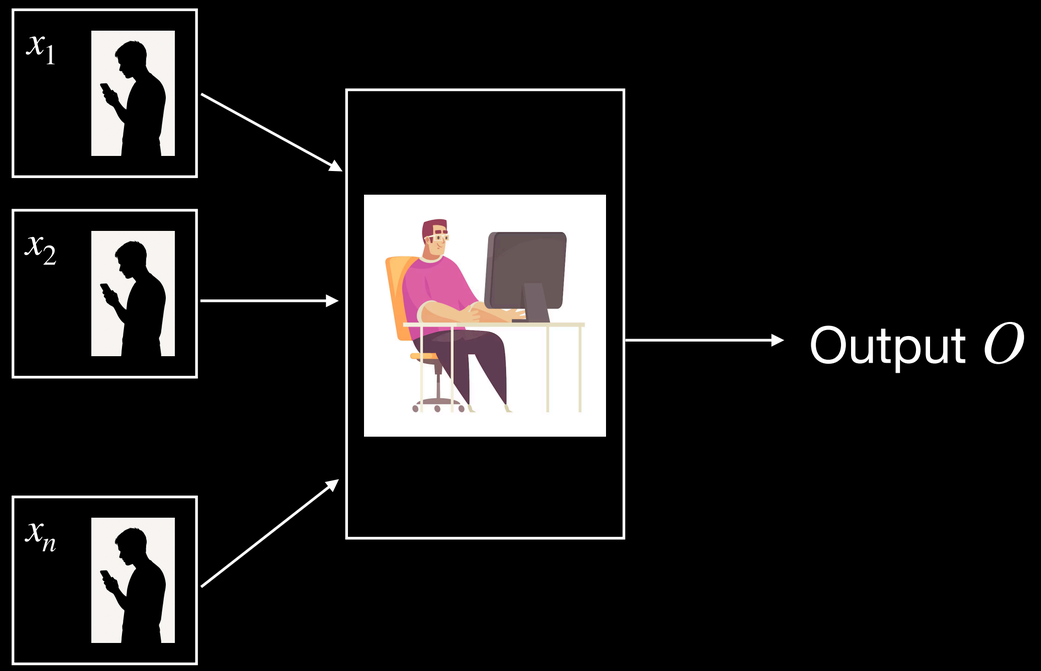
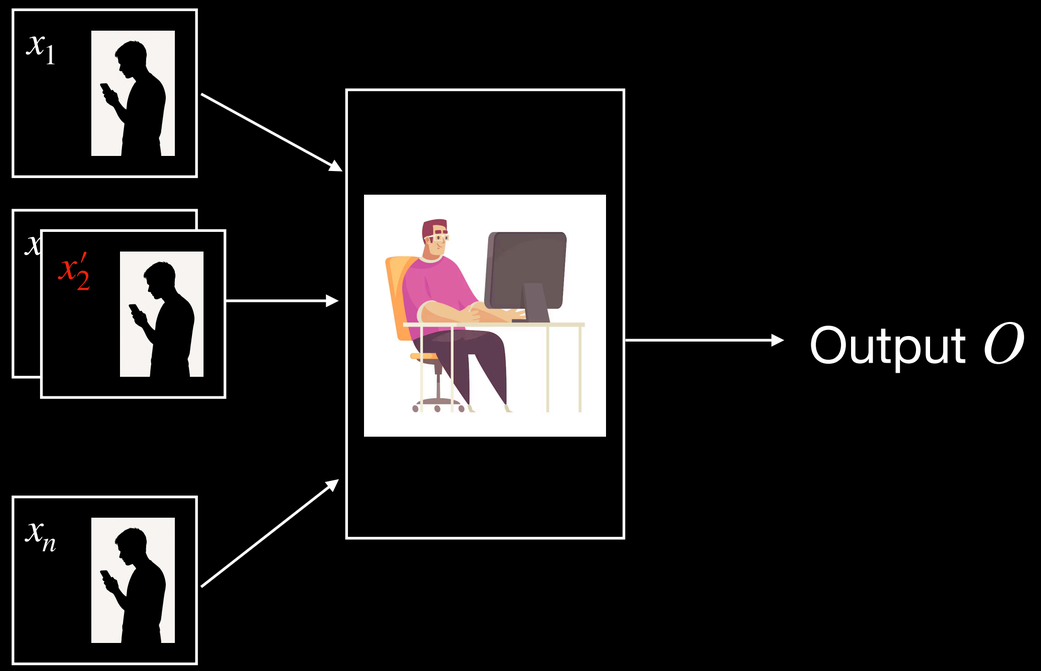
Differentially Private Quantiles with Smaller Error

