On the Power of Multiple Anonymous Messages

Badih Ghazi Noah Golowich Ravi Kumar
Rasmus Pagh Ameya Velingker

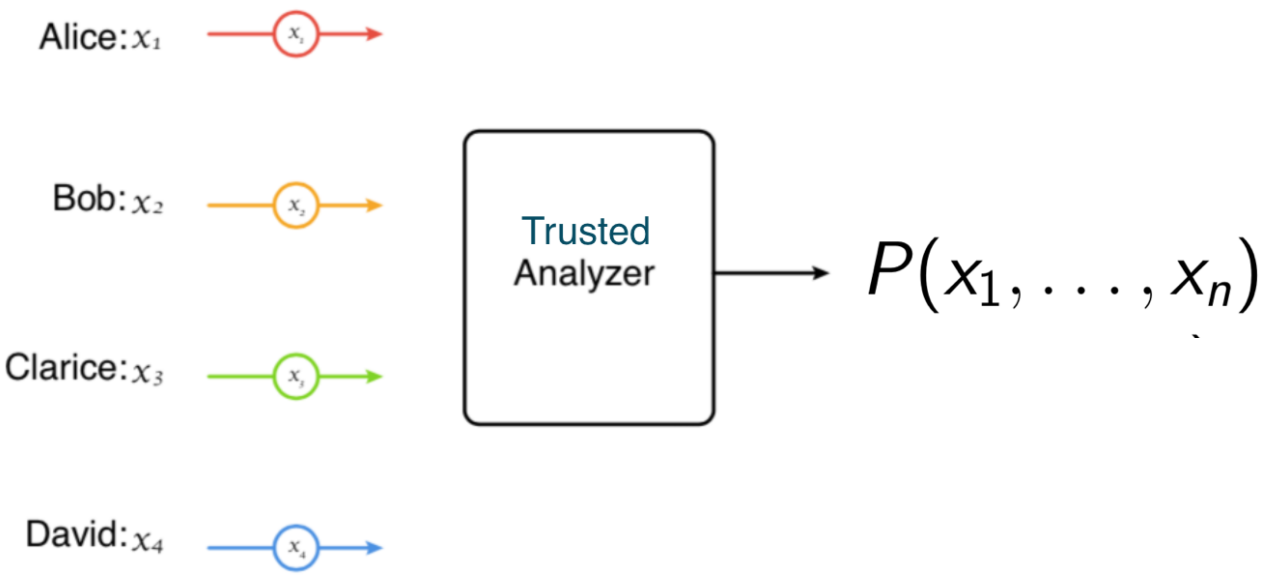
December 18, 2019

Outline

1. Review of central + local models of DP.
2. Shuffled model of DP [Bittau et al., '16].
3. Lower bounds for histogram computation & selection for *single*-message shuffled model [our paper].
   - Nearly tight, improve upon [Cheu et al., '18].
4. Upper bounds for histogram computation for *multi*-message shuffled model [our paper].
   - Applications to range queries, quantile estimation, etc.

Differential privacy – central model

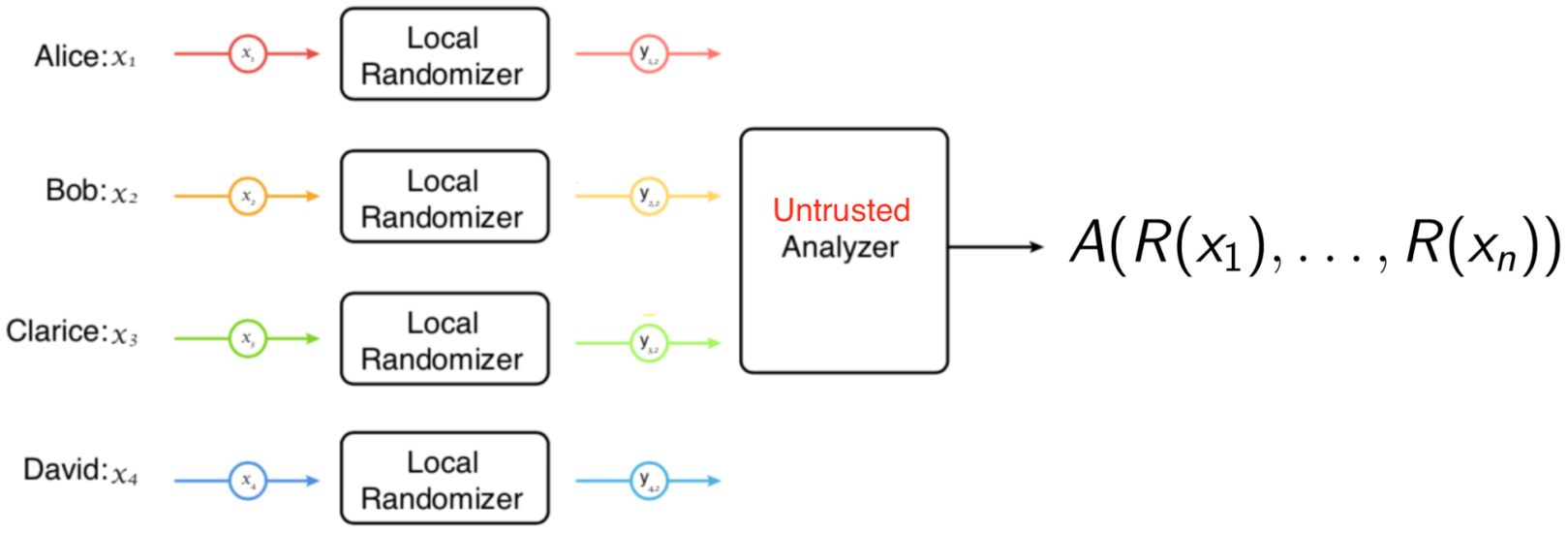
- **Universe** $\mathcal{X}$, users with data points $x_1, \ldots, x_n \in \mathcal{X}$.
- Give data to analyzer; adds noise to preserve priv.



Algorithm $P : \mathcal{X}^n \to \mathcal{Y}$ is (ϵ, δ)-**DP** if: $\forall\, \mathcal{S} \subset \mathcal{Y}$,
$\forall$ *neighboring datasets* $X = (x_1, \ldots, x_n), X' = (x_1, \ldots, x_{n-1}, x'_n)$,

$$\Pr_P[P(X) \in \mathcal{S}] \leq e^{\epsilon} \cdot \Pr_P[P(X') \in \mathcal{S}] + \delta.$$

Differential privacy – local model



- Users must add privacy-preserving noise themselves.
- Each user i (data x_i) sends output of **local randomizer** R, $R(x_i) \in \mathcal{Z}$, to central **analyzer**.

R is (ϵ, δ)**-local DP**: if for all $x, x' \in \mathcal{X}$, $\mathcal{S} \subset \mathcal{Z}$,

$$\Pr_R[R(x) \in \mathcal{S}] \leq e^{\epsilon} \Pr_R[R(x') \in \mathcal{S}] + \delta.$$

Quick interlude: ϵ, δ

- Assume, e.g., $\epsilon = 0.1$ for purposes of this talk.
  (The errors of our protocols degrade generally as $1/\epsilon$*.)*
- Assume $\delta = 1/\operatorname{poly}(n)$.
  (If $\delta > 1/n$*, then can recover constant fraction of users' data! But also want* $\log(1/\delta) = O(\log n)$*.)*

Central vs. Local models

- **Histograms**: $x_i \in \mathcal{X} = [B] := \{1, 2, \ldots, B\}$, goal is to release # of users holding each $j \in [B]$.
- How to release histograms privately? Add noise!

Central vs. Local models

- **Histograms**: $x_i \in \mathcal{X} = [B] := \{1, 2, \ldots, B\}$, goal is to release # of users holding each $j \in [B]$.
- How to release histograms privately? Add noise!
- *Central model*: Laplace mechanism
  $P(x_1, \ldots, x_n)_j = |\{i : x_i = j\}| + \mathrm{Lap}(1/\epsilon)$.
  - Additive error: $\Theta\left(\frac{\log B}{\epsilon}\right)$ for $(\epsilon, 0)$-DP.

Central vs. Local models

- **Histograms**: $x_i \in \mathcal{X} = [B] := \{1, 2, \ldots, B\}$, goal is to release # of users holding each $j \in [B]$.
- How to release histograms privately? Add noise!
- *Central model*: Laplace mechanism
  $P(x_1, \ldots, x_n)_j = |\{i : x_i = j\}| + \mathrm{Lap}(1/\epsilon)$.
  - Additive error: $\Theta\left(\frac{\log B}{\epsilon}\right)$ for $(\epsilon, 0)$-DP.
- *Local model*: RAPPOR$_\epsilon$ [Erlingsson et al., 2014]
  $R_\epsilon(x_i)$: Interpret $x_i \in \{0,1\}^B$ (unit vector e_{x_i}), flip each bit w.p. $\frac{1}{1+e^{\epsilon/2}}$, send the noisy vector.

