Tight bounds for Communication assisted agreement distillation

Jaikumar Radhakrishnan

Tata Institute of Fundamental Research, Mumbai

Joint work with Venkat Guruswami, Carnegie Mellon University

Agreement distillation

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) \in \{0,1\}^k$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y) \in \{0,1\}^k$

$(X, Y) \sim \mathrm{BSC}(\varepsilon)$: $\Pr[X_i \neq Y_i] = \varepsilon$

Goal

- $f_A(X)$ uniformly distributed in $\{0,1\}^k$
- $\Pr[f_A(X) = f_B(Y)]$ close to 1

Agreement distillation

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) \in \{0,1\}^k$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y) \in \{0,1\}^k$

$$(X, Y) \sim \text{BSC}(\varepsilon)\text{: } \Pr[X_i \neq Y_i] = \varepsilon$$

Goal

- $f_A(X)$ uniformly distributed in $\{0,1\}^k$
- $\Pr[f_A(X) = f_B(Y)]$ close to 1

Naive protocol: no communication

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) = X_1 X_2 \ldots X_k$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y) = Y_1 Y_2 \ldots Y_k$

Success probability

$$\Pr[f_A(X) = f_B(Y)] = (1-\varepsilon)^k \approx \exp(-\varepsilon k)$$

Naive protocol: no communication

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) = X_1 X_2 \dots X_k$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y) = Y_1 Y_2 \dots Y_k$

Success probability

$$\Pr[f_A(X) = f_B(Y)] = (1-\varepsilon)^k \approx \exp(-\varepsilon k)$$

Can we do better?

Yes, a little better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability least $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

But, no better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability at most $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

Can we do better?

Yes, a little better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability least $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

But, no better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability at most $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

Can we do better?

Yes, a little better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability least $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

But, no better (Bogdanov & Mossel 2011)

Alice and Bob can agree with probability at most $\approx 2^{-(\varepsilon/(1-\varepsilon))k}$.

How much can communication help?

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) = \{0,1\}^k$

$$\underset{M}{\Longrightarrow}$$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y, M) \in \{0,1\}^k$

- How many bits must Alice send Bob to ensure agreement with constant probability?
- What is the trade-off between communicaton and probability of agreement?

How much can communication help?

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) = \{0,1\}^k$

$$\underset{M}{\Longrightarrow}$$

Bob

Input: $Y \in \{0,1\}^N$

Output: $f_B(Y, M) \in \{0,1\}^k$

- How many bits must Alice send Bob to ensure agreement with constant probability?
- What is the trade-off between communicaton and probability of agreement?

The trade-off

Definition

$C^{\mathrm{BSC}(\varepsilon)}(k, \eta)$ is the minimum number of bits Alice transmits to Bob in a protocol where

- $g_A(X)$ is uniformly distributed in $\{0, 1\}^k$
- $\Pr[g_A(X) = g_B(Y, M)] \geq \eta$

$$\text{Probability of agreement} = 2^{-\gamma k}$$
$$\text{Communication} = ck$$

BM '10: If $c = 0$, then $\gamma = \varepsilon/(1 - \varepsilon)$

This work: If $c = 4\varepsilon(1 - \varepsilon)$, then $\gamma \to 0$.

The trade-off

Definition

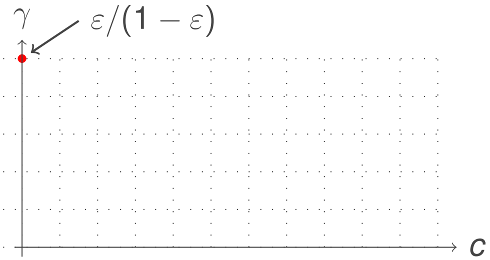
$C^{\text{BSC}(\varepsilon)}(k, \eta)$ is the minimum number of bits Alice transmits to Bob in a protocol where

- $g_A(X)$ is uniformly distributed in $\{0, 1\}^k$
- $\Pr[g_A(X) = g_B(Y, M)] \geq \eta$

$$\text{Probability of agreement} = 2^{-\gamma k}$$
$$\text{Communication} = ck$$

BM '10: If $c = 0$, then $\gamma = \varepsilon/(1-\varepsilon)$

This work: If $c = 4\varepsilon(1-\varepsilon)$, then $\gamma \to 0$.



