Finite-copy risk for polynomial energies of thermal Gaussian states

Scott Schubr
Luminary Oddities

October 30, 2026

Abstract

We study estimation of degree-n normally ordered polynomial energies from N copies of a centered thermal Gaussian state with unknown occupation covariance $A \succeq aI$, $a > 0$. A positive copy-averaged observable has an exact coefficient-dependent variance and minimizes variance at each covariance for every fixed polynomial among globally unbiased measurements and decisions. With $\gamma = 1 + a^{-1}$, its sharp worst-polynomial variance gives the arbitrary-bias minimax scale $\gamma n^2/N$. At fixed a and growing degree, the minimax relative squared risk tends exactly to $1 - e^{-\lambda}$ when this scale tends to $\lambda \in (0, \infty)$. Standard heterodyne records have an explicit optimal unbiased estimator and the corresponding minimax scale $\gamma^2 n^2/N$, yielding a matched low-brightness measurement separation. The proofs combine classical symmetry averaging and Gaussian moment methods with completeness of the copy-invariant thermal family and a supported-prior argument for negative-binomial photon data. Worked comparisons show how statistical pooling and polynomial coefficient geometry reduce fluctuations at a fixed copy budget.

Keywords. Quantum estimation; thermal Gaussian states; polynomial moments; minimax risk; symmetry averaging.
MSC 2020. 62F12, 62F10, 81P50.

1 Introduction

A thermal optical field can be interrogated through normally ordered moments. For example, a single-mode factorial photon moment has the form $\langle (b^\dagger)^n b^n \rangle$, while a two-mode intensity correlation is $\langle b_1^\dagger b_2^\dagger b_1 b_2 \rangle$. Both belong to the family $\operatorname{Tr} \rho f(b)^\dagger f(b)$, where f is a known homogeneous holomorphic polynomial. We ask how accurately such an energy can be estimated from a finite number of independent copies of an unknown thermal Gaussian state, allowing collective measurements.

The positive coherent-state representation makes the target a classical complex Gaussian moment. It does not give the observer samples of the latent coherent amplitudes: the data consist of quantum states. This distinction changes the fluctuation calculation. Moreover, high-order moments require control beyond the first term of a fixed-degree large-sample expansion. When the degree grows with the number of copies, the omitted contraction orders can contribute on the same scale as the leading term.

The main conclusion is an exact optimal *unbiased* risk for every fixed polynomial and covariance. Its formula retains all contraction orders and the geometry of the coefficient tensor. A second conclusion addresses arbitrary bias: under a known positive occupation floor a, the worst-polynomial risk is governed by $x = (1+a^{-1})n^2/N$. We determine its vanishing and collapse regimes and, for fixed a and growing degree, its exact critical limit $1 - e^{-\lambda}$ when $x \to \lambda$. Standard heterodyne replaces this governing scale by $(1 + a^{-1})^2 n^2/N$. Thus low brightness creates a regime in which the collective

quantum benchmark is consistent while every all-heterodyne procedure has worst-polynomial risk tending to one. A separate same-target photon-count comparison in Section 7.2 shows how pooling can turn an exponentially noisy direct moment average into a consistent estimate.

Prior capability and the present increment. Quantum U-statistics and their Hoeffding analysis were developed by Guţă and Butucea [3]; finite-dimensional functional-gradient and efficiency results appear in [1]. Tsang's quantum Rao–Blackwell theorem supplies invariant-state averaging [12, Section 5.3]. General quantum completeness and minimum-variance principles are established in [8, Section 5]. Normally ordered moments of positive coherent mixtures are treated in quantum semiparametric estimation [11], while photon counting for scalar thermal sensing is analyzed in [9].

Here the full unknown covariance retains its mode orientation. Copy averaging is followed by a concrete completeness proof for this thermal family, which makes the coefficient-dependent variance optimal for each polynomial. Coherent lower symbols reduce the argument to complete complex Wishart families, including singular sample scatter matrices. Finite photon cutoffs justify the treatment of arbitrary integrable real-decision POVMs. The Gaussian restriction also permits repeated copy labels, giving an unbiased observable for every positive copy count, including $N < n$.

Classical complex Wishart moment formulae [2, 7, 5] provide the combinatorial background for the exact contractions and heterodyne calculation. The direct-amplitude experiment in [10] already has an exact Gaussian-energy minimax formula and the associated power-estimation prior argument. Its observations are Gaussian amplitudes. Our scalar lower experiment instead has negative-binomial photon statistics and an occupation floor; a supported Beta prior and its tilted posteriors yield the critical quantum risk. The common decision-theoretic mechanism is credited where it enters the proof.

Two nearby inference settings clarify the scope of this comparison. Lahiry's residual-photon reduction treats displaced scalar thermal states and fixed smooth targets [6]. Polynomial coherent-displacement sensing with nonlinear quantum processing and heterodyne benchmarks is developed in [4]. The parameter here is the full centered thermal covariance, and the conclusions give exact finite-copy relative variance and a growing-degree minimax limit.

Measurement and proof route. The resource is N state copies. The quantum estimator is the spectral measurement of an observable averaged over passive copy mixing. Randomly selecting one mixing direction measures a different observable-valued experiment. In the heterodyne comparison every mode of every copy is heterodyned, followed by arbitrary classical processing. These precisely specified access models make the comparison a statistical resource benchmark.

Section 3 constructs the observable and evaluates its variance. Section 4 proves pointwise unbiased optimality. Section 5 gives the arbitrary-bias bounds and exact critical limit. Section 6 derives the heterodyne benchmark. The examples in Section 7 show how the conclusions distinguish coefficient structure and copy budgets.

2 Experiment and main results

Let d, n, N be positive integers. On d bosonic modes write $b_1, \dots, b_d$ for annihilation operators, satisfying $[b_i, b_j^\dagger] = \delta_{ij}\mathbb{I}$. For a positive definite Hermitian matrix A, define

$$\rho_A = \int_{\mathbb{C}^d} |z\rangle\langle z| \, d\,\mathrm{CN}(0, A)(z), \qquad A_{ij} = \mathbb{E}[Z_i \overline{Z_j}]. \tag{1}$$

Thus A is the occupation covariance, with vacuum noise accounted for by the commutation relations. All states in this paper have the representation (1).

