Optimal Algorithms for Mean Estimation under Local Differential Privacy

Hilal Asi — Stanford University

Vitaly Feldman — Apple

Kunal Talwar — Apple

Mean estimation under local differential privacy

v_1 v_2 $\cdots$ v_n

$v_i \in \mathrm{R}^d$

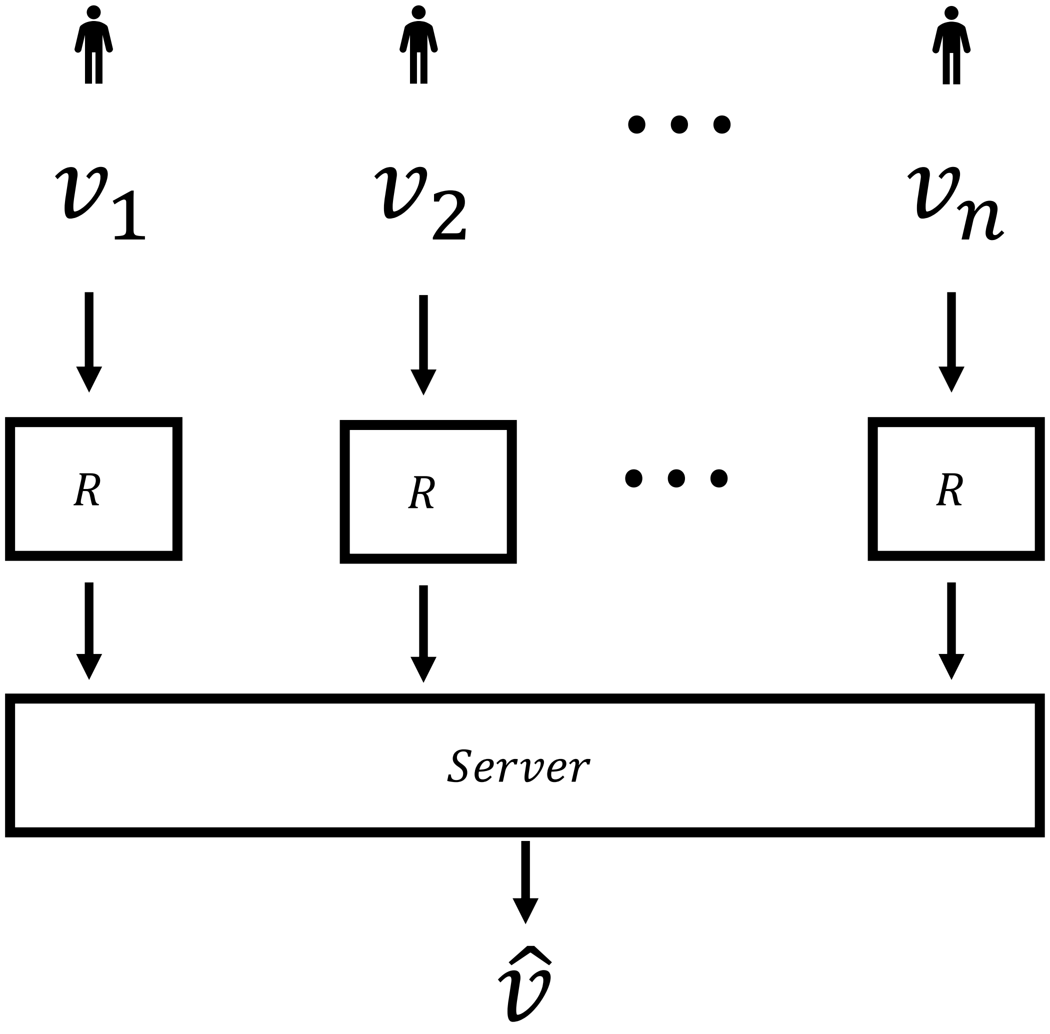
$||v_i||_2 = 1$

Goal: estimate the mean $\frac{1}{n}\sum_{i=1}^{n} v_i$ under local differential privacy

Why:

