The Wishart product inequality and spectral correlation

Caius Folkerts

Abstract. We prove the Wishart product conjecture of Genest, Ouimet and Richards for any number of disjoint principal minors. The inequality holds for arbitrary nonnegative real powers and every real shape parameter in the nonsingular range, with no sign restriction on the covariance. A stronger inequality separates any one factor from the remaining product, and the joint-to-product moment ratio increases with each exponent. Equality in the full product inequality holds precisely when the active diagonal blocks are independent. The underlying spectral monotonicity principle also gives correlation inequalities for overlapping real compressions under explicit rank hypotheses and for arbitrary increasing spectral functions of inverse-Wishart compressions. For complex and quaternion Wishart matrices, determinant products over overlapping compressions satisfy the same inequalities without rank restrictions, including for singular Gaussian Gram laws.

1. Introduction

A principal minor of a Gaussian Gram matrix is the squared volume of the parallelotope spanned by the corresponding vectors. The Wishart product inequality asks whether positive powers of disjoint such volumes have a joint moment at least as large as the product of their individual moments. Genest, Ouimet, and Richards formulated this matrix analogue of the Gaussian product inequality in [2] and proved a quantitative two-minor inequality in [3]. We prove the conjecture for any number of blocks, throughout the nonsingular Wishart range.

Our convention is that $X \sim W_p(\alpha, \Sigma)$ has density proportional to

$$(\det X)^{(\alpha-p-1)/2} \exp\{-\tfrac{1}{2} \operatorname{tr}(\Sigma^{-1} X)\}, \qquad X \in \mathcal{S}_{++}^p,$$

where $\alpha > p-1$ and $\Sigma \succ 0$. Partition $\{1, \dots, p\}$ into nonempty sets $I_1, \dots, I_d$, write $X_{ij} = X_{I_i, I_j}$, and put $\Delta_i = \det X_{ii}$.

Theorem 1.1. *For every $\alpha > p-1$, every $\Sigma \succ 0$, all $\nu_i \geq 0$, and each j,*

$$\mathbb{E}\prod_{i=1}^{d}\Delta_i^{\nu_i} \geq \mathbb{E}\Delta_j^{\nu_j}\,\mathbb{E}\prod_{i\neq j}\Delta_i^{\nu_i} \geq \prod_{i=1}^{d}\mathbb{E}\Delta_i^{\nu_i}. \tag{1}$$

All moments are finite, and equality holds throughout precisely when

$$\Sigma_{I_i,I_j} = 0 \quad \textit{whenever } i \neq j,\ \nu_i > 0,\ \nu_j > 0.$$

The shape parameter need not be integral, so a Gaussian Gram representation alone does not cover the theorem. We start instead from the exact block regression of the Wishart density. After whitening the selected block and fixing its eigenframe, increasing one eigenvalue changes only the length of one Gaussian mean row in the conditional law of the remaining blocks. A radial differential identity makes the expected determinant product nondecreasing under this change, including at nonintegral shape. The selected determinant already depends only on the eigenvalues; averaging the other conditional factor over the independent Haar frame makes it an increasing function of those same coordinates. Association of the ordered eigenvalues then gives (1).

For real Wishart matrices, the radial calculation permits overlapping compressions when the shape is large enough relative to their ranks, and it also applies to increasing convex spectral gauges and convex functions of logarithmic spectra. An explicit two-dimensional example explains why unrestricted splitting of overlapping real compressions fails near the lower shape boundary. For inverse-Wishart matrices the conditional dependence is affine and increasing, so no rank restriction is needed; combining the two laws gives beta-II product inequalities with an exact moment range. Over $\mathbb{C}$ and $\mathbb{H}$, the logarithm of a determinant is subharmonic under one row update. This yields unrestricted overlapping inequalities for a broad spectral class, even for singular integer Gram laws.

2. Regression and spectral association

We first isolate the two probabilistic inputs shared by the proofs. Eigenvalue association is classical [1, 4]; the argument below gives the form and strictness needed here.

Lemma 2.1. *Let a random vector have density proportional to*

$$\prod_{i=1}^{q} w_i(x_i)\prod_{i<j}(x_j - x_i)^{\beta}, \qquad a < x_1 < \cdots < x_q < b,$$

where $\beta > 0$, the weights are positive, and the normalizing integrals are finite. Then increasing nonnegative functions f, g satisfy

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}f(X)\mathbb{E}g(X).$$

If both functions are strictly increasing in the last coordinate, with the lower coordinates fixed, and have finite second moments, the inequality is strict.

Proof. For bounded increasing functions, argue by induction on q. In one dimension, if T' is an independent copy of T, then

$$2\operatorname{Cov}(f(T), g(T)) = \mathbb{E}[(f(T) - f(T'))(g(T) - g(T'))] \geq 0.$$

For the induction step condition on $T = X_q = t$. The lower-coordinate law μ_t has the same form, with weights $w_i(x)(t-x)^\beta$ on (a, t), and is associated by induction.

For $t_1 < t_2$, let $B = \{x_{q-1} < t_1\}$ within the support of μ_{t_2}. On B, the density ratio of $\mu_{t_2}(\cdot \mid B)$ to μ_{t_1} is a positive constant times

$$R(x) = \prod_{i<q} \left(\frac{t_2 - x_i}{t_1 - x_i}\right)^\beta,$$

which is increasing in every coordinate. Association under μ_{t_1} shows that this tilt raises increasing expectations; if R is unbounded, use $R \wedge M$ and monotone convergence. The event B is decreasing, so association under μ_{t_2} shows that conditioning on B lowers increasing expectations. Consequently

$$\begin{aligned}\mathbb{E}_{\mu_{t_1}} f(x, t_1) &\leq \mathbb{E}_{\mu_{t_2}}[f(x, t_1) \mid B] \\ &\leq \mathbb{E}_{\mu_{t_2}}[f(x, t_2) \mid B] \leq \mathbb{E}_{\mu_{t_2}} f(x, t_2).\end{aligned}$$

All evaluations lie in the ordered chamber. Thus the conditional means of f and g increase with t, and the total covariance identity completes the induction. Truncation proves the assertion for nonnegative functions.

If f is strictly increasing in its last coordinate, the middle comparison is strict; the same is true for g. Their conditional means therefore have strictly positive covariance under the positive density of T, while the conditional covariance remains nonnegative. This proves strictness. □

For $S \sim W_q(\alpha, I_q)$ the ordered eigenvalues have density

$$c \prod_{i=1}^{q} \lambda_i^{(\alpha-q-1)/2} e^{-\lambda_i/2} \prod_{i<j} (\lambda_j - \lambda_i), \qquad 0 < \lambda_1 < \cdots < \lambda_q. \tag{2}$$

The eigenframe is independently Haar distributed on $O(q)$. Indeed, in $S = O\,\mathrm{diag}(\lambda)O^T$, the off-diagonal differential coordinates contribute $(\lambda_j - \lambda_i)(O^T\,dO)_{ij}$, while the matrix density depends only on the eigenvalues. Formula (2) satisfies Lemma 2.1, and its exponential tail gives every positive moment.

Lemma 2.2. *Let $p = q + m$ and*

$$\Sigma = \begin{pmatrix} I_q & K \\ K^T & K^T K + C \end{pmatrix}, \qquad C \succ 0.$$

If $X \sim W_p(\alpha, \Sigma)$, then $S = X_{11} \sim W_q(\alpha, I_q)$ and the entire conditional matrix has the representation

$$X \mid S \overset{d}{=} \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix} + Y^T Y, \qquad Y = [S^{1/2},\, S^{1/2}K + G], \tag{3}$$

where $W \sim W_m(\alpha - q, C)$ and the q rows of G are independent $N_m(0, C)$ vectors, independently of S and W.

