Exact minimax estimation of homogeneous complex Gaussian energies

Scott Schubr
Luminary Oddities

October 26, 2026

Abstract

Let f be a fixed nonzero homogeneous holomorphic polynomial of degree n on $\mathbb{C}^d$. We estimate its Gaussian energy $F_f(A) = \mathbb{E}|f(X)|^2$ from N independent centered proper complex Gaussian observations with unknown positive semidefinite covariance A. Singular covariances are included and $F_f(A) > 0$. Writing $W = \sum_{s=1}^{N} X_s X_s^*$, $a = (N)^{\overline{n}}$ and $b = (N)^{\overline{2n}}$, we prove that the exact minimax risk under relative squared error is $1 - a^2/b$, attained by the estimator $(a/b)F_f(W)$. The unbiased estimator $F_f(W)/a$ has exact worst-case relative variance $b/a^2 - 1$, with equality at rank-one covariances. A classical complex Wishart moment operator supplies the uniform upper bound, while a scalar scale subexperiment and proper inverse-Gamma priors give the matching lower bound against all estimators. Permanent estimation is a special case. Uniform relative-error estimation has observation complexity of order $n^2\varepsilon^{-2}\log(1/\delta)$, with computation unrestricted.

1 Introduction and main result

High-order Gaussian moments encode useful positive functionals of a covariance matrix. Directly averaging $|f(X)|^2$ is an unbiased method, but its relative variance can grow rapidly with polynomial degree. The complete sample covariance retains additional cross-sample products. For a homogeneous holomorphic energy its Wishart normalization yields a dimension-free variance bound. Optimizing a scalar shrinkage of that statistic gives an exact decision-theoretic answer, uniformly over the unknown covariance shape.

We distinguish the degree n from the ambient dimension d. For $A = A^* \succeq 0$, the convention $X \sim \mathcal{CN}(0, A)$ means $X = A^{1/2}g$, where the coordinates of g are independent standard proper complex Gaussians with $\mathbb{E}|g_j|^2 = 1$. Thus $\mathbb{E}XX^* = A$ and $\mathbb{E}XX^{\mathsf{T}} = 0$. The zero mean and proper Gaussian model are known. Let

$$F_f(A) = \mathbb{E}_{X\sim\mathcal{CN}(0,A)}|f(X)|^2, \qquad \mathcal{A}_f = \{A = A^* \succeq 0 : F_f(A) > 0\}. \tag{1}$$

The set $\mathcal{A}_f$ includes singular matrices. For observations $X_1, \dots, X_N$, write

$$W = \sum_{s=1}^{N} X_s X_s^*, \qquad a = (N)^{\overline{n}} = \prod_{j=0}^{n-1}(N + j), \qquad b = (N)^{\overline{2n}}. \tag{2}$$

For a real-valued measurable decision D based on the entire sample, and possibly on independent randomization, its risk is

$$R_A(D) = \mathbb{E}_A\left(\frac{D}{F_f(A)} - 1\right)^2.$$

There is no unbiasedness or computational restriction on the decision class. Nonnegative decisions suffice: replacing a negative answer by zero cannot increase this loss.

Theorem 1.1 (Exact minimax energy risk)**.** *Let $d, n, N \geq 1$ and let f be a fixed nonzero homogeneous holomorphic polynomial of degree n on $\mathbb{C}^d$. Then*

$$\inf_D \sup_{A \in \mathcal{A}_f} R_A(D) = 1 - \frac{a^2}{b}. \tag{3}$$

The minimax value is attained by

$$D_* = \frac{a}{b} F_f(W). \tag{4}$$

The statistic $U = F_f(W)/a$ is unbiased and has exact worst-case relative variance

$$\sup_{A \in \mathcal{A}_f} \frac{\mathrm{Var}_A(U)}{F_f(A)^2} = \frac{b}{a^2} - 1 \leq \exp(n^2/N) - 1. \tag{5}$$

Both suprema are attained at every $A = \theta uu^*$ with $\theta > 0$ and $f(u) \neq 0$. The conclusions remain true if positive diagonal entries are additionally required.

