# Quantum Sub-Gaussian Mean Estimator

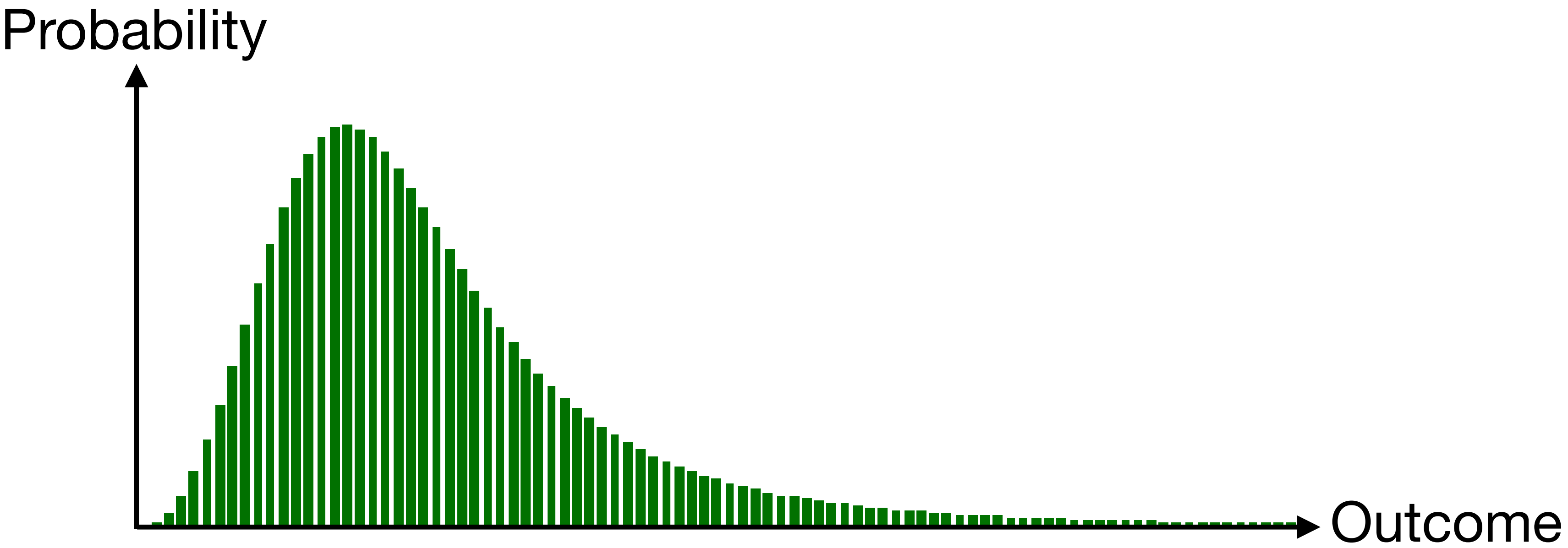
Yassine Hamoudi

ESA 2021

# Mean Estimation problem (over $\mathbb{R}$)



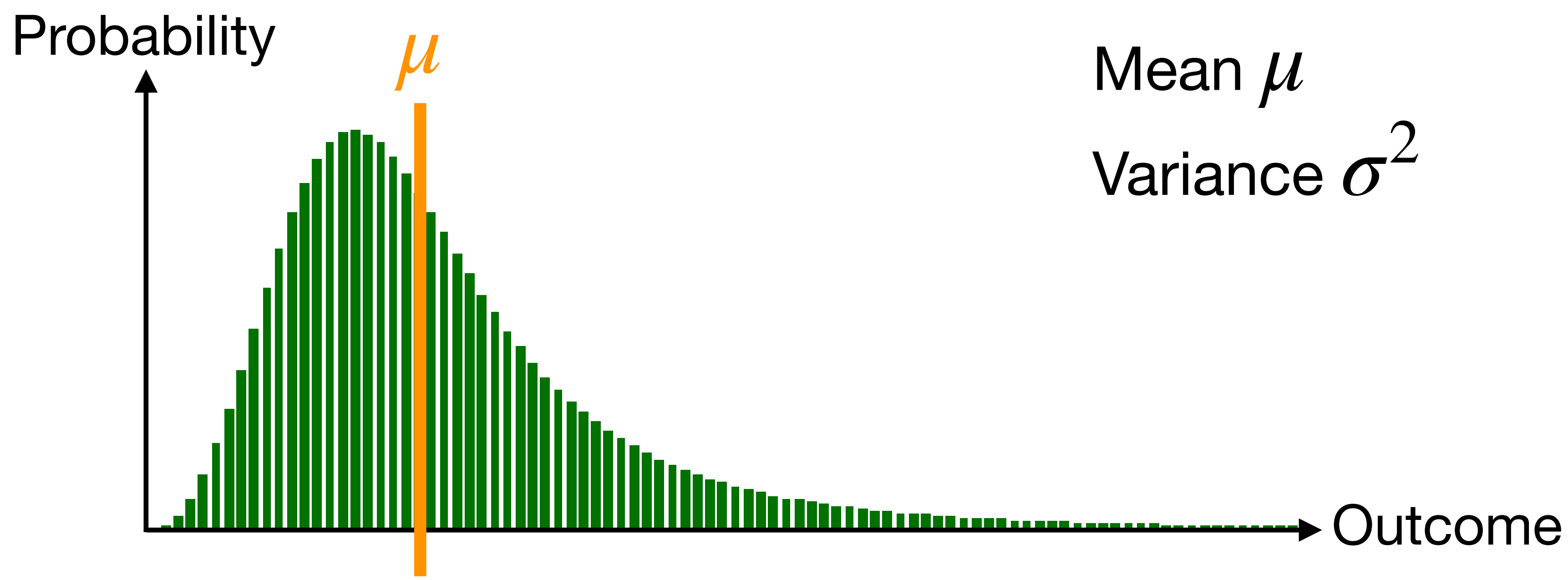
Experiment with unknown outcome distribution $D$



