Revisiting Time-Space Tradeoffs for Function Inversion

Spencer Peters

Noah S.D.

Siyao Guo

Sasha Golovnev

Function Inversion

- Given a function

$$\{1, 2, \ldots, N\} =: [N] \longrightarrow \boxed{f} \longrightarrow [N]$$

Function Inversion

- Given a function, and a point y in its range,

$$\{1, 2, \ldots, N\} =: [N] \longrightarrow \boxed{f} \longrightarrow y \in [N]$$

Function Inversion

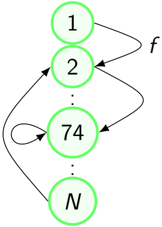
- Given a function, and a point y in its range, find x with $f(x) = y$.

$$\{1, 2, \ldots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$

Function Inversion

- Given a function, and a point y in its range, find x with $f(x) = y$.

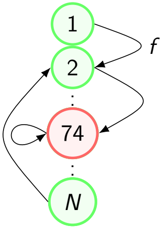
$$\{1, 2, \dots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$



Function Inversion

- Given a function, and a point y in its range, find x with $f(x) = y$.

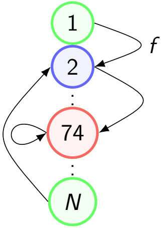
$$\{1, 2, \ldots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$



Function Inversion

- Given a function, and a point y in its range, find x with $f(x) = y$.

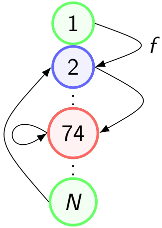
$$\{1, 2, \ldots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$



Function Inversion

- Given a function, and a point y in its range, find x with $f(x) = y$.

$$\{1, 2, \ldots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$

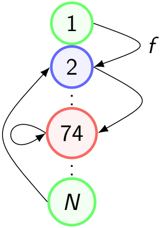


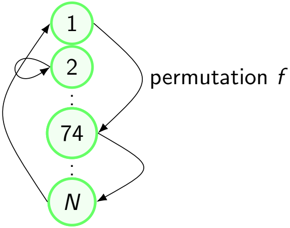
- The study of this *black-box* function inversion problem was initiated by Martin Hellman  in 1980 [Hel80].

Function Inversion

- Given a function, and a point y in its range, find x with $f(x) = y$.

$$\{1, 2, \ldots, N\} =: [N] \ni x \longrightarrow \boxed{f} \longrightarrow y \in [N]$$

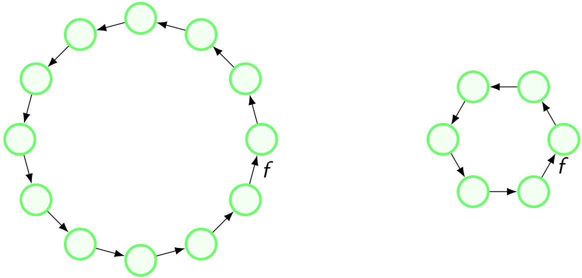




- The study of this *black-box* function inversion problem was initiated by Martin Hellman  in 1980 [Hel80].

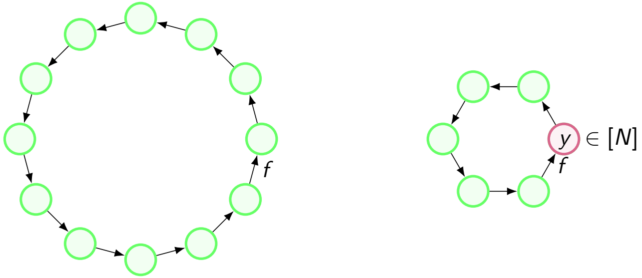
Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.



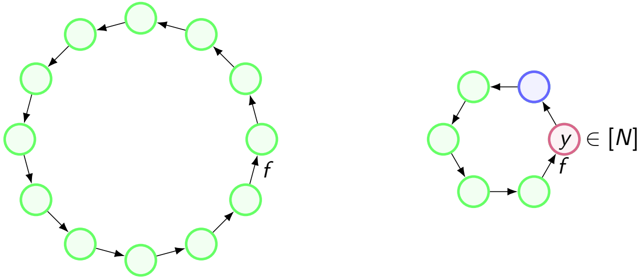
Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.



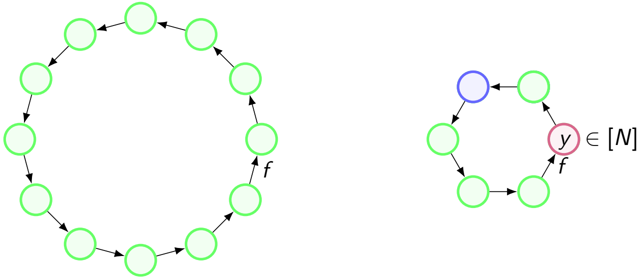
Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.



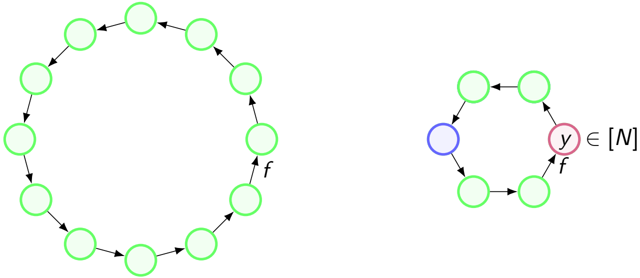
Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.



