Stability of Talagrand's Gaussian Transport-Entropy Inequality

Dan Mikulincer

Geometric and Functional Inequalities in Convexity and Probability

Weizmann Institute of Science

Based on joint work with Ronen Eldan and Alex Zhai

Geometry and Information

Throughout, $G \sim \gamma$ will denote the standard Gaussian in $\mathbb{R}^d$.

Definition (Wasserstein distance between μ and γ)

$$\mathcal{W}_2(\mu,\gamma) := \inf_{\pi} \left\{ \mathbb{E}_{\pi}\left[||x-y||^2\right] \right\}^{1/2}$$

where π ranges over all possible couplings of μ and γ.

Definition (Relative entropy between μ and γ)

$$\mathrm{Ent}(\mu||\gamma) := \mathbb{E}_{\mu}\left[\ln\left(\frac{d\mu}{d\gamma}(x)\right)\right].$$

Remark: if $X \sim \mu$ we will also write $\mathrm{Ent}(X||G), \mathcal{W}_2(X,G)$.

Geometry and Information

Throughout, $G \sim \gamma$ will denote the standard Gaussian in $\mathbb{R}^d$.

Definition (Wasserstein distance between μ and γ)

$$\mathcal{W}_2(\mu,\gamma) := \inf_{\pi} \left\{ \mathbb{E}_\pi \left[||x-y||^2 \right] \right\}^{1/2}$$

where π ranges over all possible couplings of μ and γ.

Definition (Relative entropy between μ and γ)

$$\mathrm{Ent}(\mu||\gamma) := \mathbb{E}_\mu \left[\ln\left(\frac{d\mu}{d\gamma}(x) \right) \right].$$

Remark: if $X \sim \mu$ we will also write $\mathrm{Ent}(X||G), \mathcal{W}_2(X,G)$.

Geometry and Information

Throughout, $G \sim \gamma$ will denote the standard Gaussian in $\mathbb{R}^d$.

Definition (Wasserstein distance between μ and γ)

$$\mathcal{W}_2(\mu, \gamma) := \inf_{\pi} \left\{ \mathbb{E}_\pi \left[||x - y||^2 \right] \right\}^{1/2}$$

where π ranges over all possible couplings of μ and γ.

Definition (Relative entropy between μ and γ)

$$\mathrm{Ent}(\mu||\gamma) := \mathbb{E}_\mu \left[\ln \left(\frac{d\mu}{d\gamma}(x) \right) \right].$$

Remark: if $X \sim \mu$ we will also write $\mathrm{Ent}(X||G), \mathcal{W}_2(X, G)$.

Geometry and Information

Throughout, $G \sim \gamma$ will denote the standard Gaussian in $\mathbb{R}^d$.

Definition (Wasserstein distance between μ and γ)

$$\mathcal{W}_2(\mu, \gamma) := \inf_{\pi} \left\{ \mathbb{E}_\pi \left[||x - y||^2 \right] \right\}^{1/2}$$

where π ranges over all possible couplings of μ and γ.

Definition (Relative entropy between μ and γ)

$$\mathrm{Ent}(\mu||\gamma) := \mathbb{E}_\mu \left[\ln \left(\frac{d\mu}{d\gamma}(x) \right) \right].$$

Remark: if $X \sim \mu$ we will also write $\mathrm{Ent}(X||G), \mathcal{W}_2(X, G)$.

Talagrand's Inequality

In $96'$ Talagrand proved the following inequality, which connects between geometry and information.

Theorem (Talagrand's Gaussian transport-entropy inequality)

Let μ be a measure on $\mathbb{R}^d$. Then

$$\mathcal{W}_2^2(\mu, \gamma) \leq 2\mathrm{Ent}(\mu||\gamma).$$

It is enough to consider measures such that $\mu \ll \nu$.

Talagrand's Inequality - Applications

- By considering measures of the form $\mathbb{1}_A d\gamma$ the inequality implies a (non-sharp) Gaussian isoperimetric inequality.
- The inequality tensorizes and may be used to show dimension-free Gaussian concentration bounds.
- If f is convex, then applying the inequality to $e^{-\lambda f} d\gamma$ yields a one sides Gaussian concentration for concave functions.

Talagrand's Inequality - Applications

- By considering measures of the form $\mathbb{1}_A d\gamma$ the inequality implies a (non-sharp) Gaussian isoperimetric inequality.
- The inequality tensorizes and may be used to show dimension-free Gaussian concentration bounds.
- If f is convex, then applying the inequality to $e^{-\lambda f} d\gamma$ yields a one sides Gaussian concentration for concave functions.

Talagrand's Inequality - Applications

- By considering measures of the form $\mathbb{1}_A d\gamma$ the inequality implies a (non-sharp) Gaussian isoperimetric inequality.
- The inequality tensorizes and may be used to show dimension-free Gaussian concentration bounds.
- If f is convex, then applying the inequality to $e^{-\lambda f} d\gamma$ yields a one sides Gaussian concentration for concave functions.

Talagrand's Inequality - Applications

- By considering measures of the form $\mathbb{1}_A d\gamma$ the inequality implies a (non-sharp) Gaussian isoperimetric inequality.
- The inequality tensorizes and may be used to show dimension-free Gaussian concentration bounds.
- If f is convex, then applying the inequality to $e^{-\lambda f} d\gamma$ yields a one sides Gaussian concentration for concave functions.

Gaussians

If $\gamma_{a,\Sigma} = \mathcal{N}(a, \Sigma)$, in $\mathbb{R}^d$:

- $\mathrm{Ent}(\gamma_{a,\Sigma}||\gamma) = \frac{1}{2}\left(\mathrm{Tr}(\Sigma) + ||a||_2^2 - \ln(\det(\Sigma)) - d\right)$
- $\mathcal{W}_2^2(\gamma_{a,\Sigma}, \gamma) = ||a||_2^2 + \left|\left|\sqrt{\Sigma} - \mathrm{I}_d\right|\right|_{HS}^2$

In particular, for any $a \in \mathbb{R}^d$,

$$\mathcal{W}_2^2(\gamma_{a,\mathrm{I}_d}, \gamma) = 2\mathrm{Ent}(\gamma_{a,\mathrm{I}_d}||\gamma).$$

These are the only equality cases.

