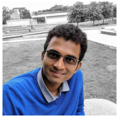
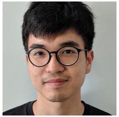
# Faster Privacy Accounting via Evolving Discretization

Badih
Ghazi


Pritish
Kamath

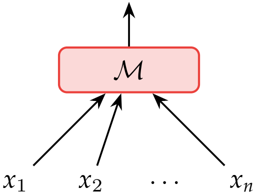
Ravi
Kumar


Pasin
Manurangsi

Google Research
Mountain View

ICML 2022

# Differential Privacy



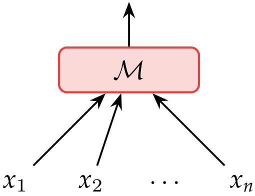
[Dwork et al. '06]

$\mathcal{M}$ satisfies $(\varepsilon, \delta)$-differential privacy if

for all *neighboring* $X$, $X'$, and all outcome events $S$,

$$\Pr[\mathcal{M}(X) \in S] \;\leq\; e^{\varepsilon} \cdot \Pr[\mathcal{M}(X') \in S] \;+\; \delta$$

# Differential Privacy



[Dwork et al. '06]

$\mathcal{M}$ satisfies $(\varepsilon, \delta)$-differential privacy if

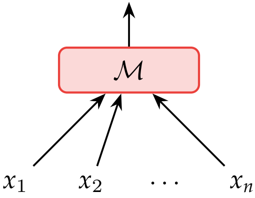
for all *neighboring* $X$, $X'$, and all outcome events $S$,

$$\Pr[\mathcal{M}(X) \in S] \;\leq\; e^{\varepsilon} \cdot \Pr[\mathcal{M}(X') \in S] \;+\; \delta$$

**Example (DP-SGD):** SGD with Gaussian noise added to each mini-batch gradient.

(Self-compositions of subsampled Gaussian mechanism.)

# Differential Privacy



[Dwork et al. '06]

$\mathcal{M}$ satisfies $(\varepsilon, \delta)$-differential privacy if
for all *neighboring* $X, X'$, and all outcome events $S$,

$$\Pr[\mathcal{M}(X) \in S] \;\leq\; e^{\varepsilon} \cdot \Pr[\mathcal{M}(X') \in S] \;+\; \delta$$

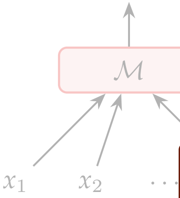
**Example (DP-SGD):** SGD with Gaussian noise added to each mini-batch gradient.

(Self-compositions of subsampled Gaussian mechanism.)

**Privacy Accounting:** Given $\mathcal{M}$ and $\varepsilon$, compute $\delta$ such that $\mathcal{M}$ satisfies $(\varepsilon, \delta)$-DP.

(Useful for computing parameters underlying $\mathcal{M}$, e.g. noise scale in DP-SGD.)

# Differential Privacy

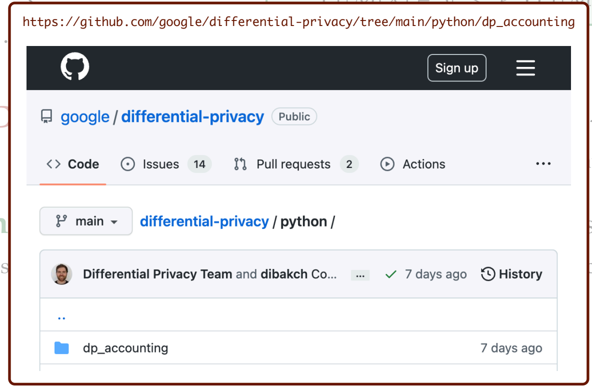


[Dwork et al. '06]

$\mathcal{M}$ satisfies $(\varepsilon, \delta)$-differential privacy if for all *neighboring* $X, X'$, and all outcome events $S$,

$$\Pr[\mathcal{M}(X) \in S] \;\le\; e^{\varepsilon} \cdot \Pr[\mathcal{M}(X') \in S] \;+\; \delta$$

https://github.com/google/differential-privacy/tree/main/python/dp_accounting



google / differential-privacy Public

differential-privacy / python /

Differential Privacy Team and dibakch Co... ... ✓ 7 days ago

..

dp_accounting 7 days ago

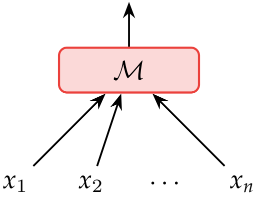
**Example (DP-SGD** -batch gradient.

ussian mechanism.)

