CRITICAL LOGARITHMIC SOBOLEV STABILITY: MOMENT BOUNDS AND HEAT FLOW

ABSTRACT. We study the transport first variation of the Gaussian logarithmic Sobolev deficit. At prescribed Fisher information and a second-moment cap, we determine the infimum of its squared norm and all attainment cases. Remainder identities identify the Gaussian limit of every minimizing sequence and its lost second moment. We also characterize the directional moment defects of sequences with vanishing first variation; their normalized escaping components need not be Gaussian. Although the instantaneous norm gives no dimension-free transport estimate, its integral along the Ornstein–Uhlenbeck flow controls Wasserstein distance, Fisher information, and the deficit with sharp constants. These estimates require no covariance or Poincaré assumption, and every centered isotropic Gaussian gives equality. No fixed observation interval suffices. Finally, we identify the Wasserstein metric slope at regular densities and extend the flow estimates to closed weighted manifolds with positive weighted Ricci curvature.

1. INTRODUCTION

A small logarithmic Sobolev deficit and a small first variation of that deficit impose different conditions on a probability density. The first variation involves third derivatives of the logarithm of the density. Its norm can vanish along a sequence even when a small amount of mass carries a fixed second moment to infinity. We determine the sharp moment constraint for this norm and show that observing it along a diffusion restores dimension-free stability without that constraint.

Let $\gamma_\kappa = N(0, \kappa^{-1}\mathrm{Id})$ on $\mathbb{R}^n$, with $\kappa > 0$. Write

$$d\pi = p(x)\,dx = e^{2h}\,d\gamma_\kappa, \qquad \int e^{2h}\,d\gamma_\kappa = 1, \tag{1.1}$$

and define b by

$$\Delta h - \kappa x \cdot \nabla h + |\nabla h|^2 + \kappa h + b(x) = 0. \tag{1.2}$$

Every smooth positive density determines h and b. Thus our estimates concern densities, rather than existence for a prescribed forcing. Unless stated otherwise, π is centered. Set

$$E = \int |\nabla h|^2\,d\pi, \qquad S = \int |x|^2\,d\pi, \qquad B = \int |\nabla b|^2\,d\pi. \tag{1.3}$$

The relative Fisher information is $I(\pi \mid \gamma_\kappa) = 4E$. For the logarithmic Sobolev deficit

$$\mathcal{D}_\kappa(\pi) = \frac{1}{2\kappa} I(\pi \mid \gamma_\kappa) - \mathrm{Ent}(\pi \mid \gamma_\kappa), \tag{1.4}$$

the standard first-variation formulas [1, 13, 17] give

$$\frac{d}{dt}\mathcal{D}_\kappa((\mathrm{Id} + t\xi)_\#\pi)\Big|_{t=0} = \int G_\pi \cdot \xi\,d\pi, \qquad G_\pi = \frac{2}{\kappa}\nabla b, \tag{1.5}$$

for $\xi \in C_c^\infty(\mathbb{R}^n; \mathbb{R}^n)$. In particular, $\|G_\pi\|_2^2 = 4B/\kappa^2$.

Theorem 1.1 (Optimal bound under a moment constraint)**.** *Fix $n \geq 1$, $\kappa > 0$, and $S_0 > 0$. Put*

$$e_{\min} = \frac{(n - \kappa S_0)_+^2}{4S_0}, \qquad e_* = \frac{\kappa^2}{4}\left(S_0 - \frac{n}{\kappa}\right). \tag{1.6}$$

2020 *Mathematics Subject Classification.* Primary 35B35, 26D10; Secondary 35J61, 49Q22, 60E15.

Key words and phrases. logarithmic Sobolev inequality, critical point, Fisher information, second moment, Gaussian mixture, stability.

Among smooth positive centered probability densities with $S \le S_0$, $E = e$, and $B < \infty$, the admissible set is nonempty exactly when $e \ge e_{\min}$, the classical Cramér–Rao threshold [8, 11]. For every such e,

$$\inf B = \frac{4}{n}(e - e_*)_+^2 \left(\kappa + \frac{4}{n}(e - e_*)_+\right). \tag{1.7}$$

If $S_0 \ge n/\kappa$, the infimum is attained only when $e = 0$, by γ_κ. If $S_0 < n/\kappa$, it is attained only when $e = e_{\min}$, by the centered Gaussian with covariance $(S_0/n)\mathrm{Id}$. In every other case there are minimizing sequences of finite Gaussian mixtures.

At the Gaussian second moment, the bound reads

$$S \le n/\kappa \quad \Longrightarrow \quad B \ge \frac{4\kappa}{n}E^2 + \frac{16}{n^2}E^3. \tag{1.8}$$

Above this moment, $0 < e \le e_*$ has zero infimum. Theorem 4.3 identifies the Gaussian limit and the lost second moment of every minimizing sequence. At $B_j \to 0$, Theorem 5.1 describes every directional moment defect by a finite measure on the sphere, and Proposition 5.2 shows why this does not determine the shapes of normalized escaping components. Theorem 3.3 gives the sharp bound at a prescribed information matrix.

The instantaneous estimate necessarily depends on dimension. Along the Ornstein–Uhlenbeck flow π_t, however, put $g(t) = \|G_{\pi_t}\|_2$. Theorem 7.1 proves

$$W_2(\pi, \gamma_\kappa) \le \int_0^\infty (e^{\kappa t} - 1) g(t)\, dt, \tag{1.9}$$

for every centered $\pi \in \mathcal{P}_2(\mathbb{R}^n)$, including singular measures. It gives companion bounds for information and deficit, with equality in all three estimates for every centered isotropic Gaussian. Corollary 7.3 shows that no fixed observation interval suffices. Theorem 9.1 identifies the actual metric slope at regular densities, and Corollary 9.2 extends the flow estimates to closed weighted manifolds.

Related work. Gross's logarithmic Sobolev inequality and Talagrand's transport inequality are basic inputs [14, 20]. The Cramér–Rao feasibility condition, score integration identities, and entropy–Fisher relation are classical [8, 19]. Bobkov–Gozlan–Roberto–Samson [4] and Eldan–Lehec–Shenfeld [11] explain their consequences for the deficit. In particular, covariance thresholds, matrix refinements, and Gaussian-mixture obstructions already occur in that literature. Our constrained problem concerns the squared first variation; its uniform third-derivative estimates, exact remainder, and minimizing-sequence limits are the additional assertions.

The diffusion identities and conditional-score formulas are standard [2, 15, 16]; higher conditional derivatives are treated in [10]. The estimates along the flow use these formulas to integrate the first-variation norm with sharp kernels. Unlike the deficit estimates of Fathi–Indrei–Ledoux [12], they assume no Poincaré inequality for the density. The global integration argument uses the logarithmic Bochner mechanism of Conrado [7], and the compactness proof invokes second-order Fisher-information regularity from Gianazza–Savaré–Toscani [13].

Wei–Wu [21] obtain Gaussian approximation for a different, H^{-1}, equation residual. Dovetta–Serra [9] prove attainment for the deficit divided by squared L^2 distance to the Gaussian cone. Section 8 distinguishes that quotient from minimization under our constraints. The metric-slope argument uses HWI [18], within the fourth-order information-flow framework of [13, 17]. Its spherical specialization concerns the slope of the sharp deficit, not a new spherical stability inequality [3, 5, 6].

2. Score identities and moment bounds

In Sections 2–6, p is smooth, strictly positive, and centered, and $E, S, B < \infty$. Unspecified integrals are over $\mathbb{R}^n$. We use the order on symmetric matrices and the Hilbert–Schmidt norm. Write $\Sigma = \int x x^T\, d\pi$ and $C = \int \nabla h \nabla h^T\, d\pi$.

Here φ_ℓ denotes the density of $N(0,\ell^{-1}\mathrm{Id})$. Substituting $p/\varphi_\kappa = e^{2h}$ in the standard entropy and Fisher first variations [1, 13, 17] gives

$$\frac{\delta \mathcal{D}_\kappa}{\delta p} = \frac{2}{\kappa} b - 1, \tag{2.1}$$

$$\nabla b = -\tfrac{1}{2}\nabla\Delta \log p - \tfrac{1}{2}(\nabla^2 \log p)\nabla \log p - \tfrac{\kappa}{2}\nabla \log p. \tag{2.2}$$

Transport differentiation gives (1.5); the absolute expression (2.2) also shows that B is translation invariant.

Write

$$s = \nabla \log p = 2\nabla h - \kappa x, \quad K = \nabla^2 \log p, \quad L = \Delta + s\cdot\nabla. \tag{2.3}$$

The operator L is symmetric for π. Define the absolute Fisher information matrix and its trace by

$$\mathcal{J} = \int s s^T \, d\pi, \qquad J = \operatorname{tr}\mathcal{J}. \tag{2.4}$$

Finiteness of E, S gives $J < \infty$. The classical score identities [8, 11] and the equation-specific balances below will be used together. Appendix A verifies Hessian integrability and all boundary limits under these assumptions.

Lemma 2.1 (Global identities)**.** *Under these assumptions, $\nabla^2 h$ and K belong to $L^2(\pi)$. Moreover,*

$$\int s\, d\pi = 0, \qquad \int x s^T \, d\pi = -\mathrm{Id}, \qquad \int K \, d\pi = -\mathcal{J}, \tag{2.5}$$

$$\int \nabla b \, d\pi = 0, \qquad \int |\nabla^2 h|^2 \, d\pi = \int \nabla h \cdot \nabla b \, d\pi, \tag{2.6}$$

$$2\int \nabla b \, s^T \, d\pi = Q - \kappa\mathcal{J}, \qquad Q = \int K^2 \, d\pi. \tag{2.7}$$

In particular, $\int \nabla h \, d\pi = 0$, and

$$\mathcal{J} = 4C + 2\kappa\mathrm{Id} - \kappa^2\Sigma, \qquad J = 4E + 2n\kappa - \kappa^2 S. \tag{2.8}$$

The information relation (2.8) is standard [11, Section 1.2]. At $B = 0$, (2.6) and centering give the classical rigidity conclusion [7]

$$B = 0 \quad \Longrightarrow \quad p = \varphi_\kappa. \tag{2.9}$$

2.1. **The scalar remainder.**

Proposition 2.2 (Exact scalar remainder)**.** *Put $\ell = J/n$. Then*

$$\begin{aligned} B - \frac{n}{4}\ell(\ell-\kappa)^2 = &\int \left|\nabla b - \frac{\ell-\kappa}{2} s\right|^2 d\pi \\ &+ \frac{\ell-\kappa}{2}\int |K + \ell\mathrm{Id}|^2 \, d\pi. \end{aligned} \tag{2.10}$$

Consequently, if $J \geq n\kappa$,

$$B \geq \frac{J(J-n\kappa)^2}{4n^2}. \tag{2.11}$$

If $J > n\kappa$, equality holds exactly for centered isotropic Gaussians of precision J/n. If $J = n\kappa$, equality holds exactly for γ_κ.

