Hypercontractivity and its Applications

Punyashloka Biswal
Advisors: James Lee and Paul Beame
University of Washington

February 10, 2010
Theory

- Problem: smoothing a function
- Log-Sobolev inequality
- Hypercontractivity

Applications

- Dictatorship testing with perfect completeness
- Integrality gap for Unique Games
Problem: smoothing a function

Given a function $f : \{-1, 1\}^n \rightarrow [-1, 1]$, we want a smoothed version $g : \{-1, 1\}^n \rightarrow [-1, 1]$.
Problem: smoothing a function

Given a function $f : \{-1, 1\}^n \to [-1, 1]$, we want a smoothed version $g : \{-1, 1\}^n \to [-1, 1]$

Values of $g$ should depend linearly on values of $f$.
Given a function $f : \{-1, 1\}^n \rightarrow [-1, 1]$, we want a smoothed version $g : \{-1, 1\}^n \rightarrow [-1, 1]$

1. Values of $g$ should depend linearly on values of $f$
2. $\mathbb{E}[g] = \mathbb{E}[f]$
Problem: smoothing a function

Given a function $f : \{-1, 1\}^n \rightarrow [-1, 1]$, we want a smoothed version $g : \{-1, 1\}^n \rightarrow [-1, 1]$

1. Values of $g$ should depend linearly on values of $f$
2. $\mathbb{E}[g] = \mathbb{E}[f]$
3. $g$ should vary less than $f$
Problem: smoothing a function

Given a function \( f : \{-1, 1\}^n \rightarrow [-1, 1] \), we want a smoothed version \( g : \{-1, 1\}^n \rightarrow [-1, 1] \)

1. Values of \( g \) should depend linearly on values of \( f \)
2. \( \mathbb{E}[g] = \mathbb{E}[f] \)
3. \( g \) should vary less than \( f \)
4. \( g(x) \) should depend on values of \( f \) near \( x \)
Global properties from local ones

\[ g: \{-1, 1\}^n \rightarrow [-1, 1] \]
Global properties from local ones

\[ g : \{-1, 1\}^n \rightarrow [-1, 1] \]

Suppose we are able to control local variation

Energy(g) = \( \frac{1}{2} \mathbf{E}_{x \sim y} [(g(x) - g(y))^2] \)
Global properties from local ones

\[ g : \{-1, 1\}^n \to [-1, 1] \]

Suppose we are able to control local variation

\[
\text{Energy}(g) = \frac{1}{2} \mathbb{E}_{x \sim y} [(g(x) - g(y))^2]
\]

What can we say about the variance?

\[
\text{Var}(g) = \mathbb{E}[(g - \mathbb{E}[g])^2]
\]
Global properties from local ones

\[ g : \{-1, 1\}^n \rightarrow [-1, 1] \]

Suppose we are able to control **local variation**

\[ \text{Energy}(g) = \frac{1}{2} E_{x \sim y} [(g(x) - g(y))^2] \]

What can we say about the **variance**?

\[ \text{Var}(g) = E[(g - E[g])^2] \]

**Poincaré**

\[ \text{Var}(g) \leq \frac{n}{2} \text{Energy}(g) \]
Global properties from local ones

\[ g : \{-1, 1\}^n \to [-1, 1] \]

Suppose we are able to control **local variation**

\[ \text{Energy}(g) = \frac{1}{2} \mathbb{E}_{x \sim y} [(g(x) - g(y))^2] \]

What can we say about the **variance**?

\[ \text{Var}(g) = \mathbb{E} \left[ (g - \mathbb{E}[g])^2 \right] \]

What can we say about the **entropy**?

\[ \text{Ent}(g) = \mathbb{E} \left[ g \ln g - \mathbb{E}[g] \ln \mathbb{E}[g] \right] \]

**Poincaré**

\[ \text{Var}(g) \leq \frac{n}{2} \text{Energy}(g) \]
Suppose we are able to control \textit{local variation}
\[
\text{Energy}(g) = \frac{1}{2} \mathbb{E}_{x \sim y} [(g(x) - g(y))^2]
\]

What can we say about the \textit{variance}?
\[
\text{Var}(g) = \mathbb{E}[(g - \mathbb{E}[g])^2]
\]

\textbf{Poincaré}
\[
\text{Var}(g) \leq \frac{n}{2} \text{Energy}(g)
\]

What can we say about the \textit{entropy}?
\[
\text{Ent}(g) = \mathbb{E}[g \ln g - \mathbb{E}[g] \ln \mathbb{E}[g]]
\]

\textbf{Log-Sobolev [Gross]}
\[
\text{Ent}(g^2) \leq n \text{Energy}(g)
\]
Some candidates for $g$

- $g(x) = f(x)^2$?
  - Not linear
Some candidates for $g$

- $g(x) = f(x)^2$?
  - Not linear

- $g(x) = \frac{1}{2} f(x)$?
  - Shrinks everything
  - Changes the expectation
Some candidates for $g$

