Toward Better Formula Lower Bounds: An Information Complexity Approach to the KRW Composition Conjecture

Dmitry Gavinsky Or Meir Omri Weinstein Avi Wigderson

Circuit lower bounds

- In complexity theory, we want to prove hardness.

Circuit lower bounds

- In complexity theory, we want to prove hardness.
- One model of computation we use is boolean circuits.

Circuit lower bounds

- In complexity theory, we want to prove hardness.
- One model of computation we use is boolean circuits.
- We would like to prove results of the form:
  - $f : \{0,1\}^n \to \{0,1\}$ does not have circuits of size $n^{O(1)}$.

Circuit lower bounds

- In complexity theory, we want to prove hardness.
- One model of computation we use is boolean circuits.
- We would like to prove results of the form:
  - $f : \{0,1\}^n \to \{0,1\}$ does not have circuits of size $n^{O(1)}$.
- We focus on explicit functions.

Circuit lower bounds

- In complexity theory, we want to prove hardness.
- One model of computation we use is boolean circuits.
- We would like to prove results of the form:
  - $f : \{0,1\}^n \to \{0,1\}$ does not have circuits of size $n^{O(1)}$.
- We focus on explicit functions.
- This talk: Fan-in is 2.

Weaker models

- Proving hardness for general circuits is hard.
- We try to prove hardness for weaker models.

Weaker models

- Proving hardness for general circuits is hard.
- We try to prove hardness for weaker models.
- This talk: Log-depth circuits and formulas.

Log-depth circuits

- The depth of a circuit is the length of the longest path from an input to an output.

Log-depth circuits

- The depth of a circuit is the length of the longest path from an input to an output.
- We study lower bounds for circuits of depth $O(\log n)$.
- Capture highly parallelizable computations.

Log-depth circuits

- The depth of a circuit is the length of the longest path from an input to an output.

- We study lower bounds for circuits of depth $O(\log n)$.
- Capture highly parallelizable computations.

- The depth complexity $\mathsf{D}(f)$ is the depth of the shallowest circuit for f.

Log-depth circuits

- The depth of a circuit is the length of the longest path from an input to an output.
- We study lower bounds for circuits of depth $O(\log n)$.
- Capture highly parallelizable computations.
- The depth complexity $\mathsf{D}(f)$ is the depth of the shallowest circuit for f.
- Would like: Explicit f with $\mathsf{D}(f) = \omega(\log n)$.

Formulas

- Formulas are circuits with fan-out 1.
- I.e., they are trees.
- Can not store intermediate results.

Formulas

- Formulas are circuits with fan-out 1.
- I.e., they are trees.
- Can not store intermediate results.
- The size of the formula is the number of its leaves.

Formulas

- Formulas are circuits with fan-out 1.
- I.e., they are trees.
- Can not store intermediate results.

- The size of the formula is the number of its leaves.
- The formula complexity $\mathsf{L}(f)$ is the size of the smallest formula for f.

Formulas

- Formulas are circuits with fan-out 1.
- I.e., they are trees.
- Can not store intermediate results.

- The size of the formula is the number of its leaves.
- The formula complexity $\mathsf{L}(f)$ is the size of the smallest formula for f.
- Would like: Explicit f with $\mathsf{L}(f) = n^{\omega(1)}$.

Models are related

- Every circuit of depth $d = O(\log n)$ can be transformed to a formula of size $2^d = \mathrm{poly}(n)$.

Models are related

- Every circuit of depth $d = O(\log n)$ can be transformed to a formula of size $2^d = \mathrm{poly}(n)$.
- Every formula of size $s = \mathrm{poly}(n)$ can be transormed to a circuit of depth $O(\log s) = O(\log n)$ (Spira's theorem).

Models are related

- Every circuit of depth $d = O(\log n)$ can be transformed to a formula of size $2^d = \text{poly}(n)$.
- Every formula of size $s = \text{poly}(n)$ can be transormed to a circuit of depth $O(\log s) = O(\log n)$ (Spira's theorem).
- The class $\mathbf{NC}_1$ can be defined as
  - The class of functions f with $\mathsf{D}(f) = O(\log n)$.
  - The class of functions f with $\mathsf{L}(f) = \text{poly}(n)$.

