Quantum Chebyshev's Inequality and Applications

Yassine Hamoudi, Frédéric Magniez

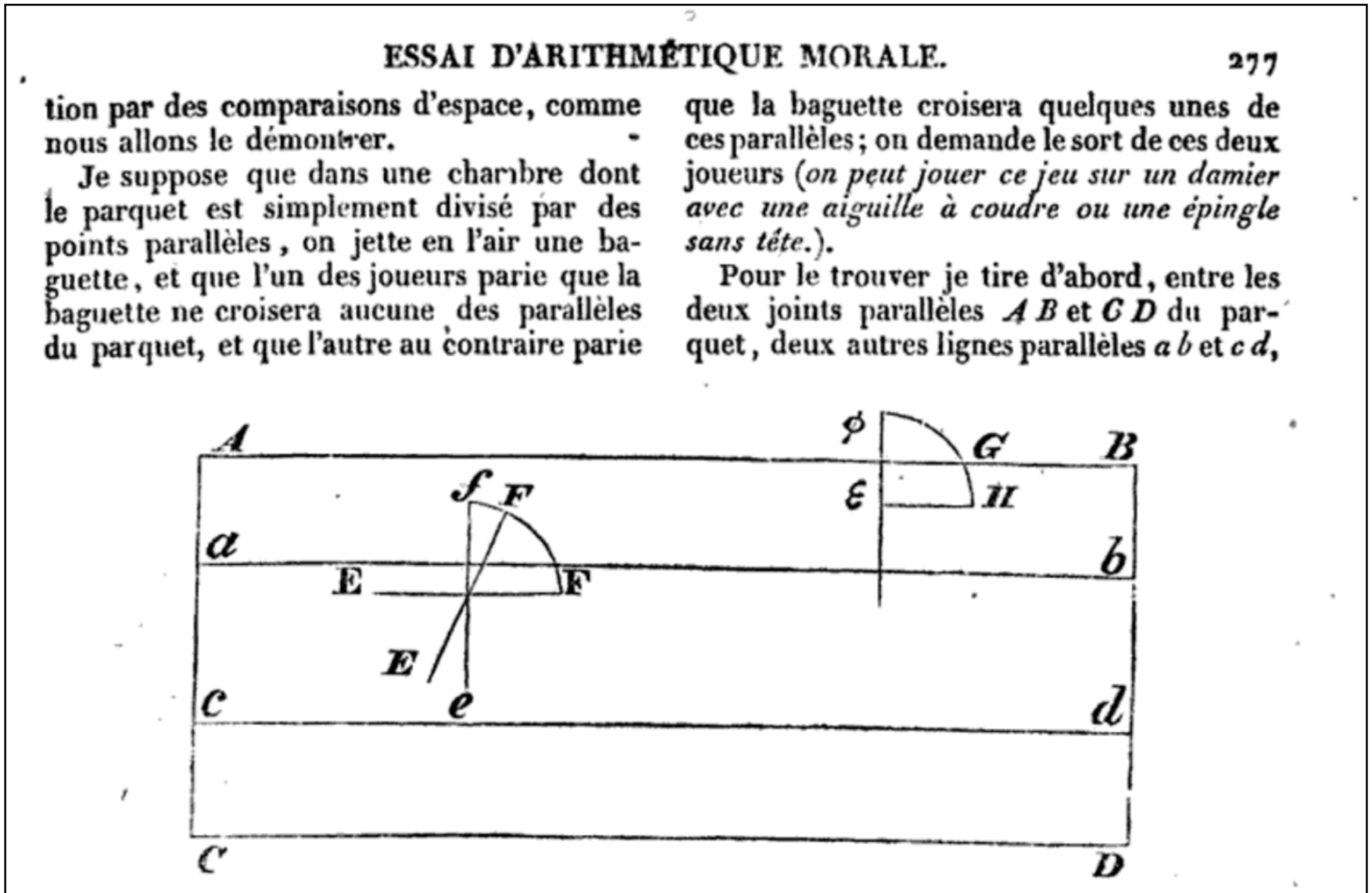
IRIF, Université Paris Diderot, CNRS

CQT 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 **$2/\pi$**.



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:

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 $O\left(\dfrac{\mathbf{Var}(X)}{\epsilon^2 \mathbf{E}(X)^2}\right) = O\left(\dfrac{1}{\epsilon^2}\left(\dfrac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2} - 1\right)\right)$)

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

Problem: approximate the number **m** of edges in an **n**-vertex graph G

Problem: approximate the number **m** of edges in an **n**-vertex graph G

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**

Problem: approximate the number **m** of edges in an **n**-vertex graph G

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**

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]

Problem: approximate the number **m** of edges in an **n**-vertex graph G

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**

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]

Consequence: $O(\sqrt{n}/\varepsilon^2)$ samples to approximate m with error ε.