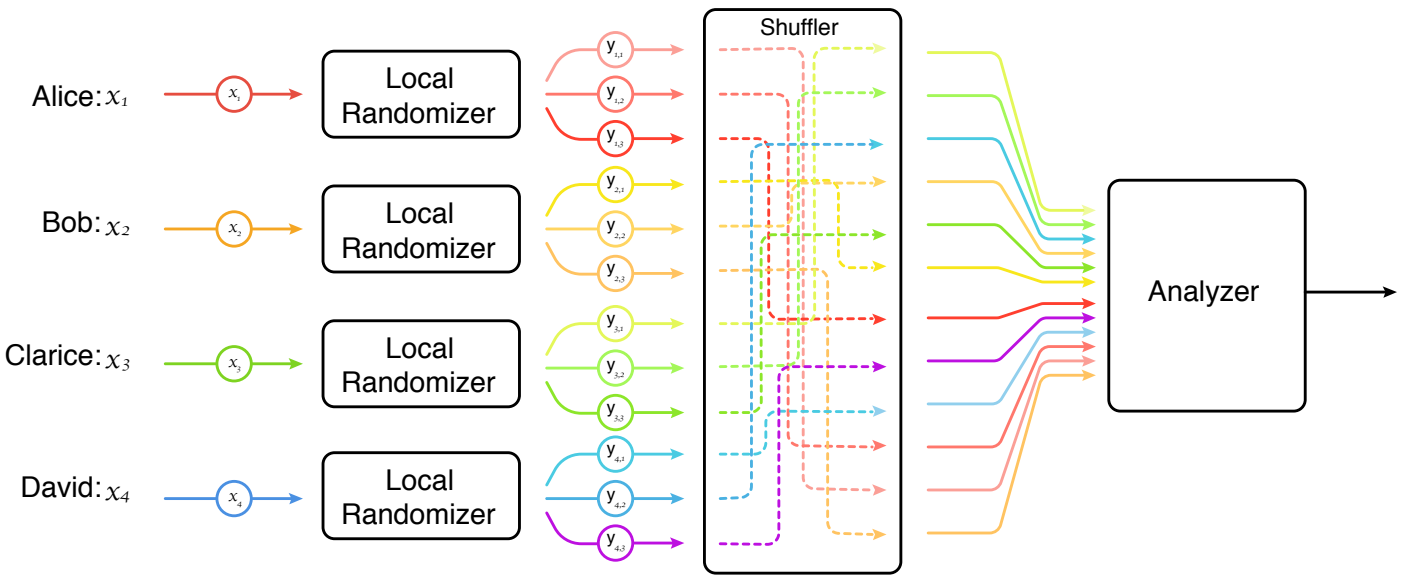
x_i	0	1	0	0	0	0
$R_\epsilon(x_i)$	1	1	0	0	1	0

  - Additive error: $\Theta\left(\frac{\sqrt{n \log B}}{\epsilon}\right)$ for $(\epsilon, 0)$-local DP.

Shuffled model [Bittau et al., 2016]

- Insert *shuffler* $S : \mathcal{Z}^* \to \mathcal{Z}^*$ between R & A:

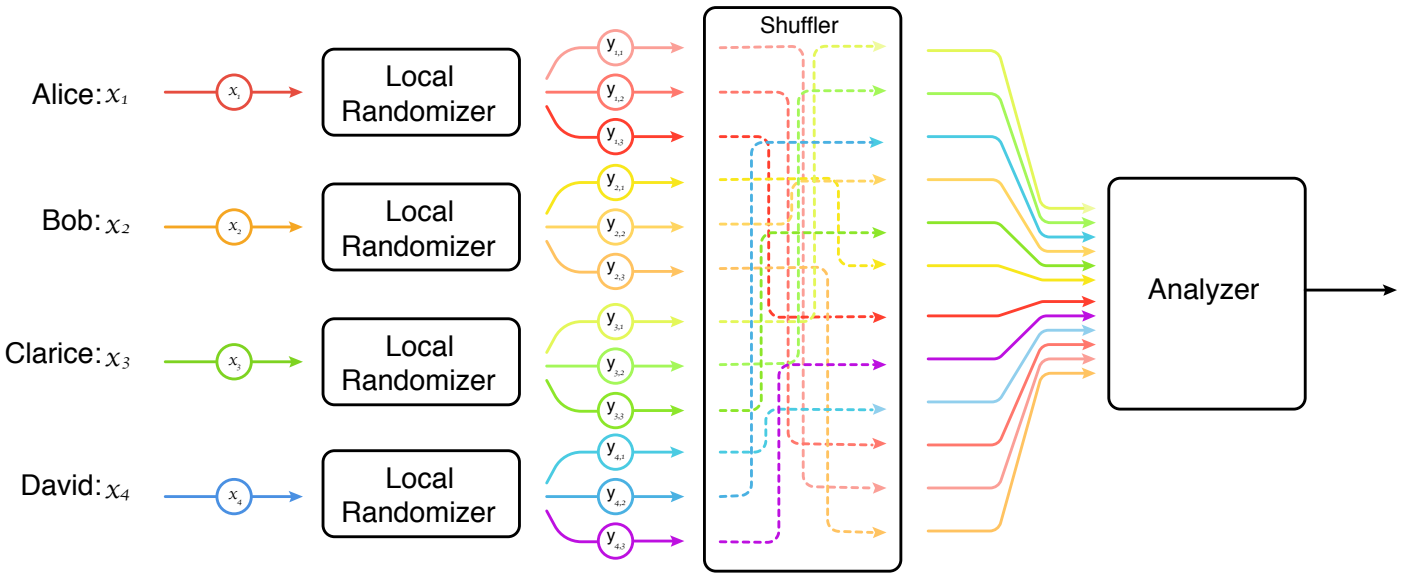
$$S(z_1, \ldots, z_m) = (z_{\pi(1)}, \ldots, z_{\pi(m)}), \quad \pi \sim S_m.$$



Shuffled model [Bittau et al., 2016]

- Insert *shuffler* $S : \mathcal{Z}^* \to \mathcal{Z}^*$ between R & A:

$$S(z_1, \ldots, z_m) = (z_{\pi(1)}, \ldots, z_{\pi(m)}), \quad \pi \sim S_m.$$



R is (ϵ, δ)-**DP in shuffled model** if:

$$(x_1, \ldots, x_n) \mapsto S(R(x_1), \ldots, R(x_n))$$

is (ϵ, δ)-DP. *(Trust shuffler, not analyzer)*

Single- vs. multi-message shuffled model

Multi-message shuffled model:

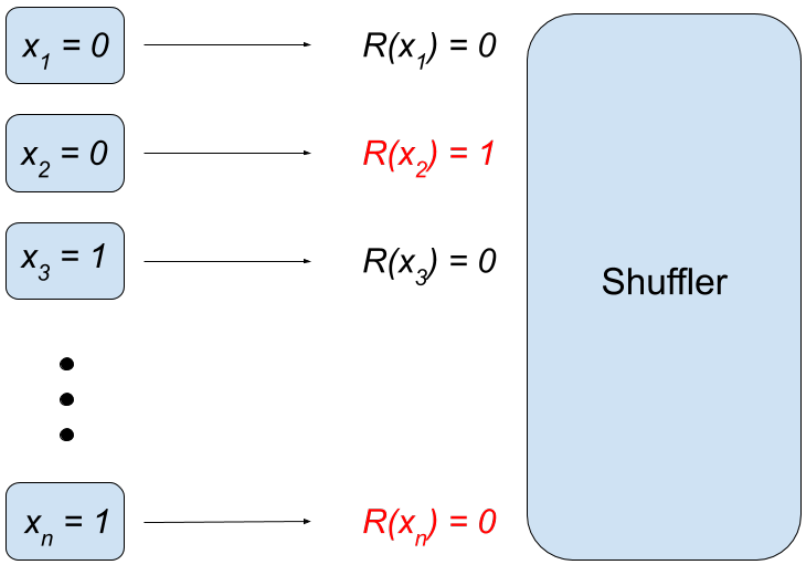
- $R(x) \in \mathcal{Z}^*$; assume for some $m_0 \in \mathbb{N}$, $R(x) \in \mathcal{Z}^{m_0}$.
- Shuffler permutes $m_0 n$ *messages* $\{R(x_i)_j\}_{j \in [m_0], i \in [n]}$.

