An Improved Upper Bound for the Most Informative Boolean Function Conjecture

Or Ordentlich, Ofer Shayevitz and Omri Weinstein
MIT TAU NYU

ISIT,
Barcelona,
July 11, 2016

The Most Informative Boolean Function Conjecture

- $\mathbf{X} \sim \mathrm{Unif}(\{0,1\}^n)$ (n i.i.d. $\mathrm{Bernoulli}\left(\frac{1}{2}\right)$ RVs)
- $Z_i \sim \mathrm{Bernoulli}(\alpha)$, i.i.d.
- $Y_i = X_i \oplus Z_i$, $\mathbf{Y} = (Y_1, \ldots, Y_n)$

The Most Informative Boolean Function Conjecture

- $\mathbf{X} \sim \mathrm{Unif}(\{0,1\}^n)$ (n i.i.d. $\mathrm{Bernoulli}\left(\frac{1}{2}\right)$ RVs)
- $Z_i \sim \mathrm{Bernoulli}(\alpha)$, i.i.d.
- $Y_i = X_i \oplus Z_i$, $\mathbf{Y} = (Y_1, \ldots, Y_n)$

Courtade-Kumar Conjecture [IT'14]

for all boolean functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha)$$

The Most Informative Boolean Function Conjecture

- $\mathbf{X} \sim \mathrm{Unif}(\{0,1\}^n)$ (n i.i.d. $\mathrm{Bernoulli}\left(\frac{1}{2}\right)$ RVs)
- $Z_i \sim \mathrm{Bernoulli}(\alpha)$, i.i.d.
- $Y_i = X_i \oplus Z_i$, $\mathbf{Y} = (Y_1, \dots, Y_n)$

Courtade-Kumar Conjecture [IT'14]

for all boolean functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha)$$

In other words: no function is better than $f(\mathbf{X}) = X_i$

The Most Informative Boolean Function Conjecture

- $\mathbf{X} \sim \mathrm{Unif}(\{0,1\}^n)$ (n i.i.d. $\mathrm{Bernoulli}\left(\frac{1}{2}\right)$ RVs)
- $Z_i \sim \mathrm{Bernoulli}(\alpha)$, i.i.d.
- $Y_i = X_i \oplus Z_i$, $\mathbf{Y} = (Y_1, \ldots, Y_n)$

Courtade-Kumar Conjecture [IT'14]

for all boolean functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha)$$

Despite considerable effort, conjecture still open

The Most Informative Boolean Function Conjecture

- $\mathbf{X} \sim \mathrm{Unif}(\{0,1\}^n)$ (n i.i.d. $\mathrm{Bernoulli}\left(\frac{1}{2}\right)$ RVs)
- $Z_i \sim \mathrm{Bernoulli}(\alpha)$, i.i.d.
- $Y_i = X_i \oplus Z_i$, $\mathbf{Y} = (Y_1, \ldots, Y_n)$

Courtade-Kumar Conjecture [IT'14]

for all boolean functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha)$$

Despite considerable effort, conjecture still open

This work:

- New upper bound $I(f(\mathbf{X}); \mathbf{Y}) \leq g(\alpha)$ that holds for all balanced functions
- $\lim_{\alpha \to 1/2} \frac{g(\alpha)}{1-h(\alpha)} = 1$

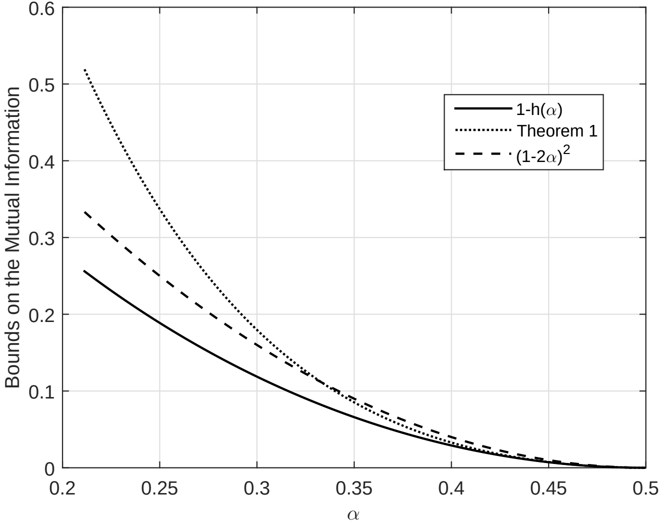
Main Result

Theorem

For any balanced function $f : \{0,1\}^n \mapsto \{-1,1\}$ and any $\frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \leq \alpha \leq \frac{1}{2}$, we have that

$$I(f(\mathbf{X}); \mathbf{Y}) \leq \frac{\log_2(e)}{2}(1-2\alpha)^2 + 9\left(1 - \frac{\log_2(e)}{2}\right)(1-2\alpha)^4.$$

Main Result



Bounds on the Mutual Information
0.6
0.5
0.4
0.3
0.2
0.1
0
0.2
0.25
0.3
0.35
0.4
0.45
0.5
α
1-h(α)
Theorem 1
(1-2α)^2

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \end{aligned}$$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \end{aligned}$$