Proof. Put $U = X_{12}$, $V = X_{22}$ and change variables to

$$Z_0 = S^{-1/2}U, \qquad W = V - U^T S^{-1} U.$$

The domain is $S \succ 0$, $W \succ 0$, $Z_0 \in \mathbb{R}^{q \times m}$, and

$$\det X = \det S \det W, \qquad dX = (\det S)^{m/2}\, dS\, dZ_0\, dW.$$

The Jacobian follows by ordering the variables as (S, Z_0, W); multiplication of each column of Z_0 by $S^{1/2}$ contributes $(\det S)^{m/2}$, and the final translation in W contributes one. Block inversion gives

$$\begin{aligned} \mathrm{tr}(\Sigma^{-1}X) = {} & \mathrm{tr}\, S + \mathrm{tr}(C^{-1}W) \\ & + \mathrm{tr}\{C^{-1}(Z_0 - S^{1/2}K)^T (Z_0 - S^{1/2}K)\}. \end{aligned}$$

The density factors into the asserted laws after setting $G = Z_0 - S^{1/2}K$. Their shapes are admissible because $\alpha - q > m - 1$. □

The way in which (3) converts an eigenvalue change into a mean change will be used repeatedly. With $S = O\,\mathrm{diag}(\lambda)O^T$ fixed, rotate the Gaussian rows by O^T; the rows of Y then have means $\sqrt{\lambda_j}$ times fixed vectors. Conditioning on all but one row leaves a fixed positive semidefinite offset plus a single varying mean row and the residual Wishart matrix.

3. **A radial identity for real Wishart matrices**

For $v, a \in \mathbb{R}^l$ write $H(v,a) = va^T + av^T$. For a twice differentiable function g on symmetric matrices define

$$\mathcal{L}_{N,v} g(V) = 2N Dg(V)[vv^T] + \sum_{a,b=1}^{l} V_{ab} D^2 g(V)[H(v,e_a), H(v,e_b)]. \tag{4}$$

The second term equals $\mathbb{E}_\xi D^2 g(V)[H(v,\xi), H(v,\xi)]$ for $\xi \sim N_l(0,V)$, which displays its covariance under congruence. Its quadratic differential form is

$$\Gamma(g,h) = \sum_{a,b} V_{ab} Dg[H(v,e_a)] Dh[H(v,e_b)].$$

In particular $\Gamma(g,g) \geq 0$ and

$$\mathcal{L}(e^h)/e^h = \mathcal{L}h + \Gamma(h,h). \tag{5}$$

For integer N, this operator has a direct meaning. If $y_1, \ldots, y_N$ are independent centered Gaussian rows with common covariance, then the expectation of $g(\sum_a (y_a + u_a v)(y_a + u_a v)^T)$ depends only on $t = \|u\|^2$. Its Euclidean Laplacian in u is both $4tF''(t) + 2NF'(t)$ and the expectation of (4). The next lemma retains this identity at real shape and allows singular covariance, as required for overlapping compressions.

Lemma 3.1 (Radial mean identity)**.** *Let $C \succeq 0$ have rank $r \geq 1$ in $\mathbb{R}^l$, let $N > r$, and let W be the Wishart matrix of shape $N-1$ and covariance C on its range. Independently take $z \sim N_l(0,C)$ and set*

$$V_t = W + (z + \sqrt{t}\, v)(z + \sqrt{t}\, v)^T.$$

Suppose that g is C^2 on a neighborhood of the positive semidefinite cone and that $g, Dg, D^2 g$ have polynomial growth there. Then $F(t) = \mathbb{E} g(V_t)$ has a finite right derivative at zero and

$$4tF''(t) + 2NF'(t) = \mathbb{E}\mathcal{L}_{N,v} g(V_t), \qquad t > 0. \tag{6}$$

If the right side is nonnegative, F is nondecreasing. The same conclusions hold for every integer $N \geq 1$ when W is the Gram matrix of $N-1$ independent Gaussian rows, with any positive semidefinite covariance.

Proof. The integer case is the Laplacian calculation just given. For a real shape, first take $g_T(V) = e^{-\operatorname{tr}(TV)}$, $T \succeq 0$. Gaussian integration on the range of C gives

$$L_t(T) := \mathbb{E} g_T(V_t) = \det(I + 2CT)^{-N/2} e^{-t a_T}, \qquad a_T = v^T T (I + 2CT)^{-1} v.$$

Put $C_T = C(I + 2TC)^{-1}$ and $w_T = (I + 2CT)^{-1}v$. Completion of the square, or differentiation in T, gives

$$\frac{\mathbb{E}[V_t g_T(V_t)]}{L_t(T)} = NC_T + tw_T w_T^T.$$

On the other hand,

$$\mathcal{L}_{N,v} g_T(V) = g_T(V)\{-2Nv^T Tv + 4v^T TVTv\}.$$

Using $T - 2TC_T T = T(I + 2CT)^{-1}$, its expectation is $L_t(T)(-2Na_T + 4ta_T^2)$, exactly $(4t\partial_t^2 + 2N\partial_t)L_t(T)$.

To pass from exponential tests to the stated class, make the same calculation with $T = -iB$, where B is real symmetric. The determinant power is defined by the analytic logarithm from the Laplace domain, equivalently by the separate eigenvalue factors of $I - 2iC^{1/2}BC^{1/2}$. With $D = C^{1/2}BC^{1/2}$,

$$a_{-iB} = -iv^T Bv + 2v^T BC^{1/2}(I - 2iD)^{-1}C^{1/2}Bv.$$

The resolvent has norm at most one and real part $(I + 4D^2)^{-1} \succeq 0$. Thus $\Re a_{-iB} \geq 0$ and $|a_{-iB}|$ has at most quadratic growth in $\|B\|$. The characteristic transform and its first two t derivatives have polynomial bounds, uniformly on compact t intervals; the operator on an exponential test grows at most quadratically in B and linearly in V_t. Fourier inversion proves (6) for smooth compactly supported functions.

For general g, first multiply by a smooth domain cutoff that equals one near the cone, allowing extension by zero, and then by compact cutoffs whose first two derivatives have uniform polynomial envelopes on the cone. Mollification passes through compact C^2 approximations when needed. All moments of V_t are bounded uniformly on compact t intervals, so dominated convergence passes the operator expectations to the limit. For the parameter derivatives, write

$$J(s) = \mathbb{E}g\{W + (z + sv)(z + sv)^T\}$$

and differentiate twice directly; only $Dg, D^2 g$ and polynomial Gaussian factors occur, so the same domination applies. The function J is even and C^2, whence $F'(0+) = J''(0)/2$ is finite. Finally,

$$(t^{N/2}F'(t))' = \tfrac{1}{4}t^{N/2-1}\mathbb{E}\mathcal{L}_{N,v}g(V_t).$$

The boundary value of $t^{N/2}F'(t)$ is zero, proving monotonicity when the operator expectation is nonnegative. □

3.1. **The determinant calculation**

Let M_i be fixed $l \times k_i$ matrices, and fix $A \succeq 0$, $E_i \succeq 0$, and $\varepsilon > 0$. Put

$$Q_i = E_i + M_i^T(A+V)M_i + \varepsilon I, \qquad \ell_i = \log\det Q_i.$$

For $a_i = M_i Q_i^{-1} M_i^T v$ and $b_i = v^T a_i$, differentiation gives

$$\mathcal{L}_{N,v}\ell_i = 2\{N - \operatorname{tr}(Q_i^{-1}M_i^T V M_i)\}b_i - 2a_i^T V a_i. \tag{7}$$

Indeed, with $\xi \sim N_l(0,V)$ the first derivative along $H(v,\xi)$ is $2a_i^T\xi$, while the second logarithmic derivative is

$$-2\{b_i\xi^T M_i Q_i^{-1} M_i^T \xi + (a_i^T\xi)^2\}.$$

Since $0 \preceq M_i^T V M_i \preceq Q_i$, formula (7) implies

$$\mathcal{L}_{N,v}\ell_i \geq 2\{N - \min(k_i, \operatorname{rank} V) - 1\}b_i. \tag{8}$$

For powers, the chain rule adds the nonnegative square

$$\frac{\mathcal{L}_{N,v}\prod_i(\det Q_i)^{\eta_i}}{\prod_i(\det Q_i)^{\eta_i}} = \sum_i \eta_i \mathcal{L}_{N,v}\ell_i + 4\left(\sum_i \eta_i a_i\right)^T V \left(\sum_i \eta_i a_i\right). \tag{9}$$

Thus products of nonnegative determinant powers have nonnegative generator whenever $N \geq 1 + \max_i \min(k_i, \operatorname{rank} V)$.