- $g(x) = f(x)^2$?
  - Not linear

- $g(x) = \frac{1}{2} f(x)$?
  - Shrinks everything
  - Changes the expectation

- $g(x) = \mathbb{E}[f]$?
  - Very lossy

P. Biswal (UW)
Some candidates for $g$

- $g(x) = f(x)^2$ ?
  - Not linear

- $g(x) = \frac{1}{2} f(x)$ ?
  - Shrinks everything
  - Changes the expectation

- $g(x) = \mathbb{E}[f]$ ?
  - Very lossy

- $g(x) = \mathbb{E}_y \text{ near } x[f(y)]$ ?
  - Like a blur kernel in graphics
  - Hmm...
The Bonami-Gross-Beckner operator

For any $\rho \in [0, 1]$,

$$T_\rho[f](x_1, \ldots, x_n) = \mathbf{E}[f(y_1, \ldots, y_n)]$$

where

$$y_i = \begin{cases} 
  x_i & \text{with probability } \frac{1+\rho}{2} \\
  -x_i & \text{with probability } \frac{1-\rho}{2}
\end{cases}$$
The Bonami-Gross-Beckner operator

For any $\rho \in [0, 1]$,

$$T_\rho[f](x_1, \ldots, x_n) = \mathbb{E}[f(y_1, \ldots, y_n)]$$

where

$$y_i = \begin{cases} x_i & \text{with probability } \frac{1+\rho}{2} \\ -x_i & \text{with probability } \frac{1-\rho}{2} \end{cases}$$

e.g., $T_0[f](x) = \mathbb{E}[f]$
For any $\rho \in [0, 1]$,

$$T_\rho[f](x_1, \ldots, x_n) = \mathbf{E}[f(y_1, \ldots, y_n)]$$

where

$$y_i = \begin{cases} 
  x_i & \text{with probability } \frac{1+\rho}{2} \\
  -x_i & \text{with probability } \frac{1-\rho}{2}
\end{cases}$$

e.g., $T_0[f](x) = \mathbf{E}[f]$, $T_1[f](x) = f(x)$
For any $f: \{-1, 1\}^n \rightarrow [-1, 1]$

$$\|f\|_p = \mathbb{E}[|f|^p]^{1/p}$$  \hspace{1cm} 1 \leq p < \infty

$$\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p = \max f$$
\( p \)-norms

- For any \( f : \{-1, 1\}^n \rightarrow [-1, 1] \)

\[
\|f\|_p = \mathbb{E}[|f|^p]^{1/p} \quad \text{for} \quad 1 \leq p < \infty
\]

\[
\|f\|_\infty = \lim_{p \to \infty} \|f\|_p = \max f
\]

- \( \|f\|_1 \leq \|f\|_2 \leq \cdots \leq \|f\|_\infty \)

- Lower norms pay more attention to the average
  Higher norms pay more attention to spikes
For any $f: \{-1, 1\}^n \rightarrow [-1, 1]$:

$$||f||_p = E[|f|^p]^{1/p}$$

$$||f||_\infty = \lim_{p \rightarrow \infty} ||f||_p = \max f$$

- $||f||_1 \leq ||f||_2 \leq \cdots \leq ||f||_\infty$

- Lower norms pay more attention to the average
- Higher norms pay more attention to spikes

**Intuition**

Noise spreads out the mass of $f$ from its spikes, so we should be able to bound the higher norms of $T_\rho[f]$.
Hypercontractivity for \(\{-1, 1\}^n\) [Gross]

For any function \(f: \{-1, 1\}^n \rightarrow [-1, 1]\) and \(1 \leq p \leq q, 0 \leq \rho \leq 1\),

\[
\rho \leq \sqrt{\frac{p-1}{q-1}} \text{ implies } \|T_\rho[f]\|_q \leq \|f\|_p
\]
Hypercontractivity for \( \{-1, 1\}^n \) [Gross]

For any function \( f: \{-1, 1\}^n \to [-1, 1] \) and \( 1 \leq p \leq q \), \( 0 \leq \rho \leq 1 \),

\[
\rho \leq \sqrt{\frac{p - 1}{q - 1}} \quad \text{implies} \quad \| T_\rho[f] \|_q \leq \| f \|_p
\]

Application

- For any unbiased boolean function \( f(x_1, \ldots, x_n) \) there is an index \( x_i \)
  such that \( f(\ldots, x_i, \ldots) \neq f(\ldots, -x_i, \ldots) \) at least \( \Omega(\frac{\log n}{n}) \) \cdot \text{Var}(f) \)
  of the time. [Kahn, Kalai, Linial]
More generally, if $X$ has notions of probability and distance, we can define

$$\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \sim y} [(f(x) - f(y))^2]$$

