Quantum Chebyshev's Inequality and Applications

Yassine Hamoudi, Frédéric Magniez

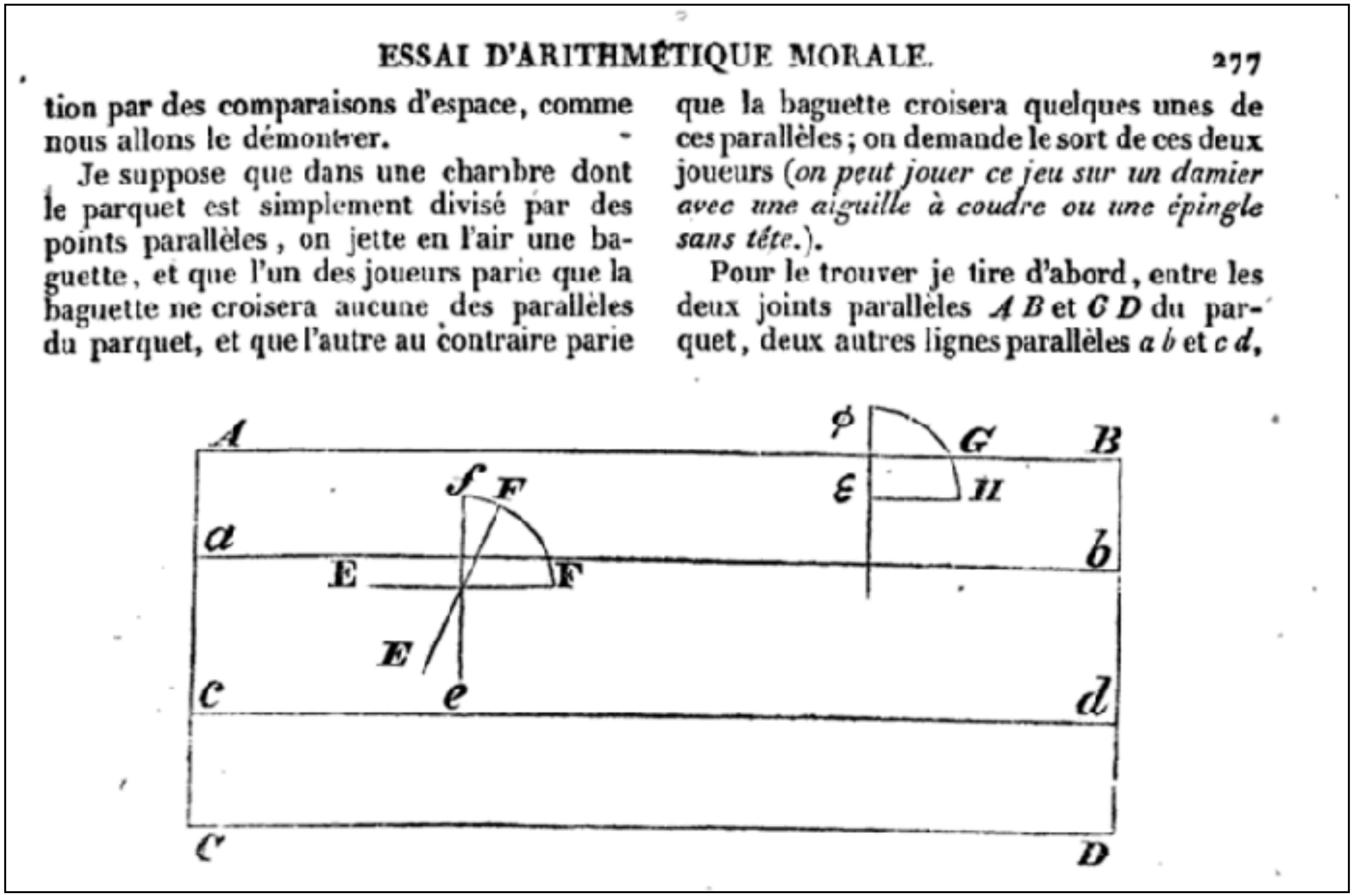
IRIF, Université Paris Diderot, CNRS

QuantAlgo 2018

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 **$2/\pi$**.



ESSAI D'ARITHMÉTIQUE MORALE. 277

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: $\dfrac{x_1 + \ldots + x_n}{n} \xrightarrow{n \to \infty} \mathbf{E}(X)$

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:

Hypothesis: $\mathbf{E}(X) \neq 0$ and $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 \neq 0$ finite

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

multiplicative error 0 < ε < 1

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

How fast does it converge to E(X) ?

Chebyshev's Inequality:

Hypothesis: $\mathbf{E}(X) \neq 0$ and $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 \neq 0$ finite

Objective: $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability

multiplicative error $0 < \varepsilon < 1$

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

How fast does it converge to E(X) ?

