On Binary Codes and Non-Interactive Simulation

Lei Yu

Joint Work with Vincent Tan
Department of ECE
National University of Singapore

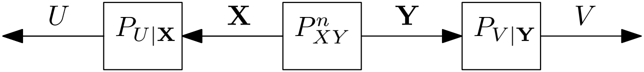
Non-Interactive Simulation Problem

- Given P_{XY}, let $(\mathbf{X}, \mathbf{Y}) \sim P_{XY}^n$ be correlated memoryless sources
  - i.e., $(\mathbf{X}, \mathbf{Y})$ are n i.i.d. copies of $(X, Y) \sim P_{XY}$

Non-Interactive Simulation Problem

- Given P_{XY}, let $(\mathbf{X}, \mathbf{Y}) \sim P_{XY}^n$ be correlated memoryless sources
  - i.e., $(\mathbf{X}, \mathbf{Y})$ are n i.i.d. copies of $(X, Y) \sim P_{XY}$
- Assume (U, V) on $\mathcal{U} \times \mathcal{V}$ are two random variables such that $U - \mathbf{X} - \mathbf{Y} - V$ forms a Markov chain, i.e.,

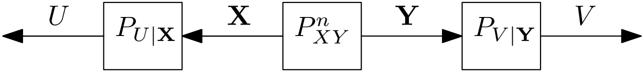
$$P_{U\mathbf{X}\mathbf{Y}V} = P_{U|\mathbf{X}} P_{XY}^n P_{V|\mathbf{Y}}$$



Non-Interactive Simulation Problem

- Given P_{XY}, let $(\mathbf{X}, \mathbf{Y}) \sim P_{XY}^n$ be correlated memoryless sources
  - i.e., $(\mathbf{X}, \mathbf{Y})$ are n i.i.d. copies of $(X, Y) \sim P_{XY}$
- Assume (U, V) on $\mathcal{U} \times \mathcal{V}$ are two random variables such that $U - \mathbf{X} - \mathbf{Y} - V$ forms a Markov chain, i.e.,

$$P_{U\mathbf{X}\mathbf{Y}V} = P_{U|\mathbf{X}} P_{XY}^n P_{V|\mathbf{Y}}$$



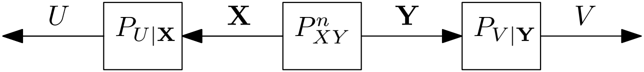
- A natural question: What are the possible joint distributions P_{UV} of (U, V)?

$$\mathcal{Q}(\mathcal{U} \times \mathcal{V} | P_{XY}) := \{P_{UV} \in \mathcal{P}(\mathcal{U} \times \mathcal{V}) : U - \mathbf{X} - \mathbf{Y} - V\}$$

Non-Interactive Simulation Problem

- Given P_{XY}, let $(\mathbf{X}, \mathbf{Y}) \sim P_{XY}^n$ be correlated memoryless sources
  - i.e., $(\mathbf{X}, \mathbf{Y})$ are n i.i.d. copies of $(X, Y) \sim P_{XY}$
- Assume (U, V) on $\mathcal{U} \times \mathcal{V}$ are two random variables such that $U - \mathbf{X} - \mathbf{Y} - V$ forms a Markov chain, i.e.,

$$P_{U\mathbf{X}\mathbf{Y}V} = P_{U|\mathbf{X}} P_{XY}^n P_{V|\mathbf{Y}}$$



- A natural question: What are the possible joint distributions P_{UV} of (U, V)?

$$\mathcal{Q}(\mathcal{U} \times \mathcal{V} | P_{XY}) := \{P_{UV} \in \mathcal{P}(\mathcal{U} \times \mathcal{V}) : U - \mathbf{X} - \mathbf{Y} - V\}$$

- This problem is termed Non-Interactive Simulation of Random Variables

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$
- Converse results derived by data processing inequalities:

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$
- Converse results derived by data processing inequalities:
  - Witsenhausen (1975) derived a converse result by maximal correlation: $\rho_{\mathrm{m}}(U;V) \le \rho_{\mathrm{m}}(X;Y)$

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$
- Converse results derived by data processing inequalities:
  - Witsenhausen (1975) derived a converse result by maximal correlation: $\rho_{\mathrm{m}}(U;V) \le \rho_{\mathrm{m}}(X;Y)$
  - Kamath-Anantharam (2016) derived a converse result by hypercontractivity: $\mathcal{R}(U;V) \supseteq \mathcal{R}(X;Y)$ ($\mathcal{R}(X;Y)$ is the hypercontractivity ribbon between X, Y)

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$
- Converse results derived by data processing inequalities:
  - Witsenhausen (1975) derived a converse result by maximal correlation: $\rho_{\mathrm{m}}(U; V) \leq \rho_{\mathrm{m}}(X; Y)$
  - Kamath-Anantharam (2016) derived a converse result by hypercontractivity: $\mathcal{R}(U; V) \supseteq \mathcal{R}(X; Y)$ ($\mathcal{R}(X; Y)$ is the hypercontractivity ribbon between X, Y)

Related Problems:

- Non-interactive correlation distillation (Mossel-O'Donnell 2005, Yang 2007): $U, V \sim \mathrm{Bern}\left(\frac{1}{2}\right)$ and maximize $\mathbb{E}UV$

Background and Motivation

Background:

- Used to define common information
  - Gács-Körner (1972) restricted U, V s.t. $\mathbb{P}(U = V) \to 1$ as $n \to \infty$
  - Wyner (1975) considered $X = Y \sim \mathrm{Bern}\left(\frac{1}{2}\right)$
- Converse results derived by data processing inequalities:
  - Witsenhausen (1975) derived a converse result by maximal correlation: $\rho_{\mathrm{m}}(U;V) \le \rho_{\mathrm{m}}(X;Y)$
  - Kamath-Anantharam (2016) derived a converse result by hypercontractivity: $\mathcal{R}(U;V) \supseteq \mathcal{R}(X;Y)$ ($\mathcal{R}(X;Y)$ is the hypercontractivity ribbon between X, Y)

Related Problems:

- Non-interactive correlation distillation (Mossel-O'Donnell 2005, Yang 2007): $U, V \sim \mathrm{Bern}\left(\frac{1}{2}\right)$ and maximize $\mathbb{E}UV$
- Noise-sensitivity of Boolean functions (Mossel-O'Donnell 2005):
  - $X \sim \mathrm{Bern}\left(\frac{1}{2}\right)$, $Y = X \oplus E$ with $E \sim \mathrm{Bern}(p)$ ind. of X
  - $U = f(\mathbf{X})$, $V = f(\mathbf{Y})$ with $f : \{-1, 1\}^n \to \{-1, 1\}$ being a balanced Boolean function (i.e., $\mathbb{P}(U = 1) = \mathbb{P}(V = 1) = \frac{1}{2}$)
  - maximize $\mathbb{P}(U = V)$ (or $\mathbb{E}UV$)

Non-Interactive Simulation: Boolean Version

- Non-Interactive simulation problem is difficult in general

Non-Interactive Simulation: Boolean Version

- Non-Interactive simulation problem is difficult in general
- So in this work, we focus on the binary case:

Non-Interactive Simulation: Boolean Version

- Non-Interactive simulation problem is difficult in general
- So in this work, we focus on the binary case:
  - X, Y, U, V are Boolean random variables taking values in $\{-1, 1\}$
  - P_{XY} is a Boolean symmetric distribution with correlation coefficient $\rho \in [0, 1]$, i.e.,

$$P_{XY} = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{bmatrix} \frac{1+\rho}{4} & \frac{1-\rho}{4} \\ \frac{1-\rho}{4} & \frac{1+\rho}{4} \end{bmatrix} \end{array}$$

Non-Interactive Simulation: Boolean Version

- For this case, P_{UV} is determined by the triple

$$(\mathbb{P}(U = 1), \mathbb{P}(V = 1), \mathbb{P}(U = V = 1))$$

Non-Interactive Simulation: Boolean Version

- For this case, P_{UV} is determined by the triple

$$(\mathbb{P}(U=1), \mathbb{P}(V=1), \mathbb{P}(U=V=1))$$

- The region of the triple above is determined by

$$p_n^+(a,b) := \max_{\substack{U,V: U-\mathbf{X}-\mathbf{Y}-V \\ \mathbb{P}(U=1)=a, \\ \mathbb{P}(V=1)=b}} \mathbb{P}(U=V=1)$$

$$p_n^-(a,b) := \min_{\substack{U,V: U-\mathbf{X}-\mathbf{Y}-V \\ \mathbb{P}(U=1)=a, \\ \mathbb{P}(V=1)=b}} \mathbb{P}(U=V=1)$$

Non-Interactive Simulation: Boolean Version

- For this case, P_{UV} is determined by the triple

$$(\mathbb{P}(U=1), \mathbb{P}(V=1), \mathbb{P}(U=V=1))$$

- The region of the triple above is determined by

$$p_n^+(a,b) := \max_{\substack{U,V: U-\mathbf{X}-\mathbf{Y}-V \\ \mathbb{P}(U=1)=a, \\ \mathbb{P}(V=1)=b}} \mathbb{P}(U=V=1)$$

$$p_n^-(a,b) := \min_{\substack{U,V: U-\mathbf{X}-\mathbf{Y}-V \\ \mathbb{P}(U=1)=a, \\ \mathbb{P}(V=1)=b}} \mathbb{P}(U=V=1)$$

- If we restrict $U = f(\mathbf{X}), V = g(\mathbf{Y})$ for $f, g : \{-1,1\}^n \to \{-1,1\}$, we obtain

$$q_n^+(a,b) := \max_{\substack{f,g: \mathbb{P}(f(\mathbf{X})=1)=a_n, \\ \mathbb{P}(g(\mathbf{Y})=1)=b_n}} \mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$$

$$q_n^-(a,b) := \min_{\substack{f,g: \mathbb{P}(f(\mathbf{X})=1)=a_n, \\ \mathbb{P}(g(\mathbf{Y})=1)=b_n}} \mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$$

where $a_n := \frac{\lfloor 2^n a \rfloor}{2^n}$ and $b_n := \frac{\lfloor 2^n b \rfloor}{2^n}$.

Replace $(P_{U|\mathbf{X}}, P_{V|\mathbf{Y}})$ with Boolean functions (f, g)

Lemma

We have

$$0 \le p_n^+(a,b) - q_n^+(a,b) \le 2^{-(n-1)}$$

$$0 \le p_n^-(a,b) - q_n^-(a,b) \le 2^{-(n-1)}.$$

In particular, if $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$*, then*

$$p_n^+(a,b) = q_n^+(a,b)$$

$$p_n^-(a,b) = q_n^-(a,b).$$

Replace $(P_{U|\mathbf{X}}, P_{V|\mathbf{Y}})$ with Boolean functions (f, g)

Lemma

We have

$$0 \leq p_n^+(a,b) - q_n^+(a,b) \leq 2^{-(n-1)}$$

$$0 \leq p_n^-(a,b) - q_n^-(a,b) \leq 2^{-(n-1)}.$$

In particular, if $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$*, then*

$$p_n^+(a,b) = q_n^+(a,b)$$

$$p_n^-(a,b) = q_n^-(a,b).$$

Proof: Observe that optimizations in $p_n^{\pm}(a,b), q_n^{\pm}(a,b)$ are linear programs. This lemma follows by the simplex method.

Replace $(P_{U|\mathbf{X}}, P_{V|\mathbf{Y}})$ with Boolean functions (f, g)

Lemma

We have

$$0 \le p_n^+(a,b) - q_n^+(a,b) \le 2^{-(n-1)}$$

$$0 \le p_n^-(a,b) - q_n^-(a,b) \le 2^{-(n-1)}.$$

In particular, if $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$*, then*

$$p_n^+(a,b) = q_n^+(a,b)$$

$$p_n^-(a,b) = q_n^-(a,b).$$

Proof: Observe that optimizations in $p_n^\pm(a,b), q_n^\pm(a,b)$ are linear programs. This lemma follows by the simplex method.

- Restricting $U = f(\mathbf{X}), V = g(\mathbf{Y})$ is asymptotically optimal in attaining $p_n^+(a,b), p_n^-(a,b)$

Connection to Coding Theory

- $A \subseteq \{-1, 1\}^n$ is called a binary code

Connection to Coding Theory

- $A \subseteq \{-1, 1\}^n$ is called a binary code
- For a Boolean function f, $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ is a binary code
  - f and A are uniquely determined by each other.

Connection to Coding Theory

- $A \subseteq \{-1,1\}^n$ is called a binary code
- For a Boolean function f, $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ is a binary code
  - f and A are uniquely determined by each other.
- In coding theory, the distance distribution between $A, B \subseteq \{-1,1\}^n$ is ,

$$P^{(A,B)}(i) := \frac{1}{|A||B|} \left|\{(\mathbf{x},\mathbf{x}') \in A \times B : d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') = i\}\right|, \quad i \in \{0,1,\ldots,n\}$$

where $d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') := \left|\left\{i : x_i \neq x'_i\right\}\right|$ denotes the Hamming distance

Connection to Coding Theory

- $A \subseteq \{-1,1\}^n$ is called a binary code
- For a Boolean function f, $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ is a binary code
  - f and A are uniquely determined by each other.
- In coding theory, the distance distribution between $A, B \subseteq \{-1,1\}^n$ is ,

$$P^{(A,B)}(i) := \frac{1}{|A||B|} \left|\{(\mathbf{x}, \mathbf{x}') \in A \times B : d_{\mathrm{H}}(\mathbf{x}, \mathbf{x}') = i\}\right|, \quad i \in \{0, 1, ..., n\}$$

where $d_{\mathrm{H}}(\mathbf{x}, \mathbf{x}') := \left|\left\{i : \ x_i \neq x_i'\right\}\right|$ denotes the Hamming distance

  - In particular, if $A = B$, then

$$P^{(A,A)}(i) := \frac{1}{|A|^2} \left|\left\{\left(\mathbf{x}, \mathbf{x}'\right) \in A^2 : d_{\mathrm{H}}\left(\mathbf{x}, \mathbf{x}'\right) = i\right\}\right|, \quad i \in \{0, 1, ..., n\}$$

is the distance distribution of a single code $A \subseteq \{-1,1\}^n$

Distance Enumerators and Average Distances

- Define the distance enumerator between $A, B \subseteq \{-1, 1\}^n$ as

$$\Gamma_z(A, B) := \frac{1}{|A||B|} \sum_{\mathbf{x}\in A} \sum_{\mathbf{x}'\in B} z^{d_{\mathrm{H}}(\mathbf{x},\mathbf{x}')} = \sum_{i=0}^{n} P^{(A,B)}(i) \cdot z^i.$$

  - Clearly, $\Gamma_z(A, B)$ is the probability-generating function of $P^{(A,B)}$.