Let

- $\mathcal{A}_{-1} \triangleq f^{-1}(-1) \triangleq \{x \;:\; f(x) = -1\}$
- $\mathcal{A}_{1} \triangleq f^{-1}(1) \triangleq \{x \;:\; f(x) = 1\}$
- $\mathbf{U}_{-1} \sim \mathrm{Unif}(\mathcal{A}_{-1})$, $\mathbf{U}_{1} \sim \mathrm{Unif}(\mathcal{A}_{1})$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \\ &= n - \tfrac{1}{2}H(\mathbf{U}_{-1} \oplus \mathbf{Z}) - \tfrac{1}{2}H(\mathbf{U}_1 \oplus \mathbf{Z}) \end{aligned}$$

Let

- $\mathcal{A}_{-1} \triangleq f^{-1}(-1) \triangleq \{x \ : \ f(x) = -1\}$
- $\mathcal{A}_1 \triangleq f^{-1}(1) \triangleq \{x \ : \ f(x) = 1\}$
- $\mathbf{U}_{-1} \sim \mathrm{Unif}(\mathcal{A}_{-1})$, $\mathbf{U}_1 \sim \mathrm{Unif}(\mathcal{A}_1)$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \\ &= n - \tfrac{1}{2}H(\mathbf{U}_{-1} \oplus \mathbf{Z}) - \tfrac{1}{2}H(\mathbf{U}_1 \oplus \mathbf{Z}) \end{aligned}$$

Mrs. Gerber's Lemma [WZ73]

$$H(\mathbf{W} \oplus \mathbf{Z}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{W})}{n}\right)\right)$$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$
\begin{aligned}
I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\
&= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \\
&= n - \tfrac{1}{2}H(\mathbf{U}_{-1} \oplus \mathbf{Z}) - \tfrac{1}{2}H(\mathbf{U}_1 \oplus \mathbf{Z}) \\
&\leq n\left(1 - h\left(\alpha * h^{-1}\left(1 - \frac{1}{n}\right)\right)\right)
\end{aligned}
$$

Mrs. Gerber's Lemma [WZ73]

$$
H(\mathbf{W} \oplus \mathbf{Z}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{W})}{n}\right)\right)
$$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X})=1)=1/2$)

$$\begin{aligned} I(f(\mathbf{X});\mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X})=-1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X})=1) \\ &= n - \tfrac{1}{2}H(\mathbf{U}_{-1}\oplus\mathbf{Z}) - \tfrac{1}{2}H(\mathbf{U}_{1}\oplus\mathbf{Z}) \\ &\leq n\left(1 - h\left(\alpha * h^{-1}\left(1-\frac{1}{n}\right)\right)\right) \\ &\overset{n\to\infty}{\to} (1-2\alpha)^2 \end{aligned}$$

Mrs. Gerber's Lemma [WZ73]

$$H(\mathbf{W}\oplus\mathbf{Z}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{W})}{n}\right)\right)$$

Simple Attempts: MGL

For simplicity assume f is balanced ($\Pr(f(\mathbf{X}) = 1) = 1/2$)

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &= H(\mathbf{Y}) - H(\mathbf{Y}|f(\mathbf{X})) \\ &= n - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = -1) - \tfrac{1}{2}H(\mathbf{Y}|f(\mathbf{X}) = 1) \\ &= n - \tfrac{1}{2}H(\mathbf{U}_{-1} \oplus \mathbf{Z}) - \tfrac{1}{2}H(\mathbf{U}_1 \oplus \mathbf{Z}) \\ &\leq n\left(1 - h\left(\alpha * h^{-1}\left(1 - \frac{1}{n}\right)\right)\right) \\ &\overset{n\to\infty}{\to} (1-2\alpha)^2 \end{aligned}$$

Mrs. Gerber's Lemma [WZ73]

$$H(\mathbf{W} \oplus \mathbf{Z}) \geq nh\left(\alpha * h^{-1}\left(\frac{H(\mathbf{W})}{n}\right)\right)$$

Weakness: MGL not tight for **W** uniform on subsets

Simple Attempts: SDPI

Let

$$\eta_{\mathsf{KL}}(P_{Y|X}, Q) \triangleq \sup_{P:0<D(P||Q)<1} \frac{D(P_{Y|X} \circ P || P_{Y|X} \circ Q)}{D(P||Q)}$$

- $\eta_{\mathsf{KL}}(P_{Y|X}, Q)$ tensorizes: $\eta_{\mathsf{KL}}(P^n_{Y|X}, Q^n) = \eta_{\mathsf{KL}}(P_{Y|X}, Q)$
- $\eta_{\mathsf{KL}}(P_{Y|X}, Q) = \sup_{\substack{U-X-Y: \\ P_{XY}=Q\times P_{Y|X}}} \frac{I(U;Y)}{I(U;X)}$

Conclusion

For all $f : \{0,1\}^n \mapsto \{-1,1\}$

$$\frac{I(f(\mathbf{X});\mathbf{Y})}{I(f(\mathbf{X});\mathbf{X})} \leq \eta_{\mathsf{KL}}(\mathrm{BSC}(\alpha), \mathrm{Bernoulli}(\tfrac{1}{2})) = (1-2\alpha)^2$$

Simple Attempts: SDPI

Let

$$\eta_{\mathrm{KL}}(P_{Y|X}, Q) \triangleq \sup_{P:0<D(P||Q)<1} \frac{D(P_{Y|X} \circ P || P_{Y|X} \circ Q)}{D(P||Q)}$$