**Privacy Accountin** sfies $(\varepsilon, \delta)$-DP.

(Us cale in DP-SGD.)

# Differential Privacy



[Dwork et al. '06]

$\mathcal{M}$ satisfies $(\varepsilon, \delta)$-differential privacy if for all *neighboring* $X$, $X'$, and all outcome events $S$,

$$\Pr[\mathcal{M}(X) \in S] \;\leq\; e^{\varepsilon} \cdot \Pr[\mathcal{M}(X') \in S] \;+\; \delta$$

**Example (DP-SGD):** SGD with Gaussian noise added to each mini-batch gradient.

(Self-compositions of subsampled Gaussian mechanism.)

**Privacy Accounting:** Given $\mathcal{M}$ and $\varepsilon$, compute $\delta$ such that $\mathcal{M}$ satisfies $(\varepsilon, \delta)$-DP.

(Useful for computing parameters underlying $\mathcal{M}$, e.g. noise scale in DP-SGD.)

**Desiderata:**

- $\delta$ value close to optimal value $\delta_{\mathcal{M}}(\varepsilon)$
- Fast computation

# Accounting using Privacy Random Variables (PRVs)

$$\mathrm{PRV}_{(P,Q)} \;:=\; \text{distributed as } \log\frac{P(\omega)}{Q(\omega)} \text{ for } \omega \sim P.$$

# Accounting using Privacy Random Variables (PRVs)

$$\mathrm{PRV}_{(P,Q)} \;:=\; \text{distributed as } \log \frac{P(\omega)}{Q(\omega)} \text{ for } \omega \sim P.$$

**PRV Accounting Approach:**

- Associate $\mathrm{PRV}_{\mathcal{M}}$ to a mechanism $\mathcal{M}$, to derive upper bounds on $\delta_{\mathcal{M}}(\varepsilon)$:

$$\delta_{\mathcal{M}}(\varepsilon) \;\le\; \mathop{\mathbb{E}}_{Y \sim \mathrm{PRV}_{\mathcal{M}}} \max\left\{0, 1 - e^{\varepsilon - Y}\right\}$$

# Accounting using Privacy Random Variables (PRVs)

$$\mathrm{PRV}_{(P,Q)} \;:=\; \text{distributed as } \log \frac{P(\omega)}{Q(\omega)} \text{ for } \omega \sim P.$$

**PRV Accounting Approach:**

- Associate $\mathrm{PRV}_{\mathcal{M}}$ to a mechanism $\mathcal{M}$, to derive upper bounds on $\delta_{\mathcal{M}}(\varepsilon)$:

$$\delta_{\mathcal{M}}(\varepsilon) \;\le\; \mathop{\mathbb{E}}_{Y \sim \mathrm{PRV}_{\mathcal{M}}} \max\left\{0, 1 - e^{\varepsilon - Y}\right\}$$

- Composition of mechanisms corresponds to addition of PRVs : $\mathrm{PRV}_{\mathcal{M} \circ \mathcal{M}'} = \mathrm{PRV}_{\mathcal{M}} + \mathrm{PRV}_{\mathcal{M}'}$.

# Accounting using Privacy Random Variables (PRVs)

$$\mathrm{PRV}_{(P,Q)} \;:=\; \text{distributed as } \log \frac{P(\omega)}{Q(\omega)} \text{ for } \omega \sim P.$$

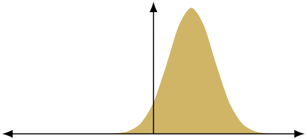
**PRV Accounting Approach:**

- Associate $\mathrm{PRV}_{\mathcal{M}}$ to a mechanism $\mathcal{M}$, to derive upper bounds on $\delta_{\mathcal{M}}(\varepsilon)$:

$$\delta_{\mathcal{M}}(\varepsilon) \;\le\; \mathop{\mathbb{E}}_{Y \sim \mathrm{PRV}_{\mathcal{M}}} \max\left\{0, 1 - e^{\varepsilon - Y}\right\}$$

- Composition of mechanisms corresponds to addition of PRVs : $\mathrm{PRV}_{\mathcal{M} \circ \mathcal{M}'} = \mathrm{PRV}_{\mathcal{M}} + \mathrm{PRV}_{\mathcal{M}'}$.