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 $\Omega \subset \mathbb{R}^+$

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 unit vector

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

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

Question: **can we estimate E(X) with less samples in the quantum setting?**

Previous Works

The Amplitude Estimation algorithm [Brassard et al.'11] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

The Amplitude Estimation algorithm [Brassard et al.'11] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega}\sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega}\sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}\,|0\rangle + \sqrt{\frac{x}{M_\Omega}}\,|1\rangle\right)$$

where $\mathbf{M}_\Omega = \max\{x \in \Omega\}$

The Amplitude Estimation algorithm [Brassard et al.'11] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}\,|0\rangle + \sqrt{\frac{x}{M_\Omega}}\,|1\rangle\right)$$

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

where $\mathbf{M_\Omega} = \max\{x \in \Omega\}$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are some unit vectors.

The Amplitude Estimation algorithm [Brassard et al.'11] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}\,|0\rangle + \sqrt{\frac{x}{M_\Omega}}\,|1\rangle\right)$$

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

where $\mathbf{M_\Omega} = \max\{x \in \Omega\}$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are some unit vectors.

Observation: 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 $\theta = \sin^{-1}(\sqrt{\mathbf{E}(X)/M_\Omega})$.

The Amplitude Estimation algorithm [Brassard et al.'11] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}\,|0\rangle + \sqrt{\frac{x}{M_\Omega}}\,|1\rangle\right)$$

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

where $\mathbf{M_\Omega} = \max\{x \in \Omega\}$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are some unit vectors.

Observation: 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 $\theta = \sin^{-1}(\sqrt{\mathbf{E}(X)/M_\Omega})$.

Algorithm: 1/ Apply Phase Estimation on G for $t \geq \Omega(\sqrt{M_\Omega}/(\epsilon\sqrt{\mathbf{E}(X)}))$ steps

to get an estimate $\widetilde{\theta}$ s.t. $|\widetilde{\theta} - |\theta|| \leq 1/t$.

2/ Output $\widetilde{\mu} = M_\Omega \cdot \sin^2(\widetilde{\theta})$ as an estimate to E(X).

The Amplitude Estimation algorithm [Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega}\sqrt{p_x}\,|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega}\sqrt{p_x}\,|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}\,|0\rangle + \sqrt{\frac{x}{M_\Omega}}\,|1\rangle\right)$$

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

where $\mathbf{M}_\Omega = \max\{x\in\Omega\}$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are some unit vectors.

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

The Amplitude Estimation algorithm [Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09]

Given $S_X|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}|\psi_x\rangle|x\rangle$ one can obtain (with 1 ancillary qubit + controlled rotation):

$$S_Y|0\rangle = \sum_{x\in\Omega} \sqrt{p_x}|\psi_x\rangle|x\rangle\left(\sqrt{1-\frac{x}{M_\Omega}}|0\rangle + \sqrt{\frac{x}{M_\Omega}}|1\rangle\right)$$

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

where $\mathbf{M_\Omega} = \max\{x \in \Omega\}$ and $|\varphi_0\rangle, |\varphi_1\rangle$ are some unit vectors.

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

…not efficient if $\mathbf{M_\Omega}$ is large (worst than the classical algorithm sometimes)

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)	$\frac{\Delta^2}{\epsilon^2}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$
[Brassard et al.'02] [Brassard et al.'11] [Wocjan et al.'09]	$\frac{\sqrt{\mathbf{M}_\Omega}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space **$\Omega \subset [0, M_\Omega]$**

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.'02] [Brassard et al.'11] [Wocjan et al.'09]	$\frac{\sqrt{\mathbf{M}_\Omega}}{\epsilon\sqrt{\mathbf{E}(X)}}$	Sample space $\Omega \subset [0, M_\Omega]$
[Montanaro'15]	$\frac{\Delta^2}{\epsilon}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$

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.'02] [Brassard et al.'11] [Wocjan et al.'09]	$\frac{\sqrt{\mathbf{M}_\Omega}}{\epsilon\sqrt{\mathbf{E}(X)}}$		Sample space **$\Omega \subset [0, M_\Omega]$**
[Montanaro'15]	$\frac{\Delta^{\mathbf{2}}}{\epsilon}$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	
Our result	$\frac{\Delta}{\epsilon} \cdot \log^{\mathbf{3}}\left(\frac{M_\Omega}{\mathbf{E}(X)}\right)$	$\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$	Sample space **$\Omega \subset [0, M_\Omega]$**