The trade-off

Definition

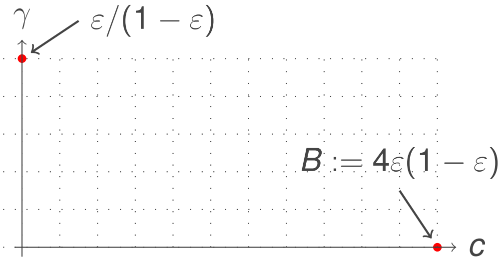
$C^{\mathrm{BSC}(\varepsilon)}(k, \eta)$ is the minimum number of bits Alice transmits to Bob in a protocol where

- $g_A(X)$ is uniformly distributed in $\{0, 1\}^k$
- $\Pr[g_A(X) = g_B(Y, M)] \geq \eta$

$$\text{Probability of agreement} = 2^{-\gamma k}$$
$$\text{Communication} = ck$$

BM '10: If $c = 0$, then $\gamma = \varepsilon/(1-\varepsilon)$

This work: If $c = 4\varepsilon(1-\varepsilon)$, then $\gamma \to 0$.



The trade-off

Definition

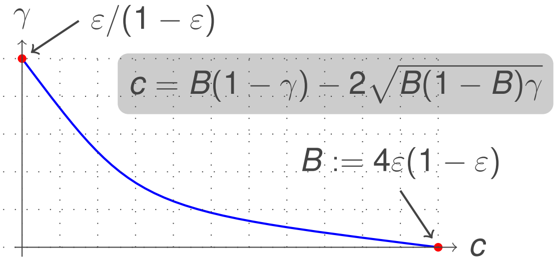
$C^{\mathrm{BSC}(\varepsilon)}(k, \eta)$ is the minimum number of bits Alice transmits to Bob in a protocol where

- $g_A(X)$ is uniformly distributed in $\{0,1\}^k$
- $\Pr[g_A(X) = g_B(Y, M)] \geq \eta$

$$\text{Probability of agreement} = 2^{-\gamma k}$$
$$\text{Communication} = ck$$

BM '10: If $c = 0$, then $\gamma = \varepsilon/(1-\varepsilon)$

This work: If $c = 4\varepsilon(1-\varepsilon)$, then $\gamma \to 0$.



Related work

Communication complexity: Canonne, Guruswami, Meka, and Sudan (2015) used capacity-achieving codes to ensure agreement with high probability with $(h(\varepsilon) + o(1))k$ bits of communication.

Information theory: The case $k = 1$ is the subject of a recent conjecture of Courtade and Kumar (2014):

The function $g_A(X) = X_1$ maximizes $I[g_A(X) : Y]$.

Chandar and Tchamkerten (2014) showed that the corresponding conjecture is false for large k.

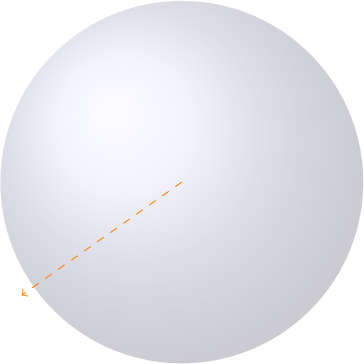
The protocol of Bogdanov and Mossel

- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

The projections along the various directions are Gaussian and approximately independent.

The protocol of Bogdanov and Mossel

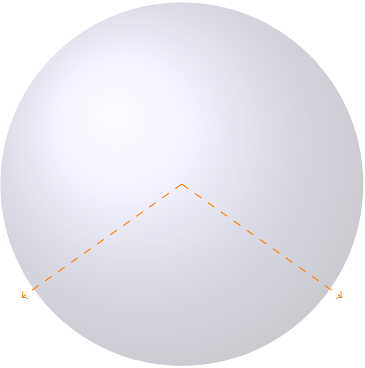


- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

The projections along the various directions are Gaussian and approximately independent.

The protocol of Bogdanov and Mossel

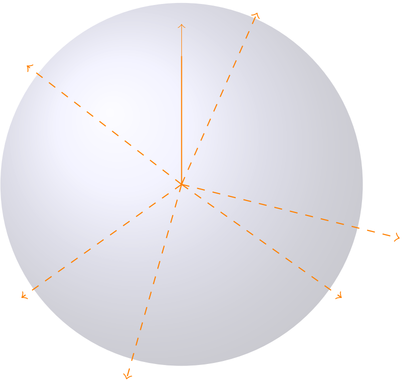


- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

The projections along the various directions are Gaussian and approximately independent.