If we can prove a Log-Sobolev inequality,

$$\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$$

Then we can define a smoothing operator $T_\rho$ for $0 \leq \rho \leq 1$ such that for any $f : X \to [-1, 1]$ and $1 \leq p \leq q$,

$$\| T_\rho f \|_q \leq \| f \|_p$$

e.g., $\mathbb{R}$ with Gaussian measure:

$$\text{Energy}(f) = \frac{1}{2} \mathbb{E}[f'(x)^2]$$

$$\text{Ent}(f^2) \leq \text{Energy}(f)$$

$$T_\rho[f](x) = \mathbb{E}_{y \sim \mathcal{N}(0, 1)} [f(\rho x + (1 - \rho^2)^{1/2} y)]$$

$$\| T_\rho[f] \|_q \leq \| f \|_p$$

e.g., $\mathbb{R}$ with Gaussian measure:

$$\text{Energy}(f) = \frac{1}{2} \mathbb{E}[f'(x)^2]$$

$$\text{Ent}(f^2) \leq \text{Energy}(f)$$

$$T_\rho[f](x) = \mathbb{E}_{y \sim \mathcal{N}(0, 1)} [f(\rho x + (1 - \rho^2)^{1/2} y)]$$

$$\| T_\rho[f] \|_q \leq \| f \|_p$$

e.g., $\mathbb{R}$ with Gaussian measure:
More generally, if $X$ has notions of probability and distance.

We can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \text{ near } y} [(f(x) - f(y))^2]$.

If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$, then we can define a smoothing operator $T_\rho$ for $0 \leq \rho \leq 1$, such that for any $f : X \to [-1, 1]$ and $1 \leq p \leq q$, we can give an explicit $\rho = \rho(p,q,C)$ such that $\|T_\rho f\|_q \leq \|f\|_p$ when $\rho < \sqrt{(q - 1)/(p - 1)}$. For example, in $\mathbb{R}$ with Gaussian measure, $\text{Energy}(f) = \frac{1}{2} \mathbb{E} \left[ f'(x)^2 \right]$ and $\text{Ent}(f^2) \leq \text{Energy}(f)$. Then $T_\rho [f](x) = \mathbb{E}_{y \sim N(0,1)} [f(\rho x + (1 - \rho)^{1/2} y)]$.
Hypercontractivity in other spaces

- More generally, if $X$ has notions of probability and distance
- We can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \text{ near } y}[(f(x) - f(y))^2]$
- If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$
More generally, if $X$ has notions of probability and distance

We can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \text{ near } y}[(f(x) - f(y))^2]$

If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$

Then we can define a smoothing operator $T_\rho$ for $0 \leq \rho \leq 1$
More generally, if $X$ has notions of probability and distance
We can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \text{ near } y} [(f(x) - f(y))^2]$
If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$
Then we can define a smoothing operator $T_{\rho}$ for $0 \leq \rho \leq 1$
Such that for any $f : X \to [-1, 1]$ and $1 \leq p \leq q$, we can give an explicit $\rho = \rho(p, q, C)$ such that $\|T_{\rho} f\|_q \leq \|f\|_q$
More generally, if $X$ has notions of probability and distance, we can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_{x \text{ near } y} [(f(x) - f(y))^2]$

If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$, then we can define a smoothing operator $T_\rho$ for $0 \leq \rho \leq 1$

Such that for any $f : X \to [-1, 1]$ and $1 \leq p \leq q$, we can give an explicit $\rho = \rho(p, q, C)$ such that $\|T_\rho f\|_q \leq \|f\|_q$

For example, $\mathbb{R}$ with Gaussian measure:

$$\text{Energy}(f) = \frac{1}{2} \mathbb{E}[(f')^2]$$

$$\text{Ent}(f^2) \leq \text{Energy}(f)$$
More generally, if $X$ has notions of probability and distance
We can define $\text{Energy}(f) = \frac{1}{2} \mathbb{E}_x \text{ near } y [(f(x) - f(y))^2]$
If we can prove a Log-Sobolev inequality $\text{Ent}(f^2) \leq C \cdot \text{Energy}(f)$
Then we can define a smoothing operator $T_\rho$ for $0 \leq \rho \leq 1$
Such that for any $f : X \to [-1, 1]$ and $1 \leq p \leq q$, we can give an explicit $\rho = \rho(p, q, C)$ such that $\| T_\rho f \|_q \leq \| f \|_q$
e.g., $\mathbb{R}$ with Gaussian measure:

$$\text{Energy}(f) = \frac{1}{2} \mathbb{E}[(f')^2]$$
$$\text{Ent}(f^2) \leq \text{Energy}(f)$$

$$T_\rho[f](x) = \mathbb{E}_{y \sim \mathcal{N}(0,1)} [f(\rho x + (1 - \rho^2)^{1/2} y)]$$
$$\| T_\rho[f] \|_q \leq \| f \|_p \text{ when } \rho < \sqrt{(q - 1)/(p - 1)}$$
Applications of HC in other spaces