Proof. Taking the trace in (2.7) gives

$$2\int \nabla b \cdot s \, d\pi = \int |K|^2 \, d\pi - \kappa J.$$

Also, (2.5) gives $\int |K+\ell \mathrm{Id}|^2 d\pi = \int |K|^2 d\pi - J^2/n$. Expand the square in (2.10) and substitute these two equalities. This proves the identity, including the signed identity when $J < n\kappa$.

For $J > n\kappa$, both remainders are nonnegative and equality forces $K = -\ell \mathrm{Id}$ everywhere. Integrating the Hessian and using normalization and centering gives the Gaussian of precision ℓ. Its score is $s = -\ell x$, and (A.2) gives $\nabla b = (\ell - \kappa)s/2$, so it attains equality. At $J = n\kappa$, equality means $B = 0$, and (2.9) applies. □

Proof of the lower bound and feasibility in Theorem 1.1. Cramér–Rao gives $SJ \geq n^2$ [8]; see also [11, equation (12)]. With (2.8), this is

$$E \geq \frac{(n-\kappa S)^2}{4S}. \tag{2.12}$$

Minimization over $0 < S \leq S_0$ gives $e \geq e_{\min}$. Conversely, Gaussians with precision $\ell \geq \max\{\kappa, n/S_0\}$ have $S = n/\ell$ and $E = n(\ell-\kappa)^2/(4\ell)$, which ranges continuously from $e_{\min}$ to infinity.

Next, $S \leq S_0$ and (2.8) give

$$J \geq 4e + 2n\kappa - \kappa^2 S_0 = n\kappa + 4(e - e_*).$$

If $e > e_*$, this lower bound exceeds $n\kappa$. The function $j \mapsto j(j-n\kappa)^2$ is strictly increasing there. Proposition 2.2 therefore gives (1.7) as a lower bound. For $e \leq e_*$ the claimed lower bound is zero.

If the positive lower bound is attained, strict monotonicity forces $S = S_0$ and equality in (2.11). Thus π is an isotropic Gaussian of precision $\ell > \kappa$, with $S_0 = n/\ell < n/\kappa$. Its energy is exactly $e_{\min}$. Conversely this Gaussian attains the bound. If a zero lower bound is attained, (2.9) forces $e = 0$ and $S = n/\kappa \leq S_0$. The matching minimizing sequences are constructed in Section 3. □

For the actual second moment, (2.11) gives

$$\frac{4\kappa}{n} q^2 + \frac{16}{n^2} q^3 \leq B, \qquad q = \left[E - \frac{\kappa^2}{4} \left(S - \frac{n}{\kappa} \right)_+ \right]_+ . \tag{2.13}$$

For $q > 0$, the moment identity gives $J \geq n\kappa + 4q$; the case $q = 0$ is immediate. In particular,

$$S \leq n/\kappa \quad \Longrightarrow \quad E \leq \min\left\{ \frac{1}{2}\sqrt{\frac{nB}{\kappa}}, \left(\frac{n^2 B}{16}\right)^{1/3} \right\}. \tag{2.14}$$

3. Minimizing sequences with prescribed moments

We use the separated Gaussian mixtures familiar from deficit instability [11, Section 6]. To control B, we need an overlap estimate for third logarithmic derivatives that is uniform in the component weights.

3.1. **An overlap estimate.** Let A take finitely many values $a_1, \dots, a_m$ with positive probabilities $w_1, \dots, w_m$, and let Z be an independent centered Gaussian with covariance $\ell^{-1}\mathrm{Id}$. Set $X = A + Z$. Its density is

$$p(x) = \sum_{j=1}^m w_j \varphi_\ell(x - a_j), \qquad \varphi_\ell(x) = \left(\frac{\ell}{2\pi}\right)^{n/2} e^{-\ell|x|^2/2}.$$

At a fixed x, write

$$r_j(x) = \Pr(A = a_j \mid X = x), \quad \bar{a}(x) = \mathbb{E}[A \mid X = x], \quad V(x) = \mathrm{Cov}(A \mid X = x),$$

and put $\theta(x) = 1 - \sum_j r_j(x)^2$. The quantity θ is small when one component accounts for most of the density.

Lemma 3.1 (Separated Gaussian components)**.** *Fix m, n and a compact interval of positive precisions ℓ. Suppose $|a_j| \le R$ and $|a_i - a_j| \ge R$ for $i \ne j$, with $R \ge 1$. Uniformly over the positive component weights, there are constants $c, C > 0$ and an integer N depending only on these fixed parameters and κ such that*

$$\int (1+|x|^2)\theta(x)p(x)\,dx \le C(1+R^2)e^{-cR^2}, \tag{3.1}$$

$$\mathbb{E}\,|s(X)+\ell Z|^2 + \mathbb{E}\left|\nabla b(X) - \frac{\ell(\kappa-\ell)}{2}Z\right|^2 \le C(1+R^N)e^{-cR^2}. \tag{3.2}$$

Here b is defined by (1.2) for this density; centering is not needed for these estimates. Consequently,

$$\mathcal{J} \longrightarrow \ell\mathrm{Id}, \qquad B \longrightarrow \frac{n}{4}\ell(\ell-\kappa)^2 \tag{3.3}$$

as $R \to \infty$, uniformly in the weights and locally uniformly in ℓ.

Proof. Put $g_j = w_j\varphi_\ell(\cdot - a_j)$. Since $p \ge g_i + g_j$,

$$p\theta = 2\sum_{i<j}\frac{g_ig_j}{p} \le 2\sum_{i<j}\sqrt{g_ig_j}.$$

Completing the square gives the exact product formula

$$\sqrt{\varphi_\ell(x-a_i)\varphi_\ell(x-a_j)} = e^{-\ell|a_i-a_j|^2/8}\varphi_\ell\left(x - \frac{a_i+a_j}{2}\right).$$

Its integral against $1+|x|^2$ is at most $C(1+R^2)e^{-cR^2}$. Summing over the fixed number of pairs proves (3.1).

Let A' be an independent conditional copy of A given $X = x$. Then

$$\operatorname{tr} V(x) = \tfrac{1}{2}\mathbb{E}[|A-A'|^2 \mid X = x] \le 2R^2\theta(x).$$

Since V is positive semidefinite, $|V| \le \operatorname{tr} V$. For the contracted third central moment

$$T(x) = \mathbb{E}[(A-\bar a)|A-\bar a|^2 \mid X = x],$$

we also have $|T| \le 2R\operatorname{tr} V \le 4R^3\theta$.

The Gaussian conditional-derivative identities [10] give

$$s = -\ell(x-\bar a), \qquad K = -\ell\mathrm{Id} + \ell^2 V, \qquad \nabla\Delta h = \tfrac{1}{2}\ell^3 T. \tag{3.4}$$

By (A.1),

$$\begin{aligned}\nabla b &= -\nabla\Delta h - (\nabla^2 h)s \\ &= \frac{\ell(\kappa-\ell)}{2}(x-\bar a) + \frac{\ell^3}{2}[V(x-\bar a) - T].\end{aligned} \tag{3.5}$$

The second term has squared norm at most $C[R^4\theta^2|x|^2 + R^6\theta^2]$. Its integral tends to zero at the rate in (3.2), because $\theta^2 \le \theta$ and (3.1) holds.

Under the joint law of (A, Z),

$$\mathbb{E}|A - \bar a(X)|^2 = \int \operatorname{tr} V\, p\,dx \le 2R^2\int \theta p\,dx.$$

Since $X - \bar a(X) - Z = A - \bar a(X)$, this proves both comparisons in (3.2). The reference vectors $-\ell Z$ and $\ell(\kappa-\ell)Z/2$ have bounded second moments, with covariance $\ell\mathrm{Id}$ and squared norm expectation $n\ell(\ell-\kappa)^2/4$. Cauchy–Schwarz now gives (3.3). □

3.2. **Exact covariance and the lower envelope.** Fix a total second moment $T > 0$ and a precision $\ell \geq n/T$. For R large, define

$$p_{R,\ell,T}(x) = \left(1 - \frac{T - n/\ell}{R^2}\right)\varphi_\ell(x) + \frac{T - n/\ell}{2nR^2}\sum_{i=1}^{n}[\varphi_\ell(x - Re_i) + \varphi_\ell(x + Re_i)]. \tag{3.6}$$

If $T = n/\ell$, this means the single Gaussian φ_ℓ. Otherwise all the weights are positive once $R^2 > T - n/\ell$. The density is centered and

$$\Sigma_{R,\ell,T} = \frac{T}{n}\mathrm{Id} \tag{3.7}$$

for every R. Its score and all derivatives grow at most polynomially in x, so E, S, B are finite. Lemma 3.1 and (2.8) show that

$$E(p_{R,\ell,T}) \longrightarrow \frac{n\ell - 2n\kappa + \kappa^2 T}{4}, \tag{3.8}$$

$$B(p_{R,\ell,T}) \longrightarrow \frac{n}{4}\ell(\ell - \kappa)^2. \tag{3.9}$$

The convergence is locally uniform in ℓ on any interval with $\ell > n/T$.