Let $f \neq 0$ be homogeneous and holomorphic of degree n, with symmetric coefficients

$$f(z) = \sum_{I \in [d]^n} c_I z_{i_1} \cdots z_{i_n}. \tag{2}$$

The target and loss are

$$F_f(A) = \operatorname{Tr} \rho_A f(b)^\dagger f(b) = \mathbb{E}|f(Z)|^2, \qquad L_A(y) = \left(\frac{y}{F_f(A)} - 1 \right)^2. \tag{3}$$

The equality follows from $b_i|z\rangle = z_i|z\rangle$ and positivity. We observe $\rho_A^{\otimes N}$, allowing any collective POVM and real-valued decision. The known parameter restriction is

$$A \succeq aI_d, \qquad a > 0, \qquad \gamma = 1 + a^{-1}. \tag{4}$$

This is a spectral floor in every mode. The estimator may depend on f, d, N, a.

For copy r, denote its annihilation operator by $b_{i,r}$. For an ordered tuple $\boldsymbol{r} \in [N]^n$ set

$$C_{\boldsymbol{r}} = \sum_I c_I \prod_{j=1}^n b_{i_j, r_j}, \qquad T_f = \frac{n!}{(N)^{\overline{n}}} \sum_{\boldsymbol{r} \in [N]^n} C_{\boldsymbol{r}}^\dagger C_{\boldsymbol{r}}. \tag{5}$$

Here $(N)^{\overline{k}} = N(N+1)\cdots(N+k-1)$, and $(N)^{\overline{0}} = 1$. Repeated indices r_j are included.

For stating the exact formula, choose a mode basis diagonalizing A, with eigenvalues λ_i, and transform c into that basis. Define

$$t_I = c_I \prod_{j=1}^n \sqrt{\lambda_{i_j}}, \qquad \psi = \frac{t}{\|t\|} \in \operatorname{Sym}^n(\mathbb{C}^d), \qquad R_k = \operatorname{Tr}_{n-k} |\psi\rangle\langle\psi|, \qquad D_A = \operatorname{diag}(1 + \lambda_i^{-1}). \tag{6}$$

The coordinate change in (6) is used in the analysis; the observable (5) is independent of the unknown A.

Theorem 2.1 (Exact variance and sharp uniform bound)**.** *For all $d, n, N \geq 1$, the positive observable T_f is unbiased for $F_f(A)$ and has finite second moment. Its relative variance is*

$$\frac{\operatorname{Var}_A(T_f)}{F_f(A)^2} = \sum_{k=1}^n \binom{n}{k}^2 \frac{k!}{(N)^{\overline{k}}} \operatorname{Tr}\left(R_k^2 D_A^{\otimes k} \right). \tag{7}$$

Under (4)*, put*

$$S(N, n, a) = \sum_{k=1}^n \binom{n}{k}^2 \frac{k! \gamma^k}{(N)^{\overline{k}}}. \tag{8}$$

Then the relative variance is at most $S(N, n, a)$. Equality holds for $d = 1$, $f(z) = z^n$, $A = a$, so the bound is sharp simultaneously in all parameters.

Theorem 2.2 (Pointwise unbiased optimality)**.** *Fix $d, n, N \geq 1$, $a > 0$ and $f \neq 0$. For every real-decision collective POVM Y that is integrable and unbiased for $F_f(A)$ at every $A \succeq aI_d$,*

$$\operatorname{Var}_A(Y) \geq \operatorname{Var}_A(T_f) \qquad (A \succeq aI_d). \tag{9}$$

Thus (7) *is the exact minimum unbiased relative variance for each polynomial and covariance. Infinite competitor variance is allowed.*

Define the worst-polynomial minimax risk

$$\mathcal{R}(N,n,a) = \sup_{d\geq 1} \sup_{0\neq f\in \mathrm{Sym}^n(\mathbb{C}^d)} \inf_{\mathsf{M},Y} \sup_{A\succeq aI_d} \mathbb{E}_A L_A(Y). \tag{10}$$

The notation for f in (10) identifies polynomials with their coefficient tensors. Define $\mathcal{R}_{\mathrm{ub}}$ by restricting the infimum to estimators unbiased at every allowed A.

Theorem 2.3 (Unbiased optimum and biased minimax order)**.** *For all $n, N \geq 1$ and $a > 0$,*

$$\mathcal{R}_{\mathrm{ub}}(N,n,a) = S(N,n,a). \tag{11}$$

With $x = \gamma n^2/N$,

$$\frac{7}{2304}\min(1,x) \;\leq\; \mathcal{R}(N,n,a) \;\leq\; \frac{S}{1+S} \;\leq\; 1-e^{-x} \;\leq\; \min(1,x). \tag{12}$$

The upper bound $S/(1+S)$ is attained as a risk bound by a spectral measurement of $T_f/(1+S)$.

Theorem 2.4 (Minimax phases and exact critical limit)**.** *For any sequence of allowed triples (N_j, n_j, a_j), write $x_j = (1+a_j^{-1})n_j^2/N_j$. Then*

$$\mathcal{R}(N_j,n_j,a_j) \longrightarrow 0 \quad\Longleftrightarrow\quad x_j \longrightarrow 0, \tag{13}$$

$$x_j \longrightarrow \infty \quad\Longrightarrow\quad \mathcal{R}(N_j,n_j,a_j) \longrightarrow 1. \tag{14}$$

For fixed $a > 0$, if $n \to \infty$ and $\gamma n^2/N \to \lambda \in (0,\infty)$, then

$$\mathcal{R}(N,n,a) \longrightarrow 1-e^{-\lambda}. \tag{15}$$

These results quantify the resources of collective quantum estimation in the declared state model. The exact formula also retains dependence on the polynomial through the reduced-tensor purities in (7).