- **Gaussian**
  
  Strong isoperimetric inequality for Gaussian space, leading to fast algorithms for graph partitioning [Sherman]
Applications of HC in other spaces

- **Gaussian**
  Strong isoperimetric inequality for Gaussian space, leading to fast algorithms for graph partitioning [Sherman]

- **Schreier graphs**
  Every monotone function from $\{-1, 1\}^n$ is $(\frac{1}{2} - \Omega(\frac{\log n}{n}))$-close to one of $\{0, 1, x_1, \ldots, x_n, \text{Maj}(x)\}$. [O’Donnell-Wimmer]
Dictatorship testing

Given a function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$,
Dictatorship testing

Given a function $f : \{-1, 1\}^n \to \{-1, 1\}$, query it at 3 points $x, y, z \in \{-1, 1\}^n$.
Given a function $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, query it at 3 points $x, y, z \in \{-1, 1\}^n$ chosen non-adaptively and then...
Dictatorship testing

Given a function \( f : \{-1, 1\}^n \rightarrow \{-1, 1\} \), query it at 3 points \( x, y, z \in \{-1, 1\}^n \) chosen non-adaptively and then

- if \( f \) is a dictator
  - i.e., it only depends on only one of its input coordinates

\[ \text{if } f \text{ is a dictator} \]

\[ \quad \text{i.e., it only depends on only one of its input coordinates} \]
Given a function $f : \{-1,1\}^n \rightarrow \{-1,1\}$, query it at 3 points $x, y, z \in \{-1,1\}^n$ chosen non-adaptively and then

1. if $f$ is a dictator
   - i.e., it only depends on only one of its input coordinates
   - you must accept with probability $\geq c$
Dictatorship testing

Given a function $f : \{-1, 1\}^n \to \{-1, 1\}$, query it at 3 points $x, y, z \in \{-1, 1\}^n$ chosen non-adaptively and then

1. if $f$ is a dictator
   - i.e., it only depends on only one of its input coordinates
   - you must accept with probability $\geq c$

2. $f$ is quasirandom
   - i.e., it is far from every small junta
Dictatorship testing

Given a function $f : \{-1, 1\}^n \to \{-1, 1\}$, query it at 3 points $x, y, z \in \{-1, 1\}^n$ chosen non-adaptively and then

1. if $f$ is a dictator
   - i.e., it only depends on only one of its input coordinates
   - you must accept with probability $\geq c$

2. $f$ is quasirandom
   - i.e., it is far from every small junta
   - you must accept with probability $\leq s$
Dictatorship testing

Given a function $f : \{-1, 1\}^n \to \{-1, 1\}$, query it at 3 points $x, y, z \in \{-1, 1\}^n$ chosen non-adaptively and then

1. If $f$ is a dictator
   - i.e., it only depends on only one of its input coordinates
   - you must accept with probability $\geq c$

2. $f$ is quasirandom
   - i.e., it is far from every small junta
   - you must accept with probability $\leq s$
Every function \( f : \{-1, 1\}^n \to \mathbb{R} \) can be written as a \textbf{multilinear polynomial}, e.g. \( f(x) = \frac{3}{4}x_1 - \frac{1}{2}x_3x_4 \).
Every function \( f : \{-1, 1\}^n \rightarrow \mathbb{R} \) can be written as a multilinear polynomial, e.g. \( f(x) = \frac{3}{4}x_1 - \frac{1}{2}x_3x_4 \)

For any set \( S \subseteq [n] \), the coefficient of \( \prod_{i \in S} x_i \) is denoted \( \hat{f}(S) \)
Quasirandomness and Fourier analysis

- Every function $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ can be written as a multilinear polynomial, e.g. $f(x) = \frac{3}{4}x_1 - \frac{1}{2}x_3x_4$

- For any set $S \subseteq [n]$, the coefficient of $\prod_{i \in S} x_i$ is denoted $\hat{f}(S)$

- If $g = T_\rho[f]$, then $\hat{g}(S) = \rho^{|S|}\hat{f}(S)$
Every function $f : \{-1, 1\}^n \to \mathbb{R}$ can be written as a multilinear polynomial, e.g. $f(x) = \frac{3}{4}x_1 - \frac{1}{2}x_3x_4$

For any set $S \subseteq [n]$, the coefficient of $\prod_{i \in S} x_i$ is denoted $\hat{f}(S)$

If $g = T_\rho[f]$, then $\hat{g}(S) = \rho^{|S|}\hat{f}(S)$

$f$ is said to be $(\epsilon, \delta)$-quasirandom if $|\hat{f}(S)| \leq \epsilon$ whenever $|S| \geq 1/\delta$.
Every function $f : \{-1, 1\}^n \to \mathbb{R}$ can be written as a multilinear polynomial, e.g. $f(x) = \frac{3}{4}x_1 - \frac{1}{2}x_3x_4$

For any set $S \subseteq [n]$, the coefficient of $\prod_{i \in S} x_i$ is denoted $\hat{f}(S)$

If $g = T_\rho[f]$, then $\hat{g}(S) = \rho^{|S|} \hat{f}(S)$

$f$ is said to be $(\epsilon, \delta)$-quasirandom if $|\hat{f}(S)| \leq \epsilon$ whenever $|S| \geq 1/\delta$.

