

Supplemental Material for An Exact Product Limit on Collective Quantum Readout

Codex 5.6 Sol

Pudim Physics Research Ledger

Program Manager: Domingos S. P. Salazar

September 29, 2026

Where does the readout gap enter, and what does each plotted curve represent? Sections S1–S3 derive the one-copy score, the fully-PPT product limit, and the two-copy surplus. Section S4 turns these results into bounds for one prescribed observable. Sections S5 and S6 give the work-stroke and equilibrium calculations; S7 maps the figures to the data and states what finite-shot samples can show. Throughout, the hypotheses factor across the registers used to define PPT, and each comparison between readout classes holds its observable fixed.

S1 The known uncertainty identity and optimal score

S1.1 Definitions and exact single-copy solution

Let $\rho, \sigma > 0$ be density matrices on a finite-dimensional Hilbert space $\mathcal{H}$, and let $\mathcal{M}(\mathcal{H})$ denote all POVMs on that space. A POVM $M = \{E_x\} \in \mathcal{M}(\mathcal{H})$ has $E_x \succeq 0$ and $\sum_x E_x = I$. It yields $p_x = \text{Tr}[\rho E_x]$ and $q_x = \text{Tr}[\sigma E_x]$. Zero effects can be omitted; otherwise full-rankness implies $q_x > 0$. Define

$$Q(M) = \sum_x p_x^2 / q_x, \quad \mathcal{Q}(\rho, \sigma) = \max_{M \in \mathcal{M}(\mathcal{H})} Q(M). \quad (\text{S1})$$

Here $Q(M) \equiv Q(M; \rho, \sigma)$ belongs to one fixed measurement. The unrestricted optimum is $\mathcal{Q}$. For a measurement class $\mathcal{C}$, define $\mathcal{Q}^{\mathcal{C}}(\rho, \sigma) := \max_{M \in \mathcal{C}} Q(M; \rho, \sigma)$; these maxima are attained. With n identical input registers, write $\mathcal{M}_n = \mathcal{M}(\mathcal{H}^{\otimes n})$ and $Q_n(M) = Q(M; \rho^{\otimes n}, \sigma^{\otimes n})$. The class optimum is $\mathcal{Q}_n^{\mathcal{C}} := \max_{M \in \mathcal{C}_n} Q_n(M)$, with $Q_n \equiv Q_n^{\mathcal{M}}$. Thus the argument of Q_n identifies a measurement, while a superscript identifies the class over which we optimize. A restriction changes that class, not the detector functional. The logarithm of $\mathcal{Q}$ is the standard measured order-two Rényi divergence, but we keep Q as the readout variable. For identical inputs, $Q_n = \mathcal{Q}(\rho^{\otimes n}, \sigma^{\otimes n})$ and $Q_0 = 1$. Temme and Verstraete [1], Lemma 5 and Eq. (44) in the arXiv PDF, establish the measured Bures χ^2 equality and the attaining score PVM. General variational expressions and projective sufficiency appear in Ref. [2]; recent smooth-entropy work also uses this quadratic quantity [3]. We give the short derivation to fix our conventions.

Let $\mathcal{J}_\sigma(A) = (\sigma A + A \sigma)/2$ on the Hilbert–Schmidt space. It is self-adjoint and positive definite. Its inverse maps positive matrices to positive matrices because

$$\mathcal{J}_\sigma^{-1}(B) = 2 \int_0^\infty e^{-t\sigma} B e^{-t\sigma} dt. \quad (\text{S2})$$

In particular, $L = \mathcal{J}_\sigma^{-1}(\rho) > 0$ is well defined. For any real outcome scores a_x ,

$$\sum_x \frac{p_x^2}{q_x} = \max_{\{a_x\}} \sum_x (2p_x a_x - q_x a_x^2). \quad (\text{S3})$$

Set $A = \sum_x a_x E_x$. The Schwarz inequality for the unital POVM map gives $A^2 \preceq \sum_x a_x^2 E_x$. The same inequality bounds $Q(M)$ above by the maximum of

$$F(A) = 2 \text{Tr}[\rho A] - \text{Tr}[\sigma A^2], \quad A = A^\dagger. \quad (\text{S4})$$

Since $\mathcal{J}_\sigma(L) = \rho$,

$$F(L) - F(A) = \text{Tr}[\sigma(A - L)^2], \quad \mathcal{Q} = F(L) = \text{Tr}[\rho L] = \text{Tr}[\sigma L^2]. \quad (\text{S5})$$

To see that the upper bound is attainable, use the spectral decomposition $L = \sum_x \ell_x P_x$. The projective measurement $\{P_x\}$ reports the eigenvalue ℓ_x ; each P_x projects onto its eigenspace. If $LP_x = \ell_x P_x$, then $\text{Tr}[\rho P_x] = \text{Tr}[\mathcal{J}_\sigma(L) P_x] = \ell_x \text{Tr}[\sigma P_x]$. The optimal classical score is therefore $p_x/q_x = \ell_x$, and this PVM attains Eq. (S5). Repeated score eigenvalues cause no problem. In an eigenbasis of σ ,

$$L_{ij} = \frac{2\rho_{ij}}{s_i + s_j}, \quad \mathcal{Q} = \sum_{ij} \frac{2|\rho_{ij}|^2}{s_i + s_j}. \quad (\text{S6})$$

No new theorem is assigned to these identities.

S1.2 Sharp Hammersley–Chapman–Robbins form

For a Hermitian observable A , put $A_c = A - \text{Tr}[\sigma A] I$. In the real inner product $\langle B, C \rangle_\sigma = \text{Tr}[B \mathcal{J}_\sigma(C)]$ on Hermitian matrices,

$$\text{Tr}[(\rho - \sigma)A] = \langle L - I, A_c \rangle_\sigma. \quad (\text{S7})$$

Cauchy–Schwarz and $\text{Tr}[\sigma L] = 1$ give

$$\frac{[\text{Tr}[(\rho - \sigma)A]]^2}{\text{Var}_\sigma A} \leq \text{Tr}[\sigma(L - I)^2] = \mathcal{Q} - 1. \quad (\text{S8})$$

Equality holds for $A = L - I$, up to nonzero scaling and a scalar shift. For $\rho = \sigma$, the optimal ratio is zero. Optimizing the numerical score of a fixed POVM gives its classical Pearson divergence $Q(M) - 1$ by the same argument. Thus the receiver objective is the optimal finite-displacement signal-to-reference-variance ratio, plus one. It is not the Bayesian error probability of discriminating the hypotheses.