Completion of the proof of Theorem 1.1. The Gaussian attainment cases were already proved. Consider first $e > e_*$, outside those cases, and choose

$$T = S_0, \qquad \ell_0 = \kappa + \frac{4}{n}(e - e_*).$$

The feasibility condition $e \geq e_{\min}$ implies $\ell_0 \geq n/T$. Equality occurs only in the already treated Gaussian boundary case, so here $\ell_0 > n/T$. Formulas (3.8)–(3.9) give energy tending to e and forcing tending to the right side of (1.7).

For $0 < e \leq e_*$, necessarily $S_0 > n/\kappa$. Choose instead

$$T = \frac{n}{\kappa} + \frac{4e}{\kappa^2} \leq S_0, \qquad \ell_0 = \kappa > n/T.$$

The same formulas give energy tending to e and $B \to 0$.

We must still arrange *exactly* $E = e$, not merely convergence to it. This can be done without altering T. For $X = A + Z$, the standard conditional-score identity [10, 16] gives

$$J(p_{R,\ell,T}) = n\ell - \ell^2\mathbb{E}[\mathrm{tr}\,\mathrm{Cov}(A \mid X)]. \tag{3.10}$$

For finite R and $\ell > n/T$, the last expectation is strictly positive because distinct components have positive densities everywhere. Thus $E(p_{R,\ell_0,T}) < e$.

For every sufficiently small fixed $\varepsilon > 0$, locally uniform convergence in (3.8) gives $E(p_{R,\ell_0+\varepsilon,T}) > e$ for all sufficiently large R. At fixed R, the energy is continuous in ℓ; this follows directly from the finite mixture formula and a locally uniform Gaussian integrable bound. The intermediate value theorem gives a root $\ell_R \in (\ell_0, \ell_0 + \varepsilon)$ with $E = e$. Choosing $\varepsilon \downarrow 0$ along a sequence yields $R_j \to \infty$ and $\ell_j \to \ell_0$, with exactly $E = e$ and $S = T \leq S_0$. The convergence in (3.9) is unchanged. This proves both sharpness and nonemptiness at every admissible e. □

These densities converge in total variation to φ_{ℓ_0}. Their fixed second moment T exceeds the limiting Gaussian moment by $T - n/\ell_0$, carried by components of total mass $(T - n/\ell_0)/R^2$.

3.3. **A prescribed information matrix.** Let $\mathcal{J}, Q$ be as in Lemma 2.1. The matrix

$$\mathcal{R} = Q - \mathcal{J}^2 = \int (K + \mathcal{J})^2\, d\pi \tag{3.11}$$

is positive semidefinite. The identity follows from $\int K d\pi = -\mathcal{J}$; no commutation of $K(x)$ and $\mathcal{J}$ is assumed.

Proposition 3.2 (Exact matrix remainder)**.** *With $M = \mathcal{J} - \kappa\mathrm{Id}$,*

$$B - \tfrac{1}{4}\operatorname{tr}(\mathcal{J}M^2) = \int \left|\nabla b - \tfrac{1}{2}Ms\right|^2 d\pi + \tfrac{1}{2}\operatorname{tr}(M\mathcal{R}). \tag{3.12}$$

In particular, if $\mathcal{J} \succeq \kappa\mathrm{Id}$, then $B \geq \frac{1}{4}\operatorname{tr}[\mathcal{J}(\mathcal{J} - \kappa\mathrm{Id})^2]$.

Proof. Expand the square and use $2\int \nabla b\, s^T d\pi = Q - \kappa\mathcal{J}$. The result is

$$\int \left|\nabla b - \tfrac{1}{2}Ms\right|^2 d\pi = B - \tfrac{1}{2}\operatorname{tr}[M(Q - \kappa\mathcal{J})] + \tfrac{1}{4}\operatorname{tr}(M^2\mathcal{J}).$$

Add $\frac{1}{2}\operatorname{tr}[M(Q - \mathcal{J}^2)]$ and use $M = \mathcal{J} - \kappa\mathrm{Id}$. This proves (3.12). If $M, \mathcal{R}$ are positive semidefinite, then $\operatorname{tr}(M\mathcal{R}) = \operatorname{tr}(M^{1/2}\mathcal{R}M^{1/2}) \geq 0$. □

Theorem 3.3 (Optimal bound at a prescribed information matrix)**.** *If $\Sigma \preceq \kappa^{-1}\mathrm{Id}$, then*

$$B \geq 4\kappa\operatorname{tr}(C^2) + 16\operatorname{tr}(C^3). \tag{3.13}$$

For every symmetric positive semidefinite matrix C_0, the right side at C_0 is the infimum of B among densities with $C = C_0$ and $\Sigma \preceq \kappa^{-1}\mathrm{Id}$. It is also the infimum when $\Sigma = \kappa^{-1}\mathrm{Id}$ is required. The infimum is attained only for $C_0 = 0$, by γ_κ.

Proof. Equation (2.8) gives $\mathcal{J} \succeq \kappa\mathrm{Id} + 4C \succeq \kappa\mathrm{Id}$. The ordered eigenvalues of $\mathcal{J}$ therefore dominate those of $\kappa\mathrm{Id} + 4C$. The scalar function $t(t - \kappa)^2$ is strictly increasing on $[\kappa, \infty)$. Applying Proposition 3.2 eigenvalue by eigenvalue gives

$$B \geq \tfrac{1}{4}\operatorname{tr}[\mathcal{J}(\mathcal{J} - \kappa\mathrm{Id})^2] \geq \tfrac{1}{4}\operatorname{tr}[(\kappa\mathrm{Id} + 4C)(4C)^2],$$

which is (3.13). This argument uses eigenvalue monotonicity, not operator monotonicity of a cubic polynomial.

Diagonalize C_0, with eigenvalues $c_1, \ldots, c_n \geq 0$. For each $c_i > 0$, apply the one-dimensional construction in Theorem 1.1 with $S_0 = 1/\kappa$ and exactly $E = c_i$. It gives centered densities of variance $1/\kappa$ and forcing tending to $4\kappa c_i^2 + 16c_i^3$. For $c_i = 0$, use the Gaussian of variance $1/\kappa$. Take their product and rotate back to the eigenbasis of C_0. The information matrices are diagonal before rotation because the one-dimensional gradients have mean zero. The forcing potentials add, so B is the sum of the component values. This proves sharpness with exactly $C = C_0$ and exactly Gaussian covariance.

Suppose equality holds and $C \neq 0$. Strict eigenvalue monotonicity forces equality of all eigenvalues in the second inequality above. Since $\mathcal{J} - (\kappa\mathrm{Id} + 4C)$ is positive semidefinite, equality of their traces forces $\mathcal{J} = \kappa\mathrm{Id} + 4C$ and hence $\Sigma = \kappa^{-1}\mathrm{Id}$. Also $\operatorname{tr}(M\mathcal{R}) = 0$. Choose a unit eigenvector v of C with eigenvalue $c > 0$. Then

$$0 = v^T\mathcal{R}v = \int |(K + \mathcal{J})v|^2\, d\pi, \qquad Kv = -(\kappa + 4c)v$$

everywhere. In coordinates with first axis v, the mixed derivatives of $\log p$ with that axis vanish and its second derivative is $-(\kappa + 4c)$. Hence the marginal along v is Gaussian with variance $(\kappa + 4c)^{-1}$. This contradicts $v^T\Sigma v = \kappa^{-1}$. For $C = 0$, $E = 0$ and p is the Gaussian. □

Unlike (1.8), the matrix bound is unchanged on adjoining independent Gaussian coordinates.

4. Limits of all minimizing sequences

We now identify every minimizing-sequence limit, using standard Fisher-information regularity.

Lemma 4.1 (Compactness with a prescribed mean Hessian)**.** *Let p_j be smooth positive centered probability densities with uniformly bounded second moments and absolute Fisher information. Suppose, for some $\ell > 0$,*

$$\int |\nabla^2 \log p_j + \ell\mathrm{Id}|^2 p_j\, dx \longrightarrow 0. \tag{4.1}$$

Then

$$p_j \longrightarrow \varphi_\ell \quad \text{in } L^1(\mathbb{R}^n), \qquad \sqrt{p_j} \longrightarrow \sqrt{\varphi_\ell} \quad \text{in } H^1(\mathbb{R}^n). \tag{4.2}$$

No uniform integrability of the second moments is asserted.

Proof. Write $K_j = \nabla^2 \log p_j$ and $u_j = \sqrt{p_j}$. The local second-order information estimate of Gianazza–Savaré–Toscani [13], rescaled from $B_1 \subset B_2$, gives

$$\int_{B_R} (|\nabla^2 \sqrt{p}|^2 + |\nabla p^{1/4}|^4)\, dx \le C_n \left(\int_{B_{2R}} |\nabla^2 \log p|^2 p\, dx + R^{-2} J \right). \tag{4.3}$$

To check the scaling, apply the local estimate to $R^n p(R\,\cdot)$: the second-order terms scale as R^4, and Fisher information as R^2. The estimate applies here because p is smooth and positive, $J < \infty$, and its weighted logarithmic Hessian is square integrable. Letting $R \to \infty$ and using (4.1) gives a uniform H^2 bound for u_j.

Rellich compactness gives a subsequence converging in H^1 on every bounded ball to some $u \ge 0$. The moment bound gives $\int_{|x|>R} u_j^2 dx \le C/R^2$, hence global L^2 convergence and $\|u\|_2 = 1$. The interpolation estimate

$$\|\nabla(u_j - u)\|_2^2 \le \|u_j - u\|_2 \, \|\Delta(u_j - u)\|_2$$

then gives global H^1 convergence. In particular $p_j \to p = u^2$ in L^1. Bounded second moments make first moments uniformly integrable, so p is centered.

It remains to identify p, including the possibility that it vanishes. The identities

$$\nabla^2 p_j = 4 \nabla u_j \nabla u_j^T + p_j K_j$$

and (4.1) show that the right sides converge in L^1 to $4\nabla u \nabla u^T - \ell p \mathrm{Id}$. Therefore $p \in W^{2,1}(\mathbb{R}^n)$ and

$$\nabla^2 p = 4 \nabla u \nabla u^T - \ell p \mathrm{Id} \tag{4.4}$$

in distributions and almost everywhere.