The theorem identifies a uniform finite-sample risk, rather than an asymptotic rate or a restriction to unbiased estimators. Its proof has two distinct parts. A symmetric-tensor consequence of the classical complex Wishart moment formula gives the upper bound for every covariance. A rank-one subexperiment has a single unknown scale; classical scale decision theory then gives the matching lower bound. Rank-one equality in a moment inequality alone would not prove minimaxity against arbitrary decisions.

Two statistical objectives. The exact-risk decision shrinks the scatter-matrix energy to minimize worst relative squared loss. The unbiased decision is used in independent batches for a prescribed relative-error probability. The common risk formula is universal across the supplied polynomial, but the decision still evaluates that polynomial's energy. For degree one the exact risk is $1/(N+1)$. Throughout W is the unnormalized scatter matrix; replacing it by W/N requires its corresponding degree normalization.

1.1 Relation to earlier work

The complex Wishart permutation-cycle moment operator is classical: Graczyk, Letac and Massam [1] give its symmetric-group formulation; their Corollary 6, equation (4.1), supplies the central coordinate moment formula. Kuriki and Numata [2] develop central and noncentral moment formulas by graph and Wick counting methods. Our proof includes the required classical central identity and its normalization.

Di Nardo [3] treats permanent representations and inherited unbiased cumulant estimators for Wishart matrices. More generally, inversion of the Wishart expectation operator and estimation of covariance functionals are established themes; see Koltchinskii [4]. Hillier and Kan [5] study expectations of equivariant matrix-valued Wishart and inverse-Wishart functions. These works provide important moment and estimation context. Here the target is a fixed, not necessarily rotation-invariant, holomorphic energy and the criterion is the exact uniform finite-sample relative squared risk.

Scalar power estimation under scale-invariant loss is also classical; see, for example, the scale-estimation framework of Stępień-Baran [6]. The proper-prior argument below is supplied in full. The

contribution studied here is the transfer from the Wishart operator to a uniform holomorphic-energy risk bound, its rank-one extremum, and the resulting exact minimax decision for the entire covariance class. The moment formulas, unbiased normalization, Gamma conjugacy and scalar scale mechanism are inherited ingredients.

Gaussian representations also underlie algorithms for approximating the permanent of a supplied matrix. Mukerji and Yang [7] obtain additive estimates from designed Gaussian fields. Here the covariance is unknown and the criterion is relative statistical risk, with computation unrestricted.

The deterministic permanental-cofactor bounds in the author's earlier work [8] concern a different matrix-norm question. They supply related permanent structure.

2 A uniform Wishart energy bound

Let $H = \mathbb{C}^d$ with its usual Hermitian inner product, conjugate-linear in the first argument. For $k \geq 1$, P_σ denotes a unitary permutation of the factors of $H^{\otimes k}$, and

$$\Pi_k = \frac{1}{k!} \sum_{\sigma \in S_k} P_\sigma$$

is the orthogonal projection onto $\mathrm{Sym}^k H$. Every degree-n homogeneous holomorphic polynomial has a unique coefficient tensor $v \in \mathrm{Sym}^n H$ satisfying

$$f(z) = \langle v, z^{\otimes n} \rangle.$$

Lemma 2.1 (Gaussian energy representation)**.** *For every Hermitian positive semidefinite B,*

$$F_f(B) = n! \langle v, B^{\otimes n} v \rangle. \tag{6}$$

In particular, F_f is a nonnegative homogeneous polynomial of degree n in B.