- $\eta_{\mathrm{KL}}(P_{Y|X}, Q)$ tensorizes: $\eta_{\mathrm{KL}}(P^n_{Y|X}, Q^n) = \eta_{\mathrm{KL}}(P_{Y|X}, Q)$
- $\eta_{\mathrm{KL}}(P_{Y|X}, Q) = \sup_{\substack{U-X-Y: \\ P_{XY}=Q\times P_{Y|X}}} \frac{I(U;Y)}{I(U;X)}$

Conclusion

For all $f : \{0,1\}^n \mapsto \{-1,1\}$

$$\frac{I(f(\mathbf{X});\mathbf{Y})}{I(f(\mathbf{X});\mathbf{X})} \leq \eta_{\mathrm{KL}}(\mathrm{BSC}(\alpha), \mathrm{Bernoulli}(\tfrac{1}{2})) = (1-2\alpha)^2$$

Weakness: not tight for $\mathbf{X}|\mathbf{U}$ uniform on subsets

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P^f_{\mathbf{y}} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^f = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X}) | \mathbf{Y}) = 1 - \mathbb{E} h(P_{\mathbf{Y}}^f)$$

$h(p) \triangleq -p \log_2(p) - (1-p) \log_2(1-p)$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) = 1 - \mathbb{E}h(P_{\mathbf{Y}}^{f})$$

Recall: $h(p) \geq 4p(1-p)$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P^f_{\mathbf{y}} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq 1 - \mathbb{E} 4 P^f_{\mathbf{Y}} (1 - P^f_{\mathbf{Y}})$$

Recall: $h(p) \geq 4p(1-p)$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X}) | \mathbf{Y}) \leq \mathbb{E}(1 - 2P_{\mathbf{Y}}^{f})^2$$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^f = \Pr(f(\mathbf{X}) = -1|\mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \le \mathbb{E}(1 - 2P_{\mathbf{Y}}^f)^2$$

Note: $\mathbb{E}(f(\mathbf{X})|\mathbf{Y} = \mathbf{y}) = 1 - 2P_{\mathbf{y}}^f$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}(\mathbb{E}(f(\mathbf{X})|\mathbf{Y}))^2$$

Note: $\mathbb{E}(f(\mathbf{X})|\mathbf{Y} = \mathbf{y}) = 1 - 2P_{\mathbf{y}}^{f}$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^f = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Conclusion

For all balanced $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq \max_{\substack{f:\{0,1\}^n \mapsto \{-1,1\} \\ \mathbb{E}f(\mathbf{X})=0, \mathbb{E}f^2(\mathbf{X})=1}} \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1|\mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X});\mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Conclusion

For all balanced $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X});\mathbf{Y}) \leq \max_{\substack{f:\{0,1\}^n \mapsto \mathbb{R} \\ \mathbb{E}f(\mathbf{X})=0, \mathbb{E}f^2(\mathbf{X})=1}} \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Rényi maximal correlation

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1|\mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X});\mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Conclusion

For all balanced $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X});\mathbf{Y}) \leq \max_{\substack{f:\{0,1\}^n \mapsto \mathbb{R} \\ \mathbb{E}f(\mathbf{X})=0, \mathbb{E}f^2(\mathbf{X})=1}} \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

$$= \max_{\substack{f:\{0,1\} \mapsto \mathbb{R} \\ \mathbb{E}f(X_1)=0, \mathbb{E}f^2(X_1)=1}} \mathbb{E}\left(\mathbb{E}(f(X_1)|Y_1)\right)^2$$

[Witsenhausen'75]

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Conclusion

For all balanced $f : \{0,1\}^n \mapsto \{-1,1\}$

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) &\leq \max_{\substack{f:\{0,1\}^n \mapsto \mathbb{R} \\ \mathbb{E}f(\mathbf{X})=0, \mathbb{E}f^2(\mathbf{X})=1}} \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2 \\ &= \max_{\substack{f:\{0,1\} \mapsto \mathbb{R} \\ \mathbb{E}f(X_1)=0, \mathbb{E}f^2(X_1)=1}} \mathbb{E}\left(\mathbb{E}(f(X_1)|Y_1)\right)^2 \\ &= (1-2\alpha)^2 \end{aligned}$$

Simple Attempts: Maximal Correlation

Assume f is balanced and let $P_{\mathbf{y}}^{f} = \Pr(f(\mathbf{X}) = -1 | \mathbf{Y} = \mathbf{y})$

$$I(f(\mathbf{X}); \mathbf{Y}) = H(f(\mathbf{X})) - H(f(\mathbf{X})|\mathbf{Y}) \leq \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2$$

Conclusion

For all balanced $f : \{0,1\}^n \mapsto \{-1,1\}$

$$\begin{aligned} I(f(\mathbf{X}); \mathbf{Y}) \leq & \max_{\substack{f:\{0,1\}^n \mapsto \mathbb{R} \\ \mathbb{E}f(\mathbf{X})=0, \mathbb{E}f^2(\mathbf{X})=1}} \mathbb{E}\left(\mathbb{E}(f(\mathbf{X})|\mathbf{Y})\right)^2 \\ & = \max_{\substack{f:\{0,1\} \mapsto \mathbb{R} \\ \mathbb{E}f(X_1)=0, \mathbb{E}f^2(X_1)=1}} \mathbb{E}\left(\mathbb{E}(f(X_1)|Y_1)\right)^2 \\ & = (1-2\alpha)^2 \end{aligned}$$

