Toward the KRW conjecture

Cubic Formula Lower Bounds via Communication Complexity

Irit Dinur Or Meir

Formulas

- (de-Morgan) Formulas are circuits with fan-out 1.
- Cannot store intermediate results.
- The formula complexity $\mathsf{L}(f)$ is the size of the smallest formula for f.

Formulas

- Would like: Explicit f with $\mathsf{L}(f) = n^{\omega(1)}$.

Formulas

- Would like: Explicit f with $\mathsf{L}(f) = n^{\omega(1)}$.
- State-of-the-art: Andreev's function [A87] has complexity $\tilde{\Omega}(n^3)$ [H98].

Formulas

- Would like: Explicit f with $\mathsf{L}(f) = n^{\omega(1)}$.
- State-of-the-art: Andreev's function [A87] has complexity $\tilde{\Omega}(n^3)$ [H98].
- Proof based on shrinkage [S61, IN93, PZ93. H98, T14].
- We now have an optimal analysis of shrinkage.

Formulas

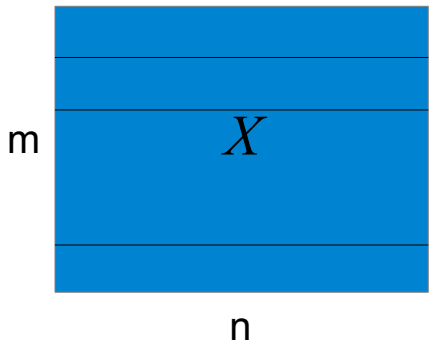
- Would like: Explicit f with $\mathsf{L}(f) = n^{\omega(1)}$.
- State-of-the-art: Andreev's function [A87] has complexity $\tilde{\Omega}(n^3)$ [H98].
- Proof based on shrinkage [S61, IN93, PZ93. H98, T14].
- We now have an optimal analysis of shrinkage.
- So shrinkage is unlikely to get us any further.

Composition

- [KRW91]: we need to understand composition.

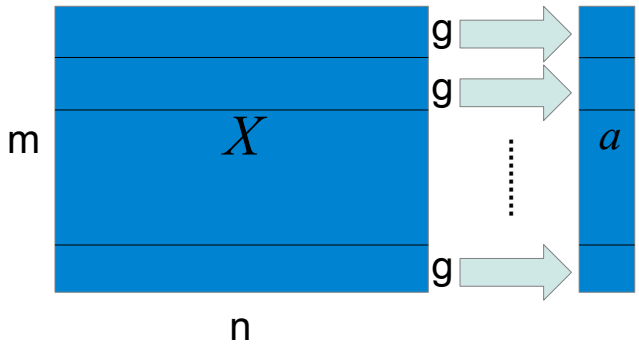
Composition

- [KRW91]: we need to understand composition.
- Let $f : \{0,1\}^m \to \{0,1\}$, $g : \{0,1\}^n \to \{0,1\}$.
- The composition $f \diamond g : \{0,1\}^{m \times n} \to \{0,1\}$ is



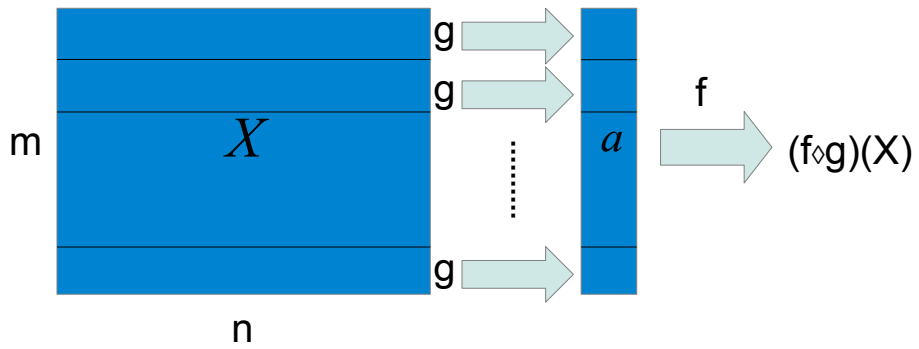
Composition

- [KRW91]: we need to understand composition.
- Let $f : \{0,1\}^m \to \{0,1\}$, $g : \{0,1\}^n \to \{0,1\}$.
- The composition $f \diamond g : \{0,1\}^{m \times n} \to \{0,1\}$ is