Single-message shuffled model:

- $R(x) \in \mathcal{Z}$ for all $x \in \mathcal{X}$ (i.e., $m_0 = 1$).

The power of a *single* anonymous message

- Histograms for $B = 2$ equivalent to bit-summation (we know n).
- **Bit-summation**: $x_i \in \{0, 1\}$; compute $\sum_{i=1}^{n} x_i$.
- [Cheu et al., 18]: bit-summation in shuffled model:
- Local randomizer: flip $x_i \in \{0, 1\}$ w.p. $p \approx \frac{\log 1/\delta}{n\epsilon^2}$.



- Error $O\left(\frac{\sqrt{\log 1/\delta}}{\epsilon}\right)$, is (ϵ, δ)-DP. (Local requires $\Omega(\sqrt{n}/\epsilon)$.)

Amplification by shuffling

More on single-message protocols:

Amplification by shuffling

More on single-message protocols:

- "Amplification by shuffling" ([Erlingsson et al., 18]; [Balle et al., 19]): R is $(\epsilon_L, 0)$-locally DP implies

$$(x_1, \ldots, x_n) \mapsto S(R(x_1), \ldots, R(x_n))$$

is (ϵ_S, δ)-DP for $\epsilon_S \ll \epsilon_L$.

Amplification by shuffling

More on single-message protocols:

- "Amplification by shuffling" ([Erlingsson et al., 18]; [Balle et al., 19]): R is $(\epsilon_L, 0)$-locally DP implies

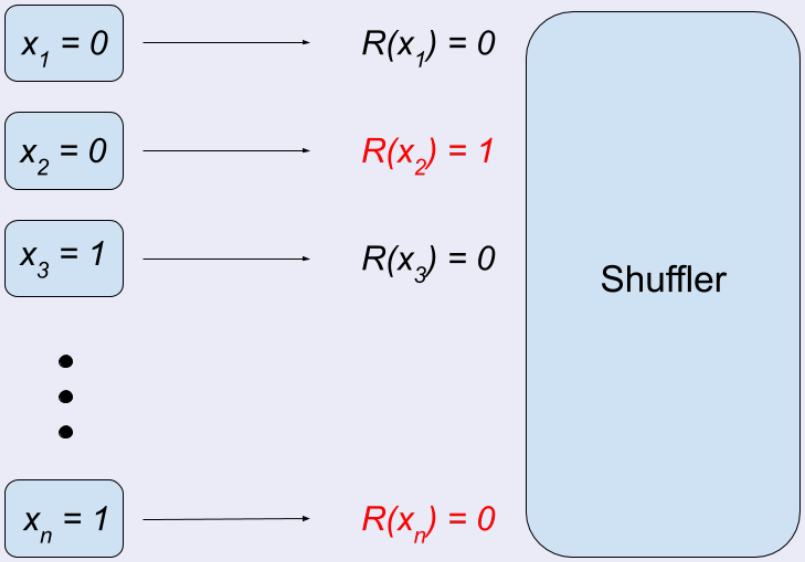
$$(x_1, \ldots, x_n) \mapsto S(R(x_1), \ldots, R(x_n))$$

is (ϵ_S, δ)-DP for $\epsilon_S \ll \epsilon_L$.

Example

Bit summation [Cheu et al]:

- $\epsilon_S \approx 0.1$,
- $\epsilon_L \approx \ln(1/p) \approx \ln(n)$,
- error $\approx n/e^{\epsilon_L} = \tilde{O}(1)$.



Amplification by shuffling is optimal

Theorem (Our paper, informal)

"Amplification by shuffling" gives optimal error for histograms in single-message shuffled model.

Amplification by shuffling is optimal

Theorem (Our paper, informal)

"Amplification by shuffling" gives optimal error for histograms in single-message shuffled model.

Amplification by shuffling [Balle et al., 19] for $(\epsilon_S = 0.1, \delta)$-DP in shuffled model:

Protocol	n size	error	regime
$\text{RAPPOR}_{\epsilon_L \approx 0.1}$	$n \gg_{\epsilon_S} \log B$	$n/10$	"small sample"

Amplification by shuffling is optimal

Theorem (Our paper, informal)

"Amplification by shuffling" gives optimal error for histograms in single-message shuffled model.

Amplification by shuffling [Balle et al., 19] for $(\epsilon_S = 0.1, \delta)$-DP in shuffled model:

Protocol	n size	error	regime
$\text{RAPPOR}_{\epsilon_L \approx 0.1}$	$n \gg_{\epsilon_S} \log B$	$n/10$	"small sample"
$\text{RAPPOR}_{\epsilon_L \approx \ln(n)}$	$n \gg \log^2 B$	$\tilde{O}_{\epsilon_S}(n^{1/4})$	"intermed. sample"

Amplification by shuffling is optimal

Theorem (Our paper, informal)

"Amplification by shuffling" gives optimal error for histograms in single-message shuffled model.

Amplification by shuffling [Balle et al., 19] for $(\epsilon_S = 0.1, \delta)$-DP in shuffled model:

Protocol	n size	error	regime
$\text{RAPPOR}_{\epsilon_L \approx 0.1}$	$n \gg_{\epsilon_S} \log B$	$n/10$	"small sample"
$\text{RAPPOR}_{\epsilon_L \approx \ln(n)}$	$n \gg \log^2 B$	$\tilde{O}_{\epsilon_S}(n^{1/4})$	"intermed. sample"
$B\text{-RR}_{\epsilon_L \approx \ln(n)}$ [Warner, 65]	$n > B^2$	$\tilde{O}_{\epsilon_S}(\sqrt{B})$	"large sample"

Amplification by shuffling is optimal

Protocol	n size	error	regime
$\text{RAPPOR}_{\epsilon_L \approx 0.1}$	$n \gg_{\epsilon_S} \log B$	$n/10$	"small sample"
$\text{RAPPOR}_{\epsilon_L \approx \ln(n)}$	$n \gg \log^2 B$	$\tilde{O}_{\epsilon_S}(n^{1/4})$	"intermed. sample"
B-$\text{RR}_{\epsilon_L \approx \ln(n)}$ [Warner, 65]	$n > B^2$	$\tilde{O}_{\epsilon_S}(\sqrt{B})$	"large sample"

Theorem (Our paper, formal)

Any $(1, o(1/n))$*-DP single-message shuffled protocol for histograms has additive error* $\tilde{\Omega}\left(\min\{n^{1/4}, \sqrt{B}\}\right)$.

If error is $< n/10$*, then* $n \geq \Omega\left(\frac{\log B}{\log \log B}\right)$.

[Cheu et al., 18]: $n \geq \Omega(\log^{1/17} B)$ for error $< n/10$.

Related lower bound: selection

- **Selection:** generalization of histograms, each user holds any number of items from $[B]$.
- User i holds $x_i \in \{0,1\}^B$, goal is to find (approximately) most common item $j^* \in [B]$, i.e.:

$$\sum_{i=1}^{n} x_{i,j^*} > \max_{j \in [B]} x_{i,j} - \frac{n}{10}.$$

	$j=1$	$j^*=2$	$j=3$	$j=4$
x_1	0	1	0	1
x_2	1	1	0	1
x_3	0	1	1	0
x_4	1	0	0	1
x_5	0	1	0	0

Bounds for selection

Best possible sample complexity:

- *Central model*: $n = \Theta\left(\frac{\log B}{\epsilon}\right)$
  - Exponential mechanism achieves this.
- *Local model*: $n = \Theta\left(\frac{B \log B}{\epsilon^2}\right)$ [Ullman, 18].
  - Randomized Response [Warner, 65] to estimate each coordinate achieves this.
- *Single-message shuffled model*:

Theorem (Our paper)

Selection in the single-message shuffled model with $(\epsilon = 1, \delta = o\left(\frac{1}{Bn}\right))$*-DP requires* $n \geq \Omega(B)$.

Compare with [Cheu et al., 18], who showed $n \geq \Omega(B^{1/17})$.

Proof idea of lower bounds

Lemma (Cheu et al., 18)

(R, S) is (ϵ, δ)-DP in 1-msg shuffled model $\Rightarrow$ R is $(\epsilon + \ln(n), \delta)$-DP in local model.

Proof idea of lower bounds

Lemma (Cheu et al., 18)

(R, S) is (ϵ, δ)-DP in 1-msg shuffled model $\Rightarrow$ R is $(\epsilon + \ln(n), \delta)$-DP in local model.

- So, suffices to lower bound error in low-privacy *local model* (for (ϵ, δ)-DP).
- By **Fano's inequality**, suffices to upper bound $I(V; R(X))$, where $(V, X) \in [B] \times [B]$:

$$V \sim \mathrm{Unif}([B]), \quad X = \begin{cases} V & : \text{w.p. } \gamma \\ \mathrm{Unif}([B]) & : \text{w.p. } 1 - \gamma \end{cases}.$$

($n\gamma$ is target error lower bound; e.g., $\gamma = \tilde{\Theta}(n^{-3/4})$ in intermed.-sample regime).