Jacob Imola, Fabrizio Boninsegna, Hannah Keller, Anders Aamand, Rasmus Pagh, Amrita Roy Chowdhury

Differential Privacy



Differential Privacy



- A change to any x_i produces negligible change to O
- For all $X \sim X'$ and $\mathcal{O}$, $\Pr[A(X) \in \mathcal{O}] \leq e^{\epsilon} \Pr[A(X') \in \mathcal{O}] + \delta$

The Quantiles Problem

Input: $x_1, \ldots, x_n \in [0,b]$, $q_1, q_2, \ldots, q_m \in (0,1)$

Quantile estimates with low error: $\text{err}_q(X, z) = \left| qn - \text{rank}_X(z) \right|$, where $\text{rank}_X(z) = \sum_{i=1}^{n} \mathbf{1}[x_i \leq z]$

The Quantiles Problem

Input: $x_1, \ldots, x_n \in [0,b]$, $q_1, q_2, \ldots, q_m \in (0,1)$

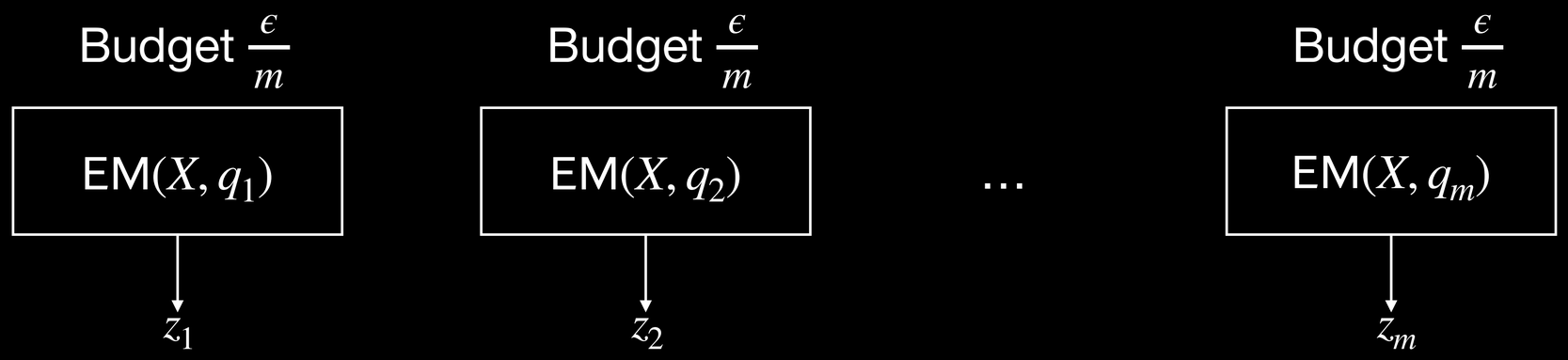
Quantile estimates with low error: $\text{err}_q(X, z) = \left| qn - \text{rank}_X(z) \right|$, where $\text{rank}_X(z) = \sum_{i=1}^{n} \mathbf{1}[x_i \leq z]$

Algorithm	Privacy	Error
AQ [KSS22]	Pure DP	$O(\frac{\log(b)\log(m)^2}{\epsilon})$
RSC [CLNSS19]	Approx DP	$O(\frac{\log^*(b)\log(1/\delta)^2}{\epsilon})$

Naive Approach

Input: $x_1, \ldots, x_n \in [0,b]$, $q_1, q_2, \ldots, q_m$

1. Sort data

2. 