Distance Enumerators and Average Distances

- Define the distance enumerator between $A, B \subseteq \{-1,1\}^n$ as

$$\Gamma_z(A,B) := \frac{1}{|A||B|}\sum_{\mathbf{x}\in A}\sum_{\mathbf{x}'\in B} z^{d_{\mathrm{H}}(\mathbf{x},\mathbf{x}')} = \sum_{i=0}^{n} P^{(A,B)}(i)\cdot z^i.$$

  - Clearly, $\Gamma_z(A,B)$ is the probability-generating function of $P^{(A,B)}$.

- The dual distance enumerator between $A, B \subseteq \{-1,1\}^n$ is defined as

$$\Pi_z(A,B) := (1+z)^n\,\Gamma_{\frac{1-z}{1+z}}(A,B).$$

Distance Enumerators and Average Distances

- Define the distance enumerator between $A, B \subseteq \{-1,1\}^n$ as

$$\Gamma_z(A,B) := \frac{1}{|A||B|}\sum_{\mathbf{x}\in A}\sum_{\mathbf{x}'\in B} z^{d_{\mathrm{H}}(\mathbf{x},\mathbf{x}')} = \sum_{i=0}^{n} P^{(A,B)}(i)\cdot z^i.$$

  - Clearly, $\Gamma_z(A,B)$ is the probability-generating function of $P^{(A,B)}$.

- The dual distance enumerator between $A, B \subseteq \{-1,1\}^n$ is defined as

$$\Pi_z(A,B) := (1+z)^n \, \Gamma_{\frac{1-z}{1+z}}(A,B).$$

- The average distance between $A, B \subseteq \{-1,1\}^n$ is defined as

$$D(A,B) := \frac{1}{|A||B|}\sum_{\mathbf{x}\in A}\sum_{\mathbf{x}'\in B} d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') = \sum_{i=0}^{n} P^{(A,B)}(i)\cdot i$$

  - Clearly, $D(A,B)$ is the mean of $P^{(A,B)}$.

Equivalence

Lemma

For $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$*, we have*

$$\mathbb{P}\left(f(\mathbf{X}) = g(\mathbf{Y}) = 1\right) = ab\,(1+\rho)^n\,\Gamma_{\frac{1-\rho}{1+\rho}}\left(A, B\right) = ab\Pi_\rho\left(A, B\right)$$

where $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ *and* $B := \{\mathbf{x} : g(\mathbf{x}) = 1\}$*.*

Equivalence

Lemma

For $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$, *we have*

$$\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1) = ab(1+\rho)^n \Gamma_{\frac{1-\rho}{1+\rho}}(A, B) = ab\Pi_\rho(A, B)$$

where $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ *and* $B := \{\mathbf{x} : g(\mathbf{x}) = 1\}$.

- Given a, b, ρ, characterizing the possible range of $\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$ is equivalent to characterizing the possible range of $\Gamma_{\frac{1-\rho}{1+\rho}}(A, B)$ or $\Pi_\rho(A, B)$

Equivalence

Lemma

For $a = \frac{M}{2^n}$ *and* $b = \frac{N}{2^n}$ *for some* $M, N \in \mathbb{N}$*, we have*

$$\mathbb{P}\left(f(\mathbf{X}) = g(\mathbf{Y}) = 1\right) = ab\,(1+\rho)^n\,\Gamma_{\frac{1-\rho}{1+\rho}}(A,B) = ab\Pi_\rho(A,B)$$

where $A := \{\mathbf{x} : f(\mathbf{x}) = 1\}$ *and* $B := \{\mathbf{x} : g(\mathbf{x}) = 1\}$*.*

- Given a, b, ρ, characterizing the possible range of $\mathbb{P}\left(f(\mathbf{X}) = g(\mathbf{Y}) = 1\right)$ is equivalent to characterizing the possible range of $\Gamma_{\frac{1-\rho}{1+\rho}}(A,B)$ or $\Pi_\rho(A,B)$
- The (Boolean function version of) non-interactive simulation problem $\Longleftrightarrow$ the problem of determining the possible range of the (dual) distance enumerator

Main Result

Assume $a = b = \frac{M}{2^n}$ for some $M \in \mathbb{N}$. Denote $q := \mathbb{P}\left(f(\mathbf{X}) = g(\mathbf{Y}) = 1\right)$.

Main Result

Assume $a = b = \frac{M}{2^n}$ for some $M \in \mathbb{N}$. Denote $q := \mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$.

Theorem (Symmetric Case: $a = b$)

$$\theta^-(a) \le q \le \theta^+(a),$$

where

$$\theta^+(a) := \min\left\{a, a^2 + \frac{a}{2}\rho + \left(\frac{a}{2} - a^2\right)\rho^2\right\}$$

$$\theta^-(a) := \max\left\{0, a^2 - \frac{a}{2}\rho - \left(\frac{a}{2} - a^2\right)\rho^2\right\}.$$

Main Result

Assume $a = b = \frac{M}{2^n}$ for some $M \in \mathbb{N}$. Denote $q := \mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$.

Theorem (Symmetric Case: $a = b$)

$$\theta^-(a) \le q \le \theta^+(a),$$

where

$$\theta^+(a) := \min\left\{a, a^2 + \frac{a}{2}\rho + \left(\frac{a}{2} - a^2\right)\rho^2\right\}$$
$$\theta^-(a) := \max\left\{0, a^2 - \frac{a}{2}\rho - \left(\frac{a}{2} - a^2\right)\rho^2\right\}.$$

In particular, for $a = \frac{1}{2}$, (Witsenhausen's result (1975))

$$\frac{1-\rho}{4} \le q \le \frac{1+\rho}{4},$$

and for $a = \frac{1}{4}$, (new)

$$\frac{1-2\rho-\rho^2}{16} \le q \le \left(\frac{1+\rho}{4}\right)^2.$$

Main Result

Assume $a = b = \frac{M}{2^n}$ for some $M \in \mathbb{N}$. Denote $q := \mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$.