Gaussians

If $\gamma_{a,\Sigma} = \mathcal{N}(a, \Sigma)$, in $\mathbb{R}^d$:

- $\mathrm{Ent}(\gamma_{a,\Sigma} || \gamma) = \frac{1}{2}\left(\mathrm{Tr}(\Sigma) + ||a||_2^2 - \ln(\det(\Sigma)) - d\right)$
- $\mathcal{W}_2^2(\gamma_{a,\Sigma}, \gamma) = ||a||_2^2 + \left|\left|\sqrt{\Sigma} - \mathrm{I}_d\right|\right|_{HS}^2$

In particular, for any $a \in \mathbb{R}^d$,

$$\mathcal{W}_2^2(\gamma_{a,\mathrm{I}_d}, \gamma) = 2\mathrm{Ent}(\gamma_{a,\mathrm{I}_d} || \gamma).$$

These are the only equality cases.

Gaussians

If $\gamma_{a,\Sigma} = \mathcal{N}(a, \Sigma)$, in $\mathbb{R}^d$:

- $\mathrm{Ent}(\gamma_{a,\Sigma}||\gamma) = \frac{1}{2}\left(\mathrm{Tr}(\Sigma) + ||a||_2^2 - \ln(\det(\Sigma)) - d\right)$
- $\mathcal{W}_2^2(\gamma_{a,\Sigma}, \gamma) = ||a||_2^2 + \left|\left|\sqrt{\Sigma} - \mathrm{I}_d\right|\right|_{HS}^2$

In particular, for any $a \in \mathbb{R}^d$,

$$\mathcal{W}_2^2(\gamma_{a,\mathrm{I}_d}, \gamma) = 2\mathrm{Ent}(\gamma_{a,\mathrm{I}_d}||\gamma).$$

These are the only equality cases.

Gaussians

If $\gamma_{a,\Sigma} = \mathcal{N}(a, \Sigma)$, in $\mathbb{R}^d$:

- $\mathrm{Ent}(\gamma_{a,\Sigma}||\gamma) = \frac{1}{2}\left(\mathrm{Tr}(\Sigma) + ||a||_2^2 - \ln(\det(\Sigma)) - d\right)$
- $\mathcal{W}_2^2(\gamma_{a,\Sigma}, \gamma) = ||a||_2^2 + \left|\left|\sqrt{\Sigma} - \mathrm{I}_d\right|\right|_{HS}^2$

In particular, for any $a \in \mathbb{R}^d$,

$$\mathcal{W}_2^2(\gamma_{a,\mathrm{I}_d}, \gamma) = 2\mathrm{Ent}(\gamma_{a,\mathrm{I}_d}||\gamma).$$

These are the only equality cases.

Stability

Define the deficit

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu,\gamma).$$

The question of stability deals with approximate equality cases.

Question

Suppose that $\delta_{\mathrm{Tal}}(\mu)$ is small, must μ be close to a translate of the standard Gaussian?

Note that the deficit is invariant to translations. So, it will be enough to consider centered measures.

Stability

Define the deficit

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu, \gamma).$$

The question of stability deals with approximate equality cases.

Question

Suppose that $\delta_{\mathrm{Tal}}(\mu)$ is small, must μ be close to a translate of the standard Gaussian?

Note that the deficit is invariant to translations. So, it will be enough to consider centered measures.

Stability

Define the deficit

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu,\gamma).$$

The question of stability deals with approximate equality cases.

Question

Suppose that $\delta_{\mathrm{Tal}}(\mu)$ is small, must μ be close to a translate of the standard Gaussian?

Note that the deficit is invariant to translations. So, it will be enough to consider centered measures.

Instability

Theorem (Fathi, Indrei, Ledoux 14')

Let μ be a centered measure on $\mathbb{R}^d$. Then

$$\delta_{\mathrm{Tal}}(\mu) \gtrsim \min\left(\frac{\mathcal{W}_{1,1}(\mu,\gamma)^2}{d}, \frac{\mathcal{W}_{1,1}(\mu,\gamma)}{\sqrt{d}}\right)$$

The 1-dimensional case was proven earlier by Barthe and Kolesnikov.

However:

Theorem

There exists a sequence of centered Gaussian mixtures $\{\mu_n\}$ on $\mathbb{R}$, such that $\delta_{\mathrm{Tal}}(\mu_n) \to 0$. but $\mathcal{W}_2^2(\mu_n,\gamma) > 1$.

Instability

Theorem (Fathi, Indrei, Ledoux 14')

Let μ *be a centered measure on* $\mathbb{R}^d$*. Then*

$$\delta_{\mathrm{Tal}}(\mu) \gtrsim \min\left(\frac{\mathcal{W}_{1,1}(\mu,\gamma)^2}{d}, \frac{\mathcal{W}_{1,1}(\mu,\gamma)}{\sqrt{d}}\right)$$

The 1-dimensional case was proven earlier by Barthe and Kolesnikov.

However:

Theorem

There exists a sequence of centered Gaussian mixtures $\{\mu_n\}$ *on* $\mathbb{R}$*, such that* $\delta_{\mathrm{Tal}}(\mu_n) \to 0$*. but* $\mathcal{W}_2^2(\mu_n,\gamma) > 1$*.*

Bounding the Deficit

In the 1-dimensional case, Talagrand actually showed

$$\delta_{\mathrm{Tal}}(\mu) = \int_{\mathbb{R}} \left(\varphi'_\mu - 1 - \ln(\varphi'_\mu)\right) d\gamma > 0,$$

where φ is the transport map $\varphi_\mu = F_\gamma^{-1} \circ F_\mu$.

For translated Gaussians, $\varphi_{\gamma_{a,1}}(x) = x + a$, which shows the equality cases.

We will take a different route.

Bounding the Deficit

In the 1-dimensional case, Talagrand actually showed

$$\delta_{\mathrm{Tal}}(\mu) = \int_{\mathbb{R}} \left(\varphi_\mu' - 1 - \ln(\varphi_\mu')\right) d\gamma > 0,$$

where φ is the transport map $\varphi_\mu = F_\gamma^{-1} \circ F_\mu$.

For translated Gaussians, $\varphi_{\gamma_{a,1}}(x) = x + a$, which shows the equality cases.

We will take a different route.

Bounding the Deficit

In the 1-dimensional case, Talagrand actually showed

$$\delta_{\mathrm{Tal}}(\mu) = \int_{\mathbb{R}} \left(\varphi_\mu' - 1 - \ln(\varphi_\mu')\right) d\gamma > 0,$$

where φ is the transport map $\varphi_\mu = F_\gamma^{-1} \circ F_\mu$.

