Differentially Private Aggregation in Shuffle Model:

Almost Central Accuracy in Almost a Single Message

Badih Ghazi
Google Research
Mountain View

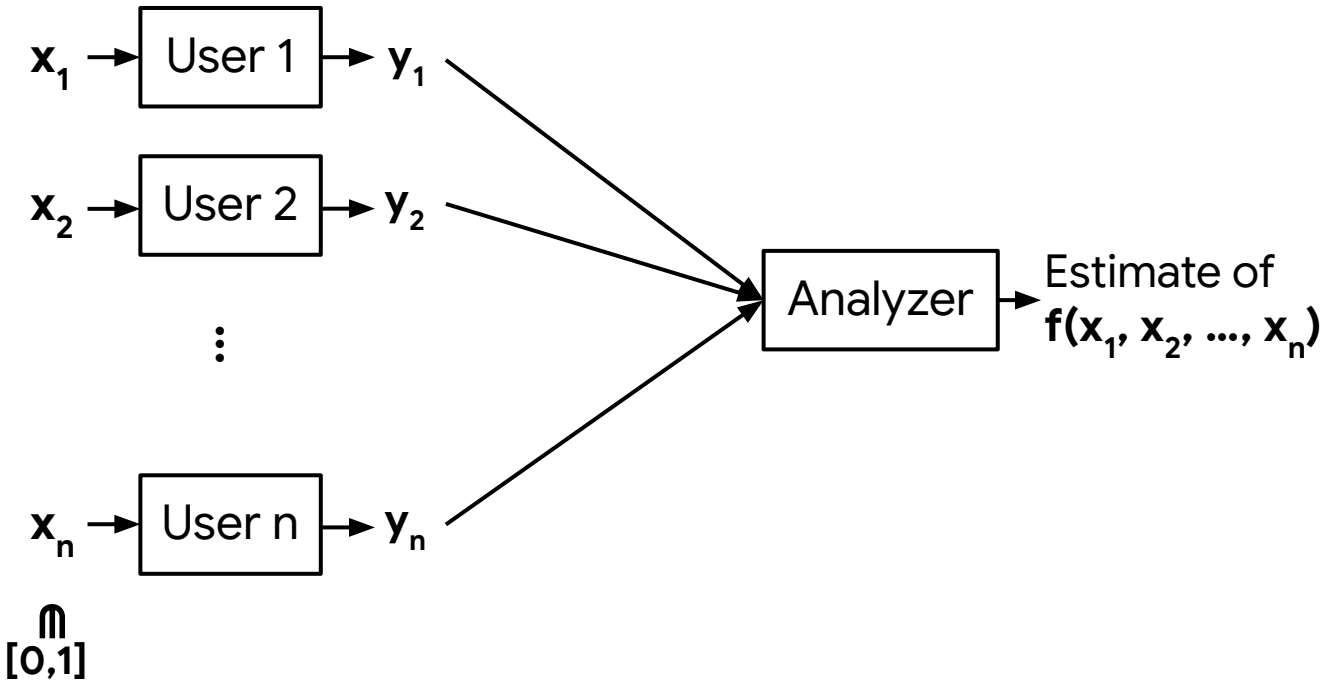
Ravi Kumar
Google Research
Mountain View