Theorem (Symmetric Case: $a = b$)

$$\theta^-(a) \le q \le \theta^+(a),$$

where

$$\theta^+(a) := \min\left\{a, a^2 + \frac{a}{2}\rho + \left(\frac{a}{2} - a^2\right)\rho^2\right\}$$

$$\theta^-(a) := \max\left\{0, a^2 - \frac{a}{2}\rho - \left(\frac{a}{2} - a^2\right)\rho^2\right\}.$$

In particular, for $a = \frac{1}{2}$, (Witsenhausen's result (1975))

$$\frac{1-\rho}{4} \le q \le \frac{1+\rho}{4},$$

and for $a = \frac{1}{4}$, (new)

$$\frac{1-2\rho-\rho^2}{16} \le q \le \left(\frac{1+\rho}{4}\right)^2.$$

- Our bounds also hold for $q := \mathbb{P}(U = V = 1)$ (stochastic version).
- Our results for asymmetric cases can be found in our paper.

Main Result

For $a = \frac{1}{2}$, (Witsenhausen's result (1975))

$$\frac{1-\rho}{4} \leq q \leq \frac{1+\rho}{4},$$

and for $a = \frac{1}{4}$, (new)

$$\frac{1-2\rho-\rho^2}{16} \leq q \leq \left(\frac{1+\rho}{4}\right)^2.$$

Main Result

For $a = \frac{1}{2}$, (Witsenhausen's result (1975))

$$\frac{1-\rho}{4} \le q \le \frac{1+\rho}{4},$$

and for $a = \frac{1}{4}$, (new)

$$\frac{1-2\rho-\rho^2}{16} \le q \le \left(\frac{1+\rho}{4}\right)^2.$$

- Both the upper and lower bounds for the case $a = \frac{1}{2}$ are sharp:
  - the upper bound is attained by $g(\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = 1\}$ (symmetric subcube functions)
  - the lower bound is attained by $g(-\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = 1\}$ (anti-symmetric subcube functions)

Main Result

For $a = \frac{1}{2}$, (Witsenhausen's result (1975))

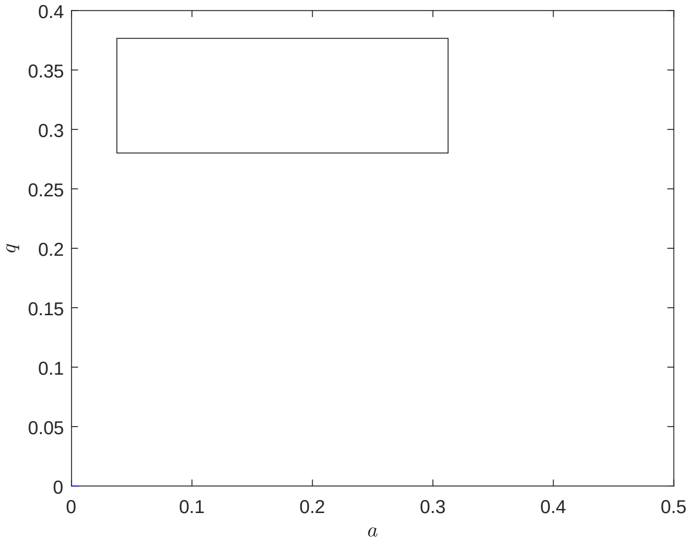
$$\frac{1-\rho}{4} \leq q \leq \frac{1+\rho}{4},$$

and for $a = \frac{1}{4}$, (new)

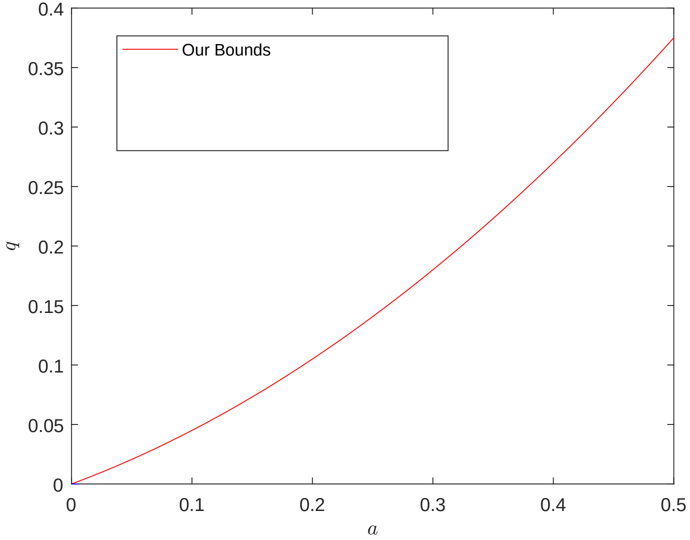
$$\frac{1-2\rho-\rho^2}{16} \leq q \leq \left(\frac{1+\rho}{4}\right)^2.$$

- Both the upper and lower bounds for the case $a = \frac{1}{2}$ are sharp:
  - the upper bound is attained by $g(\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = 1\}$ (symmetric subcube functions)
  - the lower bound is attained by $g(-\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = 1\}$ (anti-symmetric subcube functions)
- The upper bound for the case $a = \frac{1}{4}$ is sharp:
  - attained by $g(\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = x_2 = 1\}$