3 Symmetry averaging and the exact risk formula

3.1 Construction, domains and first moment

Let v be uniform on the complex unit sphere in $\mathbb{C}^N$ and write

$$b_i(v) = \sum_{r=1}^{N} v_r b_{i,r}, \qquad H_v = f(b(v))^\dagger f(b(v)). \tag{16}$$

The combined modes have canonical commutators, and their state in $\rho_A^{\otimes N}$ is again ρ_A. The sphere identity

$$\mathbb{E}_v[v_{r_1}\cdots v_{r_n}\overline{v_{s_1}\cdots v_{s_n}}] = \frac{1}{(N)^{\overline{n}}} \sum_{\pi\in S_n}\prod_{j=1}^{n}\delta_{r_j,s_{\pi(j)}} \tag{17}$$

follows by writing a standard complex Gaussian vector as its independent radius and direction. Its squared radius has nth moment $(N)^{\overline{n}}$, while the Gaussian numerator is the Wick sum in (17). Symmetry of c and commutation of annihilators give

$$T_f = \mathbb{E}_v H_v. \tag{18}$$

In particular $\mathbb{E}_A T_f = F_f(A)$, since this is true for every H_v.

For a precise spectral construction, first work on finite total-photon sectors in the dN modes. Each $C_{\boldsymbol{r}}^\dagger C_{\boldsymbol{r}}$ preserves the total number in each copy. Every sector with bounded total photon number is finite dimensional, and (5) defines a positive matrix there. The direct sum of the fixed-total-number matrices defines a positive self-adjoint operator with the finite-particle vectors as a core. Its spectral measurement is the estimator used here. Thermal states on finitely many modes have all polynomial number moments finite, so the second moment exists and the sector identities pass to the full state. Formula (18) refers to an averaged observable.

3.2 All contraction orders

We prove Theorem 2.1 by computing $\mathbb{E}_A(H_v H_w)$ for two unit vectors v, w. In an eigenbasis of A, the nonzero two-point contractions are

$$\mathbb{E}_A[b_i(v)^\dagger b_j(w)] = \delta_{ij}\lambda_i \sum_r \overline{v_r} w_r,$$
$$\mathbb{E}_A[b_i(v) b_j(w)^\dagger] = \delta_{ij}(1+\lambda_i) \sum_r v_r \overline{w_r}.$$

The product $H_v H_w$ has four groups, each of size n: creators at v, annihilators at v, creators at w, and annihilators at w. In a Wick pairing let k be the number of contractions from the first annihilator group to the second creator group. Balance forces k contractions in the reverse direction between the other two groups. The remaining $n-k$ contractions on each side are internal. The number of labeled pairings in this class is

$$\binom{n}{k}^4 (k!)^2((n-k)!)^2 = (n!)^2 \binom{n}{k}^2. \tag{19}$$

Symmetry of c makes their coefficient contractions identical.

After whitening by (6), the internal contractions yield the k-body reduction R_k. More explicitly, for ordered multi-indices $I, J \in [d]^k$,

$$(R_k)_{I,J} = \sum_{L \in [d]^{n-k}} \psi_{I,L} \overline{\psi_{J,L}}.$$

The coefficient contraction for the kth class, divided by $\|t\|^4$, is

$$\sum_{I,J} \left(\prod_{\ell=1}^{k} (1 + \lambda_{i_\ell}^{-1}) \right) |(R_k)_{I,J}|^2 = \operatorname{Tr}(R_k^2 D_A^{\otimes k}). \tag{20}$$

The cross-copy contractions contribute $|\langle v, w\rangle|^{2k}$. Since $F_f(A) = n!\|t\|^2$, equations (19)–(20) give

$$\frac{\mathbb{E}_A(H_v H_w)}{F_f(A)^2} = \sum_{k=0}^{n} \binom{n}{k}^2 |\langle v, w\rangle|^{2k} \operatorname{Tr}(R_k^2 D_A^{\otimes k}), \tag{21}$$

where the $k = 0$ term is one. All terms are nonnegative real numbers. Finite thermal fourth moments of $f(b(v))$ also bound the ordered product by state Cauchy–Schwarz, uniformly over the compact sphere, so the integrals in (21) are legitimate.

For independent uniform v, w, unitary invariance and (17) imply

$$\mathbb{E}|\langle v, w\rangle|^{2k} = \frac{k!}{(N)^{\overline{k}}}.$$

This includes $N = 1$, when both sides equal one. Integrating (21) and subtracting its $k = 0$ term proves (7). Finally $D_A \preceq \gamma I$ and $\operatorname{Tr} R_k^2 \leq 1$ prove (8). A single-mode coefficient tensor is pure at every reduction and $A = a$ makes every weight equal γ, proving simultaneous equality. This completes the proof of Theorem 2.1.

4 Pointwise unbiased optimality

The extra ingredient beyond averaging is completeness: a copy-invariant observable with zero expectation throughout the covariance family must vanish. The classical input is the open natural parameter space of the Wishart family. We give the analytic details needed to use that input for unbounded quantum decisions.

Lemma 4.1 (Complete scatter family)**.** *Let Z have N independent $\mathrm{CN}(0, V)$ columns. The laws of ZZ^*, as V varies over any nonempty open subset of positive definite Hermitian matrices, are complete for functions integrable throughout that subset.*

Proof. Relative to a covariance V_0 in the subset, the likelihood ratio is

$$\left(\frac{\det V_0}{\det V}\right)^N \exp\{-\operatorname{Tr}[(V^{-1} - V_0^{-1})ZZ^*]\}.$$

The precision differences contain a neighborhood of zero in the real vector space of Hermitian matrices. A zero-expectation function defines an exponentially integrable finite signed measure whose Laplace transform vanishes there. Uniqueness of that transform gives the zero measure. This argument applies on the common support of the scatter laws, including the rank-at-most-N cone when $N < d$; see also the exponential-family formulation in [7, Section 2]. □

Lemma 4.2 (Thermal completeness after copy averaging)**.** *Let $K = \bigoplus_{m\geq 0} K_m$ be a self-adjoint observable on the dN-mode Fock space, with finite total-photon blocks. Suppose K is invariant under passive $U(N)$ mixing of the copies and $\operatorname{Tr} \rho_A^{\otimes N}|K| < \infty$ for every $A \succeq aI_d$. If $\operatorname{Tr} \rho_A^{\otimes N} K = 0$ throughout that family, then $K = 0$.*

Proof. At $A = tI_d$, put $q = t/(1 + t)$. The state is

$$\rho_{tI}^{\otimes N} = (1 - q)^{dN} \bigoplus_{m\geq 0} q^m I_m.$$

Consequently $\sum_m q^m \operatorname{Tr}|K_m| < \infty$ for each $q > a/(1 + a)$, and $\|K_m\| \leq C_q q^{-m}$. For a coherent amplitude matrix $Z \in \mathbb{C}^{d\times N}$ the lower symbol $b_K(Z) = \langle Z|K|Z\rangle$ is therefore continuous everywhere. Indeed its degree-m contribution, after multiplication by $e^{\|Z\|_F^2}$, is bounded by $C_q q^{-m}\|Z\|_F^{2m}/m!$. With $\overline{Z}$ replaced by an independent complex matrix W, the bound is $C_q q^{-m}(\|Z\|_F\|W\|_F)^m/m!$. The series therefore converges uniformly on compact sets and defines an entire complexification.