Pasin Manurangsi
Google Research
Mountain View

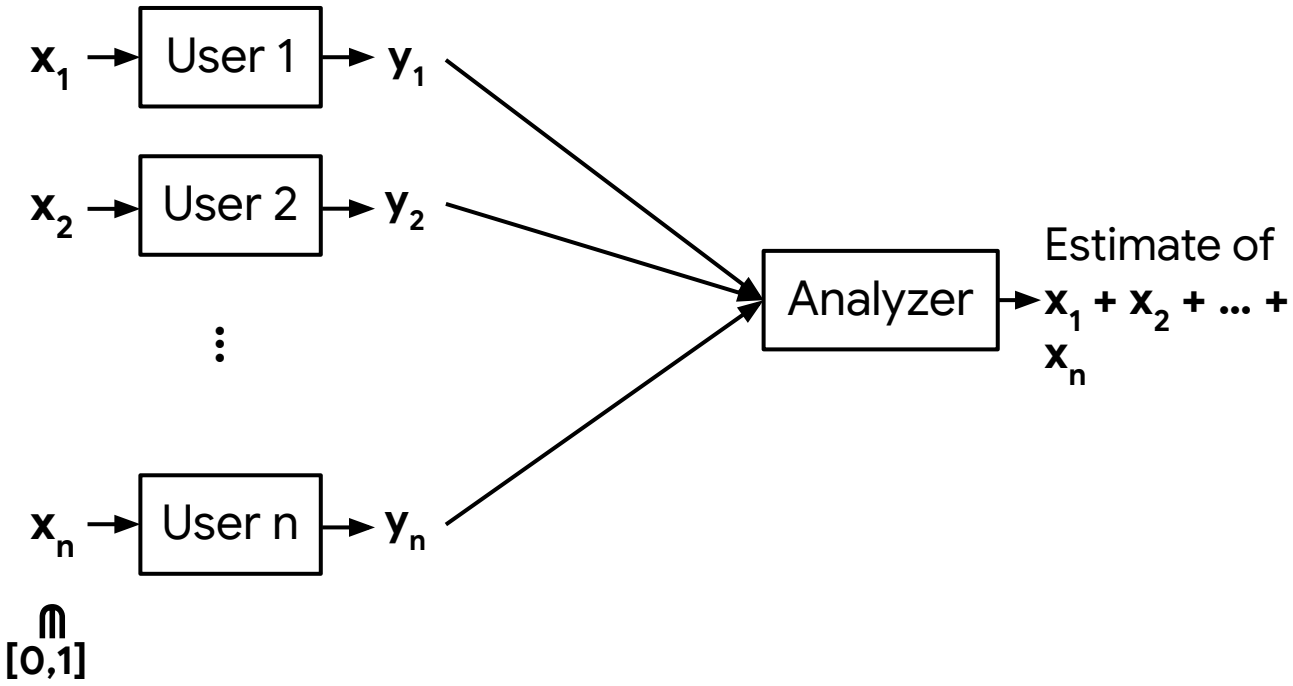
Rasmus Pagh
U. Copenhagen &
Google Research