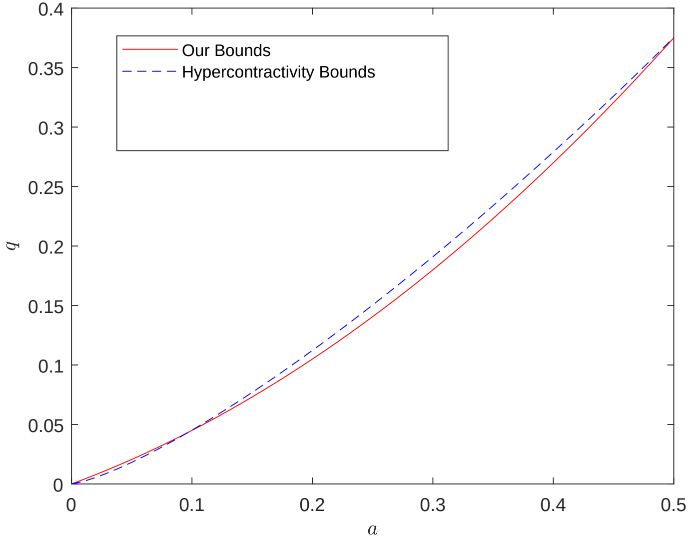
Numerical Result: Upper Bounds



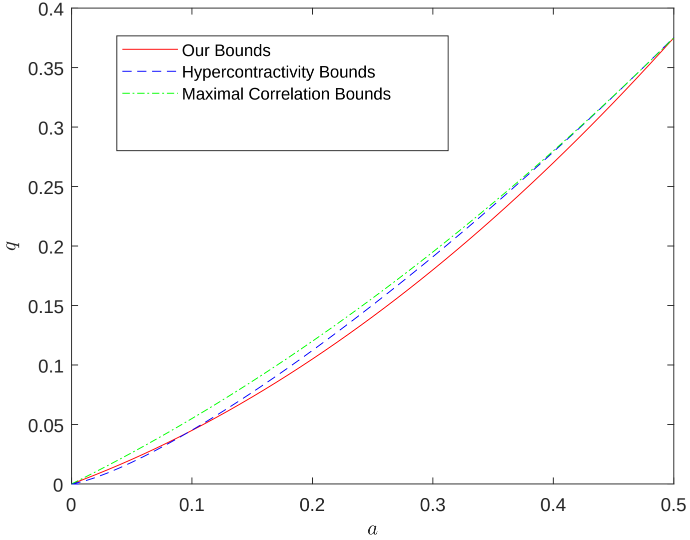
Numerical Result: Upper Bounds



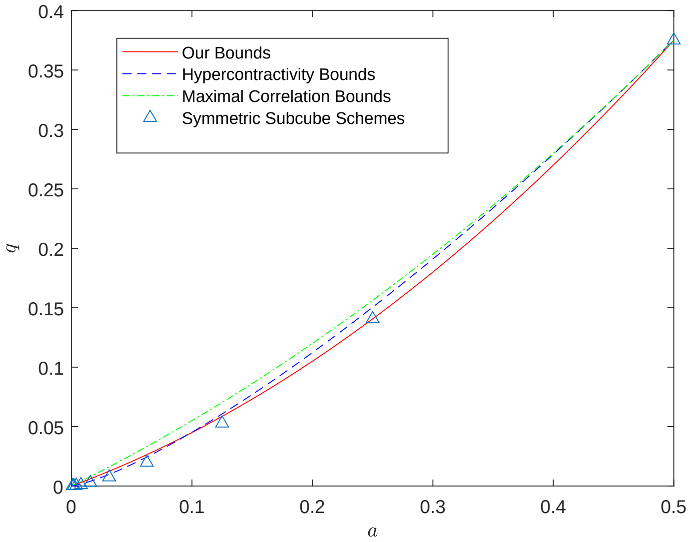
Numerical Result: Upper Bounds



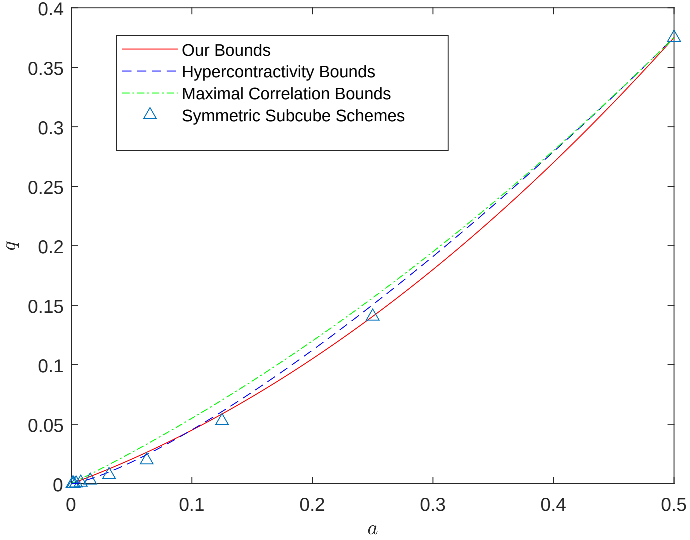
Numerical Result: Upper Bounds



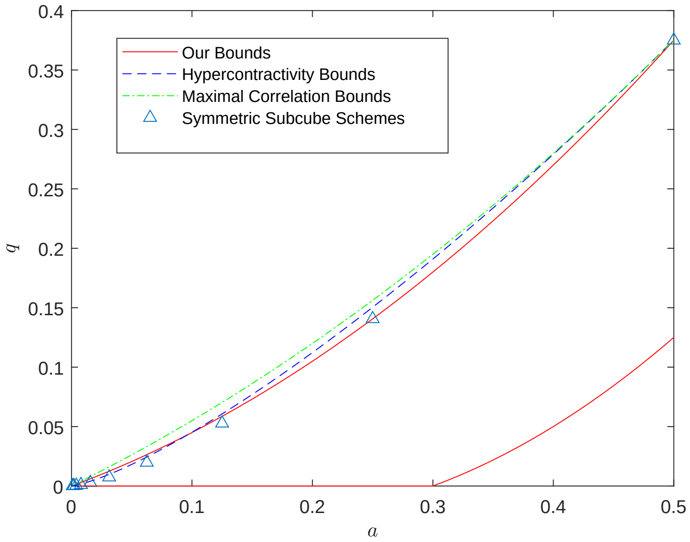
Numerical Result: Upper Bounds



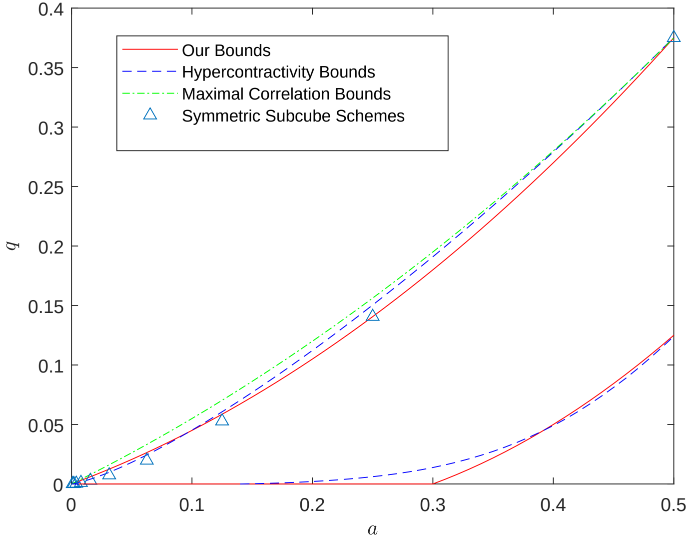
Numerical Result: Lower Bounds



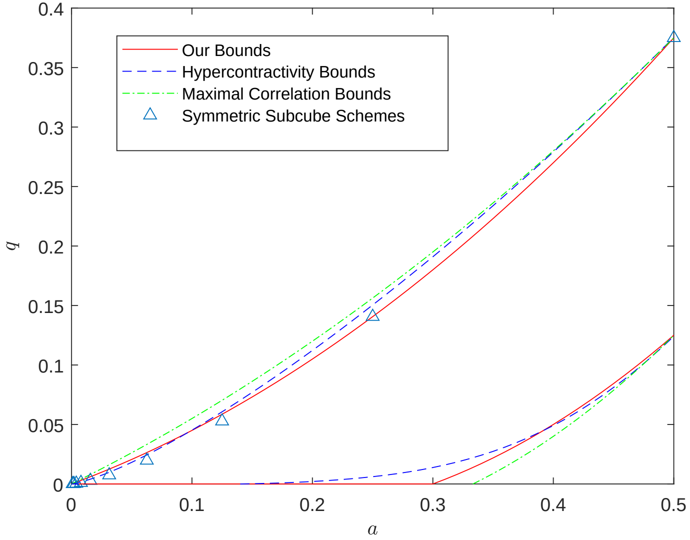
Numerical Result: Lower Bounds



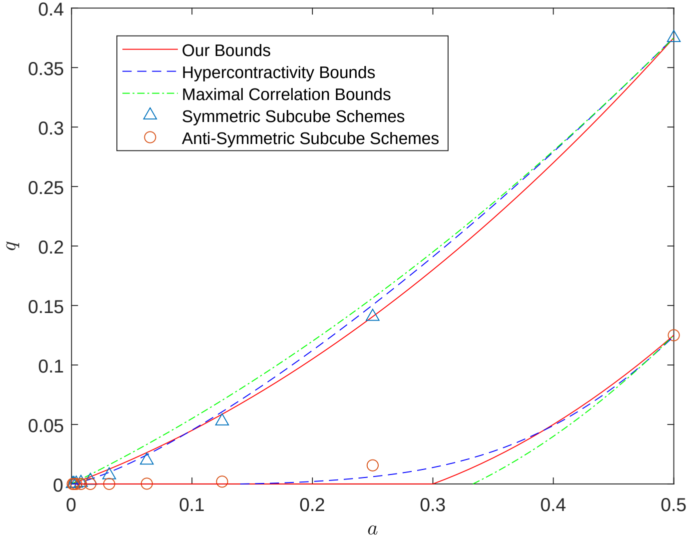
Numerical Result: Lower Bounds



Numerical Result: Lower Bounds



Numerical Result: Lower Bounds



Proof Idea – Fourier Analysis

- Consider the Fourier/Hadamard basis

$$\chi_S(\mathbf{x}) := \prod_{i \in S} x_i, \quad S \subseteq [n] := \{1, ..., n\}$$

Proof Idea – Fourier Analysis

- Consider the Fourier/Hadamard basis

$$\chi_S(\mathbf{x}) := \prod_{i \in S} x_i, \quad S \subseteq [n] := \{1, ..., n\}$$

- For a Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$, its Fourier/Hadamard transform is

$$\hat{f}_S := \mathbb{E}_{\mathbf{x} \sim \mathrm{Unif}\{-1,1\}^n}[f(\mathbf{x})\chi_S(\mathbf{x})], \quad S \subseteq [n]. \tag{1}$$

Proof Idea – Fourier Analysis

- Consider the Fourier/Hadamard basis

$$\chi_S(\mathbf{x}) := \prod_{i \in S} x_i, \quad S \subseteq [n] := \{1, \dots, n\}$$

- For a Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$, its Fourier/Hadamard transform is

$$\hat{f}_S := \mathbb{E}_{\mathbf{x} \sim \mathrm{Unif}\{-1,1\}^n} [f(\mathbf{x}) \chi_S(\mathbf{x})], \quad S \subseteq [n]. \tag{1}$$

- The inverse Fourier transform is

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \hat{f}_S \chi_S(\mathbf{x})$$

Proof Idea – Fourier Analysis

- Consider the Fourier/Hadamard basis
$$\chi_S(\mathbf{x}) := \prod_{i \in S} x_i, \quad S \subseteq [n] := \{1, ..., n\}$$
- For a Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$, its Fourier/Hadamard transform is
$$\hat{f}_S := \mathbb{E}_{\mathbf{x} \sim \mathrm{Unif}\{-1,1\}^n}[f(\mathbf{x})\chi_S(\mathbf{x})], \quad S \subseteq [n]. \tag{1}$$
- The inverse Fourier transform is
$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \hat{f}_S \chi_S(\mathbf{x})$$
- Then we can rewrite
$$\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1) = ab + \frac{1}{4} \sum_{k=1}^{n} Q(k)\rho^k$$

  where
$$Q(k) := \sum_{S \subseteq [n]: |S| = k} \hat{f}_S \hat{g}_S, \quad 1 \le k \le n \tag{2}$$

Proof Idea – Fourier Analysis

- Consider the Fourier/Hadamard basis