For almost every line parallel to a coordinate axis, Sobolev slicing makes p a $W^{2,1}_{\mathrm{loc}}$ function and hence a continuously differentiable nonnegative function on that line. On each interval where it is positive, (4.4) and $p = u^2$ give

$$p'' = (p')^2/p - \ell p, \qquad (\log p)'' = -\ell.$$

Thus it is an exponential of a quadratic polynomial on that interval. Such a function has a positive limit at every finite endpoint, so an interval of positivity cannot have a finite endpoint. Each such slice is consequently either identically zero or positive on the whole line.

The indicator of $\{p > 0\}$ is therefore constant on almost every coordinate line. Its distributional gradient vanishes, so it is constant almost everywhere on $\mathbb{R}^n$. Since $\int p = 1$, $p > 0$ almost everywhere. The measurable function $F = \log p + \ell |x|^2/2$ is affine on almost every coordinate line. It follows by induction that F agrees almost everywhere with a polynomial of degree at most one in each coordinate: choose two valid values of the first coordinate, express its slope and intercept by the corresponding sections, and repeat in the remaining coordinates.

Apply the same slicing argument to the finitely many directions $e_i + e_j$. In these directions $(\log p)'' = -\ell |e_i + e_j|^2$, so the polynomial F has zero second derivative in each of them. Its pure coordinate second derivatives already vanish, hence all mixed second derivatives vanish as well. Thus F is affine. Normalization and centering give $p = \varphi_\ell$.

Every convergent subsequence has this same limit. The compactness argument therefore proves (4.2) for the full sequence. □

Lemma 4.2 (Transport and a lost second moment)**.** *Suppose $p_j \to \varphi_\ell$ in $L^1(\mathbb{R}^n)$, $\ell > 0$, and $\int |x|^2 p_j(x) dx \to T < \infty$. Then*

$$W_2^2(p_j\, dx, \gamma_\kappa) \longrightarrow T + \frac{n}{\kappa} - \frac{2n}{\sqrt{\kappa \ell}}. \tag{4.5}$$

Proof. Take optimal couplings (X_j, Y_j) with $X_j \sim p_j$ and $Y_j \sim \gamma_\kappa$. Any weak subsequential limit couples the Gaussians of precisions ℓ and κ. Their cross moments pass to the limit. Indeed, with $T_j = \mathbb{E}|X_j|^2$ uniformly bounded,

$$\mathbb{E}[|X_j||Y_j|\mathbf{1}_{\{|Y_j|>K\}}] \le \sqrt{T_j}\left(\int_{|y|>K}|y|^2 d\gamma_\kappa\right)^{1/2},$$

whereas

$$\mathbb{E}[|X_j||Y_j|\mathbf{1}_{\{|Y_j|\le K,\ |X_j|>M\}}] \le KT_j/M.$$

With continuous cutoffs the truncated integrands are bounded and continuous, so weak convergence applies. Every limit coupling has cross moment at most $n/\sqrt{\kappa\ell}$ by Cauchy–Schwarz. Since the cost equals $T_j + n/\kappa - 2\mathbb{E}[X_j \cdot Y_j]$, this proves the lower limit in (4.5).

For the matching upper bound, couple $X_j \sim p_j$ and $Z \sim \varphi_\ell$ by their common density $\min(p_j, \varphi_\ell)$, putting $X_j = Z$ on that part. Its mass tends to one. On the remaining part, Cauchy–Schwarz bounds the absolute cross moment by the product of a bounded second-moment factor for X_j and a factor tending to zero for Z. The latter tends to zero because the residual density of Z is bounded by the fixed Gaussian density and has mass tending to zero. Thus $\mathbb{E}[X_j \cdot Z] \to n/\ell$. Taking $Y = \sqrt{\ell/\kappa}Z$ gives the required upper limit. □

Theorem 4.3 (Gaussian limit and lost second moment)**.** *Fix n, κ, S_0, and an admissible e as in Theorem 1.1. Let p_j be smooth positive centered densities with $S_j \le S_0$, $E_j \to e$, and finite B_j tending to the infimum (1.7). Put*

$$\ell_* = \kappa + \frac{4}{n}(e - e_*)_+, \qquad T_* = \min\left\{S_0, \frac{n}{\kappa} + \frac{4e}{\kappa^2}\right\}. \tag{4.6}$$

Then

$$\begin{aligned} &p_j \longrightarrow \varphi_{\ell_*} \quad \text{in } L^1, && \sqrt{p_j} \longrightarrow \sqrt{\varphi_{\ell_*}} \quad \text{in } H^1, \\ &S_j \longrightarrow T_*, && J_j \longrightarrow n\ell_*. \end{aligned} \tag{4.7}$$

In particular, the lost second moment is exactly $T_* - n/\ell_*$. Moreover,

$$W_2^2(p_j\,dx, \gamma_\kappa) \longrightarrow T_* + \frac{n}{\kappa} - \frac{2n}{\sqrt{\kappa\ell_*}}. \tag{4.8}$$

Proof. If $e > e_*$, the lower bound on J_j in Section 2 and strict monotonicity in (2.11) force $J_j \to n\ell_* > n\kappa$. The right side of (2.10) tends to zero. Since its second coefficient stays positive, it follows that $\int |K_j + \ell_* \mathrm{Id}|^2 p_j \to 0$.

If $e \le e_*$, the limiting infimum is zero. Equation (2.6) gives $\int |\nabla^2 h_j|^2 d\pi_j \le \sqrt{E_j B_j} \to 0$. Since $K_j + \kappa \mathrm{Id} = 2\nabla^2 h_j$, the same weighted Hessian conclusion holds with $\ell_* = \kappa$. Integrating K_j then gives $\mathcal{J}_j \to \kappa \mathrm{Id}$.

In both cases Lemma 4.1 applies; the Fisher information is bounded by (2.8). That identity also gives $S_j \to T_*$. This proves (4.7).

The transport conclusion follows from Lemma 4.2 with $\ell = \ell_*$ and $T = T_*$. □

In every nonattainment case $T_* > n/\ell_*$, so the second moments are not uniformly integrable. The next section describes their directional defect.

5. Second-moment defects and escaping profiles

At $B_j \to 0$, bounded second moments give a Gaussian limit. The remaining moment may escape in arbitrary directions, without determining the shape of each escaping component.

Theorem 5.1 (All second-moment defects at small forcing)**.** *Let p_j be smooth positive centered densities with $E_j < \infty$, $\sup_j S_j < \infty$, and $B_j \to 0$. After passage to a subsequence, there is a finite nonnegative Borel measure ν on $\mathbb{S}^{n-1}$ such that*

$$p_j \to \varphi_\kappa \text{ in } L^1, \qquad \sqrt{p_j} \to \sqrt{\varphi_\kappa} \text{ in } H^1, \tag{5.1}$$

and, for every continuous function a on $\mathbb{S}^{n-1}$,

$$\lim_j \int |x|^2 a(x/|x|) p_j(x)\,dx = \int |x|^2 a(x/|x|)\,d\gamma_\kappa(x) + \int a\,d\nu. \tag{5.2}$$

The value of the integrands at $x = 0$ is defined to be zero. In particular, with $M = \int \theta\theta^T\,d\nu(\theta)$,

$$\Sigma_j \to \kappa^{-1}\mathrm{Id} + M, \qquad C_j \to \frac{\kappa^2}{4} M, \qquad E_j \to \frac{\kappa^2}{4}\nu(\mathbb{S}^{n-1}). \tag{5.3}$$

Also $W_2^2(\pi_j, \gamma_\kappa) \to \nu(\mathbb{S}^{n-1})$. Every finite nonnegative ν occurs for a sequence of finite Gaussian mixtures. In dimensions $n \geq 2$ the defect need not be atomic; in every dimension its mass has no quantization.

Proof. Proposition 2.2 bounds J_j uniformly whenever $J_j \geq n\kappa$; the complementary case is already bounded. The moment identity (2.8) bounds E_j, so (2.6) gives

$$\int |K_j + \kappa\mathrm{Id}|^2 d\pi_j = 4\int |\nabla^2 h_j|^2 d\pi_j \leq 4\sqrt{E_j B_j} \longrightarrow 0.$$

Lemma 4.1 proves (5.1). Define finite measures on the sphere by

$$\lambda_j(A) = \int_{\{x\neq 0:\ x/|x|\in A\}} |x|^2 p_j(x)\,dx.$$

Their masses are bounded, so a subsequence converges weakly to λ. For nonnegative continuous a, first insert a compactly supported radial cutoff in the integral. Local L^1 convergence identifies its limit. Letting the cutoff radius tend to infinity shows that $\lambda \geq \lambda_{\gamma_\kappa}$ as measures. Set $\nu = \lambda - \lambda_{\gamma_\kappa}$. This proves (5.2). Strong H^1 convergence of square roots gives $\mathcal{J}_j = 4\int \nabla\sqrt{p_j}\nabla\sqrt{p_j}^T\,dx \to \kappa\mathrm{Id}$. Taking $a(\theta) = \theta_i\theta_k$ in (5.2) and using (2.8) yields (5.3). Lemma 4.2 gives the transport limit.

For the converse, first let $\nu = \sum_{i=1}^m a_i\delta_{\theta_i}$ with distinct directions and $a_i > 0$. For $R^2 > \sum_i a_i$, take

$$\widetilde{p}_R = \left(1 - \frac{\sum_i a_i}{R^2}\right)\varphi_\kappa + \sum_{i=1}^m \frac{a_i}{R^2}\varphi_\kappa(\,\cdot - R\theta_i).$$

Its mean is $m_R = R^{-1}\sum_i a_i\theta_i$. Center it by subtracting m_R. For this fixed configuration, the separations are bounded below by a positive multiple of R; the proof of Lemma 3.1 therefore gives $B_R \to 0$. Translation does not change B, by (2.2). Each satellite contributes $a_i\delta_{\theta_i}$ to the limiting angular second moment. For a smooth test a, the function $x \mapsto |x|^2 a(x/|x|)$ has gradient bounded by $C|x|$, so centering changes its integral by $o(1)$. Uniform approximation of continuous a by smooth tests, using the bounded second moments, gives the same conclusion for every continuous test. Finally approximate an arbitrary finite ν weakly by finite atomic measures with bounded masses, and choose R successively large enough for each finite configuration. A diagonal sequence proves the assertion. □

The moment description does not force normalized escaping profiles to be Gaussian.