Copy invariance gives $b_K(ZU) = b_K(Z)$. Two amplitude matrices have the same row Gram matrix precisely when they differ by a right unitary; this includes deficient rank. Thus $b_K(Z) = g(ZZ^*)$ for a measurable g. The positive coherent mixture yields

$$\mathbb{E}_A g(ZZ^*) = \operatorname{Tr} \rho_A^{\otimes N} K = 0, \qquad \mathbb{E}_A|g(ZZ^*)| \leq \operatorname{Tr} \rho_A^{\otimes N}|K| < \infty.$$

Positive and negative spectral parts justify these identities by Tonelli. Lemma 4.1, using $A \succ aI_d$, gives $b_K = 0$ almost everywhere for a Gaussian law on the full amplitude space. Its positive density and continuity imply $b_K = 0$ everywhere. Expanding $e^{\|Z\|_F^2} b_K(Z)$ in $\overline{Z}^\alpha Z^\beta$ now gives every matrix entry $\langle\alpha|K|\beta\rangle/\sqrt{\alpha!\beta!} = 0$. Hence every block vanishes. □

Proof of Theorem 2.2. Let $P_{\le L}$ project onto total photon number at most L, and compress the decision POVM M to this finite space. Its first moment

$$B_L = \int y\, P_{\le L} M(dy) P_{\le L}$$

exists: isotropic thermal integrability controls all diagonal absolute moments, and positivity and scalar Cauchy–Schwarz control the off-diagonal entries. Average B_L over copy $U(N)$. Its central phase removes cross-total-number blocks, giving compatible blocks K_m for $m \le L$. Trace-norm contraction and the triangle inequality give

$$\sum_m q^m \operatorname{Tr}|K_m| \le \sum_m q^m \operatorname{Tr} \int |y| P_m M(dy) P_m < \infty$$

for every allowed isotropic q. Their direct sum K is self-adjoint.

For general A, write $Q = A(I + A)^{-1}$ and choose $q > \|Q\|$ also above $a/(1 + a)$. In the total-photon decomposition,

$$\rho_A^{\otimes N} = \det(I_d + A)^{-N} \bigoplus_{m \ge 0} \operatorname{Sym}^m(Q \otimes I_N), \qquad \operatorname{Sym}^m(Q \otimes I_N) \preceq q^m I_m.$$

Here $\operatorname{Sym}^m(B)$ denotes the restriction of $B^{\otimes m}$ to the symmetric tensor space. Hence $\rho_A^{\otimes N}$ is dominated by a finite multiple of the corresponding isotropic thermal state. This proves integrability of K and absolute convergence of the block sum for its expectation. State invariance gives $\operatorname{Tr} \rho_A^{\otimes N} K = \mathbb{E}_A Y = F_f(A)$.

Apply Lemma 4.2 to $K - T_f$. The required isotropic integrability follows from the block trace-norm triangle inequality, and general integrability from the same domination. Thus $K = T_f$.

At a comparison A with finite $\mathbb{E}_A Y^2$, faithfulness gives $P_{\le L} \preceq c_{A,L} \rho_A^{\otimes N}$ for a finite $c_{A,L}$. Thus the compressed matrix second moment $C_L = \int y^2 P_{\le L} M(dy) P_{\le L}$ exists. Matrix Cauchy–Schwarz gives $B_L^2 \preceq C_L$. Jensen's inequality for the finite-dimensional copy twirl, followed by state invariance, gives

$$\sum_{m \le L} \operatorname{Tr} \rho_A^{\otimes N} K_m^2 \le \mathbb{E}_A Y^2.$$

Letting $L \to \infty$ increases the left side to $\operatorname{Tr} \rho_A^{\otimes N} T_f^2$. Subtract the common squared mean. An infinite second moment already satisfies (9). □

4.1 Sharp uniform risk and the scalar statistic

Theorem 2.2 and the scalar equality in Theorem 2.1 prove $\mathcal{R}_{\mathrm{ub}} = S$. The scalar experiment also supplies the biased lower bounds. A single photon count has probabilities

$$p_t(k) = \frac{t^k}{(1 + t)^{k+1}}, \qquad k \ge 0. \tag{22}$$

The total count K over N copies has mass $\binom{N+k-1}{k} t^k/(1 + t)^{N+k}$. Conditional on K, individual counts are uniform over compositions, independently of t. Since the states commute, every collective

POVM is a parameter-independent classical randomization of those counts. Thus K retains the scalar experiment for arbitrary decisions.

For $f(z) = z^n$, counting ordered selections of photons gives

$$T_f = \frac{n!(K)_{\underline{n}}}{(N)^{\overline{n}}}, \qquad F_f(t) = n!t^n. \tag{23}$$

The factorial identity $\mathbb{E}_t(K)_{\underline{n}} = (N)^{\overline{n}}t^n$ checks unbiasedness directly. Its second moment also gives

$$\mathbb{E}_t\left(\frac{T_f}{F_f(t)}\right)^2 = \sum_{j=0}^{n}\binom{n}{j}^2 j!\frac{(N)^{\overline{2n-j}}}{((N)^{\overline{n}})^2}t^{-j}. \tag{24}$$

This alternative expression checks the normalization of the contraction formula.