The protocol of Bogdanov and Mossel

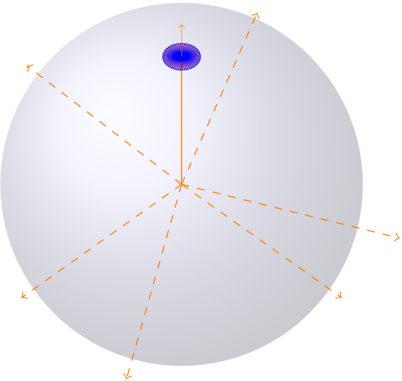


- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

The projections along the various directions are Gaussian and approximately independent.

The protocol of Bogdanov and Mossel

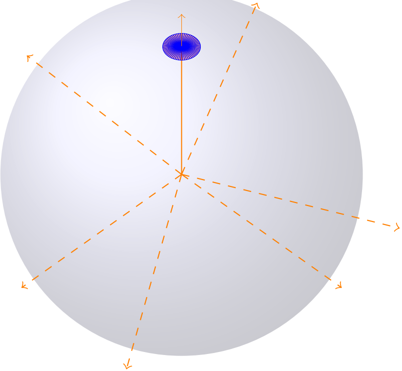


- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

The projections along the various directions are Gaussian and approximately independent.

The protocol of Bogdanov and Mossel

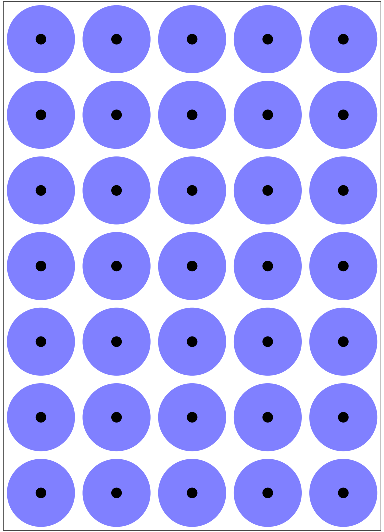


- View $\{+1, -1\}^N$ as points on an N-dimensional sphere.
- Pick 2^k well-sparated vectors, labelled by $\{0, 1\}^k$.
- Alice: $f_A(X)$ = closest vector to X
- Bob: $f_B(Y)$ = closest vector to Y

Proof idea

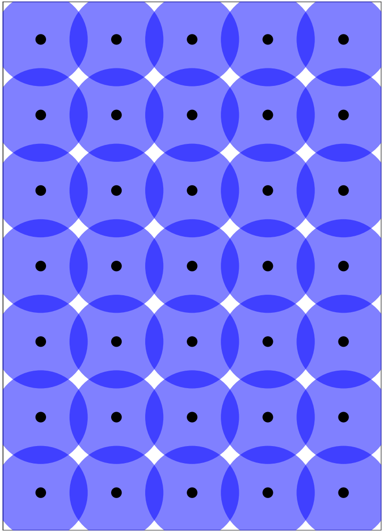
The projections along the various directions are Gaussian and approximately independent.

Alice's view



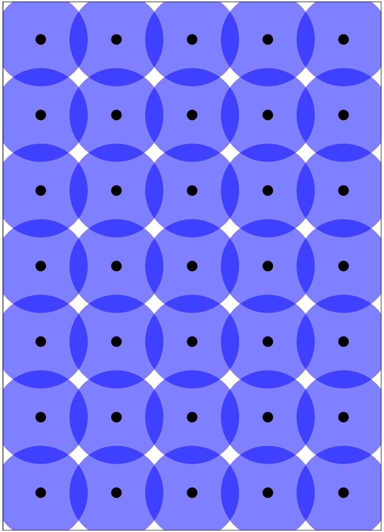
- The ambient space is $\{+1, -1\}^N$.
- The space is partitioned into disks.
- When X falls in a disk, Alice reports the label of its center.
- Each disk has volume $\approx 2^{-k}$.