Chebyshev's Inequality:

Hypothesis: $\mathbf{E}(X) \neq 0$ and $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 \neq 0$ finite

Objective: $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability

multiplicative error $0 < \varepsilon < 1$

Relative second moment

Number of samples needed: $O\left(\frac{\mathbf{E}(X^2)}{\epsilon^2 \mathbf{E}(X)^2}\right)$ (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: $\tilde{\mu} = \dfrac{x_1 + \ldots + x_n}{n}$ with $x_1, \ldots, x_n \sim X$

How fast does it converge to E(X) ?

Chebyshev's Inequality:

Hypothesis: $\mathbf{E}(X) \neq 0$ and $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 \neq 0$ finite

Objective: $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with high probability

multiplicative error 0 < ε < 1

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 $\Omega \subset R^+$

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 garbage state

($\sqrt{p_x}$ can be replaced with any α_x such that $|\alpha_x|^2 = p_x$)

Can we use quadratically less samples in the quantum setting?

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq \mathrm{Var}(X)$, **$\sigma/\varepsilon$** quantum samples vs **$\sigma^2/\varepsilon^2$** classical samples

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq \mathrm{Var}(X)$, **$\sigma/\varepsilon$** quantum samples vs **$\sigma^2/\varepsilon^2$** classical samples

??? for multiplicative error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq \mathrm{Var}(X)$, **$\sigma/\varepsilon$** quantum samples vs **$\sigma^2/\varepsilon^2$** classical samples

??? for multiplicative error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$

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

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq \mathrm{Var}(X)$, **$\sigma/\varepsilon$** quantum samples vs **$\sigma^2/\varepsilon^2$** classical samples

??? for multiplicative error approximation $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$

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

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq \mathrm{Var}(X)$, **$\sigma/\varepsilon$** quantum samples vs **$\sigma^2/\varepsilon^2$** classical samples

??? for multiplicative error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

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

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq Var(X)$, **σ/ε** quantum samples vs **σ^2/ε^2** classical samples

??? for multiplicative error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$

	Number of samples	**Conditions**
Classical samples (Chebyshev's inequality)	Δ^2/ε^2	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\sqrt{\mathbf{B}}/\left(\epsilon\sqrt{\mathbf{E}(X)}\right)$	Sample space **$\Omega \subset [0,B]$**
[Montanaro'15] [Li, Wu'17]	$\Delta^{\mathbf{2}}/\varepsilon$ or $(\Delta/\varepsilon)^*(H/L)$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ **L** $\leq E(X) \leq$ **H**

Can we use quadratically less samples in the quantum setting?

Yes! for additive error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon$

[Montanaro'15] Given $\sigma^2 \geq Var(X)$, **σ/ε** quantum samples vs **σ^2/ε^2** classical samples

??? for multiplicative error approximation $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

	Number of samples	Conditions	
Classical samples (Chebyshev's inequality)	Δ^2/ε^2	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	$\sqrt{\mathbf{B}}/\left(\epsilon\sqrt{\mathbf{E}(X)}\right)$	Sample space **$\Omega \subset [0,B]$**	
[Montanaro'15] [Li, Wu'17]	$\Delta^{\mathbf{2}}/\varepsilon$ or $(\Delta/\varepsilon)^*(H/L)$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	**L** $\leq E(X) \leq$ **H**
Our result	$(\Delta/\varepsilon)^*$**log³**$(H/E(X))$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	$E(X) \leq$ **H**

Our Approach

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]:

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

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]:

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

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 $\tilde{\theta}$ of $\pm\theta$.

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

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 $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

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

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 $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