Weakness: the approximation $h(p) \geq 4p(1-p)$ is loose

Our Approach

We will first find tighter bounds on $h(p)$

Our Approach

Taylor expansion around $1/2$:

$$h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right)$$

Our Approach

Taylor expansion around 1/2:

$$h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) = 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k}$$

$$c_k = \frac{\log_2(e)}{2k(2k-1)}$$

Our Approach

Taylor expansion around $1/2$:

$$h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) = 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k}$$

$$\geq 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \sum_{k=t}^{\infty} c_k$$

Our Approach

Taylor expansion around 1/2:

$$h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) = 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k}$$

$$\geq 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \sum_{k=t}^{\infty} c_k$$

$$= 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \left(1 - \sum_{k=1}^{t-1} c_k\right)$$

Our Approach

Taylor expansion around 1/2:

$$
\begin{aligned}
h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) &= 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k} \\
&\geq 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \sum_{k=t}^{\infty} c_k \\
&= 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \left(1 - \sum_{k=1}^{t-1} c_k\right)
\end{aligned}
$$

for $t = 1$ this gives $h(p) \geq 4p(1-p)$
Larger t, better bounds

Our Approach

Taylor expansion around 1/2:

$$
\begin{aligned}
h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) &= 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k} \\
&\geq 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \sum_{k=t}^{\infty} c_k \\
&= 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \left(1 - \sum_{k=1}^{t-1} c_k\right)
\end{aligned}
$$

Recall: $I(f(\mathbf{X}); \mathbf{Y}) = 1 - \mathbb{E} h\left(P_{\mathbf{Y}}^f\right); \quad 1 - 2P_{\mathbf{Y}}^f = \mathbb{E}\left(f(\mathbf{X}) | \mathbf{Y}\right)$

Our Approach

Taylor expansion around 1/2:

$$h(p) = h\left(\frac{1}{2} + \frac{(1-2p)}{2}\right) = 1 - \sum_{k=1}^{\infty} c_k (1-2p)^{2k}$$

$$\geq 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \sum_{k=t}^{\infty} c_k$$

$$= 1 - \sum_{k=1}^{t-1} c_k (1-2p)^{2k} - (1-2p)^{2t} \left(1 - \sum_{k=1}^{t-1} c_k\right)$$

Recall: $I(f(\mathbf{X}); \mathbf{Y}) = 1 - \mathbb{E}h\left(P_{\mathbf{Y}}^f\right); \quad 1 - 2P_{\mathbf{Y}}^f = \mathbb{E}\left(f(\mathbf{X})|\mathbf{Y}\right)$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq \sum_{k=1}^{t} c_k \mathbb{E}\left[\mathbb{E}^{2k}\left(f(\mathbf{X})|\mathbf{Y}\right)\right] + \left(1 - \sum_{k=1}^{t} c_k\right) \mathbb{E}\left[\mathbb{E}^{2t}\left(f(\mathbf{X})|\mathbf{Y}\right)\right]$$

Our Approach

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \mathbb{E}\left[\mathbb{E}^{2k}\left(f(\mathbf{X})|\mathbf{Y}\right)\right] + \left(1 - \sum_{k=1}^{t} c_k\right) \mathbb{E}\left[\mathbb{E}^{2t}\left(f(\mathbf{X})|\mathbf{Y}\right)\right]$$

Our Approach

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \mathbb{E}\left[\mathbb{E}^{2k}\left(f(\mathbf{X})|\mathbf{Y}\right)\right] + \left(1 - \sum_{k=1}^{t} c_k\right) \mathbb{E}\left[\mathbb{E}^{2t}\left(f(\mathbf{X})|\mathbf{Y}\right)\right]$$

For channel $w(x|y)$ and function $g : \mathcal{X} \mapsto \mathbb{R}$ define

$$(T_w g)(y) \triangleq \mathbb{E}\left(g(X)|Y = y\right) = \sum_{x \in \mathcal{X}} w(x|y) g(x)$$

Our Approach

$$I(f(\mathbf{X}); \mathbf{Y}) \leq \sum_{k=1}^{t} c_k \mathbb{E}\left[\mathbb{E}^{2k}\left(f(\mathbf{X})|\mathbf{Y}\right)\right] + \left(1 - \sum_{k=1}^{t} c_k\right) \mathbb{E}\left[\mathbb{E}^{2t}\left(f(\mathbf{X})|\mathbf{Y}\right)\right]$$

For channel $w(x|y)$ and function $g : \mathcal{X} \mapsto \mathbb{R}$ define

$$(T_w g)(y) \triangleq \mathbb{E}\left(g(X)|Y = y\right) = \sum_{x \in \mathcal{X}} w(x|y) g(x)$$

and for $Y \sim Q$ we can further define

$$\|T_w g\|_p \triangleq \left(\mathbb{E}\left[\mathbb{E}^p\left(g(X)|Y\right)\right]\right)^{1/p}$$