Bob's view



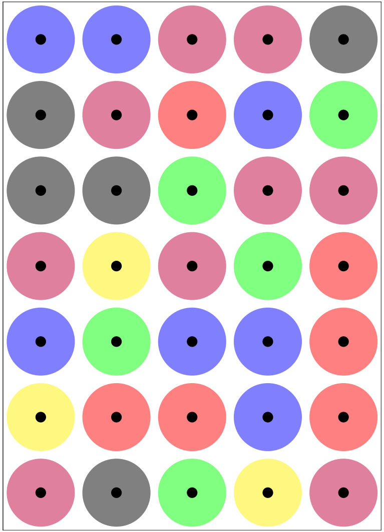
- The disks are bigger and overlap.
- Bogdanov and Mossel '10: *About* $2^{-(\varepsilon/(1-\varepsilon))k}$ *of the volume is covered by only one disk.*

Bob's view



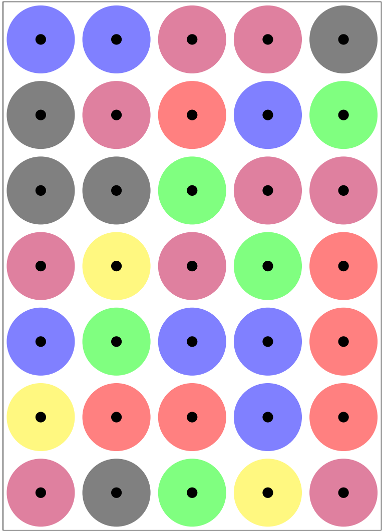
- The disks are bigger and overlap.
- Bogdanov and Mossel '10: *About $2^{-(\varepsilon/(1-\varepsilon))k}$ of the volume is covered by only one disk.*

Alice’s message



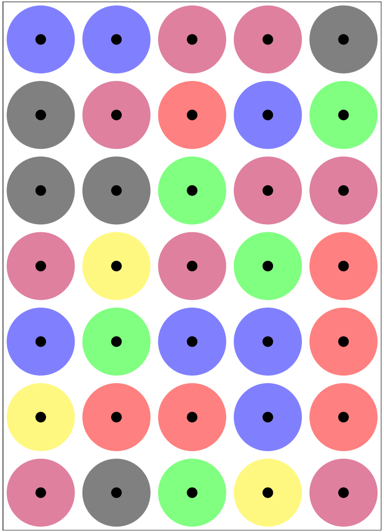
- The space is partitioned into disks.
- The disks are colored using 2^c colors.
- When X falls in a disk, Alice reports the label of its center.
- Alice sends Bob the color of the disk (c bits).

Alice's message



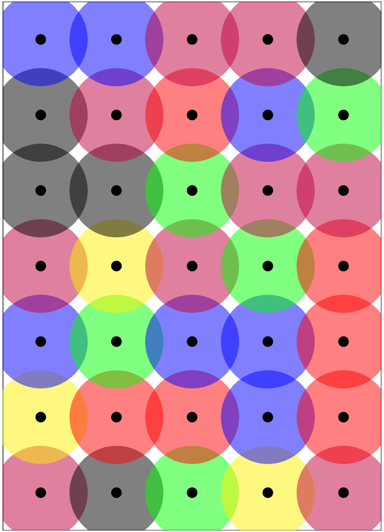
- The space is partitioned into disks.
- The disks are colored using 2^c colors.
- When X falls in a disk, Alice reports the label of its center.
- Alice sends Bob the color of the disk (c bits).

Alice's message



- The space is partitioned into disks.
- The disks are colored using 2^c colors.
- When X falls in a disk, Alice reports the label of its center.
- Alice sends Bob the color of the disk (c bits).

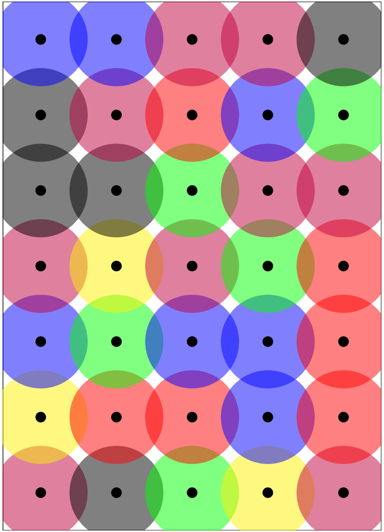
Bob’s view



- Again, the disks are bigger and overlap.
- *But, most points are covered by only one disk of a given color.*
- Bob uniquely identifies the disk (and its center).