1. Building histograms
2. Training ML models

Mean estimation under local differential privacy



$v_i \in \mathrm{R}^d$

$||v_i||_2 = 1$

***R* is *ε*-differentially private (*ε*-DP)**

For every v, v' and u:

$$\frac{P(R(v) = u)}{P(R(v') = u)} \leq e^{\varepsilon}$$

Goal: output $\hat{\boldsymbol{v}}$ such that:

1. unbiased: $\mathrm{E}[\hat{\boldsymbol{v}}] = \frac{1}{n}\sum_{i=1}^{n} v_i$

2. Min. variance: $\mathrm{E}[||\ \hat{v}\ - \frac{1}{n}\sum_{i=1}^{n} v_i\ ||_2^2]$

Error of a protocol

- Error for **randomizer R** and **server aggregator A**:

$$Err(A, R) = \max_{v_1, \ldots, v_n \in R^d} E\left[|| \ A\big(R(v_1), \ldots, R(v_n)\big) - \frac{1}{n} \sum_{i=1}^{n} v_i \ ||_2^2 \right]$$

- Minimax error:

$$\mathrm{E}rr^{\star}(n, d, \varepsilon) = \min_{R,A} Err(A, R)$$

Prior results

- Minimax rates are known asymptotically [DJW18, BDFKR18, DR18]:

$$\mathrm{Err}^{\star}(n, d, \varepsilon) = \Theta\left(\frac{d}{n\,\varepsilon}\right)$$

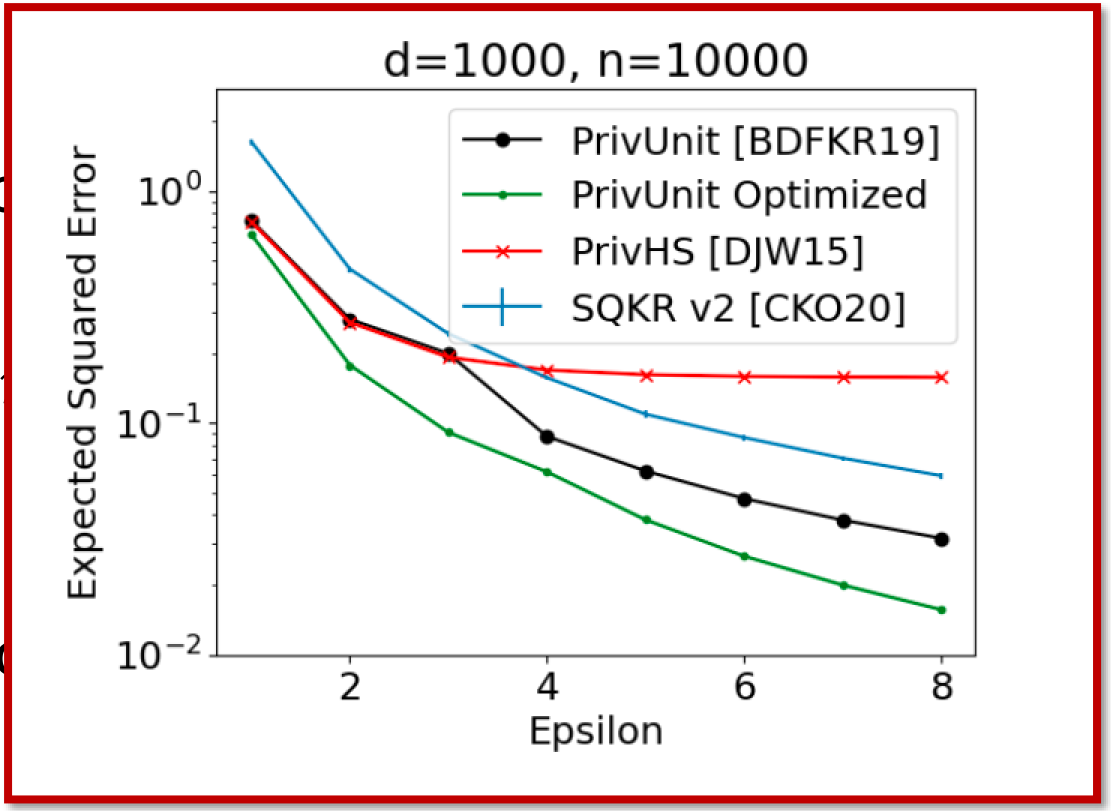
- Completely solved?
- Many (asymptotically) optimal algorithms exists [DJW18, BDFKR18, CKO20, FT21]
- Significant gaps in practice

Prior results

- Minimax rates are known asymp

$Err^\star($

- Completely solved?
- Many (asymptotically) optimal alg
- Significant gaps in practice



What is the optimal algorithm for mean estimation under local DP?

