Stein Kernels and Normal Approximation for Log-Concave Bilinear Forms

Tianle Liu*

August 30, 2026

Abstract

Jiang, Lee, and Vempala conjectured that if $X, Y \in \mathbb{R}^n$ are independent isotropic log-concave random vectors, then $W_2(\mathcal{L}(\langle X, Y\rangle), N(0,n))$ is bounded by a universal constant. Subject to Theorems 1.2 and 2.5 of Letwin's July 2026 preprint, we prove this conjecture and a rectangular bilinear-form extension. For independent isotropic log-concave $X \in \mathbb{R}^m$, $Y \in \mathbb{R}^n$, and nonzero $B \in \mathbb{R}^{m\times n}$, put

$$r_4(B) = \frac{\operatorname{Tr}(B^\top B)^2}{\operatorname{Tr}((B^\top B)^2)}.$$

We construct a nonnegative scalar Stein kernel for $X^\top BY/\|B\|_{\mathrm{HS}}$ whose squared L^2 discrepancy is at most $20/r_4(B)$, and consequently obtain the same bound for squared 2-Wasserstein distance to $N(0,1)$. The proof develops an exact covariance identity and deficit decomposition for trace observables of moment-map Stein kernels, together with a stability theorem for positive Stein kernels under log-concave approximation. Taking $B = I_n$ yields

$$W_2^2\left(\mathcal{L}\left(\frac{\langle X, Y\rangle}{\sqrt{n}}\right), N(0,1)\right) \le \frac{20}{n},$$

which is the Jiang–Lee–Vempala conjecture.

2020 Mathematics Subject Classification. Primary 60F05; Secondary 60E15, 52A23.

Keywords. Jiang–Lee–Vempala conjecture; log-concave bilinear form; spectral participation ratio; Stein kernel; Wasserstein distance; moment map.

1 Introduction

Jiang, Lee, and Vempala conjectured that there is a universal constant C such that

$$W_2(\mathcal{L}(\langle X, Y\rangle), N(0,n)) \le C \tag{1.1}$$

for every pair of independent isotropic log-concave vectors $X, Y \in \mathbb{R}^n$ [14]. Subject to the quadratic-form and moment-map matrix-energy estimates in Theorems 1.2 and 2.5 of Letwin's July 2026 preprint [25], we prove this conjecture. We in fact obtain the following rectangular extension: if $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ are independent isotropic log-concave vectors and $0 \ne B \in \mathbb{R}^{m\times n}$, then, with

$$r_4(B) := \frac{\|B\|_{\mathrm{HS}}^4}{\|B\|_{S_4}^4} = \frac{\operatorname{Tr}(B^\top B)^2}{\operatorname{Tr}((B^\top B)^2)},$$

*Department of Statistics and Operations Research, UNC Chapel Hill. Email: tianletl@unc.edu.

we have

$$W_2^2\left(\mathcal{L}\left(\frac{X^\top BY}{\|B\|_{\mathrm{HS}}}\right), N(0,1)\right) \le \frac{20}{r_4(B)}. \tag{1.2}$$

The parameter $r_4(B)$ is the participation ratio of the squared singular values of B, also called the Rényi-2 spectral effective rank; it lies between 1 and $\operatorname{rank}(B)$. Taking $m = n$ and $B = I_n$ in (1.2) gives $W_2^2(\mathcal{L}(\langle X, Y\rangle/\sqrt{n}), N(0,1)) \le 20/n$, equivalently $W_2^2(\mathcal{L}(\langle X, Y\rangle), N(0,n)) \le 20$, and thus proves the Jiang–Lee–Vempala conjecture under the cited estimates.

The proof has two structural steps. Let τ be a moment-map Stein kernel and let A be deterministic and symmetric. In the regular moment-map model, three source-coordinate integrations by parts give

$$\operatorname{Var}(\operatorname{Tr}(A\tau)) = \operatorname{Var}(Y^\top AY) - 3\mathbb{E}(AY)^\top \tau(AY) + \mathbb{E}\operatorname{Tr}(A\tau A\tau). \tag{1.3}$$

Fathi's weighted Poincaré inequality for the same kernel bounds $\mathbb{E}(AY)^\top\tau(AY)$ from below, and hence bounds the negative middle term in (1.3) from above. Polarization reveals an operator identity with an explicit nonnegative weighted-Poincaré deficit. Letwin's two estimates then give the covariance operator bound

$$\operatorname{Var}(\operatorname{Tr}(A\tau)) \le 4\operatorname{Tr}(A^2). \tag{1.4}$$

For the second step, put $S = X^\top BY$, $A = B^\top B$, and $Q = X^\top B\tau(Y)B^\top X$. The vector Stein identity for Y shows that $\mathbb{E}[Q \mid S]$ is a nonnegative scalar Stein kernel for S. The quadratic-form estimate controls the fluctuation of Q conditional on Y; (1.4), in direction A, controls its conditional mean. These contributions are at most $16\operatorname{Tr}(A^2)$ and $4\operatorname{Tr}(A^2)$. Projection onto S, normalization by $\operatorname{Var} S = \operatorname{Tr} A$, and the one-dimensional transport bound by Stein discrepancy yield (1.2).

1.1 Related literature

The geometric setting originates in the isoperimetric conjecture of Kannan, Lovász, and Simonovits [15]. It predicts a dimension-free comparison between the Cheeger constant of an isotropic log-concave measure and that of the worst linear half-space cut. Besides its intrinsic role in asymptotic convex geometry, KLS controls concentration, spectral gaps, and the efficiency of algorithms for sampling log-concave measures; see [21, 26] for background and applications.

The classical central-limit problem for convex bodies asks whether most one-dimensional marginals of a high-dimensional isotropic convex body are approximately Gaussian. Anttila, Ball, and Perissinaki established an early quantitative form [1], and Klartag proved a general central limit theorem and subsequent power-law estimates [16, 17]. Those results concern typical linear marginals of one high-dimensional law. The Jiang–Lee–Vempala conjecture instead randomizes the direction using an independent log-concave vector and asks for a dimension-free W_2 bound for their inner product [14]. The same work relates power-law forms of this conjecture to power-law estimates for the KLS constant, but its reverse implication is not an endpoint, dimension-free equivalence.

Bilinear and quadratic central limit theorems have a separate classical history. De Jong proved qualitative central limit theorems for generalized quadratic and multilinear forms built from independent scalar coordinates [6, 7]; quantitative and multidimensional versions were later developed by Döbler and Peccati [8]. Invariance principles for homogeneous sums and Gaussian-chaos fourth-moment phenomena appear in [27, 28]. For rank one and for equal-weight sums of independent Gaussian products, the natural law is variance-gamma rather than Gaussian; Stein's method for that target was developed by Gaunt [12]. These theories typically exploit coordinatewise independence, Hoeffding degeneracy, or an underlying Gaussian chaos. Here the coordinates within each factor may be strongly dependent: the moment-map kernel and the cited estimates accommodate this within-factor dependence, and the approximation rate is governed by the spectral participation ratio $r_4(B)$.

The strongest approaches to KLS have proceeded through stochastic localization. On the inverse-Cheeger normalization of the KLS constant, Eldan's reduction, combined with the then-best thin-shell estimate, gave $O(n^{1/3}\sqrt{\log n})$ [9, 13]; Lee and Vempala improved this to $O(n^{1/4})$ [24]; Chen's bound [2] was $\exp\big(O(\sqrt{\log n \, \log\log n})\big) = n^{o(1)}$; Klartag–Lehec obtained a polylogarithmic bound [20]; and Klartag sharpened it to the published $O(\sqrt{\log n})$ estimate [19]. Letwin's recent preprint improves this to $O(\log^{1/4} n)$ and proves the two estimates used below [25]. None of these KLS bounds is assumed as a black box in our argument.

Thin-shell estimates form a closely related line of work. Paouris proved sharp large-deviation estimates for isotropic log-concave measures [29], while Guédon and Milman interpolated those estimates with then-best thin-shell bounds [13]. Klartag and Lehec recently proved a dimension-free thin-shell bound by parallel coupling [22]. Chen and Klartag subsequently gave another account of the sharp result and proved a sharp Hilbert–Schmidt bound for the full third-moment tensor [3]. That tensor theorem is not used here; the matrix-energy estimate needed for general symmetric A is the one cited from Letwin.