Our Approach

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

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

Ampl-Est: $O\left(\frac{\sqrt{\mathbf{M}_\Omega}}{\epsilon\sqrt{\mathbf{E}(X)}}\right)$ 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]

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

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

If $M_\Omega \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: 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]

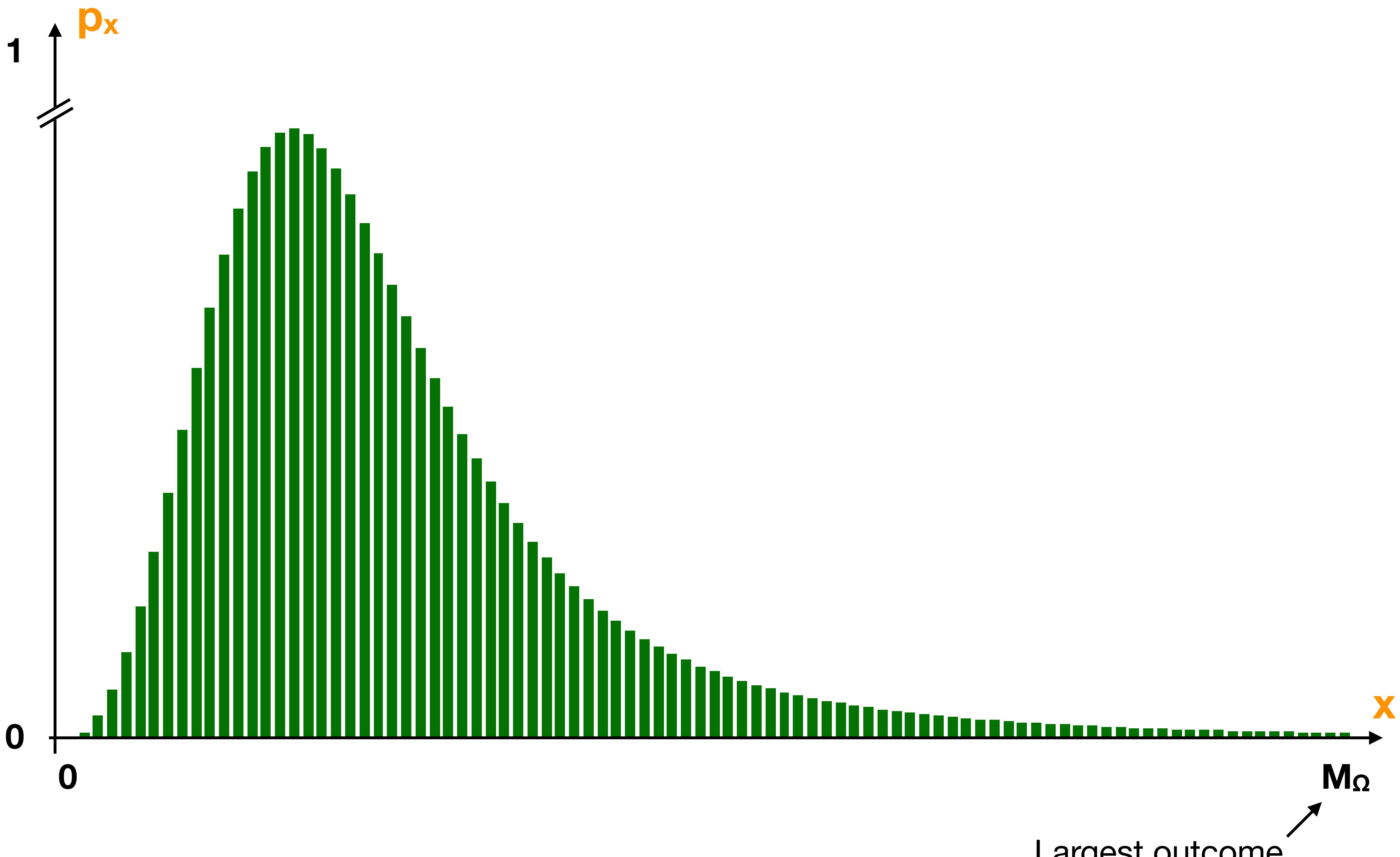
Input: Random variable **X** on sample space $\Omega \subset [0, M_\Omega]$