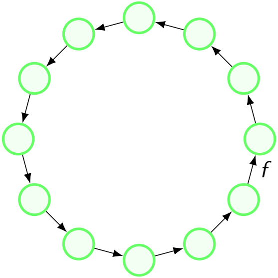
Hellman's algorithm

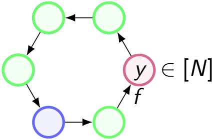
- If f is a permutation, its *graph* is a disjoint union of cycles.



Hellman's algorithm

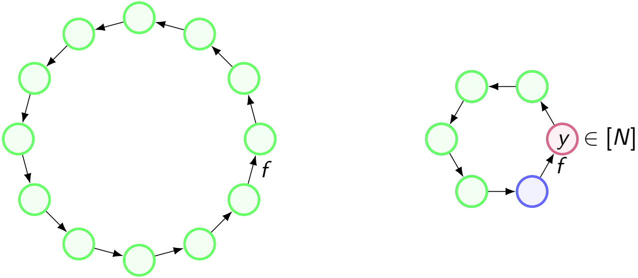
- If f is a permutation,
  its *graph* is a disjoint union of cycles.





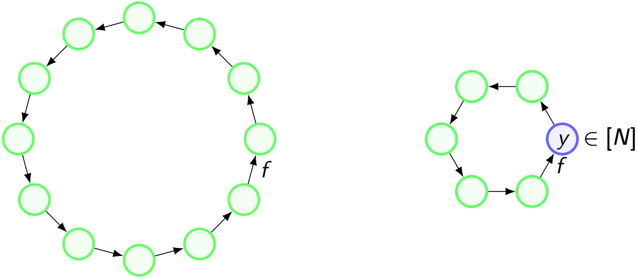
Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.



Hellman's algorithm

- If f is a permutation,
  its *graph* is a disjoint union of cycles.

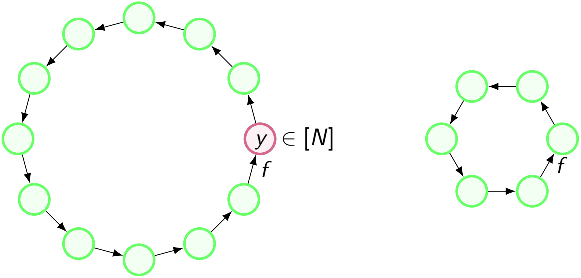


Hellman's algorithm

- What if y is on a large cycle?

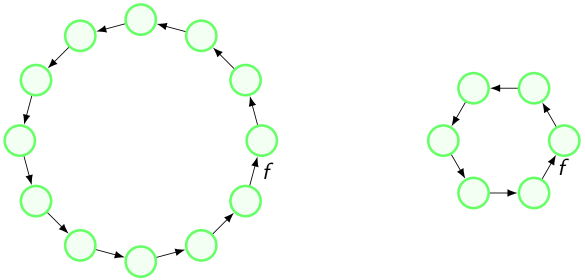
Hellman's algorithm

- What if y is on a large cycle?



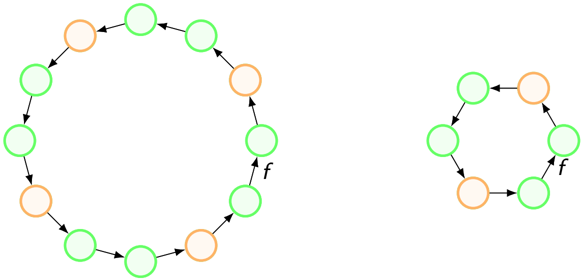
Hellman's algorithm

- What if y is on a large cycle?



Hellman's algorithm

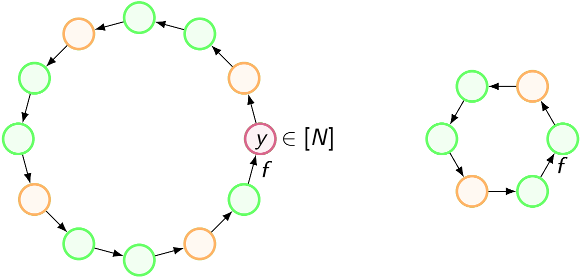
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

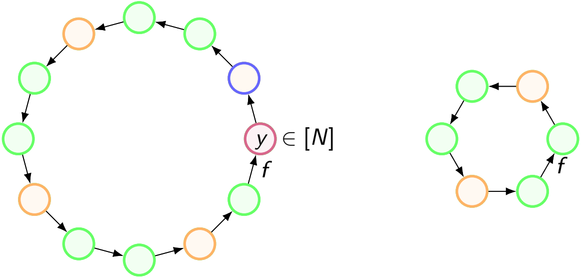
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

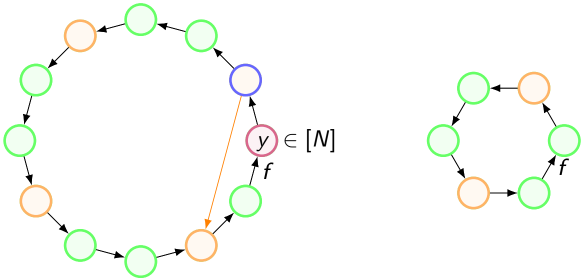
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