**Example:** Gaussian mechanism $\mathcal{M}(X) := \sum_i x_i + \zeta$ for $\zeta \sim \mathcal{N}(0, \sigma^2)$ (with $|x_i| \le 1$)



PRV for $\mathcal{M}$

# Accounting using Privacy Random Variables (PRVs)

$$\mathrm{PRV}_{(P,Q)} \;:=\; \text{distributed as } \log \frac{P(\omega)}{Q(\omega)} \text{ for } \omega \sim P.$$

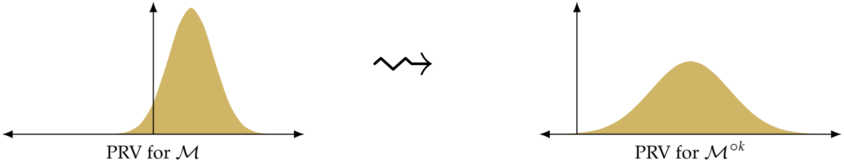
## PRV Accounting Approach:

- Associate $\mathrm{PRV}_{\mathcal{M}}$ to a mechanism $\mathcal{M}$, to derive upper bounds on $\delta_{\mathcal{M}}(\varepsilon)$:

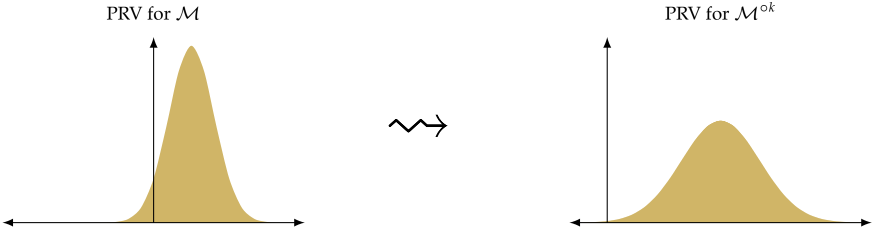
$$\delta_{\mathcal{M}}(\varepsilon) \;\le\; \mathop{\mathbb{E}}_{Y \sim \mathrm{PRV}_{\mathcal{M}}} \max\left\{0, 1 - e^{\varepsilon - Y}\right\}$$

- Composition of mechanisms corresponds to addition of PRVs : $\mathrm{PRV}_{\mathcal{M} \circ \mathcal{M}'} = \mathrm{PRV}_{\mathcal{M}} + \mathrm{PRV}_{\mathcal{M}'}$.

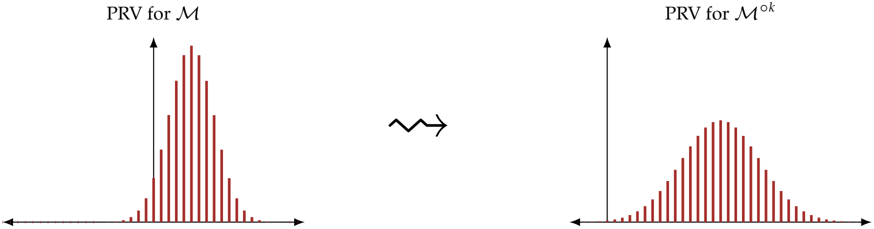
**Example:** Gaussian mechanism $\mathcal{M}(X) := \sum_i x_i + \zeta$ for $\zeta \sim \mathcal{N}(0, \sigma^2)$ (with $|x_i| \le 1$)



# Discretization of PRVs



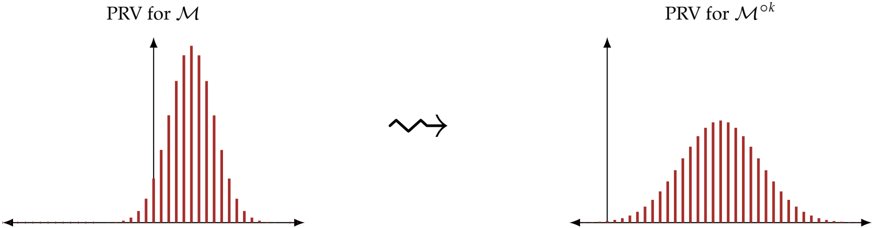
# Discretization of PRVs



To make PRV approach practical:

- Discretize into buckets [Meiser-Mohammadi ’18]