Our Approach

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \mathbb{E}\left[\mathbb{E}^{2k}\left(f(\mathbf{X})|\mathbf{Y}\right)\right] + \left(1 - \sum_{k=1}^{t} c_k\right) \mathbb{E}\left[\mathbb{E}^{2t}\left(f(\mathbf{X})|\mathbf{Y}\right)\right]$$

For channel $w(x|y)$ and function $g : \mathcal{X} \mapsto \mathbb{R}$ define

$$(T_w g)(y) \triangleq \mathbb{E}\left(g(X)|Y = y\right) = \sum_{x \in \mathcal{X}} w(x|y) g(x)$$

and for $Y \sim Q$ we can further define

$$\|T_w g\|_p \triangleq \left(\mathbb{E}\left[\mathbb{E}^p\left(g(X)|Y\right)\right]\right)^{1/p}$$

For us: $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$, $w = \mathrm{BSC}^n(\alpha)$

Our Approach

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \| T_w f \|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \| T_w f \|_{2t}^{2t}$$

For channel $w(x|y)$ and function $g : \mathcal{X} \mapsto \mathbb{R}$ define

$$(T_w g)(y) \triangleq \mathbb{E}\left(g(X)|Y = y\right) = \sum_{x \in \mathcal{X}} w(x|y) g(x)$$

and for $Y \sim Q$ we can further define

$$\| T_w g \|_p \triangleq \left(\mathbb{E}\left[\mathbb{E}^p\left(g(X)|Y\right)\right]\right)^{1/p}$$

For us: $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$, $w = \mathrm{BSC}^n(\alpha)$

Our Approach

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \|T_w f\|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \|T_w f\|_{2t}^{2t}$$

For channel $w(x|y)$ and function $g : \mathcal{X} \mapsto \mathbb{R}$ define

$$(T_w g)(y) \triangleq \mathbb{E}\left(g(X)|Y = y\right) = \sum_{x\in\mathcal{X}} w(x|y)g(x)$$

and for $Y \sim Q$ we can further define

$$\|T_w g\|_p \triangleq \left(\mathbb{E}\left[\mathbb{E}^p\left(g(X)|Y\right)\right]\right)^{1/p}$$

For us: $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$, $w = \mathrm{BSC}^n(\alpha)$

Need to upper bound $\|T_w f\|_{2k}$, for $k = 1, \ldots, t$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf \{ r \;:\; \| T_w g \|_p \le \|g\|_r \; \forall g : \mathcal{X} \mapsto \mathbb{R} \}$$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf \{ r \ : \ \| T_w g \|_p \leq \|g\|_r \ \forall g : \mathcal{X} \mapsto \mathbb{R} \}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf\{r \ : \ \|T_w g\|_p \leq \|g\|_r \ \forall g : \mathcal{X} \mapsto \mathbb{R}\}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf\{r \;:\; \|T_w g\|_p \leq \|g\|_r \;\forall g : \mathcal{X} \mapsto \mathbb{R}\}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$, and let $1 \leq r < p < \infty$. If $(1-2\delta) \leq \sqrt{\frac{r-1}{p-1}}$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\|T_w g(\mathbf{Y})\|_p \leq \|g(\mathbf{X})\|_r$$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf\{r \;:\; \|T_w g\|_p \le \|g\|_r \;\forall g : \mathcal{X} \mapsto \mathbb{R}\}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$, and let $1 \le r < p < \infty$. If $(1 - 2\delta) \le \sqrt{\frac{r-1}{p-1}}$, then for any $g : \{0, 1\}^n \mapsto \mathbb{R}$

$$\|T_w g(\mathbf{Y})\|_p \le \|g(\mathbf{X})\|_r$$

Apply with $r = 2$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf \{ r \ : \ \| T_w g \|_p \leq \|g\|_r \ \forall g : \mathcal{X} \mapsto \mathbb{R} \}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $p \geq 2$, $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$. If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| T_w g(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

Apply with $r = 2$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf \{ r \ : \ \| T_w g \|_p \leq \|g\|_r \ \forall g : \mathcal{X} \mapsto \mathbb{R} \}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $p \geq 2$, $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$. If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| T_w g(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

Denote $T_{\mathrm{BSC}^n(\delta)} = T_\delta$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf\{r \ : \ \|T_w g\|_p \le \|g\|_r \ \forall g : \mathcal{X} \mapsto \mathbb{R}\}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $p \ge 2$, $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$. If $\delta \ge \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\|T_\delta g(\mathbf{Y})\|_p \le \|g(\mathbf{X})\|_2$$

Denote $T_{\mathrm{BSC}^n(\delta)} = T_\delta$

Hypercontractivity

Given $Y \sim Q$ and channel $w = w(x|y)$, define for $p > 1$

$$S_p(w, Q) \triangleq \inf\{r \;:\; \|T_w g\|_p \le \|g\|_r \;\forall g : \mathcal{X} \mapsto \mathbb{R}\}$$

Tensorization: $S_p(w^n, Q^n) = S_p(w, Q)$ [AG'76 and others]

$\Rightarrow S_p(w^n, Q^n)$ admits a single-letter expression

Theorem: Bonami-Beckner

Let $p \ge 2$, $\mathbf{Y} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$ and $w(\mathbf{x}|\mathbf{y}) = \mathrm{BSC}^n(\delta)$. If $\delta \ge \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\|T_\delta g(\mathbf{Y})\|_p \le \|g(\mathbf{X})\|_2$$