Proof. The proper complex Gaussian Wick formula pairs each of the n unbarred coordinates with one of the n barred coordinates. Consequently

$$\mathbb{E}[z^{\otimes n} (z^{\otimes n})^*] = B^{\otimes n} \sum_{\sigma \in S_n} P_\sigma = n! B^{\otimes n} \Pi_n.$$

Taking the quadratic form at v and using $\Pi_n v = v$ proves the identity. The representation $z = B^{1/2} g$ proves it also for singular B. □

Lemma 2.2 (Classical complex Wishart tensor moment)**.** *For every integer $k \geq 1$,*

$$\mathbb{E}_A W^{\otimes k} = A^{\otimes k} \sum_{\sigma \in S_k} N^{c(\sigma)} P_\sigma, \qquad \sum_{\sigma \in S_k} N^{c(\sigma)} = (N)^{\overline{k}}, \tag{7}$$

where $c(\sigma)$ is the number of cycles of σ.

Proof. For matrix entries indexed by $(i_1, j_1), \ldots, (i_k, j_k)$, expand each $W_{i_r j_r}$ as $\sum_{s_r} X_{s_r, i_r} \overline{X_{s_r, j_r}}$. Proper complex Wick pairings are indexed by permutations $\sigma \in S_k$. A pairing contributes a covariance contraction and requires the sample labels to agree along every cycle of σ. There are $N^{c(\sigma)}$ choices. Thus the coordinate formula is

$$\mathbb{E} \prod_{r=1}^{k} W_{i_r j_r} = \sum_{\sigma \in S_k} N^{c(\sigma)} \prod_{r=1}^{k} A_{i_r j_{\sigma(r)}}.$$

This is the tensor identity with the corresponding permutation convention; inverse permutations give the same operator sum. This finite argument uses $X_s = A^{1/2} g_s$ and needs no inverse of A.

For the weight sum, insert the final element into a permutation of $k-1$ elements. Creating a singleton cycle contributes a factor N, while insertion after one of the existing $k-1$ elements preserves the number of cycles. Starting from the empty permutation of weight one gives $N(N+1)\cdots(N+k-1)$. □

Proposition 2.3 (Sharp uniform first and second moments)**.** *For all $A \in \mathcal{A}_f$,*

$$\mathbb{E}_A F_f(W) = a F_f(A), \qquad \mathbb{E}_A F_f(W)^2 \le b F_f(A)^2. \tag{8}$$

Equality in the second inequality holds at rank one whenever f is nonzero on that range.

Proof. Apply Lemma 2.2 with $k = n$ to (6). Since $P_\sigma v = v$,

$$\mathbb{E}_A F_f(W) = n! \sum_{\sigma \in S_n} N^{c(\sigma)} \langle v, A^{\otimes n} v \rangle = a F_f(A).$$

For the second moment put $w = (A^{1/2})^{\otimes n} v$ and $t = w \otimes w$. All tensor powers of the same matrix commute with permutations. The case $k = 2n$ therefore gives

$$\mathbb{E}_A F_f(W)^2 = (n!)^2 \sum_{\sigma \in S_{2n}} N^{c(\sigma)} \langle t, P_\sigma t \rangle. \tag{9}$$

The left side is real. For each summand,

$$\mathrm{Re}\langle t, P_\sigma t \rangle \le |\langle t, P_\sigma t \rangle| \le \|t\|^2 = \|w\|^4.$$

The cycle weights are positive. Taking real parts in (9), summing these bounds and using $F_f(A) = n!\|w\|^2$ proves the second assertion. Positivity of the individual contractions is neither assumed nor needed.

If $A = \theta u u^*$, write $X_s = \sqrt{\theta}\, u Z_s$, with independent $Z_s \sim \mathcal{CN}(0,1)$. Then

$$W = AG, \qquad G = \sum_{s=1}^{N} |Z_s|^2 \sim \mathrm{Gamma}(N,1), \qquad F_f(W) = F_f(A) G^n. \tag{10}$$

Since $\mathbb{E}G^n = a$ and $\mathbb{E}G^{2n} = b$, the second-moment bound is equality. Also $F_f(A) = n!\theta^n |f(u)|^2 > 0$ when $f(u) \neq 0$. □