Models are related

- Every circuit of depth $d = O(\log n)$ can be transformed to a formula of size $2^d = \mathrm{poly}(n)$.
- Every formula of size $s = \mathrm{poly}(n)$ can be transormed to a circuit of depth $O(\log s) = O(\log n)$ (Spira's theorem).
- The class $\mathbf{NC}_1$ can be defined as
  - The class of functions f with $\mathsf{D}(f) = O(\log n)$.
  - The class of functions f with $\mathsf{L}(f) = \mathrm{poly}(n)$.
- Major open problem: Prove $\mathbf{NC}_1 \neq \mathbf{P}$.

- [KRW91] suggested an approach.

- [KRW91] suggested an approach.
- Let $f : \{0,1\}^n \to \{0,1\}$, $g : \{0,1\}^m \to \{0,1\}$.
- The composition $g \circ f : \{0,1\}^{m \times n} \to \{0,1\}$ is

$$(g \circ f)(x_1, \ldots, x_m) = g(f(x_1), \ldots, f(x_m)).$$

- Clearly, $\mathsf{D}(g \circ f) \leq \mathsf{D}(g) + \mathsf{D}(f)$.

The KRW Conjecture

- [KRW91] suggested an approach.
- Let $f : \{0,1\}^n \to \{0,1\}$, $g : \{0,1\}^m \to \{0,1\}$.
- The composition $g \circ f : \{0,1\}^{m \times n} \to \{0,1\}$ is

$$(g \circ f)(x_1, \ldots, x_m) = g(f(x_1), \ldots, f(x_m)).$$

- Clearly, $\mathsf{D}(g \circ f) \leq \mathsf{D}(g) + \mathsf{D}(f)$.
- KRW conjecture: $\mathsf{D}(g \circ f) \approx \mathsf{D}(g) + \mathsf{D}(f)$.

The KRW Conjecture

- [KRW91] suggested an approach.
- Let $f : \{0,1\}^n \to \{0,1\}$, $g : \{0,1\}^m \to \{0,1\}$.
- The composition $g \circ f : \{0,1\}^{m \times n} \to \{0,1\}$ is

$$(g \circ f)(x_1, \ldots, x_m) = g(f(x_1), \ldots, f(x_m)).$$

- Clearly, $\mathsf{D}(g \circ f) \leq \mathsf{D}(g) + \mathsf{D}(f)$.
- KRW conjecture: $\mathsf{D}(g \circ f) \approx \mathsf{D}(g) + \mathsf{D}(f)$.
- Implies that $\mathbf{NC}_1 \neq \mathbf{P}$.
- Compose a random function on $\log n$ bits for $\log n$ times.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.

- The problem R_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.
- The problem R_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. $x_i \neq y_i$.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.
- The problem R_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 i s.t. $x_i \neq y_i$.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.

- The problem R_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 i s.t. $x_i \neq y_i$.
  - Communicate minimal number of bits.

KW relations

- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.

- The problem R_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 i s.t. $x_i \neq y_i$.
  - Communicate minimal number of bits.

- [KW90]: $\mathsf{D}(f) = \mathsf{C}(R_f)$.

KW relations

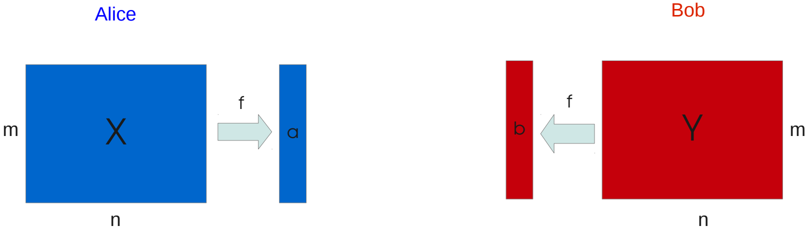
- One tool we can use is KW relations.
- Relates $\mathsf{D}(f)$ and $\mathsf{L}(f)$ to the communication complexity of a problem R_f.
- The problem R_f is defined as follows:
  - Alice gets $x \in f^{-1}(0)$.
  - Bob gets $y \in f^{-1}(1)$.
  - Clearly, $x \neq y$, so $\exists i$ s.t. $x_i \neq y_i$.
  - Want to find i s.t. $x_i \neq y_i$.
  - Communicate minimal number of bits.
- [KW90]: $\mathsf{D}(f) = \mathsf{C}(R_f)$.
- Only deterministic protocols!