Proposition 5.2 (Non-Gaussian normalized escaping profiles)**.** *Let $q = Z_v^{-1}e^v\varphi_\kappa$ be a centered non-Gaussian density, with $v \in C_c^\infty(\mathbb{R}^n)$. Fix $a > 0$ and set $\varepsilon_R = a/R^2$. Then*

$$p_R = (1 - \varepsilon_R)\varphi_\kappa + \frac{\varepsilon_R}{2}\big(q(\,\cdot - Re_1) + q(\,\cdot + Re_1)\big) \tag{5.4}$$

is centered, has bounded second moment, and satisfies

$$B_R = \varepsilon_R B(q) + o(\varepsilon_R), \qquad E_R \longrightarrow \frac{\kappa^2 a}{4}, \qquad \frac{p_R(Re_1 + \cdot)}{\varepsilon_R/2} \longrightarrow q \quad \text{in } C^3_{\mathrm{loc}}. \tag{5.5}$$

Here $B(q)$ is defined by the same critical equation. It is strictly positive.

Proof. Centered non-Gaussian examples of the stated class exist: take a nonconstant even compactly supported v. The moment calculation is exact:

$$S_R = \frac{n}{\kappa} + a + \varepsilon_R \left(\int |x|^2 q(x)\, dx - \frac{n}{\kappa} \right).$$

We record the derivative estimate needed for B_R. Let the three component densities be $q_i(x - a_i)$, where $q_0 = \varphi_\kappa$, $q_1 = q_2 = q$, and $a_i = 0, Re_1, -Re_1$. With weights w_i, let $r_i = w_i q_i(x - a_i)/p_R(x)$. Each q_i is bounded above and below by fixed positive multiples of φ_κ. Its logarithmic derivatives through order three grow at most linearly. Write $s_i = \nabla \log q_i(x - a_i)$, $K_i = \nabla^2 \log q_i(x - a_i)$, and $T_i = \nabla^3 \log q_i(x - a_i)$. The mixture score s, Hessian K, and third derivative T satisfy

$$\begin{aligned} s &= \sum_i r_i s_i, \\ K &= \sum_i r_i K_i + \sum_i r_i (s_i - s)^{\otimes 2}, \\ T &= \sum_i r_i T_i + \sum_i r_i \operatorname{Sym}_3(K_i \otimes (s_i - s)) + \sum_i r_i (s_i - s)^{\otimes 3}. \end{aligned}$$

Here Sym_3 is the sum over the three placements of the vector index. These formulas follow from $\nabla r_i = r_i(s_i - s)$. They are polynomials in the posterior weights and component derivatives, and at a vertex $r_i = 1$ they reduce to the derivatives of that component. Consequently, for some fixed integer N,

$$|\nabla b_R(x) - \nabla b(q_i)(x - a_i)| \le C(1 + |x| + R)^N (1 - r_i(x)).$$

This follows directly by subtracting the polynomial at the vertex $r_i = 1$ of the probability simplex; its coefficients have the displayed polynomial bound. Moreover,

$$p_R r_i r_k \le \tfrac{1}{2} \sqrt{w_i w_k q_i(x - a_i) q_k(x - a_k)} \qquad (i \ne k).$$

The Gaussian overlap calculation used in Lemma 3.1 thus gives

$$\sum_i w_i \int q_i(x - a_i) |\nabla b_R(x) - \nabla b(q_i)(x - a_i)|^2\, dx \le C(1 + R^{N'}) e^{-cR^2}.$$

The estimate is uniform in the three weights. Since $B(q_0) = 0$, the triangle inequality in the joint label–observation L^2 space proves $B_R = \varepsilon_R B(q) + o(\varepsilon_R)$. The same argument for the score gives $J_R = (1 - \varepsilon_R) n\kappa + \varepsilon_R J(q) + o(\varepsilon_R)$. Together with the exact moment formula and (2.8), this proves the energy limit. On every compact set around Re_1, the other two densities and their first three derivatives are exponentially small relative to ε_R. This proves the local profile limit. Finally $B(q) > 0$ by (2.9). □

The weighted norm vanishes because the component mass does, not because its normalized residual vanishes. This is compatible with an unweighted H^1 approximation that discards these components, and with the different residual used by Wei–Wu [21].

6. Transport consequences and optimal rates

For a centered density in our class, Gaussian logarithmic Sobolev and transport inequalities [14,20] give

$$\mathrm{Ent}(\pi \mid \gamma_\kappa) \le \frac{2E}{\kappa}, \qquad W_2^2(\pi,\gamma_\kappa) \le \frac{4E}{\kappa^2}. \tag{6.1}$$

They apply because $e^h \in W^{1,2}(\gamma_\kappa)$ and $S < \infty$; in particular, the relative entropy is finite.

Corollary 6.1 (Fourth-power transport bound)**.** *If $S \le n/\kappa$ and $W = W_2(\pi,\gamma_\kappa)$, then*

$$B \ge \frac{\kappa^5}{4n} W^4 + \frac{\kappa^6}{4n^2} W^6. \tag{6.2}$$

For fixed n, κ, the leading constant is sharp:

$$\lim_{\varepsilon \downarrow 0} \inf_{\substack{\Sigma = \kappa^{-1}\mathrm{Id} \\ 0 < W_2(\pi,\gamma_\kappa) \le \varepsilon}} \frac{B}{W_2^4(\pi,\gamma_\kappa)} = \frac{\kappa^5}{4n}. \tag{6.3}$$

The infimum is over the smooth positive centered densities with finite E, S, B used above.

The sixth-power term is a valid global consequence. We claim optimality of the leading fourth-power coefficient, not optimality of the entire right side at a prescribed nonzero transport distance.

Proof of Corollary 6.1. Use $E \ge \kappa^2 W^2/4$ in (1.8). To prove optimality, take the minimizing family at $S_0 = n/\kappa$ and energy $e > 0$. Its limiting precision is $\ell = \kappa + 4e/n$. Lemma 4.2 gives

$$W_2^2 \longrightarrow \frac{2n}{\kappa}\left(1 - \sqrt{\frac{\kappa}{\kappa + 4e/n}}\right), \qquad B \longrightarrow \frac{4\kappa}{n} e^2 + \frac{16}{n^2} e^3. \tag{6.4}$$

The first limit is $4e/\kappa^2 + O(e^2)$ as $e \downarrow 0$. Choose a diagonal sequence with the mixture errors smaller than $o(e^2)$ in B and $o(e)$ in W_2^2. The ratio tends to $\kappa^5/(4n)$. All these densities have covariance exactly $\kappa^{-1}\mathrm{Id}$. □

Corollary 6.2 (Failure beyond the moment threshold)**.** *For every $S_0 > n/\kappa$ and $0 < e \le e_*$, there are smooth positive centered densities with exactly $E = e$, with $S = n/\kappa + 4e/\kappa^2 \le S_0$, and with*

$$B \longrightarrow 0, \qquad W_2^2(\pi,\gamma_\kappa) \longrightarrow \frac{4e}{\kappa^2}.$$

Moreover, for every fixed $e > 0$, there is a sequence of dimensions $n \to \infty$ and densities with $E = e$, $\Sigma = \kappa^{-1}\mathrm{Id}$, such that

$$B \longrightarrow 0, \qquad W_2^2(\pi,\gamma_\kappa) \longrightarrow \frac{4e}{\kappa^2}.$$

Thus a bound depending only on the scalar B cannot be both dimension-free and vanish with B, even under exact Gaussian covariance.

Proof. For the first assertion use the low-energy construction, whose precision tends to κ, and apply Lemma 4.2. For the second use (6.4) at fixed e and let $n \to \infty$, choosing the component separation sufficiently large at each dimension. The limiting forcing is $4\kappa e^2/n + 16e^3/n^2 \to 0$, and the limiting transport cost tends to $4e/\kappa^2$. □

The analogous covariance and dimension obstructions for the deficit itself are established in [11]; here they hold even as the first variation vanishes.

For a noncentered π, center it and compare with $\gamma_{\kappa,m} = N(m, \kappa^{-1}\mathrm{Id})$, where $m = \int x\, d\pi$. The relevant energy and moment are $I(\pi \mid \gamma_{\kappa,m})/4$ and $\int |x-m|^2 d\pi$; B is unchanged. Moreover,

$$W_2^2(\pi,\gamma_{\kappa,a}) = W_2^2(\pi,\gamma_{\kappa,m}) + |m-a|^2.$$

7. Dimension-free estimates along the Ornstein–Uhlenbeck flow

For a centered probability measure $\pi \in \mathcal{P}_2(\mathbb{R}^n)$, let

$$\pi_t = \mathrm{Law}\left(e^{-\kappa t}X + \sqrt{\frac{1-e^{-2\kappa t}}{\kappa}}\,Z\right), \qquad X \sim \pi, \quad Z \sim N(0,\mathrm{Id}), \quad X \perp Z. \tag{7.1}$$

For $t > 0$, write $p_t = e^{2h_t}\varphi_\kappa$, define b_t by (1.2), and put

$$r_t = \nabla \log(p_t/\varphi_\kappa), \qquad G_t = \frac{2}{\kappa}\nabla b_t, \qquad g(t) = \|G_t\|_{L^2(\pi_t)}. \tag{7.2}$$

The measure at time zero need not have a density.

