Finding monotone patterns

Shoham Letzter ETH–ITS

joint with Omri Ben-Eliezer, Clément Canonne, Erik Waingarten

Mittagsseminar

December 2019



Property testing



- has property \mathcal{P} ,
- or is far from having property \mathcal{P} .

- has property \mathcal{P} ,
- or is far from having property \mathcal{P} .

E.g. determine, with high probability, if a graph *G* is bipartite, or cannot be made bipartite by removing at most $\varepsilon |G|^2$ edges.

- has property \mathcal{P} ,
- or is far from having property \mathcal{P} .

E.g. determine, with high probability, if a graph *G* is bipartite, or cannot be made bipartite by removing at most $\varepsilon |G|^2$ edges. (This can be done in time O(1).)

- has property \mathcal{P} ,
- or is far from having property \mathcal{P} .

E.g. determine, with high probability, if a graph *G* is bipartite, or cannot be made bipartite by removing at most $\varepsilon |G|^2$ edges. (This can be done in time O(1).)

We consider testing with **one-sided error**:

- has property \mathcal{P} ,
- or is far from having property \mathcal{P} .

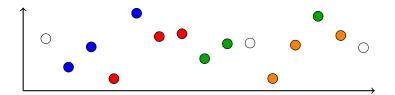
E.g. determine, with high probability, if a graph *G* is bipartite, or cannot be made bipartite by removing at most $\varepsilon |G|^2$ edges. (This can be done in time O(1).)

We consider testing with **one-sided error**: if an object is far from having \mathcal{P} , provide evidence.

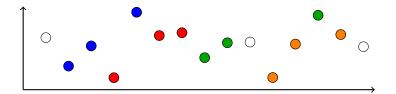
Testing for (1...k)-freeness



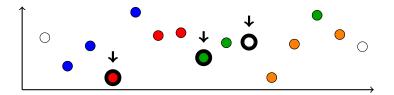
Input. $f : [n] \to \mathbb{R}$ with $\Omega(n)$ disjoint increasing k-tuples.



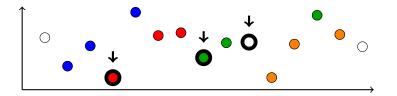
- **Input.** $f : [n] \to \mathbb{R}$ with $\Omega(n)$ disjoint increasing k-tuples.
- Aim. Find, with high probability, an increasing k-tuple.



- **Input.** $f : [n] \to \mathbb{R}$ with $\Omega(n)$ disjoint increasing k-tuples.
- Aim. Find, with high probability, an increasing k-tuple.



- Input. $f : [n] \to \mathbb{R}$ with $\Omega(n)$ disjoint increasing *k*-tuples.
- **Aim.** Find, with high probability, an increasing *k*-tuple.



We sometimes refer to an increasing k-tuple as a (1...k)-copy.

History





Ergün, Kannan, Kumar, Rubinfeld, Viswanathan '98. Optimal non-adaptive monotonicity testers make $\Theta(\log n)$ queries. (non-adaptivity: queries do not depend on previous outcomes.)

Ergün, Kannan, Kumar, Rubinfeld, Viswanathan '98. Optimal non-adaptive monotonicity testers make $\Theta(\log n)$ queries. (non-adaptivity: queries do not depend on previous outcomes.)

Fischer '09. Adaptivity does not help monotonicity testing!

Ergün, Kannan, Kumar, Rubinfeld, Viswanathan '98. Optimal non-adaptive monotonicity testers make $\Theta(\log n)$ queries. (non-adaptivity: queries do not depend on previous outcomes.)

Fischer '09. Adaptivity does not help monotonicity testing!

Newman, Rabinovich, Rajendraprasad, Sohler '17. For $k \ge 2$, there is a (non-adaptive) tester which makes $(\log n)^{O(k^2)}$ queries.

Our results



Theorem (Ben-Eliezer, Canonne, L., Waingarten)

An optimal non-adaptive algorithm for testing (1...k)-freeness makes $\Theta_k((\log n)^{\lfloor \log_2 k \rfloor})$ queries.