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

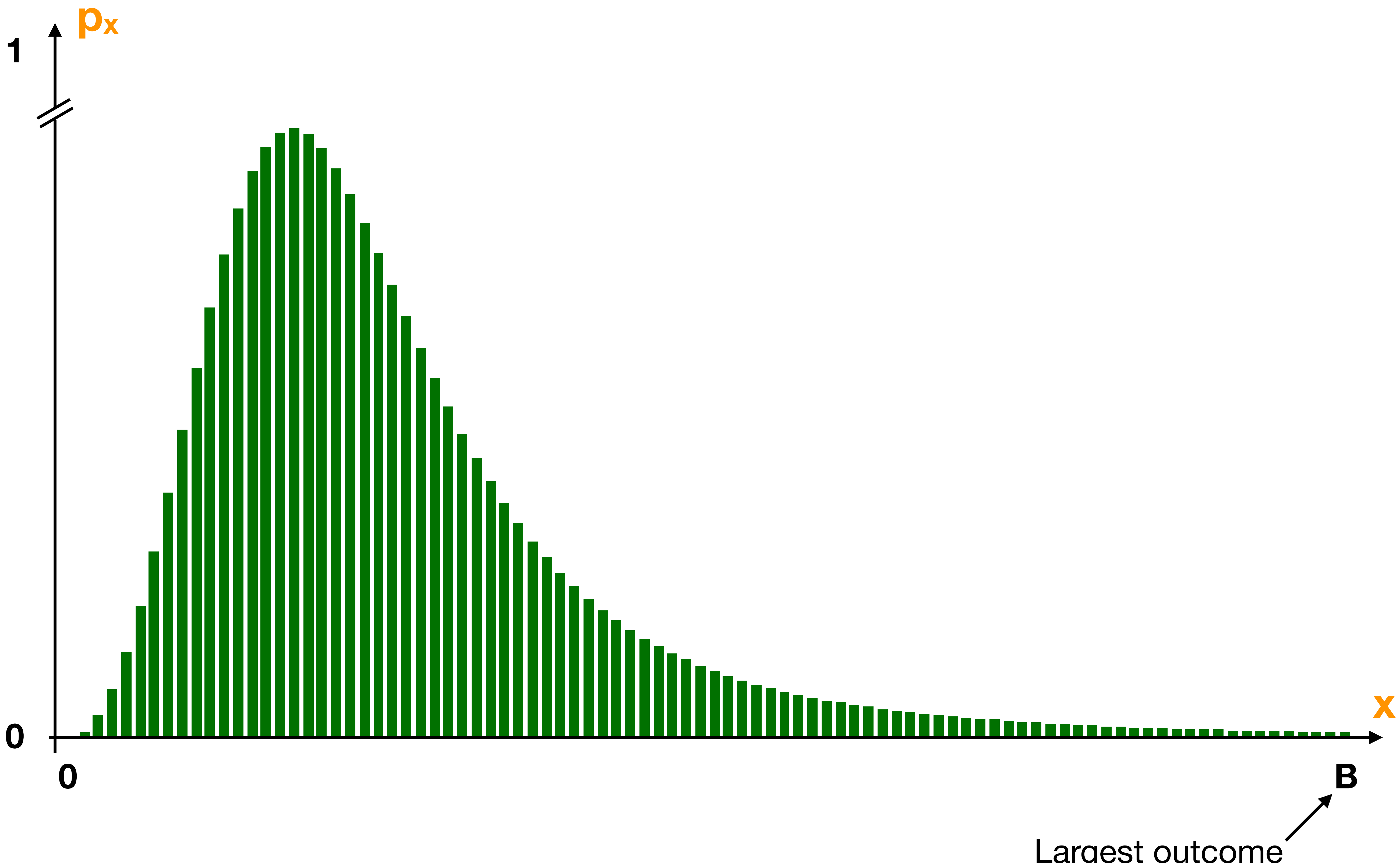
Sampler: $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ on sample space **Ω ⊂ [0,B]**

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

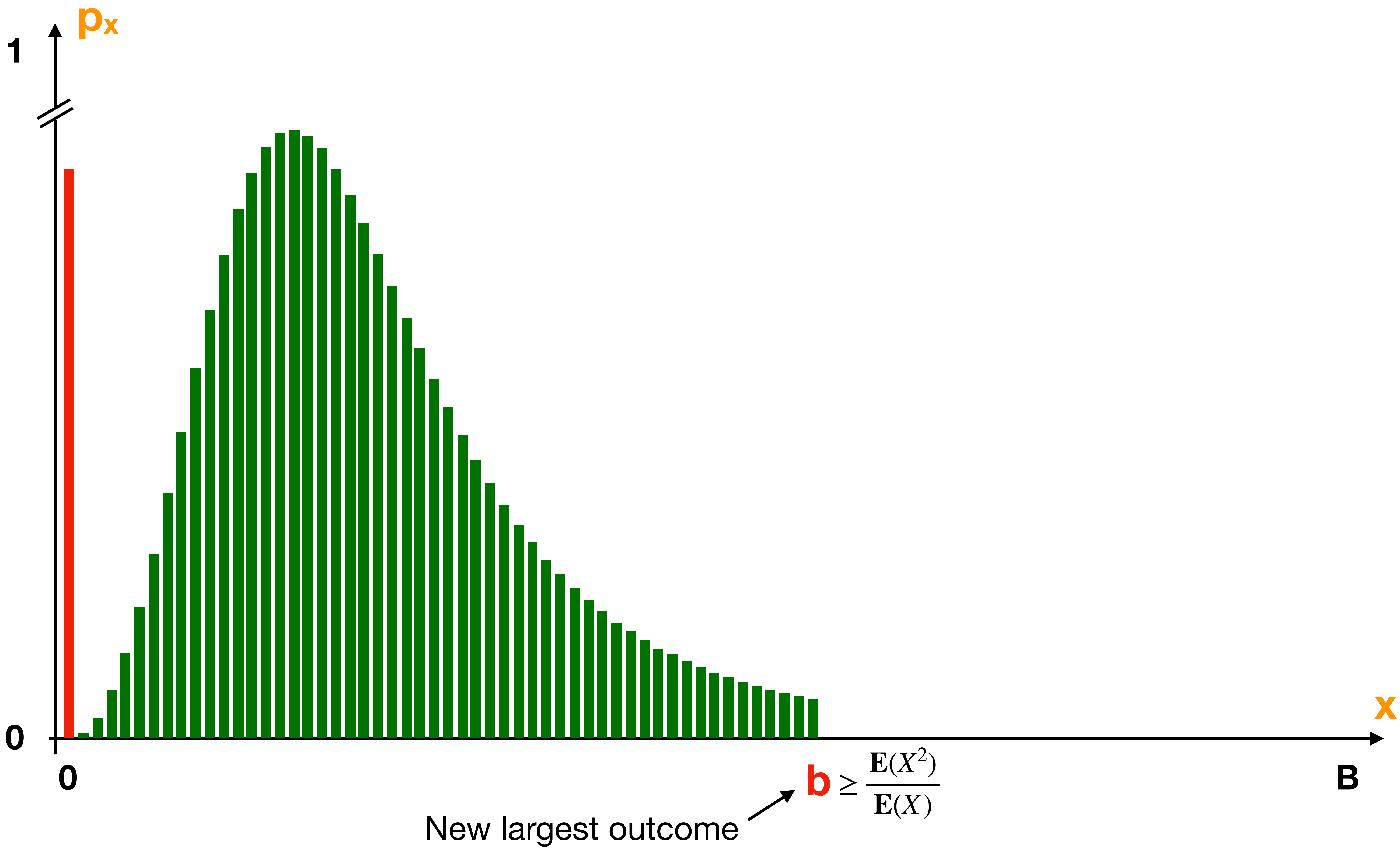
If $B \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon\mathbf{E}(X)}\right)$ ✓

