Quantum Chebyshev's Inequality and Applications

Yassine Hamoudi, Frédéric Magniez

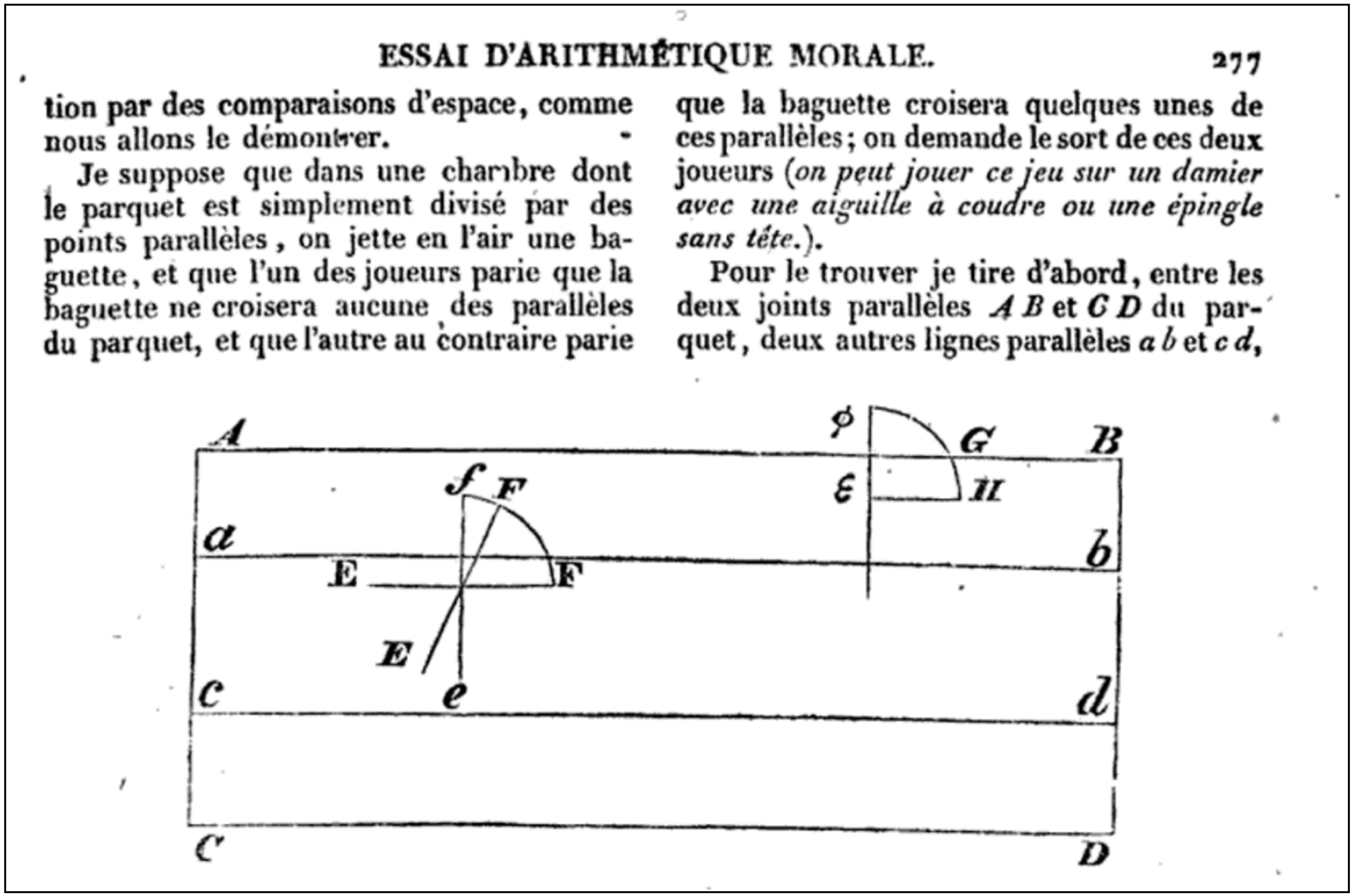
IRIF, Université Paris Diderot, CNRS

QUDATA 2019

arXiv: 1807.06456

Buffon’s needle

A needle dropped randomly on a floor with equally spaced parallel lines will cross one of the lines with probability **2/π**.



tion par des comparaisons d'espace, comme nous allons le démontrer.

Je suppose que dans une chambre dont le parquet est simplement divisé par des points parallèles, on jette en l'air une baguette, et que l'un des joueurs parie que la baguette ne croisera aucune des parallèles du parquet, et que l'autre au contraire parie que la baguette croisera quelques unes de ces parallèles; on demande le sort de ces deux joueurs (*on peut jouer ce jeu sur un damier avec une aiguille à coudre ou une épingle sans tête*).

Pour le trouver je tire d'abord, entre les deux joints parallèles *A B* et *C D* du parquet, deux autres lignes parallèles *a b* et *c d*,

Buffon, G., *Essai d'arithmétique morale*, 1777.

Monte Carlo algorithms: Use repeated **random sampling** and statistical analysis to estimate parameters of interest

Monte Carlo algorithms: Use repeated **random sampling** and statistical analysis to estimate parameters of interest

Empirical mean:

1/ Repeat the experiment n times: n i.i.d. **samples** $x_1, \ldots, x_n \sim$ **X**

2/ Output: **$(x_1 + \ldots + x_n)/n$**

Monte Carlo algorithms: Use repeated **random sampling** and statistical analysis to estimate parameters of interest

Empirical mean:

1/ Repeat the experiment n times: n i.i.d. **samples** $x_1, \ldots, x_n \sim$ **X**

2/ Output: **$(x_1 + \ldots + x_n)/n$**

Law of large numbers: $$\frac{x_1 + \ldots + x_n}{n} \xrightarrow{n \to \infty} \mathbf{E}(X)$$

Empirical mean: $\widetilde{\mu} = \frac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Empirical mean: $\widetilde{\mu} = \dfrac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Chebyshev's Inequality:

multiplicative error 0 < ε < 1

Objective: $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability ($\mathbf{E}(X), \mathbf{Var}(X) \neq 0$ finite)

Empirical mean: $\widetilde{\mu} = \dfrac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Chebyshev's Inequality:

multiplicative error 0 < ε < 1

Objective: $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$ with high probability ($\mathbf{E}(X), \mathbf{Var}(X) \neq 0$ finite)

Number of samples needed: $O\left(\dfrac{\mathbf{E}(X^2)}{\epsilon^2\mathbf{E}(X)^2}\right)$ (in fact $O\left(\frac{\mathbf{Var}(X)}{\epsilon^2\mathbf{E}(X)^2}\right) = O\left(\frac{1}{\epsilon^2}\left(\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2} - 1\right)\right)$)

Empirical mean: $\widetilde{\mu} = \dfrac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Chebyshev's Inequality:

multiplicative error 0 < ε < 1

Objective: $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability ($\mathbf{E}(X), \mathbf{Var}(X) \neq 0$ finite)

Relative second moment

Number of samples needed: $O\left(\dfrac{\mathbf{E}(X^2)}{\epsilon^2 \mathbf{E}(X)^2}\right)$ (in fact $O\left(\frac{\mathbf{Var}(X)}{\epsilon^2 \mathbf{E}(X)^2}\right) = O\left(\frac{1}{\epsilon^2}\left(\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2} - 1\right)\right)$)