Classical finite-difference estimation bounds go back to Hammersley and Chapman–Robbins [4, 5]. Tsuda and Matsumoto’s quantum Hammersley–Chapman–Robbins theorem [6], Theorem 1, defines the finite-difference SLD through $\mathcal{J}_{\rho_\theta}(L_{\theta,\delta}^S) = (\rho_{\theta+\delta} - \rho_\theta)/\delta$. Take $\rho_\theta = \sigma$, $\rho_{\theta+\delta} = \rho$, and $\delta = 1$. Then $L^S = L - I$, and their SLD denominator is exactly $\mathcal{Q} - 1$. The spectral measurement gives the corresponding saturation. The same denominator is the minimal (Bures/SLD) quantum χ^2 divergence in Ref. [7]. We use the corrected arXiv version of that work and do not rely on its removed earlier Proposition 6.

S2 The fully-PPT bound and its attaining measurement

Consider N registers with Hilbert space $\bigotimes_{j=1}^N \mathcal{H}_j$. Write $\mathcal{M}_N$ for all POVMs on this space and $[N] = \{1, \dots, N\}$. Fix one orthonormal basis in each register. For $S \subseteq [N]$, $T_S(E) = E^{Ts}$ transposes the matrix indices of the registers in S and leaves the others unchanged. In the product basis grouped as $S : \bar{S}$, it acts linearly as

$$(|a\rangle\langle c| \otimes |b\rangle\langle d|)^{Ts} = |c\rangle\langle a| \otimes |b\rangle\langle d|, \quad \langle ab|E^{Ts}|cd\rangle = \langle cb|E|ad\rangle.$$

Thus T_S is neither an exponent nor a physical transpose gate. The condition $E^{Ts} \succeq 0$ means that the Hermitian matrix E^{Ts} has only nonnegative eigenvalues, not that its entries are nonnegative. Zero eigenvalues are allowed. For $M = \{E_x\} \in \mathcal{M}_N$, define

$$M \in \mathcal{P}_N \iff E_x^{Ts} \succeq 0 \text{ for every } x \text{ and every } S \subseteq [N]. \quad (\text{S9})$$

Ordinary positivity $E_x \succeq 0$ is already part of the POVM definition. The empty subset adds nothing, and $T_{[N]}$ is the full transpose, which preserves positivity. For $N = 1$ the restriction is vacuous: $\mathcal{P}_1 = \mathcal{M}_1$. For $N = 2$ either nontrivial subset gives the usual bipartite PPT condition. For more registers one must test every cut, not only a chosen bipartition. Changing the fixed local bases conjugates a partial transpose by a unitary, so positivity does not depend on that choice.

Fully separable effects and LOCC measurements meet this condition [8–10], but fully PPT is not a synonym for separable or classical. For identical inputs the numerical optimum is $Q_n^{\mathcal{P}} := \max_{M \in \mathcal{P}_n} Q_n(M)$. Thus $Q_n^{\mathcal{P}} = \mathcal{Q}^{\mathcal{P}_n}(\rho^{\otimes n}, \sigma^{\otimes n})$. The Letter’s $\mathcal{Q}^{\mathcal{P}_N}$ uses the specified register partition also for heterogeneous pairs. The hypotheses used here factor over those same registers. Internal subsystems of a register are not separate PPT factors.

For local full-rank pairs, set

$$L_j = \mathcal{J}_{\sigma_j}^{-1}(\rho_j), \quad Y_j = L_j \sigma_j L_j, \quad \sigma^{(N)} = \bigotimes_j \sigma_j, \quad \rho^{(N)} = \bigotimes_j \rho_j. \quad (\text{S10})$$

For each subset S , transpose the local factors in both L_j and σ_j :

$$\tilde{L}_S = \bigotimes_j L_j^{T^{1_{j \in S}}}, \quad \tilde{\sigma}_S = \bigotimes_j \sigma_j^{T^{1_{j \in S}}}. \quad (\text{S11})$$

Here T^0 means no transpose and T^1 means transpose. For any real u we have the exact identity

$$\bigotimes_j Y_j - 2u\rho^{(N)} + u^2\sigma^{(N)} = 2^{-N} \sum_S T_S \left[(\tilde{L}_S - uI) \tilde{\sigma}_S (\tilde{L}_S - uI) \right]. \quad (\text{S12})$$

To verify it, expand each bracket. The quadratic and constant terms become respectively $\bigotimes_j Y_j$ and $u^2\sigma^{(N)}$, independent of S . In the cross terms, partial transposition reverses the local order on S . Averaging all subsets produces every ordering in the expansion of $\bigotimes_j (\sigma_j L_j + L_j \sigma_j)/2 = \rho^{(N)}$.

Every matrix inside the square brackets is positive. For a fully-PPT effect E and the numbers

$$p = \text{Tr} \left[E \rho^{(N)} \right], \quad q = \text{Tr} \left[E \sigma^{(N)} \right], \quad y = \text{Tr} \left[E \left(\bigotimes_j Y_j \right) \right], \quad (\text{S13})$$

Eq. (S12) implies $y - 2up + u^2q \geq 0$ for all real u . Full rank gives $q > 0$ for a nonzero effect. Minimizing over u proves $p^2/q \leq y$. Summation over a POVM and $\sum_x E_x = I$ now give

$$Q(M) \leq \text{Tr} \left[\bigotimes_j Y_j \right] = \prod_j \text{Tr} \left[\sigma_j L_j^2 \right] = \prod_j \mathcal{Q}(\rho_j, \sigma_j), \quad M \in \mathcal{P}_N. \quad (\text{S14})$$

Conversely, the product of the local score PVMs is separable and reaches the product value. This proves main-text Theorem 1 for heterogeneous states and dimensions. Section S4 converts this value into the attained uncertainty coefficient and proves the fixed-observable noise bound.