Proof idea of lower bounds

- By **Fano's inequality**, suffices to upper bound $I(V; R(X))$, where $(V, X) \in [B] \times [B]$:

$$V \sim \mathrm{Unif}([B]), \quad X = \begin{cases} V & : \text{w.p. } \gamma \\ \mathrm{Unif}([B]) & : \text{w.p. } 1 - \gamma \end{cases}.$$

- For histograms: need $I(V; R(X)) \leq \tilde{O}(1/n)$:
  - Small sample: straightforward.
  - Intermediate + large-sample: harder, have to use accuracy of (R, A).

Proof idea of lower bounds

- By **Fano's inequality**, suffices to upper bound $I(V; R(X))$, where $(V, X) \in [B] \times [B]$:

$$V \sim \mathrm{Unif}([B]), \quad X = \begin{cases} V & : \text{w.p. } \gamma \\ \mathrm{Unif}([B]) & : \text{w.p. } 1 - \gamma \end{cases}.$$

- For histograms: need $I(V; R(X)) \leq \tilde{O}(1/n)$:
  - Small sample: straightforward.
  - Intermediate + large-sample: harder, have to use accuracy of (R, A).
- For selection: distr. of V, X a bit different, but need to show $I(V; R(X)) \leq O(1/B)$.
  - Main tool: *Level-1 inequality* from analysis of boolean functions.

Outline

1. Review of central + local models of DP.
2. Shuffled model of DP [Bittau et al., '16].
3. Lower bounds for histogram computation & selection for *single*-message shuffled model [our paper].
   - Nearly tight, improve upon [Cheu et al., '18].
4. Upper bounds for histogram computation for *multi*-message shuffled model [our paper].
   - Applications to range queries, quantile estimation, etc.

Multi-message upper bounds

Recall: error $\tilde{\Theta}(\min\{n^{1/4}, \sqrt{B}\})$ w/ 1-msg; we show: can get error $\tilde{O}(1)$ with multiple messages!
(Think of $B = \text{poly}(n)$.)

Multi-message upper bounds

Recall: error $\tilde{\Theta}(\min\{n^{1/4}, \sqrt{B}\})$ w/ 1-msg; we show: can get error $\tilde{O}(1)$ with multiple messages!
(Think of $B = \text{poly}(n)$.)

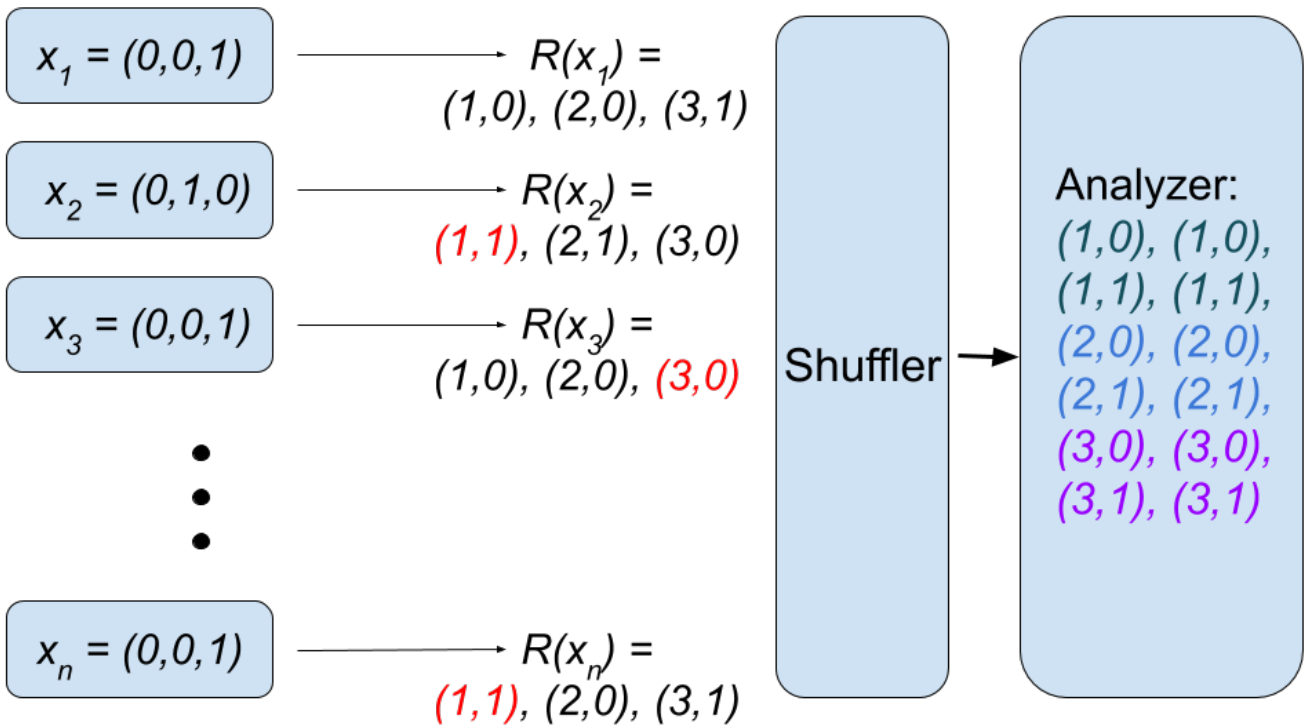
Step 1: view histogram computation as bit-summation in each coordinate.

	$j=1$	$j=2$	$j=3$	$j=4$
$x_1 = 2$	0	1	0	0
$x_2 = 4$	0	0	0	1
$x_3 = 2$	0	1	0	0
$x_4 = 1$	1	0	0	0
$x_5 = 2$	0	1	0	0
Histogram	1	3	0	1

Bit-summation: $x_i \in \mathcal{X} = \{0, 1\}$; compute $\sum_{i=1}^{n} x_i$.

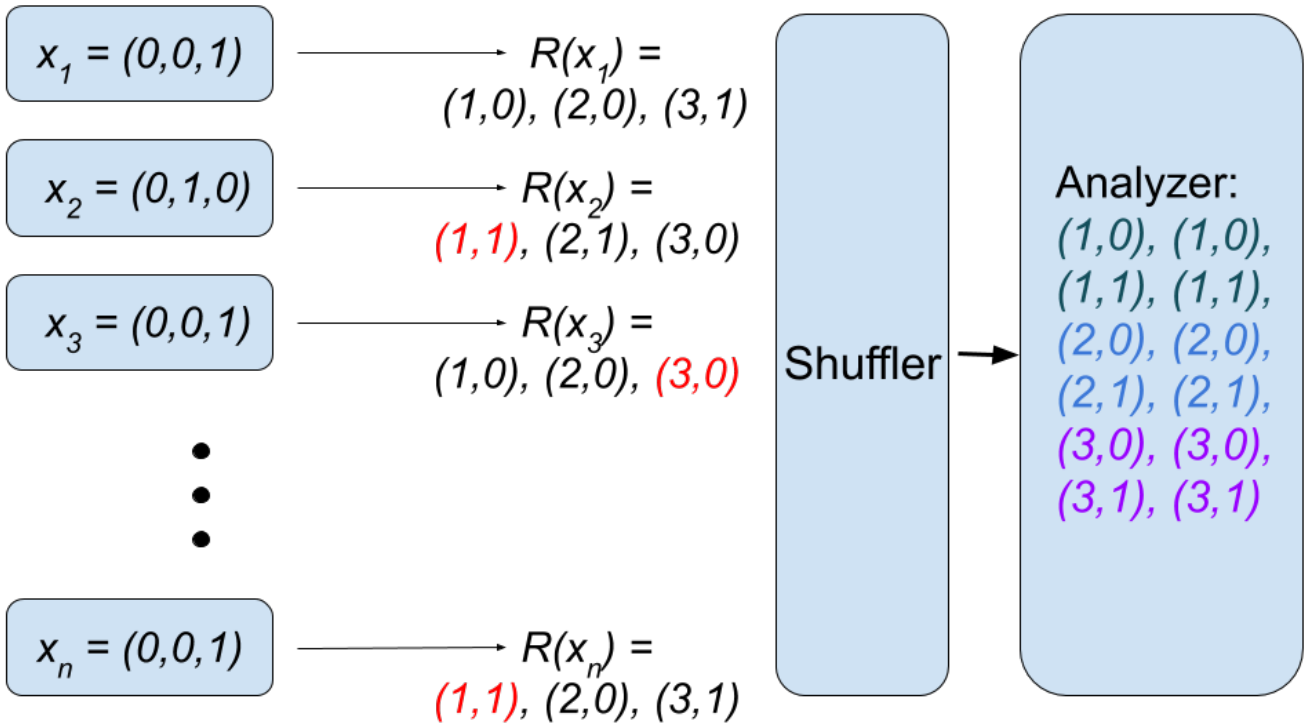
Multi-message naive protocol

- Use B copies of [Cheu et al] protocol, 1 for each $j \in [B]$:



Multi-message naive protocol

- Use B copies of [Cheu et al] protocol, 1 for each $j \in [B]$:



- Expected max error: $O\left(\frac{\sqrt{\log \frac{1}{\delta} \log B}}{\epsilon}\right)$ (noise is sub-Gaussian).
- Privacy: Follows from *composition property* of DP.
- **Communication: B messages per user.**

Multi-message upper bounds

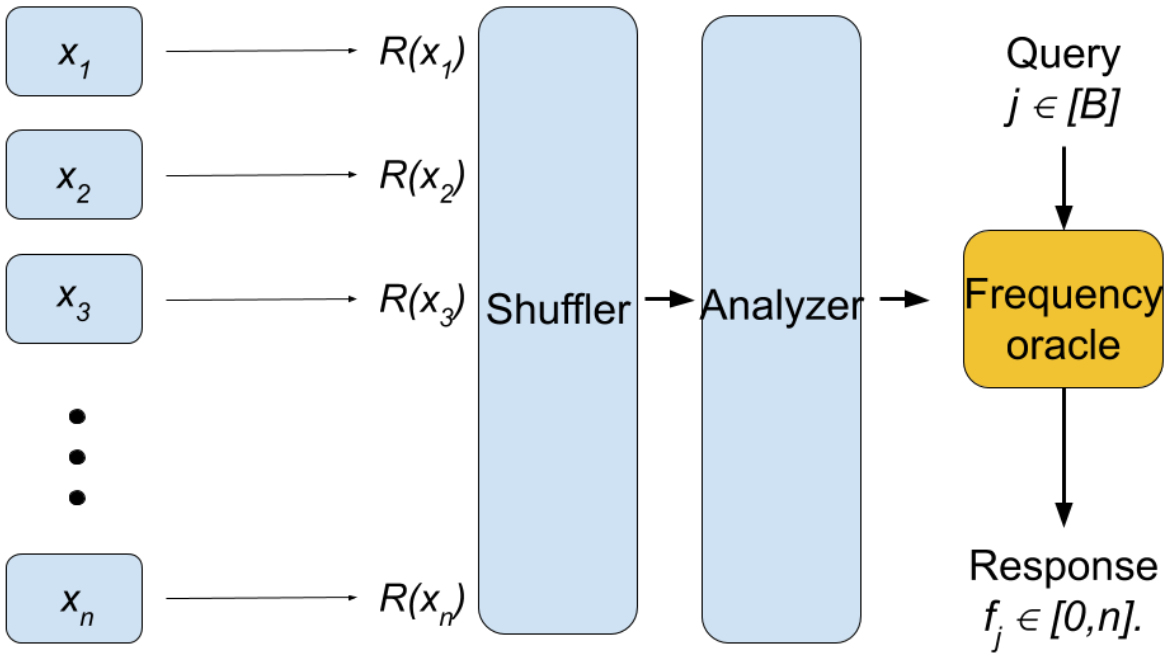
Step 2: Use either **count-min sketch** *or* **Hadamard response** *to reduce per-user communication from* $\tilde{\Theta}(B)$ *to* $\tilde{O}(1)$.

High-level tradeoffs:

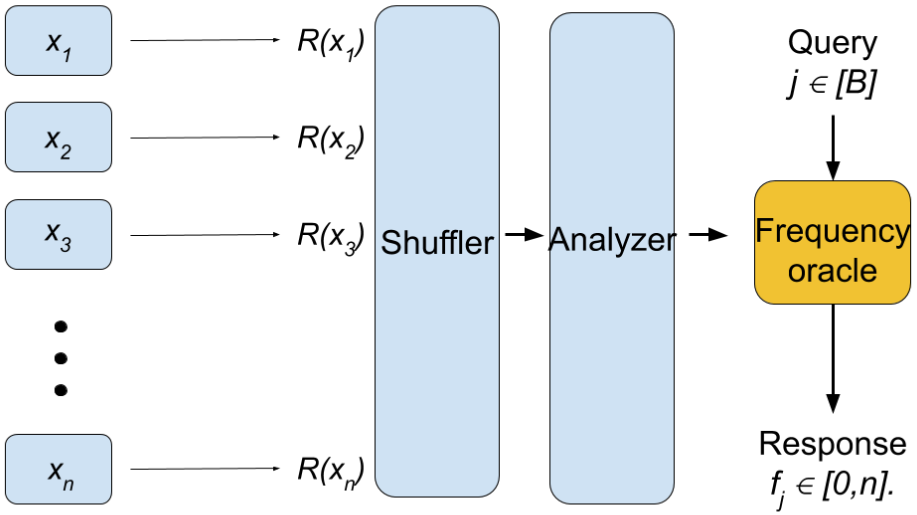
- **Count-min sketch**: requires public randomness (but adversary can see it).
- **Hadamard response**: no public randomness needed, but greater computation.

Results for multi-message protocols

Issue: B is large, so infeasible for A to compute frequencies of all $j \in [B]$.
Solution: Have A output **frequency oracle** FO.
Query time: Time taken by a single query $\mathrm{FO}(j) =?$.



Results for multi-message protocols



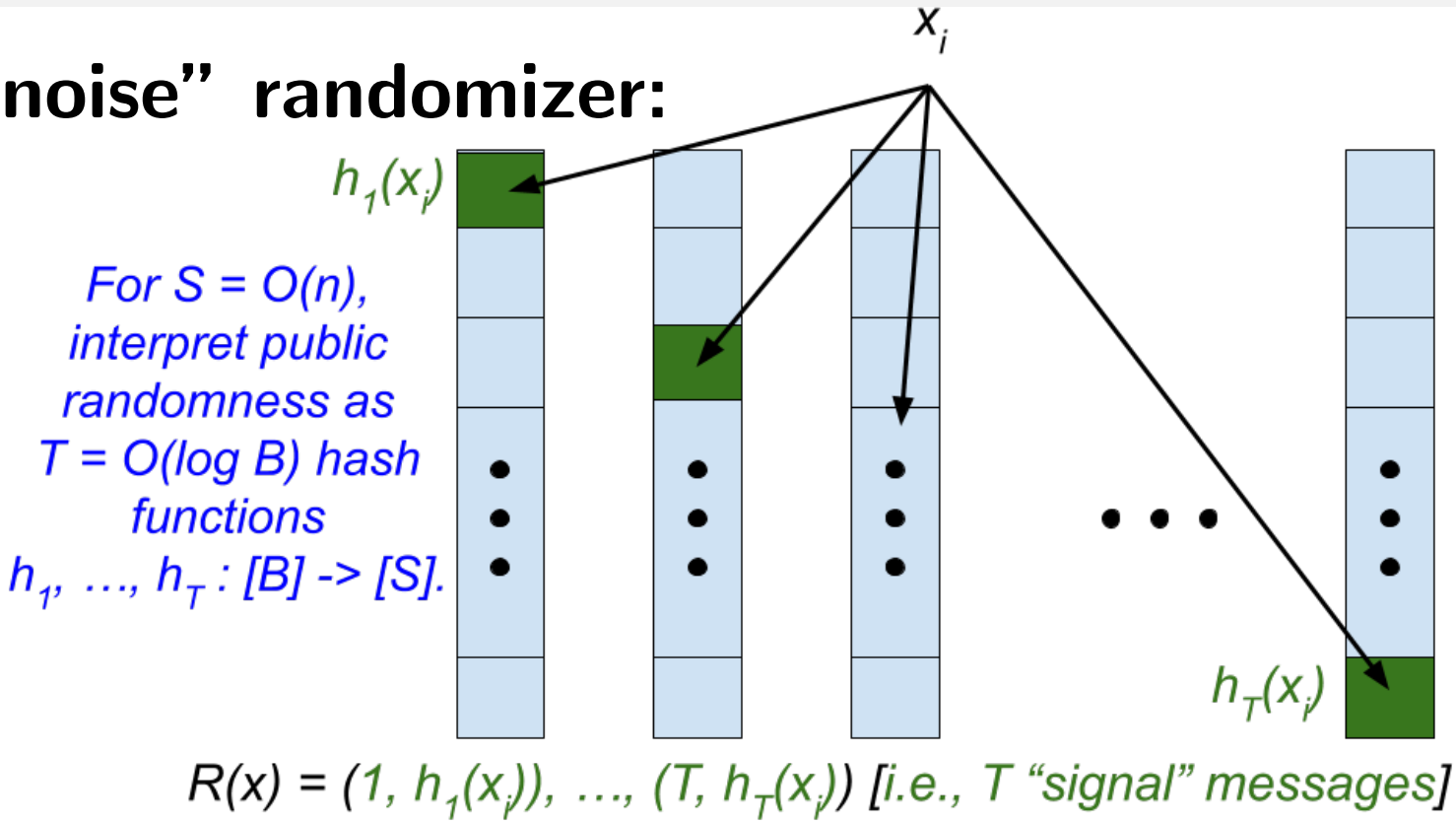
Theorem (Our paper)