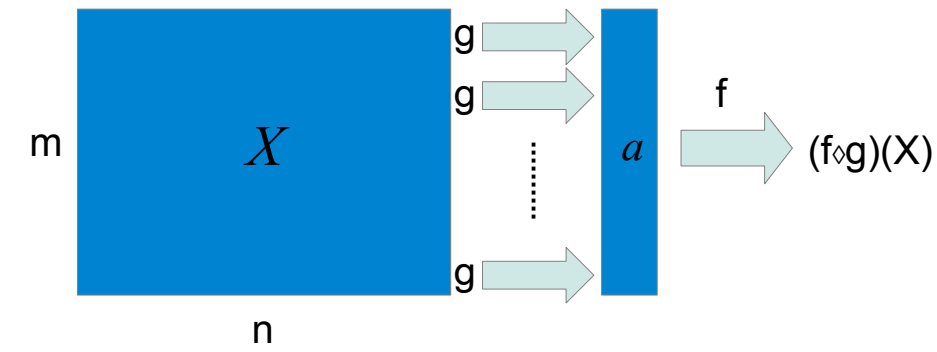


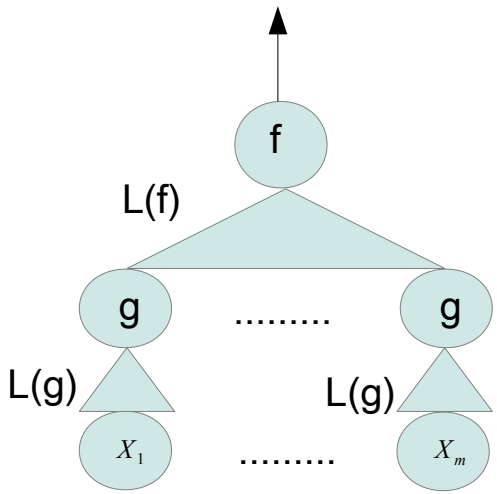
Composition

- [KRW91]: we need to understand composition.
- Let $f : \{0,1\}^m \to \{0,1\}$, $g : \{0,1\}^n \to \{0,1\}$.
- The composition $f \diamond g : \{0,1\}^{m \times n} \to \{0,1\}$ is



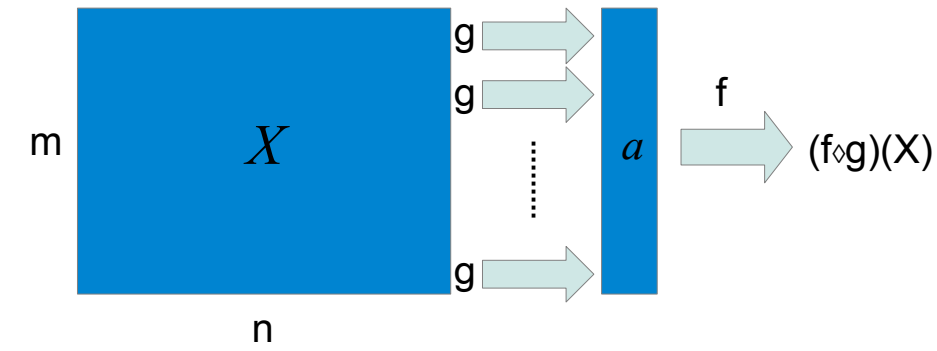
The KRW conjecture

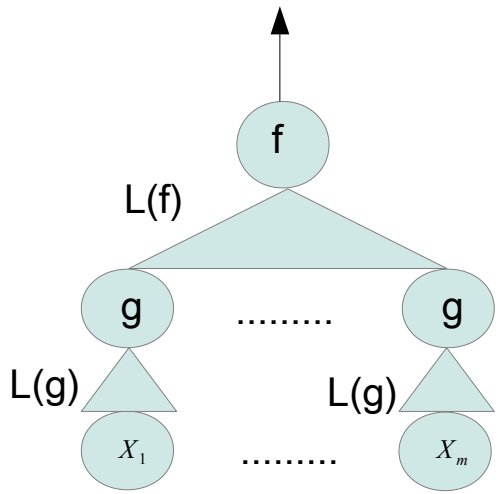




- Clearly, $\mathsf{L}(f \diamond g) \leq \mathsf{L}(f) \cdot \mathsf{L}(g)$.

The KRW conjecture





- Clearly, $\mathsf{L}(f \diamond g) \leq \mathsf{L}(f) \cdot \mathsf{L}(g)$.
- KRW conjecture*: $\forall f, g: \ \mathsf{L}(f \diamond g) \approx \mathsf{L}(f) \cdot \mathsf{L}(g)$.

The KRW conjecture — Prior Work

- KRW conjecture*: $\forall f, g: \ \mathsf{L}(f \diamond g) \approx \mathsf{L}(f) \cdot \mathsf{L}(g)$.
- Implies super-polynomial formula lower bounds.

The KRW conjecture — Prior Work

- KRW conjecture*: $\forall f, g: \ \mathsf{L}(f \diamond g) \approx \mathsf{L}(f) \cdot \mathsf{L}(g)$.
- Implies super-polynomial formula lower bounds.

- [KRW91] defined the universal relation U.
- Like a function, but simpler.
- Suggested to prove the conjecture for $U \diamond U$.

The KRW conjecture — Prior Work

- KRW conjecture*: $\forall f, g: \ \mathsf{L}(f \diamond g) \approx \mathsf{L}(f) \cdot \mathsf{L}(g)$.
- Implies super-polynomial formula lower bounds.

- [KRW91] defined the universal relation U.
- Like a function, but simpler.
- Suggested to prove the conjecture for $U \diamond U$.

- Conjecture was proved for $U \diamond U$ by [EIRS91]
- Alternative proof by [HW93].
- Recently: [GMWW14] proved conjecture for $f \diamond U$.

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g)/2^{\tilde{O}(\sqrt{m})}$$

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g)/2^{\tilde{O}(\sqrt{m})}$$

- Also, we give a structural result:

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g)/2^{\tilde{O}(\sqrt{m})}$$

- Also, we give a structural result:
  - KRW conjecture: naive formula is optimal.

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g)/2^{\tilde{O}(\sqrt{m})}$$

- Also, we give a structural result:
  - KRW conjecture: naive formula is optimal.
  - Our result: naive formula is essentially the only optimal formula.

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g)/2^{\tilde{O}(\sqrt{m})}$$

- Also, we give a structural result:
  - KRW conjecture: naive formula is optimal.
  - Our result: naive formula is essentially the only optimal formula.
- Actually, the lower bound for $f \diamond g$ already follows from [H98].
- However, our proof is very different, and seems to be more generalizable for other g's.