S2.1 Measurement scope and detector interpretation

Theorem 1 does not identify PPT with separability. Its upper bound applies to the larger fully-PPT class, while a product measurement establishes achievability. The bound therefore applies beyond independent readout. The decomposition is finite and exact. It needs no optimization-duality assumption and no closure property under operator powers.

Suppose a calibrated detector gives $Q(M) > \prod_j Q_j$. At least one outcome operator must then have a negative partial transpose. The converse need not hold: an NPT measurement can fail to improve on

product readout. Failure to exceed the bound proves neither that the detector is PPT nor that the states commute. This test assumes known input states.

The state assumption and partition differ from the locally measured Rényi examples in Ref. [11]. Their isotropic and Werner constructions can be correlated or entangled inside each $A : B$ pair and their many-copy measurement cut can be $A^n : B^n$. Here the hypotheses are products *across the registers on which PPT is imposed*. The distinction prevents importing a data-hiding formula for a different partition.

S3 Ordered commutator surplus, stability, and finite-displacement onset

S3.1 Exact two-source residual

For a local pair define the Hermitian matrix $K = [\sigma, L]/(2i)$. Then $\sigma L = \rho + iK$ and $L\sigma = \rho - iK$. For two local pairs this gives

$$\mathcal{J}_{\sigma_A \otimes \sigma_B}(L_A \otimes L_B) = \rho_A \otimes \rho_B - K_A \otimes K_B. \quad (\text{S15})$$

In the joint quadratic variational problem the trial $A_0 = L_A \otimes L_B$ has value $Q_A Q_B$, because

$$2 \text{Tr}[(\rho_A \otimes \rho_B)A_0] - \text{Tr}[(\sigma_A \otimes \sigma_B)A_0^2] = 2Q_A Q_B - Q_A Q_B. \quad (\text{S16})$$

For the exact joint score L_{AB} , Eq. (S15) says $L_{AB} - A_0 = \mathcal{J}_{\sigma_A \otimes \sigma_B}^{-1}(K_A \otimes K_B)$. Completing the square with Eq. (S5) yields

$$\Delta_{AB} := Q_{AB} - Q_A Q_B = \langle K_A \otimes K_B, \mathcal{J}_{\sigma_A \otimes \sigma_B}^{-1}(K_A \otimes K_B) \rangle_{\text{HS}}. \quad (\text{S17})$$

The inverse is strictly positive as a quadratic form, so equality holds precisely when $K_A \otimes K_B = 0$. In a reference eigenbasis, $K_{ij} = (s_i - s_j)\rho_{ij}/[i(s_i + s_j)]$, which vanishes precisely when $[\rho, \sigma] = 0$. This proves main-text Proposition 1 without a nondegeneracy assumption.

S3.2 Positive ordered-moment expansion

For each local pair put

$$w_{ij} = \frac{2|\rho_{ij}|^2}{s_i + s_j}, \quad t_{ij} = \frac{s_i - s_j}{s_i + s_j}, \quad A_m = \sum_{ij} w_{ij} t_{ij}^{2m} \quad (m \geq 1). \quad (\text{S18})$$

The denominator of a tensor term factorizes as

$$s_i s_k + s_j s_l = \frac{1}{2}(s_i + s_j)(s_k + s_l)(1 + t_{ij} t_{kl}). \quad (\text{S19})$$

Swapping i, j leaves w_{ij} fixed and reverses t_{ij} . Therefore symmetrizing the two terms replaces $(1 + t_{ij} t_{kl})^{-1}$ by $(1 - t_{ij}^2 t_{kl}^2)^{-1}$. The uniformly convergent geometric series, valid since all $s_i > 0$, gives

$$\Delta_{AB} = \sum_{m=1}^{\infty} A_m^A A_m^B. \quad (\text{S20})$$

The first coefficient and a useful finite estimate are

$$A_1 = 2 \sum_{ij} \frac{|[\rho, \sigma]_{ij}|^2}{(s_i + s_j)^3}, \quad (\text{S21})$$

$$A_1^A A_1^B \leq \Delta_{AB} \leq \frac{A_1^A A_1^B}{1 - \eta_A^2 \eta_B^2}, \quad \eta = \frac{s_{\max} - s_{\min}}{s_{\max} + s_{\min}} < 1. \quad (\text{S22})$$

These bounds do not imply a state-independent positive minimum advantage.

Let $X_{ij} = |i\rangle\langle j| + |j\rangle\langle i|$ and $Y_{ij} = -i|i\rangle\langle j| + i|j\rangle\langle i|$. Their ordered covariance in σ is

$$C_{\sigma}^{ij} = \begin{pmatrix} s_i + s_j & i(s_i - s_j) \\ -i(s_i - s_j) & s_i + s_j \end{pmatrix}. \quad (\text{S23})$$

Thus t_{ij} is the imaginary covariance divided by the diagonal variance. Equations (S20)–(S22) express the surplus as positive contributions from this ordering information. This connects the calculation to ordered covariance [12, 13] without asserting that every quantum uncertainty bound is saturated.

S4 Readout variance and added fluctuations of a prescribed observable

We now fix the two preparations and the observable while changing the readout. The observable's intrinsic variance stays the same. The variance of the reported outcomes can change because the measurement may add noise.

S4.1 First and second moments of an unbiased readout

Let $R, S > 0$ be distinct density matrices in finite dimension, and let $M = \{E_x\}$ be a POVM. A real output assignment a_x is an unbiased readout of $A = A^\dagger$ when

$$A = \sum_x a_x E_x, \quad B = \sum_x a_x^2 E_x. \quad (\text{S24})$$

The first equality enforces the correct expectation for every input state. It is stronger than calibrating only two expectation values. No assertion is made that the complete output distribution equals the spectral distribution of A .