Theorem (Ben-Eliezer, Canonne, L., Waingarten)

An optimal non-adaptive algorithm for testing (1...k)-freeness makes $\Theta_k((\log n)^{\lfloor \log_2 k \rfloor})$ queries.

Theorem (Ben-Eliezer, L., Waingarten)

An optimal adaptive algorithm for testing (1...k)-freeness makes $\Theta_k(\log n)$ queries.

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$.



For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(x, y) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$.
E.g. If $\begin{array}{l} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$,

For $x, y \in \mathbb{N}$: $b(x, y) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$. **E.g.** If $\begin{array}{l} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$, then b(x, y) = 3.

For $x, y \in \mathbb{N}$: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$.

E.g. If $\begin{array}{l} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$, then b(x, y) = 3.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$.

E.g. If
$$\begin{array}{c} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max\left\{i : \begin{array}{l} ext{the binary representations} \\ ext{of } x ext{ and } y ext{ differ in } i^{ ext{th digit}} \end{array}\right\}$.

E.g. If
$$\begin{array}{c} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}$.

Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

Proof.

Let $i = \max b(A)$.



For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max \left\{ i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array} \right\}$.

E.g. If
$$\begin{array}{c} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

Let
$$i = \max b(A)$$
. For $\sigma \in \{0, 1\}$, define
 $A_{\sigma} = \{x \in A : \text{the } i^{\text{th}} \text{ digit of } x \text{ in binary is } \sigma\}.$

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max\left\{i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array}\right\}$.

E.g. If
$$\begin{array}{l} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

Let
$$i = \max b(A)$$
. For $\sigma \in \{0, 1\}$, define
 $A_{\sigma} = \{x \in A : \text{the } i^{\text{th}} \text{ digit of } x \text{ in binary is } \sigma\}.$
 $|b(A)| \le |b(A_0)| + |b(A_1)| + 1$

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max\left\{i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array}\right\}$.

E.g. If
$$\begin{array}{c} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

Let
$$i = \max b(A)$$
. For $\sigma \in \{0, 1\}$, define
 $A_{\sigma} = \{x \in A : \text{the } i^{\text{th}} \text{ digit of } x \text{ in binary is } \sigma\}.$
 $|b(A)| \le |b(A_0)| + |b(A_1)| + 1 \le |A_0| - 1 + |A_1| - 1 + 1$

For
$$x, y \in \mathbb{N}$$
: $\boldsymbol{b}(\boldsymbol{x}, \boldsymbol{y}) = \max\left\{i : \begin{array}{l} \text{the binary representations} \\ \text{of } x \text{ and } y \text{ differ in } i^{\text{th}} \text{ digit} \end{array}\right\}$.

E.g. If
$$\begin{array}{l} x = 11 = 1011_{\text{bin}} \\ y = 13 = 1101_{\text{bin}} \end{array}$$
, then $b(x, y) = 3$.

For $A \subseteq \mathbb{N}$: $b(A) := \{b(x, y) : x, y \in A\}.$

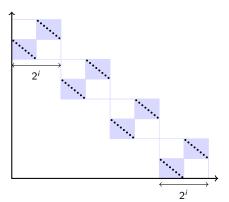
Claim

$$|b(A)| \leq |A| - 1$$
 for every finite $A \subseteq \mathbb{N}$.

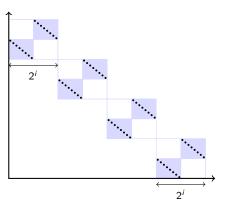
Let
$$i = \max b(A)$$
. For $\sigma \in \{0, 1\}$, define
 $A_{\sigma} = \{x \in A : \text{the } i^{\text{th}} \text{ digit of } x \text{ in binary is } \sigma\}.$
 $|b(A)| \le |b(A_0)| + |b(A_1)| + 1 \le |A_0| - 1 + |A_1| - 1 + 1 = |A| - 1.$



For $i \in [\log n]$, define $f_i : [0, n-1] \rightarrow [0, n-1]$ by



For $i \in [\log n]$, define $f_i : [0, n-1] \rightarrow [0, n-1]$ by



Observation

If (x, y) is an increasing pair in f_i , then b(x, y) = i.

Shoham Letzter



Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.



Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.

Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.

Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

 $\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) =$

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.

Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

 $\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) =$

$$\sum_{A: |A| \le \alpha \log n} \mathbb{P}(\mathbf{A} = A) \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1} \left\{ \begin{array}{l} A \text{ has an increasing} \\ \text{pair w.r.t. } f_i \end{array} \right\} \le$$

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$. Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

 $\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) = \sum_{A: |A| \le \alpha \log n} \mathbb{P}(\mathbf{A} = A) \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1} \left\{ \begin{array}{l} A \text{ has an increasing} \\ pair w.r.t. \ f_{i} \end{array} \right\} \le \sum_{A: |A| \le \alpha \log n} \mathbb{P}[\mathbf{A} = A] \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1}[i \in b(A)]$

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.

Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

$$\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) = \sum_{A: |A| \le \alpha \log n} \mathbb{P}(\mathbf{A} = A) \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1} \left\{ \begin{array}{l} A \text{ has an increasing} \\ \text{pair w.r.t. } f_i \end{array} \right\} \le \sum_{A: |A| \le \alpha \log n} \mathbb{P}[\mathbf{A} = A] \cdot \frac{1}{\log n} \cdot \underbrace{\sum_{i} \mathbb{1}[i \in b(A)]}_{= |b(A)| \le |A| \le \alpha \log n} \end{array} \right\}$$

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$. Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

 $\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) = \sum_{A: |A| \le \alpha \log n} \mathbb{P}(\mathbf{A} = A) \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1} \left\{ \begin{array}{l} A \text{ has an increasing} \\ \text{pair w.r.t. } f_i \end{array} \right\} \le \sum_{A: |A| \le \alpha \log n} \mathbb{P}[\mathbf{A} = A] \cdot \frac{1}{\log n} \cdot \underbrace{\sum_{i} \mathbb{1}[i \in b(A)]}_{=|b(A)| \le |A| \le \alpha \log n} \le \alpha.$

Choose **f** uniformly at random from $\{f_1, \ldots, f_{\log n}\}$.

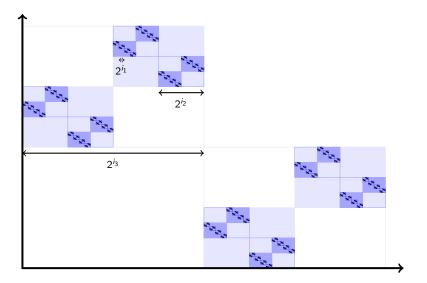
Let **A** be a random subset of [0, n-1] of size at most $\alpha \log n$.

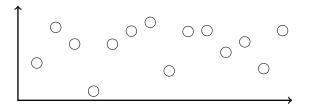
$$\mathbb{P}(\mathbf{A} \text{ has an increasing pair w.r.t. } \mathbf{f}) = \sum_{A: |A| \le \alpha \log n} \mathbb{P}(\mathbf{A} = A) \cdot \frac{1}{\log n} \cdot \sum_{i} \mathbb{1} \left\{ \begin{array}{l} A \text{ has an increasing} \\ pair w.r.t. \ f_{i} \end{array} \right\} \le \sum_{A: |A| \le \alpha \log n} \mathbb{P}[\mathbf{A} = A] \cdot \frac{1}{\log n} \cdot \underbrace{\sum_{i} \mathbb{1}[i \in b(A)]}_{= |b(A)| \le |A| \le \alpha \log n} \le \alpha.$$

Thus, to find an increasing pair with probability at least 0.99 need at least $0.99 \log n$ queries.