Main results

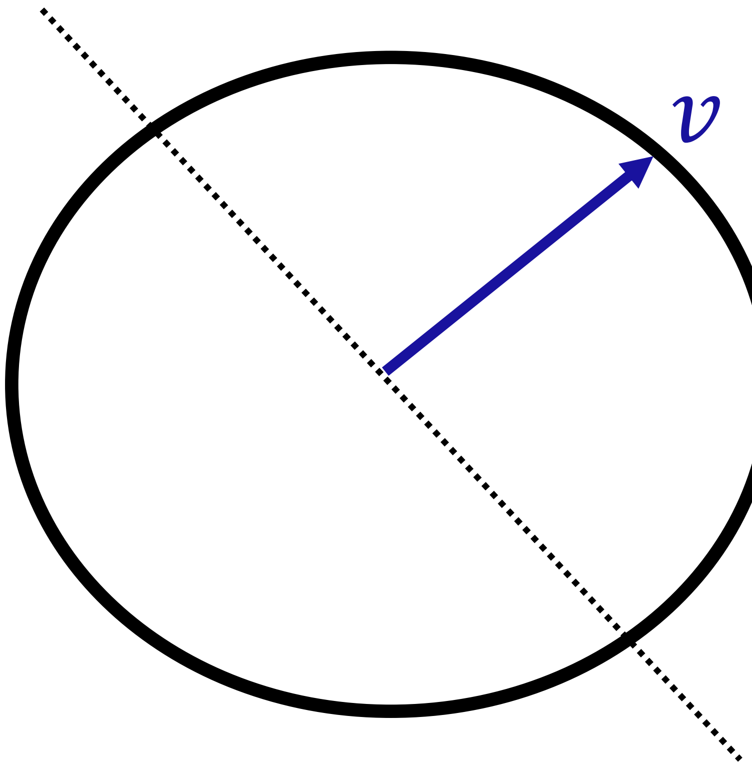
1. The optimal (unbiased) randomizer is an instance of PrivUnit
   - Optimal parameters can be found analytically

2. PrivUnitG: a Gaussian version of PrivUnit
   - Easier to analyze ➔ easier to find optimal parameters
   - Allows to estimate constants of minimax error:

$$\mathrm{E}rr^{\star}(n, d, \varepsilon) = \mathrm{c}_{d,\varepsilon} \cdot \frac{d}{n\,\varepsilon}$$