Our analytic framework is the moment-map construction of Cordero-Erausquin and Klartag [4, 18]. Fathi observed that the Hessian of the moment potential supplies a positive-semidefinite Stein kernel and a weighted Poincaré inequality [10]. Stein kernels under spectral-gap assumptions and their discrepancy consequences were further developed in [5]. Finally, the passage from a scalar L^2-Stein discrepancy to 2-Wasserstein distance is taken from Ledoux, Nourdin, and Peccati [23]. Fathi and Mikulincer proved a different stability theory for invariant measures, moment measures, and Stein kernels [11]. The compactness theorem in section 4 is tailored instead to the positive kernels and uniform matrix energies used here. The contribution of this paper is to combine these tools through the exact identity (1.3), its operator deficit decomposition, stability under log-concave approximation, and scalarization for rectangular bilinear forms.

Dependency and scope. The recent quantitative results used here are Theorems 1.2 and 2.5 of [25], restated in theorem 2.1; the regular model and its approximation also use [25, Lemma 2.2 and Lemma A.1]. As checked on 30 August 2026, these statements occur in the first arXiv version of that preprint and have not undergone peer review. Our conclusions are therefore conditional on those statements. The argument neither assumes nor proves the KLS conjecture: the scalar observables $X^\top BY$ are much more special than arbitrary test functions of a log-concave vector.

Organization. Section 2 records definitions and the exact estimates used. Section 3 proves the trace identity, its operator deficit decomposition, and covariance consequences in the regular model. Section 4 proves an abstract stability theorem and removes the regularity assumptions. Section 5 proves the bilinear-form theorem and the Jiang–Lee–Vempala conjecture. Section 6 examines the role of spectral participation in Gaussian, rank-one, and product-exponential models; section 7 concludes with the scope and limitations.

2 Preliminaries and quantitative estimates

2.1 Log-concavity, isotropy, and Stein kernels

A probability measure on $\mathbb{R}^n$ is log-concave if it has a density of the form e^{-V} on a convex support, where V is convex in the extended sense. A random vector Y is isotropic if

$$\mathbb{E}Y = 0, \qquad \mathbb{E}(YY^{\mathsf{T}}) = I_n.$$

A measurable matrix field $\tau : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is a Stein kernel for a centered law μ if

$$\mathbb{E}\langle Y, \Phi(Y)\rangle = \mathbb{E}\langle \tau(Y), \nabla\Phi(Y)\rangle_{\mathrm{HS}} \tag{2.1}$$

for every smooth compactly supported vector field Φ. Here $\langle B, C\rangle_{\mathrm{HS}} = \mathrm{Tr}(B^{\mathsf{T}}C)$. If Y is isotropic, then testing coordinatewise gives $\mathbb{E}\tau(Y) = I_n$, whenever the relevant entries are integrable.

For a centered real random variable Z, a scalar Stein kernel is a measurable function κ_Z satisfying

$$\mathbb{E}[Z\phi(Z)] = \mathbb{E}[\kappa_Z(Z)\phi'(Z)] \tag{2.2}$$

for smooth tests ϕ. If $\operatorname{Var} Z = 1$, the squared L^2 Stein discrepancy associated with this kernel is $\mathbb{E}(\kappa_Z(Z) - 1)^2$.

2.2 The regular moment-map model

We first work with a regular isotropic log-concave target

$$\mathrm{d}\mu(y) = e^{-V(y)}\mathbf{1}_K(y)\,\mathrm{d}y, \tag{2.3}$$

where $K \subset \mathbb{R}^n$ is bounded, open, and convex, and V is smooth and convex on a neighborhood of $\overline{K}$. The moment-measure theorem [4, 18] supplies a smooth strictly convex potential ψ, unique up to translation, for which

$$\mathrm{d}\nu(u) = e^{-\psi(u)}\,\mathrm{d}u, \qquad Y = \nabla\psi(U) \sim \mu, \qquad U \sim \nu. \tag{2.4}$$

The normalization in (2.4) includes $\int e^{-\psi} = 1$. Write

$$H = \nabla^2\psi(U).$$

The map $\nabla\psi : \mathbb{R}^n \to K$ is a diffeomorphism in the regular model, and the transported Hessian

$$\tau(y) = \nabla^2\psi\Big((\nabla\psi)^{-1}(y)\Big) \tag{2.5}$$

is a symmetric positive-semidefinite Stein kernel [10]. In particular, $H = \tau(Y)$ and $\mathbb{E}H = I_n$.

The same construction yields the weighted Poincaré inequality

$$\operatorname{Var}_\mu g \le \mathbb{E}\langle \tau(Y)\nabla g(Y), \nabla g(Y)\rangle, \tag{2.6}$$

initially for smooth g and then by closure; see [10]. It is also the Brascamp–Lieb inequality for $e^{-\psi}$, transported by $\nabla\psi$: for $h = g \circ \nabla\psi$, one has $\nabla h = H\nabla g(Y)$, which gives exactly (2.6).

2.3 Quadratic and moment-map estimates

The following theorem records the complete quantitative content imported from [25].

Theorem 2.1 (Letwin's quadratic and moment-map estimates). *The cited [25, Theorems 1.2 and 2.5] state the following estimates. For every isotropic log-concave $Z \in \mathbb{R}^n$ and every deterministic symmetric matrix A,*

$$\mathrm{Var}(Z^{\mathsf{T}} AZ) \le 8\,\mathrm{Tr}(A^2). \tag{2.7}$$

In the regular moment-map model above, for every deterministic symmetric matrix A,

$$\mathbb{E}\,\mathrm{Tr}(AHAH) \le 2\,\mathrm{Tr}(A^2). \tag{2.8}$$

The constant 8 in (2.7) is sharp. The all-A form of (2.8), rather than only $A = I$, is needed for the trace covariance theorem and the rectangular bilinear-form extension; the Jiang–Lee–Vempala corollary itself specializes to $A = I_n$. This theorem records the sole quantitative input from Letwin's preprint; the qualitative regularity and approximation uses were identified above. There is no KLS assumption. In particular, the weighted inequality (2.6) is tied to the non-Euclidean moment-map kernel and does not assert a dimension-free Euclidean Poincaré inequality.

2.4 Transport by scalar Stein discrepancy

We use the following standard consequence of Stein discrepancy, proved in the generality needed here by Ledoux, Nourdin, and Peccati [23].

Theorem 2.2 (Stein discrepancy controls quadratic transport). *Let Z be centered with variance 1, and suppose it admits a scalar Stein kernel $\kappa_Z \in L^2$. If $G \sim N(0,1)$, then*

$$W_2^2(\mathcal{L}(Z), \mathcal{L}(G)) \le \mathbb{E}(\kappa_Z(Z) - 1)^2. \tag{2.9}$$

No density assumption on Z is needed: the general case follows by Gaussian regularization in [23].

3 Covariance transfer for moment-map Stein kernels

Fix a deterministic symmetric matrix A. In the regular model define

$$f_A = \mathrm{Tr}(AH), \qquad q_A = Y^{\mathsf{T}} AY, \tag{3.1}$$

and the deterministic quantities

$$T_A = \mathbb{E}(AY)^{\mathsf{T}} H(AY), \qquad U_A = \mathbb{E}\,\mathrm{Tr}(AHAH). \tag{3.2}$$

Although A need not be positive-semidefinite, both T_A and U_A are nonnegative: this is immediate after inserting $H^{1/2}$.

3.1 Three integrations by parts

The source measure $\nu = e^{-\psi}\,du$ satisfies

$$\mathbb{E}_\nu \,\mathrm{div}\, F(U) = \mathbb{E}_\nu \langle F(U), \nabla\psi(U)\rangle = \mathbb{E}_\nu \langle F(U), Y\rangle \tag{3.3}$$

for compactly supported smooth vector fields F.

