Bounds on Information Combining With Quantum Side Information

QIP 2018, Delft

Christoph Hirche

Joint work with David Reeb, Leibniz University Hannover

Física Teòrica: Informació i Fenòmens Quàntics,
Universitat Autònoma de Barcelona,
Spain

Information Combining

Basic question

Given random variables X_1 and X_2:

- What do we know about $X_1 + X_2$ and in particular $H(X_1 + X_2)$?

Information Combining

Basic question

Given random variables X_1 and X_2:

- What do we know about $X_1 + X_2$ and in particular $H(X_1 + X_2)$?

More involved

Given random variables with side information (X_1, Y_1) and (X_2, Y_2):

- What about $X_1 + X_2$ given $Y_1 Y_2$ and in particular $H(X_1 + X_2 | Y_1 Y_2)$?

Information Combining

Basic question

Given random variables X_1 and X_2:

- What do we know about $X_1 + X_2$ and in particular $H(X_1 + X_2)$?

More involved

Given random variables with side information (X_1, Y_1) and (X_2, Y_2):

- What about $X_1 + X_2$ given $Y_1 Y_2$ and in particular $H(X_1 + X_2 | Y_1 Y_2)$?

Best known instance: Entropy power inequalities

$$e^{2H(X_1+X_2)} \geq e^{2H(X_1)} + e^{2H(X_2)}$$

Adding quantum side information is very difficult!

Information Combining

Basic question

Given random variables X_1 and X_2:

- What do we know about $X_1 + X_2$ and in particular $H(X_1 + X_2)$?

More involved

Given random variables with side information (X_1, Y_1) and (X_2, Y_2):

- What about $X_1 + X_2$ given $Y_1 Y_2$ and in particular $H(X_1 + X_2 | Y_1 Y_2)$?

Best known instance: Entropy power inequalities

$$e^{2H(X_1+X_2)} \geq e^{2H(X_1)} + e^{2H(X_2)}$$

Adding quantum side information is very difficult!

Here simplest setting: binary random variables.

Outline

Outline

Without conditioning

$$X_1 \sim \begin{bmatrix} p \\ 1-p \end{bmatrix}, \quad X_2 \sim \begin{bmatrix} q \\ 1-q \end{bmatrix}$$

$$\Downarrow$$

$$X_1 + X_2 \sim \begin{bmatrix} pq + (1-p)(1-q) \\ p(1-q) + q(1-p) \end{bmatrix} \equiv \begin{bmatrix} p \star q \\ 1 - p \star q \end{bmatrix}$$

Therefore

$$H(X_1 + X_2) = h(h^{-1}(H(X_1)) \star h^{-1}(H(X_2))$$

Without conditioning

$$X_1 \sim \begin{bmatrix} p \\ 1-p \end{bmatrix}, \quad X_2 \sim \begin{bmatrix} q \\ 1-q \end{bmatrix}$$

$$\Downarrow$$

$$X_1 + X_2 \sim \begin{bmatrix} pq + (1-p)(1-q) \\ p(1-q) + q(1-p) \end{bmatrix} \equiv \begin{bmatrix} p \star q \\ 1 - p \star q \end{bmatrix}$$

Therefore

$$H(X_1 + X_2) = h(h^{-1}(H(X_1)) \star h^{-1}(H(X_2)))$$

With conditioning

Classical **bounds on information combining**.
Write $H(X_i|Y_i) = H_i$,

$$h(h^{-1}(H_1) \star h^{-1}(H_2)) \leq H(X_1 + X_2 | Y_1 Y_2) \leq \log 2 - \frac{(\log 2 - H_1)(\log 2 - H_2)}{\log 2}$$

With conditioning

Classical **bounds on information combining**.
Write $H(X_i|Y_i) = H_i$,

$$h(h^{-1}(H_1)\star h^{-1}(H_2)) \leq H(X_1+X_2|Y_1Y_2) \leq \log 2 - \frac{(\log 2 - H_1)(\log 2 - H_2)}{\log 2}$$

With $H_1 = H_2 = H$,

$$\begin{aligned} 0.799\,\frac{H(\log 2 - H)}{\log 2} &\leq h(h^{-1}(H) \star h^{-1}(H)) - H \\ &\leq H(X_1 + X_2|Y_1Y_2) - H \\ &\leq \frac{H(\log 2 - H)}{\log 2} \end{aligned}$$