Theorem 7.1 (Estimates from the forcing along the flow)**.** *For every centered $\pi \in \mathcal{P}_2(\mathbb{R}^n)$,*

$$W_2(\pi,\gamma_\kappa) \le \int_0^\infty (e^{\kappa t}-1)g(t)\,dt. \tag{7.3}$$

If $I = I(\pi \mid \gamma_\kappa) < \infty$, then also

$$\sqrt{I} \le \kappa \int_0^\infty e^{\kappa t} g(t)\,dt, \tag{7.4}$$

$$\mathcal{D}_\kappa(\pi) \le \frac{\kappa}{2}\int_0^\infty\int_0^\infty e^{\kappa|t-s|}g(t)g(s)\,dt\,ds. \tag{7.5}$$

All integrals are nonnegative extended integrals. The constants in these three inequalities are sharp. Each inequality is equality for every centered isotropic Gaussian, in every dimension.

Centering is necessary: translated Gaussians have $G_t = 0$. For a measure of mean m, the same right sides control distance to $N(m,\kappa^{-1}\mathrm{Id})$, $\sqrt{I(\pi \mid \gamma_\kappa) - \kappa^2|m|^2}$, and the unchanged deficit. Translation along the flow subtracts $e^{-\kappa t}m$ and leaves $g(t)$ unchanged.

The centered endpoint below uses only a second moment.

Lemma 7.2 (Centered endpoint)**.** *For π as in (7.1),*

$$\lim_{t\to\infty} e^{\kappa t}\sqrt{I(\pi_t \mid \gamma_\kappa)} = 0. \tag{7.6}$$

At every positive time, the quantities E_t, S_t, B_t in (1.3) are finite.

Proof. Set $\tau = e^{-\kappa t}$, $\sigma^2 = (1-\tau^2)/\kappa$, and $Y = \tau X + \sigma Z$. The conditional-score representation used in [16, Section 3], rescaled to curvature κ, gives

$$r_t(Y) = \frac{\tau}{\sigma^2}\big(\mathbb{E}[X \mid Y] - \tau Y\big). \tag{7.7}$$

We claim that $\mathbb{E}[X \mid Y] \to 0$ in L^2 as $\tau \to 0$. Indeed, the joint law of (X,Y) converges in total variation to the product of the law of X and $N(0,\kappa^{-1}\mathrm{Id})$. For each fixed x, the conditional Gaussian density converges in L^1; integration in x is justified by the bound 2 on its L^1 distance. The same holds with the actual marginal law of Y in the product. Thus, for every bounded centered function $f(X)$,

$$\mathbb{E}\big|\mathbb{E}[f(X) \mid Y]\big| \longrightarrow 0.$$

The conditional expectation is uniformly bounded, so convergence also holds in L^2. Approximate each coordinate of X in L^2 by a bounded function and use contraction of conditional expectation. Since $\mathbb{E}X = 0$, this proves the claim. Also $\tau Y \to 0$ in L^2. Dividing (7.7) by τ proves (7.6).

The Gaussian conditional-derivative formulas [10, 16] also give

$$\nabla \log p_t(Y) = -\sigma^{-1}\mathbb{E}[Z \mid Y], \qquad \nabla^2 \log p_t(Y) = \sigma^{-2}\big(\mathrm{Cov}(Z \mid Y) - \mathrm{Id}\big).$$

Third derivatives of $\log p_t$ are σ^{-3} times conditional third cumulants of Z, up to sign. Conditional Jensen and Hölder inequalities therefore give all finite $L^q(\pi_t)$ norms of these derivatives. The

absolute formula (2.2) gives $B_t < \infty$. The second moment is finite by (7.1), and the score relation gives $E_t < \infty$. □

Proof of Theorem 7.1. Write $I_t = \int |r_t|^2 d\pi_t$, $v(t) = \sqrt{I_t}$, and $H_t = \int |\nabla^2 h_t|^2 d\pi_t$. The classical entropy and Fisher information identities for the Ornstein–Uhlenbeck semigroup [2, Chapter 5], also used in [16], combine with (2.6) to give

$$\begin{aligned} \frac{d}{dt}\operatorname{Ent}(\pi_t \mid \gamma_\kappa) &= -I_t, & I_t' &= -2\kappa I_t - 8H_t, \\ \mathcal{D}_\kappa(\pi_t)' &= -\frac{4}{\kappa}H_t, & \int r_t \cdot G_t \, d\pi_t &= \frac{4}{\kappa}H_t. \end{aligned} \tag{7.8}$$

These identities are used only at positive times, where Lemma 7.2 supplies the integrability. Define

$$\alpha(t) = \frac{\int r_t \cdot G_t \, d\pi_t}{v(t)} = \frac{4H_t}{\kappa v(t)}$$

when $v(t) > 0$, and set $\alpha(t) = 0$ otherwise. Then $0 \le \alpha(t) \le g(t)$ and

$$v' = -\kappa v - \kappa\alpha.$$

If $I_t = 0$ at any positive time, Gaussian deconvolution in (7.1) shows that $\pi = \gamma_\kappa$, and all conclusions are immediate. Otherwise $v > 0$. Integration backwards from the endpoint (7.6) gives the exact formula

$$v(t) = \kappa \int_t^\infty e^{\kappa(s-t)}\alpha(s)\, ds. \tag{7.9}$$

The entropy flow has metric speed $v(t)$ [1, Chapters 8 and 11]; see also [15, Section 3]. Since $\pi_t \to \pi$ in W_2 as $t \downarrow 0$ and $\pi_t \to \gamma_\kappa$ in W_2 as $t \to \infty$, Tonelli's theorem and (7.9) give

$$W_2(\pi, \gamma_\kappa) \le \int_0^\infty v(t)\, dt = \int_0^\infty (e^{\kappa s} - 1)\alpha(s)\, ds.$$

This proves (7.3). The statement remains valid when the right side is infinite.

Suppose now that $I_0 < \infty$. Strong continuity and lower semicontinuity, together with $I_t \le e^{-2\kappa t} I_0$, give $I_t \to I_0$ as $t \downarrow 0$. Multiplying (7.9) by $e^{\kappa t}$ and letting $t \downarrow 0$ proves (7.4). The entropy is continuous at zero and both entropy and information tend to zero at infinity. Hence (7.8) yields

$$\begin{aligned} \mathcal{D}_\kappa(\pi) &= \int_0^\infty v(t)\alpha(t)\, dt \\ &= \frac{\kappa}{2}\int_0^\infty \int_0^\infty e^{\kappa|t-s|}\alpha(t)\alpha(s)\, dt\, ds. \end{aligned} \tag{7.10}$$

In the second equality use (7.9) on the triangle $s > t$ and symmetry; the diagonal has Lebesgue measure zero. The inequality $\alpha \le g$ proves (7.5).

To check sharpness, take $\pi = N(0, a_0\kappa^{-1}\mathrm{Id})$ with $a_0 > 0$. Along the flow,

$$a_t = 1 + (a_0 - 1)e^{-2\kappa t}, \qquad v(t) = \sqrt{n\kappa}\,\frac{|a_t - 1|}{\sqrt{a_t}}, \qquad G_t = a_t^{-1} r_t.$$

Thus $\alpha = g$ identically. The Gaussian path changes only its scalar standard deviation and is a reparametrized W_2 geodesic. Its length is exactly $\sqrt{n/\kappa}\,|\sqrt{a_0} - 1|$. All three bounds are therefore equalities, and their multiplicative constants cannot be decreased. □

Corollary 7.3 (No fixed observation interval suffices)**.** *Fix $T < \infty$ and $a > 0$. There are centered smooth positive densities p_R with uniformly bounded second moment such that*

$$\sup_{0 \le t \le T} g_R(t) \longrightarrow 0, \qquad W_2^2(p_R\, dx, \gamma_\kappa) \longrightarrow a.$$

Proof. Use the three-component mixture (8.4) below with $\varepsilon_R = a/R^2$. Under (7.1) its weights and component covariances stay fixed, while its centers become 0 and $\pm e^{-\kappa t} R e_1$. These centers remain uniformly separated on $[0,T]$. Lemma 3.1, with $\ell = \kappa$, gives an exponentially small bound for $B_R(t)$, uniform on that interval. Lemma 4.2 gives the stated limit of W_2^2. □

The deficit dissipation is $4H_t/\kappa$, not $\|G_t\|_2^2$. The exact identity (7.10) uses the projection α of G_t onto the relative score.

8. Comparison with minimization of the deficit

8.1. **The classical constrained deficit bound.** Stam's dimensional entropy–Fisher inequality [19] reads

$$\mathcal{D}_\kappa(\pi) \geq \frac{n}{2}\left(\frac{J}{n\kappa} - 1 - \log\frac{J}{n\kappa}\right), \tag{8.1}$$

with equality exactly for isotropic Gaussians. See [4, Section 2] and [11, Section 1.2] for this normalization. Under the constraints of Theorem 1.1, its classical consequence is

$$\inf \mathcal{D}_\kappa = \frac{n}{2}(r - \log(1+r)), \qquad r = \frac{4(e-e_*)_+}{n\kappa}. \tag{8.2}$$

The lower bound follows from (2.8). The same exact-information mixtures give the upper bound: their label entropy tends to zero, and the entropy chain rule [8, 11] gives

$$\int p \log p = \int \varphi_\ell \log \varphi_\ell - \mathsf{H}(A) + \mathsf{H}(A \mid X), \qquad 0 \leq \mathsf{H}(A \mid X) \leq \mathsf{H}(A). \tag{8.3}$$

Equality in (8.1) gives the same attainment cases as Theorem 1.1. Dovetta–Serra [9] minimize a different quotient. The next calculation explains why these escaping sequences cannot minimize it.

8.2. **Distance to the Gaussian cone.** Let $a > 0$, $\varepsilon_R = a/R^2$, and, for R sufficiently large, set

$$p_R = (1-\varepsilon_R)\varphi_\kappa + \frac{\varepsilon_R}{2}(\varphi_\kappa(\cdot - Re_1) + \varphi_\kappa(\cdot + Re_1)). \tag{8.4}$$

Let

$$\mathcal{M}_\kappa = \{c\sqrt{\varphi_\kappa(\cdot - m)} : c \in \mathbb{R},\ m \in \mathbb{R}^n\}$$

be the cone of Gaussian square roots of the reference variance. At $\kappa = 2\pi$, this is the cone in [9], and their Euclidean deficit at $u = \sqrt{p}$, $\|u\|_2 = 1$, is $\pi \mathcal{D}_{2\pi}(p\,dx)$. Other values of κ are related by a fixed spatial rescaling.