# Mean Estimation problem (over $\mathbb{R}$)



Experiment with unknown outcome distribution $D$



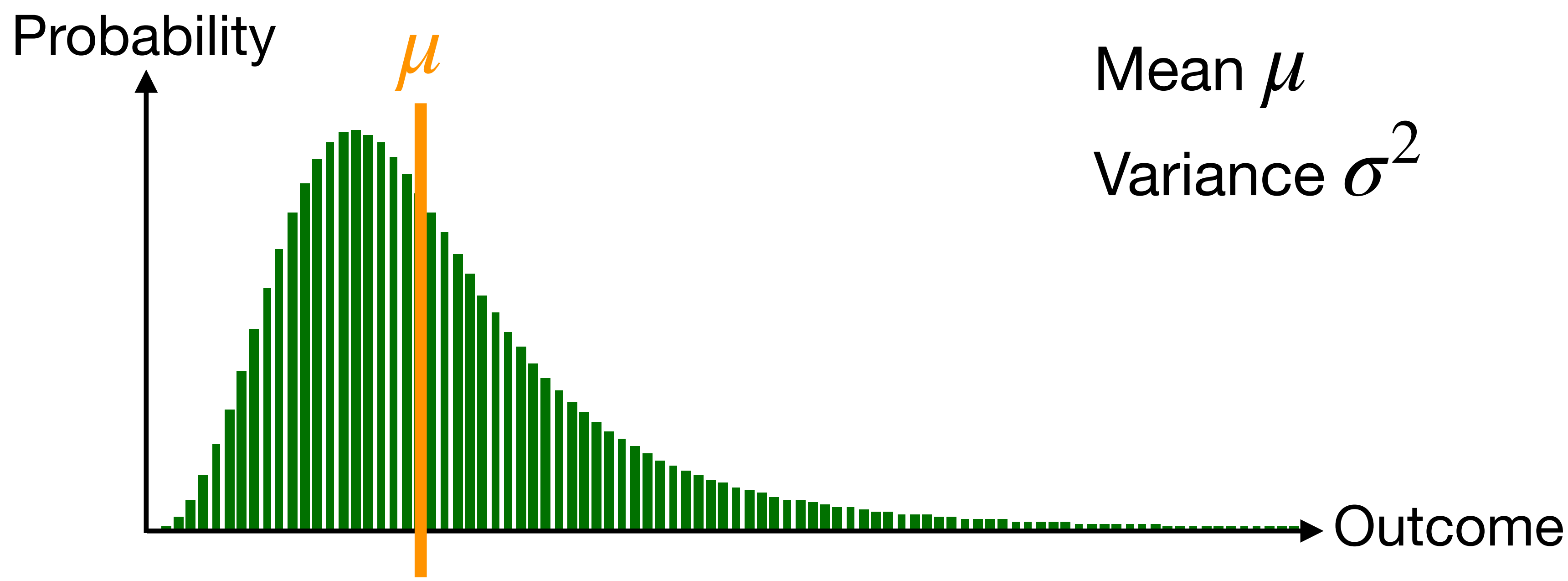

Mean $\mu$

Variance $\sigma^2$

# Mean Estimation problem (over $\mathbb{R}$)



Experiment with unknown outcome distribution $D$



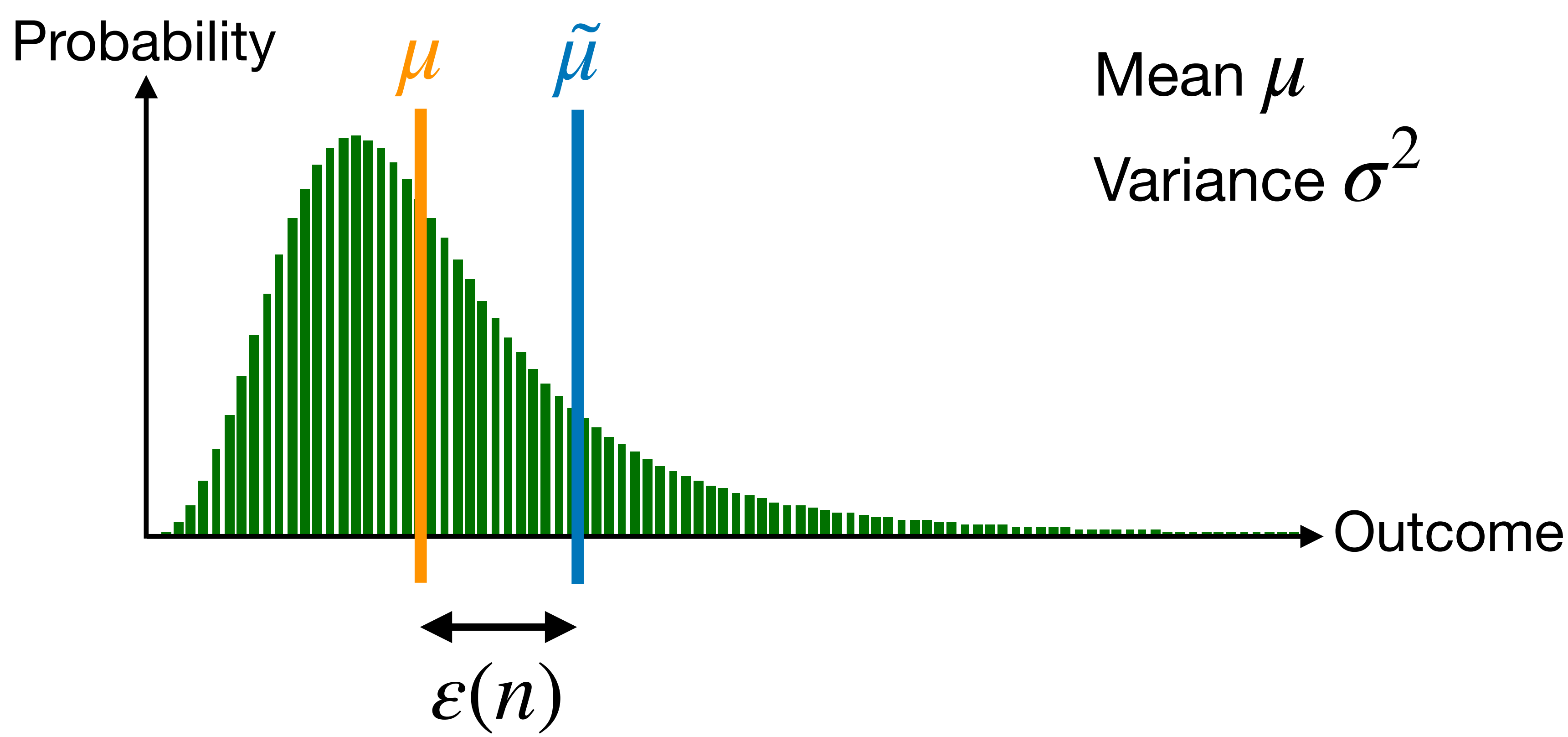

Mean $\mu$

Variance $\sigma^2$

$n$ = # of runs of the experiment $\begin{cases} \text{classical} = n \text{ i.i.d. samples from } D \\ \text{quantum} = \text{see later} \end{cases}$

# Mean Estimation problem (over $\mathbb{R}$)

Experiment with unknown outcome distribution $D$




Mean $\mu$

Variance $\sigma^2$

$n$ = # of runs of the experiment $\begin{cases} \text{classical} = n \text{ i.i.d. samples from } D \\ \text{quantum} = \text{see later} \end{cases}$

**<u>Goal:</u>** compute an estimate $\tilde{\mu}$ that minimizes the error $\varepsilon(n)$ such that

$$\Pr\left[\,|\mu - \tilde{\mu}| > \varepsilon(n)\right] < \delta \qquad \text{given } \delta \in (0,1)$$

# Optimal classical estimators



$$\Pr\Big[\, |\, \mu - \tilde{\mu}\, | > \varepsilon(n) \Big] < \delta$$

# Optimal classical estimators

$$\Pr\left[\,|\mu - \tilde{\mu}| > \varepsilon(n)\right] < \delta$$

→ If $D \in \left\{\mathcal{N}(\mu, \sigma^2)\right\}_{\mu,\sigma}$ is Gaussian then the Empirical Mean is optimal and

$$\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$$

# Optimal classical estimators

$$\Pr\left[\,|\mu - \tilde{\mu}\,| > \varepsilon(n)\right] < \delta$$

→ If $D \in \left\{\mathcal{N}(\mu, \sigma^2)\right\}_{\mu,\sigma}$ is Gaussian then the Empirical Mean is optimal and

$$\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$$

→ An estimator is **sub-Gaussian** if it achieves the above bound for **any** $D$

# Optimal classical estimators

$$\Pr\left[\,|\mu - \tilde{\mu}\,| > \varepsilon(n)\right] < \delta$$

→ If $D \in \left\{\mathcal{N}(\mu, \sigma^2)\right\}_{\mu,\sigma}$ is Gaussian then the Empirical Mean is optimal and

$$\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$$

→ An estimator is **sub-Gaussian** if it achieves the above bound for **any** $D$

❌ Empirical Mean is not sub-Gaussian

# Optimal classical estimators



$$\Pr\left[\,|\mu - \tilde{\mu}\,| > \varepsilon(n)\right] < \delta$$

→ If $D \in \left\{\mathcal{N}(\mu, \sigma^2)\right\}_{\mu,\sigma}$ is Gaussian then the Empirical Mean is optimal and

$$\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$$

→ An estimator is **sub-Gaussian** if it achieves the above bound for **any** $D$

❌ Empirical Mean is not sub-Gaussian

✔ Median-of-Means, [Catoni'12], [Lee,Valiant'21], … are sub-Gaussian

# Optimal classical estimators



$$\Pr\left[\,|\mu - \tilde{\mu}| > \varepsilon(n)\right] < \delta$$

→ If $D \in \left\{\mathcal{N}(\mu,\sigma^2)\right\}_{\mu,\sigma}$ is Gaussian then the Empirical Mean is optimal and

$$\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$$

→ An estimator is **sub-Gaussian** if it achieves the above bound for **any** $D$

❌ Empirical Mean is not sub-Gaussian

✔ Median-of-Means, [Catoni'12], [Lee,Valiant'21], … are sub-Gaussian

optimal $\sqrt{2} + o(1)$ factor in $O(\,.\,)$

# Quantum estimators



**Classical:** $\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$

**Quantum:** $\varepsilon(n) \approx \ldots$

# Quantum estimators



**Classical:** $\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$

**Quantum:** $\varepsilon(n) \approx \ldots$

$$\ldots \frac{B \log(1/\delta)}{n}$$

given $B \geq \max\{\,|x| : p_x \neq 0\}$

[Grover'98] [Abrams,Williams'99]
[Brassard,Dupuis,Gambs,Tapp'11] …

# Quantum estimators



**Classical:** $\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$

**Quantum:** $\varepsilon(n) \approx \ldots$

$$\ldots \frac{B \log(1/\delta)}{n}$$

given $B \geq \max\{\,|x| : p_x \neq 0\}$

[Grover'98] [Abrams,Williams'99]
[Brassard,Dupuis,Gambs,Tapp'11] …

$$\ldots \frac{\sqrt{B\mu} \log(1/\delta)}{n} + \frac{B \log(1/\delta)^2}{n^2}$$

[Brassard,Høyer,Mosca,Tapp'02]
**Amplitude Estimation**

# Quantum estimators

**Classical:** $\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$

**Quantum:** $\varepsilon(n) \approx \ldots$

$\ldots \frac{B \log(1/\delta)}{n}$

given $B \geq \max\{\,|x| : p_x \neq 0\}$

[Grover'98] [Abrams,Williams'99]
[Brassard,Dupuis,Gambs,Tapp'11] …

$\ldots \frac{\sqrt{B\mu} \log(1/\delta)}{n} + \frac{B \log(1/\delta)^2}{n^2}$

[Brassard,Høyer,Mosca,Tapp'02]
**Amplitude Estimation**

$\ldots \frac{\Sigma \log(1/\delta)}{n}$

given $\Sigma \geq \sigma$

[Heinrich'02][Montanaro'15]

# Quantum estimators



**Classical:** $\varepsilon(n) = O\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right)$

**Quantum:** $\varepsilon(n) \approx \ldots$

$\ldots \frac{B \log(1/\delta)}{n}$

given $B \geq \max\{\,|x| : p_x \neq 0\}$

[Grover'98] [Abrams,Williams'99]
[Brassard,Dupuis,Gambs,Tapp'11] …

$\ldots \frac{\sqrt{B\mu}\,\log(1/\delta)}{n} + \frac{B \log(1/\delta)^2}{n^2}$

[Brassard,Høyer,Mosca,Tapp'02]
**Amplitude Estimation**

$\ldots \frac{\Sigma \log(1/\delta)}{n}$

given $\Sigma \geq \sigma$

[Heinrich'02][Montanaro'15]

$\ldots \frac{\mu\Delta \log(1/\delta)}{n}$

given $\Delta \geq \sigma/\mu$

[H.,Magniez'19]

# Our result



Quantum "sub-Gaussian" estimator:

$$\varepsilon(n) = \tilde{O}\left( \frac{\sigma \log(1/\delta)}{n} \right)$$

# Our result



Quantum "sub-Gaussian" estimator:

$$\varepsilon(n) = \tilde{O}\left(\frac{\sigma \log(1/\delta)}{n}\right)$$

✓ Doesn't require any prior knowledge about $D$

✓ Subsumes all previous quantum estimators

✓ Better than any classical sub-Gaussian estimator (recall: $\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}$)

✓ Optimal (lower bound by reduction from Quantum Search)

**up to log factors*

# Quantum input model



Unknown distribution $D = (p_x)_x$

# Quantum input model



Unknown distribution $D = (p_x)_x$

<u>Model 1:</u> $n$ copies of the qsample $\sum_x \sqrt{p_x}\,|x\rangle$

# Quantum input model



Unknown distribution $D = (p_x)_x$

<u>Model 1:</u> $n$ copies of the qsample $\sum_x \sqrt{p_x} \, | x \rangle$

→ **classical** sub-Gaussian estimators are optimal in this model!