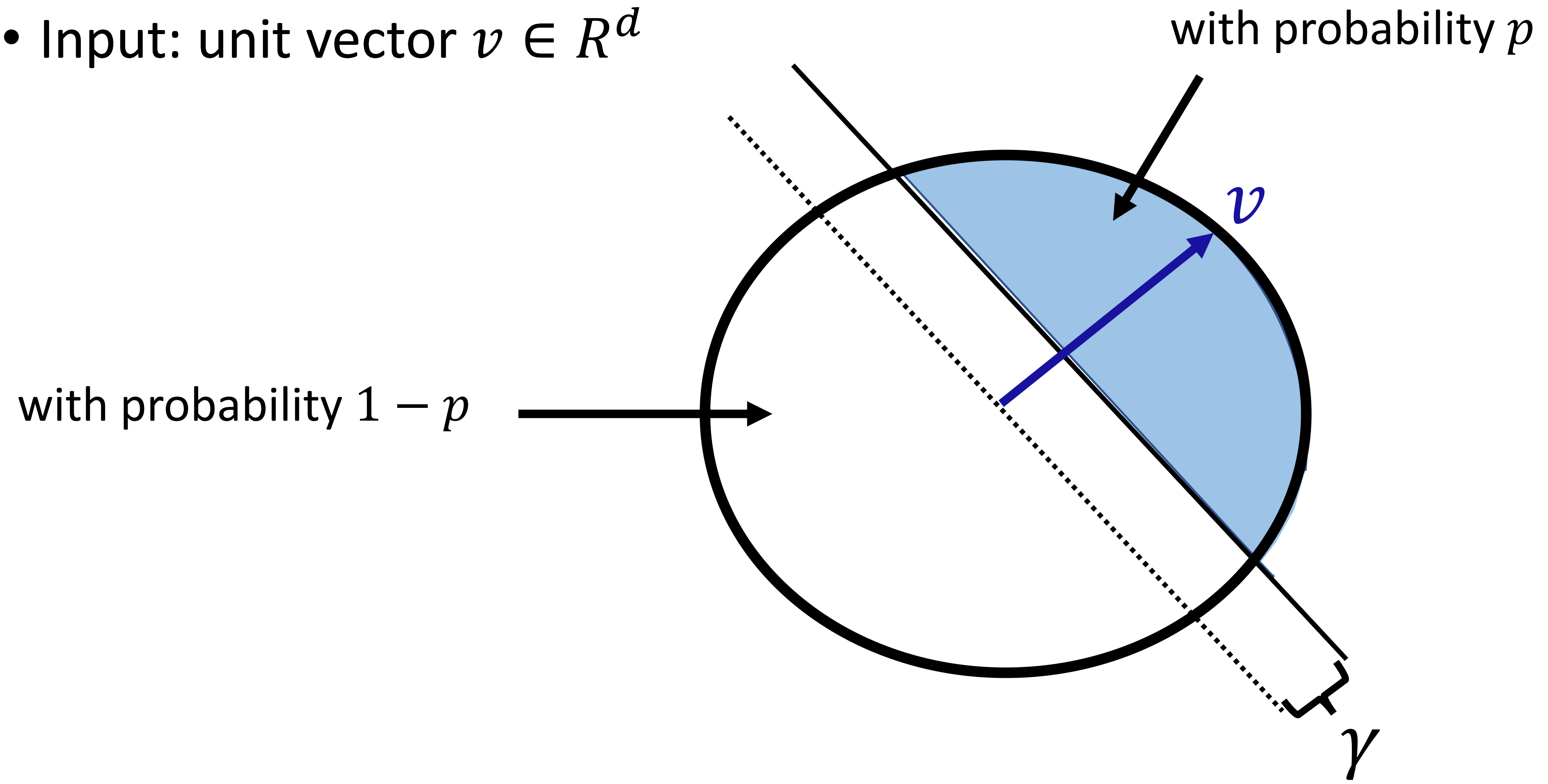
PrivUnit [BDFKR18]