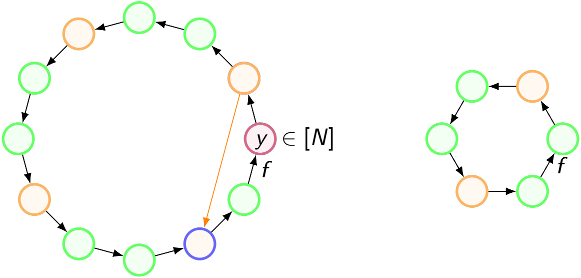
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

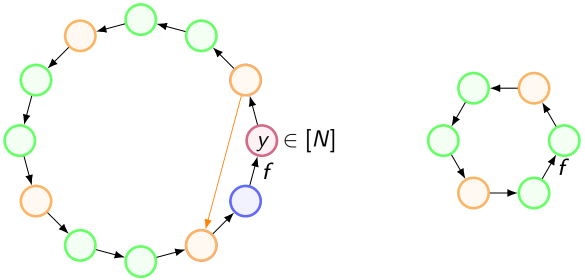
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

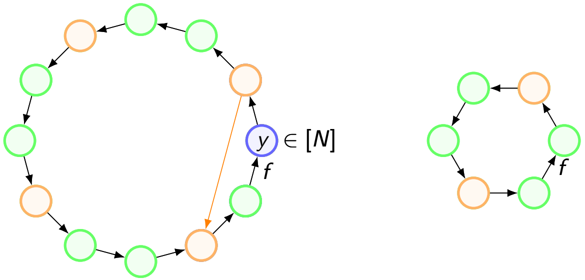
- What if y is on a large cycle?



- In a preprocessing step, store points uniformly spaced around each cycle.

Hellman's algorithm

- What if y is on a large cycle?



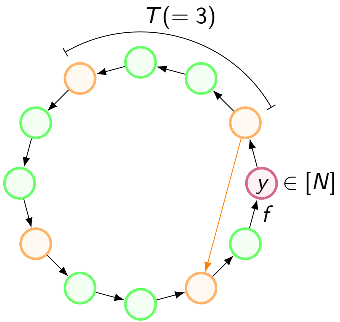
- In a preprocessing step, store points uniformly spaced around each cycle.

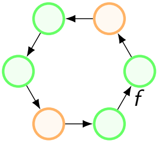
Analysis

- If we space the stored points T hops apart:

Analysis

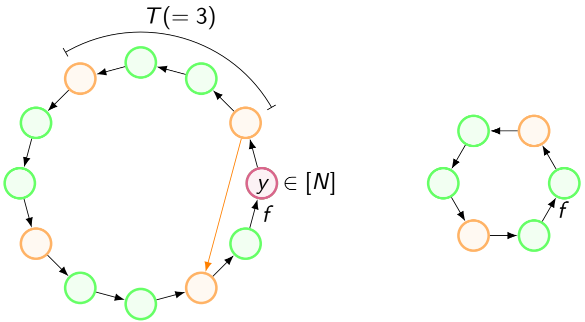
- If we space the stored points T hops apart:





Analysis

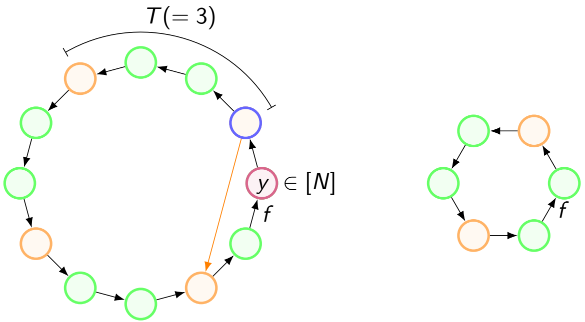
- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.

Analysis

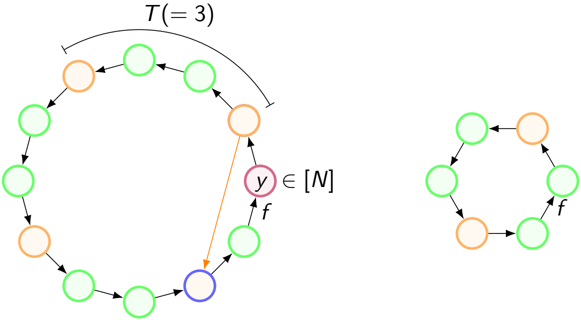
- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.

Analysis

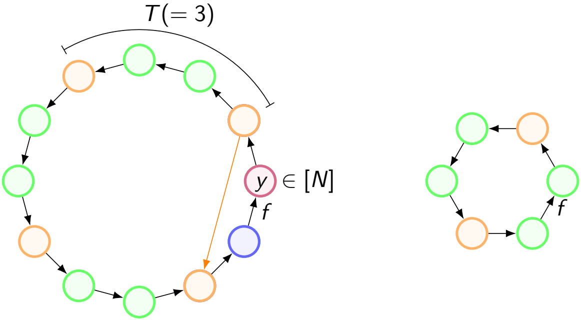
- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.