3 Exact minimaxity

3.1 Upper bound and optimal shrinkage

Proposition 2.3 proves that $U = F_f(W)/a$ is unbiased and that its worst relative variance is $b/a^2 - 1$, attained at rank one. To check the exponential envelope, observe

$$\frac{b}{a^2} = \prod_{j=0}^{n-1} \left(1 + \frac{n}{N+j}\right), \qquad \log \frac{b}{a^2} \le \frac{n^2}{N}. \tag{11}$$

For $c = a^2/b$, the proposed decision is $D_* = cU$. Thus

$$R_A(D_*) = c^2 \frac{\mathbb{E}_A U^2}{F_f(A)^2} - 2c + 1 \le c^2 \frac{b}{a^2} - 2c + 1 = 1 - \frac{a^2}{b}.$$

Equality holds at rank one. This proves the uniform upper bound and the exact worst risk of D_*. The following lower bound completes the all-estimator assertion.

3.2 A proper-prior scalar lower bound

Lemma 3.1 (Gamma power minimax risk)**.** *If $T \mid \theta \sim \mathrm{Gamma}(N, \textit{scale } \theta)$, with $\theta > 0$, then estimating θ^n under loss $(D/\theta^n - 1)^2$ has minimax risk $1 - a^2/b$. It is attained by $D = (a/b)T^n$.*

Proof. Take the proper inverse-Gamma prior

$$\pi_{\alpha,\beta}(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{-\alpha-1}e^{-\beta/\theta}, \qquad \alpha, \beta > 0.$$

Its posterior given T is inverse-Gamma with shape $s = N + \alpha$ and scale $t = T + \beta$. In particular,

$$m_j := \mathbb{E}[\theta^{-j} \mid T] = \frac{\Gamma(s+j)}{\Gamma(s)}t^{-j}, \qquad j \geq 0.$$

The posterior expected loss is $D^2 m_{2n} - 2Dm_n + 1$. Its minimum is attained at

$$D_{\alpha,\beta}(T) = \frac{\Gamma(s+n)}{\Gamma(s+2n)}(T+\beta)^n,$$

and has value

$$r_\alpha = 1 - \frac{\Gamma(s+n)^2}{\Gamma(s)\Gamma(s+2n)}.$$

This posterior minimum is constant in T and β, so it is also the proper-prior Bayes risk. All negative moments used here are finite. Integration of the nonnegative loss is justified by Tonelli's theorem.

The supremum frequentist risk of every decision is at least its integrated risk under this proper prior, hence at least r_α. Keeping $\beta > 0$ fixed and letting $\alpha \downarrow 0$ gives the lower bound $1 - a^2/b$ by continuity of Γ. This argument bounds each decision separately and makes no interchange of an infimum with a limit. Finally, $T/\theta \sim \mathrm{Gamma}(N, 1)$ gives

$$\mathbb{E}_\theta\left(\frac{(a/b)T^n}{\theta^n} - 1\right)^2 = \frac{a^2}{b^2}b - 2\frac{a}{b}a + 1 = 1 - \frac{a^2}{b}.$$

Randomization cannot improve a quadratic posterior minimum. □

Completion of Theorem 1.1. Choose a fixed u with $f(u) \neq 0$ and consider the subfamily $A_\theta = \theta uu^*$, $\theta > 0$. The target is $F_f(A_\theta) = c_0\theta^n$, where $c_0 = n!|f(u)|^2$ is known in this subexperiment. The full observations are equivalent to scalar observations $Y_s = \sqrt{\theta}Z_s$, because $Y_s = u^*X_s/\|u\|^2$ and $X_s = uY_s$. Their joint density depends on θ only through

$$T = \sum_{s=1}^{N} |Y_s|^2 \sim \mathrm{Gamma}(N, \mathrm{scale}\ \theta).$$

Conditional on T, the normalized complex sample direction is independent of θ. Conditional averaging therefore cannot increase squared relative risk. Equivalently, the proper-prior posterior calculation of Lemma 3.1 is unchanged when conditioning on the full data: its likelihood contains no additional θ dependence. Dividing any decision by c_0 reduces the target exactly to θ^n. Lemma 3.1 bounds its supremum risk from below by $1 - a^2/b$, and the full covariance class can only have larger minimax risk. The already proved upper bound gives equality.