# Quantum input model



Unknown distribution $D = (p_x)_x$

<u>Model 1:</u> $n$ copies of the qsample $\sum_x \sqrt{p_x}\,|x\rangle$

→ **classical** sub-Gaussian estimators are optimal in this model!

<u>Model 2:</u> $n$ accesses to a **unitary** $U_D$ such that $U_D|0\rangle = \sum_x \sqrt{p_x}\,|x\rangle\,|\mathrm{garbage}_x\rangle$

# Quantum input model 

Unknown distribution $D = (p_x)_x$

Model 1: $n$ copies of the qsample $\sum_x \sqrt{p_x} \, |x\rangle$

→ **classical** sub-Gaussian estimators are optimal in this model!

Model 2: $n$ accesses to a **unitary** $U_D$ such that $U_D|0\rangle = \sum_x \sqrt{p_x} \, |x\rangle |\mathrm{garbage}_x\rangle$

→ $U_D, U_D^{-1}$ and their controlled versions can be accessed

# Quantum input model



Unknown distribution $D = (p_x)_x$

Model 1: $n$ copies of the qsample $\sum_x \sqrt{p_x} |x\rangle$

→ **classical** sub-Gaussian estimators are optimal in this model!

Model 2: $n$ accesses to a **unitary** $U_D$ such that $U_D|0\rangle = \sum_x \sqrt{p_x} |x\rangle |\mathrm{garbage}_x\rangle$

→ $U_D, U_D^{-1}$ and their controlled versions can be accessed

→ Grover Search, Phase/Amplitude Estimation… can be used on $U_D$

# Quantum input model



Unknown distribution $D = (p_x)_x$

Model 1: $n$ copies of the qsample $\sum_x \sqrt{p_x} \, |x\rangle$

→ **classical** sub-Gaussian estimators are optimal in this model!

Model 2: $n$ accesses to a **unitary** $U_D$ such that $U_D \, |0\rangle = \sum_x \sqrt{p_x} \, |x\rangle \, |\mathrm{garbage}_x\rangle$

→ $U_D$, $U_D^{-1}$ and their controlled versions can be accessed

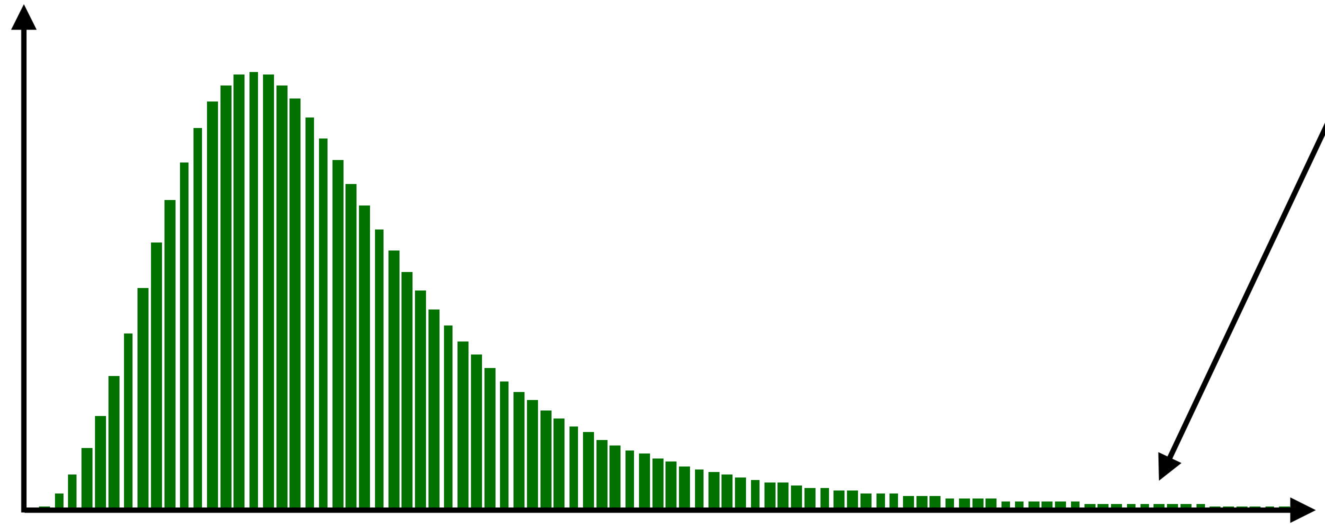
→ Grover Search, Phase/Amplitude Estimation… can be used on $U_D$

→ Similar models in other works on estimating/testing statistics [Nayak,Wu'99] [Bravyi,Harrow,Hassidim'11][Montanaro'15]…