KRW and KW

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

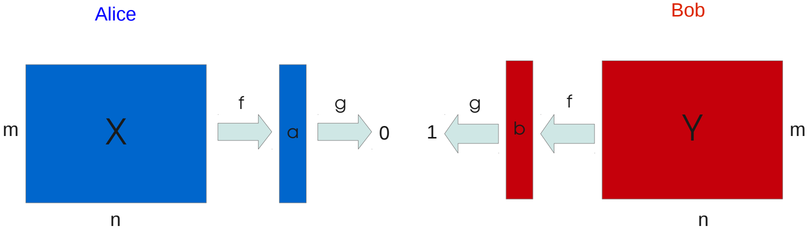
KRW and KW

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



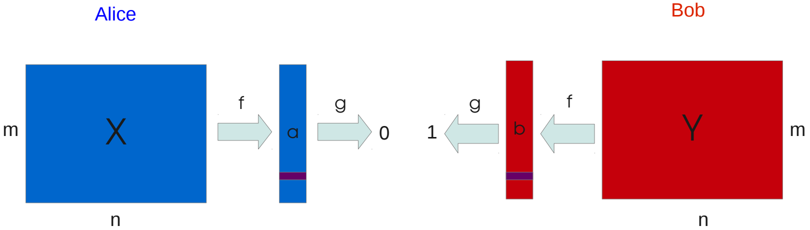
KRW and KW

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

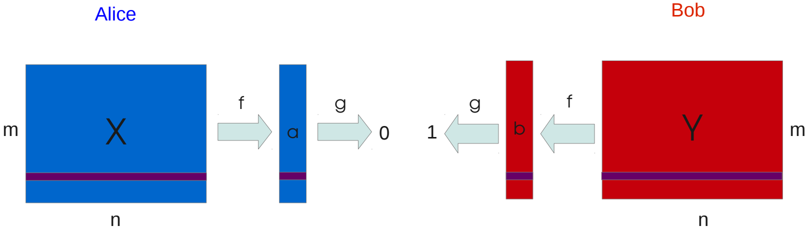


KRW and KW

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

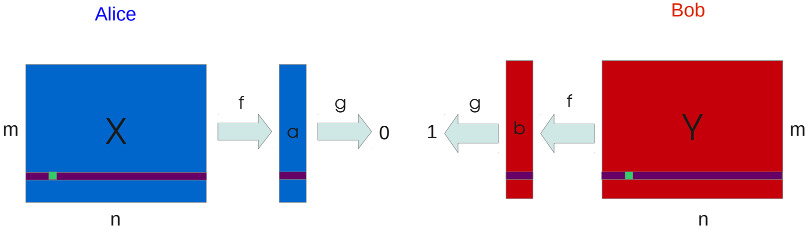


- Can we use KW relations to attack the KRW conjecture?
- How does $R_{g \circ f}$ look like?
- Recall: $g \circ f$ maps $\{0,1\}^{m \times n}$ to $\{0,1\}$.



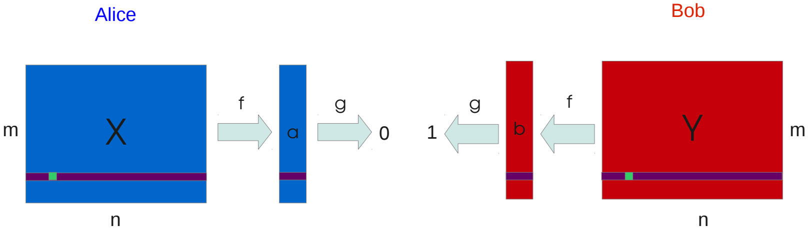
KRW and KW

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



KRW and KW

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



- KRW conjecture: the trivial protocol is essentially optimal.

The universal relation

- The KRW conjecture is hard.
- [KRW91] suggested a starting point.

The universal relation

- The KRW conjecture is hard.
- [KRW91] suggested a starting point.
- The universal relation R_{U_n} is:
  - Alice gets $x \in \{0,1\}^n$.
  - Bob gets $y \in \{0,1\}^n$.
  - $x \neq y$.
  - Wish to find i s.t. $x_i \neq y_i$.