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

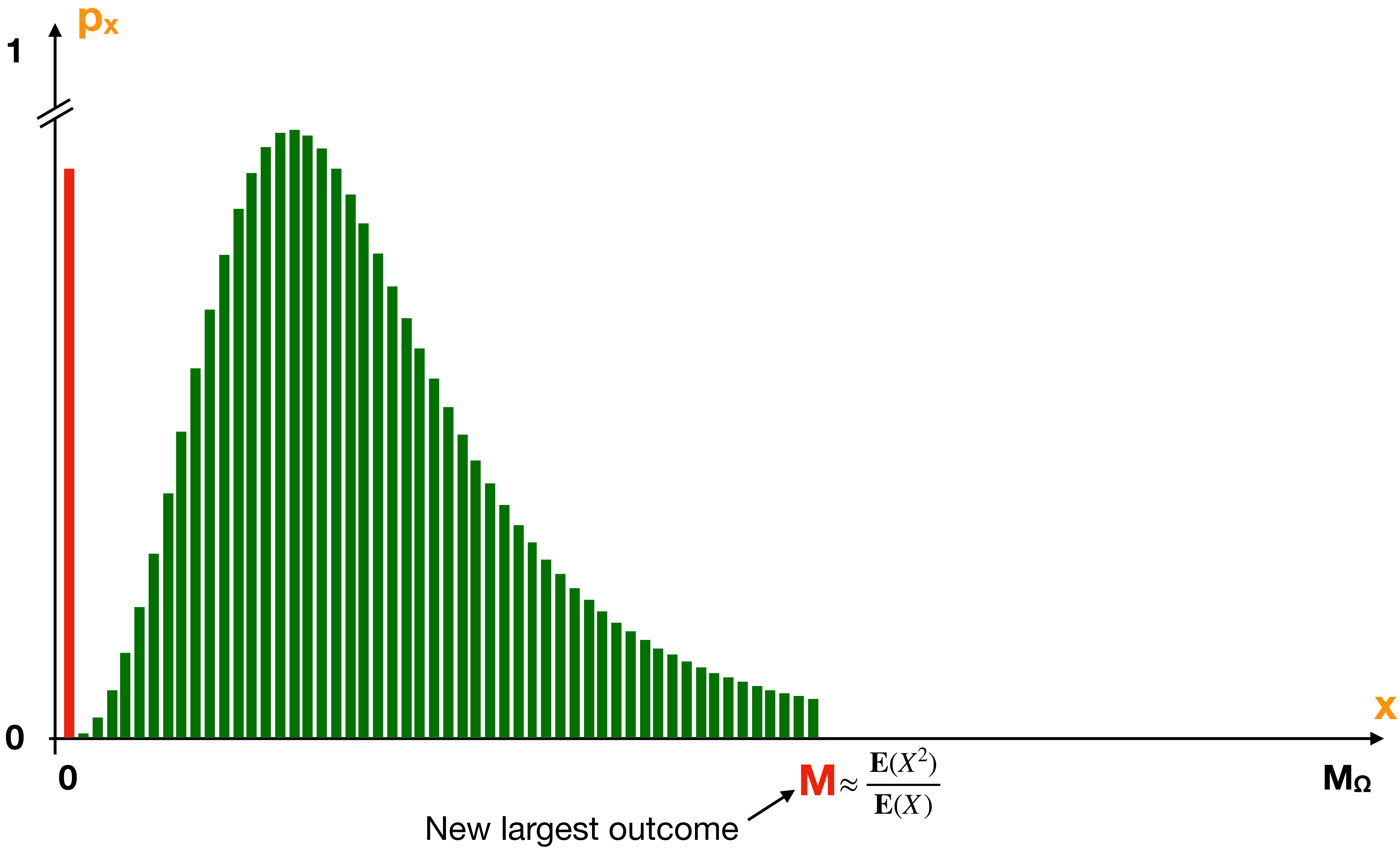
If $M_\Omega \leq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$: the number of samples is $O\left(\frac{\sqrt{\mathbf{E}(X^2)}}{\epsilon \mathbf{E}(X)}\right)$ ✓

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

Random variable X



Random variable X_M



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

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

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

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

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

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

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

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

Lemma: If $M \ge \frac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \le \mathbf{E}(X_M) \le \mathbf{E}(X)$.

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

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

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

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

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

Problem: $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ is unknown…

Input: Random variable **X** on sample space $\mathbf{\Omega} \subset \mathbf{[0,M_\Omega]}$

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

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

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

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

Problem: $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ is unknown… but we have $\Delta^2 \geq \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^{2}}$

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

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

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

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

★ **Lemma:** If $M \ge \frac{\mathbf{E}(X^2)}{\epsilon \mathbf{E}(X)}$ then $(1-\epsilon)\mathbf{E}(X) \le \mathbf{E}(X_M) \le \mathbf{E}(X)$.

Problem: $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)}$ is unknown… but we have $\Delta^2 \ge \frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}$ $\longrightarrow$ $M \approx \mathbf{E}(X) \cdot \Delta^2$?