Empirical mean: $\widetilde{\mu} = \dfrac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Chebyshev's Inequality:

multiplicative error 0 < ε < 1

Objective: $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability ($\mathbf{E}(X), \mathbf{Var}(X) \neq 0$ finite)

Relative second moment

Number of samples needed: $O\left(\dfrac{\mathbf{E}(X^2)}{\epsilon^2 \mathbf{E}(X)^2}\right)$ (in fact $O\left(\frac{\mathbf{Var}(X)}{\epsilon^2 \mathbf{E}(X)^2}\right) = O\left(\frac{1}{\epsilon^2}\left(\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2} - 1\right)\right)$)

In practice: given an upper-bound $\Delta^2 \geq \dfrac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$, take $n = \Omega\left(\dfrac{\Delta^2}{\epsilon^2}\right)$ samples

Counting with Markov chain Monte Carlo methods:

Counting vs. sampling [Jerrum, Sinclair'96] [Štefankovič et al.'09], Volume of convex bodies [Dyer, Frieze'91], Permanent [Jerrum, Sinclair, Vigoda'04]

Data stream model:

Frequency moments, Collision probability [Alon, Matias, Szegedy'99] [Monemizadeh, Woodruff'] [Andoni et al.'11] [Crouch et al.'16]

Testing properties of distributions:

Closeness [Goldreich, Ron'11] [Batu et al.'13] [Chan et al.'14], Conditional independence [Canonne et al.'18]

Estimating graph parameters:

Number of connected components, Minimum spanning tree weight [Chazelle, Rubinfeld, Trevisan'05], Average distance [Goldreich, Ron'08], Number of triangles [Eden et al. 17]

etc.

Random variable X over sample space $\Omega \subset R^+$

Classical sample: one value $x \in \Omega$, sampled with probability p_x

Random variable X over sample space $\Omega \subset \mathbf{R}^+$

Classical sample: one value $x \in \Omega$, sampled with probability p_x

Quantum sample: one (controlled-)**execution** of a **quantum sampler** S_X or S_X^{-1}, where

$$S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle\,|x\rangle$$

with $\psi_x =$ arbitrary unit vector

Can we use quadratically less samples in the quantum setting?

Can we use quadratically less samples in the quantum setting?

	Number of samples	Conditions
Classical samples (Chebyshev's inequality)	$\frac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$

Can we use quadratically less samples in the quantum setting?

	Number of samples	Conditions
Classical samples (Chebyshev's inequality)	$\dfrac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \dfrac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\dfrac{\sqrt{\mathbf{B}}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space $\Omega \subset [0,B]$

Can we use quadratically less samples in the quantum setting?

	Number of samples	Conditions
Classical samples (Chebyshev's inequality)	$\frac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\frac{\sqrt{\mathbf{B}}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space $\Omega \subset [0,B]$
[Montanaro'15]	$\frac{\Delta^{\mathbf{2}}}{\epsilon}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$

Can we use quadratically less samples in the quantum setting?

	Number of samples	Conditions
Classical samples (Chebyshev's inequality)	$\frac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\frac{\sqrt{\mathbf{B}}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space **Ω ⊂ [0,B]**
[Montanaro'15]	$\frac{\Delta^{\mathbf{2}}}{\epsilon}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Li, Wu'17]	$\frac{\Delta}{\epsilon} \cdot \frac{H}{L}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ **L** ≤ E(X) ≤ **H**

Can we use quadratically less samples in the quantum setting?

	Number of samples	Conditions	
Classical samples (Chebyshev's inequality)	$\frac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\frac{\sqrt{\mathbf{B}}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space $\Omega \subset [0,B]$	
[Montanaro'15]	$\frac{\Delta^{2}}{\epsilon}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	
[Li, Wu'17]	$\frac{\Delta}{\epsilon} \cdot \frac{H}{L}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	L ≤ E(X) ≤ H
Our result	$\frac{\Delta}{\epsilon} \cdot \log^3\left(\frac{H}{\mathbf{E}(X)}\right)$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	E(X) ≤ H

Our Approach

Amplitude Estimation Algorithm [Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]

Input: Random variable **X** on sample space **Ω** ⊂ **[0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

Amplitude Estimation Algorithm [Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]

Input: Random variable **X** on sample space **Ω** ⊂ **[0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon \mathbf{E}(X)}\right)$ ✔

Amplitude Estimation Algorithm [Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]

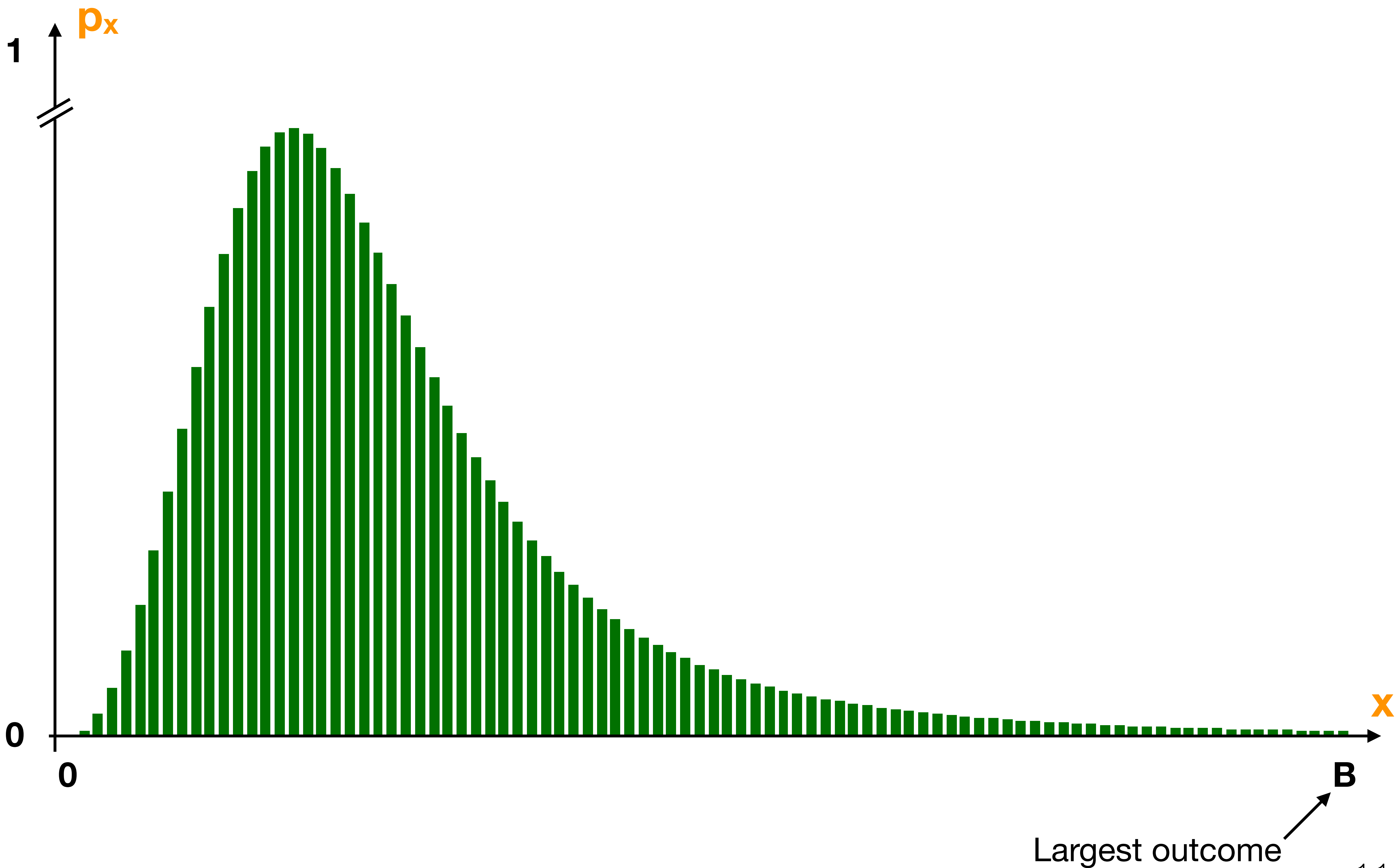
Input: Random variable **X** on sample space **Ω** $\subset$ **[0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