# Technique: Outliers and truncation

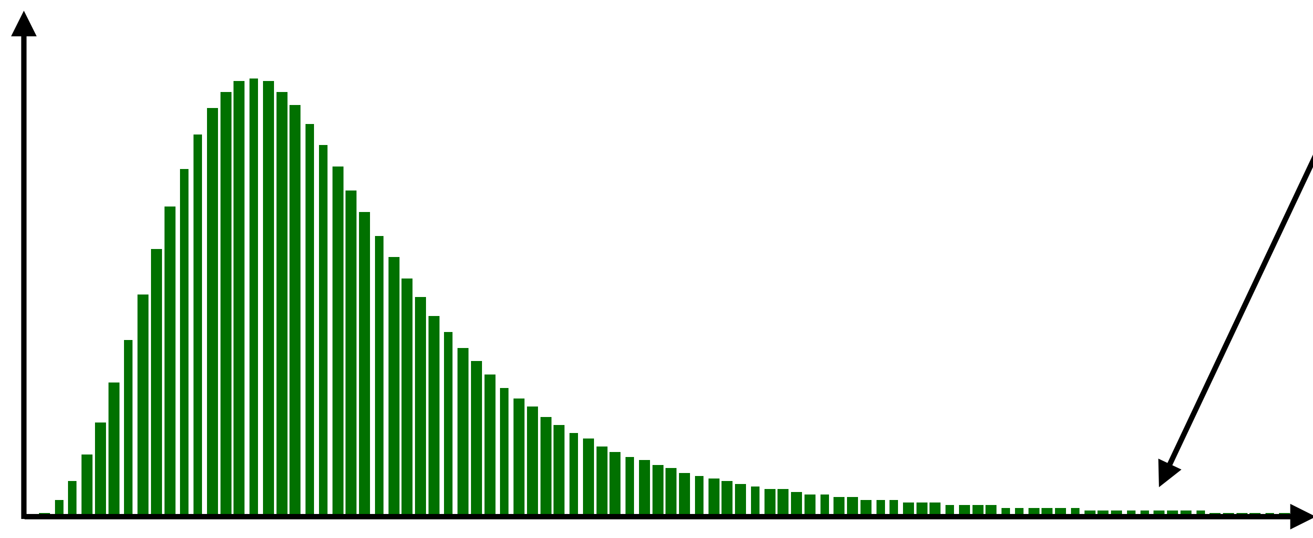


Empirical Mean and Amplitude Estimation are too sensitive to **outliers**



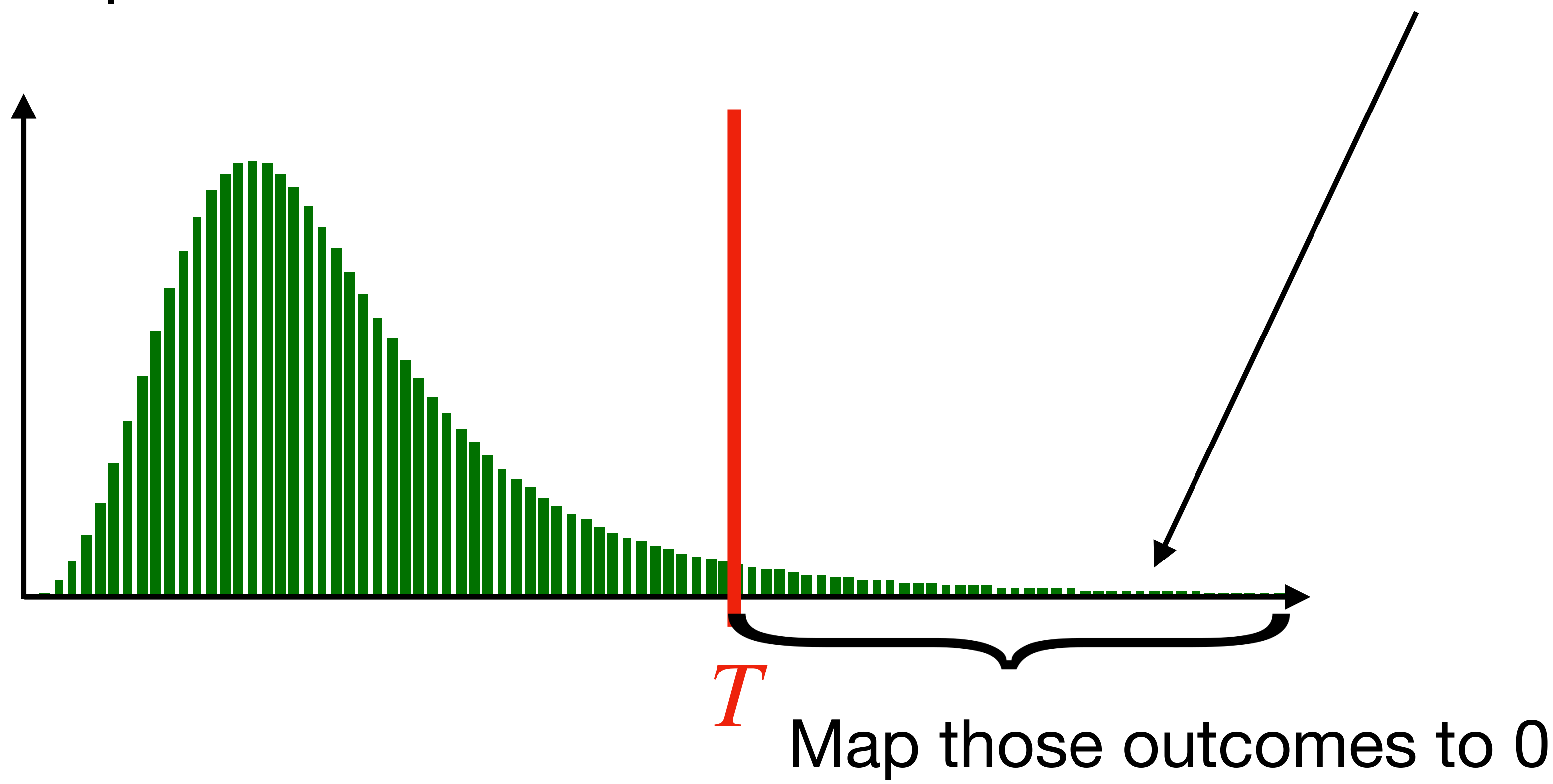
# Technique: Outliers and truncation

Empirical Mean and Amplitude Estimation are too sensitive to **outliers**



Natural approach: **truncate** the distribution at a certain level $T$

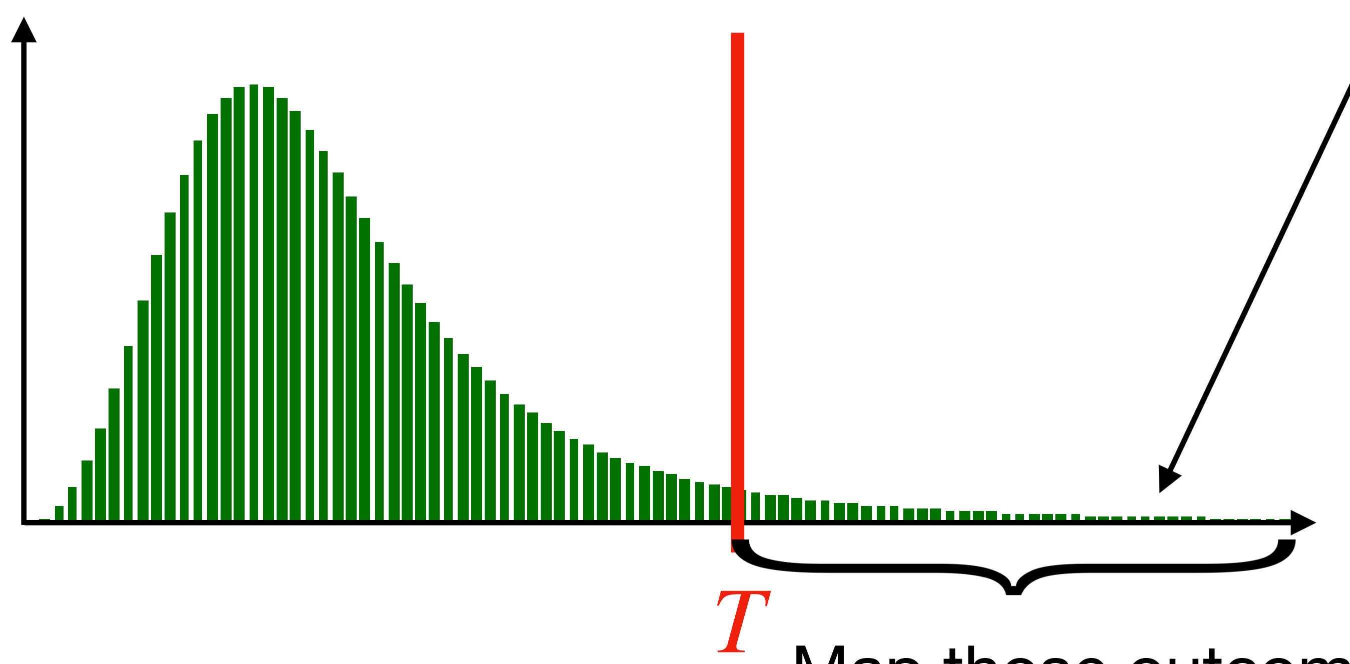
Empirical Mean and Amplitude Estimation are too sensitive to **outliers**




Natural approach: **truncate** the distribution at a certain level $T$

# Technique: Outliers and truncation

Empirical Mean and Amplitude Estimation are too sensitive to **outliers**

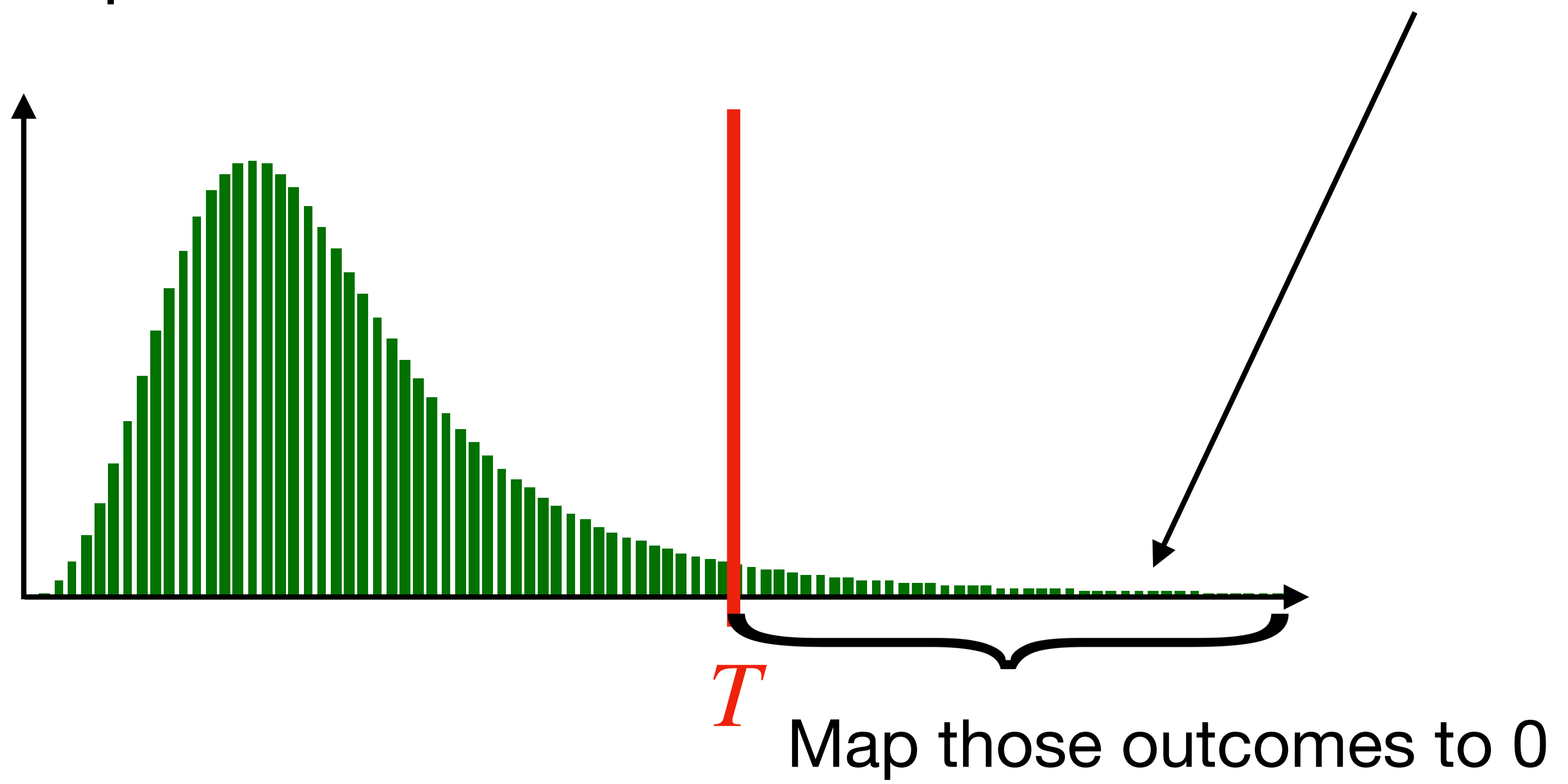



Natural approach: **truncate** the distribution at a certain level $T$

→ How to choose $T$ ?