Lemma 3.2 (One determinant under a mean shift)**.** *Let $B \succeq 0$, let z be a centered Gaussian vector with arbitrary covariance, and fix v. Then*

$$t \longmapsto \det\{B + (z + \sqrt{t}\,v)(z+\sqrt{t}\,v)^T\}$$

is stochastically nondecreasing for $t \geq 0$. If B and the covariance are positive definite, $v \neq 0$, and $\eta > 0$, the expectation of its ηth power is strictly increasing.

Proof. The adjugate identity gives, without an invertibility assumption,

$$\det(B + ww^T) = \det B + w^T \operatorname{adj}(B) w.$$

Diagonalizing B shows that $\operatorname{adj}(B) \succeq 0$. After applying its positive square root and diagonalizing the Gaussian covariance, the quadratic form becomes

$$\sum_{\lambda_j > 0} \lambda_j (G_j + \sqrt{t}\,c_j)^2 + t\sum_{\lambda_j = 0} d_j^2,$$

with independent standard normal G_j. Each noncentral square is stochastically increasing, since for $x > 0$ and $a \geq 0$,

$$\mathbb{P}\{(G+a)^2 \leq x\} = \Phi(\sqrt{x}-a) + \Phi(\sqrt{x}+a) - 1$$

has nonpositive derivative in a. Independent quantile coupling proves the stochastic comparison. Under the strictness assumptions, at least one c_j is nonzero, its noncentral square increases strictly in stochastic order, and a positive power is strictly increasing. This proves the last assertion. □

Lemma 3.3. *Suppose $C \succ 0$ on $\mathbb{R}^m$ and $N > m$. For every $A \succeq 0$, every partition $J_1, \ldots, J_s$ of $\{1, \ldots, m\}$, and $\eta_i \geq 0$, the function*

$$t \longmapsto \mathbb{E} \prod_{i=1}^{s} \det\{(A+V_t)_{J_i,J_i}\}^{\eta_i}$$

is nondecreasing.

Proof. If $s \geq 2$, each block has size at most $m-1$, so (9) is nonnegative. Apply Lemma 3.1 with an εI regularization in every block and let $\varepsilon \downarrow 0$. The regularized functions and their first two derivatives have polynomial growth for fixed ε, while their values on compact t intervals have a common polynomial bound in $1+\operatorname{tr}(A+V_t)$. Dominated convergence removes the regularization.

If $s = 1$, condition on W and apply Lemma 3.2 to $B = A+W \succ 0$. The comparison is strict for a nonzero mean direction and a positive exponent. □

4. **The product inequality and equality**

Proof of Theorem 1.1. Finiteness follows from $\det X_{ii} \leq (\operatorname{tr} X/|I_i|)^{|I_i|}$ and

$$\mathbb{E}e^{s \operatorname{tr} X} = \det(I - 2s\Sigma)^{-\alpha/2} < \infty, \qquad 0 < s < (2\lambda_{\max}(\Sigma))^{-1}.$$

To separate I_1, write $q = |I_1|$ and $m = p - q$. A block congruence whitens Σ_{11}; it multiplies the first determinant by a positive constant that cancels from the desired inequality. We may therefore use Lemma 2.2.

Define $F(S) = \mathbb{E}[\prod_{i=2}^{d} \Delta_i^{\nu_i} \mid X_{11} = S]$. For a fixed eigenframe O, the conditional lower block is

$$W + \sum_{j=1}^{q} (z_j + \sqrt{\lambda_j} v_j)(z_j + \sqrt{\lambda_j} v_j)^T,$$

where v_j^T is row j of $O^T K$ and the z_j are independent $N_m(0, C)$ vectors. Fixing all rows but j, Lemma 3.3 applies with

$$N = \alpha - q + 1 > p - q = m.$$

It follows that $F(O\,\mathrm{diag}(\lambda)O^T)$ is increasing in each λ_j, and so is its Haar average $\overline{F}(\lambda)$. Since $h(\lambda) = \prod_j \lambda_j^{\nu_1}$ is also increasing, eigenvalue association gives

$$\mathbb{E}\prod_i \Delta_i^{\nu_i} = \mathbb{E}[h(\lambda)\overline{F}(\lambda)] \geq \mathbb{E}h(\lambda)\mathbb{E}\overline{F}(\lambda).$$

This proves the first inequality in (1); relabeling treats every block. Principal submatrices have Wishart laws with the same shape, still admissible in their lower dimensions, so iteration proves the second inequality.

For equality, omit zero exponents. If the active covariance blocks are pairwise zero, the Wishart Laplace transform factors on their union, making the active diagonal blocks independent. Conversely, consider first two active blocks with $K \neq 0$. For Haar-almost every O, the last row of $O^T K$ is nonzero, since a uniform unit vector lies in the proper subspace $\ker K^T$ with probability zero. The strict part of Lemma 3.2 makes $\overline{F}$ strictly increasing in λ_q, as is the first determinant power. Their second moments are finite by the moment bound and conditional Jensen, so Lemma 2.1 gives strict inequality. For more blocks, split off every active block except a pair with nonzero cross covariance. The resulting pair inequality is strict, proving the characterization. □

4.1. **Shifts and exponent monotonicity**

The same proof permits proportional shifts; for $c_i \geq 0$, put

$$D_i = \det(X_{ii} + c_i\Sigma_{ii}).$$

The selected factor becomes a constant times $\det(S + c_i I)$, increasing in the eigenvalues, while the other shifts enter the fixed positive semidefinite offsets. Thus both inequalities in (1) hold with Δ_i replaced by D_i.

More is preserved by the spectral argument. Define

$$R(\nu) = \frac{\mathbb{E}\prod_i D_i^{\nu_i}}{\prod_i \mathbb{E}D_i^{\nu_i}}, \qquad \nu \in [0,\infty)^d.$$

Proposition 4.1. *The function R is nondecreasing in each exponent. More precisely, under the probability law tilted by $\prod_i D_i^{\nu_i}$, the ordered eigenvalues of $\Sigma_{jj}^{-1/2} X_{jj} \Sigma_{jj}^{-1/2}$ stochastically dominate their law under the single-factor tilt by $D_j^{\nu_j}$.*

Proof. The single-factor tilt changes the weights in (2) to $\lambda^{(\alpha-q-1)/2}e^{-\lambda/2}(\lambda+c_j)^{\nu_j}$ and preserves the independent Haar frame. Thus Lemma 2.1 still applies. The full tilt multiplies this eigenvalue law by the increasing conditional moment $\overline{F}$, normalized to have mean one. Association shows that such a tilt raises every bounded increasing expectation, proving stochastic domination. Apply it to D_j^t, by truncation, to obtain

$$\frac{R(\nu+te_j)}{R(\nu)} = \frac{\mathbb{E}_{\text{full tilt}}D_j^t}{\mathbb{E}_{\text{single tilt}}D_j^t} \geq 1. \qquad \square$$

5. **Overlapping compressions and spectral functions**

The proof of Theorem 1.1 separates into a regression step and an operator inequality. Keeping the regression step fixed allows the determinant calculation to control more general factors, and makes the rank restrictions transparent.

Call $f : (0,\infty) \to (0,\infty)$ *increasing and geometrically log-convex* when $h(s) = \log f(e^s)$ is finite, convex, and nondecreasing. This class contains nonnegative powers, positive polynomials with nonnegative coefficients, and products of $t^a(1+t)^b$ with $a, b \geq 0$.