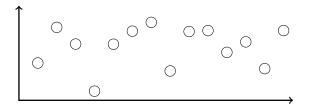
Shoham Letzter

Lower bound for $k = 2^{\kappa}$: iterated construction

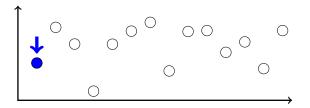




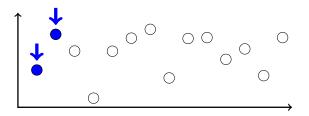




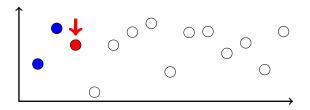
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



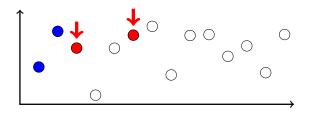
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



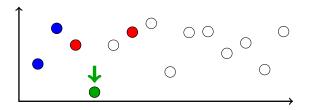
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



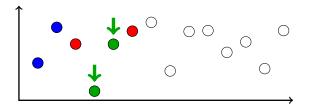
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



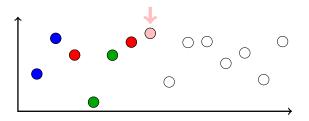
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



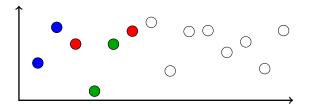
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



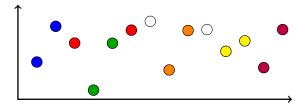
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



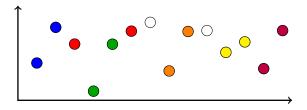
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



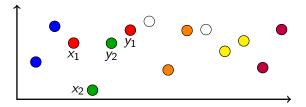
- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.



- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.

Observation

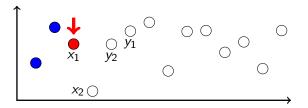
If $(x_1, y_1), (x_2, y_2) \in \mathcal{F}$, and $x_1 < x_2$, $y_1 > y_2$, then $f(y_1) > f(y_2)$.



- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.

Observation

If $(x_1, y_1), (x_2, y_2) \in \mathcal{F}$, and $x_1 < x_2$, $y_1 > y_2$, then $f(y_1) > f(y_2)$.

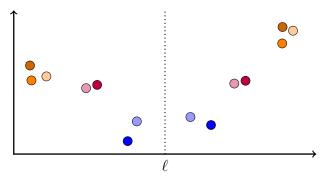


- Start with $\mathcal{F} = \emptyset$.
- For every x, from left to right, if x still unused: Let y > x be leftmost with f(y) > f(x); add (x, y) to F.

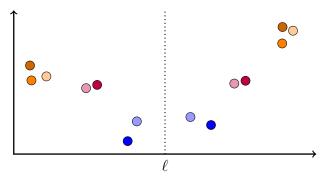
Observation

If $(x_1, y_1), (x_2, y_2) \in \mathcal{F}$, and $x_1 < x_2$, $y_1 > y_2$, then $f(y_1) > f(y_2)$.



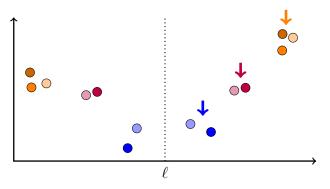


Suppose ℓ lies 'roughly in middle' of many pairs of different widths.



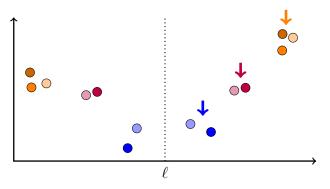
Suppose ℓ lies 'roughly in middle' of many pairs of different widths.

If hit entries of k different-widths copies to the right of ℓ, find (1...k).



Suppose ℓ lies 'roughly in middle' of many pairs of different widths.

If hit entries of k different-widths copies to the right of ℓ, find (1...k).



Suppose ℓ lies 'roughly in middle' of many pairs of different widths.

- If hit entries of k different-widths copies to the right of ℓ , find (1...k).
- Can be done with O(log n) queries: sample Θ(1) elements from [ℓ, ℓ + 2ⁱ] for every i ∈ [log n].



• $\Omega(n)$ many ℓ 's are 'roughly in middle' of many copies.



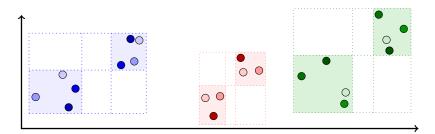
- $\Omega(n)$ many ℓ 's are 'roughly in middle' of many copies.
- If many of them are in middle of many copies of different widths, can find (1...k) with O(log n) queries.