Our results

- We prove conjecture for $f \diamond g$ where g is parity:

$$\mathsf{L}(f \diamond g) \geq \mathsf{L}(f) \cdot \mathsf{L}(g) / 2^{\tilde{O}(\sqrt{m})}$$

- Also, we give a structural result:
  - KRW conjecture: naive formula is optimal.
  - Our result: naive formula is essentially the only optimal formula.
- Actually, the lower bound for $f \diamond g$ already follows from [H98].
- However, our proof is very different, and seems to be more generalizable for other g's.
- Also: new proof of the state-of-the-art cubic lower bound of [H98].

Outline

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find such i.

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find such i.
  - Want to talk as little as possible.

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find such i.
  - Want to talk as little as possible.
- Only deterministic protocols!

Karchmer-Wigderson Relations [KW90]

- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find such i.
  - Want to talk as little as possible.
- Only deterministic protocols!
- This talk: Assume $\mathsf{C}(KW_f) = \log \mathsf{L}(f)$.

Karchmer-Wigderson Relations [KW90]

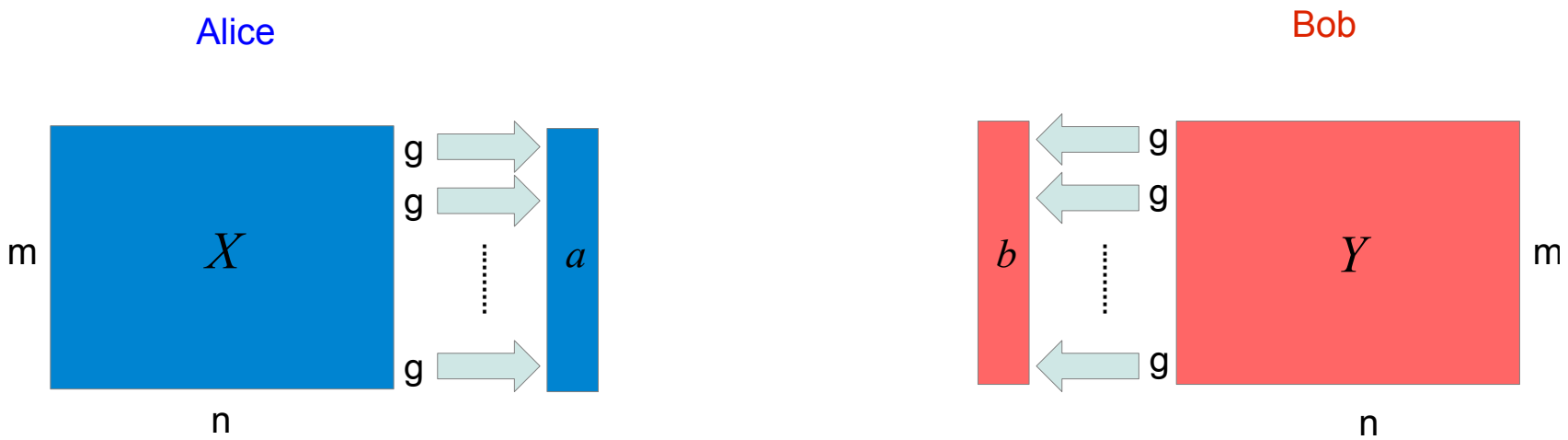
- Relate $\mathsf{L}(f)$ to complexity of a communication problem KW_f.
- The KW relation KW_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find such i.
  - Want to talk as little as possible.
- Only deterministic protocols!
- This talk: Assume $\mathsf{C}(KW_f) = \log \mathsf{L}(f)$.
- KRW conjecture: $\mathsf{C}(KW_{f \diamond g}) \approx \mathsf{C}(KW_f) + \mathsf{C}(KW_g)$.

KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.

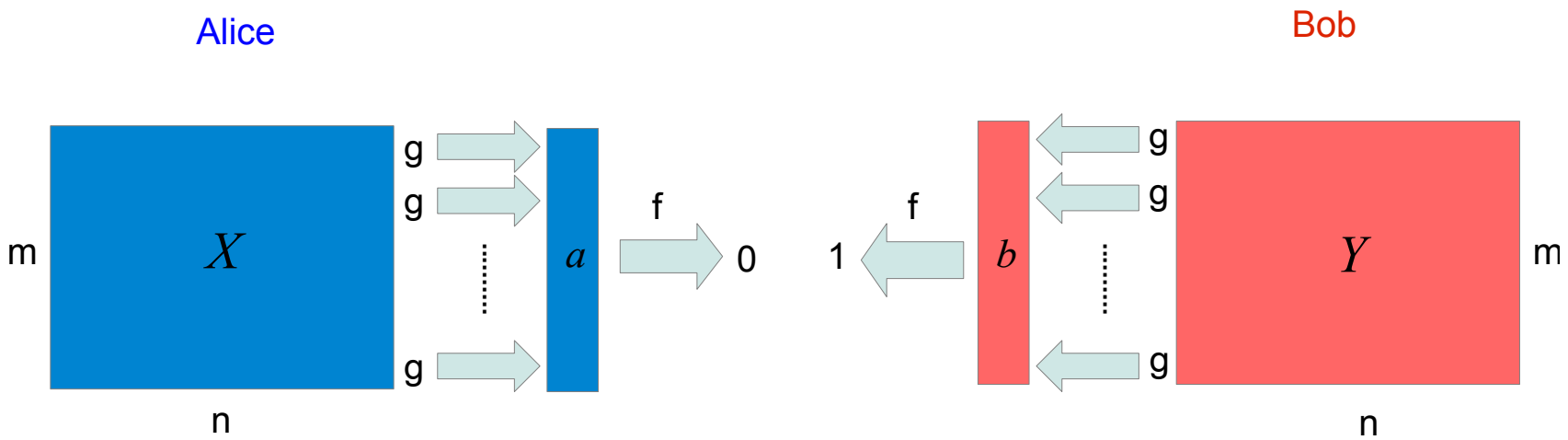
KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