How many colors must Alice use?

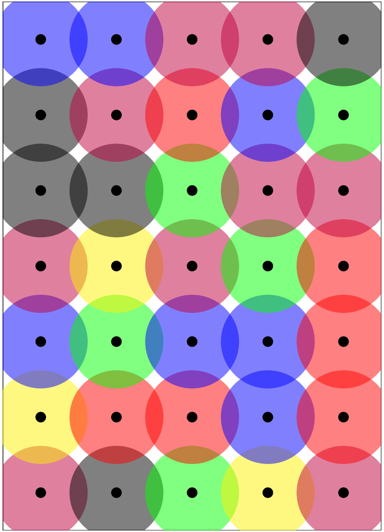
Bob's view



- Again, the disks are bigger and overlap.
- *But, most points are covered by only one disk of a given color.*
- Bob uniquely identifies the disk (and its center).

How many colors must Alice use?

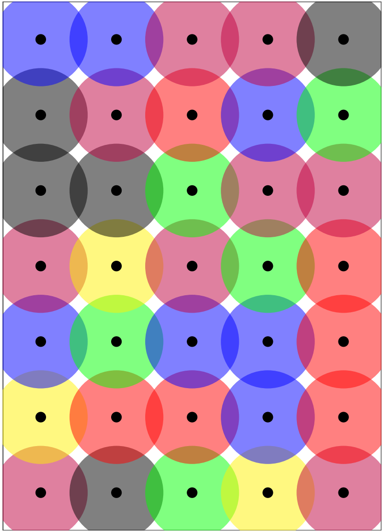
Bob's view



- Again, the disks are bigger and overlap.
- *But, most points are covered by only one disk of a given color.*
- Bob uniquely identifies the disk (and its center).

How many colors must Alice use?

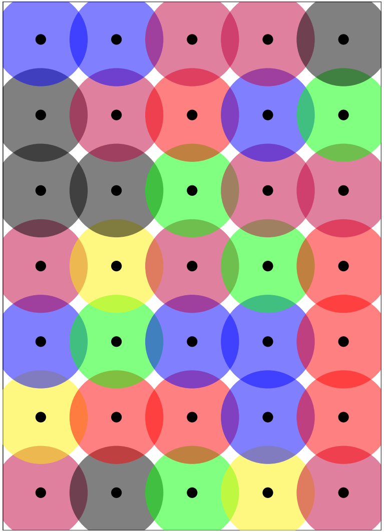
Bob's view



- Again, the disks are bigger and overlap.
- *But, most points are covered by only one disk of a given color.*
- Bob uniquely identifies the disk (and its center).

How many colors must Alice use?

Bob’s view



- Again, the disks are bigger and overlap.
- *But, most points are covered by only one disk of a given color.*
- Bob uniquely identifies the disk (and its center).

How many colors must Alice use?

The calculation

- Suppose Alice chooses $(1, 1, \ldots, 1)$
- $\mathrm{vol}(\text{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\mathrm{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\mathrm{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\mathrm{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The calculation

- Suppose Alice chooses $(1, 1, \ldots, 1)$
- $\text{vol}(\text{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\text{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\text{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\text{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The calculation

- Suppose Alice chooses $(1, 1, \ldots, 1)$
- $\mathrm{vol}(\text{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\mathrm{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\mathrm{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\mathrm{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The calculation

- Suppose Alice chooses $(1, 1, \ldots, 1)$
- $\mathrm{vol}(\mathrm{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\mathrm{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\mathrm{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\mathrm{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The calculation

- Suppose Alice chooses $(1, 1, \ldots, 1)$
- $\mathrm{vol}(\text{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\mathrm{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\mathrm{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\mathrm{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The calculation

- Suppose Alice chooses $(1, 1, \dots, 1)$
- $\mathrm{vol}(\mathrm{disk}) \approx 2^{-k}$.
- Then for a typical x in Alice's disk, $\mathrm{bias}(x) = t$, where

$$\exp(-t^2) \approx 2^{-k}.$$

- $\mathrm{bias}(y) = (1 - 2\varepsilon)t$.
- Then,

$$\mathrm{vol}(\text{expanded disk}) \approx 2^{-(1-2\varepsilon)^2 k}.$$

- So, a typical point is in

$$2^k 2^{-k(1-2\varepsilon)^2} = 2^{4\varepsilon(1-\varepsilon)k}$$

disks.