For translated Gaussians, $\varphi_{\gamma_{a,1}}(x) = x + a$, which shows the equality cases.

We will take a different route.

Bounding the Deficit - the Föllmer Drift

Our central construct will be the Föllmer drift, which is the solution to the following variational problem:

$$v_t := \arg\min_{u_t} \frac{1}{2} \int_0^1 \mathbb{E}\left[||u_t||^2 \right] dt,$$

where u_t ranges over all adapted drifts for which $B_1 + \int_0^1 u_t dt$ has the same law as μ.

We denote

$$X_t := B_t + \int_0^t v_s ds.$$

Bounding the Deficit - the Föllmer Drift

Our central construct will be the Föllmer drift, which is the solution to the following variational problem:

$$v_t := \arg\min_{u_t} \frac{1}{2} \int_0^1 \mathbb{E}\left[||u_t||^2\right] dt,$$

where u_t ranges over all adapted drifts for which $B_1 + \int_0^1 u_t dt$ has the same law as μ.
We denote

$$X_t := B_t + \int_0^t v_s ds.$$

Bounding the Deficit - the Föllmer Drift

The process v_t goes back at least to the works of Föllmer (86'). In a later work by Lehec (12') it is shown that if μ has finite entropy relative to γ, then v_t is well defined and that:

1. v_t is a martingale, with $v_t(X_t) = \nabla \ln \left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right)$.
2. $\mathrm{Ent}\left(\mu||\gamma\right) = \mathrm{Ent}\left(X_.||B_.\right) = \frac{1}{2}\int\limits_0^1 \mathbb{E}[||v_t||^2]dt$.
3. In the Wiener space, the density of X_t with respect to B_t is given by $\frac{d\mu}{d\gamma}(\omega_1)$.
4. If $G \sim \gamma$, independent from X_1,

$$X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G.$$

Bounding the Deficit - the Föllmer Drift

The process v_t goes back at least to the works of Föllmer (86'). In a later work by Lehec (12') it is shown that if μ has finite entropy relative to γ, then v_t is well defined and that:

1. v_t is a martingale, with $v_t(X_t) = \nabla \ln \left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right)$.
2. $\mathrm{Ent}(\mu||\gamma) = \mathrm{Ent}(X.||B.) = \frac{1}{2}\int\limits_0^1 \mathbb{E}[||v_t||^2]dt$.
3. In the Wiener space, the density of X_t with respect to B_t is given by $\frac{d\mu}{d\gamma}(\omega_1)$.
4. If $G \sim \gamma$, independent from X_1,

$$X_t \stackrel{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G.$$

Bounding the Deficit - the Föllmer Drift

The process v_t goes back at least to the works of Föllmer (86'). In a later work by Lehec (12') it is shown that if μ has finite entropy relative to γ, then v_t is well defined and that:

1. v_t is a martingale, with $v_t(X_t) = \nabla \ln\left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right)$.
2. $\mathrm{Ent}\left(\mu||\gamma\right) = \mathrm{Ent}\left(X_\cdot||B_\cdot\right) = \frac{1}{2}\int\limits_0^1 \mathbb{E}[||v_t||^2]dt.$
3. In the Wiener space, the density of X_t with respect to B_t is given by $\frac{d\mu}{d\gamma}(\omega_1)$.
4. If $G \sim \gamma$, independent from X_1,

$$X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G.$$

Bounding the Deficit - the Föllmer Drift

The process v_t goes back at least to the works of Föllmer (86'). In a later work by Lehec (12') it is shown that if μ has finite entropy relative to γ, then v_t is well defined and that:

1. v_t is a martingale, with $v_t(X_t) = \nabla \ln\left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right)$.
2. $\mathrm{Ent}\left(\mu||\gamma\right) = \mathrm{Ent}\left(X.||B.\right) = \frac{1}{2}\int\limits_0^1 \mathbb{E}[||v_t||^2]dt$.
3. In the Wiener space, the density of X_t with respect to B_t is given by $\frac{d\mu}{d\gamma}(\omega_1)$.
4. If $G \sim \gamma$, independent from X_1,

$$X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G.$$

Bounding the Deficit - the Föllmer Drift

The process v_t goes back at least to the works of Föllmer (86'). In a later work by Lehec (12') it is shown that if μ has finite entropy relative to γ, then v_t is well defined and that:

1. v_t is a martingale, with $v_t(X_t) = \nabla \ln\left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right)$.
2. $\mathrm{Ent}\left(\mu||\gamma\right) = \mathrm{Ent}\left(X_\cdot||B_\cdot\right) = \frac{1}{2}\int\limits_0^1 \mathbb{E}[||v_t||^2]dt$.
3. In the Wiener space, the density of X_t with respect to B_t is given by $\frac{d\mu}{d\gamma}(\omega_1)$.
4. If $G \sim \gamma$, independent from X_1,

$$X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G.$$

Proof of Talagrand's Inequality

Proof of Talagrand's Inequality (Lehec).

$$\mathcal{W}_2^2(\mu||\gamma) \leq \mathbb{E}\left[\left|\left|X_1 - B_1\right|\right|_2^2\right] = \mathbb{E}\left[\left|\left|\int_0^1 v_t dt\right|\right|_2^2\right]$$
$$\leq \int_0^1 \mathbb{E}\left[||v_t||_2^2\right] dt = 2\mathrm{Ent}(\mu||\gamma).$$

□

The goal is to make this quantitative.

Proof of Talagrand's Inequality

Proof of Talagrand's Inequality (Lehec).

$$\mathcal{W}_2^2(\mu||\gamma) \leq \mathbb{E}\left[\left|\left|X_1 - B_1\right|\right|_2^2\right] = \mathbb{E}\left[\left|\left|\int_0^1 v_t dt\right|\right|_2^2\right]$$
$$\leq \int_0^1 \mathbb{E}\left[||v_t||_2^2\right] dt = 2\mathrm{Ent}(\mu||\gamma).$$

□

The goal is to make this quantitative.