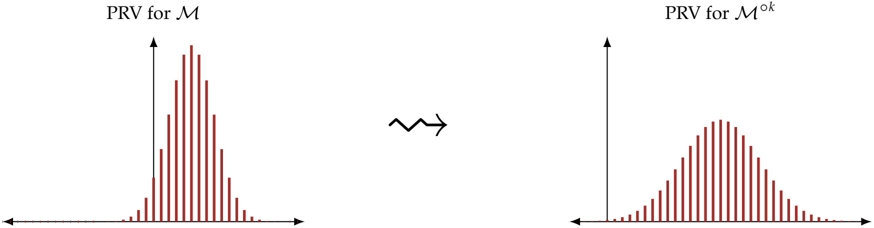
# Discretization of PRVs



To make PRV approach practical:

- Discretize into buckets [Meiser-Mohammadi ’18]
- Nearly linear-time composition using Fast Fourier Transform [Koskela et al ’20]

# Discretization of PRVs

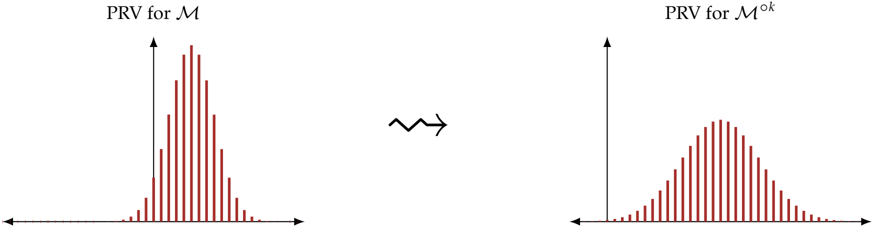


PRV for $\mathcal{M}$

PRV for $\mathcal{M}^{\circ k}$

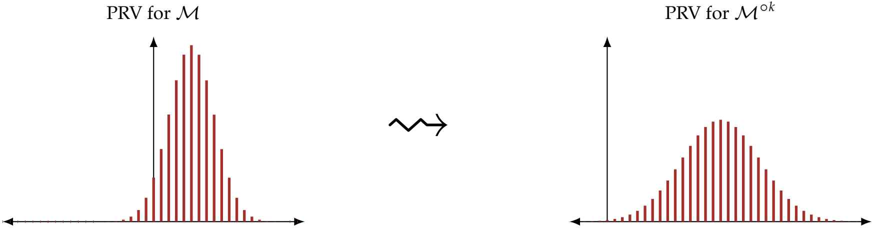
| **Reference** | **Running time for *k*-fold compositions** | |
|---|---|---|
| | Homogeneous | Heterogeneous |
| [Koskela et al. ’21] | $\widetilde{O}(k^{1.5})$ | $\widetilde{O}(k^{2.5})$ |

# Discretization of PRVs



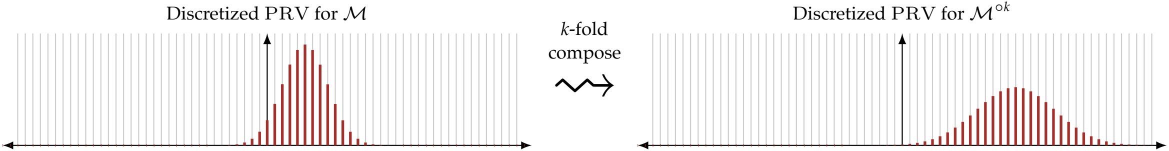
| **Reference** | **Running time for *k*-fold compositions** | |
|---|---|---|
| | Homogeneous | Heterogeneous |
| [Koskela et al. ’21] | $\widetilde{O}(k^{1.5})$ | $\widetilde{O}(k^{2.5})$ |
| [Gopi et al. ’21] | $\widetilde{O}(k^{0.5})$ | $\widetilde{O}(k^{1.5})$ |

# Discretization of PRVs



| **Reference** | **Running time for *k*-fold compositions** | |
|---|---|---|
| | Homogeneous | Heterogeneous |
| [Koskela et al. '21] | $\widetilde{O}(k^{1.5})$ | $\widetilde{O}(k^{2.5})$ |
| [Gopi et al. '21] | $\widetilde{O}(k^{0.5})$ | $\widetilde{O}(k^{1.5})$ |
| This work | $\log^{O(1)}(k)$ | $\widetilde{O}(k)$ |