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 **?**

Random variable X



Random variable $X_{<b}$



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 $|\tilde{\mu} - \mathbf{E}(X)| \le \epsilon\mathbf{E}(X)$

If $B \le \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: 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

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 $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon\mathbf{E}(X)$

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 $(1-\epsilon)\mathbf{E}(X) \leq \mathbf{E}(X_{<b}) \leq \mathbf{E}(X)$.

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 $|\tilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$

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	Estimated value	Number of samples	Estimation
$b_0 = H\Delta^2$	$\frac{\mathbf{E}(X_{<b_0})}{b_0}$	Δ	$\widetilde{\mu}_0$
$b_1 = (H/2)\Delta^2$	$\frac{\mathbf{E}(X_{<b_1})}{b_1}$	Δ	$\widetilde{\mu}_1$
$b_2 = (H/4)\Delta^2$	$\frac{\mathbf{E}(X_{<b_2})}{b_2}$	Δ	$\widetilde{\mu}_2$
...			...

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

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	Estimated value	Number of samples	Estimation
$b_0 = H\Delta^2$	$\frac{\mathbf{E}(X_{<b_0})}{b_0}$	Δ	$\widetilde{\mu}_0$
$b_1 = (H/2)\Delta^2$	$\frac{\mathbf{E}(X_{<b_1})}{b_1}$	Δ	$\widetilde{\mu}_1$
$b_2 = (H/4)\Delta^2$	$\frac{\mathbf{E}(X_{<b_2})}{b_2}$	Δ	$\widetilde{\mu}_2$
…			…

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

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

$$2 \cdot \mathbf{E}(X)\Delta^2 \leq b_i \leq 10^4 \cdot \mathbf{E}(X)\Delta^2$$

Analysis

- If $b_i \approx \mathbf{E}(X) \cdot \Delta^2$ → $\frac{\mathbf{E}(X_{<b_i})}{b_i} \overset{\star}{\approx} \frac{\mathbf{E}(X)}{b_i} \approx \frac{1}{\Delta^2}$ → Δ samples are enough

Analysis

- If $b_i \approx \mathbf{E}(X) \cdot \Delta^2$ → $\dfrac{\mathbf{E}(X_{<b_i})}{b_i} \overset{\star}{\approx} \dfrac{\mathbf{E}(X)}{b_i} \approx \dfrac{1}{\Delta^2}$ → Δ samples are enough

- If b_i is **<u>very large</u>** ? → $\dfrac{\mathbf{E}(X_{<b_i})}{b_i}$ is very small → Δ samples is not enough to distinguish $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ from 0

Analysis

- If $b_i \approx \mathbf{E}(X) \cdot \Delta^2$ → $\frac{\mathbf{E}(X_{<b_i})}{b_i} \overset{\bigstar}{\approx} \frac{\mathbf{E}(X)}{b_i} \approx \frac{1}{\Delta^2}$ → Δ samples are enough

- If b_i is **very large** → $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ is very small → Δ samples is not enough to distinguish $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ from 0

  ?