Stability for Measures with a Finite Poincaré Constant

We say that μ satisfies a Poincaré inequality, with constant $\mathrm{C_p}(\mu)$, if for every every smooth function f,

$$\mathrm{Var}_\mu(f) \leq \mathrm{C_p}(\mu)\mathbb{E}_\mu\left[||\nabla f||_2^2\right].$$

We will prove:

Theorem

Let μ be a centered measure on $\mathbb{R}^d$ with $\mathrm{C_p}(\mu) < \infty$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}\mathrm{Ent}(\mu||\gamma).$$

Stability for Measures with a Finite Poincaré Constant

We say that μ satisfies a Poincaré inequality, with constant $\mathrm{C_p}(\mu)$, if for every every smooth function f,

$$\mathrm{Var}_\mu(f) \leq \mathrm{C_p}(\mu)\mathbb{E}_\mu\left[||\nabla f||_2^2\right].$$

We will prove:

Theorem

Let μ be a centered measure on $\mathbb{R}^d$ with $\mathrm{C_p}(\mu) < \infty$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}\mathrm{Ent}(\mu||\gamma).$$

Measures with a Finite Poincaré Constant

The Poincaré constant is inequality for the following comparison lemma:

Lemma

Assume that μ *is centered and that* $\mathrm{C_p}(\mu) < \infty$*. Then*

- *For* $0 \leq t \leq \frac{1}{2}$,

$$\mathbb{E}\left[||v_t||_2^2\right] \leq \mathbb{E}\left[||v_{1/2}||_2^2\right] \frac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}.$$

- *For* $\frac{1}{2} \leq t \leq 1$,

$$\mathbb{E}\left[||v_t||_2^2\right] \geq \mathbb{E}\left[||v_{1/2}||_2^2\right] \frac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}.$$

Proof.

Recall $X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G$. Hence,

$$\mathrm{C_p}(X_t) \leq t^2 \mathrm{C_p}(\mu) + t(1-t),$$

and

$$\mathbb{E}\left[||v_t(X_t)||_2^2\right] \leq (t^2\mathrm{C_p}(\mu) + t(1-t))\mathbb{E}\left[||\nabla v_t(X_t)||_2^2\right]$$
$$= (t^2\mathrm{C_p}(\mu) + t(1-t))\frac{d}{dt}\mathbb{E}\left[||v_t(X_t)||_2^2\right].$$

$g(t) := \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \frac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}$ solves

$$f(t) = t^2\mathrm{C_p}(\mu) + t(1-t)f'(t), \text{ with } f\left(\frac{1}{2}\right) = \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right].$$

Now apply Gromwall's inequality. □

Proof.

Recall $X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G$. Hence,

$$\mathrm{C_p}(X_t) \leq t^2\mathrm{C_p}(\mu) + t(1-t),$$

and

$$\mathbb{E}\left[||v_t(X_t)||_2^2\right] \leq (t^2\mathrm{C_p}(\mu) + t(1-t))\mathbb{E}\left[||\nabla v_t(X_t)||_2^2\right]$$
$$= (t^2\mathrm{C_p}(\mu) + t(1-t))\frac{d}{dt}\mathbb{E}\left[||v_t(X_t)||_2^2\right].$$

$g(t) := \mathbb{E}\left[||v_{1/2}||_2^2\right] \frac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}$ solves

$$f(t) = t^2\mathrm{C_p}(\mu) + t(1-t)f'(t), \text{ with } f\left(\frac{1}{2}\right) = \mathbb{E}\left[||v_{1/2}||_2^2\right].$$

Now apply Gromwall's inequality. □

Proof.

Recall $X_t \stackrel{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G$. Hence,

$$\mathrm{C_p}(X_t) \leq t^2 \mathrm{C_p}(\mu) + t(1-t),$$

and

$$\begin{aligned}\mathbb{E}\left[||v_t(X_t)||_2^2\right] &\leq (t^2\mathrm{C_p}(\mu) + t(1-t))\mathbb{E}\left[||\nabla v_t(X_t)||_2^2\right] \\ &= (t^2\mathrm{C_p}(\mu) + t(1-t))\frac{d}{dt}\mathbb{E}\left[||v_t(X_t)||_2^2\right].\end{aligned}$$

$g(t) := \mathbb{E}\left[||v_{1/2}||_2^2\right] \dfrac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}$ solves

$$f(t) = t^2\mathrm{C_p}(\mu) + t(1-t)f'(t), \text{ with } f\left(\frac{1}{2}\right) = \mathbb{E}\left[||v_{1/2}||_2^2\right].$$

Now apply Gromwall's inequality. □

Proof.

Recall $X_t \overset{\text{law}}{=} tX_1 + \sqrt{t(1-t)}G$. Hence,

$$\mathrm{C_p}(X_t) \leq t^2 \mathrm{C_p}(\mu) + t(1-t),$$

and

$$\begin{aligned}\mathbb{E}\left[||v_t(X_t)||_2^2\right] &\leq (t^2\mathrm{C_p}(\mu) + t(1-t))\mathbb{E}\left[||\nabla v_t(X_t)||_2^2\right] \\ &= (t^2\mathrm{C_p}(\mu) + t(1-t))\frac{d}{dt}\mathbb{E}\left[||v_t(X_t)||_2^2\right].\end{aligned}$$

$g(t) := \mathbb{E}\left[||v_{1/2}||_2^2\right] \dfrac{(\mathrm{C_p}(\mu)+1)\,t}{(\mathrm{C_p}(\mu)-1)\,t+1}$ solves

$$f(t) = t^2\mathrm{C_p}(\mu) + t(1-t)f'(t), \text{ with } f\left(\frac{1}{2}\right) = \mathbb{E}\left[||v_{1/2}||_2^2\right].$$

