A stochastic approach for noise stability on the hypercube

Dan Mikulincer

MIT

Joint work with Ronen Eldan and Prasad Raghavendra

Outline

1. **Noise stability**
2. “Majority is Stablest”
3. A stochastic approach for noise stability via a re-normalized Brownian motion
4. Interlude - from a toy example to the Courtade-Kumar conjecture
5. Back to Majority is Stablest

Noise Operators

Consider the discrete hypercube $\mathcal{C}_n = \{-1, 1\}^n$ with its uniform probability measure μ.

For $\rho \in (0, 1)$ define the noise operator T_ρ, by

$$T_\rho f(x) = \mathbb{E}_{y \sim \rho \text{ correlated with } x} [f(y)] \,.$$

We say that y is ρ correlated with x if $\mathbb{E}[y_i x_i] = \rho$. In other words, the law of y is the unique product measure with $\mathbb{E}[y] = \rho x$.

Noise Operators

Consider the discrete hypercube $\mathcal{C}_n = \{-1, 1\}^n$ with its uniform probability measure μ.

For $\rho \in (0, 1)$ define the noise operator T_ρ, by

$$T_\rho f(x) = \mathbb{E}_{y \sim \rho \text{ correlated with } x} [f(y)] .$$

We say that y is ρ correlated with x if $\mathbb{E}[y_i x_i] = \rho$. In other words, the law of y is the unique product measure with $\mathbb{E}[y] = \rho x$.

Noise Stability

For a Boolean function $f : \mathcal{C}_n \to \{-1, 1\}$, define its noise stability by,

$$\mathrm{Stab}_\rho(f) := \mathbb{E}_\mu [f T_\rho f] .$$

Important concept in social choice theory and Boolean analysis.

Example:

Theorem (Kalai 02')

If $f : \mathcal{C}_n \to \{-1, 1\}$ is used to rank three candidates,

$$\mathbb{P}_\mu (f \text{ gives a rational outcome}) = \frac{3}{4}(1 + \mathrm{Stab}_{\frac{1}{3}}(f)).$$

Noise Stability

For a Boolean function $f : \mathcal{C}_n \to \{-1, 1\}$, define its noise stability by,

$$\mathrm{Stab}_\rho(f) := \mathbb{E}_\mu \left[f T_\rho f \right].$$

Important concept in social choice theory and Boolean analysis.

Example:

Theorem (Kalai 02')

If $f : \mathcal{C}_n \to \{-1, 1\}$ is used to rank three candidates,

$$\mathbb{P}_\mu \left(f \textit{ gives a rational outcome} \right) = \frac{3}{4}(1 + \mathrm{Stab}_{\frac{1}{3}}(f)).$$

Noise Stability

Question

Among all Boolean functions, which one maximizes the noise stability?

Easy answer: among all Boolean functions the dictator $f(x) := x_1$ has the largest noise stability.

Not a very useful fact in social choice theory.

Define the maximal influence of a Boolean function by:

$$\inf = \max_{i \in [n]} \mathbb{E}_\mu \left[(\partial_i f)^2 \right].$$

Noise Stability

Question

Among all Boolean functions, which one maximizes the noise stability?

Easy answer: among all Boolean functions the dictator $f(x) := x_1$ has the largest noise stability.

Not a very useful fact in social choice theory.

Define the maximal influence of a Boolean function by:

$$\inf = \max_{i \in [n]} \mathbb{E}_\mu \left[(\partial_i f)^2 \right].$$

Noise Stability

Question

Among all Boolean functions, which one maximizes the noise stability?

Easy answer: among all Boolean functions the dictator $f(x) := x_1$ has the largest noise stability.

Not a very useful fact in social choice theory.

Define the maximal influence of a Boolean function by:

$$\inf = \max_{i \in [n]} \mathbb{E}_\mu \left[(\partial_i f)^2 \right].$$

Noise Stability

Question

Among all Boolean functions **with small maximal influence**, which one maximizes the noise stability?