Analysis

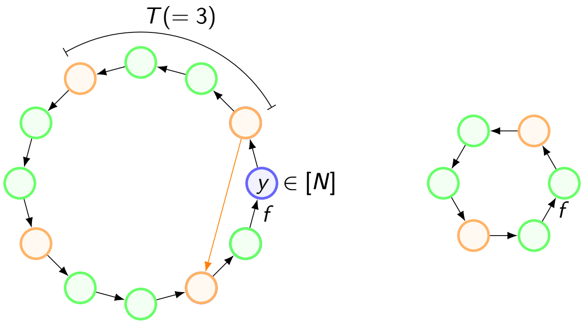
- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.

Analysis

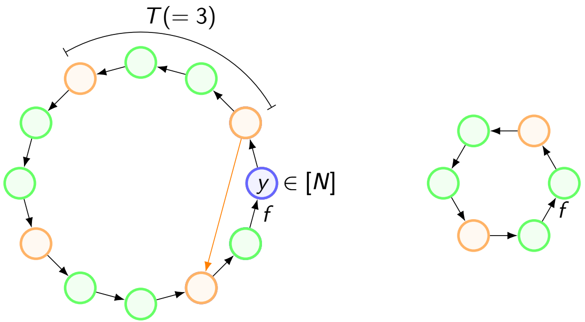
- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.

Analysis

- If we space the stored points T hops apart:



- We need T evaluations of f to invert y.
- We need to store about N/T points total.

Stepping back

- **Goal:** design a pair of algorithms $(\mathcal{P}, \mathcal{A})$ such that

$$\Pr[\alpha \leftarrow \mathcal{P}(f); x \leftarrow \mathcal{A}^f(\alpha, y); f(x) = y] \geq 9/10.$$

Stepping back

- **Goal:** design a pair of algorithms $(\mathcal{P}, \mathcal{A})$ such that

$$\Pr[\alpha \leftarrow \mathcal{P}(f); x \leftarrow \mathcal{A}^f(\alpha, y); f(x) = y] \geq 9/10.$$

- In this model, $\mathcal{P}$ and $\mathcal{A}$ have **unbounded** computational power.

Stepping back

- **Goal:** design a pair of algorithms $(\mathcal{P}, \mathcal{A})$ such that

$$\Pr[\alpha \leftarrow \mathcal{P}(f); x \leftarrow \mathcal{A}^f(\alpha, y); f(x) = y] \geq 9/10.$$

- In this model, $\mathcal{P}$ and $\mathcal{A}$ have **unbounded** computational power.
- We aim to **minimize** the bitlength S of α, and the number of queries T that $\mathcal{A}$ makes to f.

Application scenarios

- The NSA wants to break cryptography based on a widely used cryptographic function, such as AES-128.

Application scenarios

- The NSA wants to break cryptography based on a widely used cryptographic function, such as AES-128.
- Hackers want to recover passwords from a stolen database of password hashes (Rainbow Tables)

Application scenarios

- The NSA wants to break cryptography based on a widely used cryptographic function, such as AES-128.
- Hackers want to recover passwords from a stolen database of password hashes (Rainbow Tables)
- Theoretical computer scientists want better algorithms for 3-SUM [GGH+20],     

Application scenarios

- The NSA wants to break cryptography based on a widely used cryptographic function, such as AES-128.
- Hackers want to recover passwords from a stolen database of password hashes (Rainbow Tables)
- Theoretical computer scientists want better algorithms for 3-SUM [GGH$^+$20],     

  multiparty pointer jumping [CK19],  

Application scenarios

- The NSA wants to break cryptography based on a widely used cryptographic function, such as AES-128.
- Hackers want to recover passwords from a stolen database of password hashes (Rainbow Tables)
- Theoretical computer scientists want better algorithms for
  3-SUM [GGH+20],     
  multiparty pointer jumping [CK19],  
  systematic substring search [CK19], ...

Beyond Permutations

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 	1980	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 	1980	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor  	1991	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 	1980	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor  	1991	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$

- Q: Can we improve Fiat-Naor? Can we improve Yao's lower bound?

Beyond Permutations

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 	1980	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor  	1991	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$

- Q: Can we improve Fiat-Naor? Can we improve Yao's lower bound?
- A: Sort of and sort of!

Our Results

- Result 1: A simple improvement
  to Fiat and Naor's algorithm in the regime $T > S$.
- Result 2: A tight lower bound for a natural class of
  non-adaptive function inversion algorithms.
- Not in this talk: equivalences between variants of function
  inversion.

Result 1: Improving Fiat-Naor

Result		Applies To	Tradeoff	Key Point
Hellman 	1980	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 		permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 	1980	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor  	1991	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$

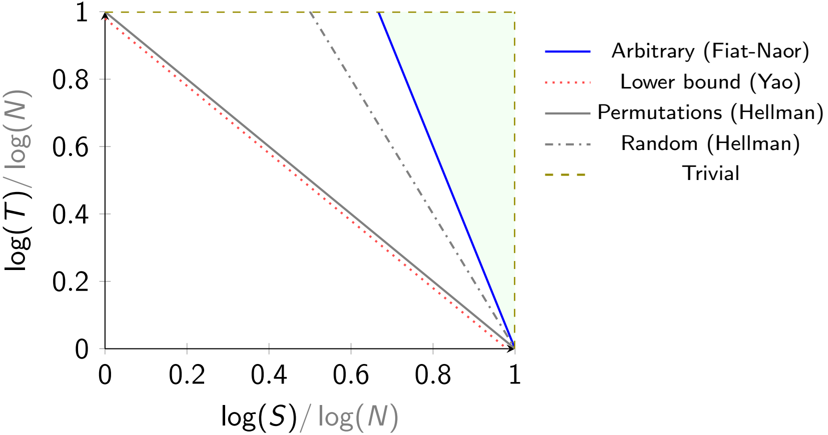
Result 1: Improving Fiat-Naor

Result	Applies To	Tradeoff	Key Point
Hellman 1980 	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 	permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 1980 	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor 1991  	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$
This work	all functions	$T \lesssim N^3/(S^2 T)$	$S = T \lesssim N^{3/4}$