$$\chi_S(\mathbf{x}) := \prod_{i \in S} x_i, \quad S \subseteq [n] := \{1, \dots, n\}$$

- For a Boolean function $f : \{-1, 1\}^n \to \{-1, 1\}$, its Fourier/Hadamard transform is

$$\hat{f}_S := \mathbb{E}_{\mathbf{x} \sim \mathrm{Unif}\{-1,1\}^n}[f(\mathbf{x})\chi_S(\mathbf{x})], \quad S \subseteq [n]. \tag{1}$$

- The inverse Fourier transform is

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} \hat{f}_S \chi_S(\mathbf{x})$$

- Then we can rewrite

$$\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1) = ab + \frac{1}{4} \sum_{k=1}^{n} Q(k)\rho^k$$

where

$$Q(k) := \sum_{S \subseteq [n] : |S| = k} \hat{f}_S \hat{g}_S, \quad 1 \le k \le n \tag{2}$$

- To bound $\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1)$, we only need to bound $\sum_{k=1}^{n} Q(k)\rho^k$

Proof Idea – Fourier Analysis

Now we bound $\sum_{k=1}^{n} Q(k)\rho^k$:

Proof Idea – Fourier Analysis

Now we bound $\sum_{k=1}^{n} Q(k)\rho^k$:

- Step 1: Bound $Q(1)$:
  - We show that

$$Q(1) = 8ab\left(\frac{n}{2} - D\left(A, B\right)\right)$$

$$\left|\frac{n}{2} - D\left(A, B\right)\right| \leq \frac{n}{2} - \frac{1}{2}\left(D\left(A, A\right) + D\left(B, B\right)\right).$$

Proof Idea – Fourier Analysis

Now we bound $\sum_{k=1}^{n} Q(k)\rho^k$:

- Step 1: Bound $Q(1)$:
  - We show that

$$Q(1) = 8ab\left(\frac{n}{2} - D(A, B)\right)$$

$$\left|\frac{n}{2} - D(A, B)\right| \le \frac{n}{2} - \frac{1}{2}\left(D(A, A) + D(B, B)\right).$$

  - Fu-Wei-Yeung (2001) showed the following (linear programming) bound on average distance

$$\min_{A:|A|=M} D(A, A) \ge \frac{n}{2} - \frac{1}{4a}$$

    where $a = \frac{M}{2^n}$.

Proof Idea – Fourier Analysis

Now we bound $\sum_{k=1}^{n} Q(k)\rho^k$:

- Step 1: Bound $Q(1)$:
  - We show that
$$Q(1) = 8ab\left(\frac{n}{2} - D(A,B)\right)$$
$$\left|\frac{n}{2} - D(A,B)\right| \le \frac{n}{2} - \frac{1}{2}\left(D(A,A) + D(B,B)\right).$$
  - Fu-Wei-Yeung (2001) showed the following (linear programming) bound on average distance
$$\min_{A:|A|=M} D(A,A) \ge \frac{n}{2} - \frac{1}{4a}$$
where $a = \frac{M}{2^n}$.
  - Combining the results above gives
$$|Q(1)| \le a + b$$

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^{k}$:

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^k$:
  - Following Pichler-Piantanida-Matz's idea (2018), we define

$$\tau^+ := \sum_{S\in\mathcal{P}} \hat{f}_S \hat{g}_S, \qquad \tau^- := \sum_{S\in\mathcal{N}} \hat{f}_S \hat{g}_S$$

where $\mathcal{P} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S \hat{g}_S \geq 0\}$ and $\mathcal{N} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S \hat{g}_S < 0\}$