Testing with perfect completeness [O'Donnell-Wu]

For every $0 < \delta < 1/8$, there is a 3-query nonadaptive test that accepts any dictator with probability 1 but accepts any $(\delta, \frac{\delta}{\log(1/\delta)})$-quasirandom function with probability $\leq \frac{5}{8} + O(\sqrt{\delta})$. 
The test

- For each $1 \leq i \leq n$, sample $(x_i, y_i, y_i)$ as follows:

| $x_1$ | $\ldots$ | $x_i$ | $\ldots$ | $x_n$ |
|------|---------|------|---------|------|
| $y_1$ | $\ldots$ | $y_i$ | $\ldots$ | $y_n$ |
| $z_1$ | $\ldots$ | $z_i$ | $\ldots$ | $z_n$ |

With probability $1 - \delta$, pick $x_i, y_i, z_i$ uniformly from the subset that satisfies $x_i y_i z_i = -1$.

With probability $\delta$, pick $x_i = y_i = z_i$ uniformly between $\{-1, 1\}$.

Query $f(x), f(y), f(z)$.

If exactly two of the values are $-1$, then reject. Otherwise accept.
The test

- For each \(1 \leq i \leq n\), sample \((x_i, y_i, z_i)\) as follows:

\[
\begin{array}{ccc}
  x_1 & \ldots & x_i & \ldots & x_n \\
  y_1 & \ldots & y_i & \ldots & y_n \\
  z_1 & \ldots & z_i & \ldots & z_n \\
\end{array}
\]

- with probability \(1 - \delta\),
pick \(x_i, y_i, z_i\) uniformly from the subset that satisfies \(x_i y_i z_i = -1\)
The test

- For each $1 \leq i \leq n$, sample $(x_i, y_i, y_i)$ as follows:

| $x_1 \ldots x_i \ldots x_n$ | $x_i \ldots x_n$ |
|------------------------------|------------------|
| $y_1 \ldots y_i \ldots y_n$ | $y_i \ldots y_n$ |
| $z_1 \ldots z_i \ldots z_n$ | $z_i \ldots z_n$ |

- with probability $1 - \delta$, pick $x_i, y_i, z_i$ uniformly from the subset that satisfies $x_i y_i z_i = -1$
- with probability $\delta$, pick $x_i = y_i = z_i$ uniformly between $\{-1, 1\}$

Query $f(x), f(y), f(z)$. If exactly two of the values are $-1$, then reject. Otherwise accept.
For each $1 \leq i \leq n$, sample $(x_i, y_i, y_i)$ as follows:

- with probability $1 - \delta$, pick $x_i, y_i, z_i$ uniformly from the subset that satisfies $x_i y_i z_i = -1$
- with probability $\delta$, pick $x_i = y_i = z_i$ uniformly between $\{-1, 1\}$

Query $f(x), f(y), f(z)$.

If exactly two of the values are $-1$, then reject. Otherwise accept.
Analysis: completeness

- \((x_i, y_i, z_i) \in \{(−1, 1, 1), (1, −1, 1), (1, 1, −1)\} \cup \{(-1, -1, -1), (1, 1, 1)\}\)

- \(x_i y_i z_i = -1\)
- \(x_i = y_i = z_i\)

- Zero, one, or three occurrences of \(-1\)!
- So if \(f(x) = x_i\), our test would pass it. \((c = 1)\)
Analysis: soundness

\[
\begin{array}{cccccccc}
  a & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
  b & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
  c & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
\end{array}
\]

\[
\begin{array}{cccccccc}
  \text{NTW}(a, b, c) & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
\end{array}
\]
Analysis: soundness

\begin{align*}
a &\quad -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
b &\quad -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
c &\quad -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
\text{NTW}(a, b, c) &\quad 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
\end{align*}

\textbf{NTW}(a, b, c) = \frac{5}{8} + \frac{1}{8}(a + b + c) + \frac{1}{8}(ab + bc + ca) - \frac{3}{8}abc

Proceed using linearity of expectation and Plancherel's theorem:

\[E[f^2] = \sum S_f(S) ^ 2\]