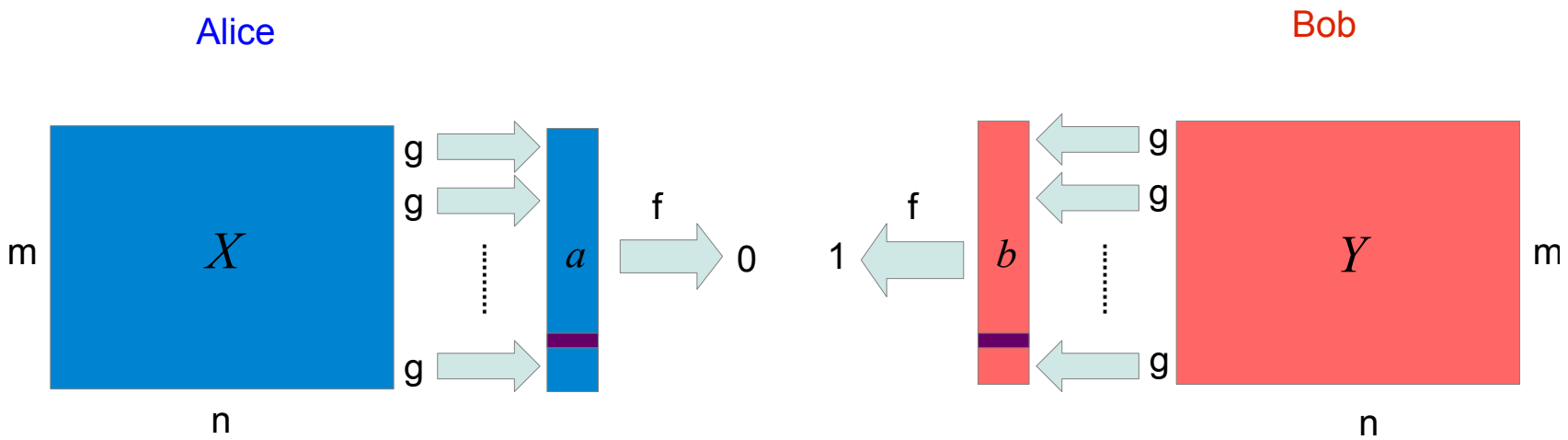
KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



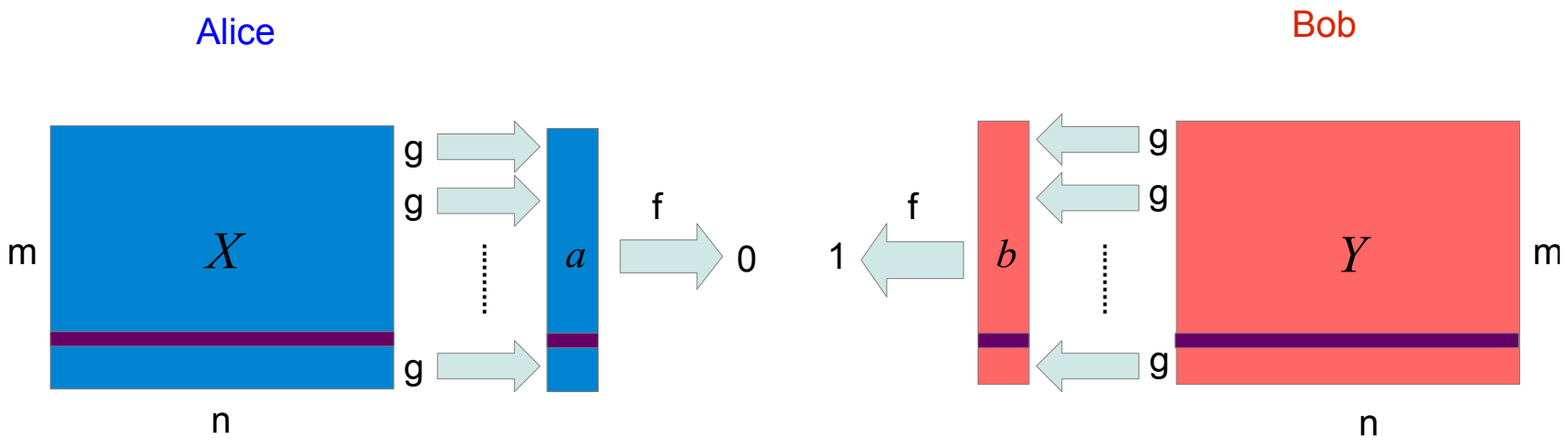
KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