# Technique: Outliers and truncation

Empirical Mean and Amplitude Estimation are too sensitive to **outliers**




Natural approach: **truncate** the distribution at a certain level $T$

→ How to choose $T$ ?

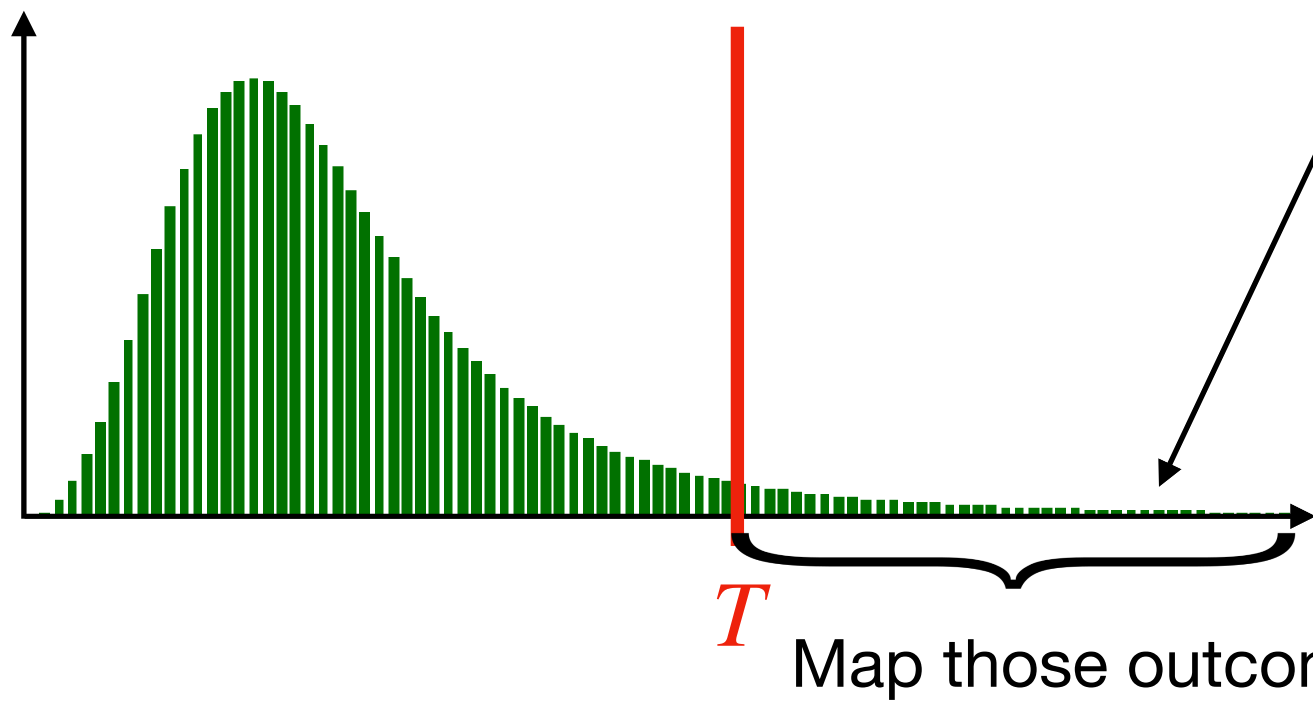
$$T \approx n\sigma$$

[Heinrich'02]
[Montanaro'15]
[H.,Magniez'19]

Require some knowledge about $\sigma$

# Technique: Outliers and truncation

Empirical Mean and Amplitude Estimation are too sensitive to **outliers**




Natural approach: **truncate** the distribution at a certain level $T$

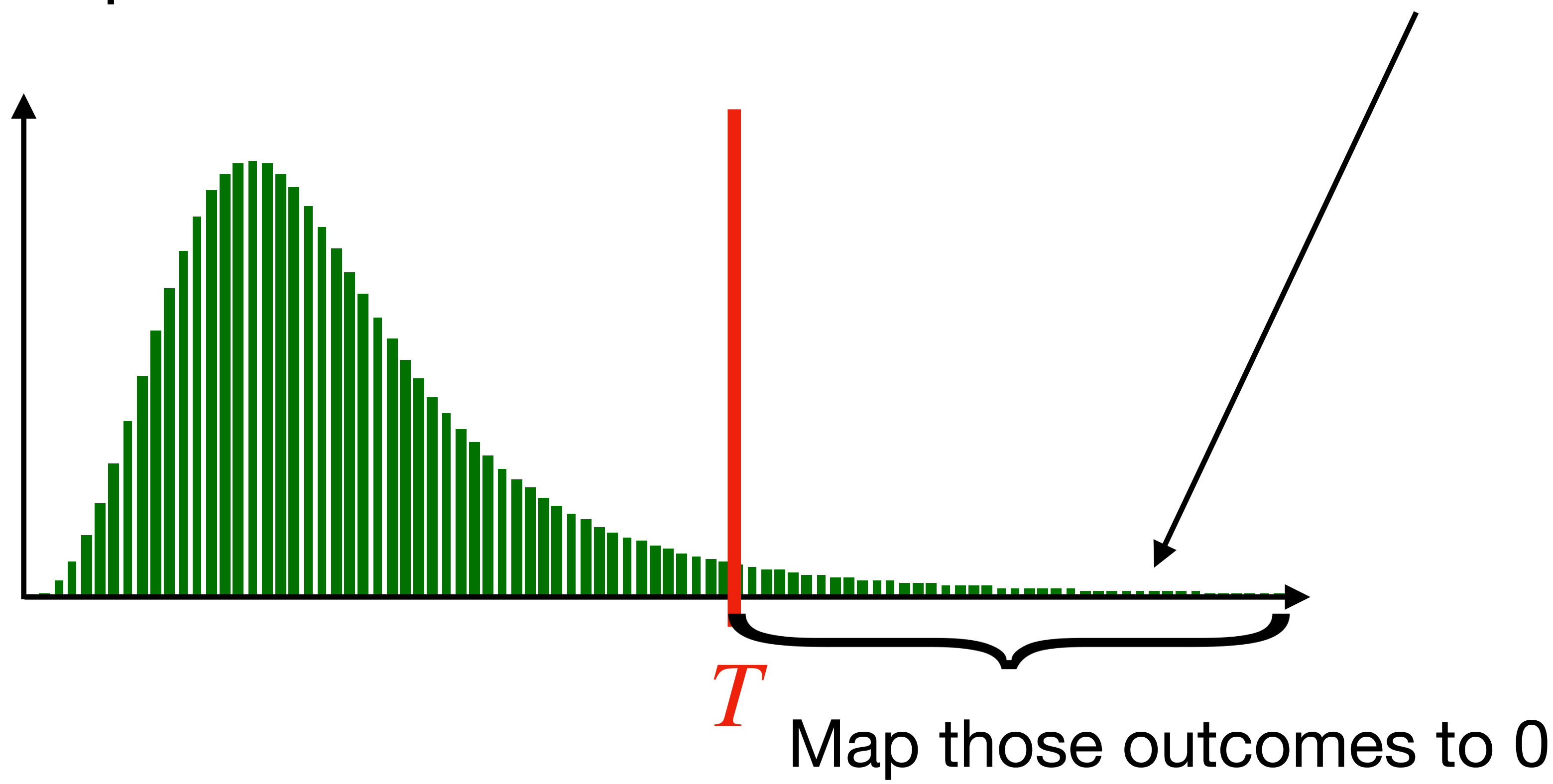
→ How to choose $T$ ?

$T \approx n\sigma$ [Heinrich'02] [Montanaro'15] [H.,Magniez'19] Require some knowledge about $\sigma$

Our work: $T \approx$ quantile satisfying $\Pr_{x \leftarrow D}[x > T] = 1/n^2$

# Technique: Outliers and truncation

Empirical Mean and Amplitude Estimation are too sensitive to **outliers**



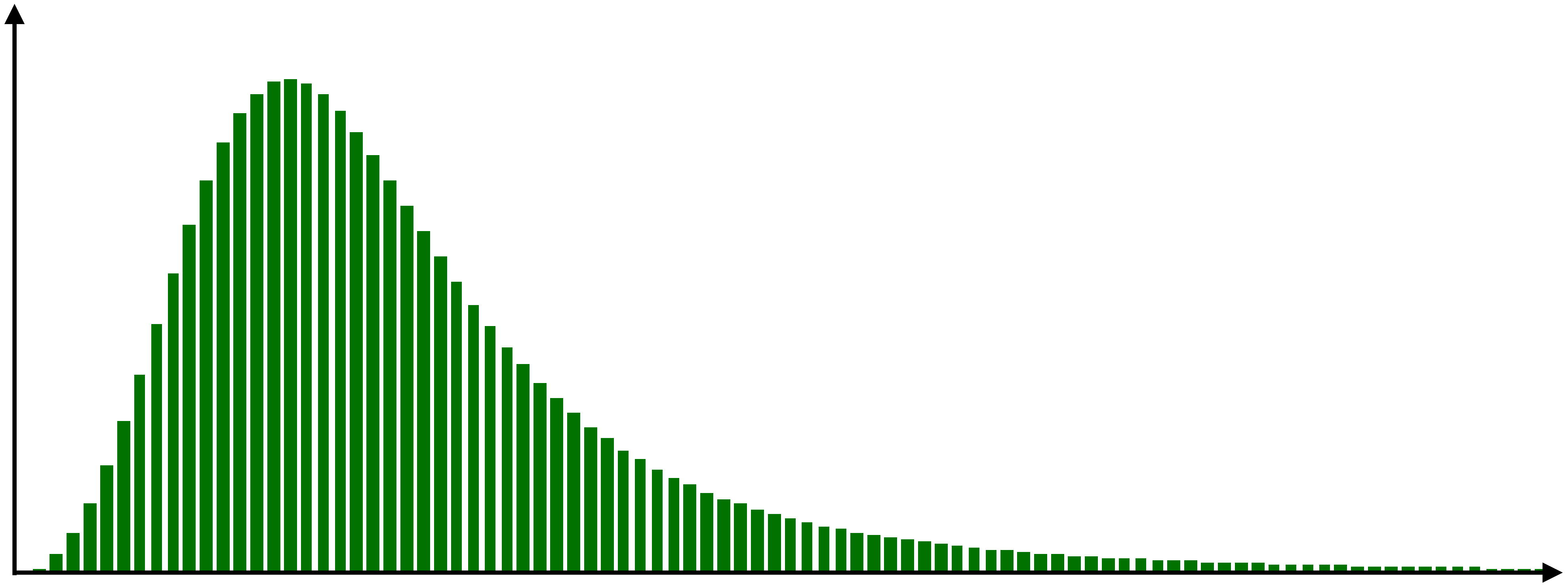

Natural approach: **truncate** the distribution at a certain level $T$

→ How to choose $T$ ?