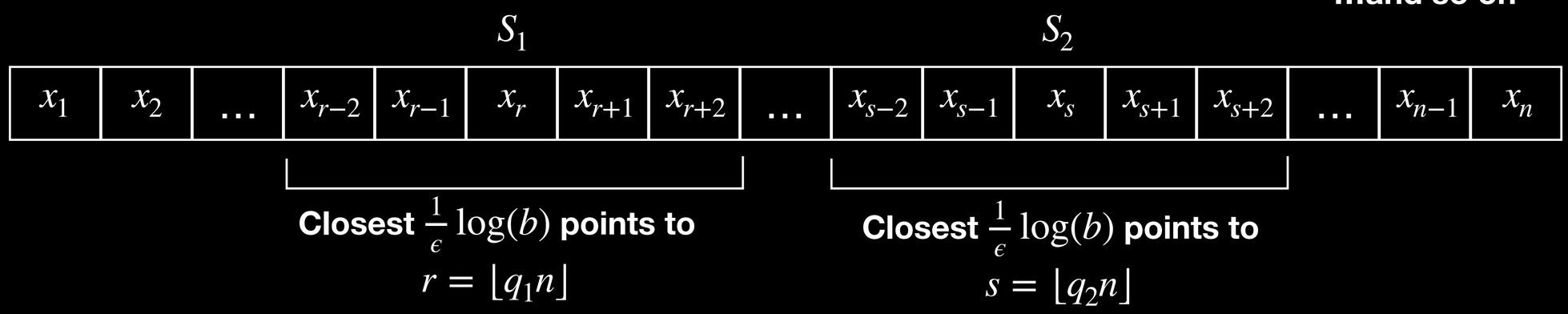
3. Output $z_1, \ldots, z_m$

SliceQuantiles: Improved DP Quantiles

Input: $x_1, \ldots, x_n \in [0,b]$, $q_1, q_2, \ldots, q_m$

1. Sort data

3. Define S_i like



4. Generate estimates $z_1 \sim \text{EM}(S_1, 0.5), \ldots, z_m \sim \text{EM}(S_m, 0.5)$

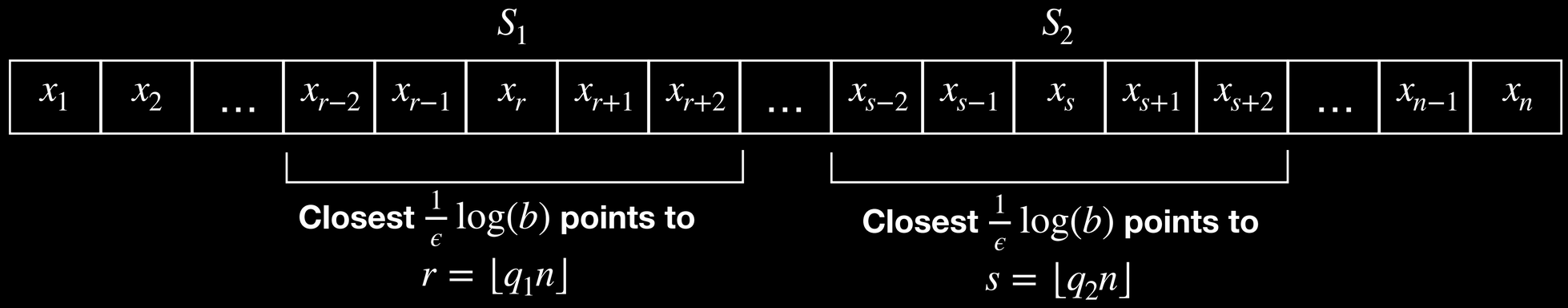
SliceQuantiles: Improved DP Quantiles

Input: $x_1, \ldots, x_n \in [0,b]$, $q_1, q_2, \ldots, q_m$

1. Sort data
2. Add correlated noise η_i to each q_i
3. Define S_i like

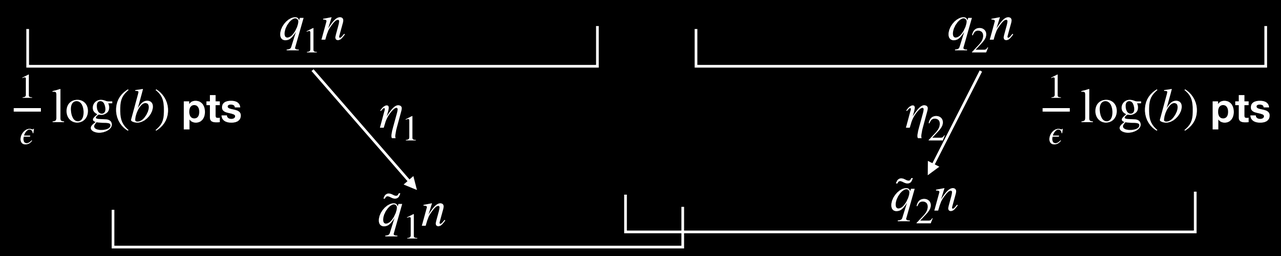
Generated to hide a shift in each $S_j, S_{j+1}, \ldots, S_m$ (Continual Counting)