- Input: unit vector $v \in R^d$

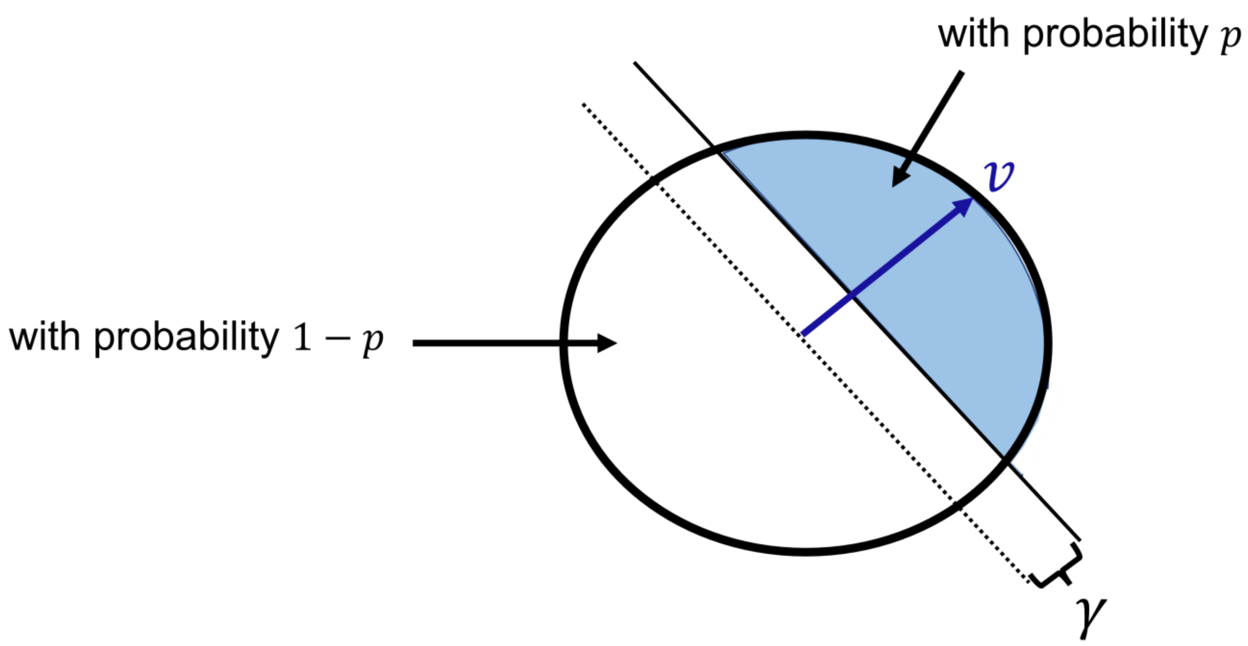


PrivUnit [BDFKR18]

- Input: unit vector $v \in R^d$



PrivUnit [BDFKR18]



$$PrivUnit(v) = C \cdot \begin{cases} Unif\ \{u: \langle u, v\rangle \geq \gamma\ \} & w.p.\ p \\ Unif\{u: \langle u, v\rangle < \gamma\} & w.p. 1-p \end{cases}$$

PrivUnit [BDFKR18]

- **Unbiased:** $E[PrivUnit(v)] = C \cdot v$

Idea: 1. $PrivUnit$ depends only on inner products with v

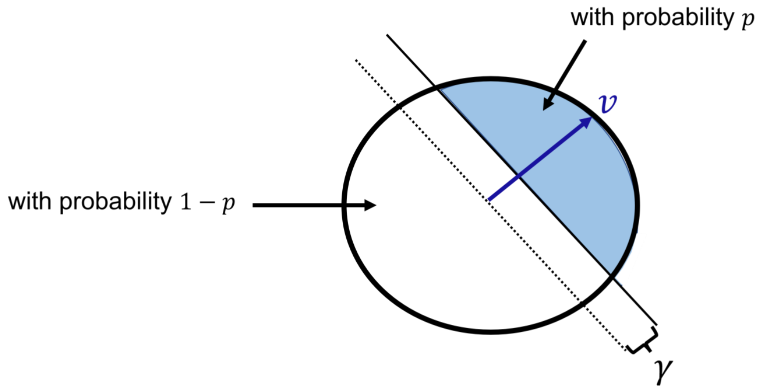
2. if $\langle u\,, v\rangle = 0$ then $\langle -u\,, v\rangle = 0$

- **Privacy:** $\varepsilon = \ln\left(\frac{p/(1-q)}{(1-p)/q}\right)$

$q = P_{U\sim Unif}(\langle U, v\rangle \leq \gamma)$

Idea: 1. $PrivUnit$ probability density has two possible values

2. Ratio between densities is bounded



Optimality of PrivUnit

Theorem [A., Feldman, Talwar 22].

For any d and ε, there is $p^\star$ and $\gamma^\star$ such that PrivUnit achieves the optimal mean squared error amongst all unbiased protocols.

For any ε-DP randomizer R and server aggregator A that is unbiased:

$$\mathrm{Err}\left(PrivUnit_{p^\star,\gamma^\star}, A^+\right) \leq Err(R, A)$$

Additive aggregation: $A^+\left(R(v_1), \dots, R(v_n)\right) = \frac{1}{n}\sum_{i=1}^{n} R(v_i)$