Need to bound \(-\frac{3}{8}E[f(x)f(y)f(z)]\)
Analysis: soundness

\begin{align*}
a & \quad -1 \quad -1 \quad -1 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1 \\
b & \quad -1 \quad -1 \quad 1 \quad 1 \quad -1 \quad -1 \quad 1 \quad 1 \\
c & \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad -1 \quad 1 \quad -1 \\
\text{NTW}(a, b, c) & \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1
\end{align*}

\textbullet \quad \text{NTW}(a, b, c) = \frac{5}{8} + \frac{1}{8}(a + b + c) + \frac{1}{8}(ab + bc + ca) - \frac{3}{8}abc

\text{Pr}[\text{accept } f] = \mathbb{E}[\text{NTW}(f(x), f(y), f(z))]
\quad = \frac{5}{8} + \frac{3}{8} \mathbb{E}[f(x)] + \frac{3}{8} \mathbb{E}[f(x)f(y)] - \frac{3}{8} \mathbb{E}[f(x)f(y)f(z)]
Analysis: soundness

\[
\begin{array}{cccccccccc}
  a & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
  b & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\
  c & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\
\end{array}
\]

\[
\text{NTW}(a, b, c) = 1 0 0 1 0 1 1 1 1 1
\]

\[\text{NTW}(a, b, c) = \frac{5}{8} + \frac{1}{8}(a + b + c) + \frac{1}{8}(ab + bc + ca) - \frac{3}{8}abc\]

\[
\Pr[\text{accept } f] = \mathbb{E}[\text{NTW}(f(x), f(y), f(z))]
\]
\[= \frac{5}{8} + \frac{3}{8} \mathbb{E}[f(x)] + \frac{3}{8} \mathbb{E}[f(x)f(y)] - \frac{3}{8} \mathbb{E}[f(x)f(y)f(z)]\]

Proceed using

- linearity of expectation
- Plancherel’s theorem: \(\mathbb{E}[f^2] = \sum_s \hat{f}(S)^2\)
- elementary algebra
Analysis: soundness

\[
\begin{array}{cccccccccc}
  a & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \\
  b & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\
  c & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\
\end{array}
\]

\[
\text{NTW}(a, b, c) = 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 1
\]

- \[\text{NTW}(a, b, c) = \frac{5}{8} + \frac{1}{8}(a + b + c) + \frac{1}{8}(ab + bc + ca) - \frac{3}{8}abc\]

\[
\Pr[\text{accept } f] = \mathbb{E}[\text{NTW}(f(x), f(y), f(z))]
\]

\[
= \frac{5}{8} + \frac{3}{8} \mathbb{E}[f(x)] + \frac{3}{8} \mathbb{E}[f(x)f(y)] - \frac{3}{8} \mathbb{E}[f(x)f(y)f(z)]
\]

- Proceed using
  - linearity of expectation
  - Plancherel's theorem: \( \mathbb{E}[f^2] = \sum_S \hat{f}(S)^2 \)
  - elementary algebra

- Need to bound \(-\frac{3}{8} \mathbb{E}[f(x)f(y)f(z)]\)
The cubic term

- The contribution due to each $A \subseteq [n]$ can be bounded by

$$4(1 - \delta)^{|A|}(\|\hat{f}(A)\|^3 + \|T_{\sqrt{\delta}} g_A\|^3)$$

where $g_A : \{-1, 1\}^{[n] \setminus A} \rightarrow \mathbb{R}$ is given by

$$\hat{g}_A(X) = \begin{cases} 0 & X = \emptyset \\ \hat{f}(A \cup X) & \text{otherwise} \end{cases}$$
The cubic term

- The contribution due to each $A \subseteq [n]$ can be bounded by

$$4(1 - \delta)^{|A|}(|\hat{f}(A)|^3 + \| T_{\sqrt{\delta}} g_A \|_3^3)$$

where $g_A : \{-1, 1\}^{[n]\setminus A} \rightarrow \mathbb{R}$ is given by

$$\hat{g}_A(X) = \begin{cases} 
0 & X = \emptyset \\
\hat{f}(A \cup X) & \text{otherwise}
\end{cases}$$

- $\sum_{|A| \leq \frac{1}{\delta} \log \frac{1}{\delta}} (1 - \delta)^{|A|} |\hat{f}(A)|^3 \leq \sqrt{\delta}$
The cubic term

- The contribution due to each $A \subseteq [n]$ can be bounded by

\[4(1 - \delta)^{|A|}(|\hat{f}(A)|^3 + \| T_{\sqrt{\delta}} g_A \|^3_3)\]

where $g_A : \{-1, 1\}^{[n]\setminus A} \rightarrow \mathbb{R}$ is given by

\[
\hat{g}_A(X) = \begin{cases} 
0 & X = \emptyset \\
\hat{f}(A \cup X) & \text{otherwise}
\end{cases}
\]

- \[\sum_{|A| \leq \frac{1}{\delta} \log \frac{1}{\delta}} (1 - \delta)^{|A|}|\hat{f}(A)|^3 \leq \sqrt{\delta}\]