$T \approx n\sigma$ [Heinrich'02] [Montanaro'15] [H.,Magniez'19] Require some knowledge about $\sigma$

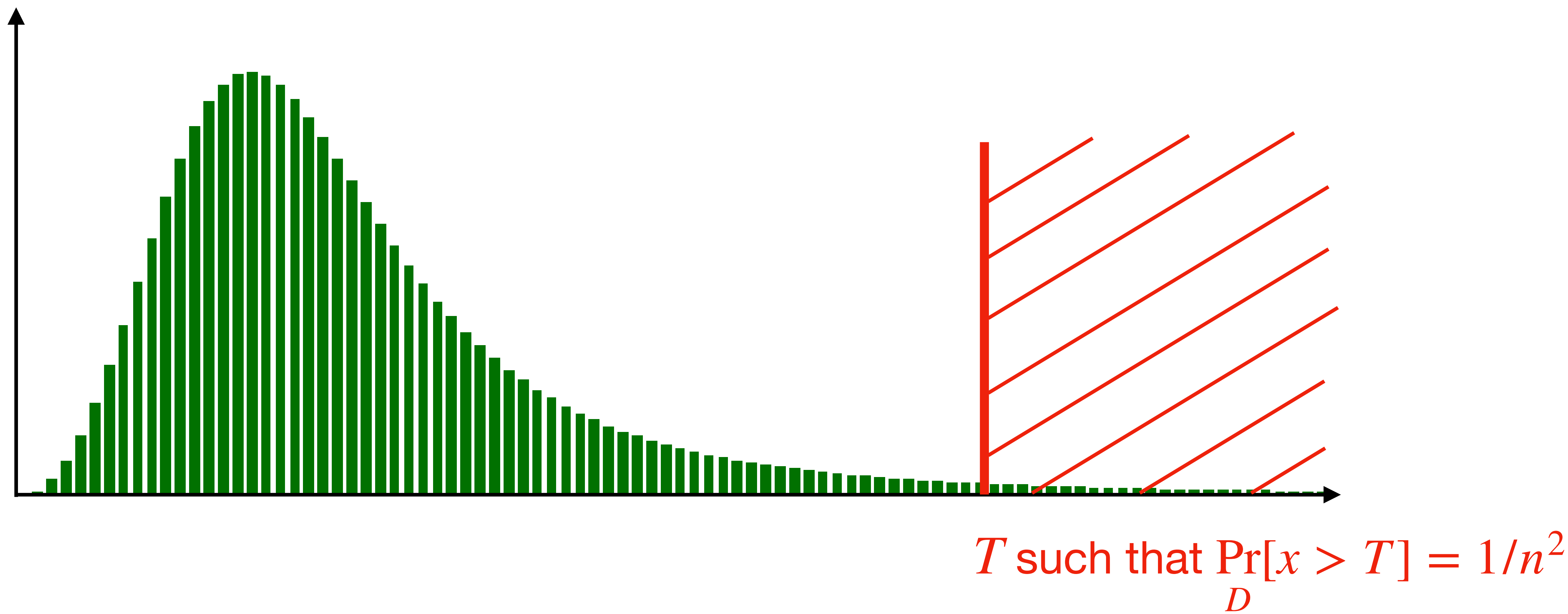
Our work: $T \approx$ quantile satisfying $\Pr_{x \leftarrow D}[x > T] = 1/n^2$ Computed in $O(n)$ time by adapting Minimum Finding [Dürr,Høyer'96]

**<u>Assumptions:</u>** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$



# Technique: Main algorithm

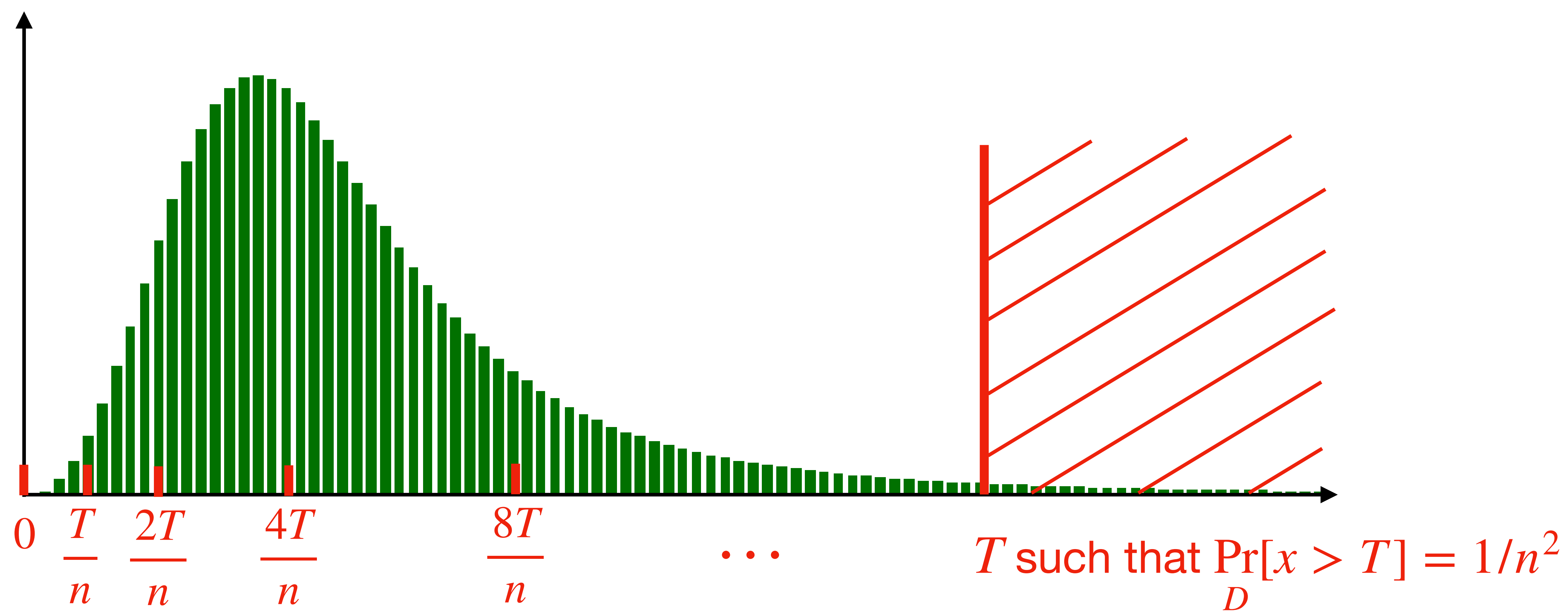
**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$