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

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

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on M

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

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on M

Threshold	Input r.v.	Number of samples	Estimation
$M_0 = M_\Omega \Delta^2$	X_{M_0}	Δ	$\widetilde{\mu}_0$
$M_1 = (M_\Omega/2)\Delta^2$	X_{M_1}	Δ	$\widetilde{\mu}_1$
$M_2 = (M_\Omega/4)\Delta^2$	X_{M_2}	Δ	$\widetilde{\mu}_2$
$\cdots$			$\cdots$

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

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

Solution: use the **Amplitude Estimation** algorithm to do a logarithmic search on M

Threshold	Input r.v.	Number of samples	Estimation
$M_0 = M_\Omega \Delta^2$	X_{M_0}	Δ	$\widetilde{\mu}_0$
$M_1 = (M_\Omega/2)\Delta^2$	X_{M_1}	Δ	$\widetilde{\mu}_1$
$M_2 = (M_\Omega/4)\Delta^2$	X_{M_2}	Δ	$\widetilde{\mu}_2$
…			…

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

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

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

Analysis

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

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

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

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

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

[Brassard et al.'02]

$\sqrt{\frac{\mathbf{E}(X_M)}{M}}$

$1/\Delta$

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

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

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

$$\sqrt{\frac{\mathbf{E}(X_M)}{M}} \qquad 1/\Delta$$

Ingredient 2: **If** $M \geq 10 \cdot \mathbf{E}(X)\Delta^2$ **then** $\frac{\mathbf{E}(X_M)}{M} \leq \frac{\mathbf{E}(X)}{M} \leq \frac{1}{10 \cdot \Delta^2}$

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

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

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

[Brassard et al.'02]

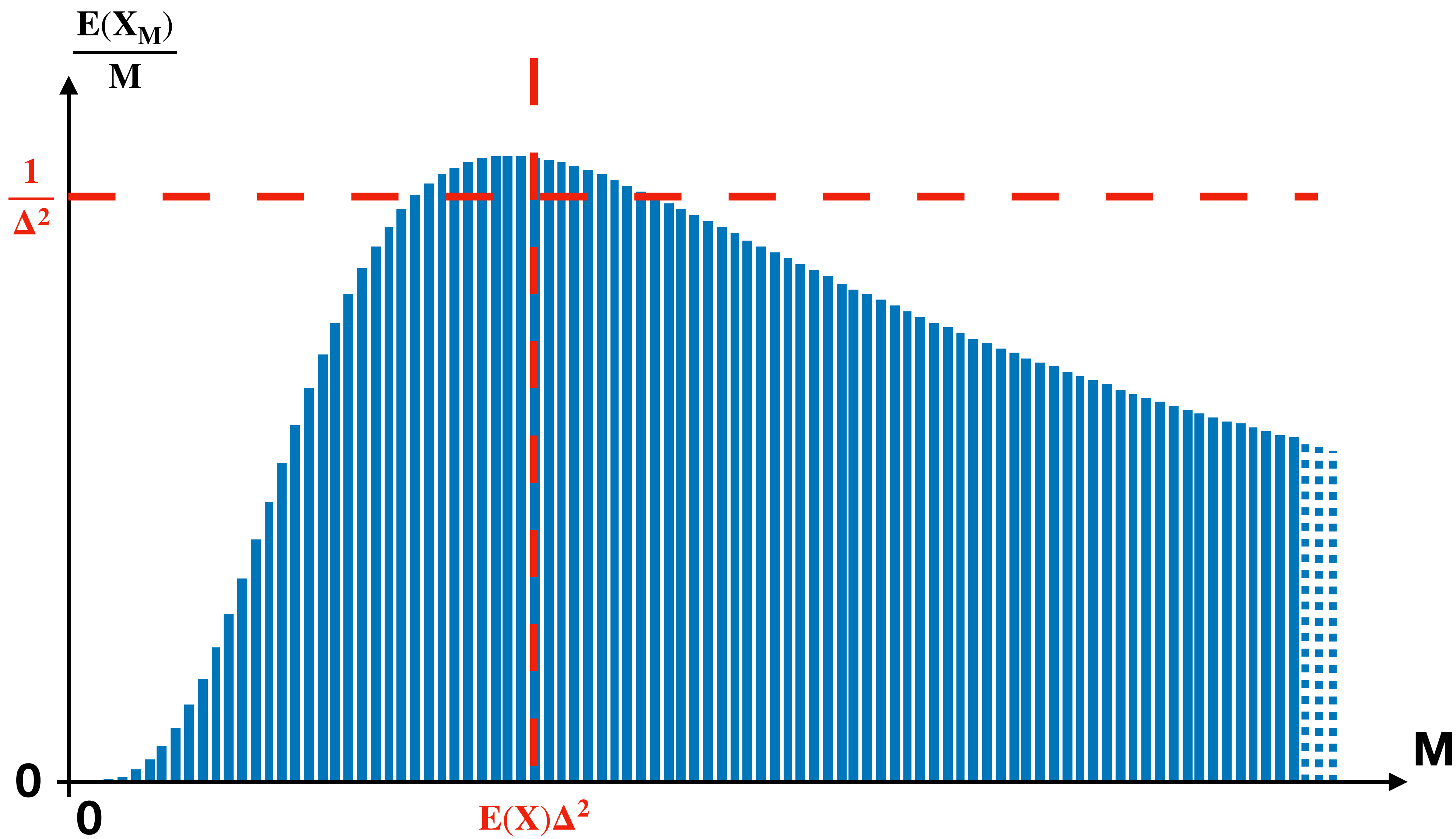
$\sqrt{\frac{\mathbf{E}(X_M)}{M}}$ $1/\Delta$