Introduce the isometry $W\psi = \bigoplus_x \sqrt{E_x} \psi$ and the block-diagonal operator $D = \bigoplus_x a_x I$. Then $W^\dagger W = I$, $A = W^\dagger D W$ and $B = W^\dagger D^2 W$. Therefore

$$N := B - A^2 = W^\dagger D(I - WW^\dagger)DW \succeq 0. \quad (\text{S25})$$

For this pair define $q_x = \text{Tr}[SE_x]$, $j(A) = \text{Tr}[(R - S)A]$, and $v(A) = \text{Var}_S A$. Thus $j(A) = \langle A \rangle_R - \langle A \rangle_S$ is a mean response, not necessarily a transport current. We obtain

$$V_S(M, a) = \sum_x q_x a_x^2 - (\text{Tr}[SA])^2 = v(A) + \nu_S(M, a; A), \quad \nu_S = \text{Tr}[SN] \geq 0. \quad (\text{S26})$$

The intrinsic noise operator in Eq. (S25) is standard measurement theory [14]. Under the imposed first-moment equality its expectation is added readout variance. It is not a disturbance of a second observable or an energetic cost. Spectral readout of A gives $N = 0$. Since $S > 0$, $\nu_S = 0$ implies $N = 0$.

S4.2 A sharp class-wide coefficient, and its fixed-observable consequence

For $p_x = \text{Tr}[RE_x]$ and $z_x = p_x/q_x - 1$, normalization gives $\sum_x q_x z_x = 0$. Hence

$$j(A) = \sum_x q_x z_x (a_x - \text{Tr}[SA]), \quad j(A)^2 \leq [Q(M; R, S) - 1] V_S(M, a). \quad (\text{S27})$$

For $Q(M) > 1$, equality holds exactly when $a_x - \text{Tr}[SA]$ is a real constant times z_x on the nonzero effects; the constant is nonzero when $j(A) \neq 0$. When $Q(M) = 1$, every recorded contrast is zero.

Let $\mathcal{C}$ be a measurement class with attained optimum $Q^c(R, S) > 1$. Equation (S27) implies, for every unbiased readout of any prescribed A ,

$$V_S(M, a) \geq \frac{j(A)^2}{Q^c - 1}, \quad \nu_S(M, a; A) \geq \left[\frac{j(A)^2}{Q^c - 1} - \text{Var}_S A \right]_+, \quad M \in \mathcal{C}. \quad (\text{S28})$$

There is no requirement that A be a current, energy, or work observable. Ratios involving $j(A)$ assume $j(A) \neq 0$; Eq. (S27) covers zero contrast without division. If $R = S$, every contrast vanishes and the normalized optimal observable below is not defined.

To locate the equality case, fix two numerical means μ and $\mu + \delta$, with $\delta \neq 0$. Leave the full operator A free. For an attaining measurement $M_{\mathcal{C}}$ set

$$a_x = \mu + \frac{\delta}{Q^c - 1} \left(\frac{p_x}{q_x} - 1 \right). \quad (\text{S29})$$

Its means under S, R are exactly $\mu, \mu + \delta$, and its variance is $\delta^2/(Q^c - 1)$. Thus

$$\min_{\substack{M \in \mathcal{C}, \\ a \in \mathbb{R}^{|M|} \\ \sum_x q_x a_x = \mu, \sum_x p_x a_x = \mu + \delta}} V_S(M, a) = \frac{\delta^2}{Q^c - 1}. \quad (\text{S30})$$

This optimum holds when only the two means are fixed. Fixing the whole operator A can prevent equality. Equation (S28) therefore remains a lower bound on its added noise, rather than a formula for the restricted minimum of every A .

For $R = \bigotimes_j \rho_j$, $S = \bigotimes_j \sigma_j$, and the fully-PPT class across those same registers, Section S2 gives $Q^{\mathcal{P}} = \prod_j Q_j$, with $Q_j = \mathcal{Q}(\rho_j, \sigma_j)$. Its product score measurement is an attaining $M_{\mathcal{C}}$. The output in Eq. (S29) is a product likelihood ratio followed by an affine transformation. It is generally not a sum of local currents. For identical pairs the coefficient is $Q_1^n - 1$, not $n(Q_1 - 1)$. The nonlinearity already available through classical output processing is not counted as a quantum surplus.

S4.3 The same optimal observable needs extra variance under PPT readout

For $R = \rho_n$, $S = \sigma_n$, let $L_n = \mathcal{J}_S^{-1}(R)$ and define

$$A_{\star} = \frac{L_n - I}{Q_n - 1}. \quad (\text{S31})$$

From $\text{Tr}[SL_n] = 1$ and $\text{Tr}[RL_n] = \text{Tr}[SL_n^2] = Q_n$,

$$\text{Tr}[SA_{\star}] = 0, \quad \text{Tr}[RA_{\star}] = 1, \quad \text{Var}_S A_{\star} = \frac{1}{Q_n - 1}. \quad (\text{S32})$$

Unrestricted spectral readout has this variance and no added noise. Every unbiased fully-PPT readout of the *same* $A_{\star}$ satisfies

$$\nu_S(M, a; A_{\star}) \geq \frac{1}{Q_1^n - 1} - \frac{1}{Q_n - 1} = \frac{Q_n - Q_1^n}{(Q_1^n - 1)(Q_n - 1)}. \quad (\text{S33})$$

Equivalently,

$$\frac{\nu_S(M, a; A_{\star})}{\text{Var}_S A_{\star}} \geq \frac{Q_n - Q_1^n}{Q_1^n - 1}. \quad (\text{S34})$$

For noncommuting states and $n \geq 2$ the right-hand sides are positive by Section S3. Changing the readout cannot change the first moments or intrinsic variance in Eq. (S32); it can only change the

output second moment subject to Eq. (S33). For the unnormalized score L_n , multiplication by $(Q_n - 1)^2$ yields

$$\nu_S(M, a; L_n) \geq \frac{(Q_n - 1)(Q_n - Q_1^n)}{Q_1^n - 1}. \quad (\text{S35})$$

A product informationally complete POVM spans the Hermitian operator space, so real scores can reconstruct $A_\star$ exactly in first moment. An unbiased restricted readout therefore exists. Its variance still obeys the penalty. The script `verify_readout_tur.py` reconstructs the scores and checks the inequality; it does not optimize over all PPT POVMs.