Quite useless for $g : \{-1,1\}^n \mapsto \{-1,1\}$ as $\|g\|_q = 1, \forall q$

Degraded Channels

- Let $w'(z|y)$ be a channel with input $\mathcal{Y}$ and output $\mathcal{Z}$
- Let $w''(x|z)$ be a channel with input $\mathcal{Z}$ and output $\mathcal{X}$
- let $w(x|y) = w'' \circ w' = \sum_{z \in \mathcal{Z}} w''(x|z)w'(z|y)$

Degraded Channels

- Let $w'(z|y)$ be a channel with input $\mathcal{Y}$ and output $\mathcal{Z}$
- Let $w''(x|z)$ be a channel with input $\mathcal{Z}$ and output $\mathcal{X}$
- let $w(x|y) = w'' \circ w' = \sum_{z \in \mathcal{Z}} w''(x|z) w'(z|y)$

Proposition

$$(T_w f)(y) = (T_{w'} g(z))(y),$$

where

$$g(z) \triangleq (T_{w''} f)(z).$$

Degraded Channels

- Let $w'(z|y)$ be a channel with input $\mathcal{Y}$ and output $\mathcal{Z}$
- Let $w''(x|z)$ be a channel with input $\mathcal{Z}$ and output $\mathcal{X}$
- let $w(x|y) = w'' \circ w' = \sum_{z \in \mathcal{Z}} w''(x|z) w'(z|y)$

Proposition

$$(T_w f)(y) = (T_{w'} g(z))(y),$$

where

$$g(z) \triangleq (T_{w''} f)(z).$$

For us: $\mathrm{BSC}(\alpha) = \mathrm{BSC}(\alpha'') \circ \mathrm{BSC}(\alpha')$ for $\alpha' * \alpha'' = \alpha$

Degraded Channels

- Let $w'(z|y)$ be a channel with input $\mathcal{Y}$ and output $\mathcal{Z}$
- Let $w''(x|z)$ be a channel with input $\mathcal{Z}$ and output $\mathcal{X}$
- let $w(x|y) = w'' \circ w' = \sum_{z \in \mathcal{Z}} w''(x|z) w'(z|y)$

Proposition

$$\left(T_\alpha f\right)(y) = \left(T_{\alpha'} g(z)\right)(y),$$

where

$$g(z) \triangleq \left(T_{\alpha''} f\right)(z).$$

For us: $\mathrm{BSC}(\alpha) = \mathrm{BSC}(\alpha'') \circ \mathrm{BSC}(\alpha')$ for $\alpha' * \alpha'' = \alpha$

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

New Bounds on $\|T_\alpha f\|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\|(T_\delta g)(\mathbf{Y})\|_p \leq \|g(\mathbf{X})\|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\| (T_\alpha f)(\mathbf{Y}) \|_p = \| (T_{\alpha'} g)(\mathbf{Y}) \|_p$$

degraded channel proposition

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\| (T_\alpha f)(\mathbf{Y}) \|_p = \| (T_{\alpha'} g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{Z}) \|_2$$

hypercontractivity

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\begin{aligned} \| (T_\alpha f)(\mathbf{Y}) \|_p &= \| (T_{\alpha'} g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{Z}) \|_2 \\ &= \| g(\mathbf{X}) \|_2 \end{aligned}$$

Z and **X** are both $\mathrm{Bernoulli}^n\left(\frac{1}{2}\right)$

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \|g(\mathbf{X})\|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\begin{aligned} \| (T_\alpha f)(\mathbf{Y}) \|_p &= \| (T_{\alpha'} g)(\mathbf{Y}) \|_p \leq \|g(\mathbf{Z})\|_2 \\ &= \|g(\mathbf{X})\|_2 = \| (T_{\alpha''} f)(\mathbf{X}) \|_2 \end{aligned}$$

definition of g

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\begin{aligned} \| (T_\alpha f)(\mathbf{Y}) \|_p &= \| (T_{\alpha'} g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{Z}) \|_2 \\ &= \| g(\mathbf{X}) \|_2 = \| (T_{\alpha''} f)(\mathbf{X}) \|_2 \leq (1 - 2\alpha'')^2 \end{aligned}$$

maximal correlation (recall: f balanced)

New Bounds on $\| T_\alpha f \|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\| (T_\delta g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{X}) \|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

$$\begin{aligned} \| (T_\alpha f)(\mathbf{Y}) \|_p &= \| (T_{\alpha'} g)(\mathbf{Y}) \|_p \leq \| g(\mathbf{Z}) \|_2 \\ &= \| g(\mathbf{X}) \|_2 = \| (T_{\alpha''} f)(\mathbf{X}) \|_2 \leq (1 - 2\alpha'')^2 \end{aligned}$$

Note: derivation only used $\mathbb{E} f(\mathbf{X}) = 0$ and $\mathbb{E} f^2(\mathbf{X}) = 1$