Easy answer: among all Boolean functions the dictator $f(x) := x_1$ has the largest noise stability.

Not a very useful fact in social choice theory.

Define the maximal influence of a Boolean function by:

$$\mathrm{inf} = \max_{i \in [n]} \mathbb{E}_\mu \left[(\partial_i f)^2 \right].$$

Outline

1. Noise stability
2. **“Majority is Stablest”**
3. A stochastic approach to noise stability via a re-normalized Brownian motion
4. Interlude - from a toy example to the Courtade-Kumar conjecture
5. Back to Majority is Stablest

Majority is Stablest

Theorem (Mossel-O'Donnel-Oleszkiewicz 05')

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, *then,*

$$\mathrm{Stab}_\rho(f) \leq \frac{2}{\pi}\arcsin(\rho) + O\left(\frac{\log\log(\frac{1}{\kappa})}{\log(\frac{1}{\kappa})}\right).$$

Define the majority function $\mathrm{Maj}_n(x) = \mathrm{sgn}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i\right)$.

- Computation: $\inf(\mathrm{Maj}_n) \leq \frac{1}{\sqrt{n}}$.
- CLT: $\mathrm{Stab}_\rho(\mathrm{Maj}_n) \xrightarrow{n\to\infty} \frac{2}{\pi}\arcsin(\rho)$.

Majority is Stablest

Theorem (Mossel-O'Donnel-Oleszkiewicz 05')

Let f be a balanced Boolean function and suppose $\inf(f) \le \kappa$, *then,*

$$\mathrm{Stab}_\rho(f) \le \frac{2}{\pi}\arcsin(\rho) + O\left(\frac{\log\log(\frac{1}{\kappa})}{\log(\frac{1}{\kappa})}\right).$$

Define the majority function $\mathrm{Maj}_n(x) = \mathrm{sgn}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} x_i\right)$.

- Computation: $\inf(\mathrm{Maj}_n) \le \frac{1}{\sqrt{n}}$.
- CLT: $\mathrm{Stab}_\rho(\mathrm{Maj}_n) \xrightarrow{n\to\infty} \frac{2}{\pi}\arcsin(\rho)$.

Majority is Stablest - Proof Sketch

1. Prove analogous result in Gaussian space:
   - Noise semi-group is replaced by Ornstein-Uhlenbeck semi-group.
   - Majority is replaced by indicator of halfspace.

   Result follows from the isoperimetric inequality.
2. Prove invariance principle for low-influence polynomials:

$$|\mathbb{E}_\mu[p] - \mathbb{E}_\gamma[p]| \leq O(2^{\mathrm{degree(p)}} \cdot \inf(p)).$$

3. Replace f by $T_\varepsilon f$, essentially a log-degree polynomial. Turns out that $\varepsilon = \Theta\left(\frac{\log\log(\frac{1}{\kappa})}{\log(\frac{1}{\kappa})}\right)$ works.

Majority is Stablest - Proof Sketch

1. Prove analogous result in Gaussian space:
   - Noise semi-group is replaced by Ornstein-Uhlenbeck semi-group.
   - Majority is replaced by indicator of halfspace.

   Result follows from the isoperimetric inequality.
2. Prove invariance principle for low-influence polynomials:

$$|\mathbb{E}_\mu[p] - \mathbb{E}_\gamma[p]| \leq O(2^{\mathrm{degree(p)}} \cdot \inf(p)).$$

3. Replace f by $T_\varepsilon f$, essentially a log-degree polynomial. Turns out that $\varepsilon = \Theta\left(\frac{\log\log(\frac{1}{\kappa})}{\log(\frac{1}{\kappa})}\right)$ works.

Majority is Stablest - Proof Sketch

1. Prove analogous result in Gaussian space:
   - Noise semi-group is replaced by Ornstein-Uhlenbeck semi-group.
   - Majority is replaced by indicator of halfspace.