5 Biased minimax bounds and asymptotic phases

5.1 Upper bound

If T is unbiased and has relative variance at most S, its rescaling $T/(1+S)$ has relative risk at most

$$\frac{S}{(1+S)^2} + \left(\frac{1}{1+S} - 1\right)^2 = \frac{S}{1+S}.$$

Moreover $\binom{n}{k}^2 k! \le n^{2k}/k!$ and $(N)^{\overline{k}} \ge N^k$. Summing the exponential series gives $S \le e^x - 1$ and proves all upper bounds in (12).

5.2 A uniform two-point lower bound

It suffices to use the scalar family (22), since it is included in (10). Compare $t_0 = a$ with $t_1 = a(1+h)$. Direct computation gives

$$\mathrm{KL}(p_{t_0}\|p_{t_1}) = (1+a)\log\left(1 + \frac{ah}{1+a}\right) - a\log(1+h). \tag{25}$$

Its derivative in h is

$$\frac{ah}{(1+h)(1+a+ah)} \le \frac{h}{\gamma},$$

so the one-copy divergence is at most $h^2/(2\gamma)$ for every $h \ge 0$. Set

$$h = \frac{1}{4}\min\left(\frac{1}{n}, \sqrt{\frac{\gamma}{N}}\right).$$

The N-copy divergence is at most $1/32$, and Pinsker's inequality bounds the total variation distance by $1/8$. The target ratio is $r = (1+h)^n$. Its relative separation

$$\delta = \frac{r-1}{r+1}$$

satisfies $\delta \ge nh/3$, because $nh \le 1/4$, $r \le e^{1/4}$ and $r - 1 \ge nh$.

Threshold any estimator at the harmonic target midpoint $2F_0F_1/(F_0+F_1)$. Each classification error entails relative squared loss at least δ^2. The sum of the testing errors is at least one minus total variation, so the maximum of the two risks is at least $\delta^2(1-1/8)/2$. Since $n^2h^2 = \min(1,x)/16$, this proves the lower bound $7\min(1,x)/2304$. A quantum POVM cannot improve the distance of the commuting family. Theorem 2.3 follows.

5.3 Consistency and complete collapse

Equation (13) follows immediately from the upper and lower bounds in (12). To prove (14), take along the given sequence

$$h = \frac{x^{1/4}}{n}, \qquad t_0 = a, \quad t_1 = a(1+h).$$

The divergence bound above gives $N\,\mathrm{KL} \le 1/(2\sqrt{x})$, hence the total variation η_x tends to zero. Meanwhile $r = (1+h)^n \ge 1 + x^{1/4}$ tends to infinity.

Let P_0, P_1 be the full photon-count laws and let their common subprobability measure have density $\min(dP_0, dP_1)$; its mass is $1 - \eta_x$. For every real y, direct minimization gives

$$\frac{1}{1+r}\left(\frac{y}{F_0} - 1\right)^2 + \frac{r}{1+r}\left(\frac{y}{F_1} - 1\right)^2 \ge \left(\frac{r-1}{r+1}\right)^2. \tag{26}$$

The minimizing y is the harmonic midpoint. Integrating (26) over the common measure bounds every estimator's maximum risk from below by

$$(1 - \eta_x)\left(\frac{r-1}{r+1}\right)^2 \longrightarrow 1.$$

Classical randomization, and therefore arbitrary quantum measurement on this scalar subfamily, preserves this lower bound. The identically zero decision has risk one, proving the collapse assertion of Theorem 2.4.

5.4 The exact critical limit

It remains to prove (15). The upper limit follows directly from the exact sum: for fixed k its summand tends to $\lambda^k/k!$, and the summands are dominated by $C^k/k!$ when $\gamma n^2/N \le C$. Hence $S \to e^\lambda - 1$ and $\limsup \mathcal{R} \le 1 - e^{-\lambda}$.

For the lower limit use the scalar experiment above and omit the target's known factor $n!$. The posterior-moment method follows the classical power-estimation argument in [10]. Here a moving truncated prior ensures that every parameter respects the occupation floor. Set

$$q_a = \frac{a}{1+a}, \quad \delta_n = n^{-1/4}, \quad m = n^{3/2}, \quad c_n = q_a + \delta_n, \quad \alpha = mc_n, \quad \beta = m(1 - c_n).$$

Write $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ for the beta function. For large n, $c_n < 1$ and $\alpha > 2n$. Give $q = t/(1+t)$ the $\mathrm{Beta}(\alpha,\beta)$ law conditioned on $E = \{q \ge q_a\}$. The posterior given K is $\mathrm{Beta}(u,v)$ conditioned on E, with $u = \alpha + K$, $v = \beta + N$. Write

$$p_j = \Pr_{\mathrm{Beta}(u-jn, v+jn)}(E), \qquad j = 0, 1, 2.$$

For $F = t^n$, the negative posterior moments are

$$\mathbb{E}[F^{-j} \mid K] = \frac{B(u-jn, v+jn)p_j}{B(u,v)p_0}.$$

Minimizing the quadratic posterior loss gives conditional risk $1 - r_K$, where

$$r_K = \frac{B(u-n, v+n)^2}{B(u,v)B(u-2n, v+2n)}\frac{p_1^2}{p_0 p_2}, \qquad 0 \le r_K \le 1. \tag{27}$$

The last bound is Cauchy–Schwarz for the two negative moments. All moments exist, since $\alpha > 2n$.

We verify that the truncation factors disappear. The probability limits here refer to the joint law of the supported prior and the photon counts. The unconditioned prior has variance $O(m^{-1})$, so $\Pr(E) \to 1$ and $q - c_n = o_P(\delta_n)$ also after conditioning. Put $t_c = c_n/(1 - c_n)$. Conditional photon variance gives

$$\mathbb{E}[K/N \mid q] = t, \qquad \mathrm{Var}(K/N \mid q) = t(1+t)/N.$$

Restricting first to a fixed neighborhood of q_a and using Chebyshev there yields $K/N = t_c + o_P(\delta_n)$, hence $K/N \to a$. As $m/N = O(n^{-1/2}) = o(\delta_n)$ and $n/N = O(n^{-1})$, all three tilted posterior means satisfy

$$\frac{u - jn}{u + v} = c_n + o_P(\delta_n).$$

Their variances are $O_P(N^{-1})$. Their distance above q_a is $\delta_n(1 + o_P(1))$, and $N\delta_n^2 \to \infty$. Chebyshev therefore gives $p_j \to 1$ in probability for every j.