Theorem 5.1. *Let $X \sim W_p(\alpha, \Sigma)$, and let B_i be full-column-rank $p \times k_i$ matrices. For $c_i \geq 0$ put*

$$D_i = \det(B_i^T X B_i + c_i B_i^T \Sigma B_i), \qquad F_i = f_i(D_i),$$

where the f_i are increasing and geometrically log-convex. A factor of proper rank $q = k_j < p$ can be split from the remaining product if only one other factor remains, or if

$$\alpha \geq \min\{q + \max_{i\neq j} k_i,\ p+1\}. \tag{10}$$

Full-rank factors split without an additional restriction. After discarding constant and full-rank factors, the product inequality holds if at most two factors remain, or if their ordered ranks satisfy

$$k_{(1)} \geq k_{(2)} \geq k_{(3)} \geq \cdots, \qquad \alpha \geq \min\{k_{(1)} + k_{(3)},\ p+1\}. \tag{11}$$

These are inequalities of extended nonnegative expectations. For two factors, both f_i may be arbitrary nonnegative nondecreasing functions with finite relevant moments.

Proof. First suppose that $h_i(s) = \log f_i(e^s)$ is smooth with bounded first and second derivatives, both nonnegative. For the regularized compressed determinants in (7), the chain rule gives

$$\mathcal{L} \log \prod_i f_i(\det Q_i) = \sum_i \{h_i'(\ell_i)\mathcal{L}\ell_i + h_i''(\ell_i)\Gamma(\ell_i, \ell_i)\}.$$

Together with (5) and (8), this is nonnegative wherever

$$N \geq 1 + \max_i \min(k_i, \operatorname{rank} V). \tag{12}$$

Bounded slopes of h_i give polynomial growth of f_i; for fixed regularization the first two matrix derivatives satisfy the same type of bound.

Select a compression of rank $q < p$, extend its columns to a basis, and whiten its covariance. The selected factor is now a function of $\det(S + c_j I)$, with the fixed determinant of its covariance absorbed into f_j, and the entire conditional matrix is (3).

Fix $S = O \operatorname{diag}(\lambda) O^T$ and rotate the regression rows by O^T, as in the proof of Theorem 1.1. If $L = [I_q, K]$, row a of the rotated matrix has the form

$$\sqrt{\lambda_a}\, e_a^T O^T L + \zeta_a^T, \qquad \zeta_a \sim N_p(0, \widetilde{C}), \qquad \widetilde{C} = \operatorname{diag}(0, C).$$

The noise rows remain independent, and their covariance has rank $m = p - q$. Fix all but row a. Together with the residual Wishart matrix, that row gives the variable V_t of Lemma 3.1, with $t = \lambda_a$, $N = \alpha - q + 1 > m$, and $\operatorname{rank} V_t \leq m + 1$. The fixed rows form a positive semidefinite offset, and every remaining factor is a compression of its sum with V_t. Condition (10) is exactly

$$N \geq 1 + \max_{i \neq j} \min(k_i, m + 1).$$

Thus (12) makes the conditional product moment increasing in each λ_a. The polynomial bounds justify removal of the regularization by dominated convergence.

With a single remaining determinant, condition also on the residual Wishart matrix. Compression leaves a fixed positive semidefinite offset and one Gaussian row, so Lemma 3.2 gives the stochastic comparison without a rank condition, even when the offset or covariance is singular. Every nonnegative nondecreasing outer function preserves it.

Haar averaging preserves this eigenvalue monotonicity, and Lemma 2.1 splits the selected factor. For a full-rank factor, its normalized compression is orthogonally equivalent to the entire whitened matrix; average the remaining, Loewner-increasing product over the eigenframe and apply association directly. Finally split the proper

factors in increasing rank order. Before the last pair, the selected rank is at most $k_{(3)}$ and the largest remaining rank at most $k_{(1)}$, so (11) suffices.

For a general finite convex increasing h, retain it on $[-n, n]$ and extend by supporting affine lines outside that interval. The resulting $h^{[n]} \le h$ is convex, increasing, and globally L_n-Lipschitz. If ρ is a smooth nonnegative mollifier supported in $[-1, 1]$, put

$$\varepsilon_n = \frac{1}{n(1 + L_n)}, \qquad h_n = h^{[n]} * \rho_{\varepsilon_n} - L_n \varepsilon_n.$$

Then $h_n \le h$, $h_n \to h$ pointwise, and h_n is smooth, convex, and increasing with bounded first two derivatives. Thus $0 < f_n \le f$ and $f_n \to f$ for $f_n(t) = e^{h_n(\log t)}$. Dominated convergence when the limiting moment is finite, and Fatou together with the upper bound when it is infinite, pass every inequality to the limit. Positive factors have positive marginal expectations, so extended products are unambiguous. The unrestricted two-factor assertion follows instead by bounded truncation from the single-determinant stochastic comparison. □

For disjoint blocks the rank conditions are automatic; if at least two factors remain after a selected block of size q, then $q + \max_{i \ne j} k_i \le p - 1 < \alpha$. Thus Theorem 1.1 extends to all the functions and shifts in Theorem 5.1.

The overlap restriction reflects a real obstruction to splitting. Let $X \sim W_2(\alpha, I_2)$ with $1 < \alpha < 2$, put $S = X_{11}$, and take

$$Q_\pm = (e_2 \pm a e_1)^T X (e_2 \pm a e_1), \qquad a > 0.$$

Regression gives $Q_\pm = W + (G \pm a\sqrt{S})^2$, where $W \sim \chi^2_{\alpha-1}$, $G \sim N(0, 1)$, and S are independent. Direct multiplication yields

$$\mathbb{E}[Q_+ Q_- \mid S] = \alpha(\alpha + 2) + 2(\alpha - 2)a^2 S + a^4 S^2,$$

so

$$\operatorname{Cov}(S, Q_+ Q_-) = 4\alpha a^2 \{\alpha - 2 + (\alpha + 2)a^2\} < 0$$

when $0 < a^2 < (2 - \alpha)/(\alpha + 2)$. Thus unrestricted splitting fails, even for three rank-one compressions at a nonsingular shape.

5.1. **Convex spectral gauges**

A convex spectral gauge is a nonzero continuous function $\Phi : \mathcal{S}^k_+ \to [0, \infty)$ that is convex, positively homogeneous, Loewner-increasing, and orthogonally invariant. Examples are the trace, the largest eigenvalue, sums of largest eigenvalues, and

Schatten norms of order at least one. General gauges do not turn congruences into scalar factors, so their covariance normalization is part of the statement.

Theorem 5.2. *For the compressions above, put*

$$C_i = B_i^T \Sigma B_i, \qquad S_i = C_i^{-1/2} B_i^T X B_i C_i^{-1/2}.$$

Let Φ_i be convex spectral gauges, $c_i \geq 0$, and f_i increasing and geometrically log-convex. For $F_i = f_i\{\Phi_i(S_i + c_i I)\}$, a proper factor of rank q splits whenever $\alpha \geq q+1$, or whenever at most one other factor remains. Full-rank factors split without a shape restriction.

After discarding constant and full-rank factors, the full product inequality holds if at most two factors remain, or if $\alpha \geq k_{(3)} + 1$. For disjoint principal blocks, every split holds throughout $\alpha > p - 1$. The assertions allow extended nonnegative expectations; for two factors, the outer functions may be arbitrary nonnegative nondecreasing functions with finite relevant moments.