Lemma 3.1 (Exact trace identity). *In the regular moment-map model, suppose that* $\mathbb{E}\,\mathrm{Tr}(H^2) < \infty$. *Then*

$$\boxed{\mathrm{Var}(f_A) = \mathrm{Var}(q_A) - 3T_A + U_A.} \tag{3.4}$$

Proof. We give the source-coordinate calculation before discussing cutoffs. Since $Y = \nabla\psi$ and $H = \nabla^2\psi$,

$$\mathrm{div}(AY) = \mathrm{Tr}(AH) = f_A, \qquad \nabla q_A = 2HAY. \tag{3.5}$$

Apply (3.3) first to the vector field $f_A AY$. This gives

$$\mathbb{E} f_A^2 = \mathbb{E}(f_A q_A) - \mathbb{E}\langle AY, \nabla f_A\rangle. \tag{3.6}$$

Applying (3.3) to $q_A AY$ and using (3.5) gives

$$\mathbb{E}(f_A q_A) = \mathbb{E} q_A^2 - 2T_A. \tag{3.7}$$

For the third identity, apply (3.3) to $AHAY$. Symmetry of the third derivatives of ψ implies

$$\mathrm{div}(AHAY) = \mathrm{Tr}(AHAH) + \langle AY, \nabla f_A\rangle. \tag{3.8}$$

Indeed, in index notation,

$$\partial_i(A_{ij}H_{jk}A_{k\ell}Y_\ell) = (AY)_k\, \partial_k f_A + \mathrm{Tr}(AHAH),$$

where symmetry of the third derivatives is used in the first term. On the other hand, symmetry of A gives $\langle AHAY, Y\rangle = (AY)^{\mathsf{T}} H(AY)$. Therefore

$$\mathbb{E}\langle AY, \nabla f_A\rangle = T_A - U_A. \tag{3.9}$$

Substitution of (3.7) and (3.9) into (3.6) yields

$$\mathbb{E} f_A^2 = \mathbb{E} q_A^2 - 3T_A + U_A. \tag{3.10}$$

The displayed calculation is formal until the cutoff passage below, which also supplies the centering needed to pass from (3.10) to (3.4).

Choose a radial $\chi \in C_c^\infty(\mathbb{R}^n)$, with $0 \le \chi \le 1$, equal to one on the unit ball and nonincreasing along rays, and put $\chi_R(u) = \chi(u/R)$, so that $\chi_R \uparrow 1$ and $|\nabla\chi_R| \le C/R$. Because the regular target K is bounded, there is an $R_K < \infty$ such that $|Y| \le R_K$. Applying (3.3) first to $\chi_R Y$ gives

$$\mathbb{E}[\chi_R \,\mathrm{Tr}\, H] = \mathbb{E}[\chi_R |Y|^2] - \mathbb{E}\langle Y, \nabla\chi_R\rangle.$$

Since $H \succeq 0$, monotone convergence and the boundedness of Y show that $\mathbb{E}\,\mathrm{Tr}\,H = n$. In particular, $\mathbb{E}\|H\|_{\mathrm{op}} \le n$. Applying the same argument to $\chi_R AY$ now gives

$$\mathbb{E} f_A = \mathbb{E} q_A = \mathrm{Tr}\, A, \tag{3.10a}$$

because $|f_A| \le \|A\|_{\mathrm{op}}\,\mathrm{Tr}\, H$ and the new boundary term is $O_A(R^{-1})$.

We next replace the three noncompact vector fields above by

$$\chi_R f_A AY, \qquad \chi_R q_A AY, \qquad \chi_R AHAY,$$

and put

$$\begin{aligned}
t_A &= (AY)^{\mathsf{T}} H(AY), \qquad u_A = \operatorname{Tr}(AHAH),\\
B_{1,R} &= \mathbb{E}\big[f_A\langle AY, \nabla\chi_R\rangle\big],\\
B_{2,R} &= \mathbb{E}\big[q_A\langle AY, \nabla\chi_R\rangle\big],\\
B_{3,R} &= \mathbb{E}\big[\langle AHAY, \nabla\chi_R\rangle\big].
\end{aligned}$$

Applying (3.3) at finite R, rather than to the uncut fields, gives

$$\begin{aligned}
\mathbb{E}[\chi_R f_A^2] &= \mathbb{E}[\chi_R f_A q_A] - \mathbb{E}[\chi_R\langle AY, \nabla f_A\rangle] - B_{1,R},\\
\mathbb{E}[\chi_R f_A q_A] &= \mathbb{E}[\chi_R q_A^2] - 2\mathbb{E}[\chi_R t_A] - B_{2,R},\\
\mathbb{E}[\chi_R\langle AY, \nabla f_A\rangle] &= \mathbb{E}[\chi_R t_A] - \mathbb{E}[\chi_R u_A] - B_{3,R}.
\end{aligned}$$

Combining these three identities first yields

$$\mathbb{E}[\chi_R f_A^2] = \mathbb{E}[\chi_R q_A^2] - 3\mathbb{E}[\chi_R t_A] + \mathbb{E}[\chi_R u_A] - B_{1,R} - B_{2,R} + B_{3,R}.$$

The boundary terms vanish without any boundedness assumption on H. Indeed,

$$|f_A| \le \|A\|_{\mathrm{op}} \operatorname{Tr} H, \qquad |AHAY| \le \|A\|_{\mathrm{op}}^2 R_K \operatorname{Tr} H,$$

while q_A and AY are bounded. Hence each $B_{j,R} = O_A(R^{-1})$ using only $\mathbb{E}\operatorname{Tr} H = n$. For the bulk terms, positivity of H and $\mathbb{E}\operatorname{Tr}(H^2) < \infty$ give integrable dominators: in particular,

$$f_A^2 \le n\|A\|_{\mathrm{op}}^2 \operatorname{Tr}(H^2), \qquad u_A \le \|A\|_{\mathrm{op}}^2 \operatorname{Tr}(H^2), \qquad t_A \le \|A\|_{\mathrm{op}}^2 R_K^2 \operatorname{Tr} H.$$

Dominated convergence therefore proves (3.10), and (3.10a) gives the centered identity (3.4). All appearances of ∇H have cancelled in the finite-R combination before $R \to \infty$. Thus the argument uses neither global boundedness of the Hessian nor an unproved third-derivative integrability estimate. Under the assumptions of theorem 3.3, the required second moment follows immediately from (2.8) with $A = I_n$. □

3.2 Covariance-operator form and trace bound

Let S_n be the Hilbert space of real symmetric $n \times n$ matrices with Hilbert–Schmidt inner product

$$\langle A, B\rangle_{\mathrm{HS}} = \operatorname{Tr}(AB).$$

For $A, B \in \mathsf{S}_n$, define the symmetric bilinear forms

$$\begin{aligned}
\Gamma_{\mathrm{tr}}(A, B) &= \operatorname{Cov}(f_A, f_B),\\
\Gamma_{\mathrm{quad}}(A, B) &= \operatorname{Cov}(q_A, q_B),\\
\mathsf{T}(A, B) &= \mathbb{E}(AY)^{\mathsf{T}} H(BY),\\
\mathsf{U}(A, B) &= \mathbb{E}\operatorname{Tr}(AHBH).
\end{aligned}$$

Thus $\mathsf{T}(A, A) = T_A$ and $\mathsf{U}(A, A) = U_A$. The symmetry of T follows from the symmetry of H, while the symmetry of U follows by transposition and cyclicity of the trace. We use the same notation for these forms and their Riesz representatives on S_n. In particular, I_{S_n} denotes the identity operator, whose associated bilinear form is

$$I_{\mathsf{S}_n}(A, B) = \operatorname{Tr}(AB).$$

For symmetric bilinear forms, $\mathsf{A} \preceq \mathsf{B}$ means that $(\mathsf{B} - \mathsf{A})(A, A) \ge 0$ for every $A \in \mathsf{S}_n$.

Proposition 3.2 (Polarized identity and weighted-Poincaré deficit). *In the regular moment-map model,*

$$\boxed{\Gamma_{\mathrm{tr}}(A,B)=\Gamma_{\mathrm{quad}}(A,B)-3\mathsf{T}(A,B)+\mathsf{U}(A,B)} \tag{3.11}$$

for every $A,B\in\mathsf{S}_n$. Moreover, the bilinear form

$$\mathsf{D}:=4\mathsf{T}-\Gamma_{\mathrm{quad}} \tag{3.12}$$

is positive-semidefinite, and

$$\boxed{\Gamma_{\mathrm{tr}}=\frac{1}{4}\Gamma_{\mathrm{quad}}+\mathsf{U}-\frac{3}{4}\mathsf{D}.} \tag{3.13}$$

Proof. Apply lemma 3.1 with $A+B$, subtract its instances with A and B, and divide by two. This gives (3.11). No further integration by parts or cutoff passage is required.

The weighted Poincaré inequality (2.6), applied to $q_A(y)=y^{\mathsf{T}}Ay$, gives

$$\Gamma_{\mathrm{quad}}(A,A)=\mathrm{Var}(q_A)\le 4\mathbb{E}(AY)^{\mathsf{T}}H(AY)=4\mathsf{T}(A,A).$$

Consequently $\mathsf{D}\succeq 0$. Solving (3.12) for T and substituting into (3.11) proves (3.13). □

The representation (3.13) separates the two quantitative inputs from the nonnegative slack in the weighted Poincaré inequality. The following statement records this transfer with arbitrary constants.