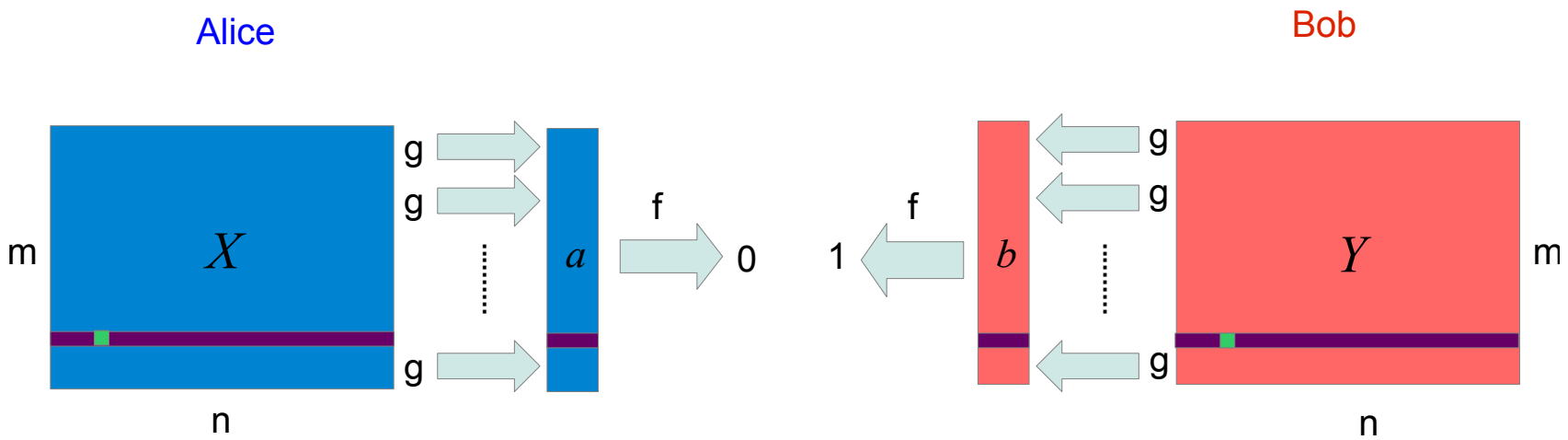
KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



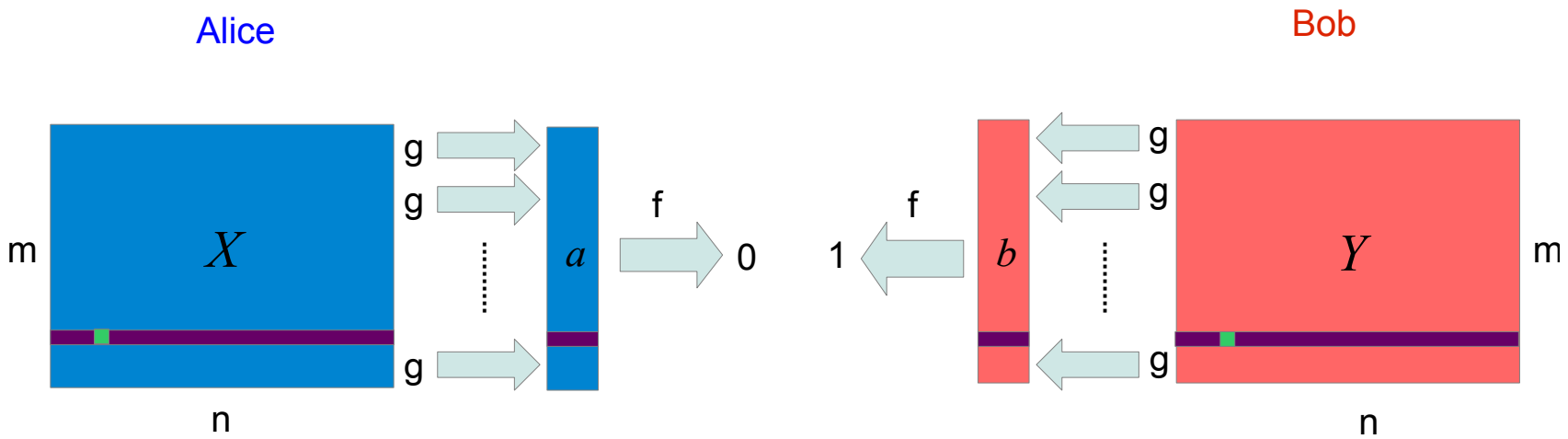
KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



KRW and KW

- Can we use KW relations to attack the KRW conjecture?
- How does $KW_{f \diamond g}$ look like?
- Recall: $f \diamond g$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



KRW conjecture: $\mathsf{C}(KW_{f \diamond g}) \approx \mathsf{C}(KW_f) + \mathsf{C}(KW_g)$

The obvious protocol is essentially optimal.

Outline

Proof strategy

- We wish to show that the obvious protocol is optimal.

Proof strategy

- We wish to show that the obvious protocol is optimal.
- Fix a protocol Π for $KW_{f \diamond g}$.
- We will show that Π must behave roughly like the obvious protocol.

Proof strategy

- We wish to show that the obvious protocol is optimal.
- Fix a protocol Π for $KW_{f \diamond g}$.
- We will show that Π must behave roughly like the obvious protocol.
- Partition each transcript π of Π to two parts: $\pi = \pi_1 \circ \pi_2$.
- π_1 and π_2 correspond to the two stages of the obvious protocol.

Proof strategy

- Show:

Proof strategy

- Show:
  1. $\forall\ \pi_1$ that has not solved KW_f, $\exists\ \pi_2$ of length $C(KW_g)$.

Proof strategy

- Show:
  1. $\forall\ \pi_1$ that has not solved KW_f, $\exists\ \pi_2$ of length $C(KW_g)$.
  2. $\exists\ \pi_1$ of length almost $\mathsf{C}(KW_f)$ that has not solved KW_f.