Amer Sinha
Google
San Bruno

Distributed Analytics

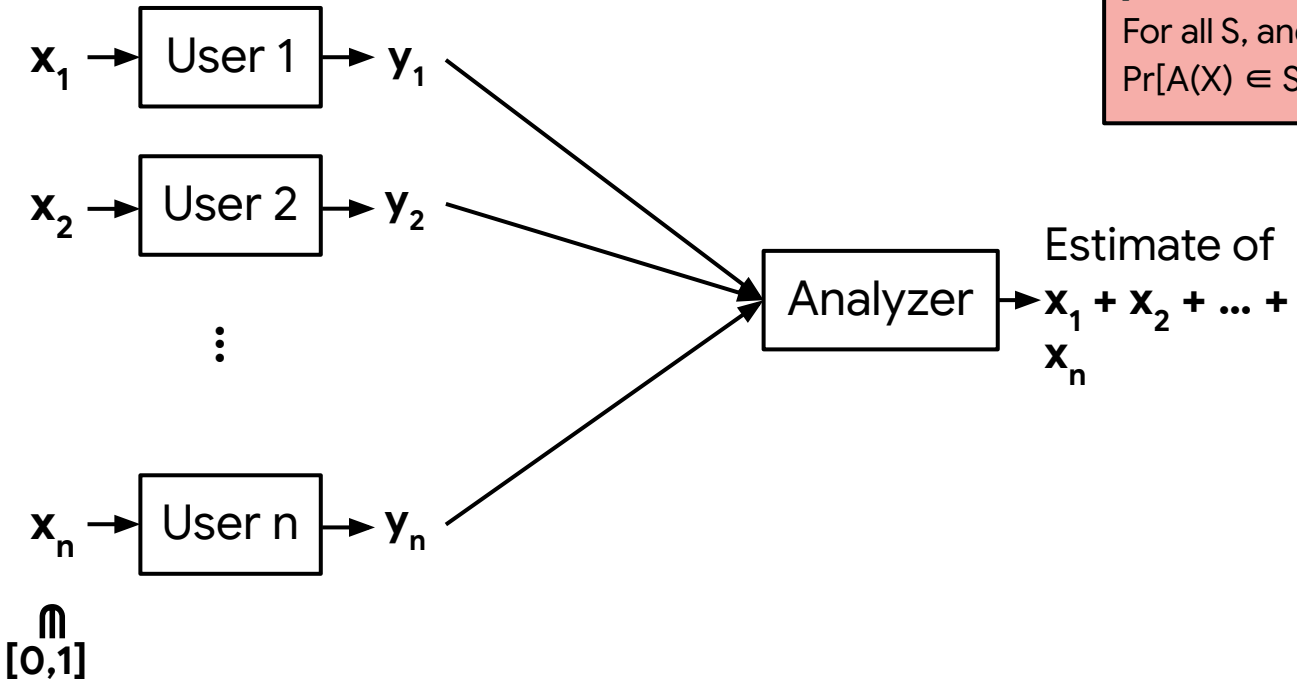


Distributed Analytics

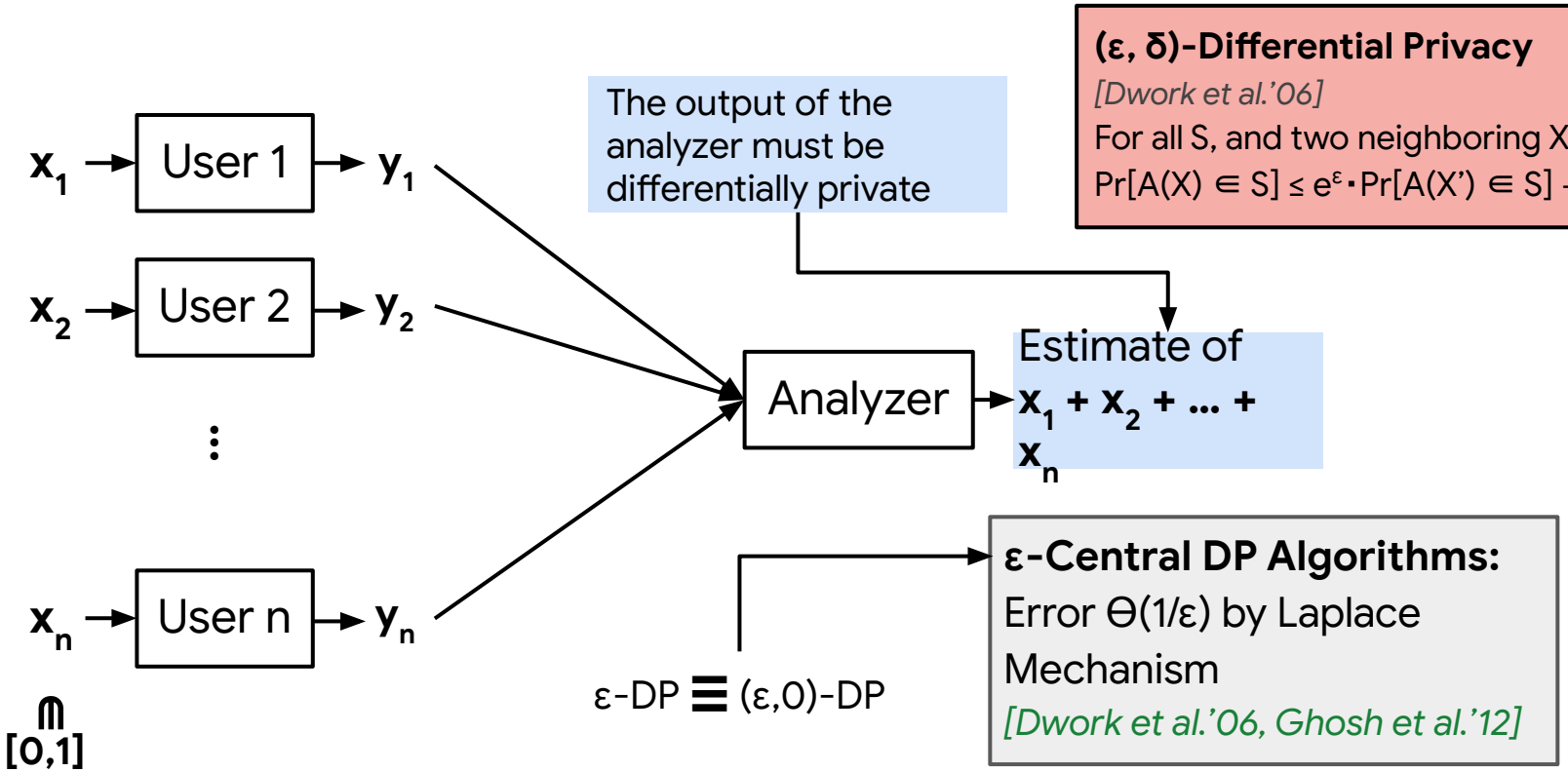


Differential Privacy *[Dwork et al.]*