[Brassard et al.'02]

The output of the **Amplitude-Estimation** algorithm is 0 w.h.p. when the **estimated value** is below the inverse-square of the **number** of samples

↖ $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ ↖ Δ

Analysis

- If $b_i \approx \mathbf{E}(X) \cdot \Delta^2$ → $\frac{\mathbf{E}(X_{<b_i})}{b_i} \overset{\star}{\approx} \frac{\mathbf{E}(X)}{b_i} \approx \frac{1}{\Delta^2}$ → Δ samples are enough

- If b_i is **very large** → $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ is very small → Δ samples is not enough to distinguish $\frac{\mathbf{E}(X_{<b_i})}{b_i}$ from 0

  ?

[Brassard et al.'02]

The output of the **Amplitude-Estimation** algorithm is 0 w.h.p. when the **estimated value** is below the inverse-square of the **number** of samples

$\frac{\mathbf{E}(X_{<b_i})}{b_i}$ ↖ (estimated value) Δ ↖ (number)

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

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

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

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

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

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]

$$\Delta/\epsilon$$

Applications

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)

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

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

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 **λ(v,w)**

    Output **0**

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]

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

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))

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   **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 ∈ 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)

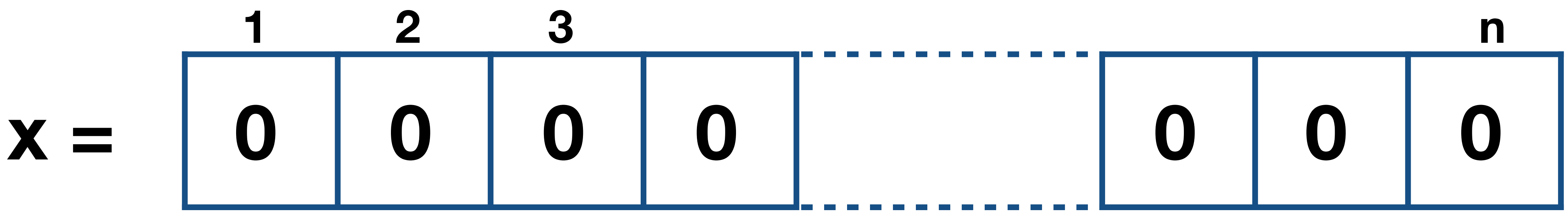
Lemma: E(X) = m and $E(X^2)/E(X)^2 \leq O(n/\sqrt{m})$, but we don't know $n/\sqrt{m}$…

[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.

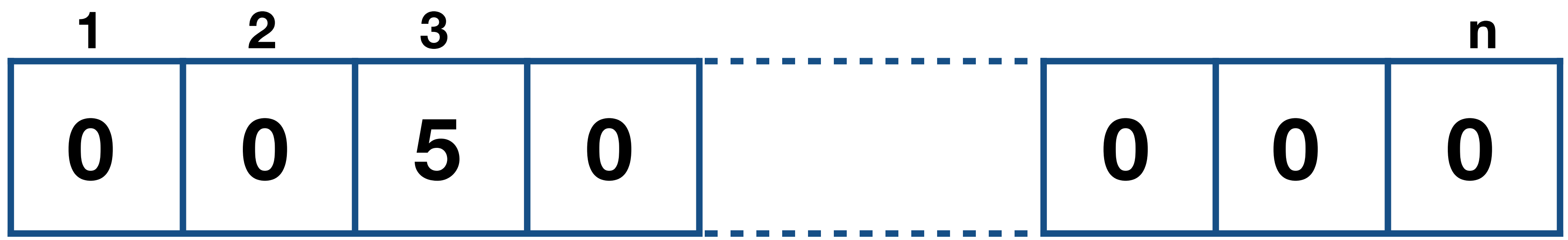
Application 2: frequency moments in the streaming model



Stream of **updates** to x:

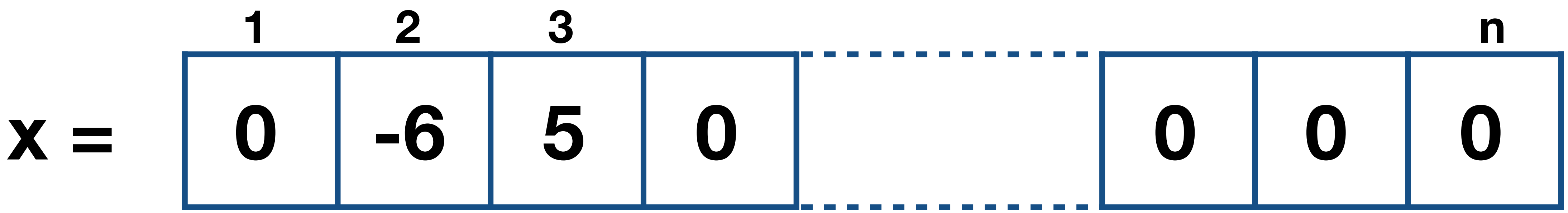
Application 2: frequency moments in the streaming model