Lower bound

Main ingredient

$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2|Y_1 Y_2) \\
&= \sum_{y_1, y_2} p_{Y_1=y_1} p_{Y_2=y_2} H(X_1 + X_2|Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1, y_2} p_{Y_1=y_1} p_{Y_2=y_2} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1=y_1} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2))) \\
&\geq h(h^{-1}(H(X_1|Y_1)) \star h^{-1}(H(X_2|Y_2))).
\end{aligned}$$

Lower bound

Main ingredient

$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2 | Y_1 Y_2) \\
&= \sum_{y_1, y_2} p_{Y_1=y_1} p_{Y_2=y_2} H(X_1 + X_2 | Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1, y_2} p_{Y_1=y_1} p_{Y_2=y_2} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1=y_1} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2))) \\
&\geq h(h^{-1}(H(X_1 | Y_1)) \star h^{-1}(H(X_2 | Y_2))).
\end{aligned}$$

Lower bound

Main ingredient

$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2 | Y_1 Y_2) \\
&= \sum_{y_1, y_2} p_{Y_1 = y_1} p_{Y_2 = y_2} H(X_1 + X_2 | Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1, y_2} p_{Y_1 = y_1} p_{Y_2 = y_2} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1 = y_1} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2))) \\
&\geq h(h^{-1}(H(X_1 | Y_1)) \star h^{-1}(H(X_2 | Y_2))).
\end{aligned}$$

Lower bound

Main ingredient

$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2 | Y_1 Y_2) \\
&= \sum_{y_1, y_2} p_{Y_1 = y_1} p_{Y_2 = y_2} H(X_1 + X_2 | Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1, y_2} p_{Y_1 = y_1} p_{Y_2 = y_2} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1 = y_1} h(h^{-1}(H(X_1 | Y_1 = y_1)) \star h^{-1}(H(X_2 | Y_2))) \\
&\geq h(h^{-1}(H(X_1 | Y_1)) \star h^{-1}(H(X_2 | Y_2))).
\end{aligned}$$

Lower bound

Main ingredient

$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2|Y_1 Y_2) \\
&= \sum_{y_1,y_2} p_{Y_1=y_1} p_{Y_2=y_2} H(X_1 + X_2|Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1,y_2} p_{Y_1=y_1} p_{Y_2=y_2} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1=y_1} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2))) \\
&\geq h(h^{-1}(H(X_1|Y_1)) \star h^{-1}(H(X_2|Y_2))).
\end{aligned}$$

Lower bound

Main ingredient

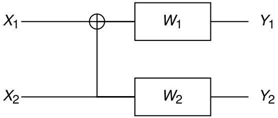
$$g_c(H_1, H_2) := h(h^{-1}(H_1) * h^{-1}(H_2))$$

is convex in H_1 for fixed H_2, and vice versa.

$$\begin{aligned}
&H(X_1 + X_2|Y_1 Y_2) \\
&= \sum_{y_1,y_2} p_{Y_1=y_1} p_{Y_2=y_2} H(X_1 + X_2|Y_1 = y_1 Y_2 = y_2) \\
&= \sum_{y_1,y_2} p_{Y_1=y_1} p_{Y_2=y_2} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2 = y_2))) \\
&\geq \sum_{y_1} p_{Y_1=y_1} h(h^{-1}(H(X_1|Y_1 = y_1)) \star h^{-1}(H(X_2|Y_2))) \\
&\geq h(h^{-1}(H(X_1|Y_1)) \star h^{-1}(H(X_2|Y_2))).
\end{aligned}$$

Channel picture

$W_1 \boxast W_2$



Notation:

$$H(X_i|Y_i) = H(W_i)$$
$$H(X_1 + X_2|Y_1 Y_2) = H(W_1 \boxast W_2)$$

Outline

Channel Duality

For every channel W we can define a dual channel $W^{\perp}$.

Additional uncertainty relation

$$H(W) = \log 2 - H(W^{\perp})$$

and symmetry relation

$$\begin{aligned} &H(W_1 \circledast W_2) - (H(W_1) + H(W_2))/2 \\ =&H(W_1^{\perp} \circledast W_2^{\perp}) - \left(H(W_1^{\perp}) + H(W_2^{\perp})\right)/2. \end{aligned}$$

Based on results in Renes, IEEE Trans. 64, 577 (2018), arXiv:1701.05583

Remember last talk!