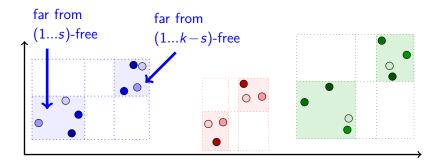
- $\Omega(n)$ many ℓ 's are 'roughly in middle' of many copies.
- If many of them are in middle of many copies of different widths, can find (1...k) with O(log n) queries.
- Otherwise, can cover Ω(n) entries with disjoint 'splittable intervals'.



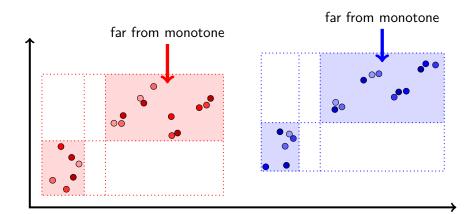
- $\Omega(n)$ many ℓ 's are 'roughly in middle' of many copies.
- If many of them are in middle of many copies of different widths, can find (1...k) with O(log n) queries.
- Otherwise, can cover Ω(n) entries with disjoint 'splittable intervals'.

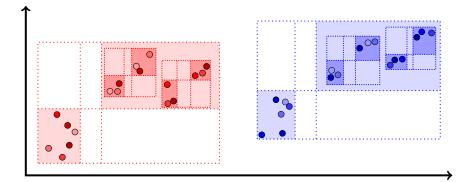


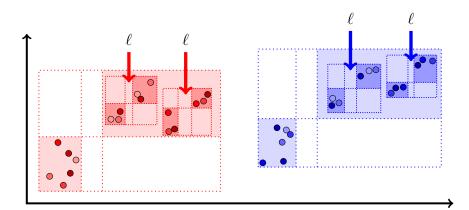
- $\Omega(n)$ many ℓ 's are 'roughly in middle' of many copies.
- If many of them are in middle of many copies of different widths, can find (1...k) with O(log n) queries.
- Otherwise, can cover Ω(n) entries with disjoint 'splittable intervals'.



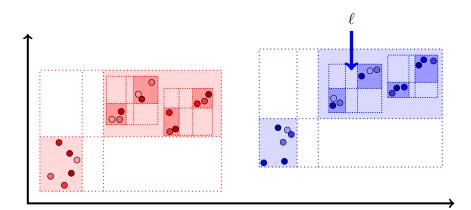




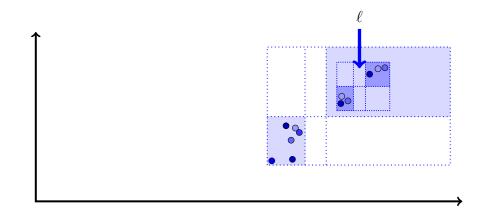




• $\Theta(1)$ queries to find ℓ as in figure.



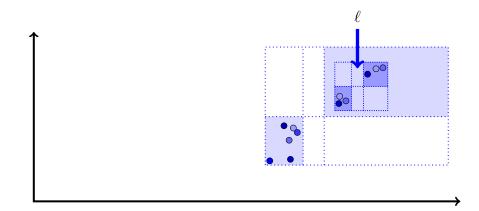
• $\Theta(1)$ queries to find ℓ as in figure.

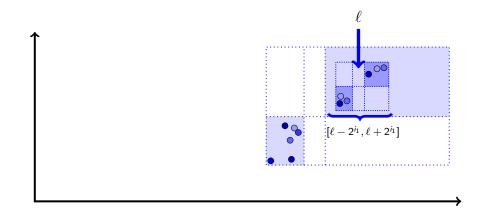


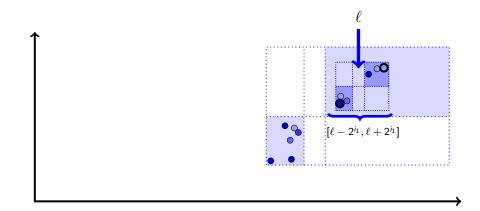
• $\Theta(1)$ queries to find ℓ as in figure.