Proof strategy

- Show:
  1. $\forall\ \pi_1$ that has not solved KW_f, $\exists\ \pi_2$ of length $C(KW_g)$.
  2. $\exists\ \pi_1$ of length almost $\mathsf{C}(KW_f)$ that has not solved KW_f.
- Together, the two items imply a lower bound of

$$\approx \mathsf{C}(KW_f) + \mathsf{C}(KW_g).$$

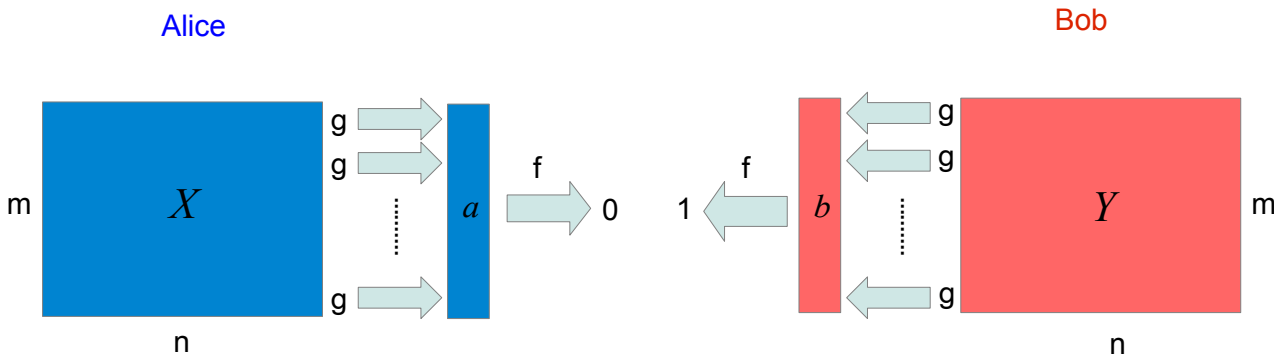
Proof strategy

- Show:
  1. $\forall\ \pi_1$ that has not solved KW_f, $\exists\ \pi_2$ of length $C(KW_g)$.
  2. $\exists\ \pi_1$ of length almost $\mathsf{C}(KW_f)$ that has not solved KW_f.
- Together, the two items imply a lower bound of

$$\approx \mathsf{C}(KW_f) + \mathsf{C}(KW_g).$$

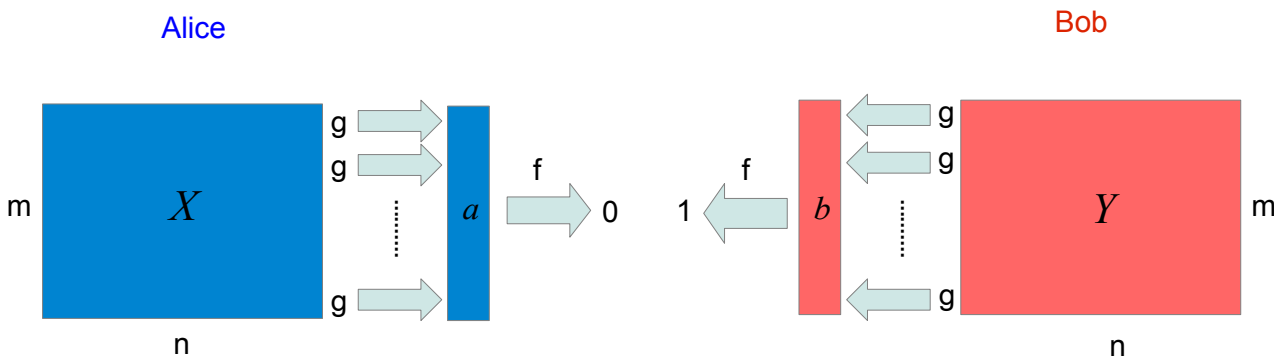
- The first item gives our structural result.
- This will be the focus of the rest of this talk.

Structural Result - Basic Idea



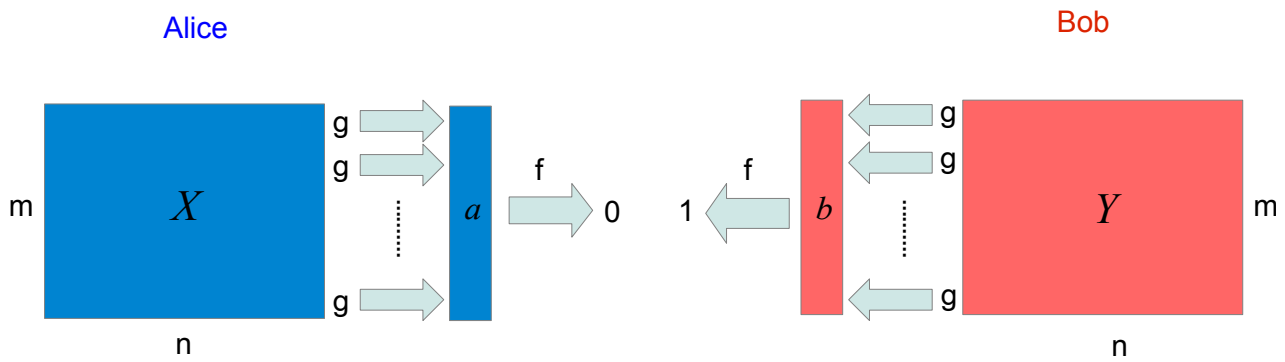
- In the 2nd stage, they need to solve KW_g on some X_i, Y_i.