For the positive-diagonal version, choose u with all coordinates nonzero and $f(u) \neq 0$. Such a vector exists because a nonzero polynomial cannot vanish on the open set $(\mathbb{C} \setminus \{0\})^d$. Then every A_θ in the same lower-bound subexperiment has positive diagonal. The upper bound remains valid on that smaller class. □

4 Confidence guarantees and permanent estimation

4.1 Minimax observation order

For a fixed f, let $\mathcal{N}_f(\varepsilon, \delta)$ be the least sample size for which there is a decision $\widehat{F}$ satisfying

$$\sup_{A \in \mathcal{A}_f} \mathbb{P}_A\{|\widehat{F}/F_f(A) - 1| > \varepsilon\} \le \delta.$$

The computation of a decision is unrestricted.

Theorem 4.1. *For every fixed nonzero homogeneous holomorphic polynomial of degree $n \ge 1$, and for $0 < \varepsilon \le 1/4$, $0 < \delta \le 1/8$,*

$$\mathcal{N}_f(\varepsilon, \delta) = \Theta(n^2 \varepsilon^{-2} \log(1/\delta)),$$

with absolute implied constants, independent of d and f. The same assertion holds with a positive-diagonal parameter restriction.

Proof. For the upper bound take independent batches of size

$$N_0 = \left\lceil \frac{n^2}{\log(1 + \varepsilon^2/8)} \right\rceil$$

and use the unbiased statistic of Theorem 1.1 in each batch. Its relative variance is at most $\varepsilon^2/8$, so Chebyshev's inequality bounds the probability of relative error greater than ε by $1/8$. Return the median of B batch statistics, with B odd. Failure requires at least half the batches to be bad. Hoeffding's inequality gives failure probability at most $\exp(-9B/32)$. Any odd $B \ge (32/9)\log(1/\delta)$ suffices. In the stated ε range, $\log(1 + \varepsilon^2/8)$ is comparable to ε^2, proving the upper order.

For the lower bound use $A_\theta = \theta u u^*$ with $f(u) \ne 0$. The target equals $c_0 \theta^n$. Take $\theta_0 = 1$ and $\theta_1 = (1 + 4\varepsilon)^{1/n}$. The relative-$\varepsilon$ answer intervals are disjoint, since

$$(1 + 4\varepsilon)(1 - \varepsilon) > 1 + \varepsilon \quad (0 < \varepsilon \le 1/4).$$

A successful estimator thus gives a binary test whose two error probabilities are at most δ. Data processing of relative entropy for the test event yields

$$\mathrm{KL}(P_0^N \| P_1^N) \ge (1 - 2\delta) \log \frac{1 - \delta}{\delta}.$$

Here P_j denotes one scalar proper complex Gaussian observation of variance θ_j; the fixed rank-one vector map loses no information. If $s = \log(\theta_1/\theta_0)$, direct integration of the scalar likelihood gives

$$\mathrm{KL}(P_0^N \| P_1^N) = N(e^{-s} - 1 + s) \le \frac{N s^2}{2} \le \frac{8 N \varepsilon^2}{n^2}.$$

Therefore every uniformly successful decision must obey

$$N \ge \frac{n^2}{8\varepsilon^2}(1 - 2\delta) \log \frac{1 - \delta}{\delta}. \tag{12}$$

For $\delta \le 1/8$ this is the matching lower order. Parameter-independent randomization adds no relative entropy. Taking u with all coordinates nonzero proves the positive-diagonal version as well. □

4.2 The permanent as an energy

For $d = n$ and $f(z) = \prod_{j=1}^{n} z_j$, the proper complex Wick formula gives

$$F_f(A) = \sum_{\sigma \in S_n} \prod_{j=1}^{n} A_{j,\sigma(j)} = \operatorname{per} A.$$

Consequently Theorem 1.1 yields the following statement.