- \[\sum_{|A| > \frac{1}{\delta} \log \frac{1}{\delta}} (1 - \delta)^{|A|}|\hat{f}(A)|^3 \leq (1 - \delta)^{1/\delta} \leq O(\delta)\]
The cubic term

- Goal: bound $\sum_A (1 - \delta)^{|A|} \| T^{\sqrt{\delta}} g_A \|^3_3$

- Using a slight variation of the hypercontractive inequality, we have for $\lambda = 1/\log_2(1/\delta) < 1/3$ that

$$\| T^{\sqrt{\delta}} g_A \|^3_3 \leq \| T^{\sqrt{\delta}} g_A \|^{3-3\lambda}_2 \| g_A \|^{3\lambda}_2$$
The cubic term

- **Goal:** bound $\sum_A (1 - \delta)^{|A|} \| T_{\sqrt{\delta}} g_A \|_3^3$

- Using a slight variation of the hypercontractive inequality, we have for $\lambda = 1 / \log_2(1/\delta) < 1/3$ that

  $$\| T_{\sqrt{\delta}} g_A \|_3^3 \leq \| T_{\sqrt{\delta}} g_A \|_2^{3-3\lambda} \| g_A \|_2^{3\lambda}$$

- **Plancherel:** $\| g_A \|_2^{3\lambda} \leq 1$
The cubic term

- **Goal:** bound $\sum_A (1 - \delta)^{|A|} \| T\sqrt{\delta} g_A \|^3_3$

- **Using a slight variation of the hypercontractive inequality,** we have for $\lambda = 1 / \log_2(1/\delta) < 1/3$ that

$$\| T\sqrt{\delta} g_A \|^3_3 \leq \| T\sqrt{\delta} g_A \|^{3-3\lambda}_2 \| g_A \|^{3\lambda}_2$$

- **Plancherel:** $\| g_A \|^{3\lambda}_2 \leq 1$

- **Algebraic manipulation:**

$$\| T\sqrt{\delta} g_A \|^{3-3\lambda}_2 \leq O(\sqrt{\delta}) \sum_{\emptyset \neq B \subseteq A} \delta^{|B|} \hat{f}(A \cup B)^2$$
The cubic term

\[
\sum_{A} (1 - \delta)^{|A|} \|T \sqrt{\delta} g_{A}\|_3^3 \leq O(\sqrt{\delta}) \sum_{A} (1 - \delta)^{|A|} \delta^{||B||} \hat{f}(A \cup B)^2
\]
The cubic term

\[ \sum_{A} (1 - \delta)^{|A|} \|T^{\sqrt{\delta}} g_A\|_3^3 \leq O(\sqrt{\delta}) \sum_{A} (1 - \delta)^{|A|} \delta^{|B|} \hat{f}(A \cup B)^2 \]

- Contribution due to each \( A \cup B \) is
  \[ \sum (1 - \delta)^{|A|} \delta^{|B|} = 1 \] (Binomial sum)
The cubic term

\[
\sum_A (1 - \delta)^{|A|} \|T_{\sqrt{\delta} g_A}\|_3^3 \leq O(\sqrt{\delta}) \sum_A (1 - \delta)^{|A|} \delta^{|B|} \hat{f}(A \cup B)^2
\]

- Contribution due to each \( A \cup B \) is
  \[
  \sum (1 - \delta)^{|A|} \delta^{|B|} = 1 \text{ (Binomial sum)}
  \]
- Total of all \( \hat{f}(A \cup B)^2 \) contributions is \( \leq 1 \) (Plancherel)
Thank You!
**Label Cover**: Given a set $V$ of variables over a domain $L$ and weighted constraints on each pair, assign values to maximize the fraction of satisfied constraints.

**Unique Label Cover**: As above, but every constraint is a bijection: a constraint on the pair $u, v \in V$ takes the form of a permutation $\pi: L \to L$.

Example: Linear equations of the form $x_u - x_v = r \pmod{p}$.
Label Cover: Given a set $V$ of variables over a domain $L$ and weighted constraints on each pair, assign values to maximize the fraction of satisfied constraints.

Unique Label Cover: As above, but every constraint is a bijection: a constraint on the pair $u, v \in V$ takes the form of a permutation $\pi : L \to L$. 
**Label Cover**: Given a set $V$ of variables over a domain $L$ and weighted constraints on each pair, assign values to maximize the fraction of satisfied constraints.

**Unique Label Cover**: As above, but every constraint is a bijection: a constraint on the pair $u, v \in V$ takes the form of a permutation $\pi : L \to L$.

**Example**: Linear equations of the form $x_u - x_v = r \pmod{p}$
**Label Cover:** Given a set $V$ of variables over a domain $L$ and weighted constraints on each pair, assign values to maximize the fraction of satisfied constraints.