Channel Duality

For every channel W we can define a dual channel $W^{\perp}$.
Additional uncertainty relation

$$H(W) = \log 2 - H(W^{\perp})$$

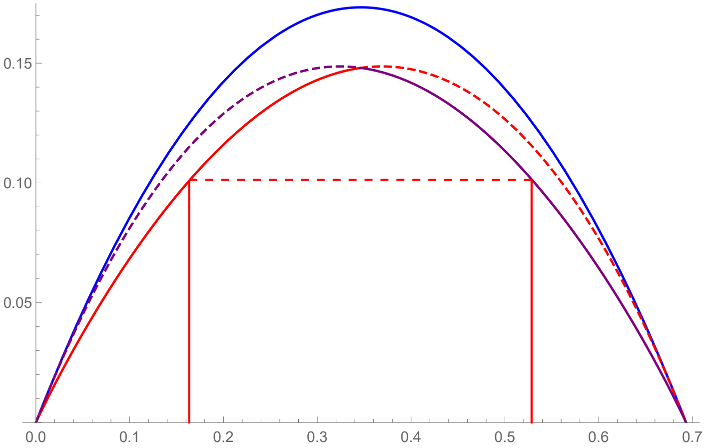
and symmetry relation

$$\begin{aligned} &H(W_1 \circledast W_2) - \left(H(W_1) + H(W_2)\right)/2 \\ =&H(W_1^{\perp} \circledast W_2^{\perp}) - \left(H(W_1^{\perp}) + H(W_2^{\perp})\right)/2. \end{aligned}$$

Remember last talk!

Based on results in Renes, IEEE Trans. 64, 577 (2018), arXiv:1701.05583

Channel Duality



CQMI

Conditional Quantum Mutual Information

$$H(X_1 + X_2|B_1B_2) - H_1 = I(X_1 + X_2 : X_2|B_1B_2)$$

Lower bounds on CQMI

$$I(A : C|B)_\tau \geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB}))$$

CQMI

Conditional Quantum Mutual Information

$$H(X_1 + X_2|B_1B_2) - H_1 = I(X_1 + X_2 : X_2|B_1B_2)$$

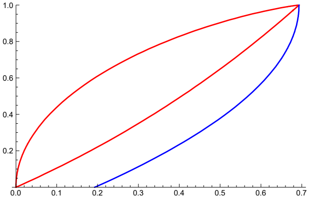
Lower bounds on CQMI

$$I(A : C|B)_\tau \geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB}))$$

Concavity of von Neumann Entropy

Let $\rho_i \in \mathcal{B}(\mathbb{C}^d)$ and $\{p_i\}_{i=1}^n$ be a probability distribution.

$$H\left(\sum_{i=1}^n p_i\rho_i\right) - \sum_{i=1}^n p_i H(\rho_i)$$
$$\geq H(\{p_i\}) - \log\left(1 + 2\sum_{1\leq i<j\leq n}\sqrt{p_i p_j}F(\rho_i,\rho_j)\right).$$



Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$H(X_1 + X_2|B_1B_2) - H$$

$$= I(A:C|B)_\tau \qquad \textit{QCMI}$$

$$\geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) \qquad \textit{Fawzi} - \textit{Renner}$$

$$\geq -2\log\cos\left[\frac{1}{2}\arccos[f^2] - \frac{1}{2}\arccos f\right] \qquad \Delta - \textit{ineq.}$$

$$\geq -2\log\cos\left[\frac{1}{2}\arccos[(1-2h_2^{-1}(\log 2 - H))^2] - \frac{1}{2}\arccos[1-2h_2^{-1}(\log 2 - H)]\right] \qquad \textit{Concavity}$$

$$\Rightarrow \begin{cases} 0.083\cdot\frac{H}{1-\log H}, & H \leq \frac{1}{2}\log 2 \\ 0.083\cdot\frac{\log 2 - H}{1-\log(\log 2 - H)}, & H > \frac{1}{2}\log 2. \end{cases} \qquad \textit{Duality/Simplify}$$

with $f := F(\rho_0, \rho_1)$.

Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$
\begin{aligned}
& H(X_1 + X_2|B_1B_2) - H \\
& = I(A:C|B)_\tau && \textit{QCMI} \\
& \geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) && \textit{Fawzi} - \textit{Renner} \\
& \geq -2\log\cos\left[\frac{1}{2}\arccos[f^2] - \frac{1}{2}\arccos f\right] && \Delta - \textit{ineq.} \\
& \geq -2\log\cos\Big[\frac{1}{2}\arccos[(1-2h_2^{-1}(\log 2 - H))^2] && \textit{Concavity} \\
& \qquad - \frac{1}{2}\arccos[1-2h_2^{-1}(\log 2 - H)]\Big] \\
& \Rightarrow \begin{cases} 0.083\cdot\frac{H}{1-\log H}, & H \leq \frac{1}{2}\log 2 \\ 0.083\cdot\frac{\log 2 - H}{1-\log(\log 2 - H)}, & H > \frac{1}{2}\log 2. \end{cases} && \textit{Duality/Simplify}
\end{aligned}
$$

with $f := F(\rho_0, \rho_1)$.

Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$\begin{aligned}
&H(X_1 + X_2|B_1B_2) - H \\
&= I(A:C|B)_\tau && \textit{QCMI} \\
&\geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) && \textit{Fawzi} - \textit{Renner} \\
&\geq -2\log\cos\left[\frac{1}{2}\arccos[f^2] - \frac{1}{2}\arccos f\right] && \Delta - \textit{ineq.} \\
&\geq -2\log\cos\Big[\frac{1}{2}\arccos[(1-2h_2^{-1}(\log 2 - H))^2] && \textit{Concavity} \\
&\qquad - \frac{1}{2}\arccos[1-2h_2^{-1}(\log 2 - H)]\Big] \\
&\Rightarrow \begin{cases} 0.083 \cdot \frac{H}{1-\log H}, & H \leq \frac{1}{2}\log 2 \\ 0.083 \cdot \frac{\log 2 - H}{1-\log(\log 2 - H)}, & H > \frac{1}{2}\log 2. \end{cases} && \textit{Duality/Simplify}
\end{aligned}$$

with $f := F(\rho_0, \rho_1)$.

Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$\begin{aligned}
&H(X_1 + X_2|B_1B_2) - H \\
&= I(A:C|B)_\tau && \textit{QCMI} \\
&\geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) && \textit{Fawzi} - \textit{Renner} \\
&\geq -2\log\cos\left[\frac{1}{2}\arccos[f^2] - \frac{1}{2}\arccos f\right] && \Delta - \textit{ineq.} \\
&\geq -2\log\cos\Big[\frac{1}{2}\arccos[(1-2h_2^{-1}(\log 2 - H))^2] && \textit{Concavity} \\
&\qquad - \frac{1}{2}\arccos[1-2h_2^{-1}(\log 2 - H)]\Big] \\
&\Rightarrow \begin{cases} 0.083\cdot\frac{H}{1-\log H}, & H \leq \frac{1}{2}\log 2 \\ 0.083\cdot\frac{\log 2 - H}{1-\log(\log 2 - H)}, & H > \frac{1}{2}\log 2. \end{cases} && \textit{Duality/Simplify}
\end{aligned}$$

with $f := F(\rho_0, \rho_1)$.

Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$
\begin{aligned}
&H(X_1 + X_2|B_1B_2) - H \\
&= I(A:C|B)_\tau && \textit{QCMI} \\
&\geq -2\log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) && \textit{Fawzi} - \textit{Renner} \\
&\geq -2\log\cos\left[\frac{1}{2}\arccos[f^2] - \frac{1}{2}\arccos f\right] && \Delta - \textit{ineq.} \\
&\geq -2\log\cos\Big[\frac{1}{2}\arccos[(1-2h_2^{-1}(\log 2 - H))^2] && \textit{Concavity} \\
&\qquad - \frac{1}{2}\arccos[1-2h_2^{-1}(\log 2 - H)]\Big] \\
&\Rightarrow \begin{cases} 0.083\cdot\frac{H}{1-\log H}, & H \leq \frac{1}{2}\log 2 \\ 0.083\cdot\frac{\log 2 - H}{1-\log(\log 2 - H)}, & H > \frac{1}{2}\log 2. \end{cases} && \textit{Duality}/\textit{Simplify}
\end{aligned}
$$

with $f := F(\rho_0, \rho_1)$.