The universal relation

- The KRW conjecture is hard.
- [KRW91] suggested a starting point.

- The universal relation R_{U_n} is:
  - Alice gets $x \in \{0,1\}^n$.
  - Bob gets $y \in \{0,1\}^n$.
  - $x \neq y$.
  - Wish to find i s.t. $x_i \neq y_i$.

- Every KW relation reduces to R_{U_n}.

The universal relation

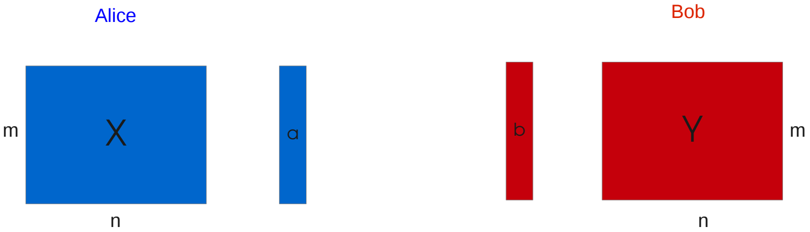
- The KRW conjecture is hard.
- [KRW91] suggested a starting point.
- The universal relation R_{U_n} is:
  - Alice gets $x \in \{0,1\}^n$.
  - Bob gets $y \in \{0,1\}^n$.
  - $x \neq y$.
  - Wish to find i s.t. $x_i \neq y_i$.
- Every KW relation reduces to R_{U_n}.
- Easy to prove: $\mathsf{C}(R_{\mathrm{U}_n}) \geq n$.

The universal relation

- The KRW conjecture is hard.
- [KRW91] suggested a starting point.
- The universal relation R_{U_n} is:
  - Alice gets $x \in \{0,1\}^n$.
  - Bob gets $y \in \{0,1\}^n$.
  - $x \neq y$.
  - Wish to find i s.t. $x_i \neq y_i$.
- Every KW relation reduces to R_{U_n}.
- Easy to prove: $\mathsf{C}(R_{\mathrm{U}_n}) \geq n$.
- [KRW91] suggested to study $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.

The composition of the universal relation

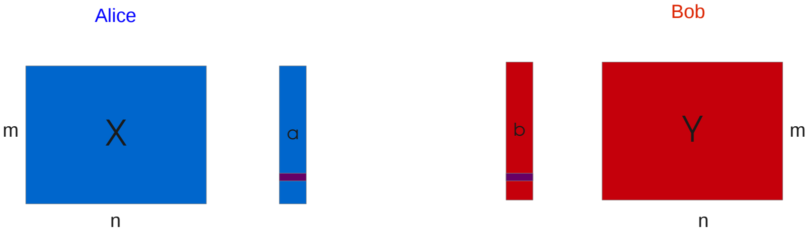
- [KRW91] suggested to study the composition $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.



- If $a_j \neq b_j$ then $X_j \neq Y_j$.

The composition of the universal relation

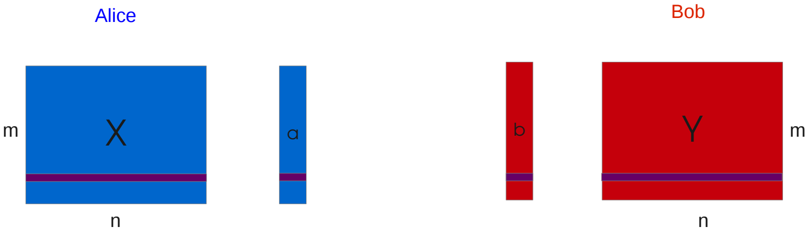
- [KRW91] suggested to study the composition $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.



- If $a_j \neq b_j$ then $X_j \neq Y_j$.

The composition of the universal relation

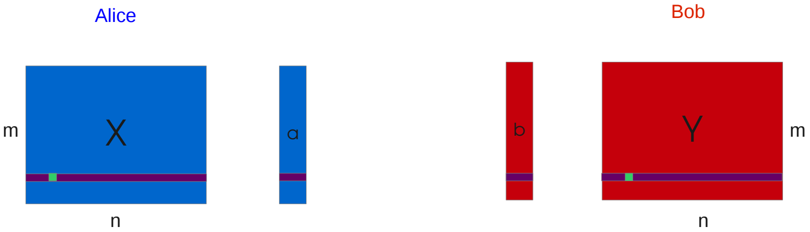
- [KRW91] suggested to study the composition $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.