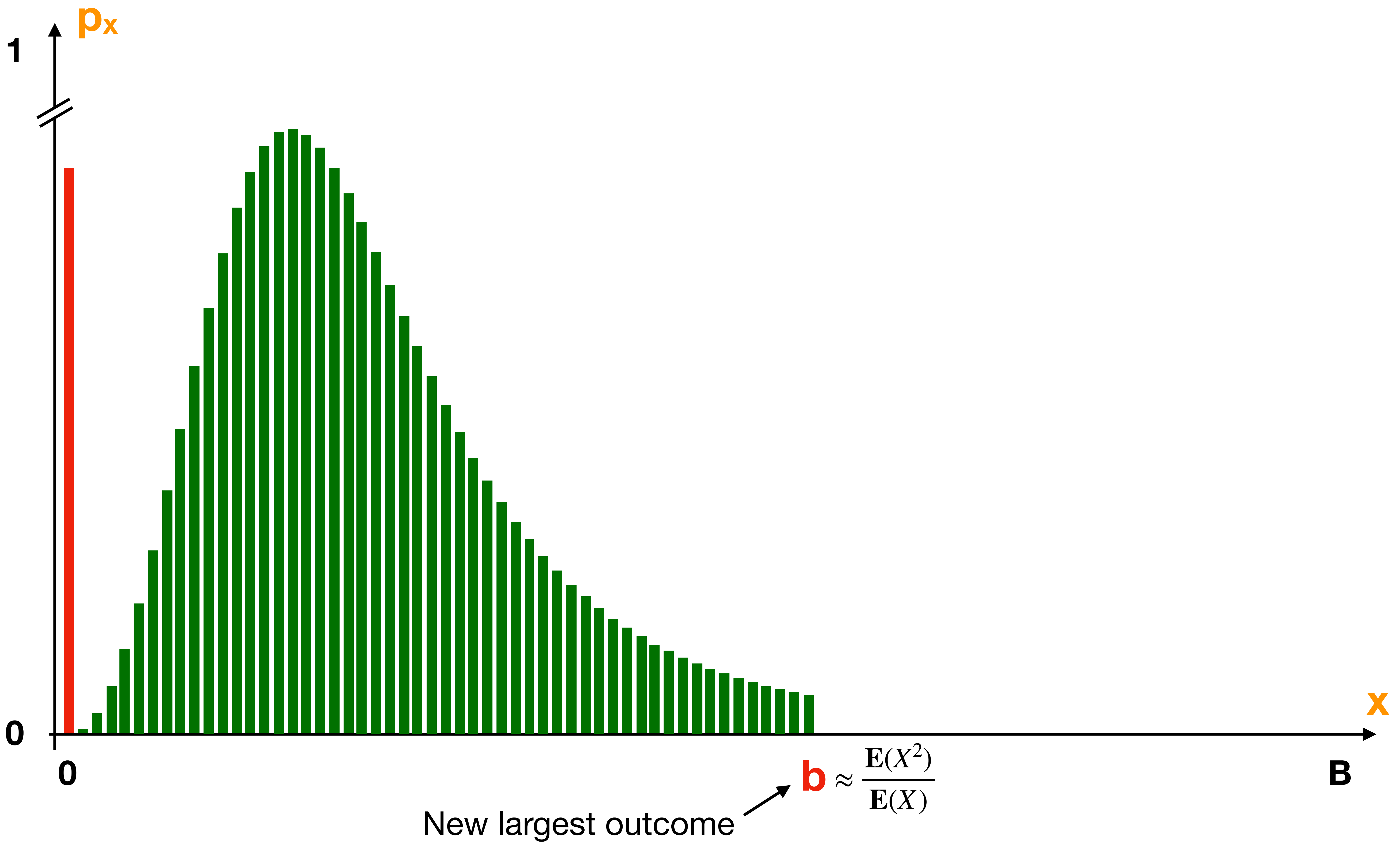
If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon \mathbf{E}(X)}\right)$ ✔

If $B \gg \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$?

Random variable X



Random variable X_b



Input: Random variable **X** on sample space **$\Omega \subset$ [0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon \mathbf{E}(X)}\right)$ ✓

If $B \gg \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: map the outcomes larger than $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ to 0 **?**

Input: Random variable **X** on sample space **$\Omega \subset$ [0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon\mathbf{E}(X)}\right)$ ✓

If $B \gg \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: map the outcomes larger than $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ to 0 ✓

Lemma: If $b \geq \frac{\mathbf{E}(X^2)}{\epsilon\mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \leq \mathbf{E}(X_b) \leq \mathbf{E}(X)$.

Input: Random variable **X** on sample space **Ω ⊂ [0,B]**

Ampl-Est: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \cdot \mathbf{E}(X)$

If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon \mathbf{E}(X)}\right)$ ✓

If $B \gg \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: map the outcomes larger than $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ to 0 ✓

★ **Lemma:** If $b \geq \frac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \leq \mathbf{E}(X_b) \leq \mathbf{E}(X)$.

Problem: given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ how to find a threshold $b \approx \mathbf{E}(X) \cdot \Delta^2$?

Problem: given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ how to find a threshold $b \approx \mathbf{E}(X) \cdot \Delta^2$?

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on b (given an upper-bound H ≥ E(X))

Problem: given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ how to find a threshold $b \approx \mathbf{E}(X) \cdot \Delta^2$?

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on b (given an upper-bound H ≥ E(X))

Threshold	Input r.v.	Number of samples	Estimation
$b_0 = H\Delta^2$	X_{b_0}	Δ	$\widetilde{\mu}_0$
$b_1 = (H/2)\Delta^2$	X_{b_1}	Δ	$\widetilde{\mu}_1$
$b_2 = (H/4)\Delta^2$	X_{b_2}	Δ	$\widetilde{\mu}_2$
$\cdots$			$\cdots$

Stopping rule: $\widetilde{\mu}_i \neq 0$ **Output:** b_i

Problem: given $\Delta^2 \geq \dfrac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ how to find a threshold $b \approx \mathbf{E}(X) \cdot \Delta^2$?