To evaluate the beta ratio put $L(s) = \log B(u - s, v + s)$. The total shape is constant, and

$$L''(s) = \psi_1(u - s) + \psi_1(v + s), \qquad 2L(n) - L(0) - L(2n) = -\int_0^n \int_0^n L''(r + s)\, dr\, ds,$$

where $\psi_1(x) = (\log \Gamma)''(x)$ is the trigamma function. Since $u/N \to a$, $v/N \to 1$ and $n/N \to 0$, its series estimate $\psi_1(x) = x^{-1} + O(x^{-2})$ gives, uniformly in the double integral,

$$2L(n) - L(0) - L(2n) = -n^2(1/u + 1/v) + o_P(1) \longrightarrow -\lambda.$$

Thus $r_K \to e^{-\lambda}$ in probability. The bound in (27) gives convergence of expectations, so this supported prior has Bayes risk tending to $1 - e^{-\lambda}$. Every scalar procedure, and hence the full quantum minimax problem, has at least this Bayes risk. Together with the upper limit this proves (15).

6 An exact heterodyne benchmark

Standard heterodyne measurements give an amplitude matrix Z whose columns are independent $\mathrm{CN}(0, A + I_d)$ variables. Define $D_k = D_A^{\otimes k}$ and the polynomial coherent symbol

$$P_f(Z) = \frac{n!}{(N)^{\overline{n}}} \sum_{\boldsymbol{r} \in [N]^n} |C_{\boldsymbol{r}}(Z)|^2, \qquad \Delta = \sum_{i,r} \partial_{z_{ir}} \partial_{\overline{z}_{ir}}, \qquad h_f = e^{-\Delta} P_f.$$

Here $C_{\boldsymbol{r}}(Z)$ is obtained from (5) by replacing $b_{i,r}$ with z_{ir}. The finite exponential removes the known unit covariance added by heterodyne detection. It is the usual Gaussian noise correction applied to this collective symbol.

Define $\mathcal{R}_H(N, n, a)$ by (10), with the infimum restricted to measurable real decisions based on these full heterodyne records, including independent randomization. Define $\mathcal{R}_{H,\mathrm{ub}}$ by imposing unbiasedness at every allowed covariance.

Theorem 6.1 (Heterodyne variance and optimality)**.** *The real statistic h_f is unbiased and has exact relative variance*

$$\frac{\mathrm{Var}_A(h_f)}{F_f(A)^2} = \sum_{k=1}^{n} \binom{n}{k}^2 \frac{k!}{(N)^{\overline{k}}} \mathrm{Tr}(R_k D_k R_k D_k). \tag{28}$$

For each f, A this is the minimum variance among all globally unbiased decisions based on the full heterodyne records. Its sharp uniform envelope is

$$S_H(N,n,a) = \sum_{k=1}^{n} \binom{n}{k}^2 \frac{k!\gamma^{2k}}{(N)^{\overline{k}}}.$$

The associated worst-polynomial arbitrary-bias minimax risk $\mathcal{R}_H$ satisfies

$$\frac{7}{2304}\min(1, x_H) \le \mathcal{R}_H \le \frac{S_H}{1+S_H} \le 1 - e^{-x_H}, \qquad x_H = \frac{\gamma^2 n^2}{N}. \tag{29}$$

It tends to zero exactly when $x_H \to 0$, and tends to one when $x_H \to \infty$.

Proof. Gaussian convolution gives $\mathbb{E}_{A+I} e^{-\Delta} P_f = \mathbb{E}_A P_f = F_f(A)$. The heat operation is finite because P_f has bidegree (n,n). Both P_f and h_f are invariant under $Z \mapsto ZU$, so they are functions of ZZ^*. The scatter matrix is sufficient for $A+I$ and complete by Lemma 4.1. Conditional Jensen and completeness prove the asserted minimum variance, including arbitrary independent randomization.

For the second moment use the sphere representation $P_f = \mathbb{E}_v |f(Zv)|^2$. For fixed v, w put $X = Zv$, $Y = Zw$ and $V = A + I$. On one factor the inverse heat is represented by the formal generating function

$$G_X(\xi, \eta) = \exp(\xi \cdot X + \eta \cdot \overline{X} - \xi \cdot \eta).$$

Use independent formal vectors ξ, η, ζ, ω for the two factors, and put $s = \sum_r v_r \overline{w_r}$. Gaussian expectation gives

$$\begin{aligned} &\mathbb{E}[G_X(\xi,\eta) G_Y(\zeta,\omega)] \\ &\quad = \exp\Big(\xi^{\mathsf{T}} A \eta + \zeta^{\mathsf{T}} A \omega + s \xi^{\mathsf{T}} V \omega + \overline{s} \zeta^{\mathsf{T}} V \eta\Big). \end{aligned}$$

Thus the internal contractions are A, while each cross direction retains V. Differentiating the four independent groups extracts the polynomial coefficients. The kth cross class has the same multiplicity (19). Whitening its internal A contractions leaves a factor D_A in *each* cross direction. Its normalized coefficient sum is therefore

$$\sum_{I,J} |(R_k)_{I,J}|^2 (D_k)_{I,I} (D_k)_{J,J} = \operatorname{Tr}(R_k D_k R_k D_k).$$

The sphere overlap again averages to $k!/(N)^{\overline{k}}$, proving (28). This finite formal calculation is also a specialization of classical complex Wishart/Wick moment machinery [2, 7, 5].

Since $D_k \preceq \gamma^k I$, each contraction is at most $\gamma^{2k} \operatorname{Tr} R_k^2 \le \gamma^{2k}$. The scalar monomial at $A = a$ attains equality. Shrinking h_f by $1 + S_H$ gives the upper bound in (29).