Now apply Gromwall's inequality. □

A Martingale Formulation

We will use the following martingale formulation:

$$Y_t := \mathbb{E}\left[X_1 | \mathcal{F}_t\right].$$

By the martingale representation theorem, for some process Γ_t, which is uniquely defined, Y_t satisfies

$$Y_t = \int_0^t \Gamma_s dB_s.$$

This implies

$$v_t = \int_0^t \frac{\Gamma_s - \mathrm{I}_d}{1 - s} dB_s.$$

A Martingale Formulation

We will use the following martingale formulation:

$$Y_t := \mathbb{E}\left[X_1 | \mathcal{F}_t\right].$$

By the martingale representation theorem, for some process Γ_t, which is uniquely defined, Y_t satisfies

$$Y_t = \int_0^t \Gamma_s dB_s.$$

This implies

$$v_t = \int_0^t \frac{\Gamma_s - \mathrm{I}_d}{1 - s} dB_s.$$

A Martingale Formulation

We will use the following martingale formulation:

$$Y_t := \mathbb{E}\left[X_1 | \mathcal{F}_t\right].$$

By the martingale representation theorem, for some process Γ_t, which is uniquely defined, Y_t satisfies

$$Y_t = \int_0^t \Gamma_s dB_s.$$

This implies

$$v_t = \int_0^t \frac{\Gamma_s - \mathrm{I}_d}{1 - s} dB_s.$$

A Martingale Formulation

It turns out that Γ_t is a positive definite matrix, hence

$$\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\int_0^1 \mathbb{E}\left[||v_s||_2^2\right] ds = \frac{1}{2}\mathrm{Tr}\int_0^1\int_0^s \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dtds$$
$$= \frac{1}{2}\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt,$$

and

$$\mathcal{W}_2^2(\mu,\gamma) \leq \mathbb{E}\left[\left|\left|\int_0^1 \Gamma_t dB_t - \int_0^1 dB_t\right|\right|_2^2\right] = \mathrm{Tr}\int_0^1 \mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right] dt.$$

A Martingale Formulation

It turns out that Γ_t is a positive definite matrix, hence

$$\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\int_0^1 \mathbb{E}\left[||v_s||_2^2\right] ds = \frac{1}{2}\mathrm{Tr}\int_0^1\int_0^s \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt ds$$
$$= \frac{1}{2}\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt,$$

and

$$\mathcal{W}_2^2(\mu,\gamma) \le \mathbb{E}\left[\left|\left|\int_0^1 \Gamma_t dB_t - \int_0^1 dB_t\right|\right|_2^2\right] = \mathrm{Tr}\int_0^1 \mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right] dt.$$

Bounding the Deficit - Martingales

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu,\gamma) \geq \mathrm{Tr}\int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t}dt$$

Integration by parts gives:

$$\delta_{\mathrm{Tal}}(\mu) \geq \mathrm{Tr}\int_0^1 t(1-t) \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2}dt$$

$$= \int_0^1 t(1-t)\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right]dt = \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right]dt$$

Bounding the Deficit - Martingales

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu,\gamma) \geq \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt$$

Integration by parts gives:

$$\delta_{\mathrm{Tal}}(\mu) \geq \mathrm{Tr} \int_0^1 t(1-t) \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt$$

$$= \int_0^1 t(1-t)\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right] dt = \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$

Bounding the Deficit - Martingales

$$\delta_{\mathrm{Tal}}(\mu) = 2\mathrm{Ent}(\mu||\gamma) - \mathcal{W}_2^2(\mu,\gamma) \geq \mathrm{Tr}\int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt$$

Integration by parts gives:

$$\delta_{\mathrm{Tal}}(\mu) \geq \mathrm{Tr}\int_0^1 t(1-t) \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt$$

$$= \int_0^1 t(1-t)\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right] dt = \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$

Applying the Lemma

$$\delta_{\mathrm{Tal}}(\mu) \geq \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \int_0^1 \frac{(2t-1)\left(\mathrm{C_p}(\mu)+1\right)t}{\left(\mathrm{C_p}(\mu)-1\right)t+1} dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}$$

If $\mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \geq \mathrm{Ent}(\mu||\gamma)$, this shows

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)} \mathrm{Ent}(\mu||\gamma).$$

The other case is easier.

Applying the Lemma

$$\delta_{\mathrm{Tal}}(\mu) \geq \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \int_0^1 \frac{(2t-1)\left(\mathrm{C_p}(\mu)+1\right)t}{\left(\mathrm{C_p}(\mu)-1\right)t+1} dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}$$

If $\mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \geq \mathrm{Ent}(\mu||\gamma)$, this shows

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)} \mathrm{Ent}(\mu||\gamma).$$

The other case is easier.

Applying the Lemma

$$\delta_{\mathrm{Tal}}(\mu) \geq \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$
$$\geq \mathbb{E}\left[||v_{1/2}||_2^2\right] \int_0^1 \frac{(2t-1)\left(\mathrm{C_p}(\mu)+1\right)t}{\left(\mathrm{C_p}(\mu)-1\right)t+1} dt$$
$$\geq \mathbb{E}\left[||v_{1/2}||_2^2\right] \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}$$

If $\mathbb{E}\left[||v_{1/2}||_2^2\right] \geq \mathrm{Ent}(\mu||\gamma)$, this shows

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)} \mathrm{Ent}(\mu||\gamma).$$

The other case is easier.

Applying the Lemma

$$\delta_{\mathrm{Tal}}(\mu) \geq \int_0^1 (2t-1)\mathbb{E}\left[||v_t||_2^2\right] dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \int_0^1 \frac{(2t-1)\left(\mathrm{C_p}(\mu)+1\right)t}{\left(\mathrm{C_p}(\mu)-1\right)t+1} dt$$

$$\geq \mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}$$

If $\mathbb{E}\left[\left|\left|v_{1/2}\right|\right|_2^2\right] \geq \mathrm{Ent}(\mu||\gamma)$, this shows

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{\ln(\mathrm{C_p}(\mu)+1)}{4\mathrm{C_p}(\mu)}\mathrm{Ent}(\mu||\gamma).$$

The other case is easier.

Further Results

Other bounds on $\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right]$, will yields different results.

For example, if $\mathrm{tr}\left(\mathrm{Cov}(\mu)\right) \leq d$, then

$$\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right] \geq \frac{\left(\mathbb{E}\left[||v_t||_2^2\right]\right)^2}{d}.$$

This gives:

Theorem

Let μ be a measure on $\mathbb{R}^d$ such that $\mathrm{tr}\left(\mathrm{Cov}(\mu)\right) \leq d$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \min\left(\frac{\mathrm{Ent}(\mu||\gamma)^2}{6d}, \frac{\mathrm{Ent}(\mu||\gamma)}{4}\right).$$

Further Results

Other bounds on $\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right]$, will yields different results.
For example, if $\mathrm{tr}\left(\mathrm{Cov}(\mu)\right) \leq d$, then

$$\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right] \geq \frac{\left(\mathbb{E}\left[||v_t||_2^2\right]\right)^2}{d}.$$

This gives:

Theorem

Let μ be a measure on $\mathbb{R}^d$ such that $\mathrm{tr}\left(\mathrm{Cov}(\mu)\right) \leq d$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \min\left(\frac{\mathrm{Ent}(\mu||\gamma)^2}{6d}, \frac{\mathrm{Ent}(\mu||\gamma)}{4}\right).$$

Further Results

Other bounds on $\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right]$, will yields different results.
For example, if $\operatorname{tr}(\operatorname{Cov}(\mu)) \leq d$, then

$$\frac{d}{dt}\mathbb{E}\left[||v_t||_2^2\right] \geq \frac{\left(\mathbb{E}\left[||v_t||_2^2\right]\right)^2}{d}.$$

This gives:

Theorem

Let μ be a measure on $\mathbb{R}^d$ such that $\operatorname{tr}(\operatorname{Cov}(\mu)) \leq d$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \min\left(\frac{\mathrm{Ent}(\mu||\gamma)^2}{6d}, \frac{\mathrm{Ent}(\mu||\gamma)}{4}\right).$$

Further Results

Two other results:

Theorem

Let μ be a measure on $\mathbb{R}^d$ and let $\{\lambda_i\}_{i=1}^d$ be the eigenvalues of $\mathrm{Cov}(\mu)$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \sum_{i=1}^{d} \frac{2(1-\lambda_i) + (\lambda_i + 1)\ln(\lambda_i)}{\lambda_i - 1} \mathbb{1}_{\{\lambda_i < 1\}}.$$

Theorem

Let μ be a measure on $\mathbb{R}^d$. There exists another measure ν such that

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{1}{3\sqrt{3}} \frac{\mathrm{Ent}(\mu||\gamma)^{3/2}}{\sqrt{d}}$$

Further Results

Two other results:

Theorem

Let μ be a measure on $\mathbb{R}^d$ and let $\{\lambda_i\}_{i=1}^d$ be the eigenvalues of $\mathrm{Cov}(\mu)$. Then

$$\delta_{\mathrm{Tal}}(\mu) \geq \sum_{i=1}^{d} \frac{2(1-\lambda_i) + (\lambda_i + 1)\ln(\lambda_i)}{\lambda_i - 1} \mathbb{1}_{\{\lambda_i < 1\}}.$$

Theorem

Let μ be a measure on $\mathbb{R}^d$. There exists another measure ν such that

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{1}{3\sqrt{3}} \frac{\mathrm{Ent}(\mu||\gamma)^{3/2}}{\sqrt{d}}$$

Further Results

Two other results:

Theorem

Let μ *be a measure on* $\mathbb{R}^d$ *and let* $\{\lambda_i\}_{i=1}^d$ *be the eigenvalues of* $\mathrm{Cov}(\mu)$*. Then*

$$\delta_{\mathrm{Tal}}(\mu) \geq \sum_{i=1}^{d} \frac{2(1-\lambda_i) + (\lambda_i + 1)\ln(\lambda_i)}{\lambda_i - 1} \mathbb{1}_{\{\lambda_i < 1\}}.$$

Theorem

Let μ *be a measure on* $\mathbb{R}^d$*. There exists another measure* ν *such that*

$$\delta_{\mathrm{Tal}}(\mu) \geq \frac{1}{3\sqrt{3}} \frac{\mathrm{Ent}(\mu||\gamma)^{3/2}}{\sqrt{d}}$$

Log-Sobolev Inequality

Definition (Fisher information of μ with respect to γ)

$$\mathrm{I}(\mu||\gamma) = \mathbb{E}_{\mu}\left[\left|\left|\nabla \ln\left(\frac{d\mu}{d\gamma}\right)\right|\right|_2^2\right].$$

In 75' Gross proved:

Theorem (Log-Sobolev inequality)

Let μ be a measure on $\mathbb{R}^d$. Then

$$2\mathrm{Ent}(\mu||\gamma) \leq \mathrm{I}(\mu||\gamma).$$

Log-Sobolev Inequality

Definition (Fisher information of μ with respect to γ)

$$\mathrm{I}(\mu||\gamma) = \mathbb{E}_\mu \left[\left|\left| \nabla \ln \left(\frac{d\mu}{d\gamma} \right) \right|\right|_2^2 \right].$$

In 75' Gross proved:

Theorem (Log-Sobolev inequality)

Let μ be a measure on $\mathbb{R}^d$. Then

$$2\mathrm{Ent}(\mu||\gamma) \leq \mathrm{I}(\mu||\gamma).$$

Define

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{I}(\mu||\gamma) - 2\mathrm{Ent}(\mu||\gamma),$$

and recall

$$v_t := v_t(X_t) = \nabla \ln \left(P_{1-t} \left(\frac{d\mu}{d\gamma}(X_t) \right) \right).$$

It follows that

$$\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt = \mathbb{E}\left[||v_1||_2^2\right] = \mathrm{I}(\mu||\gamma).$$

Since $\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt$, we get

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt.$$

Define

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{I}(\mu||\gamma) - 2\mathrm{Ent}(\mu||\gamma),$$

and recall

$$v_t := v_t(X_t) = \nabla \ln\left(P_{1-t}\left(\frac{d\mu}{d\gamma}(X_t)\right)\right).$$

It follows that

$$\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2}dt = \mathbb{E}\left[||v_1||_2^2\right] = \mathrm{I}(\mu||\gamma).$$

Since $\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t}dt$, we get

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{Tr}\int_0^1 t\cdot\frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2}dt.$$

Define

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{I}(\mu||\gamma) - 2\mathrm{Ent}(\mu||\gamma),$$

and recall

$$v_t := v_t(X_t) = \nabla \ln \left(P_{1-t} \left(\frac{d\mu}{d\gamma}(X_t) \right) \right).$$

It follows that

$$\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2 \right]}{(1-t)^2} dt = \mathbb{E}\left[||v_1||_2^2 \right] = \mathrm{I}(\mu||\gamma).$$

Since $\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2 \right]}{1-t} dt$, we get

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2 \right]}{(1-t)^2} dt.$$

Define

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{I}(\mu||\gamma) - 2\mathrm{Ent}(\mu||\gamma),$$

and recall

$$v_t := v_t(X_t) = \nabla \ln \left(P_{1-t} \left(\frac{d\mu}{d\gamma}(X_t) \right) \right).$$

It follows that

$$\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt = \mathbb{E}\left[||v_1||_2^2\right] = \mathrm{I}(\mu||\gamma).$$

Since $\mathrm{Ent}(\mu||\gamma) = \frac{1}{2}\mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{1-t} dt$, we get

$$\delta_{\mathrm{LS}}(\mu) = \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t - \mathrm{I}_d)^2\right]}{(1-t)^2} dt.$$

The Shannon-Stam Inequality

In $48'$ Shannon noted the following inequality, which was later proved by Stam, in $56'$.