Theorem 3.3 (Trace bound with modular constants). *Suppose that, for some $C_q,C_{\mathrm{en}}\ge 0$, one has*

$$\mathrm{Var}(Y^{\mathsf{T}}AY)\le C_q\,\mathrm{Tr}(A^2),\qquad \mathbb{E}\,\mathrm{Tr}(AHAH)\le C_{\mathrm{en}}\,\mathrm{Tr}(A^2)$$

for every $A\in\mathsf{S}_n$. Then

$$0\preceq\Gamma_{\mathrm{tr}}\preceq\left(\frac{C_q}{4}+C_{\mathrm{en}}\right)I_{\mathsf{S}_n}. \tag{3.14}$$

More precisely, the full slack has the positive decomposition

$$\begin{aligned}\left(\frac{C_q}{4}+C_{\mathrm{en}}\right)I_{\mathsf{S}_n}-\Gamma_{\mathrm{tr}}=&\ \frac{1}{4}\left(C_qI_{\mathsf{S}_n}-\Gamma_{\mathrm{quad}}\right)\\&+\left(C_{\mathrm{en}}I_{\mathsf{S}_n}-\mathsf{U}\right)+\frac{3}{4}\mathsf{D}.\end{aligned} \tag{3.15}$$

Under the estimates in theorem 2.1, $C_q=8$ and $C_{\mathrm{en}}=2$. Hence

$$4I_{\mathsf{S}_n}-\Gamma_{\mathrm{tr}}=\frac{1}{4}\left(8I_{\mathsf{S}_n}-\Gamma_{\mathrm{quad}}\right)+\left(2I_{\mathsf{S}_n}-\mathsf{U}\right)+\frac{3}{4}\mathsf{D}\succeq 0, \tag{3.16}$$

and, for every deterministic symmetric A,

$$\boxed{\mathrm{Var}(\mathrm{Tr}(AH))\le 4\,\mathrm{Tr}(A^2).} \tag{3.17}$$

In particular,

$$\mathrm{Var}(\mathrm{Tr}\,H)\le 4n. \tag{3.18}$$

Proof. The two assumed estimates are precisely

$$\Gamma_{\text{quad}} \preceq C_q I_{\mathsf{S}_n}, \qquad \mathsf{U} \preceq C_{\text{en}} I_{\mathsf{S}_n}.$$

Since $\mathsf{D} \succeq 0$, (3.13) gives the upper bound in (3.14); the lower bound holds because Γ_{tr} is a covariance operator. Rearranging (3.13) gives the exact decomposition (3.15). Finally, substitute the constants from (2.7) and (2.8). Evaluating the resulting operator inequality in the direction A proves (3.17), and $A = I_n$ gives (3.18). □

Equation (3.15) also records the algebraic equality conditions. Equality in (3.14) in a fixed direction A requires equality in the quadratic-form estimate, the matrix-energy estimate, and the weighted Poincaré inequality in that direction. We do not infer from this a geometric classification of extremizing measures.

Corollary 3.4 (Finite-family covariance domination). *Assume the estimates in theorem 2.1. Let $A_1, \dots, A_k \in \mathsf{S}_n$, and define*

$$K_{ij} = \operatorname{Cov}(\operatorname{Tr}(A_i H), \operatorname{Tr}(A_j H)), \qquad G_{ij} = \operatorname{Tr}(A_i A_j).$$

Then

$$0 \preceq K \preceq 4G. \tag{3.19}$$

In particular, if $A_1, \dots, A_k$ are Hilbert–Schmidt orthonormal, then $K \preceq 4I_k$. For two arbitrary directions this also gives

$$|\operatorname{Cov}(\operatorname{Tr}(AH), \operatorname{Tr}(BH))| \le 4\sqrt{\operatorname{Tr}(A^2)\operatorname{Tr}(B^2)}.$$

Proof. For $c = (c_1, \dots, c_k) \in \mathbb{R}^k$, apply (3.17) to $A = \sum_{i=1}^k c_i A_i$. This gives

$$c^{\mathsf{T}} K c \le 4 \operatorname{Tr}\left(\left(\sum_{i=1}^k c_i A_i\right)^2\right) = 4c^{\mathsf{T}} G c.$$

The lower bound follows because K is a covariance matrix. The final display follows from Cauchy–Schwarz for the centered trace observables and (3.17). □

Corollary 3.5 (Global spectral budget). *Under the same assumptions, the covariance operator satisfies*

$$\operatorname{Tr}_{\mathsf{S}_n} \Gamma_{\text{tr}} = \mathbb{E}\|H - I_n\|_{\text{HS}}^2 = \mathbb{E}\operatorname{Tr}(H^2) - n \le n. \tag{3.20}$$

Consequently, every Hilbert–Schmidt orthonormal family $A_1, \dots, A_k \in \mathsf{S}_n$ satisfies

$$\sum_{i=1}^k \operatorname{Var}(\operatorname{Tr}(A_i H)) \le \min\{4k, n\}, \qquad \|\Gamma_{\text{tr}}\|_{\text{HS}(\mathsf{S}_n)}^2 \le 4n. \tag{3.21}$$

Proof. Let $(E_\alpha)_\alpha$ be a Hilbert–Schmidt orthonormal basis of S_n. Since $\mathbb{E}H = I_n$,

$$\begin{aligned} \operatorname{Tr}_{\mathsf{S}_n} \Gamma_{\text{tr}} &= \sum_\alpha \mathbb{E}\langle E_\alpha, H - I_n\rangle_{\text{HS}}^2 \\ &= \mathbb{E}\|H - I_n\|_{\text{HS}}^2 = \mathbb{E}\operatorname{Tr}(H^2) - n. \end{aligned}$$

The case $A = I_n$ of (2.8) bounds the last expression by n. The first inequality in (3.21) follows by combining this trace budget with $\|\Gamma_{\text{tr}}\|_{\text{op}} \le 4$. If $(\lambda_\alpha)_\alpha$ are the eigenvalues of Γ_{tr}, then $0 \le \lambda_\alpha \le 4$, and therefore

$$\|\Gamma_{\text{tr}}\|_{\text{HS}(\mathsf{S}_n)}^2 = \sum_\alpha \lambda_\alpha^2 \le 4 \sum_\alpha \lambda_\alpha \le 4n.$$

□

Remark (Deficit bookkeeping for product exponentials). Let $Y_i = E_i - 1$, where the E_i are independent Exp(1) variables. The positive Stein kernel is $H = \mathrm{diag}(E_1, \dots, E_n)$. Direct calculation gives

$$\Gamma_{\mathrm{tr}}(A, B) = \sum_{i=1}^{n} A_{ii} B_{ii}.$$

Thus Γ_{tr} is the orthogonal projection of S_n onto its diagonal subspace, and $\mathrm{Tr}_{\mathsf{S}_n} \Gamma_{\mathrm{tr}} = n$. Hence the global budget (3.20) is sharp, although the directional constant in (3.17) is not.

In the direction $A = I_n$,

$$\begin{aligned} \Gamma_{\mathrm{quad}}(I_n, I_n) = 8n, \quad \mathsf{T}(I_n, I_n) = 3n, \\ \mathsf{U}(I_n, I_n) = 2n, \quad \mathsf{D}(I_n, I_n) = 4n, \quad \Gamma_{\mathrm{tr}}(I_n, I_n) = n. \end{aligned}$$

The quadratic and matrix-energy deficits in (3.16) therefore vanish in this direction, while the weighted-Poincaré contribution is $\frac{3}{4}\mathsf{D}(I_n, I_n) = 3n$. It accounts exactly for the gap between the universal upper bound $4n$ and the actual trace variance n.

The exact mixed identity in proposition 3.2 is asserted for the regular moment-map Hessian. In the next section we pass its covariance and energy consequences to a limiting positive Stein kernel; we do not claim that every mixed term in (3.11) survives for an arbitrary nonsmooth kernel.

4 Stability of positive Stein kernels

The exact identity of lemma 3.1 uses the smooth moment-map coordinates. Its consequences, however, are stable under weak approximation. We first record a moment estimate that is useful for this passage. Write S_n for the space of real symmetric $n \times n$ matrices.