There exist multi-message shuffled model (ϵ, δ)-DP protocols with the below properties:

Technique	*error*	*comm.*	*query time*	*public coins*
Count-min	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	*Yes*
Hadamard Resp.	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(n)$	*No*

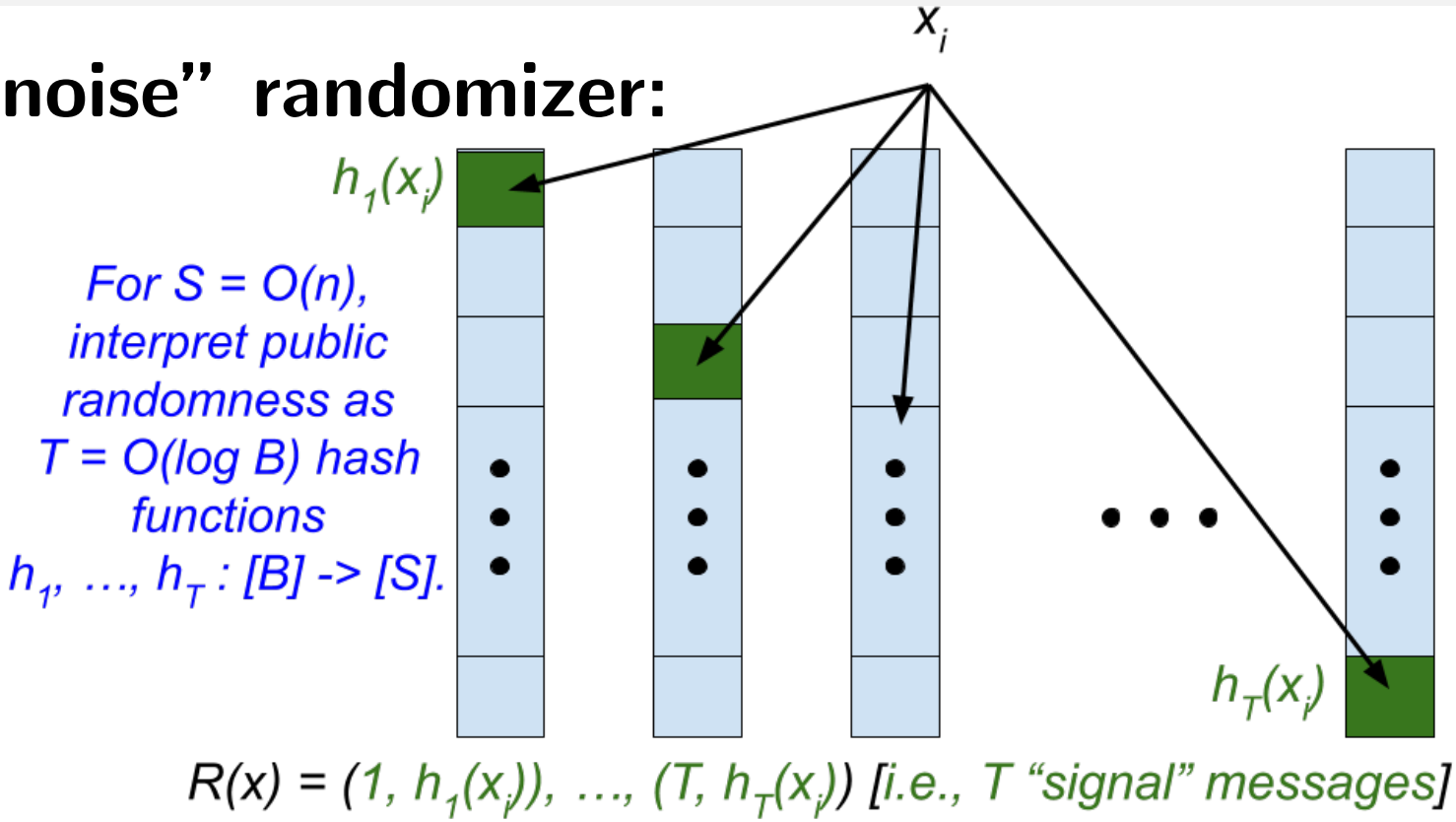
Non-DP count-min-based protocol

"No-noise" randomizer:



Non-DP count-min-based protocol

"No-noise" randomizer:

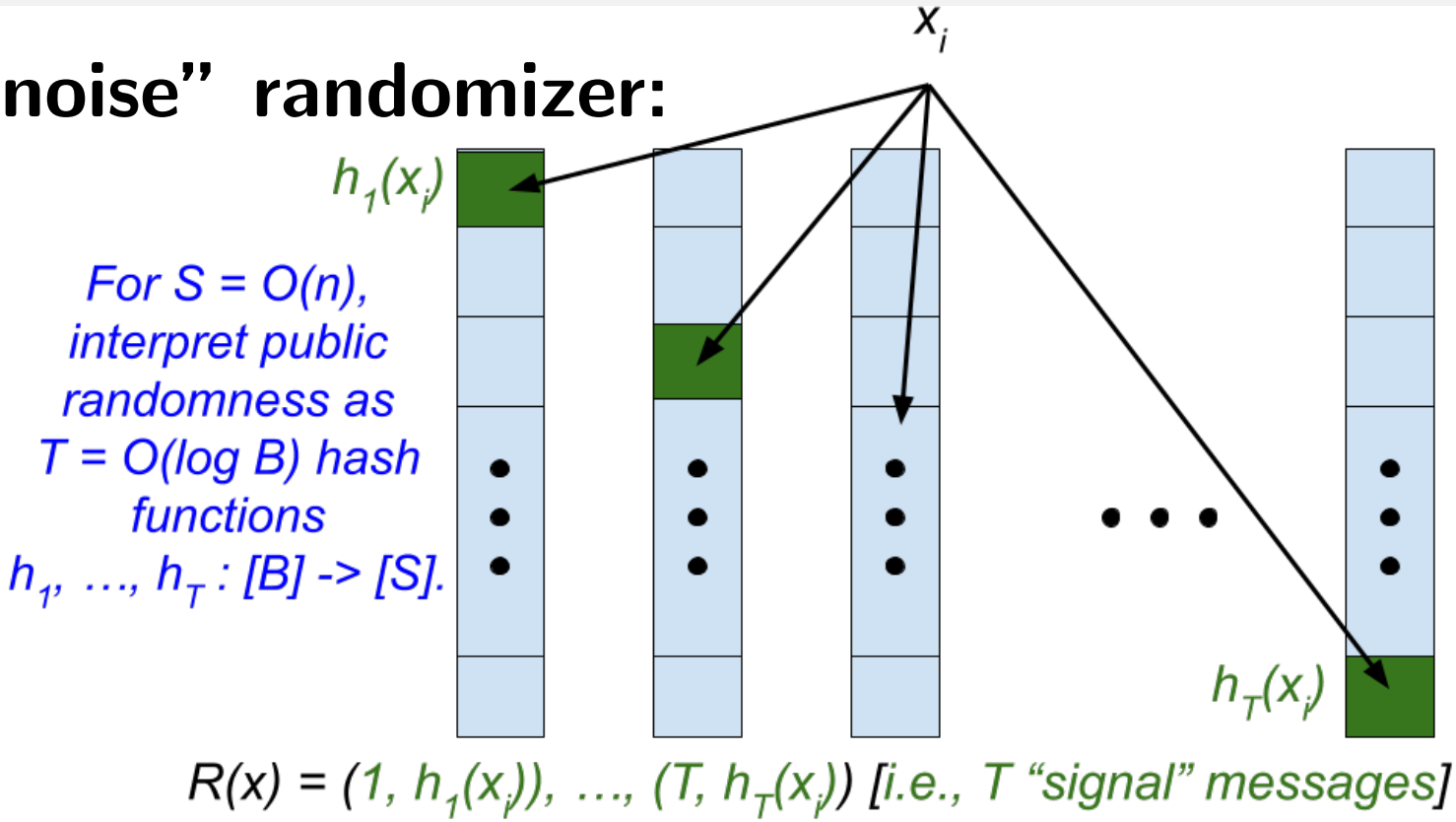


Analyzer:

- Let $C[t, s]$ be # messages of form (t, s).
- Define $\mathrm{FO}(j) := \min_{1 \le t \le T} C[t, h_t(j)]\ \forall j \in [B]$.