Theorem (Shannon-Stam Inequality)

Let X, Y be independent random vectors in $\mathbb{R}^d$ and let $G \sim \gamma$. Then, for any $\lambda \in [0,1]$,

$$\mathrm{Ent}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y||G) \leq \lambda \mathrm{Ent}(X||G) + (1-\lambda)\mathrm{Ent}(Y||G).$$

Moreover, equality holds if and only if X and Y are Gaussians with identical covariances.

Define

$$\delta_\lambda(X,Y) = \lambda \mathrm{Ent}(X||G) + (1-\lambda)\mathrm{Ent}(Y||G) - \mathrm{Ent}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y||G).$$

The Shannon-Stam Inequality

In $48'$ Shannon noted the following inequality, which was later proved by Stam, in $56'$.

Theorem (Shannon-Stam Inequality)

Let X, Y be independent random vectors in $\mathbb{R}^d$ and let $G \sim \gamma$. Then, for any $\lambda \in [0, 1]$,

$$\mathrm{Ent}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y || G) \leq \lambda \mathrm{Ent}(X||G) + (1-\lambda)\mathrm{Ent}(Y||G).$$

Moreover, equality holds if and only if X and Y are Gaussians with identical covariances.

Define

$$\delta_\lambda(X, Y) = \lambda \mathrm{Ent}(X||G) + (1-\lambda)\mathrm{Ent}(Y||G) - \mathrm{Ent}(\sqrt{\lambda}X + \sqrt{1-\lambda}Y||G).$$

Deficit of the Shannon-Stam Inequality

For simplicity we'll focus on the case $\lambda = \frac{1}{2}$.
Now, for X, Y independent random variables, take two independent Brownian motions B_t^X, B_t^Y and Γ_t^X, Γ_t^Y as above.
We get

$$\frac{X+Y}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\int_0^1 \Gamma_t^X dB_t^X + \int_0^1 \Gamma_t^Y dB_t^Y\right) \stackrel{\text{law}}{=} \int_0^1 \sqrt{\frac{(\Gamma_t^X)^2 + (\Gamma_t^Y)^2}{2}} dB_t.$$

for some Brownian motion B_t.

Deficit of the Shannon-Stam Inequality

For simplicity we'll focus on the case $\lambda = \frac{1}{2}$.
Now, for X, Y independent random variables, take two independent Brownian motions B_t^X, B_t^Y and Γ_t^X, Γ_t^Y as above.
We get

$$\frac{X+Y}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\int_0^1 \Gamma_t^X dB_t^X + \int_0^1 \Gamma_t^Y dB_t^Y\right) \overset{\text{law}}{=} \int_0^1 \sqrt{\frac{(\Gamma_t^X)^2 + (\Gamma_t^Y)^2}{2}} dB_t.$$

for some Brownian motion B_t.

Bounding the Deficit

If $H_t = \sqrt{\frac{(\Gamma_t^X)^2 + (\Gamma_t^Y)^2}{2}}$, $\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}\middle|\middle| G\right) \leq \frac{1}{2}\mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t} dt.$

Consequently,

$$2\delta_{\frac{1}{2}}(X,Y) \geq \mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^Y)^2\right]}{2(1-t)} + \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^X)^2\right]}{2(1-t)} - \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t} dt$$

$$= \mathrm{Tr}\int\limits_0^1 \frac{2\mathbb{E}[H_t] - \mathbb{E}[\Gamma_t^X] - \mathbb{E}[\Gamma_t^Y]}{1-t}.$$

Manipulating the matrix square root then shows

$$\delta_{\frac{1}{2}}(X,Y) \gtrsim \mathrm{Tr}\int\limits_0^1 \mathbb{E}\left[\frac{(\Gamma_t^X - \Gamma_t^Y)^2(\Gamma_t^X + \Gamma_t^Y)^{-1}}{(1-t)}\right] dt.$$

Bounding the Deficit

If $H_t = \sqrt{\frac{(\Gamma_t^X)^2+(\Gamma_t^Y)^2}{2}}$, $\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}||G\right) \leq \frac{1}{2}\mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t}dt.$

Consequently,

$$2\delta_{\frac{1}{2}}(X,Y) \geq \mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^Y)^2\right]}{2(1-t)} + \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^X)^2\right]}{2(1-t)} - \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t}dt$$

$$= \mathrm{Tr}\int\limits_0^1 \frac{2\mathbb{E}[H_t] - \mathbb{E}[\Gamma_t^X] - \mathbb{E}[\Gamma_t^Y]}{1-t}.$$

Manipulating the matrix square root then shows

$$\delta_{\frac{1}{2}}(X,Y) \gtrsim \mathrm{Tr}\int\limits_0^1 \mathbb{E}\left[\frac{(\Gamma_t^X - \Gamma_t^Y)^2(\Gamma_t^X + \Gamma_t^Y)^{-1}}{(1-t)}\right]dt.$$

Bounding the Deficit

If $H_t = \sqrt{\frac{(\Gamma_t^X)^2+(\Gamma_t^Y)^2}{2}}$, $\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}||G\right) \leq \frac{1}{2}\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t} dt.$

Consequently,

$$2\delta_{\frac{1}{2}}(X,Y) \geq \mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^Y)^2\right]}{2(1-t)} + \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^X)^2\right]}{2(1-t)} - \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t} dt$$

$$= \mathrm{Tr}\int_0^1 \frac{2\mathbb{E}[H_t] - \mathbb{E}[\Gamma_t^X] - \mathbb{E}[\Gamma_t^Y]}{1-t}.$$

Manipulating the matrix square root then shows

$$\delta_{\frac{1}{2}}(X,Y) \gtrsim \mathrm{Tr}\int_0^1 \mathbb{E}\left[\frac{(\Gamma_t^X - \Gamma_t^Y)^2(\Gamma_t^X + \Gamma_t^Y)^{-1}}{(1-t)}\right] dt.$$

Bounding the Deficit

If $H_t = \sqrt{\frac{(\Gamma_t^X)^2+(\Gamma_t^Y)^2}{2}}$, $\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}||G\right) \leq \frac{1}{2}\mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t}dt.$