Lemma 4.1 (Fourth moments from a positive Stein kernel). *Let Y be centered with finite second moments, and suppose that Y admits a symmetric positive-semidefinite Stein kernel $\tau \in L^2$. Then*

$$\mathbb{E}|Y|^4 \le 9n\, \mathbb{E}\|\tau(Y)\|_{\mathrm{HS}}^2. \tag{4.1}$$

Proof. Fix a coordinate i and $R > 0$, and define the odd C^1 function

$$h_R(t) = 3 \int_0^t \min\{s^2, R^2\}\, \mathrm{d}s.$$

Thus

$$h_R'(t) = 3 \min\{t^2, R^2\}, \qquad t h_R(t) \ge \min\{t^4, R^4\}, \qquad |h_R(t)| \le 3R^2 |t|.$$

To justify this test, choose $\chi \in C_c^\infty(\mathbb{R}^n)$ equal to one on the unit ball and zero outside the ball of radius two, put $\chi_L(x) = \chi(x/L)$, and apply the Stein identity to smooth approximations of $\Phi_{L,R}(x) = e_i h_R(x_i) \chi_L(x)$. Removing the smooth approximation gives

$$\begin{aligned} \mathbb{E}[Y_i h_R(Y_i) \chi_L(Y)] = {} & \mathbb{E}[\tau_{ii}(Y) h_R'(Y_i) \chi_L(Y)] \\ & + \mathbb{E}\left[h_R(Y_i) \sum_{j=1}^{n} \tau_{ij}(Y) \partial_j \chi_L(Y) \right]. \end{aligned}$$

The first two terms are dominated by integrable multiples of Y_i^2 and τ_{ii}, respectively. On the support of $\nabla \chi_L$, the absolute value of the last integrand is bounded by

$$C R^2 \|\tau(Y)\|_{\mathrm{HS}} \mathbf{1}_{\{|Y| \ge L\}},$$

which tends to zero in L^1. Hence

$$\mathbb{E}[Y_i h_R(Y_i)] = 3\mathbb{E}\Big[\tau_{ii}(Y)\min\{Y_i^2, R^2\}\Big].$$

If $M_{i,R} = \mathbb{E}\min\{Y_i^4, R^4\}$, positivity of τ and Cauchy–Schwarz give

$$M_{i,R} \le 3\big(\mathbb{E}\tau_{ii}^2\big)^{1/2} M_{i,R}^{1/2}, \qquad M_{i,R} \le 9\mathbb{E}\tau_{ii}^2.$$

Letting $R \to \infty$ and summing over coordinates yields

$$\mathbb{E}|Y|^4 \le n\sum_{i=1}^n \mathbb{E}Y_i^4 \le 9n\sum_{i=1}^n \mathbb{E}\tau_{ii}^2 \le 9n\,\mathbb{E}\|\tau\|_{\mathrm{HS}}^2.$$

□

Proposition 4.2 (Weak stability of positive L^2 Stein kernels). *Let $\mu_k \Rightarrow \mu$ weakly on $\mathbb{R}^n$, where each μ_k is centered and has finite second moments. Let $Y_k \sim \mu_k$, and suppose that μ_k admits a symmetric positive-semidefinite Stein kernel τ_k such that*

$$L := \sup_k \mathbb{E}\|\tau_k(Y_k)\|_{\mathrm{HS}}^2 < \infty. \tag{4.2}$$

Then μ is centered, $\mu_k \to \mu$ in W_2, and μ admits a symmetric positive-semidefinite Stein kernel τ satisfying

$$\mathbb{E}\tau(Y) = \mathrm{Cov}(Y), \qquad \mathbb{E}\|\tau(Y)\|_{\mathrm{HS}}^2 \le \liminf_{k\to\infty} \mathbb{E}\|\tau_k(Y_k)\|_{\mathrm{HS}}^2, \tag{4.3}$$

where $Y \sim \mu$. The same kernel may be chosen with the following additional permanence properties.

(i) *If $F : \mathsf{S}_n \to [0,\infty]$ is convex and lower semicontinuous and $\sup_k \mathbb{E}F(\tau_k(Y_k)) \le C_F$, then*

$$\mathbb{E}F(\tau(Y)) \le C_F.$$

(ii) *If, for every $A \in \mathsf{S}_n$,*

$$\mathrm{Var}(\mathrm{Tr}(A\tau_k(Y_k))) \le C_{\mathrm{tr}}\,\mathrm{Tr}(A^2),$$

then

$$\mathrm{Var}(\mathrm{Tr}(A\tau(Y))) \le C_{\mathrm{tr}}\,\mathrm{Tr}(A^2) \qquad (A \in \mathsf{S}_n). \tag{4.4}$$

(iii) *If, for every $A \in \mathsf{S}_n$,*

$$\mathbb{E}\,\mathrm{Tr}(A\tau_k(Y_k)A\tau_k(Y_k)) \le C_{\mathrm{en}}\,\mathrm{Tr}(A^2),$$

then

$$\mathbb{E}\,\mathrm{Tr}(A\tau(Y)A\tau(Y)) \le C_{\mathrm{en}}\,\mathrm{Tr}(A^2) \qquad (A \in \mathsf{S}_n). \tag{4.5}$$

(iv) *If a constant C_{wp} satisfies*

$$\mathrm{Var}_{\mu_k} g \le C_{\mathrm{wp}}\,\mathbb{E}\langle \tau_k(Y_k)\nabla g(Y_k), \nabla g(Y_k)\rangle$$

for every k and every $g \in C_c^\infty(\mathbb{R}^n)$, then

$$\mathrm{Var}_\mu g \le C_{\mathrm{wp}}\,\mathbb{E}\langle \tau(Y)\nabla g(Y), \nabla g(Y)\rangle \qquad (g \in C_c^\infty(\mathbb{R}^n)). \tag{4.6}$$

Proof. Cutoff approximations to the linear vector fields $x \mapsto e_i x_j$ give

$$\mathbb{E}\tau_k(Y_k) = \mathbb{E}(Y_k Y_k^\mathsf{T}) = \mathrm{Cov}(Y_k). \tag{4.7}$$

By lemma 4.1,

$$\sup_k \mathbb{E}|Y_k|^4 \le 9nL.$$

Thus $(|Y_k|^2)_k$ is uniformly integrable. Weak convergence implies convergence of the first and second moments, and therefore

$$W_2(\mu_k, \mu) \longrightarrow 0. \tag{4.8}$$

Choose a subsequence along which the liminf in (4.3) is attained. The joint laws of $(Y_k, \tau_k(Y_k))$ are tight. Passing to a further subsequence, let

$$(Y_k, \tau_k(Y_k)) \Rightarrow (Y, M).$$

The cone of symmetric positive-semidefinite matrices is closed, so M is symmetric and positive-semidefinite, and Portmanteau gives

$$\mathbb{E}\|M\|_{\mathrm{HS}}^2 \le \liminf_k \mathbb{E}\|\tau_k(Y_k)\|_{\mathrm{HS}}^2.$$

For $\Phi \in C_c^\infty(\mathbb{R}^n;\mathbb{R}^n)$, truncate the matrix coordinate and pass to the joint weak limit in the Stein identity. The uniform L^2 bound supplies uniform integrability, and hence

$$\mathbb{E}\langle Y, \Phi(Y)\rangle = \mathbb{E}\langle M, \nabla\Phi(Y)\rangle_{\mathrm{HS}}. \tag{4.9}$$

A weak limit of graphs need not remain a graph, so define

$$\tau(Y) = \mathbb{E}[M \mid Y]. \tag{4.10}$$

Conditional expectation preserves symmetry and positive semidefiniteness, and (4.9) becomes the Stein identity for τ. Uniform integrability of the matrix variables, together with (4.7) and (4.8), gives

$$\mathbb{E}M = \lim_k \mathbb{E}\tau_k(Y_k) = \mathrm{Cov}(Y).$$

Conditional Jensen and Portmanteau now prove (4.3). The same two tools prove part (i).

For part (ii), conditional variance contraction and Portmanteau applied to the centered square give the claim; the centering constants converge by (4.7) and the second-moment convergence already proved. Explicitly,

$$\mathrm{Var}(\mathrm{Tr}(A\tau(Y))) \le \mathrm{Var}(\mathrm{Tr}(AM)) \le \liminf_k \mathrm{Var}(\mathrm{Tr}(A\tau_k(Y_k)))\,.$$

This proves (4.4).