Non-DP count-min-based protocol

"No-noise" randomizer:



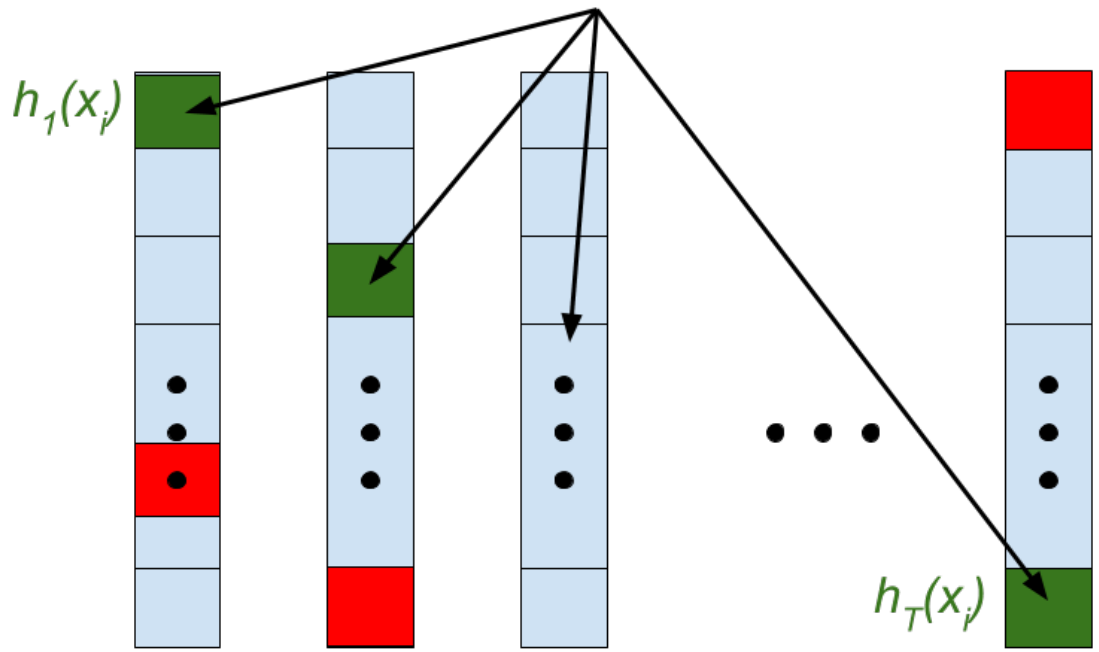
Analyzer:

- Let $C[t, s]$ be # messages of form (t, s).
- Define $\mathrm{FO}(j) := \min_{1 \le t \le T} C[t, h_t(j)]\ \forall j \in [B]$.

Lemma. With prob. $1 - 1/B^{O(1)}$, $\mathrm{FO}(j) = |\{i : x_i = j\}|$.

Private count-min sketch-based protocol

Noisy randomizer: Same as no-noise one, but also output each $(t, s) \in [T] \times [S]$ w/ prob. $p = \tilde{O}(1/n)$.

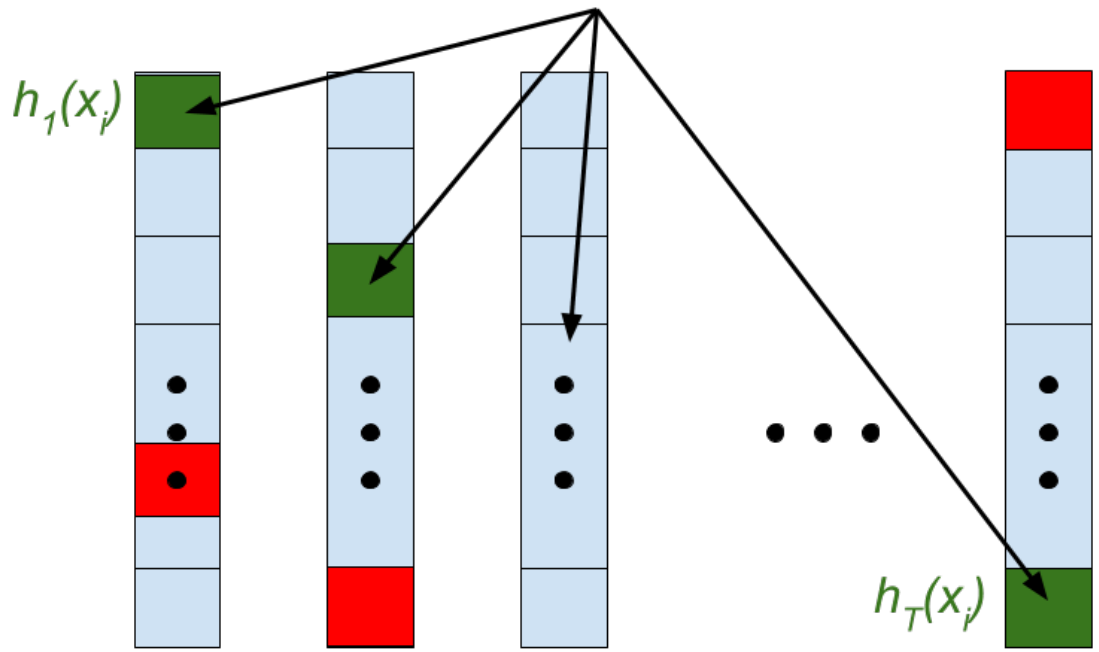


$R(x_i) = ((1, h_1(x_i)), \ldots, (T, h_T(x_i), [\text{“noise” msgs } (t,l)])$

Number of messages $\approx T + pST = \tilde{O}(1)$

Private count-min sketch-based protocol

Noisy randomizer: Same as no-noise one, but also output each $(t, s) \in [T] \times [S]$ w/ prob. $p = \tilde{O}(1/n)$.



$R(x_i) = ((1, h_1(x_i)), \ldots, (T, h_T(x_i), [\text{“noise” msgs } (t,l)])$

$\text{Number of messages} \approx T + pST = \tilde{O}(1)$

Analyzer: Same as for “no-noise” case (+ debiasing term); since count-min is *noise-stable*, get error $\tilde{O}(1)$.

Privacy analysis of count-min protocol

- Function $x \mapsto ((1, h_1(x)), \ldots, (T, h_T(x)))$ has ℓ_1-**sensitivity** $O(T) = O(\log B)$:
- *(I.e., if we change x, only $O(T)$ buckets change.)*
- Noise added to each bucket is Binom(n, p) – is sufficiently "smooth".
- Standard results: adding "smooth" noise to insensitive functions leads to privacy.

Recap: now assume no public coins...

Theorem (Our paper)

There exist multi-message shuffled model (ϵ, δ)-DP protocols with the below properties:

Technique	*error*	*comm.*	*query time*	*public coins*
Count-min	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	*Yes*
Hadamard Resp.	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(1)$	$\tilde{O}_\epsilon(n)$	*No*

Hadamard Response

[Kairouz et al., 16]; [Nguyen et al., 16]; [Acharya et al., 18]

Assume $B+1$ is power of 2, consider $B \times (B+1)$ **Hadamard matrix** H_B (*with 1st row removed*):

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \quad \leftarrow \text{e.g., } \Pr[R(2)=3] = \frac{1}{2(1+e^{\epsilon})}$$

$$R(x) = \begin{cases} \text{Unif. over } \{k \text{ s.t. } (H_B)_{x,j} = 1\} & \text{w/ prob. } \frac{e^{\epsilon}}{1+e^{\epsilon}} \\ \text{Unif. over } \{k \text{ s.t. } (H_B)_{x,j} = -1\} & \text{w/ prob. } \frac{1}{1+e^{\epsilon}} \end{cases}$$

Hadamard Response

[Kairouz et al., 16]; [Nguyen et al., 16]; [Acharya et al., 18]

Assume $B+1$ is power of 2, consider $B \times (B+1)$ **Hadamard matrix** H_B (*with 1st row removed*):

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \leftarrow \text{e.g., } \Pr[R(2)=3] = \frac{1}{2(1+e^\epsilon)}$$

$$R(x) = \begin{cases} \text{Unif. over } \{k \text{ s.t. } (H_B)_{x,j} = 1\} & \text{w/ prob. } \frac{e^\epsilon}{1+e^\epsilon} \\ \text{Unif. over } \{k \text{ s.t. } (H_B)_{x,j} = -1\} & \text{w/ prob. } \frac{1}{1+e^\epsilon} \end{cases}$$

Aside. Hadamard Response gets (optimal) error $O\left(\frac{\sqrt{n \log B}}{\epsilon}\right)$ in *local model*.