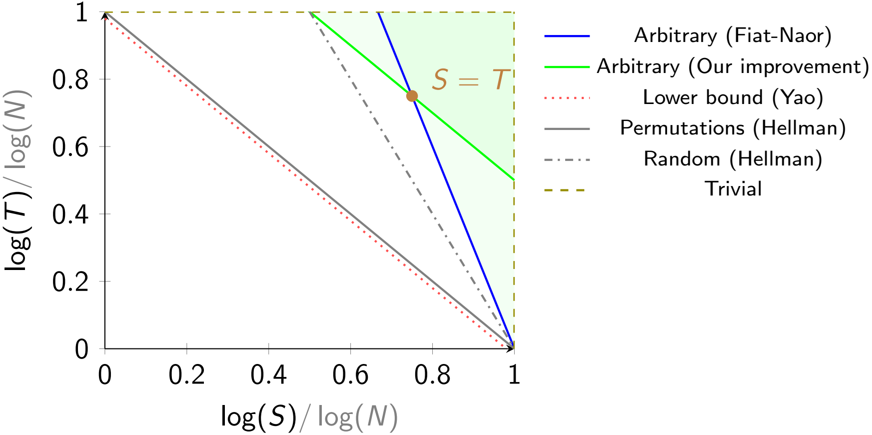
Result 1: Improving Fiat-Naor

Result	Applies To	Tradeoff	Key Point
Hellman 1980 	permutations	$T \lesssim N/S$	$S = T \lesssim \sqrt{N}$
Yao 1990 	permutations	$T \gtrsim N/S$	$S = T \gtrsim \sqrt{N}$
Hellman 1980 	random f	$T \lesssim N^2/S^2$	$S = T \lesssim N^{2/3}$
Fiat-Naor 1991  	all functions	$T \lesssim N^3/S^3$	$S = T \lesssim N^{3/4}$
This work	all functions	$T \lesssim N^3/(S^2 T)$ $T \lesssim N^{3/2}/S$	$S = T \lesssim N^{3/4}$

Result 1: Improving Fiat-Naor



Result 1: Improving Fiat-Naor

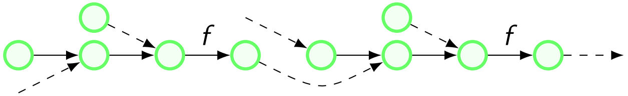


Background: beyond permutations

- Preprocessing stores the endpoints of disjoint paths.

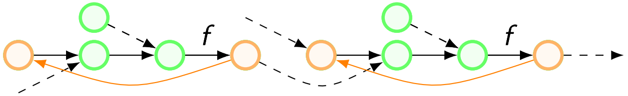
Background: beyond permutations

- Preprocessing stores the endpoints of disjoint paths.



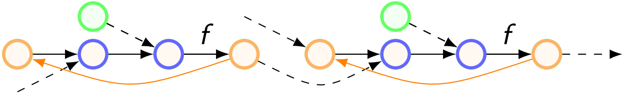
Background: beyond permutations

- Preprocessing stores the endpoints of disjoint paths.



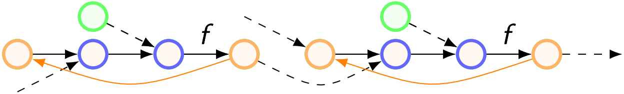
Background: beyond permutations

- Preprocessing stores the endpoints of disjoint paths.



Background: beyond permutations

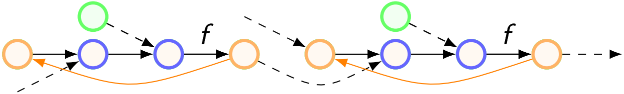
- Preprocessing stores the endpoints of disjoint paths.



- No longer possible to cover the entire range.

Background: beyond permutations

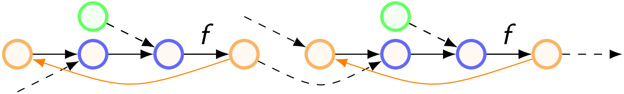
- Preprocessing stores the endpoints of disjoint paths.



- No longer possible to cover the entire range.
- But, can still cover a small fraction with disjoint paths.

Background: beyond permutations

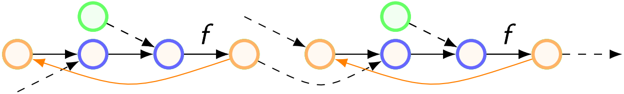
- Preprocessing stores the endpoints of disjoint paths.



- No longer possible to cover the entire range.
- But, can still cover a small fraction with disjoint paths.
- To boost the coverage, observe that inverting $g \circ f$ on $g(y)$ is often enough to invert f on y.