For the matrix energy, first suppose that $B \succeq 0$. The map

$$T \longmapsto \mathrm{Tr}(BTBT) = \|B^{1/2}TB^{1/2}\|_{\mathrm{HS}}^2$$

is continuous and convex. Conditional Jensen and Portmanteau therefore pass the assumed estimate to τ. For arbitrary $A \in \mathsf{S}_n$, write $A = A_+ - A_-$ and $|A| = A_+ + A_-$. If $T \succeq 0$, then

$$\begin{aligned}\mathrm{Tr}(|A|T|A|T) - \mathrm{Tr}(ATAT) &= 4\,\mathrm{Tr}(A_+TA_-T)\\ &= 4\|A_-^{1/2}TA_+^{1/2}\|_{\mathrm{HS}}^2 \ge 0.\end{aligned} \tag{4.11}$$

Applying the positive-semidefinite case with $B = |A|$ gives

$$\mathbb{E}\operatorname{Tr}(A\tau A\tau) \le \mathbb{E}\operatorname{Tr}(|A|\tau|A|\tau) \le C_{\mathrm{en}}\operatorname{Tr}(|A|^2) = C_{\mathrm{en}}\operatorname{Tr}(A^2),$$

which is part (iii).

Finally, fix $g \in C_c^\infty(\mathbb{R}^n)$. Weak convergence gives $\operatorname{Var}_{\mu_k} g \to \operatorname{Var}_\mu g$. The function

$$(y, T) \longmapsto \langle T\nabla g(y), \nabla g(y)\rangle$$

is continuous and bounded in absolute value by $\|\nabla g\|_\infty^2 \|T\|_{\mathrm{HS}}$. Uniform integrability of the matrix variables lets us pass to the joint limit; conditioning on Y then replaces M by $\tau(Y)$. This proves (4.6). $\square$

Remark (Why the absolute-value step is necessary). For indefinite A, the map $T \mapsto \operatorname{Tr}(ATAT)$ is not convex even on the positive-semidefinite cone. For example, take

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad T_\pm = \begin{pmatrix} 1 & \pm c \\ \pm c & 1 \end{pmatrix}, \qquad 0 < c < 1.$$

Then $T_\pm \succeq 0$, their midpoint is I_2, and

$$\operatorname{Tr}(AI_2AI_2) = 2 > 2(1-c^2) = \frac{1}{2}\operatorname{Tr}(AT_+AT_+) + \frac{1}{2}\operatorname{Tr}(AT_-AT_-).$$

Thus direct conditional Jensen would be invalid. Positivity of the kernel and (4.11) preserve the full all-symmetric estimate.

Corollary 4.3 (Log-concave approximation). *Assume the estimates in theorem 2.1. Every isotropic log-concave law μ on $\mathbb{R}^n$ admits a symmetric positive-semidefinite Stein kernel τ satisfying, simultaneously for every $A \in \mathsf{S}_n$,*

$$\mathbb{E}\tau = I_n, \tag{4.12}$$

$$\mathbb{E}\operatorname{Tr}(A\tau A\tau) \le 2\operatorname{Tr}(A^2), \tag{4.13}$$

$$\operatorname{Var}(\operatorname{Tr}(A\tau)) \le 4\operatorname{Tr}(A^2). \tag{4.14}$$

It also satisfies

$$\operatorname{Var}_\mu g \le \mathbb{E}\langle \tau(Y)\nabla g(Y), \nabla g(Y)\rangle \qquad (g \in C_c^\infty(\mathbb{R}^n)). \tag{4.15}$$

In particular,

$$\mathbb{E}\operatorname{Tr}(\tau^2) \le 2n, \qquad \operatorname{Var}(\operatorname{Tr}\tau) \le 4n. \tag{4.16}$$

Proof. The smoothing, truncation, recentering, and whitening construction of [25, Lemma A.1] provides regular isotropic log-concave laws $\mu_k \Rightarrow \mu$. Let τ_k be their moment-map Stein kernels. Fathi's construction gives the Stein identity and the weighted Poincaré inequality (2.6). Letwin's matrix-energy estimate (2.8) gives

$$\mathbb{E}\operatorname{Tr}(A\tau_k A\tau_k) \le 2\operatorname{Tr}(A^2),$$

and theorem 3.3 gives

$$\operatorname{Var}(\operatorname{Tr}(A\tau_k)) \le 4\operatorname{Tr}(A^2).$$

Taking $A = I_n$ in the energy estimate supplies the uniform L^2 hypothesis of proposition 4.2. The conclusions now follow from that proposition with

$$C_{\mathrm{en}} = 2, \qquad C_{\mathrm{tr}} = 4, \qquad C_{\mathrm{wp}} = 1.$$

$\square$

Remark. The kernel furnished by corollary 4.3 is a weak-limit kernel. It need not be the pointwise Hessian of a nonsmooth moment potential, it need not be canonical, and it may depend on the chosen subsequence. Only its Stein identity and the displayed inequalities are used below.

5 Normal approximation for log-concave bilinear forms

We now prove the rectangular statement announced in the introduction. For a nonzero matrix B, define its spectral participation ratio by

$$r_4(B) := \frac{\|B\|_{\mathrm{HS}}^4}{\|B\|_{S_4}^4} = \frac{\mathrm{Tr}(B^\mathsf{T} B)^2}{\mathrm{Tr}((B^\mathsf{T} B)^2)}. \tag{5.1}$$

Theorem 5.1 (Log-concave bilinear forms). *Assume the estimates in theorem 2.1. Let $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ be independent isotropic log-concave random vectors, let $0 \neq B \in \mathbb{R}^{m\times n}$, and set*

$$S = X^\mathsf{T} BY, \qquad \sigma^2 = \|B\|_{\mathrm{HS}}^2.$$

Then $Z = S/\sigma$ admits a nonnegative scalar Stein kernel κ_Z such that

$$\mathbb{E}(\kappa_Z(Z) - 1)^2 \leq \frac{20}{r_4(B)}. \tag{5.2}$$

Consequently, for $G \sim N(0, 1)$,

$$\boxed{W_2^2(\mathcal{L}(Z), \mathcal{L}(G)) \leq \frac{20}{r_4(B)}.} \tag{5.3}$$

Equivalently, if $G_\sigma \sim N(0, \sigma^2)$, then

$$W_2^2(\mathcal{L}(S), \mathcal{L}(G_\sigma)) \leq 20\frac{\mathrm{Tr}((B^\mathsf{T} B)^2)}{\mathrm{Tr}(B^\mathsf{T} B)}. \tag{5.4}$$

Proof. Choose for Y a positive-semidefinite Stein kernel $\tau(Y)$ furnished by corollary 4.3, and put

$$A = B^\mathsf{T} B, \qquad Q = X^\mathsf{T} B\tau(Y)B^\mathsf{T} X. \tag{5.5}$$

Independence and isotropy give

$$\mathbb{E}S = 0, \qquad \mathrm{Var}\, S = \mathrm{Tr}\, A = \sigma^2, \qquad \mathbb{E}Q = \mathrm{Tr}\, A. \tag{5.6}$$

Condition on X and apply the vector Stein identity for Y to

$$\Phi_x(y) = B^\mathsf{T} x\, \phi(x^\mathsf{T} By).$$

Starting with compactly supported smooth ϕ, radial cutoffs in y, and bounded x, one may remove both restrictions using $\tau \in L^2$, truncation of X, and the L^2 estimate for Q below. Since

$$\nabla_y \Phi_x(y) = B^\mathsf{T} xx^\mathsf{T} B\, \phi'(x^\mathsf{T} By),$$

the result is

$$\mathbb{E}[S\phi(S)] = \mathbb{E}[Q\phi'(S)]. \tag{5.7}$$

Therefore

$$\kappa_S(S) = \mathbb{E}[Q \mid S] \tag{5.8}$$

is a scalar Stein kernel for S. It is nonnegative because $\tau \succeq 0$.