Proof. The rank saving comes from representing a gauge as a supremum of positive linear forms. For a convex spectral gauge, this representation is

$$\Phi(Q) = \max_{H \in \mathcal{K}} \operatorname{tr}(HQ), \qquad Q \succeq 0, \tag{13}$$

where $\mathcal{K}$ is compact, consists of positive semidefinite matrices, and satisfies $\operatorname{tr} H \leq \Phi(I)$. To prove this, take a subgradient H at $Q \succ 0$. For $D \succeq 0$ and small $t > 0$, monotonicity and the subgradient inequality give $-t \operatorname{tr}(HD) \leq \Phi(Q - tD) - \Phi(Q) \leq 0$, hence $H \succeq 0$. Homogeneity gives $\operatorname{tr}(HQ) = \Phi(Q)$, and the subgradient inequality gives $\operatorname{tr}(HR) \leq \Phi(R)$ for every $R \succeq 0$. The set of all such H is closed with trace bounded by $\Phi(I)$, proving (13); continuity extends it to the boundary. Orthogonal averaging and the bound $Q \preceq (\operatorname{tr} Q) I$ also give

$$\frac{\Phi(I)}{k} \operatorname{tr} Q \leq \Phi(Q) \leq \Phi(I) \operatorname{tr} Q. \tag{14}$$

For $D \succeq 0$, a positive affine function $Q(V) = \varepsilon + \operatorname{tr}\{D(A + V)\}$ satisfies

$$\mathcal{L}_{N,v} \log Q = \frac{2N v^T D v}{Q} - \frac{4 v^T D V D v}{Q^2} \geq \frac{2(N-2) v^T D v}{Q}. \tag{15}$$

Indeed, $v^T D V D v \leq (v^T D v) \operatorname{tr}(DV) \leq (v^T D v) Q$. Thus the logarithm is subharmonic for this operator as soon as $N \geq 2$, independently of the compression rank.

To pass from affine functions to (13), choose a dense sequence (H_a) in $\mathcal{K}$ and set

$$\Phi_L(Q) = \left\{ \frac{1}{L} \sum_{a=1}^{L} (L^{-1} + \operatorname{tr}(H_a Q))^{L+1} \right\}^{1/(L+1)}.$$

These smooth positive functions converge to $\Phi(Q)$ and are bounded above by $L^{-1} + \Phi(Q)$. Their logarithms are convex increasing log-sum-exp functions of the logarithms in (15). The chain rule and $\Gamma \succeq 0$ therefore give $\mathcal{L} \log \Phi_L \geq 0$ when $N \geq 2$. Applying the same rule to smooth bounded-slope approximations of $\log f_i(e^s)$ proves $\mathcal{L} \prod_i f_i(\Phi_{i,L}) \geq 0$. Lemma 3.1, the upper bound on Φ_L, and polynomial domination give mean monotonicity for the limiting gauge product.

For a single remaining gauge, the condition $N \geq 2$ can be removed. The function $u \mapsto \Phi(A + uu^T)$ is even and convex, since

$$(\theta u + (1-\theta)w)(\theta u + (1-\theta)w)^T \preceq \theta uu^T + (1-\theta)ww^T.$$

Its sublevel sets are symmetric and convex. Gaussian mass of such a set decreases along a translating ray, including for a singular covariance. For completeness, intersect the set translated by $-rv$ with a centered ellipsoidal density level set E on the Gaussian support subspace. The resulting sections A_r satisfy $A_{-r} = -A_r$ and $\theta A_{r_1} + (1-\theta)A_{r_2} \subseteq A_{\theta r_1 + (1-\theta) r_2}$. Brunn–Minkowski makes their volume to the reciprocal dimension a concave even function, hence decreasing for $r \geq 0$. Integrating the ellipsoidal layers proves the Gaussian assertion; zero-dimensional support is immediate. Therefore the gauge of a shifted Gaussian row is stochastically increasing, and any nondecreasing outer function preserves the comparison.

Use the regression construction in the proof of Theorem 5.1. Its remaining row has $N = \alpha - q + 1$, so $N \geq 2$ is exactly $\alpha \geq q + 1$. Haar averaging and association prove the split; full-rank factors and the final pair are treated as there. Splitting in increasing rank order gives $\alpha \geq k_{(3)} + 1$. For disjoint blocks, $q \leq p - 2$ whenever two other blocks remain, so the shape condition is automatic. The bounded-slope approximations already constructed, together with (14), pass to general outer functions and extended expectations. □

5.2. **Convex functions of logarithmic spectra**

The determinant calculation also controls products of leading eigenvalues, and therefore every increasing convex function of the logarithmic spectrum. The two-factor range is narrower here because the one-factor quadratic-form argument does not apply to this whole class.

Theorem 5.3. *Use the normalized compressions S_i from Theorem 5.2. Let $h_i : \mathbb{R}^{k_i} \to \mathbb{R}$ be finite, convex, symmetric, and coordinatewise nondecreasing, and put*

$$F_i = \exp h_i\big(\log \lambda(S_i + c_i I)\big), \qquad c_i \geq 0.$$

A proper factor of rank q splits under (10)*, and full-rank factors split without an additional restriction. After discarding constant and full-rank factors, the product inequality holds for at most one remaining factor, or under*

$$\alpha \geq \min\{k_{(1)} + k_{(2)},\ p+1\}.$$

All expectations may be extended. For a partition into at least three blocks, every split and the product inequality hold throughout $\alpha > p-1$; for two disjoint blocks, the stated sufficient condition is $\alpha \geq p$.

Proof. Every subgradient of a finite convex increasing function has nonnegative coordinates. Choose supporting planes at a countable dense set of points and include every permutation of each plane. Their successive finite maxima h_L are symmetric, convex, and increasing, with $h_L \uparrow h$. On the chamber $s_1 \geq \cdots \geq s_k$, rearrangement writes

$$h_L(s) = \max_{\ell \leq L} \left(b_\ell + \sum_{j=1}^{k} a_{\ell j} s_j \right), \qquad a_{\ell 1} \geq \cdots \geq a_{\ell k} \geq 0.$$

With $a_{\ell,k+1} = 0$ and $T_j(Q) = \lambda_1(Q) \cdots \lambda_j(Q)$ for descending eigenvalues, this becomes

$$e^{h_L(\log \lambda(Q))} = \max_{\ell \leq L} e^{b_\ell} \prod_{j=1}^{k} T_j(Q)^{a_{\ell j} - a_{\ell,j+1}}.$$

The variational formula

$$T_j(Q) = \max_{U^T U = I_j} \det(U^T Q U) \tag{16}$$

follows from interlacing, with equality on the leading eigenspace. Replace its maximum by positive power means over finite dense subsets of the Stiefel manifold, first with $Q + \varepsilon I$. The logarithms of these means are convex increasing log-sum-exp functions of compressed log determinants. By (8), they have nonnegative generator on $\{\operatorname{rank} V \leq R\}$ whenever $N \geq \min(k, R)+1$. Nonnegative combinations and a further power-mean approximation to the outer maximum retain this property, as do products by (5).

For fixed finite approximations the regularity hypotheses of Lemma 3.1 hold. The maximum bounds and nonnegative powers give common polynomial envelopes, so dominated convergence removes the power means, frame discretizations, and ε. This proves mean monotonicity for each finite supporting-plane approximation. The regression and association argument now gives precisely (10); successive splitting, including the final pair, requires the largest-two-ranks condition. Finally $h_L \uparrow h$ and monotone convergence give the full assertion. For disjoint blocks the displayed rank

sums are at most $p-1$ when at least three blocks are present, and equal p for a partition into two blocks. □

5.3. **Singular integer Gram laws**

The radial identity also makes clear which real conclusions survive below the nonsingular shape range.

Corollary 5.4. *Let $X = Z^T Z$, where Z has $n \geq 1$ independent $N_p(0, \Sigma)$ rows, and use the determinant factors $f_i(D_i)$ of Theorem 5.1, extending f_i at zero by its finite nonnegative limit. Discard almost surely constant factors and omit full-rank factors from the rank test. The product inequality holds if at most two proper factors remain, or if*

$$n \geq k_{(1)} + k_{(3)}.$$

For two factors, arbitrary nonnegative nondecreasing outer functions with finite moments are allowed. All nonnegative determinant powers have finite moments.

Proof. For a selected proper compression of rank $q \leq n$, Gaussian regression writes the data as $[Y, YK + G]$, where Y is standard $n \times q$ Gaussian and G is independent. Conditional on $S = Y^T Y$, rotate its rows to $[S^{1/2}; 0]$. The representation (3) follows with $n - q$ residual Gaussian rows, whose Gram matrix may be singular. After fixing $q - 1$ mean rows, the integer part of Lemma 3.1 applies with $N = n - q + 1$; (9) is nonnegative under $n \geq q + \max_{i \neq j} k_i$. Splitting in increasing rank order gives the stated test.