- If $a_j \neq b_j$ then $X_j \neq Y_j$.

The composition of the universal relation

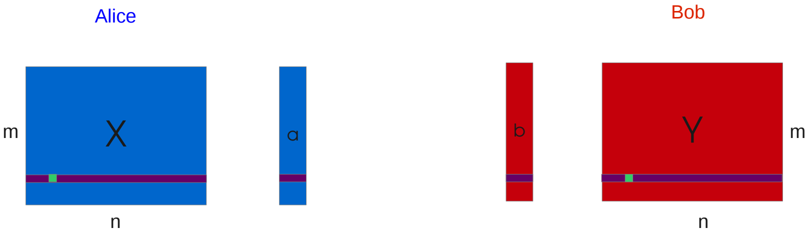
- [KRW91] suggested to study the composition $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.



- If $a_j \neq b_j$ then $X_j \neq Y_j$.

The composition of the universal relation

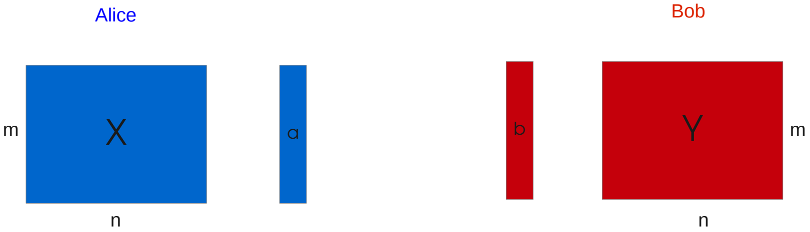
- [KRW91] suggested to study the composition $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.



- If $a_j \neq b_j$ then $X_j \neq Y_j$.
- Every KW relation $R_{g \circ f}$ reduces to $R_{\mathrm{U}_m \circ \mathrm{U}_n}$.

The composition of the universal relation

- Goal: $\mathsf{C}(R_{\mathrm{U}_m \circ \mathrm{U}_n}) \geq m + n$.
- Challenge was met by [EIRS91] and [HW93].
- To this end, they developed new techniques.



Our main result

- We analyze $R_{g \circ \mathrm{U}_n}$ for $g : \{0,1\}^m \to \{0,1\}$.

Our main result

- We analyze $R_{g \circ \mathrm{U}_n}$ for $g : \{0,1\}^m \to \{0,1\}$.
- Wish: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + n$.

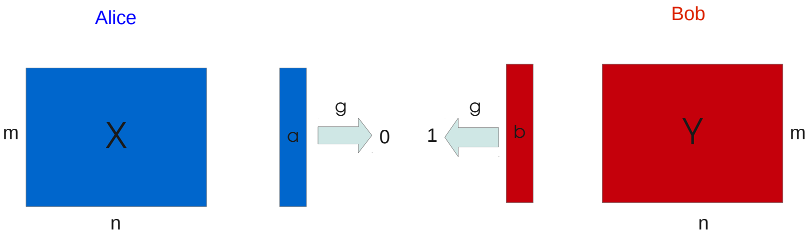
Our main result

- We analyze $R_{g \circ \mathrm{U}_n}$ for $g : \{0,1\}^m \to \{0,1\}$.
- Wish: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + n$.
- Our result: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) \geq \Omega\left(\mathsf{C}(R_g)\right) + n - O\left(\frac{m \cdot \log m}{n}\right)$.

Our main result

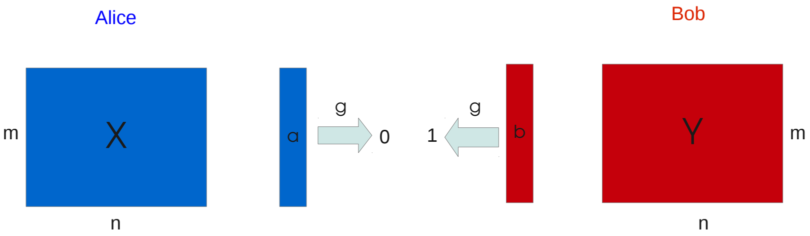
- We analyze $R_{g \circ \mathrm{U}_n}$ for $g : \{0,1\}^m \to \{0,1\}$.
- Wish: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + n$.
- Our result: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) \geq \Omega\left(\mathsf{C}(R_g)\right) + n - O\left(\frac{m \cdot \log m}{n}\right)$.
- Actually: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) \geq \log \mathsf{L}(g) + n - O\left(\frac{m \cdot \log m}{n}\right)$.