# Evolving Discretization Approach

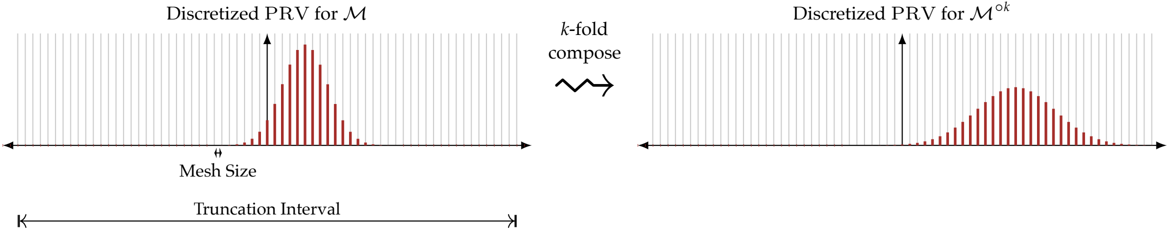


Discretized PRV for $\mathcal{M}$

$k$-fold compose

Discretized PRV for $\mathcal{M}^{\circ k}$

# Evolving Discretization Approach



Discretized PRV for $\mathcal{M}$

$k$-fold compose

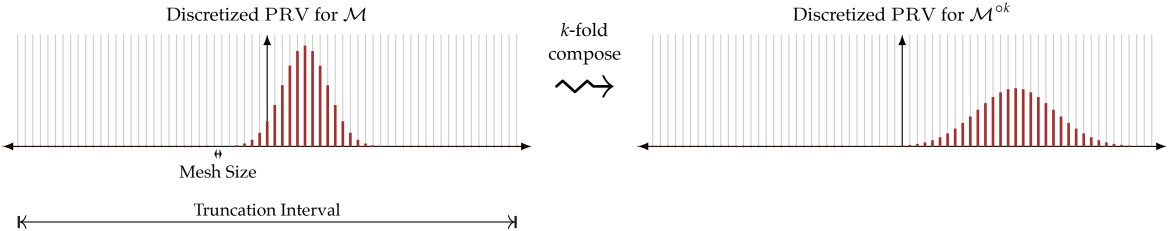
Discretized PRV for $\mathcal{M}^{\circ k}$

Mesh Size

Truncation Interval

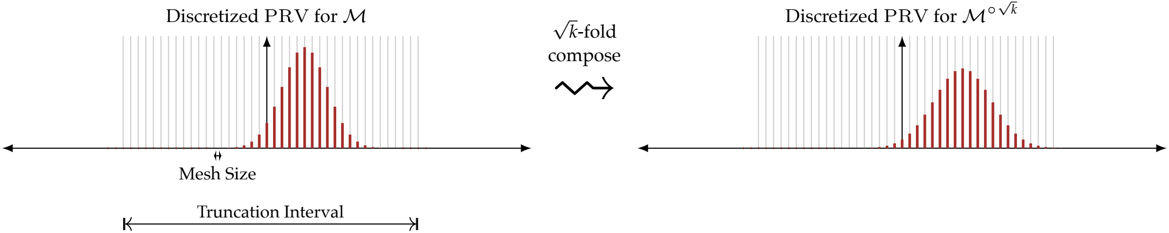
| Reference | Truncation Interval | Mesh Size | Number of Buckets |
|---|---|---|---|
| [Koskela et al. '21] | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |

# Evolving Discretization Approach


Discretized PRV for $\mathcal{M}$
$k$-fold compose
Discretized PRV for $\mathcal{M}^{\circ k}$
Mesh Size
Truncation Interval

| Reference | Truncation Interval | Mesh Size | Number of Buckets |
|---|---|---|---|
| [Koskela et al. ’21] | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |
| [Gopi et al. ’21] | $\approx O(1)$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.5}$ |