If Alice uses $\approx 2^{4\varepsilon(1-\varepsilon)k}$ colors, then most points will be covered by at most one disk of any color.

The lower bound

Theorem

$C^{\mathrm{BSC}(\varepsilon)}(k, 0.5) \geq 4\varepsilon(1-\varepsilon)k - o(k).$

Idea

The "expansion of the discs" seen by Bob is inevitable. This is formalized using a hypercontractivity inequality of the following form. For $f_{\mathrm{Alice}} : \{+1, -1\}^n \to \mathbb{R}$, let $f_{\mathrm{Bob}}(y) := \mathbb{E}[f_{\mathrm{Alice}}(X) | Y = y]$.

The lower bound

Theorem

$C^{\mathrm{BSC}(\varepsilon)}(k, 0.5) \geq 4\varepsilon(1-\varepsilon)k - o(k).$

Idea

The "expansion of the discs" seen by Bob is inevitable. This is formalized using a hypercontractivity inequality of the following form. For $f_{\mathrm{Alice}} : \{+1,-1\}^n \to \mathbb{R}$, let $f_{\mathrm{Bob}}(y) := \mathbb{E}[f_{\mathrm{Alice}}(X) | Y = y]$.

The lower bound

Theorem

$C^{\mathrm{BSC}(\varepsilon)}(k, 0.5) \geq 4\varepsilon(1-\varepsilon)k - o(k)$.

Idea

The “expansion of the discs” seen by Bob is inevitable. This is formalized using a hypercontractivity inequality of the following form. For $f_{\mathrm{Alice}} : \{+1,-1\}^n \to \mathbb{R}$, let $f_{\mathrm{Bob}}(y) := \mathbb{E}[f_{\mathrm{Alice}}(X) | Y = y]$. Then, for $q > 1$ and $p = 1 + (1-2\varepsilon)^2(q-1)$,

$$\|f_{\mathrm{Bob}}\|_q \leq \|f_{\mathrm{Alice}}\|_p,$$

where for $\alpha, \beta : \{+1,-1\}^N \to \mathbb{R}$,

$$\|\alpha\|_p = \mathbb{E}_X[|\alpha(X)|^p]^{1/p} \qquad \|\beta\|_q = \mathbb{E}_Y[|\beta(Y)|^q]^{1/q}.$$

The analysis

- Suppose Bob receives y.
- Let $\beta(z|y) = \Pr[g_A(X) = z | Y = y]$.
- Without a message, the best strategy for Bob is to output the z for which $\beta(z|y)$ is maximum.

$$\Pr[\text{Success} \mid Y = y] \leq \max_z \beta(z|y).$$

- Suppose Alice sends c-bit messages; so, there are at most $t = 2^c$ possible transcripts. Then,

$$\Pr[\text{Success} \mid Y = y] \leq \sum_{i=1}^{2^c} \beta(z_i|y),$$

where $z_1, z_2, \ldots, z_t \in \{0,1\}^k$ give the top t values for $\beta(z|y)$.

The analysis

- Suppose Bob receives y.
- Let $\beta(z|y) = \Pr[g_A(X) = z | Y = y]$.
- Without a message, the best strategy for Bob is to output the z for which $\beta(z|y)$ is maximum.

$$\Pr[\text{Success} \mid Y = y] \le \max_z \beta(z|y).$$

- Suppose Alice sends c-bit messages; so, there are at most $t = 2^c$ possible transcripts. Then,

$$\Pr[\text{Success} \mid Y = y] \le \sum_{i=1}^{2^c} \beta(z_i|y),$$

where $z_1, z_2, \ldots, z_t \in \{0,1\}^k$ give the top t values for $\beta(z|y)$.

The analysis

- Suppose Bob receives y.
- Let $\beta(z|y) = \Pr[g_A(X) = z | Y = y]$.
- Without a message, the best strategy for Bob is to output the z for which $\beta(z|y)$ is maximum.

$$\Pr[\text{Success} \mid Y = y] \leq \max_z \beta(z|y).$$

- Suppose Alice sends c-bit messages; so, there are at most $t = 2^c$ possible transcripts. Then,

$$\Pr[\text{Success} \mid Y = y] \leq \sum_{i=1}^{2^c} \beta(z_i|y),$$

where $z_1, z_2, \ldots, z_t \in \{0,1\}^k$ give the top t values for $\beta(z|y)$.