For the final pair, even when $q > n$, condition on the selected Gram matrix and fix an eigenframe of its positive eigenvalues. Writing $S = \sum_{a=1}^{r} \lambda_a u_a u_a^T$, choose its data factor to have rows $\sqrt{\lambda_a} u_a^T$ followed by $n - r$ zero rows. Every other factor differs by an orthogonal row transformation, absorbed by the independent Gaussian regression noise. Varying one positive eigenvalue therefore changes only one mean row, so the single-determinant comparison proves conditional monotonicity without a rank condition. The positive eigenvalues have the eigenvalue law of $W_r(\max(n, q), I_r)$, where $r = \min(n, q)$, by comparing $Y^T Y$ and $Y Y^T$; they are associated and independent of the Haar frame. This proves the pair inequality. Full-rank normalized factors are spectral in the whole whitened Gram matrix and split directly by the same argument.

An unshifted determinant is zero with positive probability only when its compression rank exceeds n, in which case it is identically zero. All other determinant arguments are positive almost surely, so the regularization and outer-function approximations from Theorem 5.1 apply. Constant zero factors make both sides zero;

all other extended products are unambiguous. Polynomial determinant growth gives moment finiteness for powers. □

6. **Inverse-Wishart and beta-II inequalities**

The inverse-Wishart law has a more direct regression geometry; conditional on a principal block, the full matrix is an affine increasing function of that block with independent coefficients. This eliminates the radial rank restrictions and permits arbitrary monotone spectral factors.

Write

$$Y \sim \mathrm{IW}_p(\beta, \Omega) \quad \Longleftrightarrow \quad Y^{-1} \sim W_p(\beta, \Omega^{-1}), \qquad \beta > p - 1.$$

Theorem 6.1. *Let B_i be arbitrary full-column-rank $p \times k_i$ matrices and put*

$$\Omega_i = B_i^T \Omega B_i, \qquad Z_i = \Omega_i^{-1/2} B_i^T Y B_i \Omega_i^{-1/2}.$$

Suppose the nonnegative measurable functions Φ_i of the ordered eigenvalues of Z_i are all coordinatewise nondecreasing, or all coordinatewise nonincreasing. Whenever the relevant moments are finite,

$$\mathbb{E} \prod_i \Phi_i(\lambda(Z_i)) \geq \mathbb{E}\Phi_j(\lambda(Z_j))\, \mathbb{E} \prod_{i \neq j} \Phi_i(\lambda(Z_i)) \geq \prod_i \mathbb{E}\Phi_i(\lambda(Z_i))$$

for every j. There is no restriction on overlap or ranks.

Proof. Whiten the full scale. Then $C_i = \Omega^{1/2} B_i \Omega_i^{-1/2}$ has orthonormal columns, and the factors are spectral functions of $C_i^T Z C_i$ for $Z \sim \mathrm{IW}_p(\beta, I)$. Rotate the selected compression to the first q coordinates and put $m = p - q$. For $q < p$, the block variables

$$Z = \begin{pmatrix} S & U \\ U^T & V \end{pmatrix}, \qquad H = S^{-1}U, \qquad R = V - U^T S^{-1} U$$

give the representation

$$Z = \begin{pmatrix} 0 & 0 \\ 0 & R \end{pmatrix} + [I_q, H]^T S [I_q, H], \tag{17}$$

where

$$S \sim \mathrm{IW}_q(\beta - m, I), \quad R \sim \mathrm{IW}_m(\beta, I), \quad H = G R^{1/2},$$

and S, R, G are independent, with G a standard $q \times m$ Gaussian matrix. Indeed,

$$\det Z = \det S \det R, \quad dZ = (\det S)^m \, dS \, dH \, dR,$$

and

$$\operatorname{tr} Z^{-1} = \operatorname{tr} S^{-1} + \operatorname{tr} R^{-1} + \operatorname{tr}(HR^{-1}H^T).$$

The inverse-Wishart density consequently factors into the stated laws.

For fixed (R, H), (17) is affine and Loewner-increasing in S. The conditional expectation of the remaining product is therefore Loewner-monotone in the same direction as its factors. The ordered eigenvalues of S have density

$$c \prod_{a=1}^{q} \lambda_a^{-(\gamma+q+1)/2} e^{-1/(2\lambda_a)} \prod_{a<b} (\lambda_b - \lambda_a), \qquad \gamma = \beta - m,$$

with an independent Haar eigenframe. Lemma 2.1 applies. Average the conditional moment over the frame and compare it with the selected spectral factor to obtain the split. For decreasing functions use bounded association on their negatives, then truncate the original nonnegative factors. A full-rank selected factor is treated directly on the whole whitened matrix. Iterating the splits proves the product inequality. □

Corollary 6.2. *For $c_i \geq 0$, put $D_i = \det(B_i^T Y B_i + c_i \Omega_i)$. The preceding inequalities hold for $D_i^{\nu_i}$ if all $\nu_i \geq 0$ and*

$$2 \sum_i \nu_i < \beta - p + 1, \tag{18}$$

or if all $\nu_i \leq 0$, with no further condition. For nonnegative powers, (18) *is the exact joint finite-moment condition. In the unshifted case,*

$$\mathbb{E} \det(B_i^T Y B_i)^{\nu_i} = \det(\Omega_i)^{\nu_i} 2^{-k_i \nu_i} \frac{\Gamma_{k_i}((\beta - p + k_i)/2 - \nu_i)}{\Gamma_{k_i}((\beta - p + k_i)/2)}.$$

Proof. The spectral factors $\prod_j (\lambda_j + c_i)^{\nu_i}$ have the required common monotonicity. For positive powers, put $s = \sum_i \nu_i$. After orthonormalizing each compression and bounding its fixed shift by a scalar matrix, interlacing gives

$$\prod_i D_i^{\nu_i} \leq C \det(I + Y)^s.$$

With $X = Y^{-1}$ the right side is $C \det(I + X)^s \det(X)^{-s}$; the Wishart exponential tail controls infinity, while the matrix-gamma integral at the boundary is finite exactly when $\beta - 2s > p - 1$.

For necessity, keep all but the smallest eigenvalue of X in a fixed compact interval away from zero, and let the smallest eigenvalue λ tend to zero. Restrict its eigenvector to a set of positive Haar measure whose projections onto all active compression ranges are bounded away from zero. Each active compressed determinant is then at least

a positive constant times λ^{-1}, including fixed positive semidefinite shifts, while the eigenvalue density is bounded below by a constant times $\lambda^{(\beta-p-1)/2}$. Integrability forces $s < (\beta - p + 1)/2$.

For nonpositive powers, shifts decrease the factors, and each unshifted negative determinant power is bounded by a polynomial in $\|Y^{-1}\|$; all joint moments are therefore finite. Finally $B_i^T Y B_i$ has law $\mathrm{IW}_{k_i}(\beta - p + k_i, \Omega_i)$, either from (17) and rotation or by whitening. Its normalizing integral gives the gamma ratio, with $\Gamma_k(a) = \pi^{k(k-1)/4} \prod_{j=1}^{k} \Gamma(a - (j-1)/2)$. □

6.1. **Mixing the two laws**

The matrix beta-II law, in the normalization used here, has density

$$\frac{\Gamma_p((\alpha+\beta)/2)\det(\Omega)^{\beta/2}}{\Gamma_p(\alpha/2)\Gamma_p(\beta/2)}(\det X)^{(\alpha-p-1)/2}\det(\Omega+X)^{-(\alpha+\beta)/2}, \qquad X \succ 0, \tag{19}$$

where $\alpha, \beta > p - 1$. It has the hierarchy

$$Y \sim \mathrm{IW}_p(\beta, \Omega), \qquad X \mid Y \sim W_p(\alpha, Y).$$

To verify the constants, the joint density is proportional to

$$(\det X)^{(\alpha-p-1)/2}(\det Y)^{-(\alpha+\beta+p+1)/2}e^{-\operatorname{tr}((\Omega+X)Y^{-1})/2}.$$

Integrating Y contributes $2^{(\alpha+\beta)p/2}\Gamma_p((\alpha+\beta)/2)\det(\Omega+X)^{-(\alpha+\beta)/2}$, giving (19).