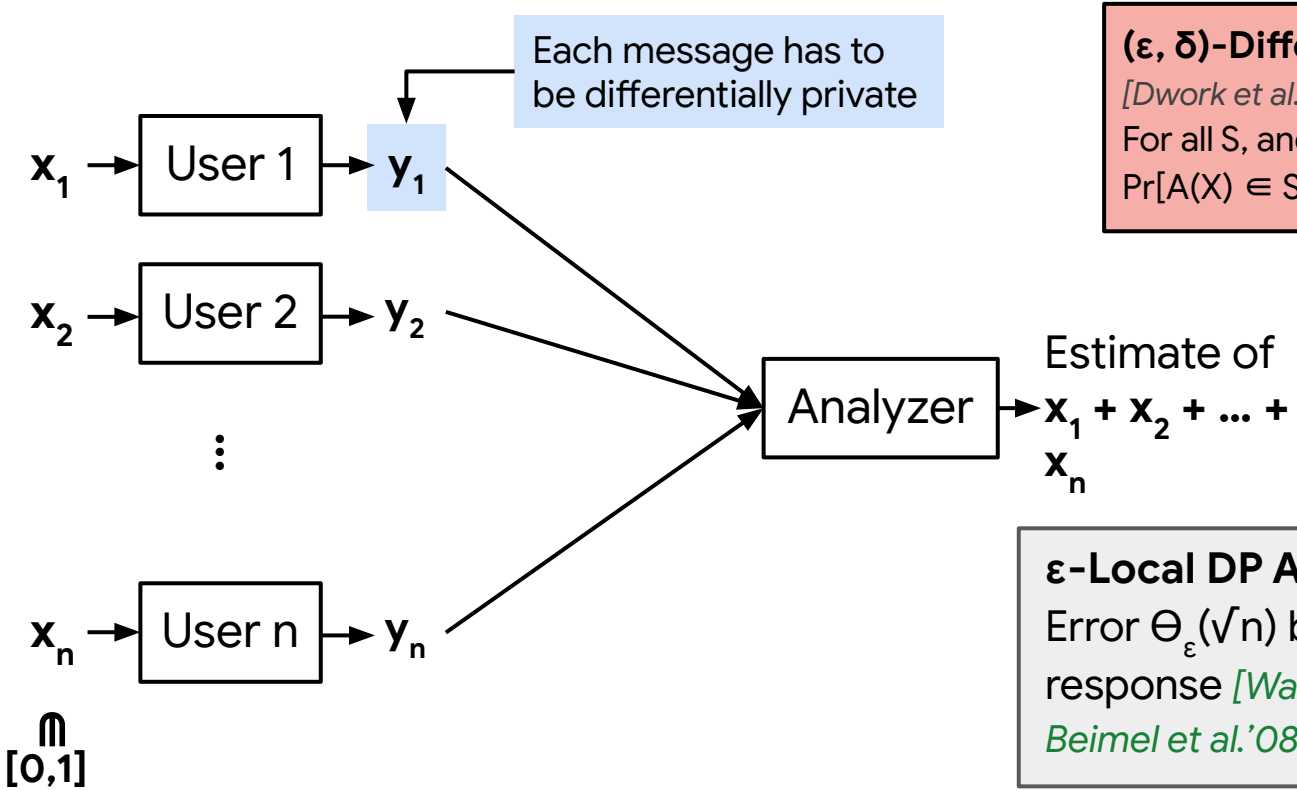
(ε, δ)-Differential Privacy
[Dwork et al.'06]
For all S, and two neighboring X, X'
$\Pr[A(X) \in S] \leq e^{\varepsilon} \cdot \Pr[A(X') \in S] + \delta$