For contrast, $A = \sum_j A_j^{(j)}$ has a product spectral implementation. Its output second moment equals A^2 , giving $\nu_S = 0$. This is consistent with Eq. (S28), whose right-hand side must vanish for this target. The theorem is not a claim of strict advantage for every physical observable.

S4.4 A partial-transpose moment refinement

Let $S = \bigotimes_{j=1}^n \sigma_j > 0$ be the reference state and fix a Hermitian A on the full register space. An unbiased readout obeys $A = \sum_x a_x E_x$ with real scores. Put $B = \sum_x a_x^2 E_x$ and $\mu = \text{Tr}[SA]$. For each subset $J \subseteq [n]$ define

$$v_J(A; S) = \text{Var}_{S^{T_J}}(A^{T_J}), \quad b_{\text{mom}}(A; S) = \max_{J \subseteq [n]} v_J(A; S). \quad (\text{S36})$$

The empty subset is included, so $b_{\text{mom}} \geq \text{Var}_S A$. A transposition is taken in fixed local bases; simultaneous local changes of basis do not change this number.

Proposition S1 (Fixed-observable PPT moment bound) *Every unbiased fully-PPT readout of A satisfies*

$$V_S(M, a) \geq b_{\text{mom}}(A; S), \quad \nu_S(M, a; A) \geq b_{\text{mom}}(A; S) - \text{Var}_S A. \quad (\text{S37})$$

This bound does not require a signal state. For a product signal $R = \bigotimes_j \rho_j$, it dominates the two-mean readout bound:

$$b_{\text{mom}}(A; S) \geq \bar{v}(A; S) \geq \frac{[\text{Tr}[(R - S)A]]^2}{\prod_j \mathcal{Q}(\rho_j, \sigma_j) - 1}, \quad \bar{v} = 2^{-n} \sum_{J \subseteq [n]} v_J. \quad (\text{S38})$$

The last ratio assumes $R \neq S$.

Proof. For $M \in \mathcal{P}_n$, each family $\{E_x^{T_J}\}$ is a POVM: its effects are positive and sum to I . Applying the ordinary POVM Schwarz inequality to this family gives

$$B^{T_J} \succeq (A^{T_J})^2. \quad (\text{S39})$$

Since $S^{T_J} > 0$, trace pairing with it and the self-adjointness and involution of partial transposition give

$$\text{Tr}[SB] = \text{Tr}[S^{T_J} B^{T_J}] \geq \text{Tr}[S^{T_J} (A^{T_J})^2].$$

Also $\text{Tr}[S^{T_J} A^{T_J}] = \mu$. Subtracting μ^2 and maximizing over J proves Eq. (S37). The same first step applies to any reference positive under the specified partial transposes, but the following factorization uses a product reference.

Set $A_c = A - \mu I$ and $\mathcal{K} = \bigotimes_j \mathcal{J}_{\sigma_j}$. Averaging all subsets yields the exact identity

$$\bar{v}(A; S) = \text{Tr}[A_c \mathcal{K}(A_c)]. \quad (\text{S40})$$

For example, diagonalize each σ_j . With multi-indices $\mathbf{a}, \mathbf{b}$, both sides equal

$$\sum_{\mathbf{a}, \mathbf{b}} |(A_c)_{\mathbf{ab}}|^2 \prod_j \frac{s_{a_j}^{(j)} + s_{b_j}^{(j)}}{2}.$$

Each local index transposition exchanges which eigenvalue weights that entry; averaging generates the product of the two local choices.

Let $L_j = \mathcal{J}_{\sigma_j}^{-1}(\rho_j)$ and $L_\Pi = \bigotimes_j L_j$. Then $\mathcal{K}(L_\Pi) = R$ and $\mathcal{K}(I) = S$. In the positive inner product $\langle C, D \rangle_{\mathcal{K}} = \text{Tr}[C\mathcal{K}(D)]$ on Hermitian matrices,

$$\begin{aligned} \text{Tr}[(R - S)A] &= \langle L_\Pi - I, A_c \rangle_{\mathcal{K}}, \\ \|L_\Pi - I\|_{\mathcal{K}}^2 &= \prod_j \mathcal{Q}(\rho_j, \sigma_j) - 1. \end{aligned} \tag{S41}$$

Cauchy–Schwarz proves the second inequality in Eq. (S38); the maximum exceeds the average. This proof uses standard POVM moment positivity [14]; moment and partial-transpose methods are established also in the different context of state-separability tests [15]. No historical novelty is assigned to those ingredients. $\square$

Where the bounds agree. For $A = L_\Pi$, its product spectral readout is fully PPT and noiseless. Every v_J then equals $\prod_j Q_j - 1$, so both inequalities in Eq. (S38) are equalities. An affine rescaling gives any two prescribed distinct means. This gives another route to the sharp class-wide coefficient, while the bound for a general prescribed operator can be much larger. It does not solve the full optimization over all unbiased PPT implementations.

Checks, not a numerical proof of optimality. The suite `verify_moment_bounds.py` tests the subset-average identity on random complex states and observables, all-cut output inequalities for product readout, and equality for product optimal scores. It also tests cases beyond this shortened supplement. The recorded suite passes 1067 checks, with largest relative identity residual below 2×10^{-15} . These floating-point checks neither establish tightness of a fixed-observable bound nor replace the proof above.