The analysis

- Suppose Bob receives y.
- Let $\beta(z|y) = \Pr[g_A(X) = z | Y = y]$.
- Without a message, the best strategy for Bob is to output the z for which $\beta(z|y)$ is maximum.

$$\Pr[\text{Success} \mid Y = y] \leq \max_z \beta(z|y).$$

- Suppose Alice sends c-bit messages; so, there are at most $t = 2^c$ possible transcripts. Then,

$$\Pr[\text{Success} \mid Y = y] \leq \sum_{i=1}^{2^c} \beta(z_i|y),$$

where $z_1, z_2, \ldots, z_t \in \{0,1\}^k$ give the top t values for $\beta(z|y)$.

The final bound

Claim

$$\Pr[\text{Success}] \leq \left(\sum_z \mathop{\mathbb{E}}_Y \left[\beta(z|Y)^q\right] \right)^{1/q} \cdot t^{1-1/q}.$$

- Using the hypercontractivity inequality with $q = 1 + \delta$ (so, $p = 1 + (1 - 2\varepsilon)^2\delta$), we obtain

$$t \geq \Pr[\text{Success}]^{(1+\delta)/\delta} \cdot 2^{4\varepsilon(1-\varepsilon)k/(1+(1-2\varepsilon)^2\delta)}.$$

- Set $\Pr[\text{Success}] = 2^{-\gamma k}$.
  Choose $\delta > 0$ optimally for each choice of γ.

The final bound

Claim

$$\Pr[\text{Success}] \leq \left(\sum_z \mathop{\mathbb{E}}_Y [\beta(z|Y)^q]\right)^{1/q} \cdot t^{1-1/q}.$$

- Using the hypercontractivity inequality with $q = 1 + \delta$ (so, $p = 1 + (1-2\varepsilon)^2\delta$), we obtain

$$t \geq \Pr[\text{Success}]^{(1+\delta)/\delta} \cdot 2^{4\varepsilon(1-\varepsilon)k/(1+(1-2\varepsilon)^2\delta)}.$$

- Set $\Pr[\text{Success}] = 2^{-\gamma k}$.
  Choose $\delta > 0$ optimally for each choice of γ.

The final bound

Claim

$$\Pr[\text{Success}] \le \left(\sum_z \mathop{\mathbb{E}}_Y \left[\beta(z|Y)^q \right] \right)^{1/q} \cdot t^{1-1/q}.$$

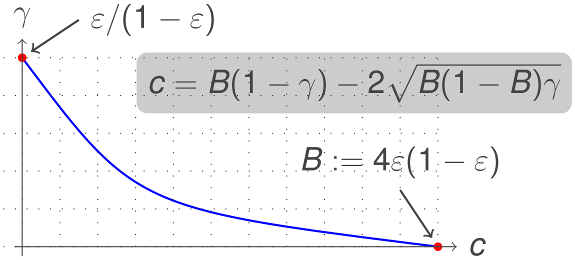
- Using the hypercontractivity inequality with $q = 1 + \delta$ (so, $p = 1 + (1-2\varepsilon)^2\delta$), we obtain

$$t \ge \Pr[\text{Success}]^{(1+\delta)/\delta} \cdot 2^{4\varepsilon(1-\varepsilon)k/(1+(1-2\varepsilon)^2\delta)}.$$

- Set $\Pr[\text{Success}] = 2^{-\gamma k}$.
  Choose $\delta > 0$ optimally for each choice of γ.

The result

$$\text{Probability of agreement} = 2^{-\gamma k}$$
$$\text{Communication} = ck$$



The Binary Erasure Channel

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) \in \{0,1\}^k$

Bob

Input: $Y \in \{0,1,\star\}^N$

Output: $f_B(Y) \in \{0,1\}^k$

$(X, Y) \sim \mathrm{BEC}(\varepsilon)$: $\Pr[Y_i = \star \mid X_i = x] = \varepsilon$; $\Pr[Y_i = x \mid X_i = x] = 1 - \varepsilon$.

Probability of agreement $= 2^{-\gamma k}$

Communication $= ck$

- $c = \varepsilon(1-\gamma) - 2\sqrt{\varepsilon(1-\varepsilon)\gamma}$.
- The lower bound is based on a new hypercontractivity inequality due to Nair and Wang (2015).