# Evolving Discretization Approach



Discretized PRV for $\mathcal{M}$

Mesh Size

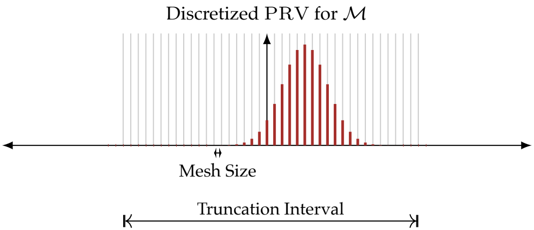
Truncation Interval

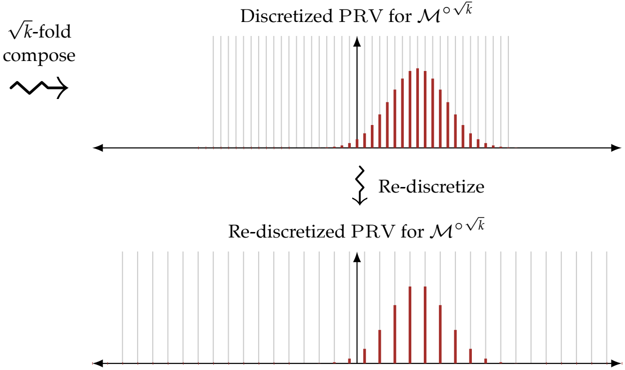
$\sqrt{k}$-fold compose

Discretized PRV for $\mathcal{M}^{\circ \sqrt{k}}$

| Reference | | Truncation Interval | Mesh Size | Number of Buckets |
|---|---|---|---|---|
| [Koskela et al. '21] | | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |
| [Gopi et al. '21] | | $\approx O(1)$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.5}$ |
| This work | Stage 1 | $\approx \frac{1}{k^{0.25}}$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.25}$ |
| | | | | |

# Evolving Discretization Approach

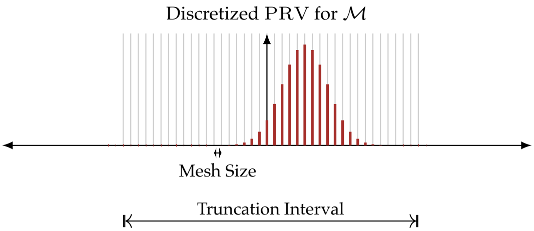




$\sqrt{k}$-fold compose
Discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$
Re-discretize
Re-discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$

| Reference | | Truncation Interval | Mesh Size | Number of Buckets |
|---|---|---|---|---|
| [Koskela et al. '21] | | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |
| [Gopi et al. '21] | | $\approx O(1)$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.5}$ |
| This work | Stage 1 | $\approx \frac{1}{k^{0.25}}$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.25}$ |
| | Stage 2 | $\approx O(1)$ | $\approx \frac{1}{k^{0.25}}$ | $\approx k^{0.25}$ |

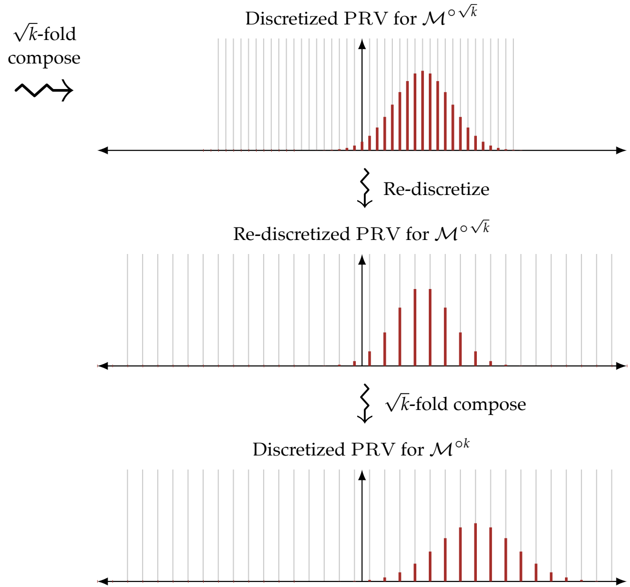
# Evolving Discretization Approach



Discretized PRV for $\mathcal{M}$

Mesh Size

Truncation Interval



$\sqrt{k}$-fold compose

Discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$

Re-discretize

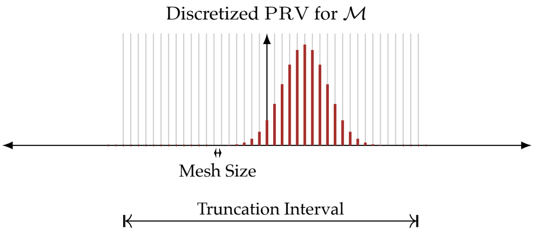
Re-discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$