   Result follows from the isoperimetric inequality.
2. Prove invariance principle for low-influence polynomials:

$$|\mathbb{E}_{\mu}[p] - \mathbb{E}_{\gamma}[p]| \leq O(2^{\mathrm{degree(p)}} \cdot \mathrm{inf}(p)).$$

3. Replace f by $T_{\varepsilon} f$, essentially a log-degree polynomial.
   Turns out that $\varepsilon = \Theta\left(\frac{\log\log(\frac{1}{\kappa})}{\log(\frac{1}{\kappa})}\right)$ works.

Quantitative Majority is Stablest

We prove a quantitative version of the Majority theorem.

Theorem

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, then,

$$\mathrm{Stab}_\rho(f) \leq \frac{2}{\pi}\arcsin(\rho) + \mathrm{poly}(\kappa).$$

- The main idea is to realize $(\mathrm{Stab}_\rho(f))_{\rho \geq 0}$ as a measurement of some stochastic process.
- Allows using stochastic analysis to bypass the invariance principle.
- For the proof we introduce a new martingale embedding of μ as a re-normalized Brownian motion.

Quantitative Majority is Stablest

We prove a quantitative version of the Majority theorem.

Theorem

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, *then,*

$$\mathrm{Stab}_\rho(f) \leq \frac{2}{\pi} \arcsin(\rho) + \mathrm{poly}(\kappa).$$

- The main idea is to realize $(\mathrm{Stab}_\rho(f))_{\rho \geq 0}$ as a measurement of some stochastic process.
- Allows using stochastic analysis to bypass the invariance principle.
- For the proof we introduce a new martingale embedding of μ as a re-normalized Brownian motion.

Outline

1. Noise stability
2. "Majority is Stablest"
3. **A stochastic approach to noise stability via a re-normalized Brownian motion**
4. Interlude - from a toy example to the Courtade-Kumar conjecture
5. Back to Majority is Stablest

Noise Stability - an Observation

If $f : \mathcal{C}_n \to \mathbb{R}$, we extend it harmonically to $f : [-1,1]^n \to \mathbb{R}$. In particular, $T_\rho f(x) = f(\rho x)$. So, if $\mu_\rho = \mathrm{Uniform}(\{-\sqrt{\rho}, \sqrt{\rho}\}^n)$,

$$\mathrm{Stab}_\rho(f) = \mathbb{E}_\mu[f(x) \cdot f(\rho x)] = \mathbb{E}_\mu[f(\sqrt{\rho}x) \cdot f(\sqrt{\rho}x)] = \mathbb{E}_{\mu_\rho}[f^2].$$

Now, if ν is any measure on $[-1,1]$, an orthogonal decomposition of $L^2(\mu)$ can be used to show

$$\mathrm{Stab}_\nu(f) := \mathbb{E}_{\nu^{\otimes n}}[f^2] = \mathrm{Stab}_{\mathrm{Var}(\nu)}(f).$$

Noise Stability - an Observation

If $f : \mathcal{C}_n \to \mathbb{R}$, we extend it harmonically to $f : [-1,1]^n \to \mathbb{R}$. In particular, $T_\rho f(x) = f(\rho x)$. So, if $\mu_\rho = \mathrm{Uniform}(\{-\sqrt{\rho}, \sqrt{\rho}\}^n)$,

$$\mathrm{Stab}_\rho(f) = \mathbb{E}_\mu[f(x) \cdot f(\rho x)] = \mathbb{E}_\mu[f(\sqrt{\rho}x) \cdot f(\sqrt{\rho}x)] = \mathbb{E}_{\mu_\rho}[f^2].$$

Now, if ν is any measure on $[-1,1]$, an orthogonal decomposition of $L^2(\mu)$ can be used to show

$$\mathrm{Stab}_\nu(f) := \mathbb{E}_{\nu^{\otimes n}}[f^2] = \mathrm{Stab}_{\mathrm{Var}(\nu)}(f).$$