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on b (given an upper-bound H ≥ E(X))

Threshold	Input r.v.	Number of samples	Estimation
$b_0 = H\Delta^2$	X_{b_0}	Δ	$\widetilde{\mu}_0$
$b_1 = (H/2)\Delta^2$	X_{b_1}	Δ	$\widetilde{\mu}_1$
$b_2 = (H/4)\Delta^2$	X_{b_2}	Δ	$\widetilde{\mu}_2$
$\cdots$			$\cdots$

Stopping rule: $\widetilde{\mu}_i \neq 0$ **Output:** b_i

Theorem: the first non-zero $\widetilde{\mu}_i$ is obtained w.h.p. when:

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10 \cdot \mathbf{E}(X)\Delta^2$$

Theorem: the first non-zero $\widetilde{\mu}_i$ is obtained w.h.p. when:

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10 \cdot \mathbf{E}(X)\Delta^2$$

Theorem: the first non-zero $\widetilde{\mu}_i$ is obtained w.h.p. when:

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10 \cdot \mathbf{E}(X)\Delta^2$$

Ingredient 1: The output of **Amplitude-Estimation** is 0 w.h.p. if and only if the **normalized estimated mean** is below the **inverse-square number** of samples.

[Brassard et al.'02]

$$\frac{\mathbf{E}(X_b)}{b}$$

$$1/\Delta^2$$

Analysis

Theorem: the first non-zero $\widetilde{\mu}_i$ is obtained w.h.p. when:

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10 \cdot \mathbf{E}(X)\Delta^2$$

Ingredient 1: The output of **Amplitude-Estimation** is 0 w.h.p. if and only if the **normalized estimated mean** is below the **inverse-square number** of samples.

[Brassard et al.'02]

$$\frac{\mathbf{E}(X_b)}{b} \qquad 1/\Delta^2$$

Ingredient 2: **If** $b \geq 10 \cdot \mathbf{E}(X)\Delta^2$ **then** $\dfrac{\mathbf{E}(X_b)}{b} \leq \dfrac{\mathbf{E}(X)}{b} \leq \dfrac{1}{10 \cdot \Delta^2}$

Analysis

Theorem: the first non-zero $\widetilde{\mu}_i$ is obtained w.h.p. when:

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10 \cdot \mathbf{E}(X)\Delta^2$$

Ingredient 1: The output of **Amplitude-Estimation** is 0 w.h.p. if and only if the **normalized estimated mean** is below the **inverse-square number** of samples.
[Brassard et al.'02]

$$\frac{\mathbf{E}(X_b)}{b} \qquad 1/\Delta^2$$

Ingredient 2: **If** $b \geq 10 \cdot \mathbf{E}(X)\Delta^2$ **then** $\dfrac{\mathbf{E}(X_b)}{b} \leq \dfrac{\mathbf{E}(X)}{b} \leq \dfrac{1}{10 \cdot \Delta^2}$

Ingredient 3: **If** $b \approx \mathbf{E}(X) \cdot \Delta^2$ **then** $\dfrac{\mathbf{E}(X_b)}{b} \overset{\star}{\approx} \dfrac{\mathbf{E}(X)}{b} \approx \dfrac{1}{\Delta^2}$

Applications

Application 1: approximating graph parameters

Input: graph G=(V,E) with **n** vertices, **m** edges, **t** triangles

Query access: unitaries $O_{\mathrm{deg}}|v\rangle|0\rangle = |v\rangle|\mathrm{deg}(v)\rangle$ *(degree query)*

$O_{\mathrm{pair}}|v\rangle|w\rangle|0\rangle = |v\rangle|w\rangle|(v,w) \in E\ ?\rangle$ *(pair query)*

$O_{\mathrm{ngh}}|v\rangle|i\rangle|0\rangle = |v\rangle|i\rangle|v_i\rangle$ *(neighbor query)*

i^{th} neighbor of v

Application 1: approximating graph parameters

Input: graph G=(V,E) with **n** vertices, **m** edges, **t** triangles

Query access: unitaries

$$O_{\text{deg}}|v\rangle|0\rangle = |v\rangle|\deg(v)\rangle$$ *(degree query)*

$$O_{\text{pair}}|v\rangle|w\rangle|0\rangle = |v\rangle|w\rangle|(v,w) \in E\ ?\rangle$$ *(pair query)*

$$O_{\text{ngh}}|v\rangle|i\rangle|0\rangle = |v\rangle|i\rangle|v_i\rangle$$ *(neighbor query)*

i^{th} neighbor of v

Result: $\widetilde{\Theta}\left(\frac{\sqrt{n}}{m^{1/4}}\right)$ degree/neighbor quantum queries to approximate **m**

$\widetilde{\Theta}\left(\frac{\sqrt{n}}{t^{1/6}} + \frac{m^{3/4}}{\sqrt{t}}\right)$ degree/pair/neighbor quantum queries to approximate **t**

Application 1: approximating graph parameters

Input: graph G=(V,E) with **n** vertices, **m** edges, **t** triangles

Query access: unitaries $O_{\deg}|v\rangle|0\rangle = |v\rangle|\deg(v)\rangle$ *(degree query)*

$O_{\text{pair}}|v\rangle|w\rangle|0\rangle = |v\rangle|w\rangle|(v,w) \in E\ ?\rangle$ *(pair query)*

$O_{\text{ngh}}|v\rangle|i\rangle|0\rangle = |v\rangle|i\rangle|v_i\rangle$ *(neighbor query)*

i^{th} neighbor of v

Result: $\widetilde{\Theta}\left(\frac{\sqrt{n}}{m^{1/4}}\right)$ degree/neighbor quantum queries to approximate **m**

(vs. $\widetilde{\Theta}\left(\frac{n}{\sqrt{m}}\right)$ classical degree/neighbor queries)

[Goldreich, Ron'08] [Seshadhri'15]

$\widetilde{\Theta}\left(\frac{\sqrt{n}}{t^{1/6}} + \frac{m^{3/4}}{\sqrt{t}}\right)$ degree/pair/neighbor quantum queries to approximate **t**

(vs. $\widetilde{\Theta}\left(\frac{n}{t^{1/3}} + \frac{m^{3/2}}{t}\right)$ classical degree/pair/neighbor queries)

[Eden, Levi, Ron'15] [Eden, Levi, Ron, Seshadhri'17]

Application 2: frequency moments in the streaming model

Input: (finite) stream of updates $\mathbf{x_i} \leftarrow \mathbf{x_i} + \delta$ on **x** = (0,…,0) of **dimension n**

Output: (at the end of the stream) approximate of $F_k = \sum_{i=1}^{n} |x_i|^k$ (moment of order k ≥ 3)

Application 2: frequency moments in the streaming model

Input: (finite) stream of updates $\mathbf{x_i} \leftarrow \mathbf{x_i} + \delta$ on **x** = (0,…,0) of **dimension n**

Output: (at the end of the stream) approximate of $F_k = \sum_{i=1}^{n} |x_i|^k$ (moment of order k ≥ 3)

Algorithm with smallest possible **memory M**
using **P passes** over the same stream?