Proposition 8.1 (Divergence of the deficit quotient)**.** *For* (8.4),

$$\mathcal{D}_\kappa(p_R\,dx) \sim \varepsilon_R \log(1/\varepsilon_R), \qquad \operatorname{dist}_{L^2}^2(\sqrt{p_R}, \mathcal{M}_\kappa) \sim \varepsilon_R. \tag{8.5}$$

Consequently their quotient tends to infinity, although $B_R \to 0$ and $W_2^2(p_R\,dx, \gamma_\kappa) \to a$.

Proof. For the latent center $A \in \{0, Re_1, -Re_1\}$, (3.10) and (8.3) give the exact identity

$$\mathcal{D}_\kappa(p_R\,dx) = \mathsf{H}(A) - \mathsf{H}(A \mid X) - \frac{\kappa}{2}\mathbb{E}\,\operatorname{tr}\operatorname{Cov}(A \mid X).$$

The last term is exponentially small by Lemma 3.1. The conditional entropy is also exponentially small: for three probabilities r_i, the elementary bound $-\sum_i r_i \log r_i \leq C(1 - \sum_i r_i^2)^{1/2}$, followed by Cauchy–Schwarz and (3.1), gives this conclusion. The label entropy is

$$\mathsf{H}(A) = -(1-\varepsilon_R)\log(1-\varepsilon_R) - \varepsilon_R \log(\varepsilon_R/2) \sim \varepsilon_R \log(1/\varepsilon_R).$$

This proves the first assertion.

Put $A_R(m) = \int \sqrt{p_R(x)\varphi_\kappa(x-m)}\,dx$. The Gaussian-cone distance formula [9, Remark 2.2] gives

$$\operatorname{dist}_{L^2}^2(\sqrt{p_R}, \mathcal{M}_\kappa) = 1 - \sup_m A_R(m)^2.$$

The Gaussian affinity formula and $\sqrt{u+v+w} \le \sqrt{u}+\sqrt{v}+\sqrt{w}$ imply

$$A_R(m) \le \sqrt{1-\varepsilon_R}e^{-\kappa|m|^2/8} + \sqrt{\varepsilon_R/2}\left(e^{-\kappa|m-Re_1|^2/8}+e^{-\kappa|m+Re_1|^2/8}\right).$$

For $|m| \le R/3$, this is at most $\sqrt{1-\varepsilon_R}+Ce^{-cR^2}$. For $|m| > R/3$, it is at most $Ce^{-cR^2}+\sqrt{2\varepsilon_R}$, which is smaller for large R. Conversely, $A_R(0) \ge \sqrt{1-\varepsilon_R}$. Therefore $\sup_m A_R(m)^2 = 1-\varepsilon_R+O(e^{-cR^2})$. The second assertion follows. The limits of B_R and W_2^2 follow from Lemmas 3.1 and 4.2. □

The quotient diverges because mixing entropy has order $\varepsilon\log(1/\varepsilon)$, whereas squared distance to the Gaussian cone has order ε.

9. Metric slope and weighted manifolds

The first-variation field in (1.5) describes smooth transport perturbations. A metric-slope identity must also control arbitrary nearby measures. The following argument uses HWI for this comparison; it does not assume displacement convexity of Fisher information [13,17].

9.1. **The metric slope at regular densities.** Let (M,g) be a connected closed Riemannian manifold and $d\mu = e^{-f}dv_g$ a smooth probability measure. For $c>0$, define

$$\mathcal{D}_c^\mu(\rho\mu) = \frac{1}{2c}\int |\nabla\log\rho|^2\rho\,d\mu - \int \rho\log\rho\,d\mu.$$

Information is understood as $4\int|\nabla\sqrt{\rho}|^2d\mu$, and the functional is $+\infty$ when this is infinite. The descending metric slope is

$$|\partial\mathcal{D}_c^\mu|(\pi) = \limsup_{\substack{\eta\to\pi \text{ in } W_2\\ \eta\neq\pi}} \frac{(\mathcal{D}_c^\mu(\pi)-\mathcal{D}_c^\mu(\eta))_+}{W_2(\pi,\eta)}.$$

No smoothness is imposed on the competing measures η.

Theorem 9.1 (Metric slope at a positive smooth density)**.** *Suppose M is closed and ρ is smooth, positive, and normalized. Put*

$$h = \tfrac{1}{2}\log\rho, \qquad b = -\Delta_f h - |\nabla h|^2 - ch.$$

Then

$$|\partial\mathcal{D}_c^\mu|^2(\rho\mu) = \frac{4}{c^2}\int|\nabla b|^2\rho\,d\mu. \tag{9.1}$$

The same assertion holds for $\mathcal{D}_\kappa$ on $\mathcal{P}_2(\mathbb{R}^n)$ at every density

$$p(x) = Z^{-1}\exp\left(-\tfrac{1}{2}(x-m)^TA(x-m)+v(x)\right), \quad A\succ 0, \quad v\in C_c^\infty(\mathbb{R}^n), \tag{9.2}$$

with $c=\kappa$. In both cases the slope is taken over the full Wasserstein space, including nonsmooth finite-information competitors.

Proof. Write $\pi=\rho\mu$ and $F = \delta\mathcal{D}_c^\mu/\delta\pi = 2b/c-1$. For any competing probability measure η of finite information, expansion of relative Fisher information and relative entropy gives

$$\mathcal{D}_c^\mu(\eta)-\mathcal{D}_c^\mu(\pi) = \int F\,d(\eta-\pi) + \frac{1}{2c}I(\eta\mid\pi) - \mathrm{Ent}(\eta\mid\pi). \tag{9.3}$$

To see the information expansion, write $d\eta=\sigma\,d\mu$ and expand $\nabla\log\sigma = \nabla\log(\sigma/\rho)+\nabla\log\rho$; integration of the cross term leaves $-2\Delta_f\log\rho-|\nabla\log\rho|^2$ as the linear coefficient. The entropy expansion is the definition of relative entropy. Approximation in H^1 of the square-root density extends the identities to finite-information competitors. On a closed manifold the coefficients are

bounded; for (9.2) they have at most quadratic growth, so the finite second moment controls the additional terms.

At the base measure π, the weighted Ricci tensor is bounded below by $-Kg$ for some $K \geq 0$. The HWI inequality of Otto–Villani [18], in its weighted-manifold form [2, Chapter 9], implies, with $W = W_2(\eta, \pi)$,

$$\mathrm{Ent}(\eta \mid \pi) \leq W\sqrt{I(\eta \mid \pi)} + \frac{K}{2}W^2.$$

Completing the square gives

$$\frac{1}{2c}I(\eta \mid \pi) - \mathrm{Ent}(\eta \mid \pi) \geq -\frac{c+K}{2}W^2. \tag{9.4}$$

Also $\nabla^2 F \geq -Lg$ for some $L \geq 0$. Integrating the second-order lower bound for F along minimizing geodesics in an optimal coupling, and using Cauchy–Schwarz, yields

$$\int F\, d(\eta - \pi) \geq -\|\nabla F\|_{L^2(\pi)}W - \frac{L}{2}W^2.$$

Together with (9.3)–(9.4), this proves $|\partial \mathcal{D}_c^\mu|(\pi) \leq \|\nabla F\|_2$. The curvature and Hessian bounds are automatic on the closed manifold. For (9.2), $-\log p$ and F are quadratic outside a compact set, so the same bounds hold globally.

For the reverse inequality, transport π by the smooth field $-\nabla F$. On a closed manifold its short-time flow gives

$$\frac{d}{dt}\mathcal{D}_c^\mu(\pi_t)\Big|_{t=0} = -\|\nabla F\|_2^2, \qquad W_2(\pi_t, \pi) \leq t\|\nabla F\|_2 + o(t).$$

On $\mathbb{R}^n$, first use compactly supported gradient fields approximating $-\nabla F$ in $L^2(\pi)$, and then pass to the limit. Such approximation follows by cutting off the potential F, which has quadratic growth while π has Gaussian tails. This proves the reverse inequality and (9.1). □

9.2. **Energy dissipation and the sphere.** The corresponding smooth evolution is

$$\partial_t \rho = \mathrm{div}_\mu(\rho \nabla F_\rho), \qquad F_\rho = -\frac{1}{c}\Delta_f \log\rho - \frac{1}{2c}|\nabla \log\rho|^2 - \log\rho - 1, \tag{9.5}$$

where $\mathrm{div}_\mu X = e^f \,\mathrm{div}(e^{-f}X)$. This is a fourth-order equation, distinct from the Ornstein–Uhlenbeck flow used in Section 7. Information-driven fourth-order Wasserstein flows are classical [17]; the negative entropy term in (9.5) is part of the present functional.

For a smooth positive normalized solution on a closed manifold, the standard metric-speed and chain-rule identities [1, Chapter 8], together with Theorem 9.1, give

$$\mathcal{D}_c^\mu(\pi_t) + \frac{1}{2}\int_s^t |\dot\pi_r|_{W_2}^2\, dr + \frac{1}{2}\int_s^t |\partial\mathcal{D}_c^\mu|^2(\pi_r)\, dr = \mathcal{D}_c^\mu(\pi_s). \tag{9.6}$$

Indeed, the velocity is $-\nabla F_\rho$, whose norm is the metric speed, and $d\mathcal{D}_c^\mu/dt = -\|\nabla F_\rho\|_2^2$. This assertion is restricted to the interval of smooth positive existence.

In Euclidean space the Gaussian family gives global solutions. For $N(m, \Sigma)$,

$$G(x) = \kappa^{-1}\Sigma^{-1}(\kappa\mathrm{Id} - \Sigma^{-1})(x - m).$$

The linear velocity $-G$ fixes the mean and yields

$$\dot\Sigma = \frac{2}{\kappa}\Sigma^{-1} - 2\mathrm{Id}. \tag{9.7}$$

The eigenvectors are fixed, and each eigenvalue satisfies $\dot\sigma = 2/(\kappa\sigma) - 2$. Positive eigenvalues stay positive and converge to κ^{-1}. Direct differentiation gives (9.6) along these curves.