Noise Stability - an Observation

If $f : \mathcal{C}_n \to \mathbb{R}$, we extend it harmonically to $f : [-1,1]^n \to \mathbb{R}$. In particular, $T_\rho f(x) = f(\rho x)$. So, if $\mu_\rho = \mathrm{Uniform}(\{-\sqrt{\rho}, \sqrt{\rho}\}^n)$,

$$\mathrm{Stab}_\rho(f) = \mathbb{E}_\mu[f(x) \cdot f(\rho x)] = \mathbb{E}_\mu[f(\sqrt{\rho}x) \cdot f(\sqrt{\rho}x)] = \mathbb{E}_{\mu_\rho}[f^2].$$

Now, if ν is any measure on $[-1,1]$, an orthogonal decomposition of $L^2(\mu)$ can be used to show

$$\mathrm{Stab}_\nu(f) := \mathbb{E}_{\nu^{\otimes n}}[f^2] = \mathrm{Stab}_{\mathrm{Var}(\nu)}(f).$$

A Re-normalized Brownian Motion

Consider the following martingale,

$$dX(t) = \sigma_t dB(t) \text{ with } \sigma_t = \mathrm{diag}(\sqrt{(1 - X_i(t))(1 + X_i(t))}),$$

and define $\nu_t = \mathrm{Law}(X_1(t))$.

Lemma

$\mathrm{Var}(\nu_t) = 1 - e^{-t}$.

Proof.

$X_1(t)^2 = \text{martingale} + (1 - X_1(t)^2)dt$. So,
$\frac{d}{dt}\mathbb{E}\left[X_1(t)^2\right] = 1 - \mathbb{E}\left[X_1(t)^2\right]$. Now solve an ODE. □

If $Y(t) \sim X(\infty)|X(t)$ then $\mathbb{E}[Y(t)] = X(t)$ and $\mathrm{Cov}(Y(t)) = \sigma_t$.

A Re-normalized Brownian Motion

Consider the following martingale,

$$dX(t) = \sigma_t dB(t) \text{ with } \sigma_t = \mathrm{diag}(\sqrt{(1 - X_i(t))(1 + X_i(t))}),$$

and define $\nu_t = \mathrm{Law}(X_1(t))$.

Lemma

$\mathrm{Var}(\nu_t) = 1 - e^{-t}$.

Proof.

$X_1(t)^2 = \mathrm{martingale} + (1 - X_1(t)^2)dt$. So,
$\frac{d}{dt}\mathbb{E}\left[X_1(t)^2\right] = 1 - \mathbb{E}\left[X_1(t)^2\right]$. Now solve an ODE. □

If $Y(t) \sim X(\infty)|X(t)$ then $\mathbb{E}[Y(t)] = X(t)$ and $\mathrm{Cov}(Y(t)) = \sigma_t$.

A Re-normalized Brownian Motion

Consider the following martingale,

$$dX(t) = \sigma_t dB(t) \text{ with } \sigma_t = \mathrm{diag}(\sqrt{(1 - X_i(t))(1 + X_i(t))}),$$

and define $\nu_t = \mathrm{Law}(X_1(t))$.

Lemma

$\mathrm{Var}(\nu_t) = 1 - e^{-t}$.

Proof.

$X_1(t)^2 = \mathrm{martingale} + (1 - X_1(t)^2)dt$. So,
$\frac{d}{dt}\mathbb{E}\left[X_1(t)^2\right] = 1 - \mathbb{E}\left[X_1(t)^2\right]$. Now solve an ODE. □

If $Y(t) \sim X(\infty)|X(t)$ then $\mathbb{E}[Y(t)] = X(t)$ and $\mathrm{Cov}(Y(t)) = \sigma_t$.