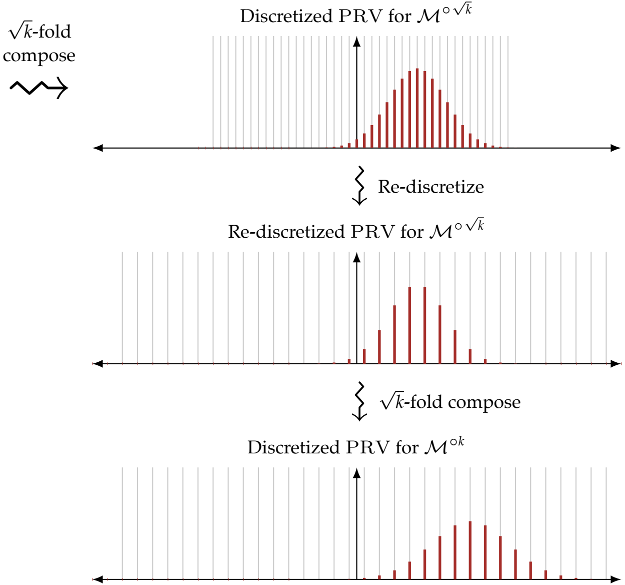
$\sqrt{k}$-fold compose

Discretized PRV for $\mathcal{M}^{\circ k}$

| **Reference** | | **Truncation Interval** | **Mesh Size** | **Number of Buckets** |
|---|---|---|---|---|
| [Koskela et al. '21] | | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |
| [Gopi et al. '21] | | $\approx O(1)$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.5}$ |
| This work | Stage 1 | $\approx \frac{1}{k^{0.25}}$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.25}$ |
| | Stage 2 | $\approx O(1)$ | $\approx \frac{1}{k^{0.25}}$ | $\approx k^{0.25}$ |

# Evolving Discretization Approach




$\sqrt{k}$-fold compose
Discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$
Re-discretize
Re-discretized PRV for $\mathcal{M}^{\circ\sqrt{k}}$
$\sqrt{k}$-fold compose
Discretized PRV for $\mathcal{M}^{\circ k}$

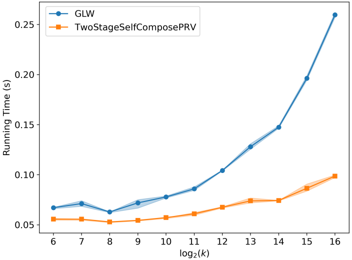
| **Reference** | | **Truncation Interval** | **Mesh Size** | **Number of Buckets** |
|---|---|---|---|---|
| [Koskela et al. ’21] | | $\approx k^{0.5}$ | $\approx \frac{1}{k}$ | $\approx k^{1.5}$ |
| [Gopi et al. ’21] | | $\approx O(1)$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.5}$ |
| This work | Stage 1 | $\approx \frac{1}{k^{0.25}}$ | $\approx \frac{1}{k^{0.5}}$ | $\approx k^{0.25}$ |
| | Stage 2 | $\approx O(1)$ | $\approx \frac{1}{k^{0.25}}$ | $\approx k^{0.25}$ |

Recursively with $O(\log k)$ stages, the running time is $\log^{O(1)}(k)$.

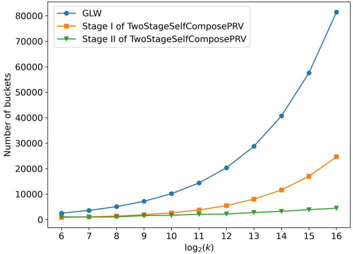
# Evaluation of Two-Stage Algorithm

**DP-SGD application :** Compositions of Poisson subsampled Gaussian mechanism

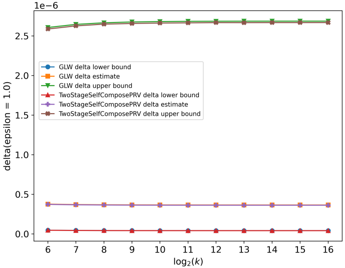
Comparison against [Gopi et al. '21]



Running Times



Number of Buckets



Delta Estimates

# Summary & Future Directions

**Summary:**

- Evolving discretization can speed up privacy accounting with privacy random variables.

# Summary & Future Directions

**Summary:**

- Evolving discretization can speed up privacy accounting with privacy random variables.

**Future Directions:**

- Make the recursive algorithm practical, by tightening parameters.
- Make heterogeneous composition more practical in the case where mechanisms have very different privacy profiles

# Summary & Future Directions

**Summary:**

- Evolving discretization can speed up privacy accounting with privacy random variables.

**Future Directions:**

- Make the recursive algorithm practical, by tightening parameters.
- Make heterogeneous composition more practical in the case where mechanisms have very different privacy profiles

## Thanks!