Conditional on Y, apply (2.7) to X with the symmetric matrix $B\tau(Y)B^{\mathsf{T}}$. By cyclicity of the trace and (4.13),

$$\begin{aligned}\mathbb{E}\operatorname{Var}_X(Q \mid Y) &\leq 8\mathbb{E}\operatorname{Tr}(A\tau(Y)A\tau(Y)) \\ &\leq 16\operatorname{Tr}(A^2).\end{aligned} \tag{5.9}$$

Isotropy of X gives

$$\mathbb{E}_X[Q \mid Y] = \operatorname{Tr}(A\tau(Y)),$$

and (4.14) gives

$$\operatorname{Var}(\mathbb{E}_X[Q \mid Y]) \leq 4\operatorname{Tr}(A^2). \tag{5.10}$$

The law of total variance now yields

$$\mathbb{E}(Q - \sigma^2)^2 \leq 20\operatorname{Tr}(A^2). \tag{5.11}$$

Since conditional expectation is an L^2 contraction,

$$\mathbb{E}(\kappa_S(S) - \sigma^2)^2 \leq 20\operatorname{Tr}(A^2).$$

The rescaled kernel $\kappa_Z(Z) = \kappa_S(S)/\sigma^2$ therefore satisfies (5.2). Apply theorem 2.2 to obtain (5.3); scaling W_2 by σ gives (5.4). □

Corollary 5.2 (Jiang–Lee–Vempala conjecture). *Assume the estimates in theorem 2.1. If $X, Y \in \mathbb{R}^n$ are independent isotropic log-concave random vectors and $G \sim N(0,1)$, then*

$$\boxed{W_2^2\left(\mathcal{L}\left(\frac{\langle X, Y\rangle}{\sqrt{n}}\right), \mathcal{L}(G)\right) \leq \frac{20}{n}.} \tag{5.12}$$

Equivalently, if $G_n \sim N(0,n)$, then

$$W_2^2(\mathcal{L}(\langle X, Y\rangle), \mathcal{L}(G_n)) \leq 20. \tag{5.13}$$

Thus the dimension-free central limit conjecture of Jiang, Lee, and Vempala [14] follows from the cited estimates.

Proof. Take $m = n$ and $B = I_n$ in theorem 5.1; then $r_4(I_n) = n$ and $\sigma^2 = n$. □

Remark (Input accounting). The proof uses a positive-semidefinite Stein kernel for one factor, its full matrix-energy bound (4.13), the trace covariance bound (4.14), the quadratic-form estimate for the other factor, and the scalar Stein-discrepancy-to-W_2 inequality. The extension from regular targets uses [25, Lemma A.1] but adds no quantitative constant. The argument does not use KLS, a density formula for the bilinear form, or a separate central limit theorem. The sharp Hilbert–Schmidt estimate for the full third-moment tensor proved by Chen and Klartag [3] is part of the surrounding theory but is not used here. Independence is essential in (5.6), (5.9), and the conditional-mean identity; isotropy fixes the target variance and that conditional mean.

Remark (Spectral participation and the strength of the Stein statement). Since $1 \leq r_4(B) \leq \operatorname{rank}(B)$, the theorem gives a central limit regime precisely when the singular-value mass is not concentrated in a few directions. The trivial independent coupling gives $W_2^2(\mathcal{L}(Z), N(0,1)) \leq 2$, so the transport conclusion may be sharpened to $\min\{2, 20/r_4(B)\}$. Estimate (5.2) contains more information: it supplies an approximate integration-by-parts identity for every smooth scalar test. No claim is made here about total variation, relative entropy, or a multivariate approximation.

6 Models and sharpness checks

The following computations are not needed for the proof. They show why the spectral participation ratio is the natural parameter, compare the universal constant with exact model calculations, and verify that the order in (5.3) cannot in general be improved.

6.1 Rank one and the necessity of spectral participation

Suppose first that X and Y are standard Gaussian and that $B = ab^{\mathsf{T}}$ has rank one. After normalization,

$$\frac{X^{\mathsf{T}}BY}{\|B\|_{\mathrm{HS}}} \overset{d}{=} GH,$$

where G, H are independent standard Gaussians. This law is not normal: it has fourth moment 9, rather than 3. Ordinary rank growth does not by itself repair this obstruction. Indeed, let B_k have singular values $1, k^{-1}, \dots, k^{-1}$, with $k-1$ copies of k^{-1}. Then $\operatorname{rank}(B_k) = k$, but $r_4(B_k) \to 1$, and the normalized Gaussian bilinear form converges in L^2 to GH. Thus the singular values must become diffuse. For rank-one B, $r_4(B) = 1$, so theorem 5.1 correctly gives only a constant-order estimate. Sums of such products also explain the relevance of variance-gamma approximation in low-rank regimes [12].

6.2 Gaussian bilinear forms

Let $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ be independent standard Gaussian vectors. If (s_i) are the nonzero singular values of B, orthogonal invariance gives

$$\frac{X^{\mathsf{T}}BY}{\|B\|_{\mathrm{HS}}} \overset{d}{=} \frac{\sum_i s_i G_i H_i}{(\sum_i s_i^2)^{1/2}}, \tag{6.1}$$

with independent standard Gaussian families (G_i) and (H_i). Since the Stein kernel of Y is I_n, the pre-kernel in (5.5) is $Q = X^{\mathsf{T}}BB^{\mathsf{T}}X$, and

$$\operatorname{Var} Q = 2\operatorname{Tr}((B^{\mathsf{T}}B)^2).$$

The same projection argument as in theorem 5.1 therefore gives the sharper model-specific estimate

$$\mathbb{E}(\kappa_Z(Z) - 1)^2 \le \frac{2}{r_4(B)}, \qquad W_2^2(\mathcal{L}(Z), N(0,1)) \le \frac{2}{r_4(B)}. \tag{6.2}$$

The fourth moment is exact:

$$\mathbb{E}Z^4 = 3 + \frac{6}{r_4(B)}. \tag{6.3}$$

This confirms both the effective-rank scale and the fact that the universal constant 20 is not expected to be optimal.

For completeness, the Gaussian model also tests the trace identity. If $Y \sim N(0, I_n)$, then $\tau(Y) = I_n$, and for every symmetric A,

$$\operatorname{Var}(Y^{\mathsf{T}}AY) = 2\operatorname{Tr}(A^2), \qquad T_A = \operatorname{Tr}(A^2), \qquad U_A = \operatorname{Tr}(A^2).$$

Thus the right side of (3.4) is

$$2\operatorname{Tr}(A^2) - 3\operatorname{Tr}(A^2) + \operatorname{Tr}(A^2) = 0,$$

which is exactly $\operatorname{Var}(\operatorname{Tr}(A\tau)) = 0$. This confirms both the negative sign and the coefficient 3 in the trace identity.

6.3 Rectangular products of centered exponentials

Let the coordinates of $X \in \mathbb{R}^m$ and $Y \in \mathbb{R}^n$ be independent centered unit-rate exponentials, and take the two vectors independent of one another. Thus $X_i = E_i - 1$, $Y_j = F_j - 1$, and $\tau(Y) = \operatorname{diag}(F_1, \ldots, F_n)$. Write b_j for the columns and r_i for the rows of B, and set

$$c_{\text{col}} = \sum_j \|b_j\|^4, \qquad c_{\text{row}} = \sum_i \|r_i\|^4, \qquad c_4 = \sum_{i,j} b_{ij}^4.$$

For $A = B^\mathsf{T} B$, a direct calculation from $\mathbb{E}X_i^4 = \mathbb{E}Y_j^4 = 9$ gives

$$\operatorname{Var} Q = 2\operatorname{Tr}(A^2) + 3c_{\text{col}} + 6c_{\text{row}} + 6c_4 \le 17\operatorname{Tr}(A^2). \tag{6.4}$$

Indeed,

$$\mathbb{E}\operatorname{Tr}(B\tau B^\mathsf{T} B\tau B^\mathsf{T}) = \operatorname{Tr}(A^2) + c_{\text{col}},$$

the non-Gaussian diagonal correction to the conditional quadratic variance is $6(c_{\text{row}} + c_4)$, and the variance of the conditional mean is c_{col}. Each of $c_{\text{col}}, c_{\text{row}}, c_4$ is at most $\operatorname{Tr}(A^2)$. Equality in the final bound of (6.4) holds when B is diagonal, so the universal estimate $20\operatorname{Tr}(A^2)$ is close to the exact pre-kernel variance on this extremal log-concave model.

Take, in particular, $B = I_r$. For

$$Z_r = \frac{1}{\sqrt{r}} \sum_{i=1}^r X_i Y_i,$$

independence gives the exact moment identities

$$\mathbb{E}Z_r^3 = \frac{4}{\sqrt{r}}, \qquad \mathbb{E}Z_r^4 = 3 + \frac{78}{r}. \tag{6.5}$$

For any coupling of Z_r with $G \sim N(0,1)$, factor $Z_r^3 - G^3 = (Z_r - G)(Z_r^2 + Z_r G + G^2)$ and apply Cauchy–Schwarz. By (6.5),

$$\mathbb{E}(Z_r^2 + Z_r G + G^2)^2 \le 3\left(\mathbb{E}Z_r^4 + \sqrt{\mathbb{E}Z_r^4 \mathbb{E}G^4} + \mathbb{E}G^4\right) \le C$$

for a universal C, independently of the coupling. Consequently there is a universal $c > 0$ such that

$$W_2^2(\mathcal{L}(Z_r), N(0,1)) \ge \frac{c}{r}. \tag{6.6}$$

Thus the order $1/r_4(B)$ in theorem 5.1 is sharp in general, even though its numerical constant is not.