# Technique: Main algorithm



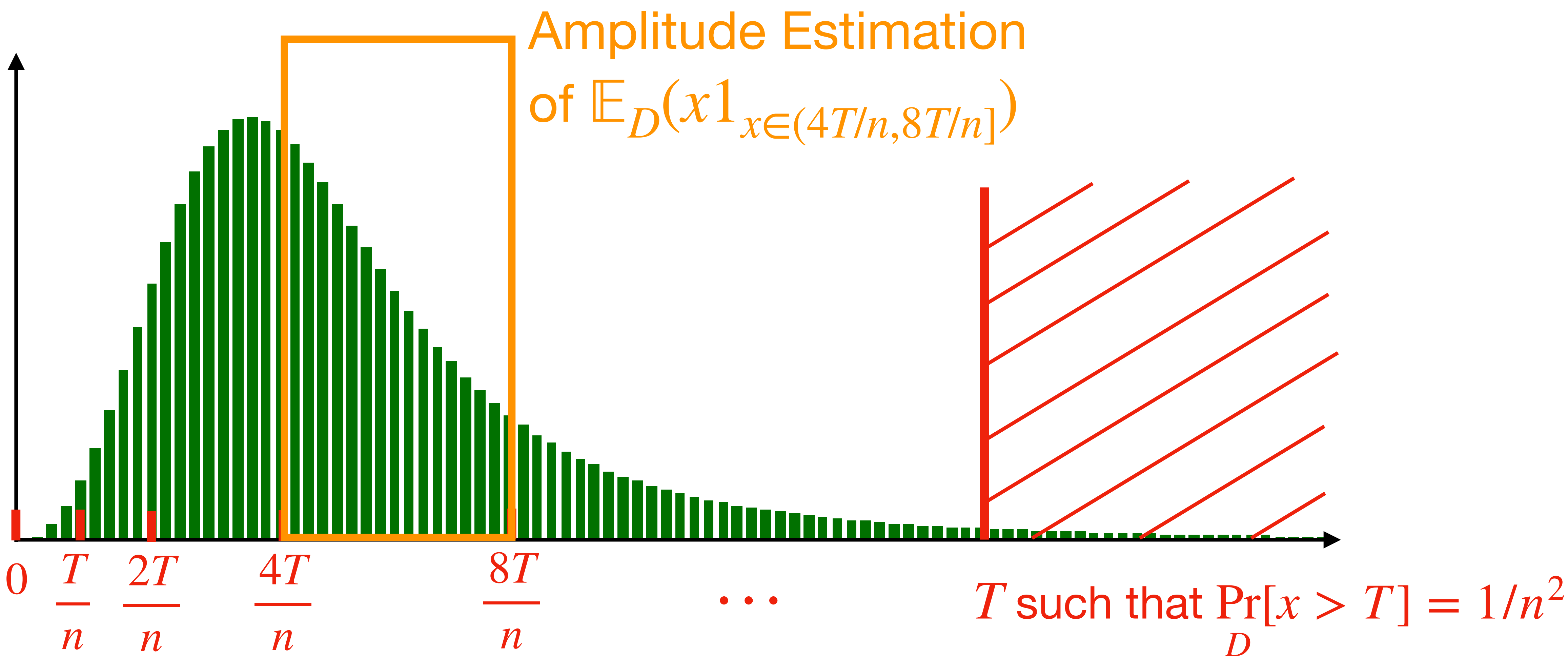
**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$



# Technique: Main algorithm

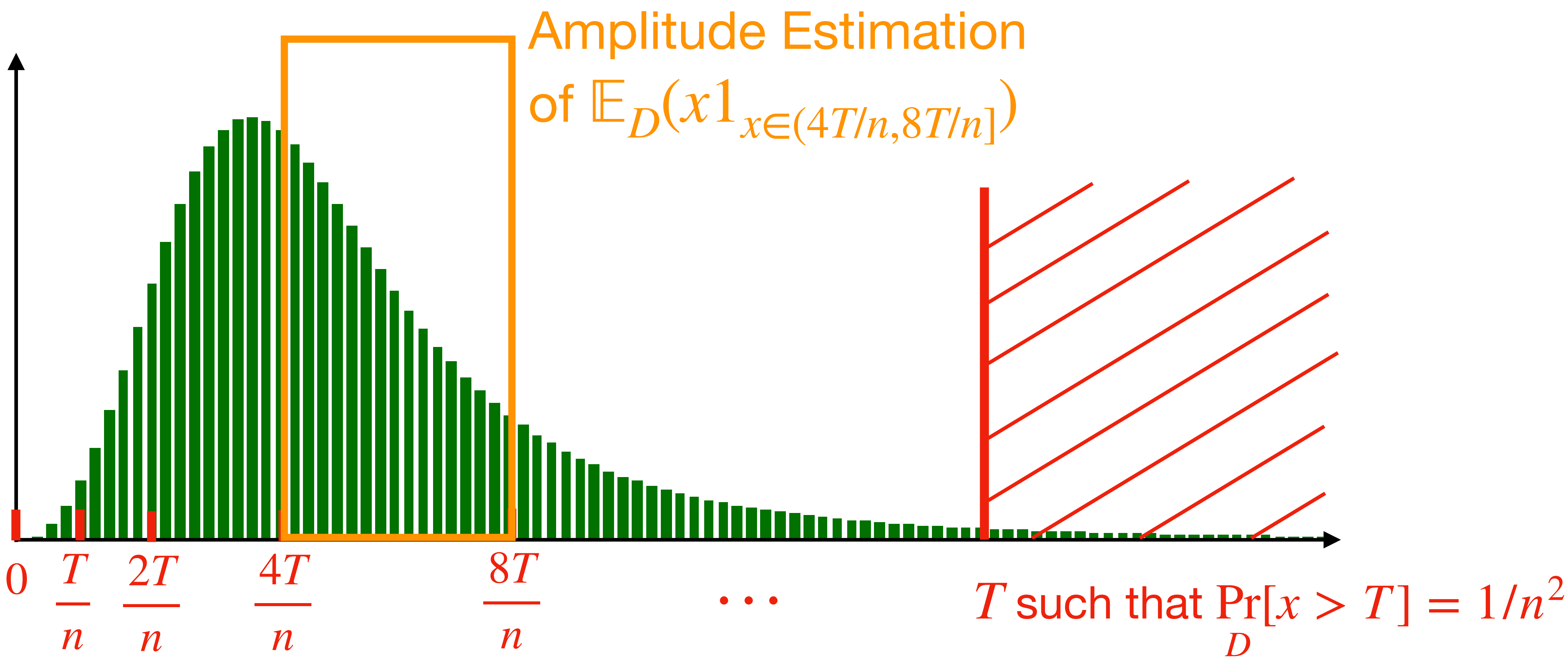


**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$



# Technique: Main algorithm

**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$

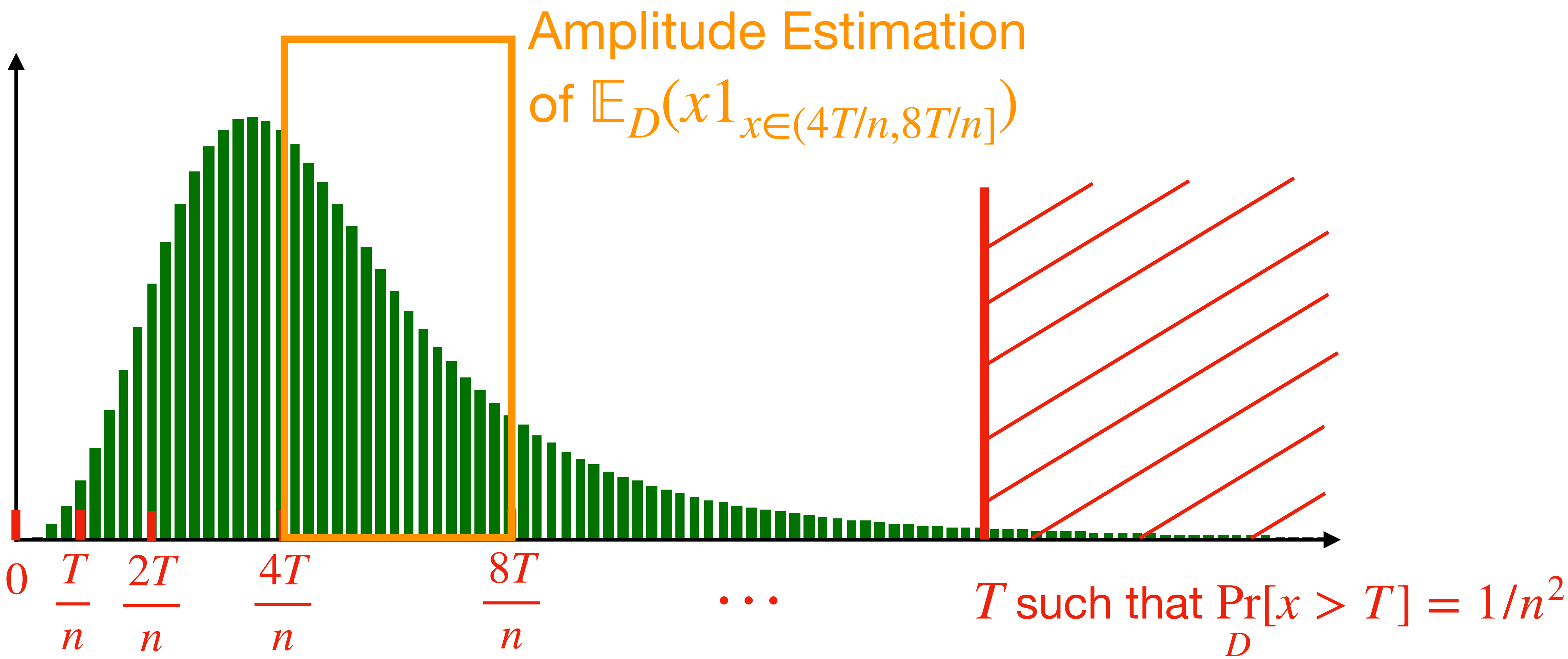



$$\tilde{\mu} = \tilde{\mu}_1 + \tilde{\mu}_2 + \tilde{\mu}_3 + \tilde{\mu}_4 + \cdots + 0$$