General Proof Strategy

Let $f : \mathcal{C}_n \to \{-1, 1\}$ and define the martingale $N_t = f(X(t))$.

Observe,

$$\mathbb{E}\left[[N]_t\right] = \mathbb{E}[N_t^2] = \mathbb{E}_{\nu_t^{\otimes n}}\left[f^2\right]$$
$$= \mathrm{Stab}_{\mathrm{Var}(\nu_t)}(f) = \mathrm{Stab}_{1-e^{-t}}(f).$$

The proof goes by finding a "model process" M_t to represent $\mathrm{Stab}_\rho(\mathrm{Maj})$ and a coupling which affords an almost-sure path-wise inequality,

$$[N]_t \leq [M]_t.$$

General Proof Strategy

Let $f : \mathcal{C}_n \to \{-1, 1\}$ and define the martingale $N_t = f(X(t))$. Observe,

$$\mathbb{E}\left[[N]_t\right] = \mathbb{E}[N_t^2] = \mathbb{E}_{\nu_t^{\otimes n}}\left[f^2\right]$$
$$= \mathrm{Stab}_{\mathrm{Var}(\nu_t)}(f) = \mathrm{Stab}_{1-e^{-t}}(f).$$

The proof goes by finding a "model process" M_t to represent $\mathrm{Stab}_\rho(\mathrm{Maj})$ and a coupling which affords an almost-sure path-wise inequality,

$$[N]_t \leq [M]_t.$$

General Proof Strategy

Let $f : \mathcal{C}_n \to \{-1, 1\}$ and define the martingale $N_t = f(X(t))$. Observe,

$$\mathbb{E}\left[[N]_t\right] = \mathbb{E}[N_t^2] = \mathbb{E}_{\nu_t^{\otimes n}}\left[f^2\right]$$
$$= \mathrm{Stab}_{\mathrm{Var}(\nu_t)}(f) = \mathrm{Stab}_{1-e^{-t}}(f).$$

The proof goes by finding a "model process" M_t to represent $\mathrm{Stab}_\rho(\mathrm{Maj})$ and a coupling which affords an almost-sure path-wise inequality,

$$[N]_t \leq [M]_t.$$

Outline

1. Noise stability
2. “Majority is Stablest”
3. A stochastic approach to noise stability via a re-normalized Brownian motion
4. **Interlude - from a toy example to the Courtade-Kumar conjecture**
5. Back to Majority is Stablest

Interlude - Toy Example

Theorem

Among all Boolean functions, the dictator maximizes noise stability.

- Let $f : \mathcal{C}_n \to \{-1, 1\}$ and let $g : \mathcal{C}_n \to \{-1, 1\}$, $g(x) = x_1$.
- Define the martingales $N_t = f(X(t)), M_t = g(X(t)) = X_1(t)$.
- The theorem will follow, if we can find a coupling of N_t and M_t, such that $[N]_t \leq [M]_t$ almost surely.

Interlude - Toy Example

Theorem

Among all Boolean functions, the dictator maximizes noise stability.

- Let $f : \mathcal{C}_n \to \{-1, 1\}$ and let $g : \mathcal{C}_n \to \{-1, 1\}$, $g(x) = x_1$.
- Define the martingales $N_t = f(X(t)), M_t = g(X(t)) = X_1(t)$.
- The theorem will follow, if we can find a coupling of N_t and M_t, such that $[N]_t \leq [M]_t$ almost surely.