The Binary Erasure Channel

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) \in \{0,1\}^k$

Bob

Input: $Y \in \{0,1,\star\}^N$

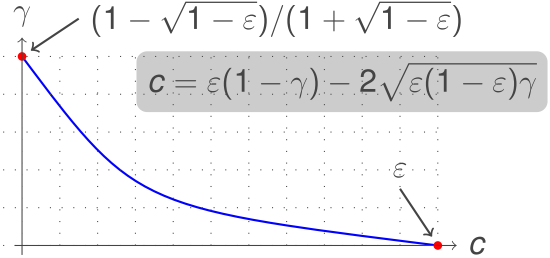
Output: $f_B(Y) \in \{0,1\}^k$

$(X, Y) \sim \mathrm{BEC}(\varepsilon)$: $\Pr[Y_i = \star \mid X_i = x] = \varepsilon$; $\Pr[Y_i = x \mid X_i = x] = 1 - \varepsilon$.

Probability of agreement $= 2^{-\gamma k}$

Communication $= ck$

- $c = \varepsilon(1-\gamma) - 2\sqrt{\varepsilon(1-\varepsilon)\gamma}$.
- The lower bound is based on a new hypercontractivity inequality due to Nair and Wang (2015).



The Binary Erasure Channel

Alice

Input: $X \in \{0,1\}^N$

Output: $f_A(X) \in \{0,1\}^k$

Bob

Input: $Y \in \{0,1,\star\}^N$

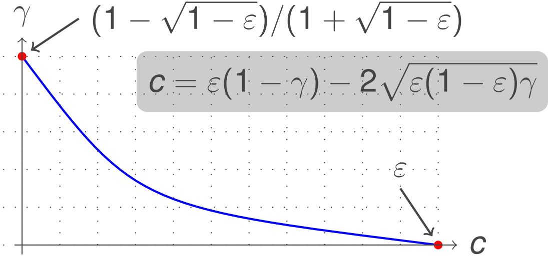
Output: $f_B(Y) \in \{0,1\}^k$

$(X,Y) \sim \mathrm{BEC}(\varepsilon)$: $\Pr[Y_i = \star \mid X_i = x] = \varepsilon$; $\Pr[Y_i = x \mid X_i = x] = 1 - \varepsilon$.

Probability of agreement $= 2^{-\gamma k}$

Communication $= ck$

- $c = \varepsilon(1-\gamma) - 2\sqrt{\varepsilon(1-\varepsilon)\gamma}$.
- The lower bound is based on a new hypercontractivity inequality due to Nair and Wang (2015).



Open problems

- Our lower bounds apply to commmunication over multiple rounds, if we assume that Alice's output depends on her input alone.
  We do not know if the same lower bound holds without this assumption.

- We examined the two well-studied channels, BSC(ε) and BEC(ε), for agreement probability parametrized as $1 - 2^{\gamma k}$. *We do not know if a general bound applicable to all channels can be stated in terms of some information-theoretic parameter of the channel.* For agreement close to 1 (error close to zero), the works of Zhao & Chia (2011) and Anantharam *et al.* (2013) address this question.

Thank you!

Open problems

- Our lower bounds apply to commmunication over multiple rounds, if we assume that Alice's output depends on her input alone.
  We do not know if the same lower bound holds without this assumption.

- We examined the two well-studied channels, BSC(ε) and BEC(ε), for agreement probability parametrized as $1 - 2^{\gamma k}$. *We do not know if a general bound applicable to all channels can be stated in terms of some information-theoretic parameter of the channel.* For agreement close to 1 (error close to zero), the works of Zhao & Chia (2011) and Anantharam *et al.* (2013) address this question.

Thank you!

Open problems

- Our lower bounds apply to commmunication over multiple rounds, if we assume that Alice's output depends on her input alone.
  We do not know if the same lower bound holds without this assumption.

- We examined the two well-studied channels, BSC(ε) and BEC(ε), for agreement probability parametrized as $1 - 2^{\gamma k}$. *We do not know if a general bound applicable to all channels can be stated in terms of some information-theoretic parameter of the channel.* For agreement close to 1 (error close to zero), the works of Zhao & Chia (2011) and Anantharam *et al.* (2013) address this question.

Thank you!