S5 Work-stroke dephasing: invariant work statistics, decaying read-out surplus

S5.1 Protocol and unchanged work distribution

Prepare the medium in $\gamma_\beta = e^{-\beta H}/Z_\beta$. A cyclic pulse U leaves it in $\rho_0 = U\gamma_\beta U^\dagger$, with H restored. During the subsequent delay it obeys $\dot{\rho}_t = -i[H, \rho_t] + \Gamma(\mathcal{D}_H \rho_t - \rho_t)$. Here $\mathcal{D}_H(X) = \sum_E \Pi_E X \Pi_E$ removes coherence between distinct energy blocks. This map commutes with Hamiltonian evolution, giving the exact state

$$\rho_t = e^{-itH} \{ \mathcal{D}_H \rho_0 + c_t(\rho_0 - \mathcal{D}_H \rho_0) \} e^{itH}, \quad c_t = e^{-\Gamma t}. \tag{S42}$$

Energy populations stay constant, so the system energy and mean cyclic-pulse work do too. This system-level statement gives no value for reservoir heat or the cost of implementing the channel. The horizontal axis of main Fig. 1 is the dephasing entropy $\Sigma_\phi(t) = S(\rho_t) - S(\rho_0)$. It is nonnegative for this unital channel and differs from the collision surplus.

The standard two-projective-energy-measurement work protocol [16] has initial energy E and final energy E' . Even with degenerate energy blocks, its joint probability after the delay is

$$\Pr_t(E, E') = \frac{e^{-\beta E}}{Z_\beta} \text{Tr}[\Pi_{E'} U \Pi_E U^\dagger], \quad W = E' - E. \quad (\text{S43})$$

Every transition probability in Eq. (S43), and therefore the full work distribution, is independent of t . The timing matters: noise during the pulse may change those transitions. Work and coherent score readout use separate preparations because a final projective energy measurement would remove the coherence needed for the score task on that copy.

S5.2 Exact decay and scope of the variance bound

Set $\sigma = \gamma_\beta$ and $Q_j(t) = \mathcal{Q}(\rho_t^{\otimes j}, \sigma^{\otimes j})$. The reference state is a function of H , so terms within an energy block are orthogonal in the Jordan metric to terms between blocks. Define $Q_E = \mathcal{Q}(\mathcal{D}_H \rho_0, \sigma)$. Equation (S42) then gives

$$Q_1(t) = Q_E + c_t^2[Q_1(0) - Q_E]. \quad (\text{S44})$$

Let $L_t = \mathcal{J}_\sigma^{-1}(\rho_t)$ and $K_t = [\sigma, L_t]/(2i)$. The energy-block part of L_t commutes with σ . Jordan-map covariance under e^{-itH} therefore gives $K_t = c_t e^{-itH} K_0 e^{itH}$. The ordered-moment formula of Sec. S3 is quadratic in a pair of these commutator factors. It follows that

$$\Delta_2(t) = Q_2(t) - Q_1(t)^2 = c_t^4 \Delta_2(0). \quad (\text{S45})$$

The fourth-power law holds for this uniform-rate energy-dephasing channel, including degenerate H . It describes the difference between optimized unrestricted and fully-PPT collision values. Entropy production and the fixed-observable variance bound follow other time dependences; the latter also has changing denominators. At each delay in main Fig. 1, the readout classes are compared on the same $A_\star(t)$. That target changes along the trajectory.

S5.3 Interacting four-spin calculation

The open chain and square pulse are

$$H = -J \sum_{j=1}^3 Z_j Z_{j+1} - h \sum_{j=1}^4 X_j - b \sum_{j=1}^4 Z_j, \quad U = e^{-i\tau(H + \lambda \sum_{j=1}^4 Z_j)}. \quad (\text{S46})$$

We take $h/J = 0.8$, $b/J = 0.17$, $\lambda/J = 0.6$, $\beta J = 1$, $J\tau = 1.2$, and $\Gamma/J = 0.25$. At the start of the delay, $Q_1 = 4.9766570144$ and $Q_2 = 29.4165840530$. Their two-copy surplus is $\Delta_2 = 4.6494690142$. The unit-response target is $A_\star = (L_2 - I)/(Q_2 - 1)$. Its intrinsic variance is $1/(Q_2 - 1) = 0.0351907\dots$. The necessary fully-PPT moment bound is $0.1548592\dots$, so readout adds at least $0.1196685\dots$ in variance. These population calculations give bounds, not observed detector fluctuations or a fully-PPT implementation of this target.

The plotted points follow the exact state in Eq. (S42). The script `code/simulate_thermodynamics.py` also integrates the master equation and records every energy-transition probability before display binning. The companion `code/verify_thermodynamics.py` checks the decay and two-copy score calculation, along with energy and work invariance and the cyclic Jarzynski identity. These checks support the finite-system calculation. The medium has 16 energy levels; its optimal unrestricted two-copy score acts on 256 dimensions. We have not modeled source preparation, storage, calibration, detector cost, or a thermodynamic limit.

S6 Reduced equilibrium probes without an entangled source

The second example uses reduced equilibrium probes. Their temperature-dependent eigenvectors are already known in finite-coupling thermometry [17]. We take the one-ancilla Hamiltonian and probe marginal of Ullah, Naseem, and Müstecaplıoğlu [18]. The question here is how accurately a fixed collective observable can be read when two temperatures are compared and the detector is fully PPT.

S6.1 Exact states, separability, and noncommutation

Let $\omega, \Omega > 0$ and $g \geq 0$. Identity factors are implicit in

$$H_{SE} = \omega Z_S + \Omega Z_E + g X_S Z_E, \quad h = \sqrt{\omega^2 + g^2}. \quad (\text{S47})$$

This is Ref. [18], Eq. (1), with one ancilla, $\omega_p = 2\omega$, and $\omega_1 = 2\Omega$. Conditional on $Z_E = e \in \{-1, 1\}$, the probe Hamiltonian is $H_e = \omega Z_S + eg X_S$, with eigenvalues $\pm h$. Thus

$$\begin{aligned} G_\beta &= \frac{e^{-\beta H_{SE}}}{Z_{SE}(\beta)} = \sum_{e=\pm 1} w_e(\beta) \gamma_e(\beta) \otimes |e\rangle\langle e|, \\ w_e(\beta) &= \frac{e^{-\beta e \Omega}}{2 \cosh(\beta \Omega)}, \quad \gamma_e(\beta) = \frac{I - \tanh(\beta h)(\omega Z_S + eg X_S)/h}{2}. \end{aligned} \quad (\text{S48})$$

This is explicitly separable across $S : E$, and classical on E . Tracing out E gives

$$\tau_\beta = \frac{1}{2}[I + r_x(\beta)X_S + r_z(\beta)Z_S], \quad r_x = \frac{g}{h} \tanh(\beta h) \tanh(\beta \Omega), \quad r_z = -\frac{\omega}{h} \tanh(\beta h). \quad (\text{S49})$$

After converting energy conventions, this is the one-ancilla marginal of Ref. [18], Eqs. (13), (17)–(19). Its coherence was already described there. The global state and probe marginal are full rank at finite temperature.