Corollary 4.2. *For unknown Hermitian PSD A with $\operatorname{per} A > 0$, the decision $(a/b) \operatorname{per} W$ has exact minimax relative squared risk $1 - a^2/b$. The unbiased statistic $\operatorname{per} W/a$ has exact worst relative variance $b/a^2 - 1$. Both conclusions remain valid if positive diagonal entries are required.*

Positive diagonal is a sufficient condition for $\operatorname{per} A > 0$: in a Gram representation each coordinate linear form is nonzero, so their product is a nonzero polynomial and its squared Gaussian norm is positive. Rank-one matrices with all coordinates nonzero attain the two extrema.

5 Interpretation and reproducibility

For $n = 1$, the minimax risk is $1/(N+1)$, recovering the familiar scale shrinkage. More generally, the exact value lies between zero and one for every $N, n \geq 1$. Formula (11) explains the degree-squared sample scale. The biased optimum in relative squared risk and the batch-unbiased estimate used for confidence amplification address different objectives.

The experiment uses centered proper complex Gaussian observations, an unknown overall scale, and the positive semidefinite covariance class, including singular matrices. Both lower bounds use its rank-one scale subfamily. Relative loss is defined on $F_f(A) > 0$. The functional class consists of Gaussian holomorphic energies, whose rank-one behavior is essential to the exact risk formula. Known-diagonal, strictly positive definite, unknown-mean and nonproper-Gaussian experiments have different parameter or observation structures.

In the permanent application, evaluating $\operatorname{per} W$ can be computationally expensive. The statistical results hold uniformly over all covariances in the stated parameter class.

The accompanying exact moment checker evaluates finite permanent-coordinate Wick expansions exactly with Gaussian-integer arithmetic and rational comparisons. It checks unbiasedness, the uniform second-moment bound, rank-one equality and the corresponding shrinkage-risk bound in 21 cases of dimensions at most three. These diagnostics support normalization and reproducibility; the tensor and decision proofs above establish the general theorems. Instructions are in `README.md`.

Funding and competing interests. This research received no external funding. The author declares no competing interests.

Assistance disclosure. OpenAI ChatGPT and Codex tools assisted mathematical exploration and derivation, literature discovery, manuscript drafting, verification-code development and checking. The author is responsible for the statements, references, code and conclusions.

References

[1] P. Graczyk, G. Letac and H. Massam, The complex Wishart distribution and the symmetric group, *Annals of Statistics* **31** (2003), 287–309. doi:10.1214/aos/1046294466.

[2] S. Kuriki and Y. Numata, Graph presentations for moments of noncentral Wishart distributions and their applications, *Annals of the Institute of Statistical Mathematics* **62** (2010), 645–672. doi:10.1007/s10463-010-0279-4.

[3] E. Di Nardo, On a symbolic representation of non-central Wishart random matrices with applications, *Journal of Multivariate Analysis* **125** (2014), 121–135. arXiv:1312.4395.

[4] V. Koltchinskii, Asymptotically efficient estimation of smooth functionals of covariance operators, *Journal of the European Mathematical Society* **23** (2021), 765–843. doi:10.4171/JEMS/1023.

[5] G. Hillier and R. Kan, On the expectations of equivariant matrix-valued functions of Wishart and inverse Wishart matrices, *Scandinavian Journal of Statistics* **51** (2024), 697–723. doi:10.1111/sjos.12707.

[6] A. Stępień-Baran, Sequential estimation of powers of a scale parameter from delayed observations, *Applicationes Mathematicae* **36** (2009), 13–28. doi:10.4064/am36-1-2.

[7] T. Mukerji and W.-S. Yang, Additive estimates of the permanent using Gaussian fields, preprint, arXiv:2212.10672v2 (2024). `https://arxiv.org/abs/2212.10672v2`.

[8] S. Schubr, Asymptotically optimal bounds for permanental cofactor matrices, preprint, Luminary Oddities, October 25, 2026. `https://research.luminaryoddities.com/asymptotically-optimal-bounds-permanental-cofactor-matrices/v/1`.