…and so on

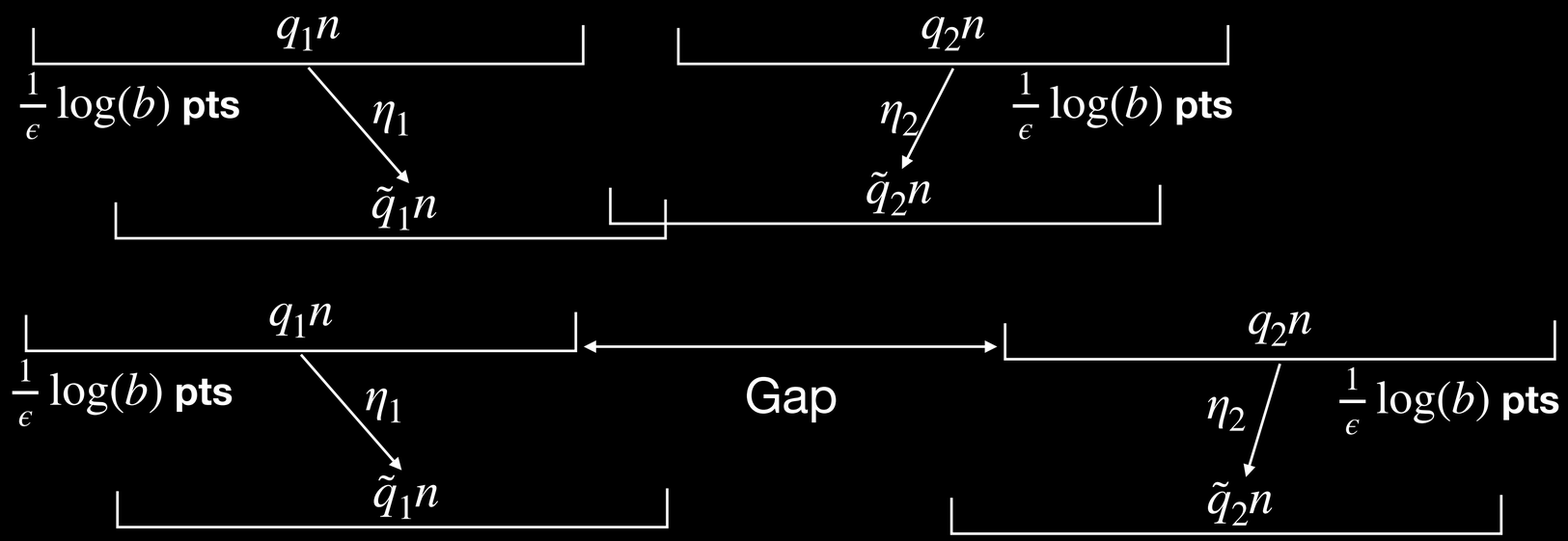


4. Generate estimates $z_1 \sim \text{EM}(S_1, 0.5), \ldots, z_m \sim \text{EM}(S_m, 0.5)$

Quantile Gaps Needed to Avoid Collisions



Quantile Gaps Needed to Avoid Collisions



- Gap $|nq_i - nq_j| \geq \frac{1}{\epsilon}\log(b) + \log(m)\log(\frac{m}{\delta})$
- Conjecture it can be removed with better analysis

The Quantiles Problem

Algorithm	Privacy	Error	Notes
AQ [KSS22]	Pure DP	$O(\frac{\log(b)\log(m)^2}{\epsilon})$	
RSC [CLNSS19]	Approx DP	$O(\frac{\log^*(b)\log(1/\delta)^2}{\epsilon})$	
SliceQuantiles	Pure DP	$O(\frac{\log(b)}{\epsilon})$	$\lvert q_i - q_j \rvert \geq \frac{m\log(m)\log(b)}{n\epsilon}$
SliceQuantiles	Approx DP	$O(\frac{\log(b)}{\epsilon})$	$\lvert q_i - q_j \rvert \geq \frac{\log(1/\delta)\log(m)+\log(b)}{n\epsilon}$