Application 2: frequency moments in the streaming model

Input: (finite) stream of updates $\mathbf{x_i} \leftarrow \mathbf{x_i} + \delta$ on **x** = (0,…,0) of **dimension n**

Output: (at the end of the stream) approximate of $F_k = \sum_{i=1}^{n} |x_i|^k$ (moment of order k ≥ 3)

Algorithm with smallest possible **memory M** using **P passes** over the same stream?

Result: $\mathbf{M = \widetilde{O}\left(\frac{n^{1-2/k}}{P^2}\right)}$ qubits of memory

(vs. $M = \widetilde{\Theta}\left(\frac{n^{1-2/k}}{P}\right)$ classical bits of memory)

[Monemizadeh, Woodruff'10]
[Andoni, Krauthgamer, Onak'11]

Conclusion

The **mean** of a random variable X can be estimated with **multiplicative error ε** using $\widetilde{O}\left(\frac{\Delta}{\epsilon} \cdot \log^3\left(\frac{H}{E(X)}\right)\right)$ **quantum samples**, given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ and $H \geq \mathbf{E}(X)$.

arXiv: 1807.06456

The **mean** of a random variable X can be estimated with **multiplicative error ε** using $\widetilde{O}\left(\frac{\Delta}{\epsilon} \cdot \log^3\left(\frac{H}{E(X)}\right)\right)$ **quantum samples**, given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ and $H \geq \mathbf{E}(X)$.

Lower bound: $\Omega\left(\frac{\Delta - 1}{\epsilon}\right)$ quantum samples

arXiv: 1807.06456

The **mean** of a random variable X can be estimated with **multiplicative error ε** using $\widetilde{O}\left(\frac{\Delta}{\epsilon} \cdot \log^3\left(\frac{H}{E(X)}\right)\right)$ **quantum samples**, given $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ and $H \geq \mathbf{E}(X)$.

Lower bound: $\Omega\left(\frac{\Delta - 1}{\epsilon}\right)$ quantum samples

or $\Omega\left(\frac{\Delta^2 - 1}{\epsilon^2}\right)$ copies of the state $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$

arXiv: 1807.06456

Extra slides

Subroutine: the Amplitude Estimation algorithm

Sampler: $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle\,|x\rangle$ on sample space **$\Omega \subset$ [0,B]**

Result: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

Subroutine: the Amplitude Estimation algorithm

Sampler: $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ on sample space **Ω ⊂ [0,B]**

Result: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$

Reduction to a Bernoulli sampler [Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]:

$$\sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle|0\rangle \xrightarrow{\text{Controlled rotation}} \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{B}}\,|0\rangle + \sqrt{\frac{x}{B}}\,|1\rangle\right)$$

$$\xrightarrow{\text{Reordering}} \sqrt{1-\frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_0\rangle|\mathbf{0}\rangle + \sqrt{\frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_1\rangle|\mathbf{1}\rangle = S_Y|0\rangle$$

Subroutine: the Amplitude Estimation algorithm

Sampler: $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ on sample space $\Omega \subset [0,B]$

Result: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

Expectation of a Bernoulli sampler [Brassard et al.'02]:

$$\mathbf{S_Y}|\mathbf{0}\rangle = \sqrt{\mathbf{1} - \frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_0\rangle|\mathbf{0}\rangle + \sqrt{\frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_1\rangle|\mathbf{1}\rangle$$

Subroutine: the Amplitude Estimation algorithm

Sampler: $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ on sample space **Ω ⊂ [0,B]**

Result: $O\left(\frac{\sqrt{B}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ quantum samples to obtain $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

Expectation of a Bernoulli sampler [Brassard et al.'02]:

$$\mathbf{S_Y}|\mathbf{0}\rangle = \sqrt{\mathbf{1} - \frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_0\rangle|\mathbf{0}\rangle + \sqrt{\frac{\mathbf{E(X)}}{\mathbf{B}}}\,|\varphi_1\rangle|\mathbf{1}\rangle$$

Step 0: the Grover's operator $\mathbf{G} = \mathbf{S_Y^{-1}}(I - 2|0\rangle\langle 0|)\mathbf{S_Y}(I - 2I\otimes|1\rangle\langle 1|)$ has eigenvalues $e^{\pm 2i\theta}$, where $\theta = \sin^{-1}(\sqrt{\mathbf{E}(X)/B})$.

Step 1: use the Phase Estimation Algorithm on G for $t \geq \Omega(\sqrt{B}/(\epsilon\sqrt{\mathbf{E}(X)}))$ steps (i.e. using t quantum samples), to get an estimate $\widetilde{\theta}$ of $\pm\theta$.

Step 2: output $\sin^2(\widetilde{\theta})$ as an estimate to E(X)/B. $\quad(\widetilde{\mu} = B\cdot\sin^2(\widetilde{\theta}))$

No a priori information on $E(X^2)/E(X)^2$

Result: There is an **optimal** algorithm that approximates the mean of any quantum sampler S_X over $\Omega \subset [0,B]$ with

$$\widetilde{\Theta}\left(\frac{\sqrt{B}}{\sqrt{\epsilon \mathbf{E}(X)}} + \frac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}\right)$$

quantum samples, when there is no a priori information on X.

→ Quantization of [Dagum, Karp, Luby, Ross'00]

Lemma: If $b \geq \frac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \leq \mathbf{E}(X_{<b}) \leq \mathbf{E}(X)$.

Lemma: If $b \geq 10^4 \cdot \mathbf{E}(X)\Delta^2$ then $\frac{\mathbf{E}(X_{<b})}{b} \leq \frac{1}{10^4 \cdot \Delta^2}$

Lemma: If $b \geq \dfrac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \leq \mathbf{E}(X_{<b}) \leq \mathbf{E}(X)$.