Background: beyond permutations

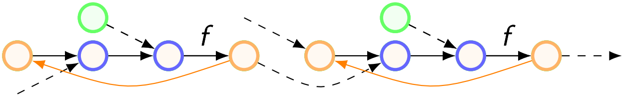
- Preprocessing stores the endpoints of disjoint paths.



- No longer possible to cover the entire range.
- But, can still cover a small fraction with disjoint paths.
- To boost the coverage, observe that inverting $g \circ f$ on $g(y)$ is often enough to invert f on y.
- So, can repeatedly apply the basic scheme to many compositions $g_i \circ f$, for suitably chosen "rerandomization" functions g_i.

Background: beyond permutations

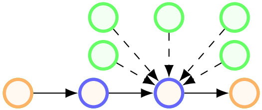
- Preprocessing stores the endpoints of disjoint paths.



- No longer possible to cover the entire range.
- But, can still cover a small fraction with disjoint paths.
- To boost the coverage, observe that inverting $g \circ f$ on $g(y)$ is often enough to invert f on y.
- So, can repeatedly apply the basic scheme to many compositions $g_i \circ f$, for suitably chosen "rerandomization" functions g_i.
- For *random* functions, Hellman showed (heuristically) this can be made to work. 

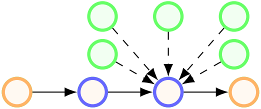
Background: Fiat-Naor

- Hellman's argument fails for arbitrary functions.



Background: Fiat-Naor

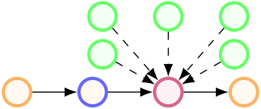
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.

Background: Fiat-Naor

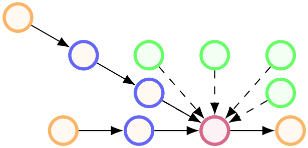
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.

Background: Fiat-Naor

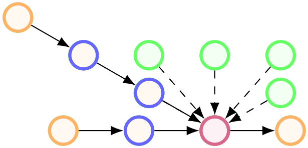
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.

Background: Fiat-Naor

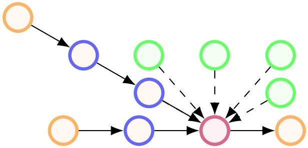
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.
- Paths collide at these points, causing all sorts of problems.

Background: Fiat-Naor

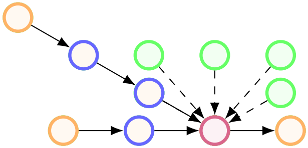
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.
- Paths collide at these points, causing all sorts of problems.
- Fiat and Naor deal with this by storing $\alpha = (\alpha', L)$, where L contains junction points along with their inverses.

Background: Fiat-Naor

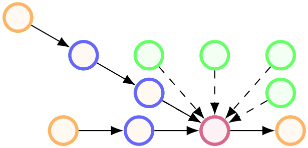
- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.
- Paths collide at these points, causing all sorts of problems.
- Fiat and Naor deal with this by storing $\alpha = (\alpha', L)$, where L contains junction points along with their inverses.
- Intuitively, α' is the data structure for a *restriction* of f that avoids the junction points in L.

Background: Fiat-Naor

- Hellman's argument fails for arbitrary functions.



- Arbitrary functions can have "junction points" with many inverses.
- Paths collide at these points, causing all sorts of problems.
- Fiat and Naor deal with this by storing $\alpha = (\alpha', L)$, where L contains junction points along with their inverses.
- Intuitively, α' is the data structure for a *restriction* of f that avoids the junction points in L.
- More precisely, the "rerandomization" functions are sampled using rejection sampling so that their range is $[N] - L$.

Our Improvement (1)

- Recall that Fiat and Naor's preprocessing and online algorithms must agree on a list L of "junction points".

Our Improvement (1)

- Recall that Fiat and Naor's preprocessing and online algorithms must agree on a list L of "junction points".
- We observe that the tradeoff $T \lesssim N^3/S^3$ comes from

$$T \lesssim \frac{1}{|L|} \cdot \frac{N^3}{S^2}.$$

Our Improvement (1)

- Recall that Fiat and Naor's preprocessing and online algorithms must agree on a list L of "junction points".
- We observe that the tradeoff $T \lesssim N^3/S^3$ comes from

$$T \lesssim \frac{1}{|L|} \cdot \frac{N^3}{S^2}.$$

- Fiat and Naor get $|L| \simeq S$, but this is the hard limit, since L needs to fit into S-bit advice α.

Our Improvement (1)

- Recall that Fiat and Naor's preprocessing and online algorithms must agree on a list L of "junction points".
- We observe that the tradeoff $T \lesssim N^3/S^3$ comes from

$$T \lesssim \frac{1}{|L|} \cdot \frac{N^3}{S^2}.$$

- Fiat and Naor get $|L| \simeq S$, but this is the hard limit, since L needs to fit into S-bit advice α.
- Or does it?

Our Improvement (2)

- Fiat and Naor's list L actually consists of images $f(x_i)$ of random points $x_i \sim [N]$.

Our Improvement (2)

- Fiat and Naor's list L actually consists of images $f(x_i)$ of random points $x_i \sim [N]$.
- Our idea: Instead of reading L from α, $\mathcal{A}$ recovers L by evaluating f on the same random points x_i.