Structural Result - Basic Idea



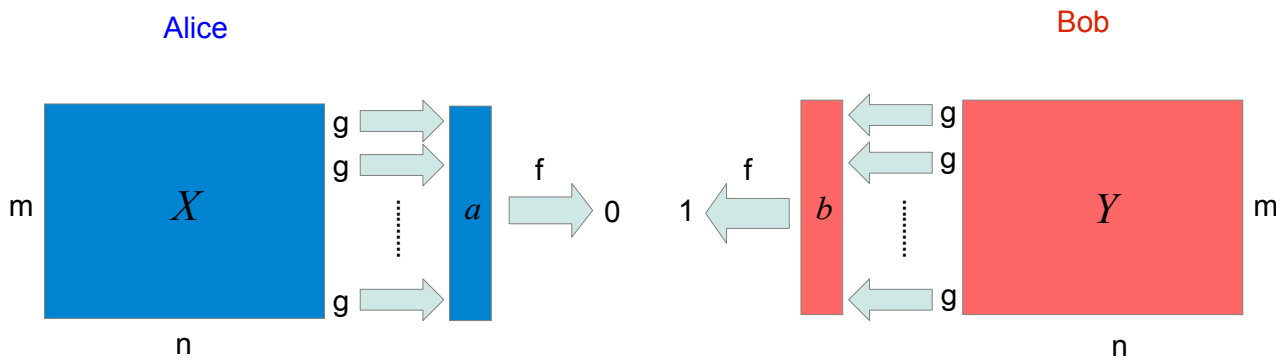
- In the 2nd stage, they need to solve KW_g on some X_i, Y_i.
- In principle, this requires $\mathsf{C}(KW_g)$ bits.

Structural Result - Basic Idea



- In the 2nd stage, they need to solve KW_g on some X_i, Y_i.
- In principle, this requires $\mathsf{C}(KW_g)$ bits.
- However, they already communicated almost $\mathsf{C}(KW_f)$ bits in the 1st stage.

Structural Result - Basic Idea



- In the 2nd stage, they need to solve KW_g on some X_i, Y_i.
- In principle, this requires $\mathrm{C}(KW_g)$ bits.
- However, they already communicated almost $\mathrm{C}(KW_f)$ bits in the 1st stage.
- Need to show: does not make the 2nd stage easier.

Structural Result - Basic Idea

- The players communicated less than

$$\mathsf{C}(KW_f) \leq m$$

bits in the first stage.

Structural Result - Basic Idea

- The players communicated less than

$$\mathsf{C}(KW_f) \leq m$$

  bits in the first stage.

- On average: less than 1 bit per row.

Structural Result - Basic Idea

- The players communicated less than

$$\mathsf{C}(KW_f) \leq m$$

  bits in the first stage.

- On average: less than 1 bit per row.

- Thus, on the typical row, they did not make much progress.

Structural Result - Basic Idea

- The players communicated less than

$$\mathsf{C}(KW_f) \leq m$$

  bits in the first stage.

- On average: less than 1 bit per row.

- Thus, on the typical row, they did not make much progress.

- If they solve a typical row in the 2nd stage, they still need to send $\approx \mathsf{C}(KW_g)$ bits.

Structural Result - Basic Idea

- There might be non-typical rows on which Alice and Bob communicated a lot.

Structural Result - Basic Idea

- There might be non-typical rows on which Alice and Bob communicated a lot.
- Might be easy to solve KW_g on those rows.

Structural Result - Basic Idea

- There might be non-typical rows on which Alice and Bob communicated a lot.
- Might be easy to solve KW_g on those rows.
- Ensure: Players do not solve KW_g on non-typical rows.

Structural Result - Basic Idea

- There might be non-typical rows on which Alice and Bob communicated a lot.
- Might be easy to solve KW_g on those rows.
- Ensure: Players do not solve KW_g on non-typical rows.
- Force $X_i = Y_i$ on non-typical rows.

Structural Result - Basic Idea

- Forcing $X_i = Y_i$ is hardest part of the proof.

Structural Result - Basic Idea

- Forcing $X_i = Y_i$ is hardest part of the proof.
- In particular, need to force $g(X_i) = g(Y_i)$.

Structural Result - Basic Idea

- Forcing $X_i = Y_i$ is hardest part of the proof.
- In particular, need to force $g(X_i) = g(Y_i)$.
- Issue: What if the players already found that $g(X_i) \neq g(Y_i)$ in the first stage?

Structural Result - Basic Idea

- Forcing $X_i = Y_i$ is hardest part of the proof.
- In particular, need to force $g(X_i) = g(Y_i)$.
- Issue: What if the players already found that $g(X_i) \neq g(Y_i)$ in the first stage?
- This cannot happen, since we know they did not solve KW_f in the first stage.

Credit

- This proof strategy was developed by Edmonds, Impagliazzo, Rudich and Sgall [EIRS91].
- Used it to prove the KRW conjecture for the composition of universal relations $U \diamond U$.

Credit

- This proof strategy was developed by Edmonds, Impagliazzo, Rudich and Sgall [EIRS91].
- Used it to prove the KRW conjecture for the composition of universal relations $U \diamond U$.

- Our novelty: extending this strategy to $f \diamond g$, where $g = \oplus$.

Credit

- This proof strategy was developed by Edmonds, Impagliazzo, Rudich and Sgall [EIRS91].
- Used it to prove the KRW conjecture for the composition of universal relations $U \diamond U$.
- Our novelty: extending this strategy to $f \diamond g$, where $g = \oplus$.
- This setting is considerably more challenging.

Main Difficulty

- Universal relations are "information-theoretic objects".

Main Difficulty

- Universal relations are "information-theoretic objects".
- Direct correspondence between information and complexity:
  - If Alice sends a message that has t bits of information,
  - the complexity of the problem decreases by exactly t bits.

Main Difficulty

- Universal relations are "information-theoretic objects".
- Direct correspondence between information and complexity:
  - If Alice sends a message that has t bits of information,
  - the complexity of the problem decreases by exactly t bits.
- KW relations are "computational objects".

Main Difficulty

- Universal relations are "information-theoretic objects".
- Direct correspondence between information and complexity:
  - If Alice sends a message that has t bits of information,
  - the complexity of the problem decreases by exactly t bits.
- KW relations are "computational objects".
- There are examples KW_f in which
  - Alice sends little information,
  - and decreases the complexity by a lot.