New Bounds on $\|T_\alpha f\|_p$, $p \geq 2$

If $\delta \geq \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$, then for any $g : \{0,1\}^n \mapsto \mathbb{R}$

$$\|(T_\delta g)(\mathbf{Y})\|_p \leq \|g(\mathbf{X})\|_2$$

- Set $\alpha' = \frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right)$; and assuming $\alpha \geq \alpha'$
- Set $\alpha'' = \frac{1}{2}\left(1 - \sqrt{p-1}(1-2\alpha)\right)$ such that $\alpha = \alpha' * \alpha''$
- Set $g(z) = (T_{\alpha''} f)(z)$ such that $(T_\alpha f)(y) = (T_{\alpha'} g)(y)$

Theorem

Let $\mathbf{X} \sim \mathrm{Bernoulli}^n(\frac{1}{2})$, $p \geq 2$ and $\frac{1}{2}\left(1 - \sqrt{\frac{1}{p-1}}\right) \leq \alpha \leq \frac{1}{2}$. For any $f : \{0,1\}^n \mapsto \mathbb{R}$ with $\mathbb{E}f(\mathbf{X}) = 0$ and $\mathbb{E}f^2(\mathbf{X}) = 1$

$$\|(T_\alpha f)(\mathbf{X})\|_p \leq (p-1)(1-2\alpha)^2.$$

Main result

We have found:

- $I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \| T_\alpha f \|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \| T_\alpha f \|_{2t}^{2t}$
- $\| (T_\alpha f)(\mathbf{X}) \|_p \leq (p-1)(1-2\alpha)^2$ for $p \geq 2$

Main result

We have found:

- $I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \| T_\alpha f \|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \| T_\alpha f \|_{2t}^{2t}$
- $\| (T_\alpha f)(\mathbf{X}) \|_p \leq (p-1)(1-2\alpha)^2$ for $p \geq 2$

Theorem

For any balanced function $f : \{0,1\}^n \mapsto \{-1,1\}$,any integer $t \geq 1$ and any $\frac{1}{2}\left(1 - \frac{1}{\sqrt{2t-1}}\right) \leq \alpha \leq \frac{1}{2}$, we have that

$$I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t-1} \frac{\log_2(e)}{2k(2k-1)} (2k-1)^k (1-2\alpha)^{2k} + \left(1 - \sum_{k=1}^{t-1} \frac{\log_2(e)}{2k(2k-1)}\right) (2t-1)^t (1-2\alpha)^{2t}.$$

Main result

We have found:

- $I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \| T_\alpha f \|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \| T_\alpha f \|_{2t}^{2t}$
- $\| (T_\alpha f)(\mathbf{X}) \|_p \leq (p-1)(1-2\alpha)^2$ for $p \geq 2$

Theorem

For any balanced function $f : \{0,1\}^n \mapsto \{-1,1\}$ and any $0 \leq \alpha \leq \frac{1}{2}$, we have that

$$I(f(\mathbf{X});\mathbf{Y}) \leq (1-2\alpha)^2.$$

$$t = 1$$

Main result

We have found:

- $I(f(\mathbf{X});\mathbf{Y}) \leq \sum_{k=1}^{t} c_k \| T_\alpha f \|_{2k}^{2k} + \left(1 - \sum_{k=1}^{t} c_k\right) \| T_\alpha f \|_{2t}^{2t}$
- $\| (T_\alpha f)(\mathbf{X}) \|_p \leq (p-1)(1-2\alpha)^2$ for $p \geq 2$

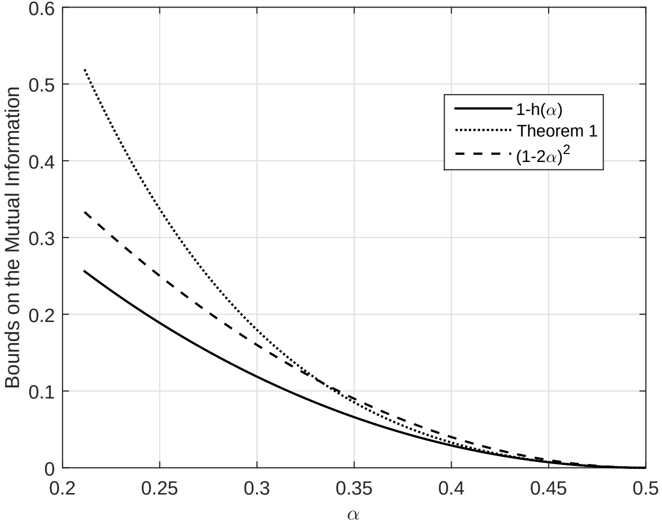
Theorem

For any balanced function $f : \{0,1\}^n \mapsto \{-1,1\}$ and any $\frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right) \leq \alpha \leq \frac{1}{2}$, we have that

$$I(f(\mathbf{X});\mathbf{Y}) \leq \frac{\log_2(e)}{2}(1-2\alpha)^2 + 9\left(1 - \frac{\log_2(e)}{2}\right)(1-2\alpha)^4.$$

$$t = 2$$

Main result



0.6
0.5
0.4
0.3
0.2
0.1
0
Bounds on the Mutual Information
0.2
0.25
0.3
0.35
0.4
0.45
0.5
α
1-h(α)
Theorem 1
(1-2α)^2

Properties of the Bound

Our bound (with $t = 2$) has the optimal slope at $\alpha = 1/2$ and approaches $1 - h(\alpha)$ from above

Properties of the Bound

Our bound (with $t = 2$) has the optimal slope at $\alpha = 1/2$ and approaches $1 - h(\alpha)$ from above