On the unit sphere $\mathbb{S}^n$ with uniform probability measure, the sharp logarithmic Sobolev coefficient is $c = n$ [3,5,6], not the Ricci coefficient $n-1$. Theorem 9.1 therefore identifies the squared metric slope of the sharp spherical deficit as

$$\frac{4}{n^2}\int \left|\nabla\left(\Delta h + |\nabla h|^2 + nh\right)\right|^2 e^{2h}\, d\mu.$$

Equations (9.5) and (9.6) apply with $c = n$. The constructive spherical stability inequalities are proved in [5,6].

Corollary 9.2 (Heat-flow control on closed weighted manifolds)**.** *Let M be closed and connected, let $d\mu = e^{-f} dv_g$ be a smooth probability measure, and assume $\mathrm{Ric} + \nabla^2 f \geq \kappa g$ for $\kappa > 0$. Let π_t be the weighted heat flow from any probability measure π. At $t > 0$, put*

$$g(t) = |\partial \mathcal{D}_\kappa^\mu|(\pi_t).$$

Then (7.3) holds with μ in place of γ_κ. If $I(\pi \mid \mu) < \infty$, the information and deficit bounds (7.4)–(7.5) hold as well, with the same replacements. No centering condition is needed.

Proof. At every positive time the heat density is smooth and strictly positive, so Theorem 9.1 identifies g with the first-variation norm. Set $r_t = \nabla \log(d\pi_t/d\mu)$, $v(t) = \sqrt{I_t}$, and $G_t = \nabla F_{\rho_t}$. The usual integrated Bochner identity [2] gives

$$-\frac{d}{dt}\mathcal{D}_\kappa^\mu(\pi_t) = \frac{1}{\kappa}\int \left(|\nabla^2 \log \rho_t|^2 + (\mathrm{Ric}_f - \kappa g)(r_t, r_t)\right) d\pi_t \geq 0.$$

The left side is $\int r_t \cdot G_t\, d\pi_t$. Thus $\alpha(t) = \int r_t \cdot G_t\, d\pi_t / v(t)$ is nonnegative, bounded by $g(t)$, and again satisfies $v' = -\kappa v - \kappa\alpha$.

The weighted Bochner bound gives $\lambda_1 \geq \kappa$ for the first positive eigenvalue of $-\Delta_f$ [2]. Equality would force a nonconstant eigenfunction to have zero Hessian, impossible on a closed manifold. Thus $\lambda_1 > \kappa$. Spectral decay after a positive regularization time gives $\rho_t \to 1$ uniformly and $\|\nabla \rho_t\|_2 = O(e^{-\lambda_1 t})$. Hence $e^{\kappa t} v(t) \to 0$, and the proof of Theorem 7.1, including its time-zero limits, applies. □

On $\mathbb{S}^n$, $n \geq 2$, the corollary uses $\kappa = n-1$, not the sharp spherical coefficient n.

The slope theorem assumes a regular base measure. We make no claim about slopes at arbitrary nonsmooth densities or global weak solutions of the fourth-order flow.

Appendix A. Cutoffs and global integration

We verify the global identities under $E, S, B < \infty$. The pointwise Bochner and score formulas are classical [2,8]; the issue is their integrability and boundary terms.

Proof. Choose $\eta_R \in C_c^\infty(\mathbb{R}^n)$ with $0 \leq \eta_R \leq 1$, $\eta_R = 1$ on the ball of radius R, support in the ball of radius $2R$, and $|\nabla \eta_R| \leq c/R$. For $H_R = \int \eta_R^2 |\nabla^2 h|^2 d\pi$ and $Y_R = \int |\nabla \eta_R|^2 |\nabla h|^2 d\pi$, the Gaussian Bochner formula [2], tested with η_R^2, gives

$$H_R = \int \eta_R^2 \nabla h \cdot \nabla b\, d\pi - \int \eta_R \nabla \eta_R \cdot \nabla |\nabla h|^2\, d\pi \leq \sqrt{EB} + 2\sqrt{H_R Y_R}.$$

Since $Y_R \leq c^2 E/R^2$, Young's inequality bounds H_R. Fatou gives $H := \int |\nabla^2 h|^2 d\pi < \infty$, and the boundary term is at most $2\sqrt{HY_R} \to 0$. Dominated convergence proves (2.6); $K = 2\nabla^2 h - \kappa \mathrm{Id}$ then belongs to $L^2(\pi)$.

The classical score identities [8] follow by the same cutoff: $s, K, x \in L^2(\pi)$ and $|\nabla \eta_R| \leq c/R$ make their boundary terms vanish. This gives (2.5).

Differentiating (1.2) gives

$$L(\partial_i h) = -\partial_i b. \tag{A.1}$$

Testing against η_R and using $\nabla^2 h \in L^2(\pi)$ proves $\int \partial_i b \, d\pi = 0$. Since $Lx = s$, (2.3) and (A.1) yield

$$Ls = -\kappa s - 2\nabla b. \tag{A.2}$$

Test the ith component against $\eta_R^2 s_j$. The cutoff term tends to zero: its absolute value is at most

$$\frac{2c}{R} \|s_j\|_{L^2(\pi)} \|\nabla s_i\|_{L^2(\pi)}.$$

The other terms converge by $s, K, \nabla b \in L^2(\pi)$. Thus $\int \nabla s_i \cdot \nabla s_j d\pi = \kappa \int s_i s_j d\pi + 2\int (\partial_i b) s_j d\pi$, which is (2.7).

Finally, centering gives $\int \nabla h = 0$, and the score identities justify the standard relation (2.8). □

References

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, second edition, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 2008. doi:10.1007/978-3-7643-8722-8.

[2] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*, Grundlehren der mathematischen Wissenschaften, vol. 348, Springer, Cham, 2014. doi:10.1007/978-3-319-00227-9.

[3] W. Beckner, *Sharp Sobolev inequalities on the sphere and the Moser–Trudinger inequality*, Ann. of Math. (2) **138** (1993), no. 1, 213–242. doi:10.2307/2946638.

[4] S. G. Bobkov, N. Gozlan, C. Roberto, and P.-M. Samson, *Bounds on the deficit in the logarithmic Sobolev inequality*, J. Funct. Anal. **267** (2014), no. 11, 4110–4138. doi:10.1016/j.jfa.2014.09.016.

[5] G. Brigati, J. Dolbeault, and N. Simonov, *Logarithmic Sobolev and interpolation inequalities on the sphere: constructive stability results*, Ann. Inst. H. Poincaré Anal. Non Linéaire **41** (2024), no. 5, 1289–1321. doi:10.4171/AIHPC/106.

[6] G. Brigati, J. Dolbeault, and N. Simonov, *Corrigendum to "Logarithmic Sobolev and interpolation inequalities on the sphere: constructive stability results"*, Ann. Inst. H. Poincaré Anal. Non Linéaire **41** (2024), no. 5, 1323–1324. doi:10.4171/AIHPC/118.

[7] F. Conrado, *Rigidity for the logarithmic Sobolev inequality on complete metric measure spaces*, Arch. Math. (Basel) **121** (2023), 279–286. doi:10.1007/s00013-023-01906-6.

[8] A. Dembo, T. M. Cover, and J. A. Thomas, *Information theoretic inequalities*, IEEE Trans. Inform. Theory **37** (1991), no. 6, 1501–1518. doi:10.1109/18.104312.

[9] S. Dovetta and E. Serra, *Existence of optimizers for the sharp stability constant in the logarithmic Sobolev inequality*, arXiv:2606.23225 (2026). doi:10.48550/arXiv.2606.23225.

[10] A. Dytso, H. V. Poor, and S. Shamai (Shitz), *Conditional mean estimation in Gaussian noise: a meta derivative identity with applications*, IEEE Trans. Inform. Theory **69** (2023), no. 3, 1883–1898. doi:10.1109/TIT.2022.3216012.

[11] R. Eldan, J. Lehec, and Y. Shenfeld, *Stability of the logarithmic Sobolev inequality via the Föllmer process*, Ann. Inst. Henri Poincaré Probab. Stat. **56** (2020), no. 3, 2253–2269. doi:10.1214/19-AIHP1038.

[12] M. Fathi, E. Indrei, and M. Ledoux, *Quantitative logarithmic Sobolev inequalities and stability estimates*, Discrete Contin. Dyn. Syst. **36** (2016), no. 12, 6835–6853. doi:10.3934/dcds.2016097.

[13] U. Gianazza, G. Savaré, and G. Toscani, *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation*, Arch. Ration. Mech. Anal. **194** (2009), 133–220. doi:10.1007/s00205-008-0186-5.

[14] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. doi:10.2307/2373688.

[15] M. Ledoux, I. Nourdin, and G. Peccati, *Stein's method, logarithmic Sobolev and transport inequalities*, Geom. Funct. Anal. **25** (2015), no. 1, 256–306. doi:10.1007/s00039-015-0312-0.

[16] M. Ledoux, I. Nourdin, and G. Peccati, *A Stein deficit for the logarithmic Sobolev inequality*, Sci. China Math. **60** (2017), no. 7, 1163–1180. doi:10.1007/s11425-016-0134-7.

[17] D. Matthes, R. J. McCann, and G. Savaré, *A family of nonlinear fourth order equations of gradient flow type*, Comm. Partial Differential Equations **34** (2009), no. 11, 1352–1397. doi:10.1080/03605300903296256.

[18] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. doi:10.1006/jfan.1999.3557.

[19] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control **2** (1959), no. 2, 101–112. doi:10.1016/S0019-9958(59)90348-1.

[20] M. Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600. doi:10.1007/BF02249265.

[21] J. Wei and Y. Wu, *Sharp stability of the logarithmic Sobolev inequality in the critical point setting*, in *Potentials and Partial Differential Equations: The Legacy of David R. Adams*, S. Lenhart and J. Xiao (eds.), De Gruyter, Berlin–Boston, 2023, pp. 77–102. doi:10.1515/9783110792720-004.