Main Difficulty

- We deal with this difficulty separately for KW_f and KW_g.

Main Difficulty

- We deal with this difficulty separately for KW_f and KW_g.
- For KW_f, we prove a general "fortification lemma", which creates a correspondence between information and complexity.

Main Difficulty

- We deal with this difficulty separately for KW_f and KW_g.
- For KW_f, we prove a general "fortification lemma", which creates a correspondence between information and complexity.
- For KW_g, we use the fact that the lower bound for parity has an information-theoretic proof.

Main Difficulty

- We deal with this difficulty separately for KW_f and KW_g.
- For KW_f, we prove a general "fortification lemma", which creates a correspondence between information and complexity.
- For KW_g, we use the fact that the lower bound for parity has an information-theoretic proof.
- Therefore, can analyze directly how information and complexity interact for parity.

Another difficulty: the "randomized barrier"

- Every KW relation has a randomized protocol of complexity $2 \log n$.
- Hence, cannot use techniques that work against randomized protocols.

Another difficulty: the "randomized barrier"

- Every KW relation has a randomized protocol of complexity $2 \log n$.
- Hence, cannot use techniques that work against randomized protocols.

- KW relations have no hard distributions (with complexity $> 2 \log n$).

Another difficulty: the "randomized barrier"

- Every KW relation has a randomized protocol of complexity $2 \log n$.
- Hence, cannot use techniques that work against randomized protocols.

- KW relations have no hard distributions (with complexity $> 2 \log n$).

- Hard to use information-theoretic techniques.
- Also, most rectangle bounds don't work [KKN95].

Another difficulty: the "randomized barrier"

- Every KW relation has a randomized protocol of complexity $2 \log n$.
- Hence, cannot use techniques that work against randomized protocols.

- KW relations have no hard distributions (with complexity $> 2 \log n$).

- Hard to use information-theoretic techniques.
- Also, most rectangle bounds don't work [KKN95].

- Our work is the first to bypass this barrier.

Outline

Fortification

- "Dream version": If Alice sends t bits of information, then complexity decreases by t.

Fortification

- "Dream version": If Alice sends t bits of information, then complexity decreases by t.
- Counter-example:
  - Suppose exactly half of Alice's inputs start with 0.
  - Suppose all of Bob's inputs start with 0.

Fortification

- "Dream version": If Alice sends t bits of information, then complexity decreases by t.
- Counter-example:
  - Suppose exactly half of Alice's inputs start with 0.
  - Suppose all of Bob's inputs start with 0.
- If Alice says "my first bit is 1" then
  - she sent 1 bit of information,
  - but complexity decreased to 0.

Fortification

- "Dream version": If Alice sends t bits of information, then complexity decreases by t.
- Counter-example:
  - Suppose exactly half of Alice's inputs start with 0.
  - Suppose all of Bob's inputs start with 0.
- If Alice says "my first bit is 1" then
  - she sent 1 bit of information,
  - but complexity decreased to 0.
- Can fix it: discard in advance all Alice's inputs that start with 1.

Fortification

- Recall: we model states of the protocol by rectangles $A \times B$.

Fortification

- Recall: we model states of the protocol by rectangles $A \times B$.
- We say a rectangle $A \times B$ is "fortified" if
  - whenever Alice's sends t bits of information,
  - the complexity decreases by at most $t + O(\log n)$.

Fortification

- Recall: we model states of the protocol by rectangles $A \times B$.
- We say a rectangle $A \times B$ is "fortified" if
  - whenever Alice's sends t bits of information,
  - the complexity decreases by at most $t + O(\log n)$.

Fortification lemma

Every rectangle $A \times B$ has a fortified subrectangle whose complexity is close to that of $A \times B$.

Averaging Argument for Information

"On the typical row, they did not make much progress."

Averaging Argument for Information

"On the typical row, they did not make much progress."

- This requires an "averaging argument" for information.

Averaging Argument for Information

"On the typical row, they did not make much progress."

- This requires an "averaging argument" for information.
- How do we model information?

Averaging Argument for Information

"On the typical row, they did not make much progress."

- This requires an "averaging argument" for information.
- How do we model information?

- For entropy: follows immediately from sub-additivity.

Averaging Argument for Information

"On the typical row, they did not make much progress."

- This requires an "averaging argument" for information.
- How do we model information?
- For entropy: follows immediately from sub-additivity.
- [EIRS91]: proved it for "predictability".

Averaging Argument for Information

"On the typical row, they did not make much progress."

- This requires an "averaging argument" for information.
- How do we model information?
- For entropy: follows immediately from sub-additivity.
- [EIRS91]: proved it for "predictability".
- This work: we prove it for min-entropy.

Summary

- Result: We (re-)prove the KRW conjecture for the special case $f \diamond g$ where g is parity.

Summary

- Result: We (re-)prove the KRW conjecture for the special case $f \diamond g$ where g is parity.
- Corollary: an alternative proof for the cubic lower bound for formulas.

Summary

- Result: We (re-)prove the KRW conjecture for the special case $f \diamond g$ where g is parity.
- Corollary: an alternative proof for the cubic lower bound for formulas.
- New tools: fortification lemma, averaging argument for min-entropy.

Summary

- Result: We (re-)prove the KRW conjecture for the special case $f \diamond g$ where g is parity.
- Corollary: an alternative proof for the cubic lower bound for formulas.
- New tools: fortification lemma, averaging argument for min-entropy.
- Future direction: replace $\bigoplus$ with a general function?

The 1-out-of-m problem [BBKW14]

- Alice and Bob are given m distinct instances of KW_g.
- They need to solve one of them.
- Prove: not easier than solving a single instance.

Thank you!