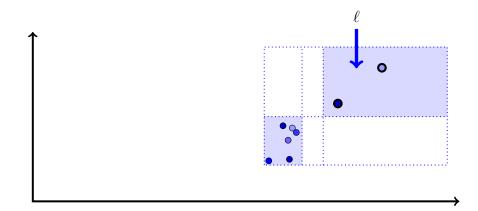
Shoham Letzter

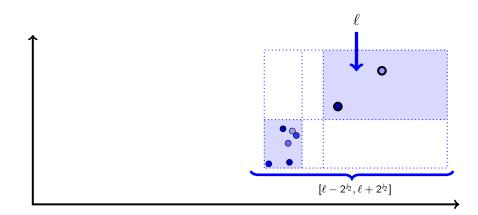
13 / 14



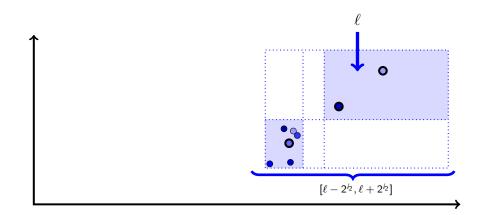




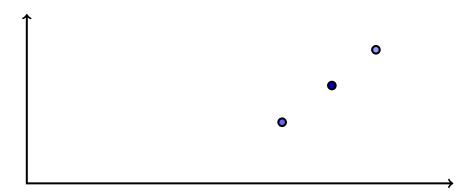




Θ(1) queries to find ℓ as in figure.
For each i ∈ [log n]: make Θ(1) queries in [ℓ − 2ⁱ, ℓ + 2ⁱ].



Θ(1) queries to find ℓ as in figure.
For each i ∈ [log n]: make Θ(1) queries in [ℓ − 2ⁱ, ℓ + 2ⁱ].



Shoham Letzter



k not fixed?



k not fixed?

Testing for other permutations.

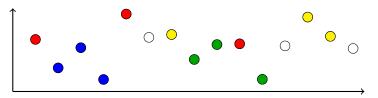


k not fixed?

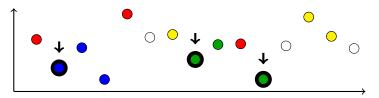
• Testing for other permutations. E.g. $\pi = (231)$.



- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.

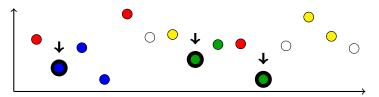


- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.



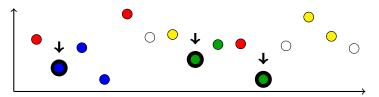


- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.



NRRS '17. If π not monotone, need $\Omega(\sqrt{n})$ non-adaptive queries.

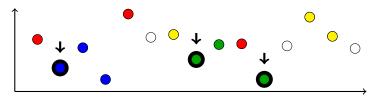
- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.



NRRS '17. If π not monotone, need $\Omega(\sqrt{n})$ non-adaptive queries.

Can π -freeness be tested adaptively in polylog *n* queries?

- *k* not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.

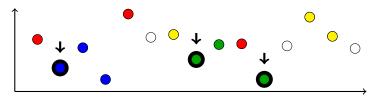


NRRS '17. If π not monotone, need $\Omega(\sqrt{n})$ non-adaptive queries.

Can π -freeness be tested adaptively in polylog *n* queries?

■ Finding a π-copy (length k) in a permutation of length n:
 Fox, '13. 2^{O(k²)} n.

- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.

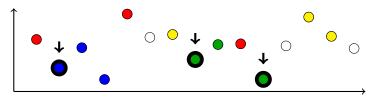


NRRS '17. If π not monotone, need $\Omega(\sqrt{n})$ non-adaptive queries.

Can π -freeness be tested adaptively in polylog *n* queries?

■ Finding a π-copy (length k) in a permutation of length n:
 Fox, '13. 2^{O(k²)} n. Better algorithms?

- k not fixed?
- Testing for other permutations. E.g. $\pi = (231)$.



NRRS '17. If π not monotone, need $\Omega(\sqrt{n})$ non-adaptive queries.

Can π -freeness be tested adaptively in polylog *n* queries?

Finding a π-copy (length k) in a permutation of length n:
 Fox, '13. 2^{O(k²)} n. Better algorithms?