Theorem 6.3. *For the beta-II law* (19)*, the determinant factors* $f_i\{\det(B_i^T X B_i)\}$ *satisfy the product inequality under the rank condition of Theorem 5.1. In particular, disjoint principal blocks permit every pair of admissible shapes* $\alpha, \beta > p - 1$*. The functions may be increasing and geometrically log-convex, and the expectations extended.*

Proof. Let T_i have the determinant law of $W_{k_i}(\alpha, I)$ and define $m_i(t) = \mathbb{E} f_i(tT_i)$. This is a positive nondecreasing function, possibly infinite. Conditional on Y, each compressed determinant has marginal law $\det(B_i^T Y B_i)T_i$. Applying Theorem 5.1 conditionally gives

$$\mathbb{E}\left[\prod_i f_i\{\det(B_i^T X B_i)\} \mid Y\right] \geq \prod_i m_i\{\det(B_i^T Y B_i)\}.$$

The factors on the right are increasing spectral functions of normalized inverse-Wishart compressions. Apply Theorem 6.1 to $m_i \wedge L$, then let $L \uparrow \infty$. Monotone convergence and the tower property give the desired product of marginal expectations. □

The moment range also passes exactly through the mixture. For a full-column-rank $p \times k$ matrix B,

$$\mathbb{E}\det(B^T X B)^\nu = \det(B^T \Omega B)^\nu \frac{\Gamma_k(\alpha/2+\nu)}{\Gamma_k(\alpha/2)} \times \frac{\Gamma_k((\beta-p+k)/2-\nu)}{\Gamma_k((\beta-p+k)/2)}, \qquad 0 \le \nu < \frac{\beta-p+1}{2}.$$

For any family of nonnegative powers, joint finiteness is equivalent to (18), with no restriction on overlap. Indeed, realize $X = Y^{1/2} Z Y^{1/2}$ with independent $Z \sim W_p(\alpha, I)$. The bound $X \preceq \|Z\| Y$ and the finite moments of $\|Z\|$ give sufficiency, while the positive-probability event $Z \succeq I$ gives necessity from the inverse-Wishart criterion.

7. **Complex and quaternion Wishart matrices**

Over the complex and quaternion fields, the needed mean monotonicity already holds for a single Gaussian row. A common subharmonicity calculation treats both fields and explains why the real rank conditions disappear.

Let $\mathbb{F}$ be $\mathbb{C}$ or $\mathbb{H}$, and put $b = \dim_{\mathbb{R}} \mathbb{F} \in \{2, 4\}$. For a positive semidefinite Hermitian matrix, $\det_{\mathbb{F}}$ denotes the product of its real eigenvalues; over $\mathbb{H}$ this is the Moore determinant. We use the Wishart density

$$c(\det_{\mathbb{F}} X)^{\frac{b}{2}(\alpha-p+1)-1} e^{-\frac{b}{2}\Re\,\mathrm{tr}(\Sigma^{-1}X)}, \qquad X \succ 0, \quad \alpha > p-1. \tag{20}$$

Each real component of a standard Gaussian entry has variance $1/b$.

Theorem 7.1. *Let X have density (20), or let $X = Z^* Z$ where Z has any integer $n \ge 1$ independent $\mathbb{F}$-Gaussian rows of covariance $\Sigma \succ 0$. For arbitrary full-column-rank matrices $B_i \in \mathbb{F}^{p \times k_i}$, put*

$$C_i = B_i^* \Sigma B_i, \qquad S_i = C_i^{-1/2} B_i^* X B_i C_i^{-1/2}.$$

Let $h_i : \mathbb{R}^{k_i} \to \mathbb{R}$ be finite, convex, symmetric, and coordinatewise nondecreasing. Define

$$\Phi_i(S) = \exp h_i(\log \lambda(S)) \quad (S \succ 0), \qquad \Phi_i(S) = \lim_{\varepsilon \downarrow 0} \Phi_i(S + \varepsilon I) \quad (S \succeq 0).$$

For every $c_i \ge 0$ and every selected index j,

$$\mathbb{E}\prod_i \Phi_i(S_i + c_i I) \ge \mathbb{E}\Phi_j(S_j + c_j I)\, \mathbb{E}\prod_{i \ne j} \Phi_i(S_i + c_i I) \ge \prod_i \mathbb{E}\Phi_i(S_i + c_i I).$$

There is no restriction on overlap or ranks. The expectations may be extended; an identically zero factor makes both sides zero.

This includes increasing geometrically log-convex functions of

$$\det_{\mathbb{F}}(B_i^* X B_i + c_i B_i^* \Sigma B_i),$$

since the covariance determinant can be absorbed into the outer function. In particular all nonnegative determinant powers satisfy the inequalities, with finite moments. The theorem also includes nonnegative powers of elementary symmetric and Schur-positive polynomials in the eigenvalues, the functions $(\operatorname{tr} S^r)^{a/r}$ for $r > 0$, $a \geq 0$, and all nonnegative powers of unitarily invariant norms of $S^{1/2}$.

7.1. **Subharmonicity of a row update**

We use row vectors in this section. For $Q \succ 0$ and $w \in \mathbb{F}^{1\times k}$,

$$\det_{\mathbb{F}}(Q + w^* w) = \det_{\mathbb{F}} Q \, (1 + w Q^{-1} w^*). \tag{21}$$

The complex case is the determinant lemma. For quaternions, the real-algebra representation

$$\Psi(Z + Wj) = \begin{pmatrix} Z & W \\ -\overline{W} & \overline{Z} \end{pmatrix}$$

preserves multiplication and adjoints, and each eigenvalue of a Hermitian matrix occurs twice in its complex image. Hence $\det_{\mathbb{C}} \Psi(Q) = (\det_{\mathbb{H}} Q)^2$. The complex rank-two determinant lemma gives (21) after taking positive square roots. The same argument gives the Schur-complement identity and, for invertible square R,

$$\det_{\mathbb{F}}(R^* Q R) = \det_{\mathbb{F}}(R^* R) \det_{\mathbb{F}} Q.$$

Vary one row as $w(u) = a + uv$, with the scalar on the left. The nonconstant factor in (21) is

$$q(u) = c + 2\Re(u\eta) + d|u|^2,$$

where $c = 1 + aQ^{-1}a^* > 0$, $d = vQ^{-1}v^* \geq 0$, and $|\eta|^2 \leq d(c-1)$. If $d = 0$ the function is constant; otherwise completing the square gives $q(u) = d|u - u_0|^2 + e$ with $e \geq 1$. On $\mathbb{F} \simeq \mathbb{R}^b$,

$$\Delta \log q(u) = \frac{2(b-2)d}{q(u)} + \frac{4de}{q(u)^2} \geq 0. \tag{22}$$

Thus every compressed log determinant is subharmonic under a row update. The first term in (22) is precisely where the distinction from the real case enters.

Lemma 7.2. *Suppose Φ_i are continuous nonnegative functions of positive semidefinite Hermitian matrices, with polynomial upper growth, such that*

$$u \longmapsto \log \Phi_i\{A + (a + uv)^*(a + uv)\}$$

is subharmonic for every $A \succeq 0$ and every pair of rows a, v, allowing the value $-\infty$. Then for a centered $\mathbb{F}$-Gaussian row z, possibly with singular covariance, the expectation of a product of these factors under the row update $z + \sqrt{t}\, v$ is nondecreasing in t. Fixed compressions and positive semidefinite shifts are allowed in every factor.