Ingredient 2: **If** $M \geq 10 \cdot \mathbf{E}(X)\Delta^2$ **then** $\frac{\mathbf{E}(X_M)}{M} \leq \frac{\mathbf{E}(X)}{M} \leq \frac{1}{10 \cdot \Delta^2}$

Ingredient 3: **If** $M \approx \mathbf{E}(X) \cdot \Delta^2$ **then** $\frac{\mathbf{E}(X_M)}{M} \overset{\bigstar}{\approx} \frac{\mathbf{E}(X)}{M} \approx \frac{1}{\Delta^2}$

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

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



Final algorithm:

Step 1: Logarithmic search on M until **Amplitude-Estimation**$(X_M, \Delta) \neq 0$

$\longrightarrow$ $2 \cdot \mathbf{E}(X)\Delta^2 \leq M \leq 10^4 \cdot \mathbf{E}(X)\Delta^2$ with high probability

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

Final algorithm:

Step 1: Logarithmic search on M until **Amplitude-Estimation**$(X_M, \Delta) \neq 0$

$\longrightarrow$ $2 \cdot \mathbf{E}(X)\Delta^2 \leq M \leq 10^4 \cdot \mathbf{E}(X)\Delta^2$ with high probability

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

Step 2: Set threshold $N = M/\epsilon$ and output $\widetilde{\mu} = N \cdot$ **Amplitude-Estimation**$(X_N, \Delta/\epsilon^{3/2})$

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

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

Final algorithm:

Step 1: Logarithmic search on M until **Amplitude-Estimation**$(X_M, \Delta) \neq 0$

$\longrightarrow$ $2 \cdot \mathbf{E}(X)\Delta^2 \leq M \leq 10^4 \cdot \mathbf{E}(X)\Delta^2$ with high probability

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

Step 2: Set threshold $N = M/\epsilon$ and output $\widetilde{\mu} = N \cdot$ **Amplitude-Estimation**$(X_N, \Delta/\epsilon^{3/2})$

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

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

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

$$\Delta/\epsilon$$

Optimality

Lower bounds

For any **Δ, ε** there exists two samplers $\begin{cases} S_X|0\rangle = \sqrt{1-p}\rangle|0\rangle + \sqrt{p}|1\rangle \\ S_Y|0\rangle = \sqrt{1-q}\rangle|0\rangle + \sqrt{q}|1\rangle \end{cases}$

with $\mathbf{E}(Y) \geq (1+2\epsilon)\cdot\mathbf{E}(X)$ and $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}, \frac{\mathbf{E}(Y^2)}{\mathbf{E}(Y)^2} \in [\Delta^2, 2\Delta^2]$

such that distinguishing between X and Y requires:

$$\boldsymbol{\Omega}\left(\frac{\boldsymbol{\Delta - 1}}{\epsilon}\right)$$

Quantum samples
from S_X / S_Y

For any **Δ, ε** there exists two samplers $\begin{cases} S_X|0\rangle = \sqrt{1-p}\rangle|0\rangle + \sqrt{p}|1\rangle \\ S_Y|0\rangle = \sqrt{1-q}\rangle|0\rangle + \sqrt{q}|1\rangle \end{cases}$

with $\mathbf{E}(Y) \geq (1+2\epsilon)\cdot\mathbf{E}(X)$ and $\frac{\mathbf{E}(X^2)}{\mathbf{E}(X)^2}, \frac{\mathbf{E}(Y^2)}{\mathbf{E}(Y)^2} \in [\Delta^2, 2\Delta^2]$

such that distinguishing between X and Y requires:

$$\Omega\left(\frac{\Delta-1}{\epsilon}\right)$$

Quantum samples from S_X / S_Y

or

$$\Omega\left(\frac{\Delta^2-1}{\epsilon^2}\right)$$

Copies of the states $S_X|0\rangle$ / $S_Y|0\rangle$

Applications

A generic quantization method

Randomized algorithm A
with output X = A()