Non-DP Hadamard Response variant

(Towards Hadamard Response variant for shuffled model)

"No-noise" randomizer: Let $T = \log_2 n$,
$R(x_i)$ is *T random indices of Hadamard codeword*:

$$R(x_i) = (a_{i,1}, \ldots, a_{i,T}), \qquad a_{i,t} \sim \mathrm{Unif}(\{k : (H_B)_{x_i,k} = 1\}).$$

$R(x_i)$ is 1 message of $\log(B+1)\log(n)$ bits.

Non-DP Hadamard Response variant

(Towards Hadamard Response variant for shuffled model)

"No-noise" randomizer: Let $T = \log_2 n$,
$R(x_i)$ is *T random indices of Hadamard codeword*:

$$R(x_i) = (a_{i,1}, \ldots, a_{i,T}), \qquad a_{i,t} \sim \text{Unif}(\{k : (H_B)_{x_i,k} = 1\}).$$

$R(x_i)$ is 1 message of $\log(B+1)\log(n)$ bits.

Analyzer: Given query $j \in [B]$, output *number of messages $(a_{i,1}, \ldots, a_{i,T})$ "consistent with" j*:

$$\text{FO}(j) := |\{i : H_{j,a_{i,t}} = 1 \ \ \forall t \in [T]\}|.$$

Non-DP Hadamard Response variant

(Towards Hadamard Response variant for shuffled model)

"No-noise" randomizer: Let $T = \log_2 n$,
$R(x_i)$ is *T random indices of Hadamard codeword*:

$$R(x_i) = (a_{i,1}, \ldots, a_{i,T}), \qquad a_{i,t} \sim \mathrm{Unif}(\{k : (H_B)_{x_i,k} = 1\}).$$

$R(x_i)$ is 1 message of $\log(B+1)\log(n)$ bits.

Analyzer: Given query $j \in [B]$, output *number of messages $(a_{i,1}, \ldots, a_{i,T})$ "consistent with" j*:

$$\mathrm{FO}(j) := |\{i : H_{j,a_{i,t}} = 1 \ \ \forall t \in [T]\}|.$$

Lemma. $\mathbb{E}[\mathrm{FO}(j) - |\{i : x_i = j\}|] = O(1)$.

(Proof follows from *orthogonality* of any 2 rows of H_B.)

DP Hadamard Response Variant

Noisy randomizer: Let $T = \log_2 n$, $\rho = O\left(\frac{\log 1/\delta}{\epsilon^2}\right)$.

$R(x_i) = (a_i, \tilde{a}_i^1, \ldots, \tilde{a}_i^\rho)$: 1 "signal msg", ρ "noise msgs"

$a_i = (a_{i,1}, \ldots, a_{i,T})$ $\quad a_{i,t} \sim \text{Unif}(\{k : (H_B)_{x_i,k} = 1\})$

$\tilde{a}_i^g \sim \text{Unif}([B+1]^T)$ $\quad g = 1, 2, \ldots, \rho.$

DP Hadamard Response Variant

Noisy randomizer: Let $T = \log_2 n, \rho = O\left(\frac{\log 1/\delta}{\epsilon^2}\right)$.

$R(x_i) = (a_i, \tilde{a}_i^1, \ldots, \tilde{a}_i^\rho)$: 1 "signal msg", ρ "noise msgs"

$a_i = (a_{i,1}, \ldots, a_{i,T})$ $\quad a_{i,t} \sim \text{Unif}(\{k : (H_B)_{x_i,k} = 1\})$

$\tilde{a}_i^g \sim \text{Unif}([B+1]^T)$ $\quad g = 1, 2, \ldots, \rho.$

Analyzer: Similar to no-noise case (as is error guarantee).

Privacy: Number of noisy messages $\tilde{a}_i^g$ consistent with each row of H_B is Binom $\left(n, \frac{\log(1/\delta)}{\epsilon^2}\right)$ (which is "smooth").

Application of histograms: range queries

Range queries: Again $\mathcal{X} = [B]$, but want to answer all $O(B^2)$ queries of the form:

Given as input $1 \leq j_1 \leq j_2 \leq B$, how many $x_i \in [j_1, j_2]$?

Corollary (Our paper)

Can answer all range queries on $[B]$ *(simultaneously) with:*

- *Error* $\tilde{O}_\epsilon(1)$ *(max error over all range queries).*
- *Per-user communication* $\tilde{O}_\epsilon(1)$.

Application of histograms: range queries

Range queries: Again $\mathcal{X} = [B]$, but want to answer all $O(B^2)$ queries of the form:

Given as input $1 \leq j_1 \leq j_2 \leq B$, how many $x_i \in [j_1, j_2]$?

Corollary (Our paper)

Can answer all range queries on $[B]$ (simultaneously) with:

- *Error $\tilde{O}_\epsilon(1)$ (max error over all range queries).*
- *Per-user communication $\tilde{O}_\epsilon(1)$.*

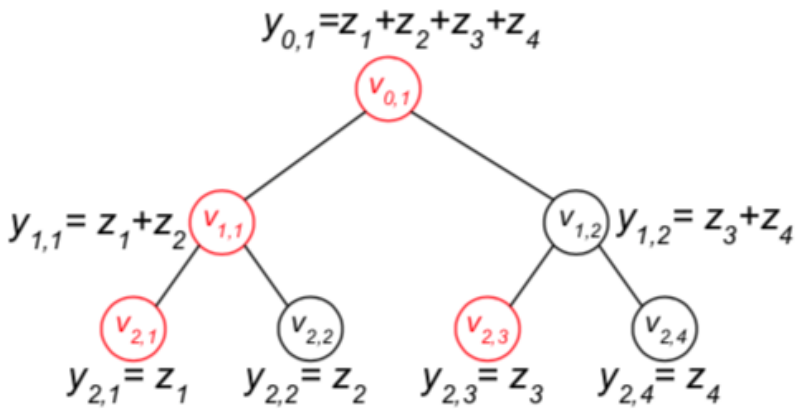
Same holds for d-dimensional range queries ($d = O(1)$).

Reduction: range queries to histograms

[Li et al., 10], [Chan et al., 11]

Range query trees:

- Suppose z_j users hold $j \in [B]$;
- Binary tree, each node v stores sum of all z_j in leaves that are descendents of v.
- Can compute all range queries from set of B red nodes:



- Think of each user at $j \in [B]$ holding the $\leq \log B$ items corresponding to red nodes above j.

M-estimation of median

- Now suppose $x_1, \ldots, x_n \in [0, 1]$.
- $\text{Median}(x_1, \ldots, x_n) = \arg\min_i \sum_{i'=1}^{n} |x_i - x_{i'}|$.
- Measure error of estimate y of $\text{Median}(x_1, \ldots, x_n)$:

$$M(y) := \sum_{i=1}^{n} |y - x_i|.$$

M-estimation of median

- Now suppose $x_1, \ldots, x_n \in [0, 1]$.
- $\text{Median}(x_1, \ldots, x_n) = \arg\min_i \sum_{i'=1}^{n} |x_i - x_{i'}|$.
- Measure error of estimate y of $\text{Median}(x_1, \ldots, x_n)$:

$$M(y) := \sum_{i=1}^{n} |y - x_i|.$$

Corollary (Our paper)

There is a shuffled-model protocol with per-user communication $\tilde{O}(1)$, and outputs $y \in [0, 1]$ so that

$$M(y) \leq M(\text{Median}(x_1, \ldots, x_n)) + \tilde{O}(1).$$

[Duchi et al., 13]: Best M-error in local model is $\Omega_\epsilon(\sqrt{n})$.

Outline

1. Review of central + local models of DP.
2. Shuffled model of DP [Bittau et al., '16].
3. Lower bounds for histogram computation & selection for *single*-message shuffled model [our paper].
   - Nearly tight, improve upon [Cheu et al., '18].
4. Upper bounds for histogram computation for *multi*-message shuffled model [our paper].
   - Applications to range queries, quantile estimation, etc.

Open questions

- Multi-message histograms: get no public coins and $\tilde{O}(1)$ computation per frequency query? (i.e., best of both Hadamard response + count-min)
- Separation in error between central + multi-message shuffled models?
- One candidate: Multi-message selection: [Cheu et al., 18] noted a protocol with $n = O(\sqrt{B})$ – optimal?
- Lower bound for *single-message* M-estimation of median?