# Technique: Main algorithm



**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$

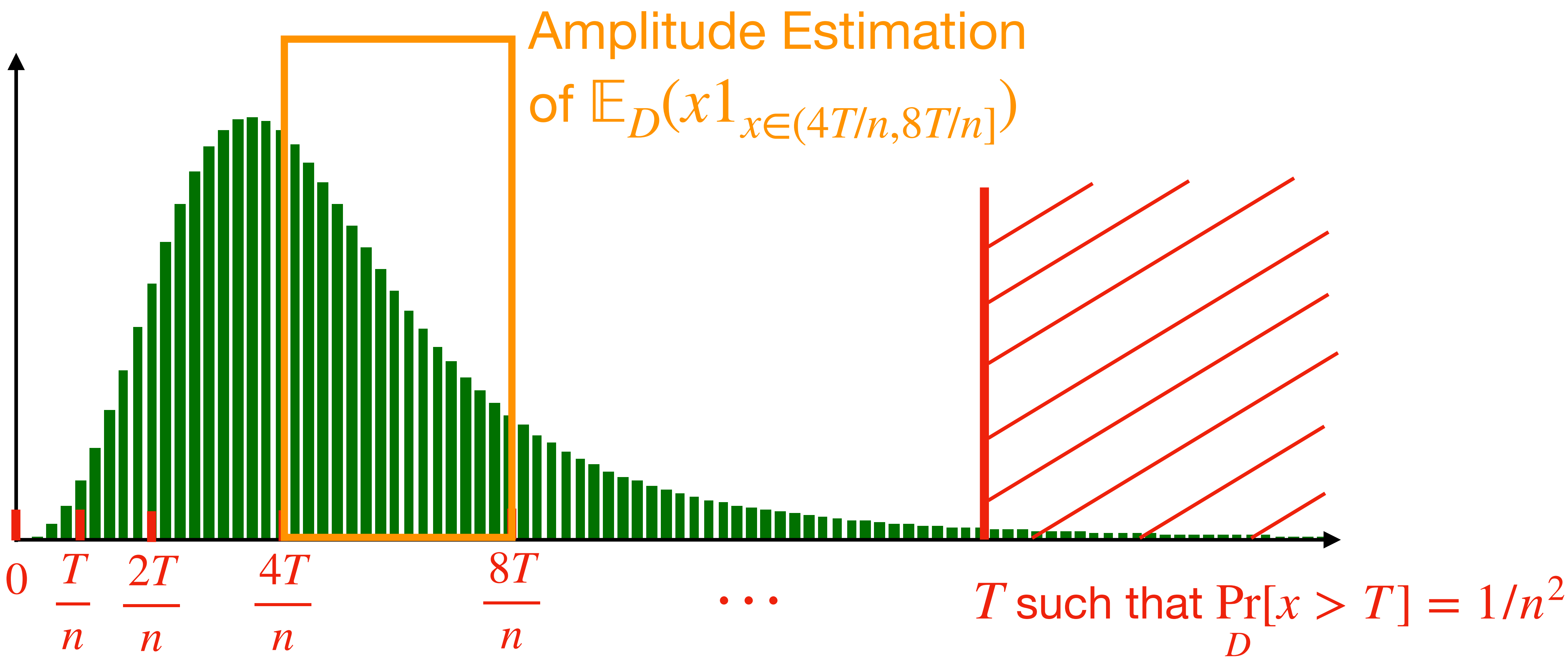



$$\tilde{\mu} = \tilde{\mu}_1 + \tilde{\mu}_2 + \tilde{\mu}_3 + \tilde{\mu}_4 + \cdots + 0$$

$$\varepsilon(n) \approx \frac{\sqrt{\mathbb{E}_D(x^2)}}{n} + \frac{T}{n^2} + \sqrt{\mathbb{E}_D(x^2) \cdot \Pr_D[x > T]}$$

# Technique: Main algorithm

**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$



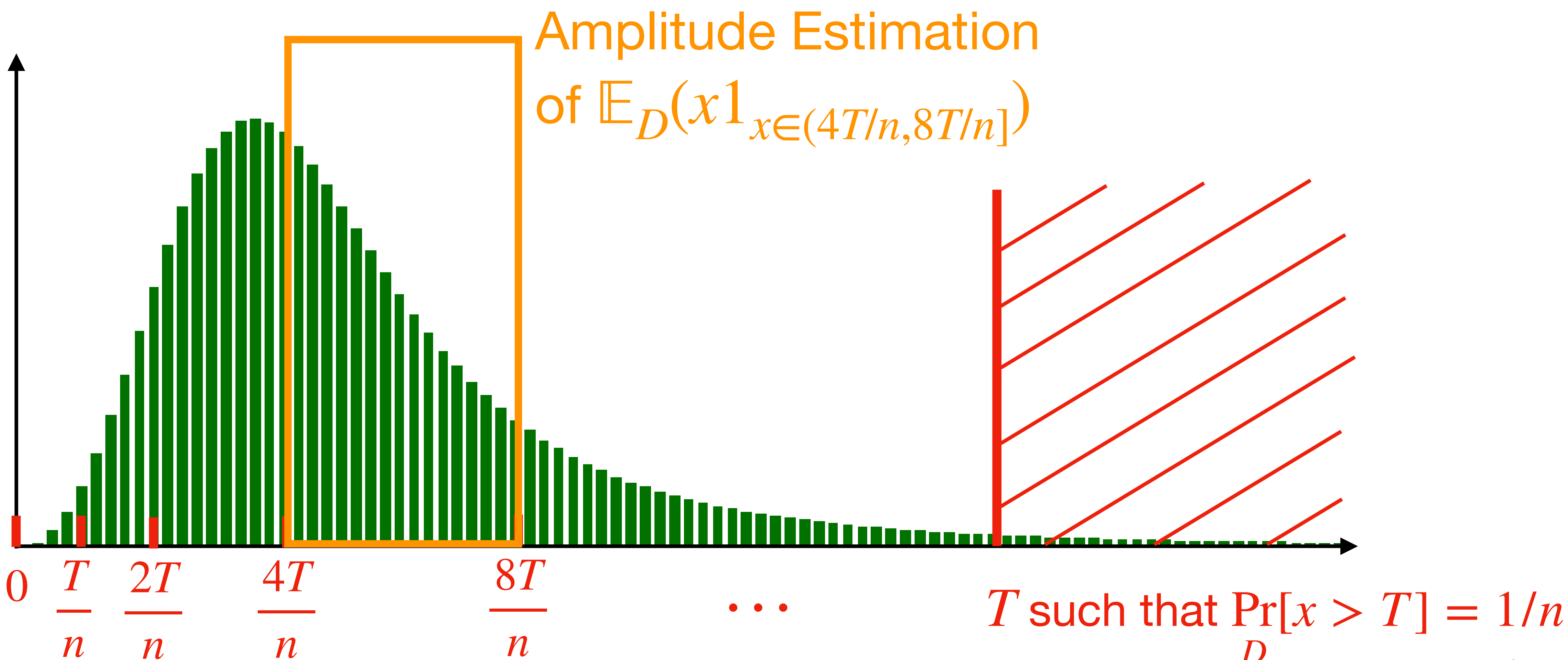

$$\tilde{\mu} = \tilde{\mu}_1 + \tilde{\mu}_2 + \tilde{\mu}_3 + \tilde{\mu}_4 + \cdots + 0$$

$$\varepsilon(n) \approx \frac{\sqrt{\mathbb{E}_D(x^2)}}{n} + \frac{T}{n^2} + \sqrt{\mathbb{E}_D(x^2) \cdot \Pr_D[x > T]}$$

$$\leq n\sqrt{\mathbb{E}_D(x^2)}$$

# Technique: Main algorithm

**Assumptions:** $\delta = 1/3$ + non-negative outcomes + $\sigma^2$ replaced with $\mathbb{E}_D(x^2)$




$$\tilde{\mu} = \tilde{\mu}_1 + \tilde{\mu}_2 + \tilde{\mu}_3 + \tilde{\mu}_4 + \cdots + 0$$

$$\varepsilon(n) \approx \frac{\sqrt{\mathbb{E}_D(x^2)}}{n} + \frac{T}{n^2} + \sqrt{\mathbb{E}_D(x^2) \cdot \Pr_D[x > T]}$$

$\leq n\sqrt{\mathbb{E}_D(x^2)}$

- **Estimating expected values of observables**

  → observable $O$, state $|\psi\rangle$

  → $\mu = \langle\psi|O|\psi\rangle$ and $\sigma^2 = \langle\psi|(O-\mu)^2|\psi\rangle$

- **Estimating expected values of observables**

  → observable $O$, state $|\psi\rangle$

  → $\mu = \langle\psi|O|\psi\rangle$ and $\sigma^2 = \langle\psi|(O-\mu)^2|\psi\rangle$

  → if $|\psi\rangle = e^{-iHt}|0\rangle$ is obtained by Hamiltonian sim. then $n =$ # runs of Hs

# Applications



- **Estimating expected values of observables**

  → observable $O$, state $|\psi\rangle$

  → $\mu = \langle\psi|O|\psi\rangle$ and $\sigma^2 = \langle\psi|(O-\mu)^2|\psi\rangle$

  → if $|\psi\rangle = e^{-iHt}|0\rangle$ is obtained by Hamiltonian sim. then $n =$ # runs of Hs

- **Monte Carlo estimation**

  → integration

  → partition functions

  → mathematical finance

# Applications

- **Estimating expected values of observables**

  → observable $O$, state $|\psi\rangle$

  → $\mu = \langle\psi|O|\psi\rangle$ and $\sigma^2 = \langle\psi|(O-\mu)^2|\psi\rangle$

  → if $|\psi\rangle = e^{-iHt}|0\rangle$ is obtained by Hamiltonian sim. then $n =$ # runs of Hs

- **Monte Carlo estimation**

  → integration

  → partition functions

  → mathematical finance

- …

# Conclusion



- **Sub-Gaussian estimators**

$$\Theta\left(\sqrt{\frac{\sigma^2 \log(1/\delta)}{n}}\right) \quad \text{vs} \quad \tilde{\Theta}\left(\frac{\sigma \log(1/\delta)}{n}\right)$$

Classical Quantum

- **Open questions**

  - Improve the log-factors

  - Extend to the **multivariate** setting (recent work: [Cornelissen,Jerbi'21])

  - Find other applications