Lower Bound: Summary

Here for the simple $H_1 = H_2 = H$ case:

$$
\begin{aligned}
& H(X_1 + X_2 | B_1 B_2) - H \\
& = I(A : C|B)_\tau && \textit{QCMI} \\
& \geq -2 \log F(\tau_{ACB}, \mathcal{R}'_{B\to AB}(\tau_{CB})) && \textit{Fawzi} - \textit{Renner} \\
& \geq -2 \log \cos \left[\frac{1}{2} \arccos[f^2] - \frac{1}{2} \arccos f \right] && \Delta - \textit{ineq.} \\
& \geq -2 \log \cos \Big[\frac{1}{2} \arccos[(1 - 2h_2^{-1}(\log 2 - H))^2] && \textit{Concavity} \\
& \qquad - \frac{1}{2} \arccos[1 - 2h_2^{-1}(\log 2 - H)] \Big] \\
& \Rightarrow \begin{cases} 0.083 \cdot \frac{H}{1 - \log H}, & H \leq \frac{1}{2} \log 2 \\ 0.083 \cdot \frac{\log 2 - H}{1 - \log(\log 2 - H)}, & H > \frac{1}{2} \log 2. \end{cases} && \textit{Duality/Simplify}
\end{aligned}
$$

with $f := F(\rho_0, \rho_1)$.

Outline

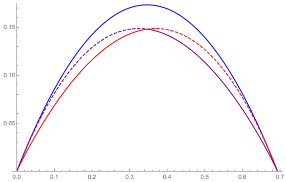
Conjectured Bounds

Let $\rho^{X_1B_1}$ and $\rho^{X_2B_2}$ be cq.-states with H_1 and H_2. Then:

$$H(X_1 + X_2|B_1B_2) - (H_1 + H_2)$$
$$\geq \begin{cases} h(h^{-1}(H_1) * h^{-1}(H_2)) - (H_1 + H_2) & H_1 + H_2 \leq \log 2 \\ h(h^{-1}(\log 2 - H_1) * h^{-1}(\log 2 - H_2)) - \log 2 & H_1 + H_2 \geq \log 2 \end{cases}$$

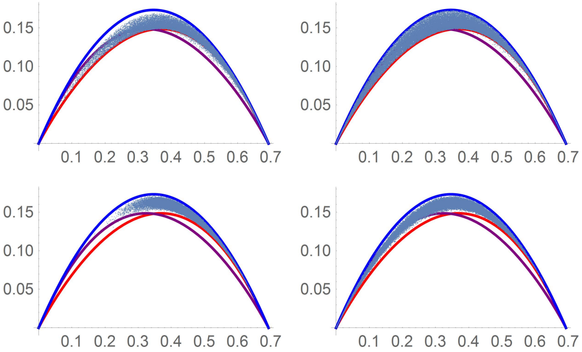
and

$$H(X_1 + X_2|B_1B_2) \leq \log 2 - \frac{(\log 2 - H_1)(\log 2 - H_2)}{\log 2}.$$



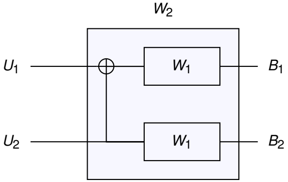
Evidence

- States with Equality.
- Numerics:

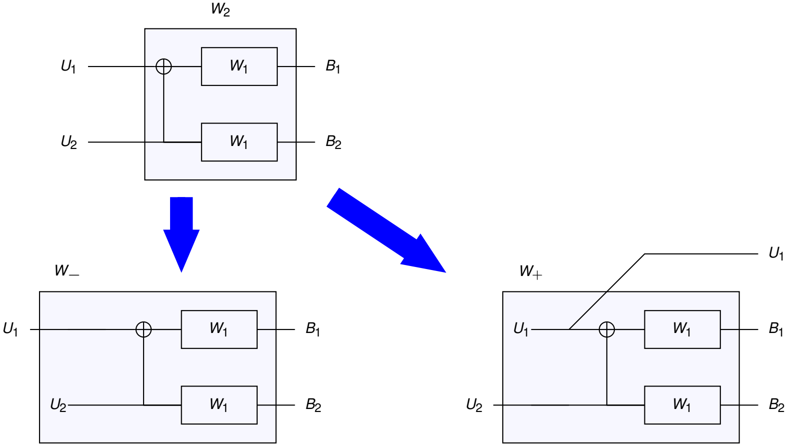


Outline

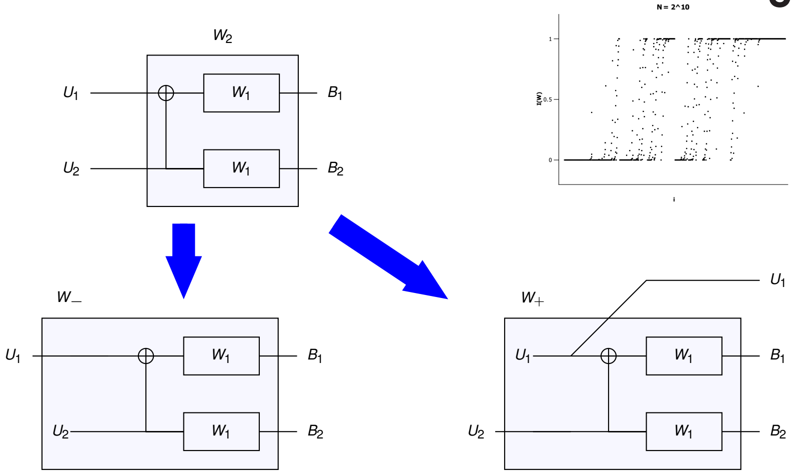
What are polar codes?



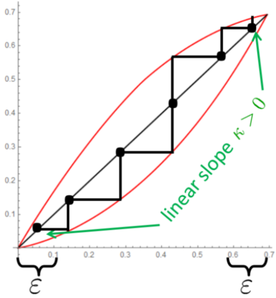
What are polar codes?

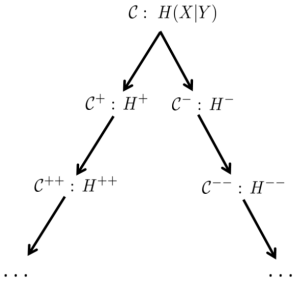


What are polar codes?



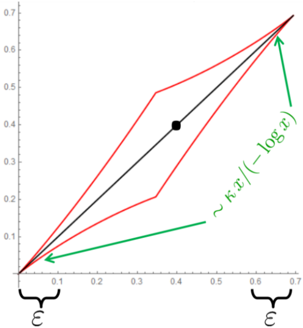
Speed of Polarization

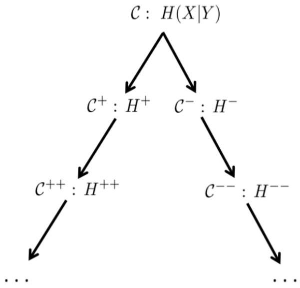




$\#$ steps to reach $[0, \epsilon] \cup [\log 2 - \epsilon, \log 2]$: $n \approx \frac{1}{\kappa} \log \frac{1}{\epsilon}$
$\Rightarrow$ Rate $R = I(W) - \epsilon$ with polynomial blocklength $\approx poly(1/\epsilon)$.

Speed of Polarization





$\#$ steps to reach $[0,\epsilon] \cup [\log 2 - \epsilon, \log 2]$: $n \approx \frac{1}{\kappa}(\log \frac{1}{\epsilon})^2$

$\Rightarrow$ Rate $R = I(W) - \epsilon$ with subexponential blocklength $\approx (1/\epsilon)^{\log(1/\epsilon)}$.

Non-stationary channels

Also:

Bounds for $H_1 \neq H_2$ give

- a conceptually simple proof of polarization (without martingales),
- that also works for non-stationary channels.

Outline

Wrap-up

- Lower bound on $H(X_1 + X_2|B_1 B_2)$.
- Conjectures for optimal lower and upper bounds.
- Applications to Polar codes.

To do!

- Prove conjectures!
  - Especially linear behavior!
- Extensions to other input alphabets.
- Much more!

Wrap-up

- Lower bound on $H(X_1 + X_2|B_1 B_2)$.
- Conjectures for optimal lower and upper bounds.
- Applications to Polar codes.

To do!

- Prove conjectures!
  - Especially linear behavior!
- Extensions to other input alphabets.
- Much more!

Recovery bounds