|P\|_F \|Q\|_F$. Its Schatten p quasinorm is at most

$$N^{1/p-1}N^{-1/2} = N^{1/p-3/2}.$$

Both local maximally entangled inputs attain equality. This calculation allows arbitrary ancillary dimensions.

If Y has operator Schmidt rank at most D, the row $\langle 1,1|Y$ is a sum of at most D product bras. Contracting the output is norm contractive, so

$$\|T_N - Y\| \ge \|\langle\Phi_N| - \langle 1,1|Y\| .$$

The smallest vector distance from Φ_N to a Schmidt-rank-D vector is $\sqrt{1-D/N}$, by truncating its N equal Schmidt coefficients. This proves (21). □

For $0 < p < 2/3$, normalize $X_N = N^{-\kappa_p}T_N$. Then $j_p(X_N) = 1$, while choosing $N = 2D$ for any integer $D \ge 1$ gives

$$\inf_{\mathrm{OSR}(Y)\le D} \|X_N - Y\| \ge 2^{-\kappa_p - 1/2} D^{-\kappa_p}. \tag{22}$$

Thus no larger power of D can hold uniformly with a constant depending only on p. The upper constant one in (2) is not asserted to be optimal. At $p = 2/3$, $j_p(T_N) = 1$ and (21) rules out any dimension-independent error function tending to zero. For $p > 2/3$, monotonicity of sequence quasinorms gives $j_p(T_N) \le 1$, so the same obstruction applies.

6 Scope and supplementary checks

Theorem 1 gives a complete dimension-independent positive-range guarantee for arbitrary complex operators at $0 < \alpha < 1/3$. Its proof uses a global input maximum only for existence and first-order inequalities. It does not show that this maximum, or the resulting approximant, can be found efficiently. Compression of dynamics under additional locality assumptions and practical algorithms remain separate questions.

The accompanying package tests complex filter inversion, Schmidt-vector conventions, and both input derivatives in twelve small matrix cases, giving twenty-four finite-difference comparisons. It also includes two local stationarity diagnostics and eight deliberately incorrect substitutions detected by the checks. One control is exact rational arithmetic: a falsely stronger exponent predicts squared error at most $1/81$ in a case where (21) requires at least $1/32$. The optimization diagnostics are not certificates of global maximality. These finite checks support implementation details; the all-dimensional claim rests on the analytic proof above.

This preprint has received sequential internal proof and source review. Independent expert review and formal verification have not been performed. The author directed the research and manuscript preparation; AI assistance contributed to derivation, source comparison, computation, checking, and drafting.

References

[1] D. Kim, Y. Kimura, H. Mackay, Y. Mitsuhashi, H. Nishikawa, C. Rubiliani, C. Shang, A. Ukai, and T. Kuwahara, Spectral Small-Incremental Entangling: Breaking Quasipolynomial Complexity Barriers in Long-Range Interacting Systems, *Phys. Rev. X* **16**, 041035 (2026), doi:10.1103/8fcr-5tf4. See Definition 2 and Conjecture 2, p. 22; discussion item (3) and Note added, p. 46; reference 93, p. 49.

[2] Kim, Kusunoki, and Kuwahara, unpublished work, cited as reference 93 in Ref. [1]. No title or public manuscript identifier is supplied there; the full work was unavailable for comparison.

[3] H. Mackay, D. Kim, T. V. Vu, and T. Kuwahara, Optimal low-rank compression of quantum dynamics, arXiv:2609.27497v1 (2026).

[4] N. Dowling, Classical Simulability from Operator Entanglement Scaling, arXiv:2603.05656v4 (2026).