Consequently,

$$2\delta_{\frac{1}{2}}(X,Y) \geq \mathrm{Tr}\int\limits_0^1 \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^Y)^2\right]}{2(1-t)} + \frac{\mathbb{E}\left[(\mathrm{I}_d - \Gamma_t^X)^2\right]}{2(1-t)} - \frac{\mathbb{E}\left[(\mathrm{I}_d - H_t)^2\right]}{1-t}dt$$

$$= \mathrm{Tr}\int\limits_0^1 \frac{2\mathbb{E}[H_t] - \mathbb{E}[\Gamma_t^X] - \mathbb{E}[\Gamma_t^Y]}{1-t}.$$

Manipulating the matrix square root then shows

$$\delta_{\frac{1}{2}}(X,Y) \gtrsim \mathrm{Tr}\int\limits_0^1 \mathbb{E}\left[\frac{(\Gamma_t^X - \Gamma_t^Y)^2(\Gamma_t^X + \Gamma_t^Y)^{-1}}{(1-t)}\right]dt.$$

Deficit of Log-Concave Measures

Fact: if X is log-concave, then $\Gamma_t^X \preceq \frac{1}{t}\mathrm{I}_d$ almost surely.

So, if both X and Y are log-concave,

$$\delta_{\frac{1}{2}}(X, Y) \gtrsim \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \Gamma_t^Y)^2\right]}{1-t} dt.$$

In particular,

$$\delta_{\frac{1}{2}}(X, G) \gtrsim \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \mathrm{I}_d)^2\right]}{1-t} dt.$$

Deficit of Log-Concave Measures

Fact: if X is log-concave, then $\Gamma_t^X \preceq \frac{1}{t}\mathrm{I}_d$ almost surely.

So, if both X and Y are log-concave,

$$\delta_{\frac{1}{2}}(X, Y) \gtrsim \mathrm{Tr}\int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \Gamma_t^Y)^2\right]}{1-t}dt.$$

In particular,

$$\delta_{\frac{1}{2}}(X, G) \gtrsim \mathrm{Tr}\int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \mathrm{I}_d)^2\right]}{1-t}dt.$$

Deficit of Log-Concave Measures

Fact: if X is log-concave, then $\Gamma_t^X \preceq \frac{1}{t}\mathrm{I}_d$ almost surely.
So, if both X and Y are log-concave,

$$\delta_{\frac{1}{2}}(X, Y) \gtrsim \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \Gamma_t^Y)^2\right]}{1-t} dt.$$

In particular,

$$\delta_{\frac{1}{2}}(X, G) \gtrsim \mathrm{Tr} \int_0^1 t \cdot \frac{\mathbb{E}\left[(\Gamma_t^X - \mathrm{I}_d)^2\right]}{1-t} dt.$$

The Entropic Central Limit Theorem

Let $\{X_i\}$ be *i.i.d.* copies of X and $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$.

Set $H_t = \sqrt{\frac{\sum (\Gamma_t^i)^2}{n}}$. Then

$$S_n \overset{\text{law}}{=} \int_0^1 H_t dB_t.$$

Using this, we show

$$\mathrm{Ent}(S_n || G) \leq C_X \mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(H_t - \mathbb{E}[H_t])^2\right]}{1-t} dt,$$

where $C_X > 0$, depends on X. This can be used to prove the entropic central limit theorem.

The Entropic Central Limit Theorem

Let $\{X_i\}$ be *i.i.d.* copies of X and $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$.

Set $H_t = \sqrt{\frac{\sum (\Gamma_t^i)^2}{n}}$. Then

$$S_n \overset{\text{law}}{=} \int_0^1 H_t dB_t.$$

Using this, we show

$$\mathrm{Ent}(S_n || G) \leq C_X \mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(H_t - \mathbb{E}[H_t])^2\right]}{1-t} dt,$$

where $C_X > 0$, depends on X. This can be used to prove the entropic central limit theorem.

The Entropic Central Limit Theorem

Let $\{X_i\}$ be *i.i.d.* copies of X and $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$.

Set $H_t = \sqrt{\frac{\sum(\Gamma_t^i)^2}{n}}$. Then

$$S_n \stackrel{\text{law}}{=} \int_0^1 H_t dB_t.$$

Using this, we show

$$\mathrm{Ent}(S_n||G) \leq C_X \mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(H_t - \mathbb{E}[H_t])^2\right]}{1-t} dt,$$

where $C_X > 0$, depends on X. This can be used to prove the entropic central limit theorem.

The Entropic Central Limit Theorem

Let $\{X_i\}$ be *i.i.d.* copies of X and $S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$.

Set $H_t = \sqrt{\frac{\sum(\Gamma_t^i)^2}{n}}$. Then

$$S_n \overset{\text{law}}{=} \int_0^1 H_t dB_t.$$

Using this, we show

$$\mathrm{Ent}(S_n || G) \leq C_X \mathrm{Tr} \int_0^1 \frac{\mathbb{E}\left[(H_t - \mathbb{E}[H_t])^2\right]}{1-t} dt,$$

where $C_X > 0$, depends on X. This can be used to prove the entropic central limit theorem.

Quantitative Entropic Central Limit Theorem

For a more quantitative result we have the formula

$$\mathrm{Ent}(S_n||G) \leq \frac{\mathrm{poly}(\mathrm{C_p}(X))}{n}\mathrm{Tr}\int_0^1 \frac{\mathbb{E}\left[\left(\Gamma_t^2 - \mathbb{E}\left[H_t^2\right]\right)^2\right]}{1-t}dt,$$

$$= \frac{\mathrm{poly}(\mathrm{C_p}(X))}{n}\mathrm{Tr}\int_0^1 \frac{\mathrm{Var}(\Gamma_t^2)}{1-t}dt,$$

valid for X which satisfies a Poincaré inequality. For X log-concave, $\Gamma_t \preceq \frac{1}{t}\mathrm{I}_d$, and

$$\mathrm{Tr}\int_0^1 \frac{\mathrm{Var}(\Gamma_t^2)}{1-t}dt \leq \mathrm{Tr}\int_0^1 \frac{1}{t^2}\frac{\mathbb{E}\left[\left(\Gamma_t - \mathrm{I}_d\right)^2\right]}{1-t}dt.$$

Thank You