x =



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

Application 2: frequency moments in the streaming model



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)**

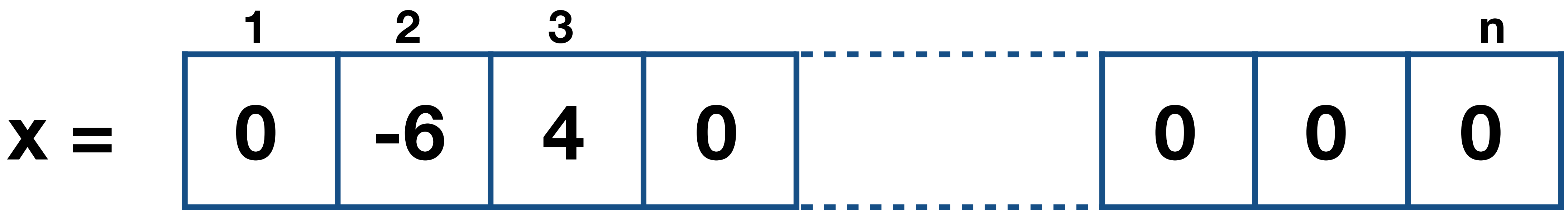
Application 2: frequency moments in the streaming model

	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



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 **space 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 **space 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 **space memory M** approximating F_k?

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

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

||

1 sample from a random variable X with

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

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

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

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

||

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

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

* 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

Conclusion

The **mean** of any quantum sampler S_X is 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 $H \geq \mathbf{E}(X)$.

[Nayak, Wu'99] : corresponding **lower bound**

Applications:

- Frequency moments: $\mathbf{P^2M = \widetilde{O}(n^{1-2/k})}$

Lower bound: **?**

- Edge counting: $\widetilde{\Theta}\left(\frac{\sqrt{n}}{m^{1/4}}\right)$
- Triangle counting: $\widetilde{\Theta}\left(\frac{\sqrt{n}}{t^{1/6}} + \frac{m^{3/4}}{\sqrt{t}}\right)$

Lower bounds with a property testing to communication complexity reduction method (reduction to Disjointness)

[Blais et al'12] [Eden, Rosenbaum'17]

arXiv: 1807.06456

Extra slides

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

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

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

quantum samples, and no a priori information on X.

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

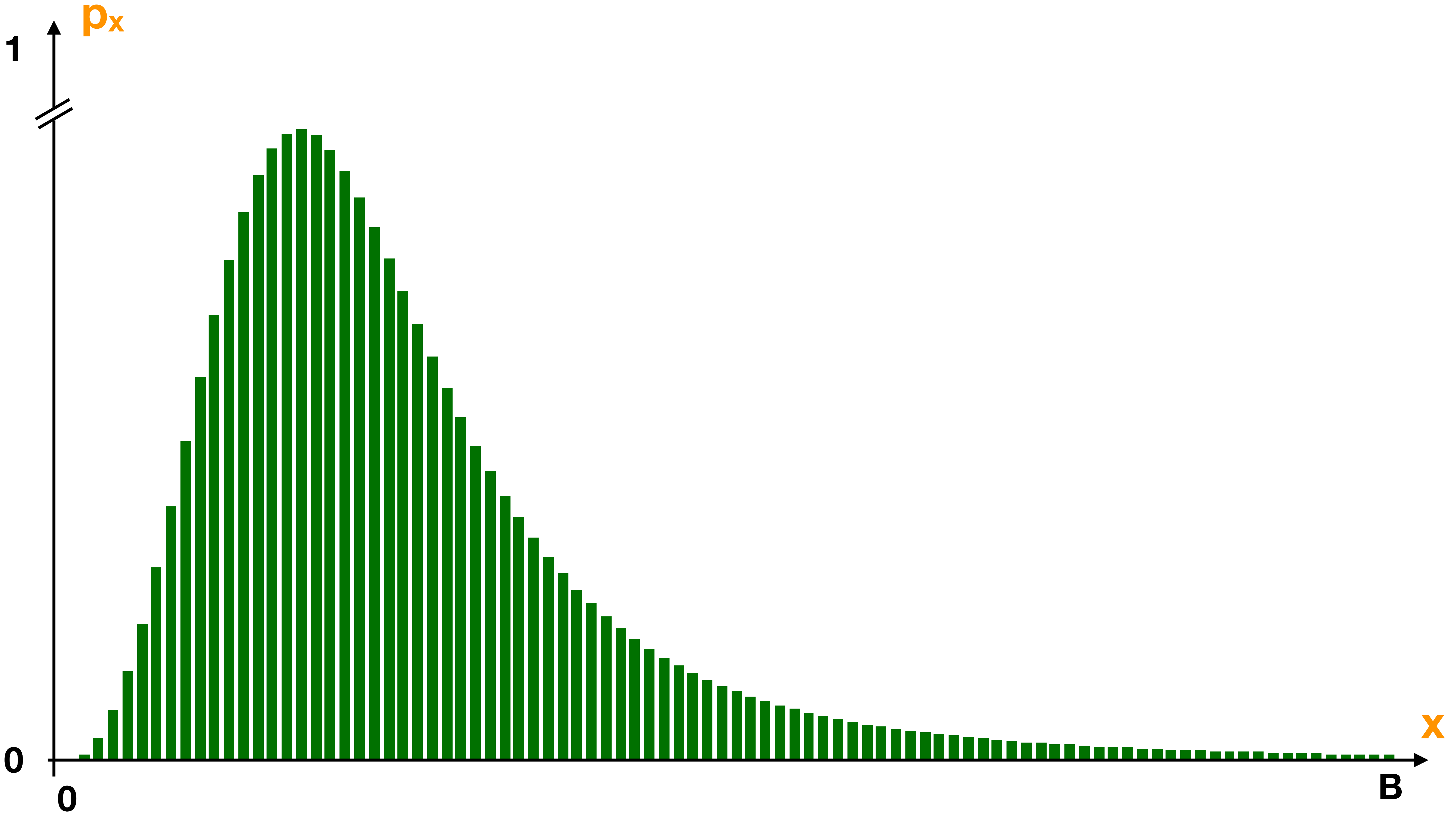
Proof:

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

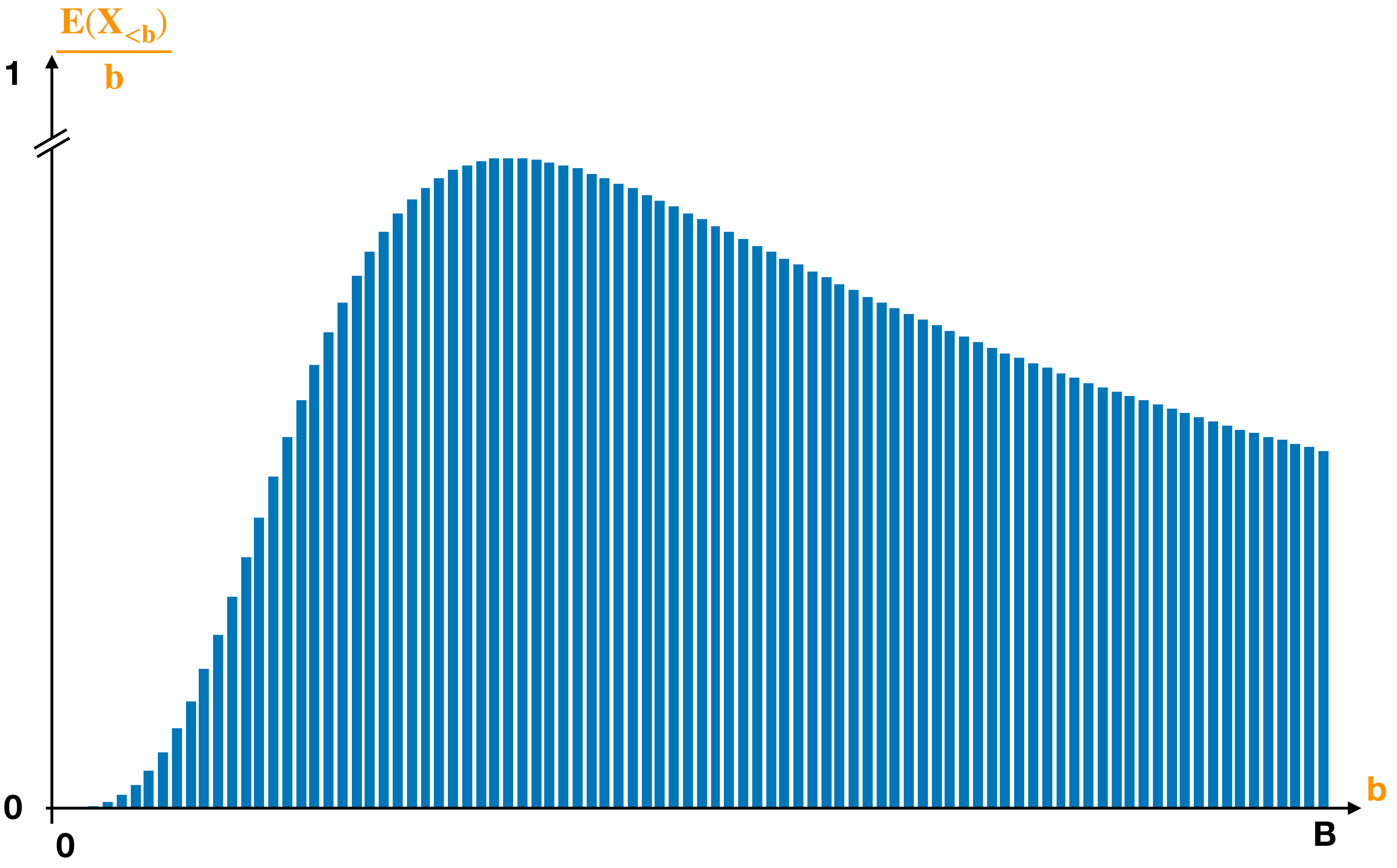
Lemma: If $b \geq 10^4 \cdot \mathbf{E}(X)\Delta^2$ then $\frac{\mathbf{E}(X_{<b})}{b} \leq \frac{1}{10^4 \cdot \Delta^2}$