Our Improvement (2)

- Fiat and Naor's list L actually consists of images $f(x_i)$ of random points $x_i \sim [N]$.
- Our idea: Instead of reading L from α, $\mathcal{A}$ recovers L by evaluating f on the same random points x_i.
- This allows $|L| \simeq T$, so we can get $T \lesssim N^3/(S^2 T)$, or

$$T \lesssim N^{3/2}/S.$$

Our Improvement (2)

- Fiat and Naor's list L actually consists of images $f(x_i)$ of random points $x_i \sim [N]$.
- Our idea: Instead of reading L from α, $\mathcal{A}$ recovers L by evaluating f on the same random points x_i.
- This allows $|L| \simeq T$, so we can get $T \lesssim N^3/(S^2 T)$, or

$$T \lesssim N^{3/2}/S.$$

- That's it!
- But I've cheated here...

Our Improvement (2)

- Fiat and Naor's list L actually consists of images $f(x_i)$ of random points $x_i \sim [N]$.
- Our idea: Instead of reading L from α, $\mathcal{A}$ recovers L by evaluating f on the same random points x_i.
- This allows $|L| \simeq T$, so we can get $T \lesssim N^3/(S^2 T)$, or

$$T \lesssim N^{3/2}/S.$$

- That's it!
- But I've cheated here...
- How do $\mathcal{A}$ and $\mathcal{P}$ agree on the *same* list of random values x_i?

Along the Way: Shared Randomness

- We show that, in the preprocessing model,
  one can assume *shared randomness* without loss of generality.

Along the Way: Shared Randomness

- We show that, in the preprocessing model,
one can assume *shared randomness* without loss of generality.
- The proof adapts *Newman's lemma* [New91]  from communication complexity.

Along the Way: Shared Randomness

- We show that, in the preprocessing model,
  one can assume *shared randomness* without loss of generality.
- The proof adapts *Newman's lemma* [New91]  from communication complexity.
- In practice, can instantiate a random oracle.

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.
- Corrigan-Gibbs and Kogan   [CK19]: any small improvement $\implies$ new lower bounds in circuit complexity.

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.
- Corrigan-Gibbs and Kogan   [CK19]: any small improvement $\implies$ new lower bounds in circuit complexity.
- Even improving Yao's bound just for *non-adaptive* algorithms would do it!

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.
- Corrigan-Gibbs and Kogan   [CK19]: any small improvement $\implies$ new lower bounds in circuit complexity.
- Even improving Yao's bound just for *non-adaptive* algorithms would do it!
- $\mathcal{A}$ is non-adaptive if its evaluation points $x_1, \ldots, x_T$ are chosen up front, before any evaluations of f are seen.

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.
- Corrigan-Gibbs and Kogan   [CK19]: any small improvement $\implies$ new lower bounds in circuit complexity.
- Even improving Yao's bound just for *non-adaptive* algorithms would do it!
- $\mathcal{A}$ is non-adaptive if its evaluation points $x_1, \ldots, x_T$ are chosen up front, before any evaluations of f are seen.
- Non-adaptive algorithms seem very weak. Hellman's algorithm is *very* adaptive.

Result 2: Background

- Recall that Yao's lower bound (for inverting *arbitrary* functions) hasn't been improved in 30+ years.
- Corrigan-Gibbs and Kogan   [CK19]: any small improvement $\implies$ new lower bounds in circuit complexity.
- Even improving Yao's bound just for *non-adaptive* algorithms would do it!
- $\mathcal{A}$ is non-adaptive if its evaluation points $x_1, \ldots, x_T$ are chosen up front, before any evaluations of f are seen.
- Non-adaptive algorithms seem very weak. Hellman's algorithm is *very* adaptive.
- Corrigan-Gibbs and Kogan speculated that there is no non-adaptive algorithm with

$$S = o(N \log N) \text{ and } T = o(N).$$

A Lower Bound

- We observe that there IS in fact a very simple algorithm, that (barely!) outperforms the trivial inverter.

A Lower Bound

- We observe that there IS in fact a very simple algorithm, that (barely!) outperforms the trivial inverter.
- For each range element, preprocessing stores a $(\log(N) - \log(T))$-bit prefix of one of its inverses.

A Lower Bound

- We observe that there IS in fact a very simple algorithm, that (barely!) outperforms the trivial inverter.
- For each range element, preprocessing stores a $(\log(N) - \log(T))$-bit prefix of one of its inverses.
- This implicitly defines T candidate inverses for the online algorithm to check, achieving the tradeoff $S = O(N \log(N/T))$.

A Lower Bound

- We observe that there IS in fact a very simple algorithm, that (barely!) outperforms the trivial inverter.
- For each range element, preprocessing stores a $(\log(N) - \log(T))$-bit prefix of one of its inverses.
- This implicitly defines T candidate inverses for the online algorithm to check, achieving the tradeoff $S = O(N \log(N/T))$.
- When the online algorithm just (non-adaptively) checks T candidate inverses determined by the preprocessing α and the challenge y, we call it a *guess-and-check* algorithm.