Interlude - Quadratic Variation

By Itô's formula

$$dM_t = \nabla g(X(t))\sigma_t dB_t = \sqrt{(1 - X_1(t))(1 + X_1(t))}dB_t.$$

Hence,

$$\frac{d}{dt}[M]_t = (1 - X_1(t))(1 + X_1(t)) = (1 - M_t^2).$$

In a similar way,

$$\frac{d}{dt}[N]_t = \|\nabla f(X(t))\sigma_t\|_2^2 = \sum_i (1 - X_i(t))(1 + X_i(t))\partial_i f(X(t)).$$

An application of Parseval's inequality gives,

$$\frac{d}{dt}[N]_t \leq (1 - f(X(t)))(1 + f(X(t))) = (1 - N_t^2).$$

Interlude - Quadratic Variation

By Itô's formula

$$dM_t = \nabla g(X(t))\sigma_t dB_t = \sqrt{(1 - X_1(t))(1 + X_1(t))} dB_t.$$

Hence,

$$\frac{d}{dt}[M]_t = (1 - X_1(t))(1 + X_1(t)) = (1 - M_t^2).$$

In a similar way,

$$\frac{d}{dt}[N]_t = \|\nabla f(X(t))\sigma_t\|_2^2 = \sum_i (1 - X_i(t))(1 + X_i(t))\partial_i f(X(t)).$$

An application of Parseval's inequality gives,

$$\frac{d}{dt}[N]_t \leq (1 - f(X(t)))(1 + f(X(t))) = (1 - N_t^2).$$

Interlude - a Coupling

By the Dambis-Dubins-Schwartz theorem, there exists a Brownian motion W_t, such that,

$$W_{[N]_t} = N_t \text{ and } W_{[M]_t} = M_t.$$

Reversing roles, for $\tau \geq 0$, write,

$$W_\tau = N_{T_1(\tau)} = M_{T_2(\tau)}.$$

So, keeping in mind that T_1 is the inverse function of $t \to [N]_t$

$$T_2'(\tau) = \frac{1}{1 - M^2_{T_2(\tau)}} = \frac{1}{1 - W^2_\tau} = \frac{1}{1 - N^2_{T_1(\tau)}} \leq T_1'(\tau).$$

Hence, almost surely, $T_2(\tau) \leq T_1(\tau) \implies [M]_\tau \geq [N]_\tau$.

Interlude - a Coupling

By the Dambis-Dubins-Schwartz theorem, there exists a Brownian motion W_t, such that,

$$W_{[N]_t} = N_t \text{ and } W_{[M]_t} = M_t.$$

Reversing roles, for $\tau \geq 0$, write,

$$W_\tau = N_{T_1(\tau)} = M_{T_2(\tau)}.$$

So, keeping in mind that T_1 is the inverse function of $t \to [N]_t$

$$T_2'(\tau) = \frac{1}{1 - M^2_{T_2(\tau)}} = \frac{1}{1 - W^2_\tau} = \frac{1}{1 - N^2_{T_1(\tau)}} \leq T_1'(\tau).$$

Hence, almost surely, $T_2(\tau) \leq T_1(\tau) \implies [M]_\tau \geq [N]_\tau$.

Interlude - Beyond the Toy Example

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be any convex function and fix $t \geq 0$.

$$\begin{aligned}\mathbb{E}\left[\varphi(M_t)\right] &= \mathbb{E}\left[\varphi(W_{[M]_t})\right] = \mathbb{E}\left[\mathbb{E}\left[\varphi(W_{[M]_t})|W_{[N]_t}\right]\right] \\ &\geq \mathbb{E}\left[\varphi\left(\mathbb{E}\left[W_{[M]_t}|W_{[N]_t}\right]\right)\right] = \mathbb{E}\left[\varphi\left(W_{[N]_t}\right)\right] \\ &= \mathbb{E}\left[\varphi(N_t)\right].\end{aligned}$$

Choose $\varphi(x) = x\log(x) + (1-x)\log(1-x)$, to get,

$$\mathbb{E}\left[\varphi(N_t)\right] = \mathbb{E}\left[\varphi\left(\mathbb{E}\left[f(X(\infty))|X(t)\right]\right)\right] = -\mathrm{Ent}(f(X(\infty))|X(t)).$$

Interlude - Beyond the Toy Example

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be any convex function and fix $t \geq 0$.