Proof idea

Step 1: optimal randomizer R has output space R^d and is unbiased

Step 2: PrivUnit is optimal amongst real-valued randomizers

Proof idea

Step 1: optimal randomizer R has output space R^d and is unbiased

Idea: use the server aggregation with **fake** inputs to transform the output space

$$\hat{R}(v) = E[A(R(v), R(v_2'), \ldots, R(v_n'))]$$

$v_2', \ldots, v_n'$ are iid uniform over the sphere

Step 2: PrivUnit is optimal amongst real-valued randomizers

Idea:

prove several structural properties of the optimal algorithm

$\downarrow$

Optimal algorithm is an instance of $PrivUnit$

Step 2: PrivUnit is optimal amongst real-valued randomizers

1. **Rotational symmetry:** $R(v)$ and $R(v')$ are the same up to rotations

   ➔ Enough to study the randomizer for a fixed v

2. **Linear program** formulation of the best randomizer: $p_j = P(R(v) = u_j)$

$$\min \quad \sum_{j=1}^{M} p_j ||u_j - v||_2^2$$

$$\text{s.t.} \quad \sum_{j=1}^{M} p_j = 1\,, p_j \geq 0 \qquad \text{probability distribution}$$

$$\sum_{j=1}^{M} u_j p_j = v \qquad \text{unbiased}$$

$$e^{-\varepsilon} \leq \frac{p_j}{p_{j'}} \leq e^{\varepsilon} \qquad \text{privacy}$$

Step 2: PrivUnit is optimal amongst real-valued randomizers

1. **Rotational symmetry:** $R(v)$ and $R(v')$ are the same up to rotations

   ➔ Enough to study the randomizer for a fixed v

2. **Linear program** formulation of the best randomizer: $p_j = P(R(v) = u_j)$

$$\min \quad \sum_{j=1}^{M} p_j ||u_j - v||_2^2$$

Key Lemma

R is symmetric ε-DP local randomizer **iff** $1 \leq \frac{\mathrm{p}(R(v)=u)}{\mathbf{p}} \leq e^{\varepsilon}$

$1 \leq \frac{p_j}{p} \leq e^{\varepsilon}$ 🡐 $e^{-\varepsilon} \leq \frac{p_j}{p_{j'}} \leq e^{\varepsilon}$ privacy

Step 2: PrivUnit is optimal amongst real-valued randomizers

1. **Rotational symmetry:** $R(v)$ and $R(v')$ are the same up to rotations

   ➔ Enough to study the randomizer for a fixed v

2. **Linear program** formulation of the best randomizer: $p_j = P(R(v) = u_j)$

$$\min \quad \sum_{j=1}^{M} p_j ||u_j - v||_2^2$$

$$\text{s.t.} \quad \sum_{j=1}^{M} p_j = 1\,, p_j \geq 0 \qquad \text{probability distribution}$$

$$\sum_{j=1}^{M} u_j p_j = v \qquad \text{unbiased}$$

$$1 \leq \frac{p_j}{p} \leq e^{\varepsilon} \qquad \text{privacy}$$

$p_j \in \{1, e^{\varepsilon}\} \cdot p$ ⇐

M linearly independent constraints must be satisfied

Step 2: PrivUnit is optimal amongst real-valued randomizers

1. **Rotational symmetry:** $R(v)$ and $R(v')$ are the same up to rotations

➔ Enough to study the randomizer for a fixed v

2. **Linear program** formulation of the best randomizer

➔ Optimal randomizer has two values for the density:

$$P\big(R(v) = u_j\big) \in \{1, e^{\varepsilon}\} \cdot p$$

Step 2: PrivUnit is optimal amongst real-valued randomizers

1. **Rotational symmetry:** $R(v)$ and $R(v')$ are the same up to rotations

   ➔ Enough to study the randomizer for a fixed v

2. **Linear program** formulation of the best randomizer

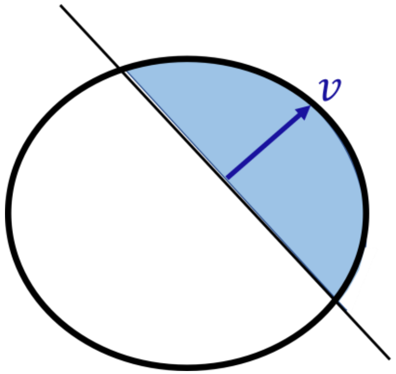
   ➔ Optimal randomizer has two values for the density:

$$P\big(R(v) = u_j\big) \in \{1, e^{\varepsilon}\} \cdot p$$

3. $\mathrm{u}_1^{\mathrm{T}} v > \mathrm{u}_2^{\mathrm{T}} v$ implies $P(R(v) = u_1) \geq P(R(v) = u_2)$