A Lower Bound

- We observe that there IS in fact a very simple algorithm, that (barely!) outperforms the trivial inverter.
- For each range element, preprocessing stores a $(\log(N) - \log(T))$-bit prefix of one of its inverses.
- This implicitly defines T candidate inverses for the online algorithm to check, achieving the tradeoff $S = O(N \log(N/T))$.
- When the online algorithm just (non-adaptively) checks T candidate inverses determined by the preprocessing α and the challenge y, we call it a *guess-and-check* algorithm.
- We show that **the simple algorithm above is asymptotically optimal among guess-and-check algorithms**.

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

$$\begin{aligned} &\implies S + N \log T = \Omega(N \log N) \\ &\implies S = \Omega(N \log(N/T)). \end{aligned}$$

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

$$\begin{aligned} &\implies S + N \log T = \Omega(N \log N) \\ &\implies S = \Omega(N \log(N/T)). \end{aligned}$$

- Proof:
  - Encoder computes $\alpha \leftarrow \mathcal{P}(f)$.

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

$$\begin{aligned} &\implies S + N \log T = \Omega(N \log N) \\ &\implies S = \Omega(N \log(N/T)). \end{aligned}$$

- Proof:
  - Encoder computes $\alpha \leftarrow \mathcal{P}(f)$.
  - For each $y \in [N]$, encoder runs $\mathcal{A}(\alpha, y)$ and receives $x_1, \ldots, x_T$. It writes down the $i_y \in [T]$ that satisfies $f(x_{i_y}) = y$.

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

$$\implies S + N \log T = \Omega(N \log N)$$
$$\implies S = \Omega(N \log(N/T)).$$

- Proof:
  - Encoder computes $\alpha \leftarrow \mathcal{P}(f)$.
  - For each $y \in [N]$, encoder runs $\mathcal{A}(\alpha, y)$ and receives $x_1, \ldots, x_T$. It writes down the $i_y \in [T]$ that satisfies $f(x_{i_y}) = y$.
  - Encoding is $(\alpha, i_1, \ldots, i_N)$.

A Lower Bound

- The proof is a compression argument. For simplicity, assume there exists a guess-and-check algorithm that always succeeds, with parameters S, T. Then
- Theorem: Can encode any permutation f using $S + N \log T$ bits.

$$\begin{aligned} &\implies S + N \log T = \Omega(N \log N) \\ &\implies S = \Omega(N \log(N/T)). \end{aligned}$$

- Proof:
  - Encoder computes $\alpha \leftarrow \mathcal{P}(f)$.
  - For each $y \in [N]$, encoder runs $\mathcal{A}(\alpha, y)$ and receives $x_1, \dots, x_T$. It writes down the $i_y \in [T]$ that satisfies $f(x_{i_y}) = y$.
  - Encoding is $(\alpha, i_1, \dots, i_N)$.
  - For each y, decoder again runs $\mathcal{A}(\alpha, y)$ and receives $x_1, \dots, x_T$. It sets $f^{-1}(y) = x_{i_y}$.

Open Problems

- In some sense, only one open problem–

Open Problems

- In some sense, only one open problem–**close the gap!**

Open Problems

- In some sense, only one open problem–**close the gap!**
- Moonshots:
  - Improve Yao's lower bound against (general) non-adaptive algorithms?

Open Problems

- In some sense, only one open problem–**close the gap!**
- Moonshots:
  - Improve Yao's lower bound against (general) non-adaptive algorithms?
  - $S^2 T = N^2$ algorithm for worst-case function inversion?

Open Problems

- In some sense, only one open problem–**close the gap!**
- Moonshots:
  - Improve Yao's lower bound against (general) non-adaptive algorithms?
  - $S^2 T = N^2$ algorithm for worst-case function inversion?
- Possibly more tractable:
  - Better algorithms for inverting a small fraction of the range? That is, improving on De, Trevisan and Tulsiani [DTT10]?

Open Problems

- In some sense, only one open problem–**close the gap!**
- Moonshots:
  - Improve Yao's lower bound against (general) non-adaptive algorithms?
  - $S^2 T = N^2$ algorithm for worst-case function inversion?
- Possibly more tractable:
  - Better algorithms for inverting a small fraction of the range? That is, improving on De, Trevisan and Tulsiani [DTT10]?
  - Better algorithms for inverting injective functions?

Thank you!

I'm happy to take additional questions offline.

You can ping me at `speters@cs.cornell.edu`.

References I

Henry Corrigan-Gibbs and Dmitry Kogan.
The function-inversion problem: Barriers and opportunities.
In *TCC*, 2019.

Anindya De, Luca Trevisan, and Madhur Tulsiani.
Time space tradeoffs for attacks against one-way functions and PRGs.
In *CRYPTO*, 2010.

Alexander Golovnev, Siyao Guo, Thibaut Horel, Sunoo Park, and Vinod Vaikuntanathan.
Data structures meet cryptography: 3SUM with preprocessing.
In *STOC*, 2020.

References II

Alexander Golovnev, Siyao Guo, Spencer Peters, and Noah Stephens-Davidowitz.
Revisiting time-space tradeoffs for function inversion.
Available at https://eccc.weizmann.ac.il/report/2022/145/, 2022.

Martin Hellman.
A cryptanalytic time-memory trade-off.
IEEE Transactions on Information Theory, 26(4):401–406, 1980.

Ilan Newman.
Private vs. common random bits in communication complexity.
Information processing letters, 39(2):67–71, 1991.