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^k$:
  - Following Pichler-Piantanida-Matz's idea (2018), we define

$$\tau^+ := \sum_{S\in\mathcal{P}} \hat{f}_S\hat{g}_S, \qquad \tau^- := \sum_{S\in\mathcal{N}} \hat{f}_S\hat{g}_S$$

    where $\mathcal{P} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S \geq 0\}$ and $\mathcal{N} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S < 0\}$
  - Then

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S\hat{g}_S\rho^{|S|} \in \left[\tau^-\rho^2, \tau^+\rho^2\right]$$

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^k$:
  - Following Pichler-Piantanida-Matz's idea (2018), we define

$$\tau^+ := \sum_{S\in\mathcal{P}} \hat{f}_S\hat{g}_S, \qquad \tau^- := \sum_{S\in\mathcal{N}} \hat{f}_S\hat{g}_S$$

    where $\mathcal{P} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S \geq 0\}$ and $\mathcal{N} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S < 0\}$
  - Then

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S\hat{g}_S\rho^{|S|} \in \left[\tau^-\rho^2, \tau^+\rho^2\right]$$

  - Now we only need to bound τ^+, τ^-:

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^k$:
  - Following Pichler-Piantanida-Matz's idea (2018), we define

$$\tau^+ := \sum_{S\in\mathcal{P}} \hat{f}_S\hat{g}_S, \qquad \tau^- := \sum_{S\in\mathcal{N}} \hat{f}_S\hat{g}_S$$

    where $\mathcal{P} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S \geq 0\}$ and $\mathcal{N} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S < 0\}$
  - Then

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S\hat{g}_S\rho^{|S|} \in \left[\tau^-\rho^2, \tau^+\rho^2\right]$$

  - Now we only need to bound τ^+, τ^-:
    - We show $\tau^+ - \tau^- \leq 4\sqrt{a\bar{a}b\bar{b}} - Q(1)$ by using Parseval's Theorem ($\sum_{S:|S|\geq 0} \hat{f}_S^2 = 1$)
    - We show $-4ab - Q(1) \leq \tau^+ + \tau^- \leq 4a\bar{b} - Q(1)$

Proof Idea – Fourier Analysis

- Step 2: Bound $\sum_{k=2}^{n} Q(k)\rho^k$:
  - Following Pichler-Piantanida-Matz's idea (2018), we define

$$\tau^+ := \sum_{S\in\mathcal{P}} \hat{f}_S\hat{g}_S, \qquad \tau^- := \sum_{S\in\mathcal{N}} \hat{f}_S\hat{g}_S$$

  where $\mathcal{P} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S \geq 0\}$ and $\mathcal{N} := \{S \subseteq [n] : |S| \geq 2, \hat{f}_S\hat{g}_S < 0\}$
  - Then

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S\hat{g}_S\rho^{|S|} \in [\tau^-\rho^2, \tau^+\rho^2]$$

  - Now we only need to bound τ^+, τ^-:
    - We show $\tau^+ - \tau^- \leq 4\sqrt{a\bar{a}b\bar{b}} - Q(1)$ by using Parseval's Theorem ($\sum_{S:|S|\geq 0} \hat{f}_S^2 = 1$)
    - We show $-4ab - Q(1) \leq \tau^+ + \tau^- \leq 4a\bar{b} - Q(1)$
- Finally, combining Steps 1 and 2 yields our bounds: $\theta^-(a) \leq q \leq \theta^+(a)$, where

$$\theta^+(a) = \min\left\{a, a^2 + \frac{a}{2}\rho + \left(\frac{a}{2} - a^2\right)\rho^2\right\}$$

$$\theta^-(a) = \max\left\{0, a^2 - \frac{a}{2}\rho - \left(\frac{a}{2} - a^2\right)\rho^2\right\}.$$

Why our proof works?

- In our Step 2, we use the following bounds:

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S \hat{g}_S \rho^{|S|} \in \left[\tau^- \rho^2, \tau^+ \rho^2\right]$$

  - This implies that we discard $Q(k), k \geq 3$

Why our proof works?

- In our Step 2, we use the following bounds:

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S \hat{g}_S \rho^{|S|} \in \left[\tau^- \rho^2, \tau^+ \rho^2\right]$$

  - This implies that we discard $Q(k), k \geq 3$

- Conjecture: Given $a = b = 2^{-m}$, the optimal f, g are subcube functions, i.e., $g(\pm\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = \ldots = x_m = 1\}$

Why our proof works?

- In our Step 2, we use the following bounds:

$$\sum_{k=2}^{n} Q(k)\rho^k = \sum_{S\subseteq[n]:|S|\geq 2} \hat{f}_S \hat{g}_S \rho^{|S|} \in \left[\tau^- \rho^2, \tau^+ \rho^2\right]$$

  - This implies that we discard $Q(k), k \geq 3$

- Conjecture: Given $a = b = 2^{-m}$, the optimal f, g are subcube functions, i.e., $g(\pm\mathbf{x}) = f(\mathbf{x}) = 1\{x_1 = \ldots = x_m = 1\}$

- Subcube functions satisfy $Q(k) = 0, k \geq m+1$

A New Bound on Average Distances

- Until now, we have shown the equivalence

$$\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1) = ab(1+\rho)^n \Gamma_{\frac{1-\rho}{1+\rho}}(A, B) = ab\Pi_\rho(A, B)$$

  - Non-interactive simulation is equivalent to some coding-theoretic problem

A New Bound on Average Distances

- Until now, we have shown the equivalence

$$\mathbb{P}\left(f(\mathbf{X}) = g(\mathbf{Y}) = 1\right) = ab\left(1+\rho\right)^{n}\Gamma_{\frac{1-\rho}{1+\rho}}\left(A, B\right) = ab\Pi_{\rho}\left(A, B\right)$$

  - Non-interactive simulation is equivalent to some coding-theoretic problem
- We have applied coding-theoretic results to non-interactive simulation

A New Bound on Average Distances

- Until now, we have shown the equivalence

$$\mathbb{P}(f(\mathbf{X}) = g(\mathbf{Y}) = 1) = ab\,(1+\rho)^n\,\Gamma_{\frac{1-\rho}{1+\rho}}(A,B) = ab\Pi_\rho(A,B)$$

  - Non-interactive simulation is equivalent to some coding-theoretic problem
- We have applied coding-theoretic results to non-interactive simulation
- Next, in turn, we apply techniques for non-interactive simulation to a coding-theoretic problem
  - Specifically, apply hypercontractivity inequalities to bound average distances
- Recall that: The average distance between A, B is defined as

$$D(A,B) := \frac{1}{|A||B|}\sum_{\mathbf{x}\in A}\sum_{\mathbf{x}'\in B} d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') = \sum_{i=0}^{n} P^{(A,B)}(i)\cdot i$$

Main Result: A New Bound on Average Distances

By hypercontractivity inequalities, we obtain:

Theorem

For $1 \le M \le 2^n$*, we have*

$$\min_{A:|A|=M} D(A,A) \ge \frac{n}{2} - \psi(a),$$

where $a := \frac{M}{2^n}$ *and*

$$\psi(a) := \inf_{t>0, t\neq 1} \frac{(ta+\overline{a})\left[at\log t - (ta+\overline{a})\log(ta+\overline{a})\right]}{a^2(t-1)^2}.$$