A generic quantization method

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

A generic quantization method

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

↓

Reversible algorithm C with random
seed r as input and output X = C(r)

A generic quantization method

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random seed r as input and output X = B(r)

↓

Reversible algorithm C with random seed r as input and output X = C(r)

↓

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

A generic quantization method

Randomized algorithm A
with output X = A()

Deterministic algorithm B with random
seed r as input and output X = B(r)

Reversible algorithm C with random
seed r as input and output X = C(r)

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

samples to approximate E(X)

$$\frac{\Delta^2}{\epsilon^2}$$

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

First obstacle: time complexity

Randomized algorithm A
with output X = A()

T_{max} = maximum running time of A

T_{avg} = average running time of A

Deterministic algorithm B with random seed r as input and output X = B(r)

Reversible algorithm C with random seed r as input and output X = C(r)

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

First obstacle: time complexity

Randomized algorithm A
with output X = A()

Deterministic algorithm B with random seed r as input and output X = B(r)

Reversible algorithm C with random seed r as input and output X = C(r)

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

$\mathbf{T_{max}}$ **=** maximum running time of A

$\mathbf{T_{avg}}$ **=** average running time of A

N samples in maximum time $\mathbf{N*T_{max}}$

N samples in average time $\mathbf{N*T_{avg}}$

First obstacle: time complexity

Randomized algorithm A
with output X = A()

Deterministic algorithm B with random seed r as input and output X = B(r)

Reversible algorithm C with random seed r as input and output X = C(r)

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

T_{max} = maximum running time of A

T_{avg} = average running time of A

N samples in maximum time **$N*T_{max}$**

N samples in average time **$N*T_{avg}$**

N quantum samples in maximum/average time O(**$N*T_{max}$**)

First obstacle: time complexity

New tool: Variable-Time Amplitude Estimation

(≠ Variable-Time Amplitude Amplification)

New tool: Variable-Time Amplitude Estimation

(≠ Variable-Time Amplitude Amplification)

Randomized algorithm A
with output X in time T_{max},T_{avg}

↓

Estimate of E(X) in (average) time:

$$\frac{\Delta^2}{\epsilon^2} \cdot T_{avg}$$

First obstacle: time complexity

New tool: Variable-Time Amplitude Estimation

(≠ Variable-Time Amplitude Amplification)

Randomized algorithm A
with output X in time T_{max}, T_{avg}

Quantum sampler S_X

Estimate of E(X) in (average) time:

$$\frac{\Delta^2}{\epsilon^2} \cdot T_{avg}$$

Estimate of E(X) in time:

$$\frac{\Delta}{\epsilon^2} \cdot T_{avg,2} \cdot \mathrm{polylog}\left(\frac{M_\Omega}{\mathbf{E}(X)}, T_{max}\right)$$

where $T_{avg,2}$ = L_2-average running time of A

Application: triangle counting

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

Query access: unitaries

$$O_{\mathrm{deg}}|v\rangle|0\rangle = |v\rangle|\deg(v)\rangle$$ *(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: triangle counting

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

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

Deterministic algorithm B with random
seed r as input and output X = B(r)

Reversible algorithm C with random
seed r as input and output X = C(r)

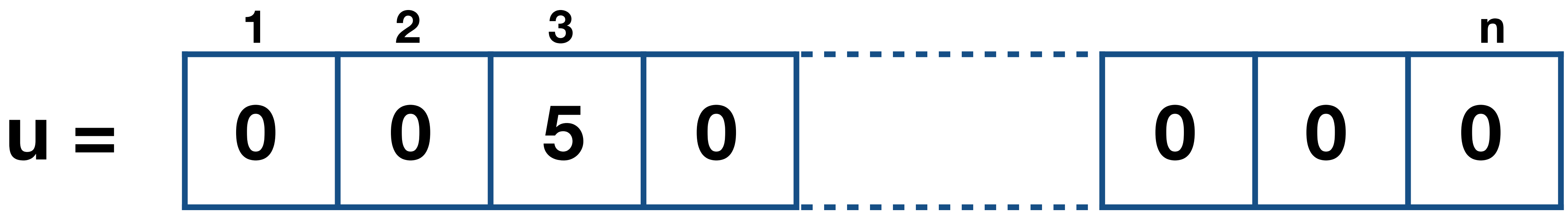
Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