For balanced boolean functions $\| T_\alpha f \|_2^2 = (1 - 2\alpha)^2$ for dictatorship and $\| T_\alpha f \|_2^2 \leq (1 - 2\alpha)^2 (1 - c_{n,\alpha})$ otherwise

Properties of the Bound

Our bound (with $t = 2$) has the optimal slope at $\alpha = 1/2$ and approaches $1 - h(\alpha)$ from above

For balanced boolean functions $\|T_\alpha f\|_2^2 = (1-2\alpha)^2$ for dictatorship and $\|T_\alpha f\|_2^2 \leq (1-2\alpha)^2(1-c_{n,\alpha})$ otherwise

For α close enough to $1/2$ our bound is less than $1 - h(\alpha)$ for any function that is not dictatorship

Corollary

For all $\frac{1}{2}(1-2^{-(n+2)}) \leq \alpha \leq \frac{1}{2}$ and all balanced functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha) \qquad (*)$$

Properties of the Bound

Our bound (with $t = 2$) has the optimal slope at $\alpha = 1/2$ and approaches $1 - h(\alpha)$ from above

For balanced boolean functions $\| T_\alpha f \|_2^2 = (1 - 2\alpha)^2$ for dictatorship and $\| T_\alpha f \|_2^2 \leq (1 - 2\alpha)^2 (1 - c_{n,\alpha})$ otherwise

For α close enough to $1/2$ our bound is less than $1 - h(\alpha)$ for any function that is not dictatorship

Corollary

For all $\frac{1}{2}(1 - 2^{-(n+2)}) \leq \alpha \leq \frac{1}{2}$ and all balanced functions $f : \{0,1\}^n \mapsto \{-1,1\}$

$$I(f(\mathbf{X}); \mathbf{Y}) \leq 1 - h(\alpha) \qquad (*)$$

It is now known that (*) holds for all boolean functions and $\frac{1}{2}(1 - \delta) < \alpha \leq \frac{1}{2}$ ($\delta > 0$ and indep. of n) [Samorodnitsky'15]

Weakness of Moments-Based Bounding Technique

Our bound is based on bounding $\|T_\alpha f\|_{2k}$ for all $k = 1, 2, \ldots$

Weakness of Moments-Based Bounding Technique

Our bound is based on bounding $\|T_\alpha f\|_{2k}$ for all $k = 1, 2, \ldots$

In particular, we upper bounded

$$\rho(2k, \alpha) \triangleq \sup_n \max_{f:\{0,1\}^n \mapsto \{-1,1\}} \|T_\alpha f\|_{2k}$$

Weakness of Moments-Based Bounding Technique

Our bound is based on bounding $\|T_\alpha f\|_{2k}$ for all $k = 1, 2, \ldots$

In particular, we upper bounded

$$\rho(2k, \alpha) \triangleq \sup_n \max_{f:\{0,1\}^n \mapsto \{-1,1\}} \|T_\alpha f\|_{2k}$$

To prove the conjecture using this approach, we must show that $\rho(2k, \alpha) = (1-2\alpha)^{2k}$
In other words, that for any k dictatorship maximizes $\|T_\alpha f\|_{2k}$ provided that n is large enough

Weakness of Moments-Based Bounding Technique

Our bound is based on bounding $\|T_\alpha f\|_{2k}$ for all $k = 1, 2, \ldots$

In particular, we upper bounded

$$\rho(2k, \alpha) \triangleq \sup_n \max_{f:\{0,1\}^n \mapsto \{-1,1\}} \|T_\alpha f\|_{2k}$$

To prove the conjecture using this approach, we must show that $\rho(2k, \alpha) = (1 - 2\alpha)^{2k}$
In other words, that for any k dictatorship maximizes $\|T_\alpha f\|_{2k}$ provided that n is large enough

It is easy to verify that for fixed n and $k \to \infty$ we can find g (e.g., majority) with $\|T_\alpha g\|_{2k} > (1 - 2\alpha)^{2k}$

Weakness of Moments-Based Bounding Technique

Our bound is based on bounding $\| T_\alpha f \|_{2k}$ for all $k = 1, 2, \ldots$

In particular, we upper bounded

$$\rho(2k, \alpha) \triangleq \sup_n \max_{f:\{0,1\}^n \mapsto \{-1,1\}} \| T_\alpha f \|_{2k}$$

To prove the conjecture using this approach, we must show that $\rho(2k, \alpha) = (1 - 2\alpha)^{2k}$
In other words, that for any k dictatorship maximizes $\| T_\alpha f \|_{2k}$ provided that n is large enough

It is easy to verify that for fixed n and $k \to \infty$ we can find g (e.g., majority) with $\| T_\alpha g \|_{2k} > (1 - 2\alpha)^{2k}$

Less plausible to believe that $\rho(2k, \alpha) = (1 - 2\alpha)^{2k}$ for all k...

Summary

- We have studied the most informative boolean function conjecture
- Derived a new upper bound on $I(f(\mathbf{X}); \mathbf{Y}))$ for balanced f
- Bound becomes tight as channel becomes noisier
- Main ingredient was to bound high moments of $T_\alpha f$
- This was done by
  - Markov operator for degraded channels
  - hypercontractivity
  - maximal correlation