$$\begin{aligned}\mathbb{E}\left[\varphi(M_t)\right] &= \mathbb{E}\left[\varphi(W_{[M]_t})\right] = \mathbb{E}\left[\mathbb{E}\left[\varphi(W_{[M]_t})|W_{[N]_t}\right]\right] \\ &\geq \mathbb{E}\left[\varphi\left(\mathbb{E}\left[W_{[M]_t}|W_{[N]_t}\right]\right)\right] = \mathbb{E}\left[\varphi\left(W_{[N]_t}\right)\right] \\ &= \mathbb{E}\left[\varphi(N_t)\right].\end{aligned}$$

Choose $\varphi(x) = x\log(x) + (1-x)\log(1-x)$, to get,

$$\mathbb{E}\left[\varphi(N_t)\right] = \mathbb{E}\left[\varphi\left(\mathbb{E}\left[f(X(\infty))|X(t)\right]\right)\right] = -\mathrm{Ent}(f(X(\infty))|X(t)).$$

Interlude - Beyond the Toy Example

Define the mutual information $\mathrm{I}(X;Y) := \mathrm{Ent}(X) - \mathrm{Ent}(X|Y)$.

Theorem (Most informative $X(t)$ bit)

Among all Boolean functions, the dictator maximizes the mutual information,

$$\mathrm{I}(f(X(\infty)); X(t)).$$

Compare this with the 'most informative bit' conjecture of Courtade and Kumar.

Conjecture

Among all Boolean functions, the dictator maximizes the mutual information,

$$\mathrm{I}(f(X); Y),$$

where X and Y are ρ-correlated copies of uniform vectors on $\mathcal{C}_n$.

Interlude - Beyond the Toy Example

Define the mutual information $\mathrm{I}(X;Y) := \mathrm{Ent}(X) - \mathrm{Ent}(X|Y)$.

Theorem (Most informative $X(t)$ bit)

Among all Boolean functions, the dictator maximizes the mutual information,

$$\mathrm{I}(f(X(\infty)); X(t)).$$

Compare this with the 'most informative bit' conjecture of Courtade and Kumar.

Conjecture

Among all Boolean functions, the dictator maximizes the mutual information,

$$\mathrm{I}(f(X); Y),$$

where X and Y are ρ-correlated copies of uniform vectors on $\mathcal{C}_n$.

Most informative bit theorem

- Note that while $X(\infty)$ and $X(t)$ are correlated vectors, in general

$$(X(\infty), X(t)) \neq (X, Y),$$

for a ρ-correlated pair (X, Y).

- Thus while the theorem is in the spirit of the Courtade-Kumar conjecture, it proves it with respect to a different noise model.
- Interestingly, the analog of the relation $\mathbb{E}[\varphi(M_t)] \geq \mathbb{E}[\varphi(N_t)]$, for general convex φ is known to be under for the 'usual' noise semi-group.

Most informative bit theorem

- Note that while $X(\infty)$ and $X(t)$ are correlated vectors, in general

$$(X(\infty), X(t)) \neq (X, Y),$$

for a ρ-correlated pair (X, Y).

- Thus while the theorem is in the spirit of the Courtade-Kumar conjecture, it proves it with respect to a different noise model.
- Interestingly, the analog of the relation $\mathbb{E}[\varphi(M_t)] \geq \mathbb{E}[\varphi(N_t)]$, for general convex φ is known to be under for the 'usual' noise semi-group.

Outline

1. Noise stability
2. "Majority is Stablest"
3. A stochastic approach to noise stability via a re-normalized Brownian motion
4. Interlude - from a toy example to the Courtade-Kumar conjecture
5. **Back to Majority is Stablest**

Back to Majority is Stablest

Theorem

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, *then,*

$$\mathrm{Stab}_\rho(f) \leq \frac{2}{\pi}\arcsin(\rho) + \mathrm{poly}(\kappa).$$

Main ingredients:

- A martingale $N_t := f(X_t)$.
- A martingale M_t to represent noise stability of majority, or $\frac{2}{\pi}\arcsin(\rho)$.
- A differential *equality* for $[M]_t$.
- A differential *inequality* for $[N]_t$.