Differential Privacy: **Central Model** *[Dwork et al.]*



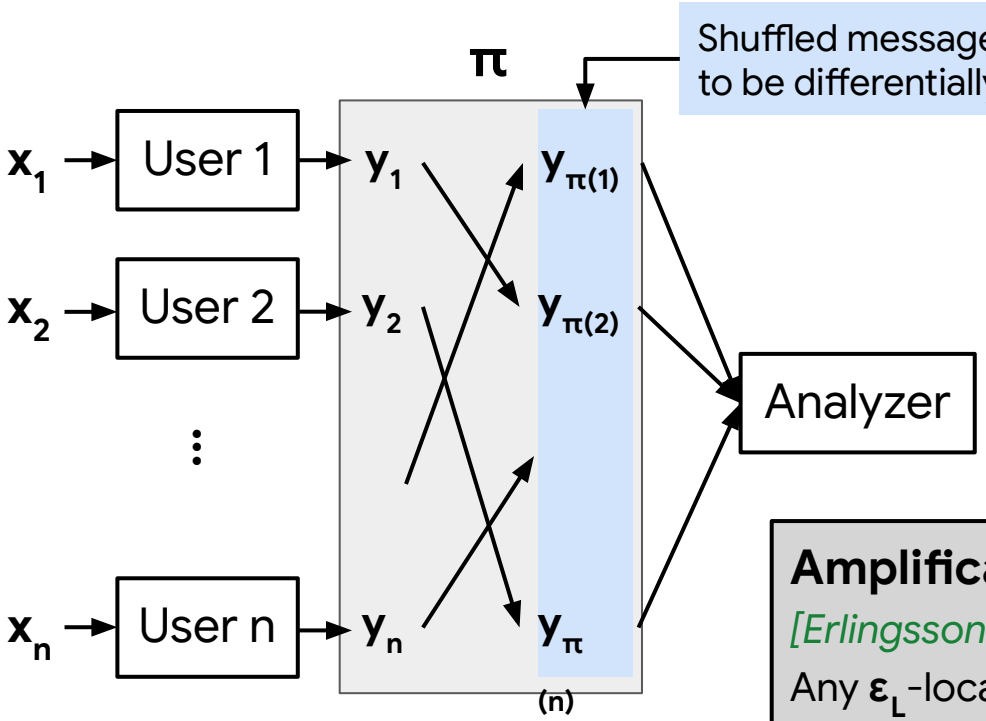
Differential Privacy: **Local Model** *[Kasiviswanathan et al.]*



(ε, δ)-Differential Privacy
[Dwork et al.'06]
For all S, and two neighboring X, X'
$\Pr[A(X) \in S] \le e^{\varepsilon} \cdot \Pr[A(X') \in S] + \delta$

ε-Local DP Algorithms:
Error $\Theta_{\varepsilon}(\sqrt{n})$ by randomized response *[Warner'65, Beimel et al.'08, Chan et al.'12]*

Differential Privacy: **Shuffled Model** *[Bittau et al., Erlingsson et al.]*



Shuffled messages have to be differentially private ≡ Multiset of messages have to be differentially private