   - Otherwise can improve the error

Instance of PrivUnit



PrivUnitG

Idea: approximate the uniform distribution in PrivUnit using Gaussian

Let $U \sim Unif\{u \in \mathrm{R}^d : \ ||u||_2 = 1\}$

$$PrivUnit(v) = C \cdot \begin{cases} U \mid \langle U, v\rangle \geq \gamma & w.p.\ p \\ U \mid \langle U, v\rangle < \gamma & w.p. 1-p \end{cases}$$

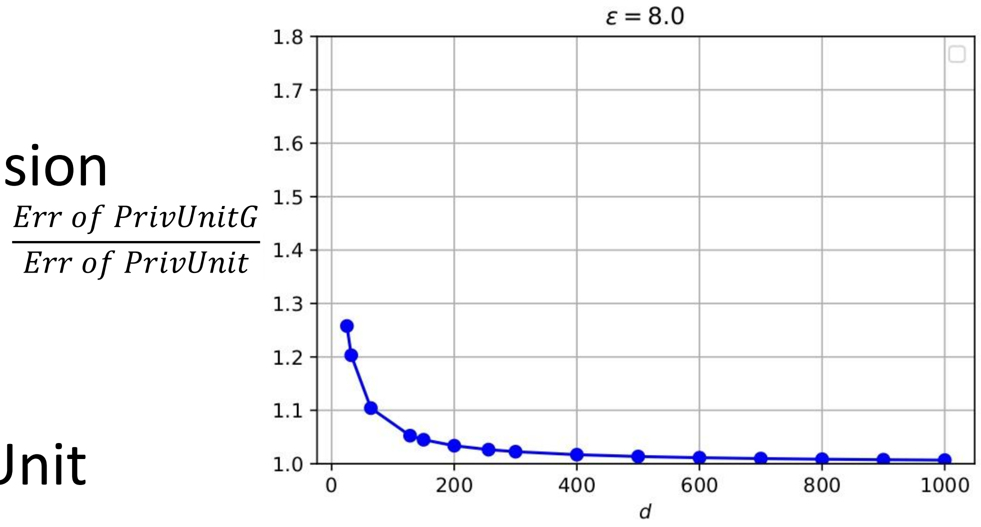
Let $W \sim Normal\left(0, \frac{1}{d}\right)$

Easier to analyze

$$PrivUnitG(v) = C \cdot \begin{cases} W \mid \langle W, v\rangle \geq \gamma & w.p.\ p \\ W \mid \langle W, v\rangle < \gamma & w.p. 1-p \end{cases}$$

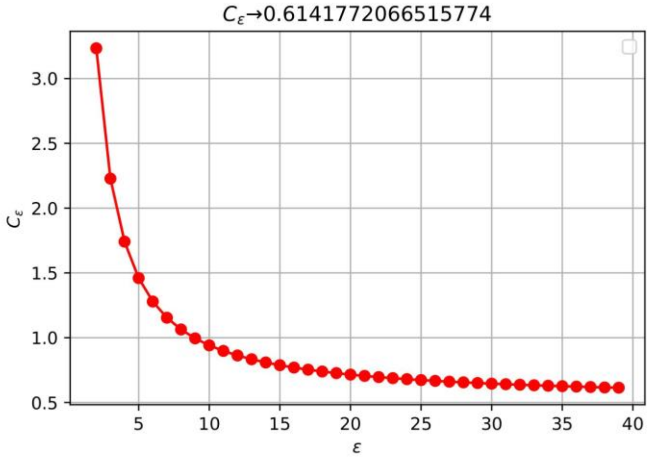
Implications

Optimal hyperparameters are independent of dimension



Error of PrivUnitG is at most $\left(1 + \frac{1}{\sqrt{d}}\right) \cdot$ Error of PrivUnit

Allows to estimate constants of the optimal squared error



$$\mathrm{Err}^{\star}(n, d, \varepsilon) = \mathrm{c}_{d,\varepsilon} \cdot \frac{d}{n\,\varepsilon}$$

$$\lim_{\varepsilon \to \infty} \lim_{d \to \infty} c_{d,\varepsilon} = c^{\star} \approx 0.614$$

Thanks!