Second obstacle: reversibility and streaming algorithms

	1	2	3					n
u =	**0**	**0**	**0**	**0**	…	**0**	**0**	**0**

Stream of **updates** to u:

Second obstacle: reversibility and streaming algorithms



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

Second obstacle: reversibility and streaming algorithms

u =

1	2	3					n
0	**-6**	**5**	**0**	…	**0**	**0**	**0**

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

Second obstacle: reversibility and streaming algorithms

u =

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

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

Second obstacle: reversibility and streaming algorithms

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

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

Goal: approximate some function f(u) of the **final** vector u

(example: f(u) = # of distinct elements in u)

Second obstacle: reversibility and streaming algorithms

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

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

Goal: approximate some function f(u) of the **final** vector u

(example: f(u) = # of distinct elements in u)

Algorithm with smallest possible **memory M ≪ n** using **P passes** over the same stream to approximate f(u)?

Second obstacle: reversibility and streaming algorithms

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

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

Goal: approximate some function f(u) of the **final** vector u

(example: f(u) = # of distinct elements in u)

Algorithm with smallest possible **memory M** $\ll$ **n** using **P passes** over the same stream to approximate f(u)?

Standard method (Alon, Matias, Szegedy'99):

Design an algorithm A with memory **M** that produces in **1 pass** a sample X = A(1 pass) such that **E(X) = f(u)** and $E(X^2)/E(X)^2 \leq P$

$\longrightarrow$ the average of P samples over P passes is a good approximation of f(u)

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

Deterministic algorithm B with random
seed r as input and output X = B(r)

Reversible algorithm C with random
seed r as input and output X = C(r)

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

↓

Reversible algorithm C with random
seed r as input and output X = C(r)

↓

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

- Our algorithm needs S_X^{-1}, which requires to run C^{-1}.

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

↓

Reversible algorithm C with random
seed r as input and output X = C(r)

↓

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle |C(r)\rangle$

- Our algorithm needs S_X^{-1}, which requires to run C^{-1}.
- Usually, C^{-1} needs to read the input of A in **reverse** order.

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

↓

Reversible algorithm C with random
seed r as input and output X = C(r)

↓

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle|C(r)\rangle$

- Our algorithm needs S_X^{-1}, which requires to run C^{-1}.
- Usually, C^{-1} needs to read the input of A in **reverse** order.
- If A is a streaming algorithm, it means reading the stream in the **reverse** direction!

Second obstacle: reversibility and streaming algorithms

Randomized algorithm A
with output X = A()

↓

Deterministic algorithm B with random
seed r as input and output X = B(r)

↓

Reversible algorithm C with random
seed r as input and output X = C(r)

↓

Quantum sampler $S_X|0\rangle = \frac{1}{\sqrt{R}} \sum_{r \in [R]} |r\rangle|C(r)\rangle$

- Our algorithm needs S_X^{-1}, which requires to run C^{-1}.
- Usually, C^{-1} needs to read the input of A in **reverse** order.
- If A is a streaming algorithm, it means reading the stream in the **reverse** direction!

We showed that **linear sketch** streaming algorithms can be made reversible efficiently.

(= C^{-1} and S_X^{-1} can be implemented with one pass in the **direct** direction)

Application: frequency moments in the streaming model

u =

1	2	3		…			n
0	**-6**	**5**	**0**	…	**0**	**0**	**0**

Frequency moment of order k ≥ 3: $f_k(u) = \sum_{i=1}^{n} |u_i|^k$

Application: frequency moments in the streaming model

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

Frequency moment of order k ≥ 3: $f_k(u) = \sum_{i=1}^{n} |u_i|^k$

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

Application: frequency moments in the streaming model

	1	2	3		…			n
u =	**0**	**-6**	**5**	**0**	…	**0**	**0**	**0**

Frequency moment of order k ≥ 3: $f_k(u) = \sum_{i=1}^{n} |u_i|^k$

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

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

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

Application: frequency moments in the streaming model

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

Frequency moment of order k ≥ 3: $f_k(u) = \sum_{i=1}^{n} |u_i|^k$

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

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

[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

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

Conclusion

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

Open questions:

- Can we improve the complexity to O(Δ/ε) ?
- Sample space Ω with negative values?
- Lower bound for the Frequency Moments estimation problem?

  (would follow from an $\Omega(t + nm^{-2}/t)$ lower bound for the 2-player **t**-round cc of L_∞ problem)

- Other applications ?

arXiv: 1807.06456

Extra slides

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

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