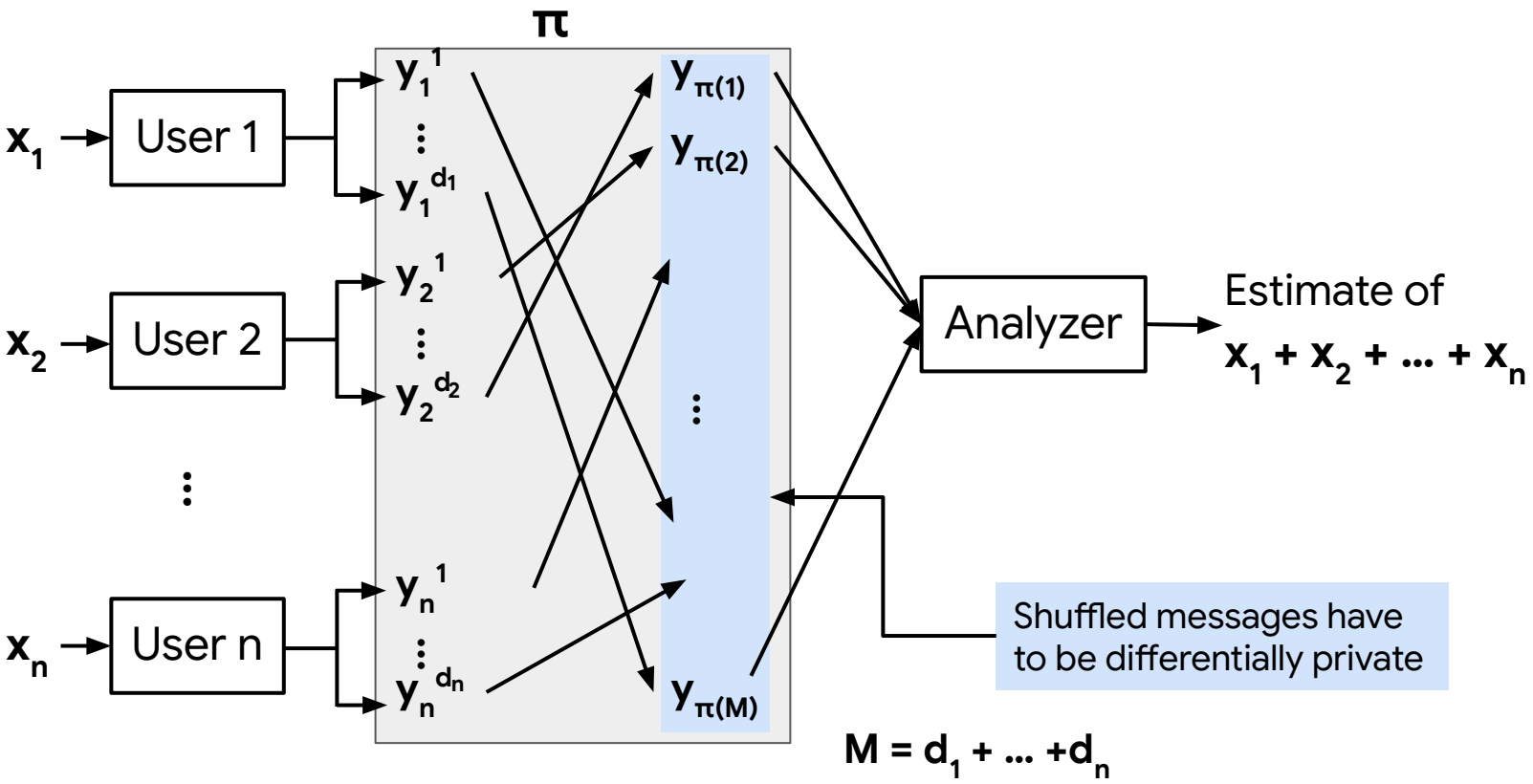
- **Anonymous Model** *[Ishai et al.]*
- **Encode-Shuffle-Analyze Architecture** *[Bittau et al.]*
- **Shuffled/Shuffle Model** *[Cheu et al., Erlingsson et al.]*

Amplification by Shuffling (Informal)
[Erlingsson et al., Balle et al.]
Any ε_L-local DP algorithm is **(ε, δ)**-shuffled DP for $\varepsilon \ll \varepsilon_L$ and for reasonable value of **δ**.

Real Summation

		Error	# messages per user	Bits per message
ε-Central DP	Laplace Mechanism *[Dwork et al.'06, Ghosh et al.'12]*	$\Theta(1/\varepsilon)$	1	O(log n)
ε-Local DP	Randomized Response *[Warner'65, Beimel et al.'08]*	$\Theta_\varepsilon(\sqrt{n})$	1	O(log n)
(ε,δ)-Shuffled DP	*[Balle et al.' 19]*	$\Theta_\varepsilon(n^{1/6})$	1	O(log n)

Shuffled Model: Multi-Message Setting



Real Summation

		Error	# messages per user	Bits per message
ε-Central DP	Laplace Mechanism *[Dwork et al.'06, Ghosh et al.'12]*	$\Theta(1/\varepsilon)$	1	$O(\log n)$
ε-Local DP	Randomized Response *[Warner'65, Beimel et al.'08]*	$\Theta_\varepsilon(\sqrt{n})$	1	$O(\log n)$
(ε,δ)-Shuffled DP	*[Balle et al.' 19]*	$\Theta_\varepsilon(n^{1/6})$	1	$O(\log n)$
	[Cheu et al.'19]	$O\left(\frac{\log(1/\delta)}{\varepsilon}\right)$	$O_\varepsilon(\sqrt{n})$	1
	[Balle et al.'20, Ghazi et al.' 20]	$O(1/\varepsilon)$	$O\left(1+\frac{\log(1/\delta)}{\log n}\right)$	$O(\log n)$
	This work	$O(1/\varepsilon)$	$1+O_\varepsilon\left(\frac{\log(1/\delta)}{\sqrt{n}}\right)$	$O(\log n)$
ε-Shuffled DP	*[Ghazi et al.'20]*	$O(1/\varepsilon^{1.5})$	$O_\varepsilon(\log^3 n)$	$O(\log\log n)$

Longer talk available here