For the lower bound the scalar heterodyne law is $\mathrm{CN}(0, t+1)$. At $t_0 = a$, $t_1 = a(1+h)$ its divergence is $\log(1+u) - u/(1+u) \le u^2/2$, where $u = h/\gamma$. Repeating the harmonic-midpoint test with $h = \frac{1}{4}\min(n^{-1}, \gamma/\sqrt{N})$ gives (29). For collapse take $h = x_H^{1/4}/n$. Then $N \,\mathrm{KL} \le 1/(2\sqrt{x_H})$ while the target ratio is at least $1 + x_H^{1/4}$. Integrating (26) over the common part of the two laws proves convergence to one. The zero decision supplies the matching upper bound. □

Pointwise comparison for the same energy. Let V_Q, V_H be the minimum unbiased relative variances in (7) and (28). Put $m_A = 1+\lambda_{\max}(A)^{-1}$ and $M_A = 1+\lambda_{\min}(A)^{-1}$. For each contraction order,

$$m_A^k \operatorname{Tr}(R_k^2 D_k) \le \operatorname{Tr}(R_k D_k R_k D_k) \le M_A^k \operatorname{Tr}(R_k^2 D_k).$$

Indeed, if d_I are the diagonal entries of D_k, symmetry of $|(R_k)_{I,J}|^2$ writes the quantum contraction as $\frac{1}{2}\sum_{I,J}|(R_k)_{I,J}|^2(d_I+d_J)$. The heterodyne contraction replaces $(d_I+d_J)/2$ by $d_I d_J$, whose ratio lies between m_A^k and M_A^k. In particular $V_H \ge m_A V_Q > V_Q$. This compares pointwise optimal *globally unbiased* procedures for the same f, A.

Corollary 6.2 (Matched measurement resources)**.** *Along any sequence with $\gamma \to \infty$ and*

$$\gamma n^2 \ll N \ll \gamma^2 n^2,$$

the quantum minimax risk tends to zero and the all-heterodyne minimax risk tends to one.

This compares arbitrary collective POVMs with standard heterodyne detection followed by unrestricted classical processing. The low-brightness photon-count advantage at degree one is established in [9]; the formulas above retain every degree and the full coefficient geometry.

7 Worked examples and resource consequences

7.1 Photon counts and fewer copies than the degree

For $f(z) = z^n$ in one mode, the target is $n!t^n$ and the total photon count K over N copies has the negative-binomial law used above. The estimator is the falling-factorial statistic (23). At degree two it becomes

$$T_f = \frac{2K(K-1)}{N(N+1)}, \qquad F_f(t) = 2t^2,$$
$$\frac{\operatorname{Var}_t(T_f)}{F_f(t)^2} = \frac{4\gamma_t}{N} + \frac{2\gamma_t^2}{N(N+1)}, \qquad \gamma_t = 1 + t^{-1}.$$

For $N = 1$ this is the single-copy observable $K(K-1)$. Its expectation already gives the degree-two target because the thermal distribution relates all factorial moments to the same occupation parameter t. Thus $N < n$ is compatible with unbiasedness in this restricted state family. Multiple copies reduce its fluctuations, with both displayed orders contributing to the exact variance. At degree one the sample mean photon count has relative variance γ_t/N.

7.2 Pooling versus direct moment averaging

For the same scalar target $F_f(t) = n!t^n$, separately measuring each copy's factorial moment and averaging gives $\overline{H} = N^{-1}\sum_{r=1}^{N}(K_r)_{\underline{n}}$. At $t = a$, independence and Theorem 2.1 with one copy give

$$\frac{\operatorname{Var}_a(\overline{H})}{F_f(a)^2} = \frac{S(1,n,a)}{N} \ge \frac{\gamma^n}{N}, \qquad \frac{\operatorname{Var}_a(T_f)}{F_f(a)^2} = S(N,n,a) \le e^{\gamma n^2/N} - 1.$$

The lower bound retains only the $k = n$ term. With fixed a and $N = n^3$, the pooled relative variance tends to zero while the direct average's variance diverges exponentially. For example, at $a = 1$, $n = 20$ and $N = 8000$, the pooled variance is below 0.106, whereas the direct average's variance is at least 131.072. Both scalar statistics are computed from separate photon counts: the gain comes from using the Gaussian model to pool the data through their total K. The formulas compare the same target, state family and copy budget, while the multimode result additionally describes an optimal collective observable.

7.3 A two-mode product and the role of coefficient geometry

Take $f(z) = z_1 z_2$ and $A = tI_2$. The target is $F_f(tI_2) = t^2$. In the tensor convention (2), the only nonzero coefficients are $c_{12} = c_{21} = 1/2$, so

$$\psi = \frac{|1,2\rangle + |2,1\rangle}{\sqrt{2}}, \qquad R_1 = \frac{I_2}{2}, \qquad \operatorname{Tr} R_1^2 = \frac{1}{2}, \qquad \operatorname{Tr} R_2^2 = 1.$$

Substitution in (7) gives

$$\frac{\operatorname{Var}_{tI_2}(T_f)}{F_f(tI_2)^2} = \frac{2\gamma_t}{N} + \frac{2\gamma_t^2}{N(N+1)}. \tag{30}$$

At the same occupation, the single-mode monomial z^2 has leading term $4\gamma_t/N$. The product energy has half that leading variance, while its second-order term is unchanged. For $t = 1$ the exact relative variances are

N	1	2	3
$z_1 z_2$	8	10/3	2
z^2	12	16/3	10/3

By Theorem 2.2, these values are the optimal globally unbiased variances at the stated isotropic covariance. The reduced matrices belong to the polynomial coefficient tensor; their purities encode how its indices are distributed across modes. This is the observable-specific information retained by the exact formula and lost on replacing every purity by its worst-case value one. The uniform minimax result instead takes the supremum over polynomials and allowed covariances.

7.4 Sufficient copy budgets and growing degree

For a desired relative squared-error bound $0 < r < 1$, Theorem 2.3 gives the sufficient condition

$$N \geq \frac{(1 + a^{-1})n^2}{-\log(1 - r)}. \tag{31}$$

Indeed this makes $x \leq -\log(1 - r)$, and the observable $T_f/(1 + S)$ has risk at most $1 - e^{-x} \leq r$. Evaluating the finite sum S can give a tighter sufficient budget. Formula (31) is a conservative design guarantee under the stated collective measurement model.