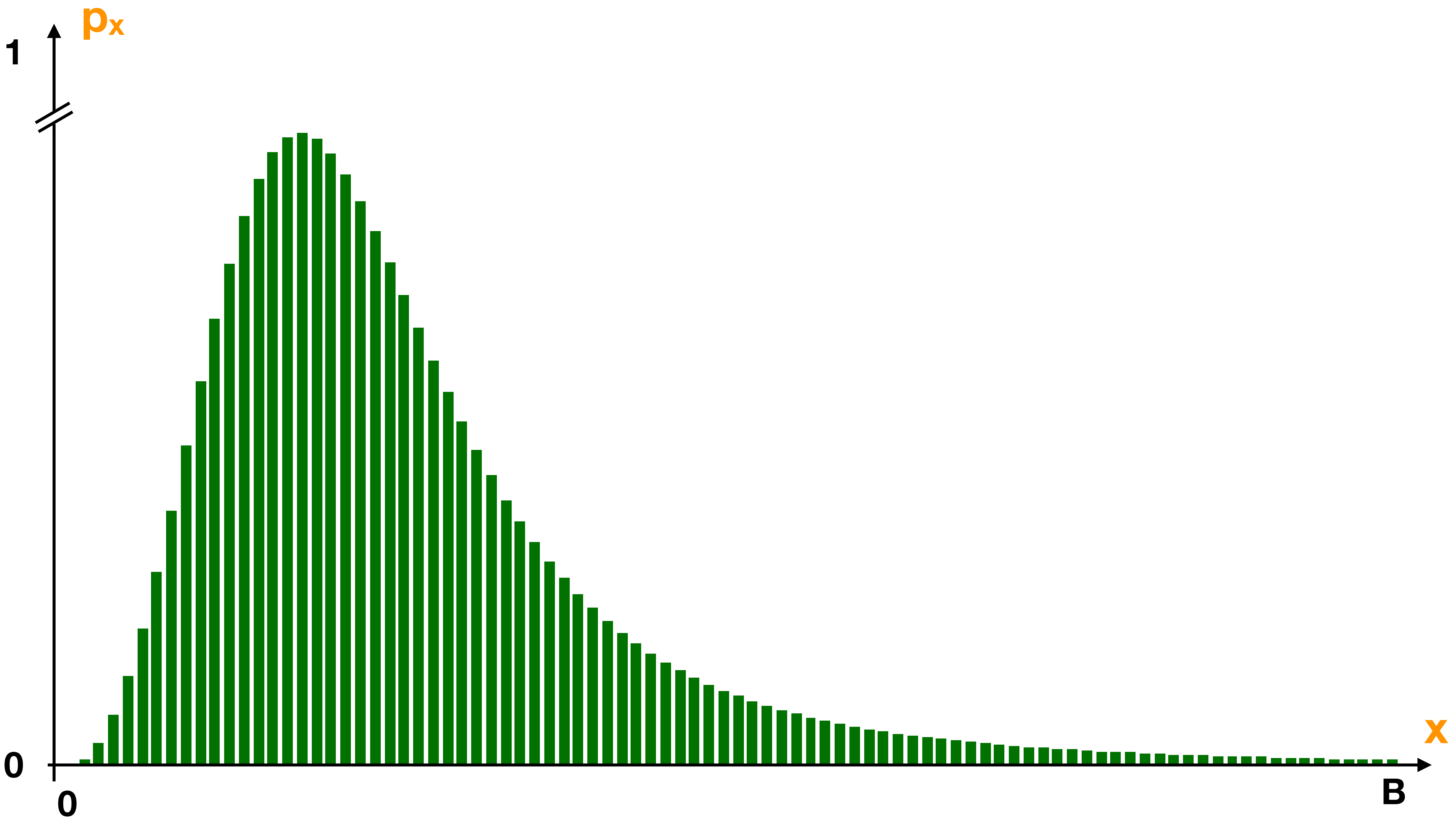
Proof:

- $\mathbf{E}(X_{\geq b}) \leq \dfrac{\mathbf{E}(X^2)}{b} \leq \epsilon \mathbf{E}(X)$
- $\mathbf{E}(X_{<b}) = \mathbf{E}(X) - \mathbf{E}(X_{\geq b}) \geq (1-\epsilon)\mathbf{E}(X)$

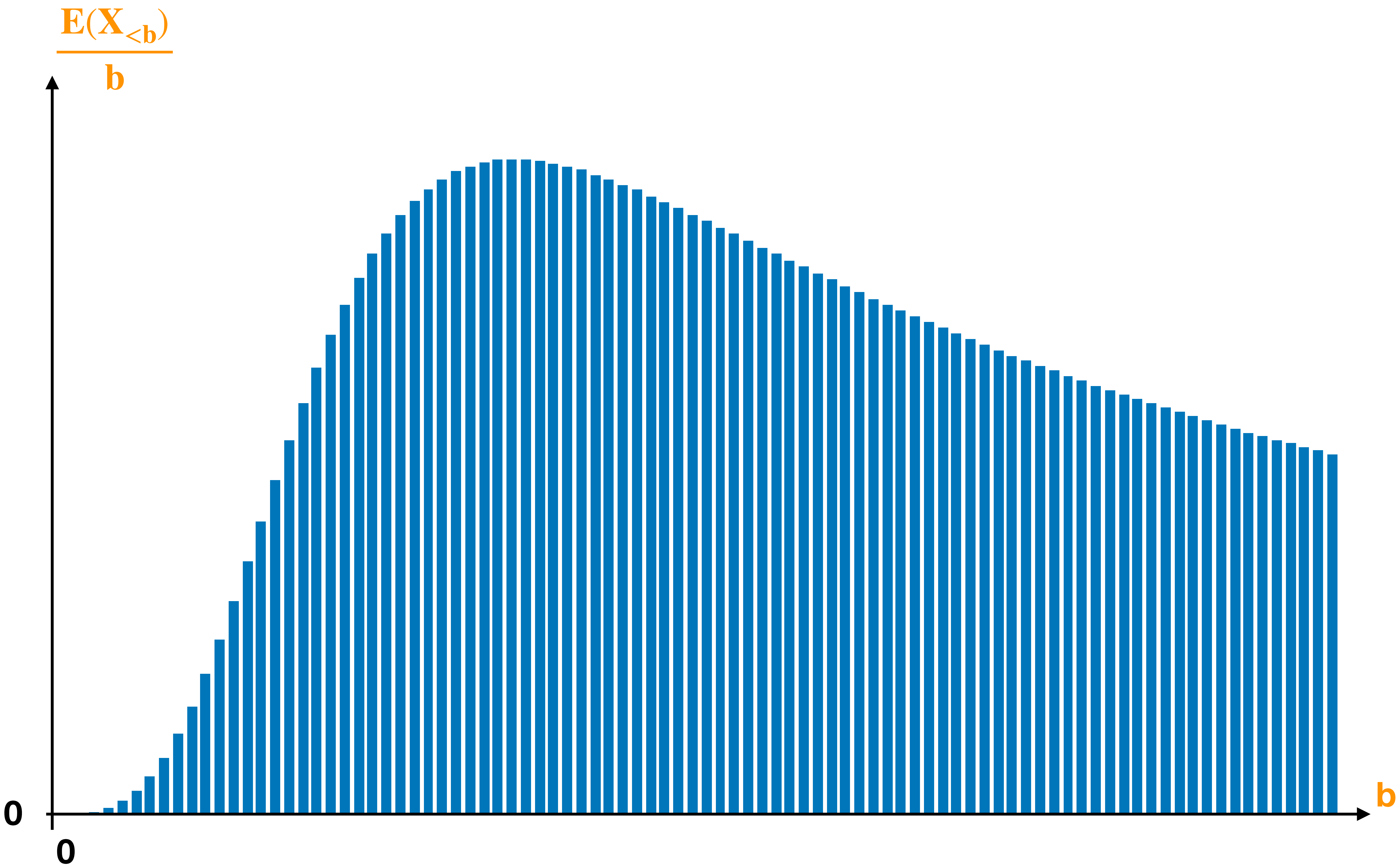
Lemma: If $b \geq 10^4 \cdot \mathbf{E}(X)\Delta^2$ then $\dfrac{\mathbf{E}(X_{<b})}{b} \leq \dfrac{1}{10^4 \cdot \Delta^2}$

Proof: $\dfrac{\mathbf{E}(X_{<b})}{b} \leq \dfrac{\mathbf{E}(X)}{10^4 \mathbf{E}(X)\Delta^2} \leq \dfrac{1}{10^4 \cdot \Delta^2}$