Constructing the Martingale M_t

There are infinitely many martingales M_t, which satisfy

$$\mathbb{E}\left[[M]_t\right] = \mathbb{E}\left[M_t^2\right] = \frac{2}{\pi}\arcsin(1 - e^{-t}) = \frac{2}{\pi}\arcsin(\rho).$$

We require one whose paths interact well with the paths of $f(X(t))$ when f has low influence.

One possibility is to take $M_t = f(\mathrm{Maj}_n)$. However, that depends on the dimension.

Instead, we take a limiting object of $\mathrm{Maj}_n(x) = \mathrm{sign}\left(\frac{1}{\sqrt{n}}\sum x_i\right)$ in Gaussian space.

Constructing the Martingale M_t

There are infinitely many martingales M_t, which satisfy

$$\mathbb{E}\left[[M]_t\right] = \mathbb{E}\left[M_t^2\right] = \frac{2}{\pi}\arcsin(1-e^{-t}) = \frac{2}{\pi}\arcsin(\rho).$$

We require one whose paths interact well with the paths of $f(X(t))$ when f has low influence.

One possibility is to take $M_t = f(\mathrm{Maj}_n)$. However, that depends on the dimension.

Instead, we take a limiting object of $\mathrm{Maj}_n(x) = \mathrm{sign}\left(\frac{1}{\sqrt{n}}\sum x_i\right)$ in Gaussian space.

Constructing the Martingale M_t

Let Φ stand for the Gaussian CDF and define the *Gaussian isoperimetric profile*:

$$I(x) := \Phi' \circ \Phi^{-1}(x).$$

Now, define M_t by, $dM_t = I(M_t)dB_t$.

It can be shown that $\mathbb{E}\left[[M]_t\right]$ encodes the limit of $\mathrm{Stab}_\rho(\mathrm{Maj}_n)$. Evidently, we have the differential equality

$$\frac{d}{dt}[M]_t = I(M_t)^2$$

Constructing the Martingale M_t

Let Φ stand for the Gaussian CDF and define the *Gaussian isoperimetric profile*:

$$I(x) := \Phi' \circ \Phi^{-1}(x).$$

Now, define M_t by, $dM_t = I(M_t)dB_t$.

It can be shown that $\mathbb{E}\left[[M]_t\right]$ encodes the limit of $\mathrm{Stab}_\rho(\mathrm{Maj}_n)$. Evidently, we have the differential equality

$$\frac{d}{dt}[M]_t = I(M_t)^2$$

Level 1 Inequality

Lemma

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, *then, if* $N_t = f(X(t))$,

$$\frac{d}{dt}[N]_t \lesssim (I(N_t) + \sqrt{\kappa})^2.$$

For the proof, we use the representation,

$$\frac{d}{dt}[N]_t = \|\nabla f(X(t))\sigma_t\|_2^2 = \int_{t \geq \alpha} t d\nu(t),$$

where ν is a marginal of $\frac{X(\infty)|X(t)-X(t)}{\sigma_t}$ in direction ∇f.

Level 1 Inequality

Lemma

Let f be a balanced Boolean function and suppose $\inf(f) \leq \kappa$, *then, if* $N_t = f(X(t))$,

$$\frac{d}{dt}[N]_t \lesssim (I(N_t) + \sqrt{\kappa})^2.$$

For the proof, we use the representation,

$$\frac{d}{dt}[N]_t = \|\nabla f(X(t))\sigma_t\|_2^2 = \int_{t \geq \alpha} t d\nu(t),$$

where ν is a marginal of $\frac{X(\infty)|X(t) - X(t)}{\sigma_t}$ in direction ∇f.

Thank You