7 Conclusion and qualifications

Under the quantitative estimates stated in theorem 2.1, this paper proves the Jiang–Lee–Vempala dimension-free central limit conjecture and a rectangular extension. For $X^\mathsf{T} BY$, the normalized squared Stein discrepancy and squared 2-Wasserstein error are at most $20/r_4(B)$; the choice $B = I_n$ gives the conjectured dimension-free bound for $\langle X, Y\rangle$. The argument derives the required trace covariance bound from an exact moment-map identity and a nonnegative deficit decomposition, then preserves the covariance, matrix-energy, and weighted Poincaré estimates through a general stability theorem for positive Stein kernels.

The conclusion does not prove KLS. The reverse implication in [14, Theorem 7] is an exponent-level statement with fixed $\varepsilon \in (0, 1/2)$ and an arbitrary loss $\delta > 0$. For any $\eta > 0$, the present $O(1)$ bound is also $O(n^{\eta})$, so that result with $\varepsilon = 1/2 - \eta$ yields only a subpolynomial conclusion for the KLS constant: for every $\alpha > 0$, it is $O_{\alpha}(n^{\alpha})$, not a dimension-free Poincaré bound. Letwin's explicit polylogarithmic estimate is already stronger than this consequence; recovering KLS would require an endpoint conversion controlling general nonlinear observables.

The Gaussian and product-exponential models show that $1/r_4(B)$ is the correct general order, while also exhibiting slack in the numerical constant 20. No optimality claim is made for that constant. For nonsmooth targets, corollary 4.3 constructs a limiting positive-semidefinite Stein kernel with the required bounds; it does not assert a canonical kernel or a classical Hessian formula. The exact mixed identity itself remains a statement about the regular moment-map Hessian.

AI disclosure

GPT 5.6 Sol assisted with literature triage, algebraic cross-checking, drafting, and production checks, while Claude Opus 5 provided independent proof verification and writing feedback. The author independently verified the mathematical arguments and citations and takes full responsibility for the content of the manuscript.

References

[1] Milla Anttila, Keith Ball, and Irini Perissinaki. The central limit problem for convex bodies. *Transactions of the American Mathematical Society*, 355(12):4723–4735, 2003. doi: 10.1090/S0002-9947-03-03085-X. See the erratum in Trans. Amer. Math. Soc. 356 (2004), 2137, doi:10.1090/S0002-9947-03-03578-5.

[2] Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geometric and Functional Analysis*, 31(1):34–61, 2021. doi: 10.1007/s00039-021-00558-4.

[3] Yuansi Chen and Bo'az Klartag. Digesting the proof of the sharp thin-shell inequality, 2026. URL https://arxiv.org/abs/2607.23307. arXiv:2607.23307v1, submitted 25 July 2026.

[4] Dario Cordero-Erausquin and Bo'az Klartag. Moment measures. *Journal of Functional Analysis*, 268(12):3834–3866, 2015. doi: 10.1016/j.jfa.2015.04.001.

[5] Thomas A. Courtade, Max Fathi, and Ashwin Pananjady. Existence of Stein kernels under a spectral gap, and discrepancy bounds. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 55(2):777–790, 2019. doi: 10.1214/18-AIHP898.

[6] Peter de Jong. A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields*, 75(2):261–277, 1987. doi: 10.1007/BF00354037.

[7] Peter de Jong. A central limit theorem for generalized multilinear forms. *Journal of Multivariate Analysis*, 34(2):275–289, 1990. doi: 10.1016/0047-259X(90)90040-O.

[8] Christian Döbler and Giovanni Peccati. Quantitative de Jong theorems in any dimension. *Electronic Journal of Probability*, 22(2):1–35, 2017. doi: 10.1214/16-EJP19.

[9] Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geometric and Functional Analysis*, 23(2):532–569, 2013. doi: 10.1007/s00039-013-0214-y.

[10] Max Fathi. Stein kernels and moment maps. *The Annals of Probability*, 47(4):2172–2185, 2019. doi: 10.1214/18-AOP1305.

[11] Max Fathi and Dan Mikulincer. Stability estimates for invariant measures of diffusion processes, with applications to stability of moment measures and Stein kernels. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze (5)*, 23(3):1417–1445, 2022. doi: 10.2422/2036-2145.202011_016.

[12] Robert E. Gaunt. Variance-Gamma approximation via Stein's method. *Electronic Journal of Probability*, 19(38):1–33, 2014. doi: 10.1214/EJP.v19-3020.

[13] Olivier Guédon and Emanuel Milman. Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures. *Geometric and Functional Analysis*, 21(5):1043–1068, 2011. doi: 10.1007/s00039-011-0136-5.

[14] Haotian Jiang, Yin Tat Lee, and Santosh S. Vempala. A generalized central limit conjecture for convex bodies. In *Geometric Aspects of Functional Analysis*, volume 2266 of *Lecture Notes in Mathematics*, pages 1–41. Springer, Cham, 2020. doi: 10.1007/978-3-030-46762-3_1.

[15] Ravi Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13:541–559, 1995. doi: 10.1007/BF02574061.

[16] Bo'az Klartag. A central limit theorem for convex sets. *Inventiones Mathematicae*, 168(1): 91–131, 2007. doi: 10.1007/s00222-006-0028-8.

[17] Bo'az Klartag. Power-law estimates for the central limit theorem for convex sets. *Journal of Functional Analysis*, 245(1):284–310, 2007. doi: 10.1016/j.jfa.2006.12.005.

[18] Bo'az Klartag. Logarithmically-concave moment measures I. In *Geometric Aspects of Functional Analysis*, volume 2116 of *Lecture Notes in Mathematics*, pages 231–260. Springer, Cham, 2014. doi: 10.1007/978-3-319-09477-9_16.

[19] Bo'az Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inveniendi Analytica*, 2023:Paper No. 4, 17 pp., 2023.

[20] Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and KLS isoperimetry up to polylog. *Geometric and Functional Analysis*, 32(5):1134–1159, 2022. doi: 10.1007/s00039-022-00612-9.

[21] Bo'az Klartag and Joseph Lehec. Isoperimetric inequalities in high-dimensional convex sets. *Bulletin of the American Mathematical Society*, 62(4):575–642, 2025. doi: 10.1090/bull/1869.

[22] Bo'az Klartag and Joseph Lehec. Thin-shell bounds via parallel coupling, 2025. URL https://arxiv.org/abs/2507.15495. arXiv:2507.15495, revised February 2026.

[23] Michel Ledoux, Ivan Nourdin, and Giovanni Peccati. Stein's method, logarithmic Sobolev and transport inequalities. *Geometric and Functional Analysis*, 25(1):256–306, 2015. doi: 10.1007/s00039-015-0312-0.

[24] Yin Tat Lee and Santosh S. Vempala. Eldan's stochastic localization and the KLS conjecture: isoperimetry, concentration and mixing. *Annals of Mathematics*, 199(3):1043–1092, 2024. doi: 10.4007/annals.2024.199.3.2.

[25] Brayden Letwin. The KLS constant is $O(\log^{1/4} n)$, 2026. URL https://arxiv.org/abs/2607.24164. arXiv:2607.24164v1, submitted 27 July 2026.

[26] László Lovász and Santosh S. Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures & Algorithms*, 30(3):307–358, 2007. doi: 10.1002/rsa.20135.

[27] Ivan Nourdin and Giovanni Peccati. The optimal fourth moment theorem. *Proceedings of the American Mathematical Society*, 143(7):3123–3133, 2015. doi: 10.1090/S0002-9939-2015-12417-3.

[28] Ivan Nourdin, Giovanni Peccati, and Gesine Reinert. Invariance principles for homogeneous sums: Universality of Gaussian Wiener chaos. *The Annals of Probability*, 38(5):1947–1985, 2010. doi: 10.1214/10-AOP531.

[29] Grigoris Paouris. Concentration of mass on convex bodies. *Geometric and Functional Analysis*, 16(5):1021–1049, 2006. doi: 10.1007/s00039-006-0584-5.