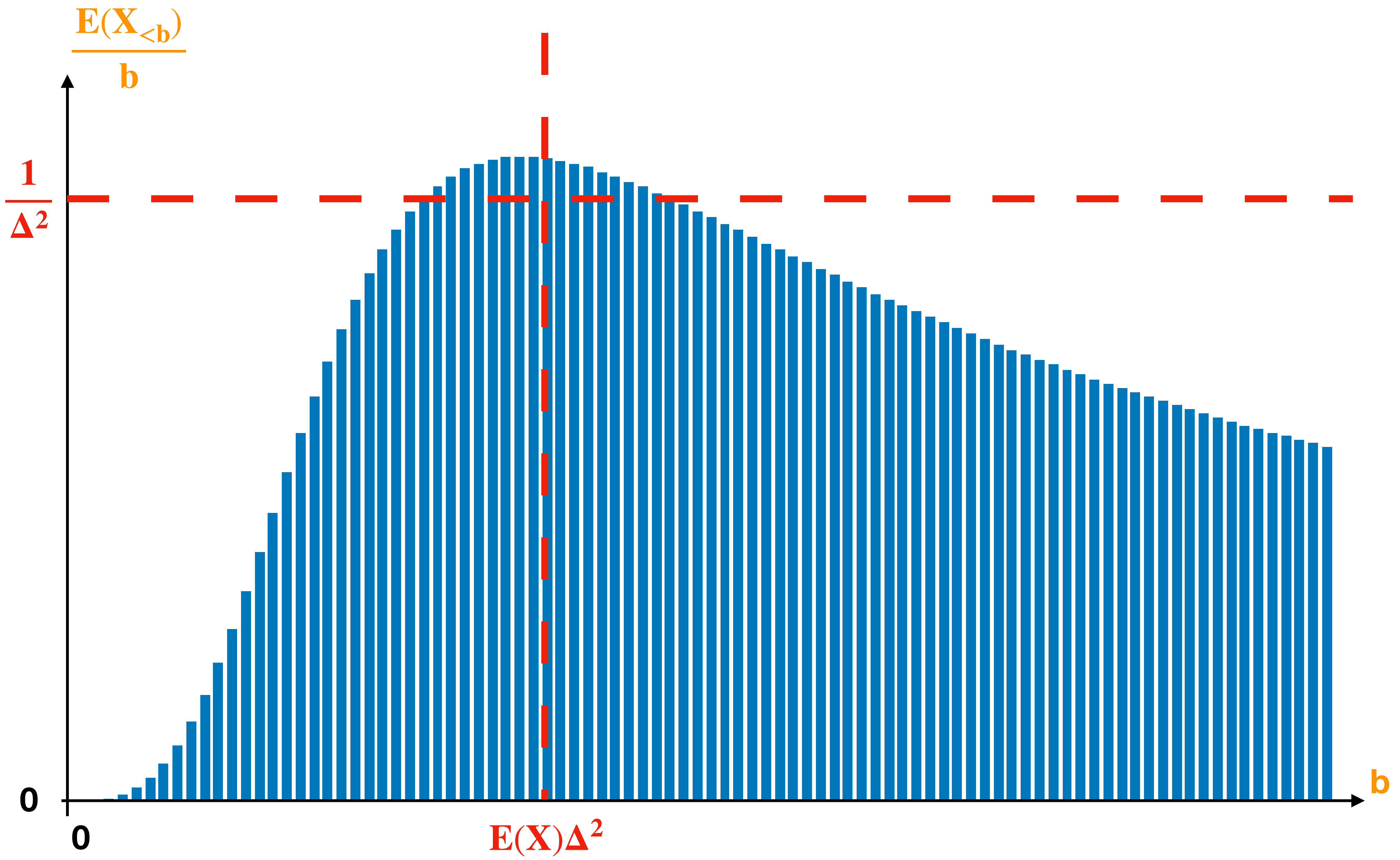
Example



Example



Example



Final algorithm:

Step 1: Logarithmic search on b until **Amplitude-Estimation**$(S_{X_{<b}}, \Delta) \neq 0$

⟶ get $2 \cdot \mathbf{E}(X)\Delta^2 \leq b \leq 10^4 \cdot \mathbf{E}(X)\Delta^2$ with high probability

$$\Delta \cdot \log^3\left(\frac{H}{\mathbf{E}(X)}\right)$$

Step 2: Set threshold $d = b/\epsilon$ and output **Amplitude-Estimation**$(S_{X_{<d}}, \Delta/\epsilon^{3/2}) \neq 0$

⟶ get $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$ with high probability

$$\Delta/\epsilon^{3/2}$$

Step 2bis: Slightly refined algorithm, adapted from [Heinrich'01, Montanaro'15]

$$\Delta/\epsilon$$

Application 1: counting the number of edges in a graph

Estimator X :=

1. Sample a vertex $v \in V$ uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If $\deg(v) < \deg(w)$ (or $\deg(v) = \deg(w)$ and $v <_{lex} w$)

   Output $\mathbf{n*deg(v)}$

   Else

   Output $\mathbf{0}$

λ(v,w)

Application 1: counting the number of edges in a graph

Estimator X :=

1. Sample a vertex $v \in V$ uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If $\deg(v) < \deg(w)$ (or $\deg(v) = \deg(w)$ and $v <_{lex} w$)

    Output $n*\deg(v)$

   Else

    Output 0

$\lambda(v,w)$

Lemma: $E(X) = m$ and $E(X^2)/E(X)^2 \leq O(\sqrt{n})$. (when $m \geq \Omega(n)$)

[Goldreich, Ron'08]
[Seshadhri'15]

Estimator X :=

1. Sample a vertex v ∈ V uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If deg(v) < deg(w) (or deg(v) = deg(w) and v $<_{lex}$ w)

   Output **n*deg(v)**

   Else

   Output **0**

λ(v,w)

Lemma: E(X) = m and $E(X^2)/E(X)^2 \leq O(\sqrt{n})$. (when m ≥ Ω(n))

[Goldreich, Ron'08]
[Seshadhri'15]

$$S_X|0\rangle = \sum_{v \in V} \sum_{w \in N(v)} \frac{1}{\sqrt{n \cdot \deg(v)}} |v\rangle |w\rangle |\lambda(v,w)\rangle$$

Application 1: counting the number of edges in a graph

Estimator X :=

1. Sample a vertex v ∈ V uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If deg(v) < deg(w)   (or deg(v) = deg(w) and v $<_{lex}$ w)

    Output **n*deg(v)**

  Else

    Output **0**

λ(v,w)

Lemma: E(X) = m and $E(X^2)/E(X)^2 \leq O(\sqrt{n})$.   (when m ≥ Ω(n))

[Goldreich, Ron'08]
[Seshadhri'15]

$$S_X|0\rangle = \sum_{v \in V} \sum_{w \in N(v)} \frac{1}{\sqrt{n \cdot \deg(v)}} |v\rangle |w\rangle |\lambda(v,w)\rangle$$

Result: $O(n^{1/4}/\varepsilon)$ quantum samples (= quantum queries) to approximate m.