Proof. For fixed z, the logarithm of the product is subharmonic in $u \in \mathbb{F}$. Exponentiation preserves the submean inequality by Jensen, so the product itself is subharmonic. Polynomial growth gives a locally uniform integrable envelope, and Gaussian averaging preserves continuity and the submean inequality.

A centered $\mathbb{F}$-Gaussian row is invariant under multiplication on the left by every unit scalar, while a common unit scalar does not change its Gram matrix. The averaged function is therefore radial in u. Its boundary value on a ball is constant, and the subharmonic maximum principle bounds every interior value by that boundary value. It follows that the average increases with $|u|$, as claimed. □

7.2. **Spectral functions and boundary faces**

We next pass from compressed determinants to the spectral class of Theorem 7.1. The approximation must preserve values on singular matrices, because the integer Gram laws may live entirely on a lower-rank stratum.

Fix one $h : \mathbb{R}^k \to \mathbb{R}$ as in the theorem. For $J \subseteq \{1, \dots, k\}$ define

$$h_J(x_J) = \lim_{t \to -\infty} h(x_J, t\mathbf{1}_{J^c}).$$

This limit is convex and bounded above on compact sets, and is either finite everywhere or identically $-\infty$. Indeed, if it is $-\infty$ at one point, express any other point as the midpoint of that point and a third point and use convexity. When finite, its supporting planes extend to affine minorants of h by putting zero slopes on J^c. All slopes are nonnegative by monotonicity.

For every finite face function choose supporting planes at a countable dense set of points and include all coordinate permutations. Enumerate them in finite permutation orbits. The successive finite maxima, exponentiated on logarithmic eigenvalues, give continuous functions Φ_L with

$$0 \le \Phi_L(S) \uparrow \Phi(S) \qquad (S \succeq 0). \tag{23}$$

On positive definite matrices this is the usual supporting-plane representation. At a singular matrix, the planes with zero slopes on the zero-eigenvalue coordinates recover the corresponding h_J. Their limit equals the regularized value in the theorem; sandwich the positive eigenvalues of $S + \varepsilon I$ between their limits and slightly larger fixed values, and use continuity of finite h_J; if $h_J \equiv -\infty$, the upper limit after exponentiation is zero. Including face planes is essential for this boundary convergence.

Each Φ_L is symmetric, Loewner-increasing, and of polynomial upper growth. To check its row property, sort the slopes of a supporting plane as $a_1 \geq \cdots \geq a_k \geq 0$ and the eigenvalues in descending order. Its permutation maximum is

$$e^b \prod_{j=1}^{k} \lambda_j^{a_j} = e^b \prod_{j=1}^{k} (\lambda_1 \cdots \lambda_j)^{a_j - a_{j+1}}, \qquad a_{k+1} = 0.$$

The variational formula (16) holds over both fields, with U^* in place of U^T and $\det_{\mathbb{F}}$ in place of det. At a point where a leading eigenvalue product is positive, choose a maximizing compression and apply the submean inequality for its log determinant, bounding boundary values by the maximum. At a zero the inequality is automatic; continuity of the eigenvalue product supplies upper semicontinuity. Thus the logarithm of each leading eigenvalue product is subharmonic under row updates. Nonnegative sums and finite maxima preserve this property, so each Φ_L satisfies Lemma 7.2. Positive regularization and its decreasing limit justify the assertion when the fixed offset is singular.

7.3. **Regression and completion of the proof**

Proof of Theorem 7.1. First take the polynomial-growth factors $\Phi_{i,L}$. Select a proper compression of rank q, extend its columns to a basis, and whiten its covariance. The scale becomes

$$\begin{pmatrix} I_q & K \\ K^* & C + K^*K \end{pmatrix}, \qquad C \succ 0.$$

For the nonsingular law, the full conditional matrix is

$$X \mid S \overset{d}{=} \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix} + Y^*Y, \qquad Y = [S^{1/2}, S^{1/2}K + G], \tag{24}$$

where $S \sim W_q^{\mathbb{F}}(\alpha, I)$, $W \sim W_{p-q}^{\mathbb{F}}(\alpha - q, C)$, and G has q independent centered $\mathbb{F}$-Gaussian rows. The Schur substitution $X_{12} = SK + S^{1/2}G$ has real Jacobian $(\det_{\mathbb{F}} S)^{b(p-q)/2}$. The determinant and trace identities factor the density exactly as in Lemma 2.2, proving (24) at every admissible real shape.

For an integer Gram law, write the Gaussian data directly as $[Y_0, Y_0K + G]$. Conditional on $S = Y_0^*Y_0$, its factor Y_0 is determined up to a left unitary transformation, which is absorbed by the independent law of G. This remains true when $q > n$ and S is singular.

Fix an eigenframe of S and all but one Gaussian mean row. The positive eigenvalues of S appear as squared lengths of the mean rows; the remaining conditional matrix is a fixed positive semidefinite offset plus one row with mean $\sqrt{\lambda_a}$ times a fixed vector. Every other normalized compression is a fixed compression of this matrix. Lemma 7.2 shows that its product moment increases with each positive eigenvalue.

In the nonsingular case the ordered eigenvalue density is

$$c \prod_a \lambda_a^{\frac{b}{2}(\alpha-q+1)-1} e^{-b\lambda_a/2} \prod_{a<c} (\lambda_c - \lambda_a)^b.$$

The power b of the Vandermonde comes from the b real coordinates of each off-diagonal spectral differential. In the integer case the positive eigenvalues have the law of $W_r^{\mathbb{F}}(\max(n,q), I)$, with $r = \min(n,q)$. In either case the eigenframe is independently Haar distributed and Lemma 2.1 applies. Averaging the conditional moment over that frame proves the selected split. A full-rank normalized factor is spectral in the entire whitened matrix and splits directly by the same association argument. Iteration gives the product inequality.

The polynomial-growth approximations have finite moments by the exponential density tail or the Gaussian row representation. Now let $L \to \infty$ in (23). Monotone convergence passes every split and product to its limit, including singular boundary strata and infinite moments. A compression has fixed rank almost surely; on that stratum its face function is either finite everywhere or identically $-\infty$. Thus a nonzero factor has positive marginal expectation, and the convention for identically zero factors covers the only possible ambiguity. □

To verify the spectral examples stated after the theorem, a symmetric polynomial with nonnegative monomial coefficients becomes a sum of exponentials of affine forms with nonnegative slopes after $\lambda_j = e^{x_j}$. Its logarithm is convex, symmetric, and increasing; Schur-positive polynomials have such coefficients. The same log-sum-exp calculation treats $(\sum_j \lambda_j^r)^{a/r}$. For a symmetric gauge norm γ, duality on the positive orthant gives

$$\gamma(t) = \max_{a \geq 0,\ \gamma^*(a) \leq 1} \sum_j a_j t_j,$$

so $\log \gamma(e^{x_1/2}, \dots, e^{x_k/2})$ is a supremum of convex log-sum-exp functions. Products, nonnegative sums, and increasing geometrically log-convex compositions preserve

the class. Norms and fixed polynomials have polynomial upper growth, giving finite moments for the uncomposed examples.

References

[1] R. L. Dykstra and J. E. Hewett, *Positive dependence of the roots of a Wishart matrix*, Ann. Statist. **6** (1978), 235–238; DOI 10.1214/aos/1176344083.

[2] C. Genest, F. Ouimet, and D. Richards, *On the Gaussian product inequality conjecture for disjoint principal minors of Wishart random matrices*, Electron. J. Probab. **29** (2024), paper no. 166; DOI 10.1214/24-EJP1222.

[3] C. Genest, F. Ouimet, and D. Richards, *An explicit Wishart moment formula for the product of two disjoint principal minors*, Proc. Amer. Math. Soc. **153** (2025), 1299–1311; DOI 10.1090/proc/17077.

[4] S. Karlin and Y. Rinott, *Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions*, J. Multivariate Anal. **10** (1980), 467–498; DOI 10.1016/0047-259X(80)90065-2.