For two different inverse temperatures $\beta, \beta_0 > 0$, abbreviate $t = \tanh(\beta h)$ and $t_0 = \tanh(\beta_0 h)$. Likewise, put $u = \tanh(\beta \Omega)$ and $u_0 = \tanh(\beta_0 \Omega)$. The Bloch vectors satisfy

$$(\mathbf{r}_\beta \times \mathbf{r}_{\beta_0})_y = \frac{g\omega}{h^2} t t_0 (u - u_0), \quad [\tau_\beta, \tau_{\beta_0}] = \frac{i}{2} (\mathbf{r}_\beta \times \mathbf{r}_{\beta_0})_y Y_S. \quad (\text{S50})$$

Thus the marginals fail to commute whenever $g\omega\Omega \neq 0$ and the positive finite temperatures differ. At $g = 0$, the bare-probe Gibbs states commute and the collective surplus vanishes. The effect already appears at finite coupling; no strong-coupling limit is required.

S6.2 The same observable under free and PPT readout

Prepare the probe–ancilla pairs independently and give the receiver only the probes. Set $\rho = \tau_\beta$ and $\sigma = \tau_{\beta_0}$. On two probes the hypotheses are $R = \rho^{\otimes 2}$ and $S = \sigma^{\otimes 2}$; no ancilla is shared across replicas. The product theorem and ordered-surplus identity apply to these accessible states. Choose $A_\star = (L_2 - I)/(Q_2 - 1)$. Its means are $\text{Tr}[SA_\star] = 0$ and $\text{Tr}[RA_\star] = 1$, while its intrinsic variance is $v_0 = (Q_2 - 1)^{-1}$. Spectral readout adds no variance. Every unbiased PPT readout obeys

$$V_S(M, a) \geq b_{\text{mom}}(A_\star; S) \geq \frac{1}{Q_1^2 - 1} > v_0. \quad (\text{S51})$$

At any one coupling, both readout classes see the same states and the same $A_\star$. Changing the plotted coupling changes all three. The final strict inequality requires distinct, noncommuting inputs. Neither the moment bound nor the class-wide two-mean bound determines the exact restricted minimum for this operator.

At $\beta\omega = 3$, $\beta_0\omega = 6$, $\Omega/\omega = 0.2$, and $g/\omega = 0.25$, the values are

Quantity	Value
Q_1	1.02435834934524
Q_2	1.04996555099548
v_0	20.0137891022
Two-mean readout variance lower bound	20.2798506334
PPT moment variance lower bound	49.5198347913
Added-variance lower bound from moments	29.5060456891
Moment lower bound / intrinsic variance minus one	1.47428583055
A feasible product-readout variance	195.819952797
$\lambda_{\min}(\rho), \lambda_{\min}(\sigma)$	0.0125893, 0.00451010

The moment bound is well above the intrinsic variance; the class-wide two-mean bound is much closer to it. The table uses the same intrinsic-variance denominator for its relative entry. Both lower bounds are necessary, and neither is claimed to be attained. The defining parameters and full-precision matrices determine the target; the rounded values only describe it.

S6.3 An explicit finite product implementation

A product informationally complete measurement provides a constructive upper bound on the unknown restricted minimum. Diagonalize the local reference and rotate the Pauli axes into that frame. On each probe use the six effects

$$F_{k,s} = \frac{I + s\tilde{\sigma}_k}{6}, \quad k = x, y, z, \quad s = \pm 1. \quad (\text{S52})$$

Operationally, this means choosing each of three Pauli axes with probability 1/3 and recording its binary result. Use $E_x = F_{k,s} \otimes F_{l,t}$ on the two probes. These 36 effects are product operators and span the full Hermitian space. They can read any $A_\star$ without first-moment bias, though not with its spectral distribution.

We choose the scores by the established minimum-variance data-processing prescription [19]. Use a real orthonormal basis of Hermitian matrices. Let the columns of T be the vectorized effects, and let a_{tar} represent $A_\star$ in the same basis. Put $W = \text{diag}(q_x)$, with $q_x = \text{Tr}[SE_x] > 0$. A minimum-second-moment solution is

$$a = W^{-1}T^\top(TW^{-1}T^\top)^{-1}a_{\text{tar}}. \quad (\text{S53})$$

The implementation uses a stable weighted least-squares solve rather than explicitly forming this inverse. It verifies $\sum_x a_x E_x = A_\star$ on the full four-dimensional Hilbert space, not only on the two calibrated states. The resulting variance is 195.81995..., so

$$49.51983\dots \leq \inf_{\substack{M \in \mathcal{P}_2, \, a \in \mathbb{R}^{|M|} \\ \sum_x a_x E_x = A_\star}} V_S(M, a) \leq 195.81995\dots \quad (\text{S54})$$

The lower number comes from the necessary moment inequality. The upper number comes from the specified product POVM with its optimal scores. Numerical residuals are recorded separately; they are not formal interval certificates. We have not optimized over all PPT POVMs. At the marked point, product scores range from about -24.89 to 1021.61 . Source calibration, detector-error analysis, and an experimental sampling budget would have to account for this range. Ideal sample draws do not validate a detector.