(when m ≥ Ω(n))

Estimator X :=

1. Sample a vertex $v \in V$ uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If deg(v) < deg(w) (or deg(v) = deg(w) and $v <_{lex} w$)

 Output **n*deg(v)**

 Else **λ(v,w)**

 Output **0**

Lemma: E(X) = m and $E(X^2)/E(X)^2 \leq O(n/\sqrt{m})$, but we don't know $n/\sqrt{m}$…

[Goldreich, Ron'08]
[Seshadhri'15]

$$S_X|0\rangle = \sum_{v\in V}\sum_{w\in N(v)} \frac{1}{\sqrt{n\cdot \deg(v)}}|v\rangle|w\rangle|\lambda(v,w)\rangle$$

Application 1: counting the number of edges in a graph

Estimator X :=

1. Sample a vertex $v \in V$ uniformly at random
2. Sample a neighbor w of v uniformly at random
3. If deg(v) < deg(w) (or deg(v) = deg(w) and $v <_{lex} w$)

   Output **n*deg(v)**

   Else

   Output **0**

λ(v,w)

Lemma: $E(X) = m$ and $E(X^2)/E(X)^2 \leq O(n/\sqrt{m})$, but we don't know $n/\sqrt{m}$…
[Goldreich, Ron'08]
[Seshadhri'15]

$$S_X|0\rangle = \sum_{v \in V} \sum_{w \in N(v)} \frac{1}{\sqrt{n \cdot \deg(v)}} |v\rangle|w\rangle|\lambda(v,w)\rangle$$

Result: $\Theta(n^{1/2}/m^{1/4})$ quantum samples (= quantum queries) to approximate m.

Application 2: frequency moments in the streaming model

$$\mathbf{x} = \begin{array}{cccccccc} 1 & 2 & 3 & & & & & n \\ \boxed{0} & \boxed{0} & \boxed{0} & \boxed{0} & \cdots & \boxed{0} & \boxed{0} & \boxed{0} \end{array}$$

Stream of **updates** to x:

Application 2: frequency moments in the streaming model

	1	2	3					n
x =	**0**	**0**	**5**	**0**	…	**0**	**0**	**0**

Stream of **updates** to x: **(3,+5)**

Application 2: frequency moments in the streaming model

	1	2	3					n
x =	**0**	**-6**	**5**	**0**	...	**0**	**0**	**0**

Stream of **updates** to x: **(3,+5) ; (2,-6)**

Application 2: frequency moments in the streaming model

	1	2	3					n
x =	**0**	**-6**	**4**	**0**	…	**0**	**0**	**0**

Stream of **updates** to x: **(3,+5) ; (2,-6) ; (3,-1)**

Application 2: frequency moments in the streaming model

	1	2	3					n
x =	**0**	**-6**	**4**	**0**		**0**	**0**	**0**

Stream of **updates** to x: **(3,+5) ; (2,-6) ; (3,-1)**

Frequency moment of order k ≥ 3: $F_k = \sum_{i=1}^{n} |x_i|^k$

Application 2: frequency moments in the streaming model

$$\mathbf{x} = \begin{array}{cccccccc} 1 & 2 & 3 & & & & & n \\ \boxed{0} & \boxed{-6} & \boxed{4} & \boxed{0} & \cdots & \boxed{0} & \boxed{0} & \boxed{0} \end{array}$$

Stream of **updates** to x: **(3,+5) ; (2,-6) ; (3,-1)**

Frequency moment of order k ≥ 3: $F_k = \sum_{i=1}^{n} |x_i|^k$

Best **P-pass** algorithm with **memory M** approximating F_k?

Application 2: frequency moments in the streaming model

	1	2	3					n
x =	0	-6	4	0	...	0	0	0

Stream of **updates** to x: **(3,+5) ; (2,-6) ; (3,-1)**

Frequency moment of order k ≥ 3: $F_k = \sum_{i=1}^{n} |x_i|^k$

Best **P-pass** algorithm with **memory M** approximating F_k?

Classically: PM = $\Theta(n^{1-2/k})$

1 pass + memory $\mathbf{M = \frac{n^{1-2/k}}{P}}$

||

1 sample from a random variable X with

$\mathbf{E(X) \approx F_k}$ and $\mathbf{E(X^2)/E(X)^2 \leq P \cdot F_k^2}$

[Monemizadeh, Woodruff'10]
[Andoni, Krauthgamer, Onak'11]

Application 2: frequency moments in the streaming model

	1	2	3		...			n
x =	0	-6	4	0	...	0	0	0

Stream of **updates** to x: **(3,+5) ; (2,-6) ; (3,-1)**

Frequency moment of order k ≥ 3: $F_k = \sum_{i=1}^{n} |x_i|^k$

Best **P-pass** algorithm with **memory M** approximating F_k?

Classically: $PM = \Theta(n^{1-2/k})$

1 pass + memory $\mathbf{M = \frac{n^{1-2/k}}{P}}$

||

1 sample from a random variable X with

$\mathbf{E(X) \approx F_k}$ and $\mathbf{E(X^2)/E(X)^2 \leq P \cdot F_k^2}$

[Monemizadeh, Woodruff'10]
[Andoni, Krauthgamer, Onak'11]

Quantumly: $P^2M = O(n^{1-2/k})$

1 pass + memory $\mathbf{M = \frac{n^{1-2/k}}{P^2}}$

||

1 **quantum** sample* S_X from a r.v. X with

$\mathbf{E(X) \approx F_k}$ and $\mathbf{E(X^2)/E(X)^2 \leq (P \cdot F_k)^2}$

* $\mathbf{S_X^{-1}}$ can be done in one pass also

Application 3: counting the number of triangles in a graph

More complicated than edges… [Eden, Levi, Ron'15] [Eden, Levi, Ron, Seshadhri'17]

Main subroutine: estimator X for the number of triangles adjacent to any vertex v

Application 3: counting the number of triangles in a graph

More complicated than edges… [Eden, Levi, Ron'15] [Eden, Levi, Ron, Seshadhri'17]

Main subroutine: estimator X for the number of triangles adjacent to any vertex v

1 classical sample = O(1) queries **in expectation**

but O($\sqrt{m}$) in the **worst case**

Application 3: counting the number of triangles in a graph

More complicated than edges… [Eden, Levi, Ron'15] [Eden, Levi, Ron, Seshadhri'17]

Main subroutine: estimator X for the number of triangles adjacent to any vertex v

1 classical sample = O(1) queries **in expectation**
but $O(\sqrt{m})$ in the **worst case**

Variable-time Amplitude Estimation: estimate the amplitude when some "branches" of the computation stop earlier than the others

Application 3: counting the number of triangles in a graph

More complicated than edges… [Eden, Levi, Ron'15] [Eden, Levi, Ron, Seshadhri'17]

Main subroutine: estimator X for the number of triangles adjacent to any vertex v

1 classical sample = O(1) queries **in expectation**

but O(√m) in the **worst case**

Variable-time Amplitude Estimation: estimate the amplitude when some "branches" of the computation stop earlier than the others

Result:

$$\widetilde{\Theta}\left(\frac{\sqrt{n}}{t^{1/6}} + \frac{m^{3/4}}{\sqrt{t}}\right)$$ quantum queries for triangle counting

vs. $$\widetilde{\Theta}\left(\frac{n}{t^{1/3}} + \frac{m^{3/2}}{t}\right)$$ classical queries