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

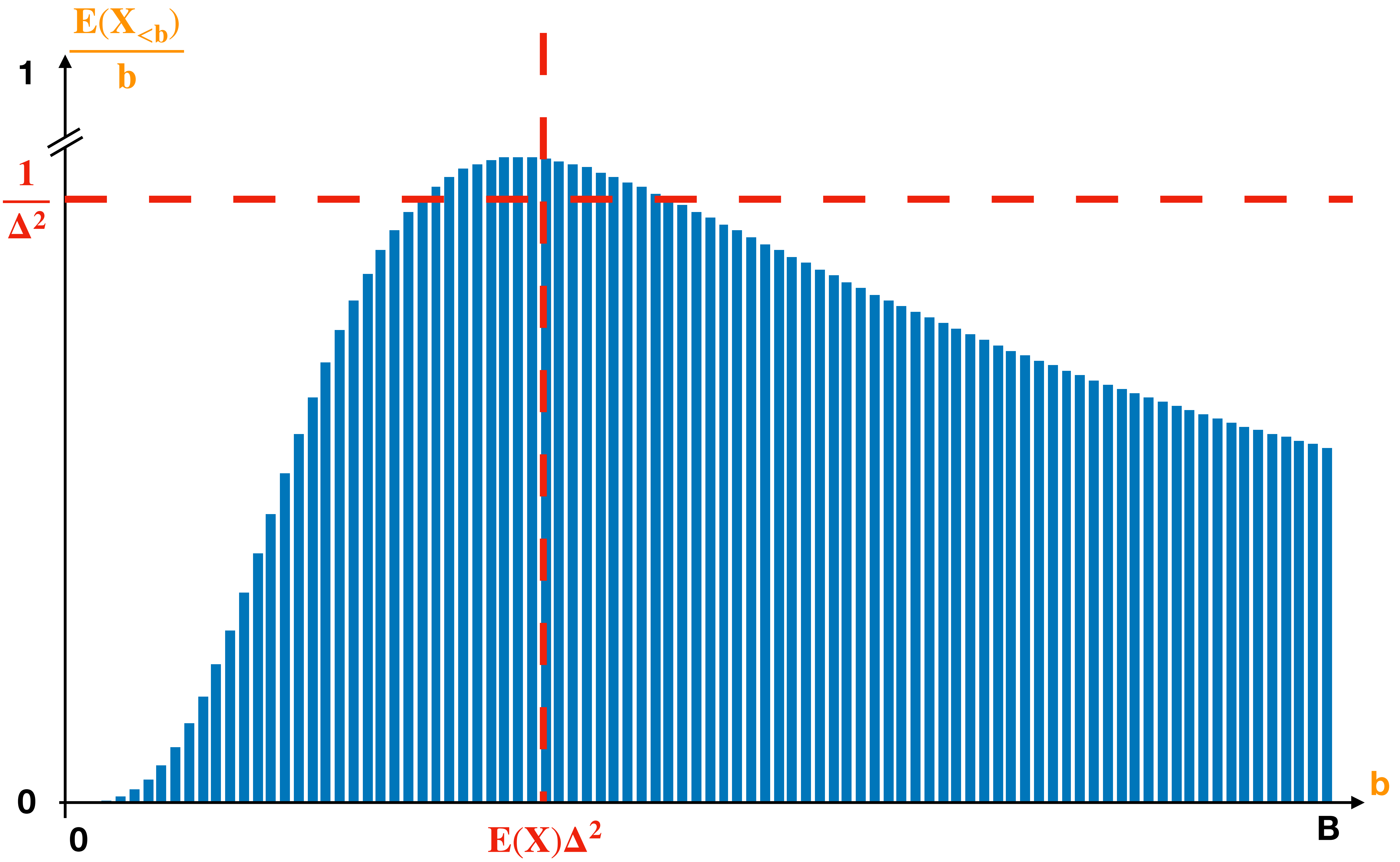
Example



Example



Example



Definition: An algorithm solves the **Mean Estimation problem** for parameters ε,Δ if, for any sampler S_X satisfying $E(X^2)/E(X)^2 \in [\Delta, 2\Delta]$, it outputs a value $\widetilde{\mu}$ satisfying $|\widetilde{\mu} - \mathbf{E}(X)| \leq \epsilon \mathbf{E}(X)$ with probability 2/3.

[Nayak, Wu'99] Any algorithm solving the **Mean Estimation problem** for parameters $0 < \varepsilon < 1/6$, $\Delta > 1$ on the sample space $\Omega = \{0,1\}$ must use **$\Omega((\Delta-1)/\varepsilon)$** quantum samples.