For example, at degree $n = 10$ and floor $a = 1$, choosing $N = 2000$ gives $x = 0.1$ and a uniform relative mean-square bound $1 - e^{-0.1} < 0.096$. At floor $a = 0.1$, $N = 11000$ gives the same bound. The comparison displays the occupation cost while keeping degree, loss and measurement class fixed. The rate lower bound and Theorem 2.4 establish the corresponding order and vanishing-risk criterion for the hardest polynomial.

For fixed a and growing degree, Theorem 2.4 gives the exact limiting risk $1 - e^{-\lambda}$ at $\gamma n^2/N \to \lambda$. Thus retaining all contraction orders determines the crossover risk itself, in addition to the copy scale for consistency.

7.5 Coefficient geometry and the measurement comparison

At $A = tI_d$, put $g = 1 + t^{-1}$. Each kth heterodyne variance term is g^k times its quantum counterpart. At degree one, $f(z) = z_1 + iz_2$ and $A = \operatorname{diag}(1, 4)$ give $F_f = 5$ and relative variances $7/(5N)$ and $49/(25N)$. This anisotropic example also illustrates why the two-sided contraction in (28) matters.

For a higher-degree example take $d = n$, $f(z) = \prod_{i=1}^{n} z_i$ and $A = tI_n$. Across k versus $n-k$ tensor slots, the coefficient vector has a uniform Schmidt decomposition over the $\binom{n}{k}$ mode subsets. Thus $\operatorname{Tr} R_k^2 = \binom{n}{k}^{-1}$ and

$$V_Q = \sum_{k=1}^{n} \frac{(n)_{\underline{k}} g^k}{(N)^{\overline{k}}}, \qquad V_Q + 1 = (N-1) \int_0^1 (1-u)^{N-2} (1+gu)^n \, du \quad (N > 1). \tag{32}$$

The integral follows by expanding $(1+gu)^n$ and using the beta integral. For $N = 1$, $V_Q = (1+g)^n - 1$. Theorem 2.2 makes these exact optimal unbiased variances under the full covariance model. The finite sum also gives a direct resource bound:

$$V_Q \le \sum_{k=1}^{\infty} (gn/N)^k = \frac{gn}{N - gn} \qquad (N > gn),$$

since $(n)_{\underline{k}} \le n^k$ and $(N)^{\overline{k}} \ge N^k$. Thus at this isotropic covariance, $N/n \to \infty$ suffices for unbiased relative consistency, while the worst-polynomial uniform scale is quadratic in n.

For fixed t, the substitution $u = v/n$ and dominated convergence give $V_Q \to g/(c-g)$ when $N/n \to c > g$. At $N = gn + O(1)$, substitute $u = v/\sqrt{n}$: the exponent tends to $-g(1+g)v^2/2$, giving

$$\frac{V_Q}{\sqrt{n}} \longrightarrow \sqrt{\frac{\pi g}{2(1+g)}}.$$

The domination follows from the strict maximum at zero of $g\log(1-u) + \log(1+gu)$, with quadratic decay near zero. These formulas exhibit how a product polynomial's coefficient geometry changes the optimal unbiased copy scale relative to the monomial envelope. Heterodyne has the same finite sum with g replaced by g^2.

Computational verification. The accompanying exact-arithmetic programs check the quantum contraction formula against an independent normal-order expansion, the product example (30), and the scalar identity (24). A separate polynomial expansion and inverse-heat calculation checks (28) in complex anisotropic examples. These finite checks support the normalizations; the analytic proofs give the general results.

Acknowledgment. OpenAI Codex was used for mathematical, computational and editorial assistance in preparing this work.

References

[1] A. Dasgupta, N. A. Warsi and P. Chatterjee, *Uniqueness, Cramér–Rao efficiency and concentration bounds for quantum U-statistics*, arXiv:2609.08745v2 (2026). `https://arxiv.org/abs/2609.08745v2`.

[2] P. Graczyk, G. Letac and H. Massam, *The complex Wishart distribution and the symmetric group*, Ann. Statist. **31** (2003), 287–309. `https://doi.org/10.1214/aos/1046294466`.

[3] M. Guţă and C. Butucea, *Quantum U-statistics*, J. Math. Phys. **51** (2010), 102202. `https://doi.org/10.1063/1.3476776`.

[4] S. A. Khan, S. Prabhu, L. G. Wright and P. L. McMahon, *Quantum computational sensing using quantum signal processing, quantum neural networks, and Hamiltonian engineering*, npj Quantum Inf. **12** (2026), 119. https://doi.org/10.1038/s41534-026-01235-w.

[5] S. Kuriki and Y. Numata, *Graph presentations for moments of noncentral Wishart distributions and their applications*, Ann. Inst. Statist. Math. **62** (2010), 645–672. https://doi.org/10.1007/s10463-010-0279-4.

[6] S. Lahiry, *Classical Sufficiency in Quantum Statistical Experiments*, arXiv:2609.29348v1 (2026). https://arxiv.org/abs/2609.29348v1.

[7] G. Letac and H. Massam, *All invariant moments of the Wishart distribution*, Scand. J. Statist. **31** (2004), 295–318. https://doi.org/10.1111/j.1467-9469.2004.01-043.x.

[8] A. Łuczak, *On a general concept of sufficiency in von Neumann algebras*, Probab. Math. Statist. **35** (2015), 313–324. https://www.math.uni.wroc.pl/~pms/files/35.2/Article/35.2.9.pdf.

[9] R. Nair and M. Tsang, *Quantum optimality of photon counting for temperature measurement of thermal astronomical sources*, Astrophys. J. **808** (2015), 125. https://doi.org/10.1088/0004-637X/808/2/125.

[10] S. Schubr, *Exact minimax estimation of homogeneous complex Gaussian energies*, author-published edition 2, Luminary Oddities (2026). https://research.luminaryoddities.com/exact-minimax-homogeneous-complex-gaussian-energies/v/2.

[11] M. Tsang, F. Albarelli and A. Datta, *Quantum semiparametric estimation*, Phys. Rev. X **10** (2020), 031023. https://doi.org/10.1103/PhysRevX.10.031023.

[12] M. Tsang, *Operational meanings of a generalized conditional expectation in quantum metrology*, Quantum **7** (2023), 1162. https://doi.org/10.22331/q-2023-11-03-1162.