Main Result: A New Bound on Average Distances

By hypercontractivity inequalities, we obtain:

Theorem

For $1 \leq M \leq 2^n$*, we have*

$$\min_{A:|A|=M} D(A,A) \geq \frac{n}{2} - \psi(a),$$

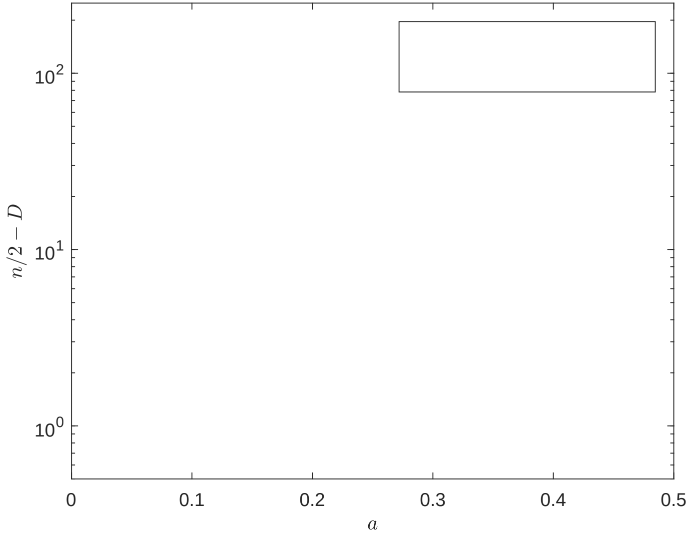
where $a := \frac{M}{2^n}$ *and*

$$\psi(a) := \inf_{t>0, t\neq 1} \frac{(ta+\overline{a})\left[at\log t - (ta+\overline{a})\log(ta+\overline{a})\right]}{a^2(t-1)^2}.$$

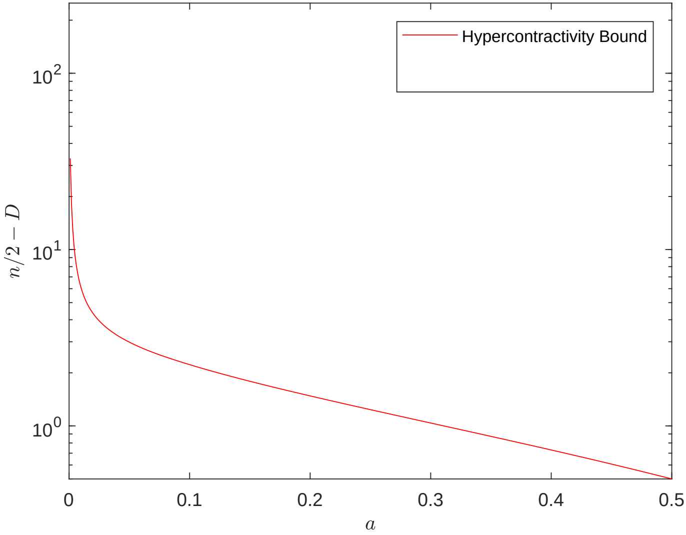
- Best known result: Fu-Wei-Yeung (2001) showed the following (linear programming) bound

$$\min_{A:|A|=M} D(A,A) \geq \frac{n}{2} - \frac{1}{4a}$$

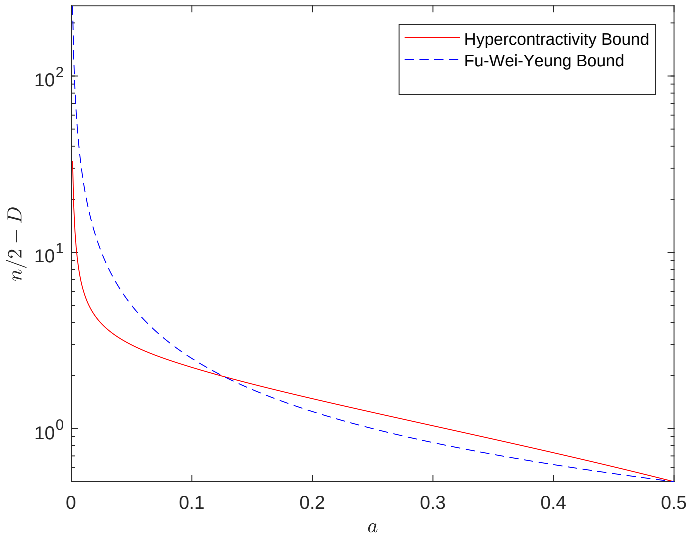
Numerical Result



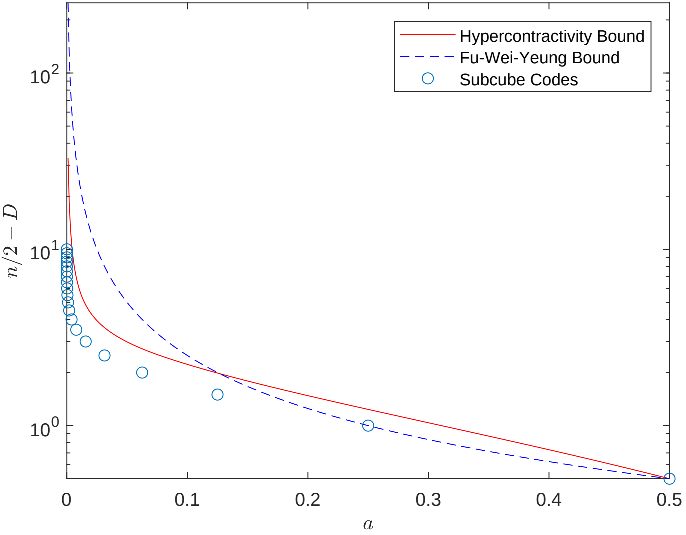
Numerical Result



Numerical Result



Numerical Result



Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful

Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful
- For non-interactive simulation problem, data processing inequalities (DPIs) are very useful

Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful
- For non-interactive simulation problem, data processing inequalities (DPIs) are very useful

In this work:

Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful
- For non-interactive simulation problem, data processing inequalities (DPIs) are very useful

In this work:

- Equivalence: non-interactive simulation problem $\Longleftrightarrow$ some coding-theoretic problem

Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful
- For non-interactive simulation problem, data processing inequalities (DPIs) are very useful

In this work:

- Equivalence: non-interactive simulation problem $\Longleftrightarrow$ some coding-theoretic problem
- We applied Fourier analysis (combined with linear programming) to the non-interactive simulation problem
  - Our bounds are sharp for some cases and tighter than existing results for some other cases

Conclusion

- For coding-theoretic problems, Fourier analysis and linear programming techniques are very useful
- For non-interactive simulation problem, data processing inequalities (DPIs) are very useful

In this work:

- Equivalence: non-interactive simulation problem $\Longleftrightarrow$ some coding-theoretic problem
- We applied Fourier analysis (combined with linear programming) to the non-interactive simulation problem
  - Our bounds are sharp for some cases and tighter than existing results for some other cases
- In turn, applied DPIs (hypercontractivity) to the minimal average-distance problem
  - Our bound is tighter than the best known result for some cases