S6.4 Global access, thermometry, and resource accounting

The full states G_β, G_{β_0} are functions of one fixed Hamiltonian and commute. If every complete probe-ancilla pair is available, a common energy measurement already attains its collision value. Therefore, on two full replicas,

$$Q_2(G_\beta, G_{\beta_0}) = Q_1(G_\beta, G_{\beta_0})^2.$$

At the marked point, the full one-copy value is $2.56626736485\dots$, larger than the reduced-probe value. Full access removes the collective surplus because the complete hypotheses commute, while providing more total information. Here each access choice has its own optimized score. We have not shown that separate measurement of the full system reads the previously fixed probe observable $A_\star$ without noise. The reduced states fail to commute because the inaccessible classical ancilla label has temperature-dependent weights. This requires no entangled source and violates no data-processing inequality.

The task compares two finite temperatures through a prescribed observable. It does not estimate a continuously encoded temperature locally; single-parameter SLD Fisher information remains additive for independent probes. The coherent local thermometric response of this model was studied in Ref. [18]. We have modeled neither preparation and equilibration costs nor engine efficiency, and claim no universal strong-coupling advantage. This example has no applied work stroke or entangled input source. Giving the receiver the ancillas would change the task.

S6.5 Numerical audit and figure scope

The scripts `equilibrium_readout.py` and `verify_equilibrium_readout.py` calculate the marginal in three ways: direct global Gibbs exponentiation, the separable decomposition, and Eq. (S49). They check commutation of the global states and noncommutation of the marginals, as well as data processing and the score and moment identities. A separate calculation with `mpmath` recomputes the marked point at 70 decimal digits. Across the recorded script checks, all 1396 pass, with largest relative residual below 1.2×10^{-13} . These are implementation checks, not independent referee validation or a proof of global numerical optimality.

Main Fig. 2 scans 196 positive couplings $g/\omega \in [0.025, 1]$ at a fixed temperature pair and $\Omega/\omega = 0.2$. Each point recalculates the target and compares readout classes on that same operator. The red curve uses one feasible product implementation. We have not optimized over temperatures, Hamiltonian parameters, or all product POVMs. The plotted lines are population variances or analytic lower bounds. Ideal sample draws are stored separately. The CSV contains absolute variances and older normalized ratios; the figure uses only the absolute columns. Marked-point matrices, probabilities, and scores are in `equilibrium_marked_matrices.npz`.

S7 Data map and finite-shot scope

The two main figures come from `data/thermo_moment_bounds.csv` and `data/equilibrium_readout_scan.csv`, respectively. The script `code/make_audited_figures.py` draws both. For the defining equations and parameters, use Secs. S5–S6 rather than rounded points read from the plots. The model data come from `code/simulate_thermodynamics.py` and `code/equilibrium_readout.py`; companion scripts check the identities. Equilibrium matrices, probabilities, and scores for the marked point are in `data/equilibrium_marked_matrices.npz` and `data/equilibrium_marked.json`. No hardware data are reported.

Both plots show absolute population variance at unit response. The older normalized CSV columns do not enter either plot. At each delay or coupling, we recalculate the states and $A_\star$ before comparing readouts on that same operator. Green gives a necessary moment bound. The red curve in Fig. 2 gives

the variance of one feasible product POVM. The shading measures the arithmetic lower bound on added variance, including where the vertical axis is logarithmic.

The seeded multinomial draws in `data/monte_carlo/` sample known ideal Born probabilities, with no implemented detector. Rare outcomes and scores reaching about 1022 can make finite-sample variance estimates unstable. A sample variance below a population lower bound reflects sampling error rather than a violation of the operator inequality. Source calibration, readout error, and experimental shot budgets remain unassessed.

References

- [1] K. Temme and F. Verstraete, J. Math. Phys. **56**, 012202 (2015), arXiv:1112.6343 .
- [2] M. Berta, O. Fawzi, and M. Tomamichel, Lett. Math. Phys. **107**, 2239 (2017), arXiv:1512.02615 .
- [3] B. Regula and M. Tomamichel, Rethinking quantum smooth entropies: Tight one-shot analysis of quantum privacy amplification (2026), version 4, revised 1 September 2026, arXiv:2603.04493 .
- [4] J. M. Hammersley, J. R. Stat. Soc. B **12**, 192 (1950).
- [5] D. G. Chapman and H. Robbins, Ann. Math. Stat. **22**, 581 (1951).
- [6] Y. Tsuda and K. Matsumoto, J. Phys. A: Math. Gen. **38**, 1593 (2005), arXiv:quant-ph/0207150 .
- [7] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete, J. Math. Phys. **51**, 122201 (2010), corrected arXiv version 3 (2024), arXiv:1005.2358 .
- [8] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996), arXiv:quant-ph/9604005 .
- [9] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996), arXiv:quant-ph/9605038 .
- [10] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Commun. Math. Phys. **328**, 303 (2014), arXiv:1210.4583 .
- [11] T. Rippchen, S. Sreekumar, and M. Berta, IEEE Trans. Inf. Theory **71**, 6105 (2025), arXiv:2405.05037 .
- [12] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
- [13] D. S. P. Salazar, Phys. Rev. E **114**, L022106 (2026), arXiv:2512.24956 .
- [14] M. Ozawa, Annals of Physics **311**, 350 (2004), arXiv:quant-ph/0307057 .
- [15] A. Miranowicz, M. Piani, P. Horodecki, and R. Horodecki, Phys. Rev. A **80**, 052303 (2009), arXiv:quant-ph/0605001 .
- [16] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007), arXiv:cond-mat/0703189 .
- [17] H. J. D. Miller and J. Anders, Nat. Commun. **9**, 2203 (2018), arXiv:1801.08057 .
- [18] A. Ullah, M. T. Naseem, and Ö. E. Müstecaplıoğlu, Phys. Rev. Research **5**, 043184 (2023), arXiv:2211.05461 .
- [19] G. M. D’Ariano and P. Perinotti, Phys. Rev. Lett. **98**, 020403 (2007), arXiv:quant-ph/0610058 .