Our main result



- If $a_j \neq b_j$ then $X_j \neq Y_j$.

Our main result



- If $a_j \neq b_j$ then $X_j \neq Y_j$.
- Every KW game $R_{g \circ f}$ reduces to $R_{g \circ \mathrm{U}_n}$.

Our approach

- Our approach is based on information complexity [CSWY01, BBCR10].
- Lower bound communication complexity by analyzing the information that protocol gives on players' inputs.

Our approach

- Our approach is based on information complexity [CSWY01, BBCR10].
- Lower bound communication complexity by analyzing the information that protocol gives on players' inputs.

- $\log \mathsf{L}(g)$ can be viewed as information complexity of R_g.
- This is why we have $\log \mathsf{L}(g)$ in our bound.

Our approach

- Our approach is based on information complexity [CSWY01, BBCR10].
- Lower bound communication complexity by analyzing the information that protocol gives on players' inputs.

- $\log \mathsf{L}(g)$ can be viewed as information complexity of R_g.
- This is why we have $\log \mathsf{L}(g)$ in our bound.

- Maybe "correct" KRW conjecture is $\mathsf{L}(g \circ f) \approx \mathsf{L}(g) \cdot \mathsf{L}(f)$.

Our approach

- Wish to prove: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + \mathsf{C}(R_{\mathrm{U}_n})$.

Our approach

- Wish to prove: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + \mathsf{C}(R_{\mathrm{U}_n})$.
- Would like:
  - Must speak $\mathsf{C}(R_g)$ bits about R_g.
  - Must speak $\mathsf{C}(R_{\mathrm{U}_n})$ bits about R_{U_n}.

Our approach

- Wish to prove: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + \mathsf{C}(R_{\mathrm{U}_n})$.
- Would like:
  - Must speak $\mathsf{C}(R_g)$ bits about R_g.
  - Must speak $\mathsf{C}(R_{\mathrm{U}_n})$ bits about R_{U_n}.
- How do we perform such a decomposition?

Our approach

- Wish to prove: $\mathsf{C}(R_{g \circ \mathrm{U}_n}) = \mathsf{C}(R_g) + \mathsf{C}(R_{\mathrm{U}_n})$.
- Would like:
  - Must speak $\mathsf{C}(R_g)$ bits about R_g.
  - Must speak $\mathsf{C}(R_{\mathrm{U}_n})$ bits about R_{U_n}.
- How do we perform such a decomposition?

One key idea

When measuring information instead of communication, can use the chain rule to do the decomposition.

Other results

- Basic observations for analyzing KW relations with information complexity.

Other results

- Basic observations for analyzing KW relations with information complexity.
- Next milestone – $\oplus_m \circ f$?

Other results

- Basic observations for analyzing KW relations with information complexity.
- Next milestone – $\oplus_m \circ f$?
  - Constructing a candidate hard distribution.

Other results

- Basic observations for analyzing KW relations with information complexity.
- Next milestone – $\oplus_m \circ f$?
  - Constructing a candidate hard distribution.
  - Almost tight result for $R_{\oplus_m \circ U_n}$.

Other results

- Basic observations for analyzing KW relations with information complexity.
- Next milestone – $\oplus_m \circ f$?
  - Constructing a candidate hard distribution.
  - Almost tight result for $R_{\oplus_m \circ U_n}$.
- Alternative proof for main result using a counting argument.

Other results

- Basic observations for analyzing KW relations with information complexity.
- Next milestone – $\oplus_m \circ f$?
  - Constructing a candidate hard distribution.
  - Almost tight result for $R_{\oplus_m \circ U_n}$.
- Alternative proof for main result using a counting argument.
- Another open problem: What about $R_{U_m \circ f}$?