**Unique Label Cover:** As above, but every constraint is a bijection: a constraint on the pair $u, v \in V$ takes the form of a permutation $\pi : L \to L$.

**Example:** Linear equations of the form $x_u - x_v = r \pmod{p}$

**Hardness**
- Easy when there exists a perfect solution
- But if there exists a solution satisfying 99% of the constraints, we don’t even know how to find a 1% satisfying solution
**Unique Games**

- **Label Cover**: Given a set $V$ of variables over a domain $L$ and weighted constraints on each pair, assign values to maximize the fraction of satisfied constraints.

- **Unique Label Cover**: As above, but every constraint is a bijection: a constraint on the pair $u, v \in V$ takes the form of a permutation $\pi : L \rightarrow L$.

- **Example**: Linear equations of the form $x_u - x_v = r \pmod{p}$

- **Hardness**
  - Easy when there exists a perfect solution
  - But if there exists a solution satisfying 99% of the constraints, we don’t even know how to find a 1% satisfying solution

- **Unique Games Conjecture**
An SDP for ULC

maximize $E_{e \{u,v\}} \sum_{i \in L} \langle u_i, v_{\pi_e(i)} \rangle$

subject to $\langle u_i, v_j \rangle \geq 0$  $\forall u, v \in V, \forall i, j \in L$

$\sum_{i \in L} \langle v_i, v_i \rangle = 1$  $\forall v \in V$

$\langle \sum_{i \in L} u_i, \sum_{j \in L} v_j \rangle = 1$  $\forall u, v \in L$

$\langle v_i, v_j \rangle = 0$  $\forall v \in V, \forall i \neq j \in L$
An SDP for ULC

maximize $E_{e\{u,v\}} \sum_{i \in L} \langle u_i, v_{\pi_e(i)} \rangle$

subject to $\langle u_i, v_j \rangle \geq 0 \quad \forall u, v \in V, \forall i, j \in L$

$\sum_{i \in L} \langle v_i, v_i \rangle = 1 \quad \forall v \in V$

$\langle \sum_{i \in L} u_i, \sum_{j \in L} v_j \rangle = 1 \quad \forall u, v \in L$

$\langle v_i, v_j \rangle = 0 \quad \forall v \in V, \forall i \neq j \in L$

Integrality gap

For domain size $2^k$ and any value $0 < \eta < \frac{1}{2}$, there is a ULC instance whose integer optimum is $\leq 2^{-k\eta}$ but whose SDP admits solutions of value $\geq 1 - \eta$. 

[Khot-Vishnoi]
Gap instance

- Take $V = \text{all functions } f : \{-1, 1\}^n \to \{-1, 1\}$
- Take $L = \text{all monomials } \prod_{i \in S} x_i$
- Hard constraints:
  - If $f = g \chi$ for some monomial $\chi$, then $\text{Label}(f) = \text{Label}(g) \chi$ must hold
- Fix one $f$ from each group tied by hard constraints
Gap instance

- Take $V = \text{all functions } f: \{-1, 1\}^n \to \{-1, 1\}$
- Take $L = \text{all monomials } \prod_{i \in S} x_i$
- Hard constraints:
  - If $f = g \chi$ for some monomial $\chi$, then $\text{Label}(f) = \text{Label}(g) \chi$ must hold
- Fix one $f$ from each group tied by hard constraints
- Soft constraints
  - Weight $= \Pr_{h, h'}[\{f, g\} = \{h, h'\}]$ where $h, h'$ are $(1 - 2\eta)$-correlated
  - Permutation: $\frac{\text{Label}(f \chi)}{\chi} = \frac{\text{Label}(g \psi)}{\psi}$
Soundness

- Objective value is precisely $\Pr[\text{Label}(h) = \text{Label}(h')]$
- Let $\phi: V \rightarrow \{0, 1\}$ indicate the set that received some label $\chi$
Soundness

- Objective value is precisely $\Pr[\text{Label}(h) = \text{Label}(h')]$
- Let $\phi: V \rightarrow \{0, 1\}$ indicate the set that received some label $\chi$

$$\Pr[\text{Label}(h) = \text{Label}(h') = \chi] = \mathbb{E}[\phi(h)\phi(h')] = \mathbb{E}[h, T_{1-2\eta} h] = \|T\sqrt{1-2\eta} h\|_2^2$$
Soundness

- Objective value is precisely $\Pr[\text{Label}(h) = \text{Label}(h')]$
- Let $\phi : V \to \{0, 1\}$ indicate the set that received some label $\chi$

\[
\Pr[\text{Label}(h) = \text{Label}(h') = \chi] = \mathbb{E}[\phi(h)\phi(h')] = \mathbb{E}[h, T_{1-2\eta} h] = \|T_{\sqrt{1-2\eta}} h\|_2^2
\]

- By hypercontractivity, $\leq \|h\|_2^2(1-\eta) = 1/2^{1+\eta}$