Quantum Chebyshev's Inequality

Yassine Hamoudi and Frédéric Magniez

IRIF, Université Paris Diderot, CNRS

1 Problem

Given a non-negative r.v. X with mean $\mu = \mathrm{E}(X)$, how many samples from X are necessary to compute an **ε-relative error** estimate $\tilde{\mu}$ such that

$$|\tilde{\mu} - \mu| \leq \epsilon\mu$$

with probability 2/3, provided we *know* an upper-bound $\Delta^2 \geq \frac{\mathrm{E}(X^2)}{\mathrm{E}(X)^2}$.

2 Classical setting

The optimal strategy is to compute the empirical mean $\tilde{\mu} = \frac{x_1 + \ldots + x_n}{n}$ of $n = \Omega((\Delta^2 - 1)/\epsilon^2)$ samples $x_1, \ldots, x_n \overset{\text{iid}}{\sim} X$.

⟶ The correctness is a direct consequence of **Chebyshev's inequality**.

3 Quantum setting

A quantum sample is defined as one (controlled-)execution of a unitary operator S_X or S_X^{-1} that satisfies

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

with $\psi_x =$ arbitrary unit vector.

Main result: there is a quantum algorithm that estimates μ with relative error ε using $\tilde{O}\left(\frac{\Delta}{\epsilon} \cdot \log^3\left(\frac{H}{\mathrm{E}(X)}\right)\right)$ quantum samples, given $H \geq \mathrm{E}(X)$.

Previous results:

	Number of samples	Additional inputs
Classical samples	$(\Delta^2 - 1)/\epsilon^2$	
[Brassard et al.'11] [Wocjan et al.'09] [Montanaro'15]	*Amplitude-Estimation* $\sqrt{B}/\left(\epsilon\sqrt{\mathrm{E}(X)}\right)$	Sample space $\Omega \subset [0,B]$
[Montanaro'15]	Δ^2/ϵ	
[Li, Wu'17]	$(\Delta/\epsilon) \cdot (H/L)$	$L \leq \mathrm{E}(X) \leq H$

Another model of quantum sampling?

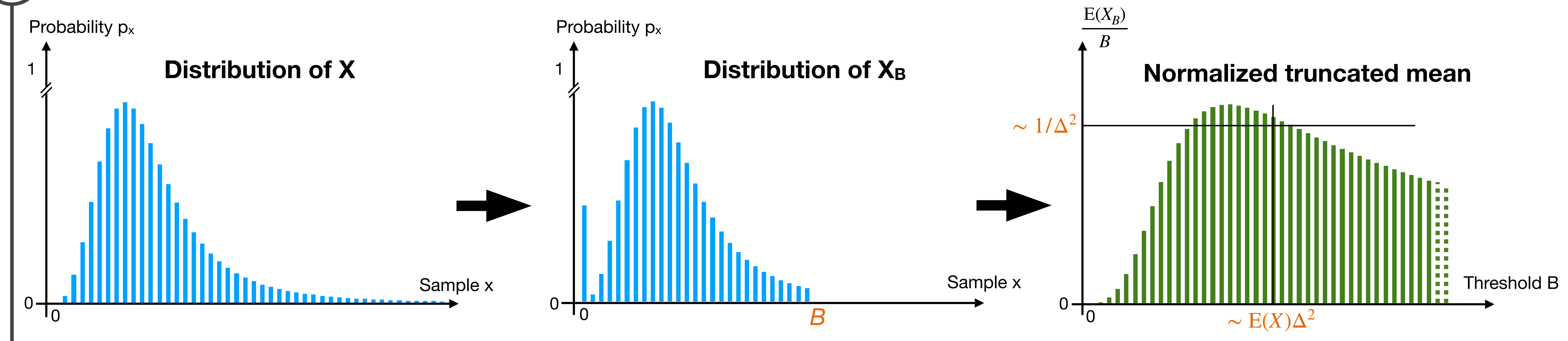
If we only have access to copies of a quantum state $\sum_{x \in \Omega} \sqrt{p_x}|\psi_x\rangle|x\rangle$ (instead of access to a unitary S_X preparing it) then one can show that

$$\Omega\left((\Delta^2 - 1)/\epsilon^2\right)$$

copies are necessary to estimate the mean with relative error ε.

⟶ No speed-up

4 Two fundamental properties



1/ The outcomes of X that are larger than $\epsilon^{-1} \cdot \mathrm{E}(X)\Delta^2$ can be replaced with 0 without changing the mean much.

- $B \geq \epsilon^{-1} \cdot \mathrm{E}(X)\Delta^2 \implies (1-\epsilon)\mathrm{E}(X) \leq \mathrm{E}(X_B) \leq \mathrm{E}(X)$

? *$\sqrt{B}/(\epsilon\sqrt{\mathrm{E}(X_B)}) = \Delta/\epsilon^{3/2}$ samples with Amplitude-Estimation... but $B = \epsilon^{-1}\mathrm{E}(X)\Delta^2$ is unknown*

2/ The "normalized truncated mean" $\mathrm{E}(X_B)/B$ is small when $B \gtrsim \mathrm{E}(X)\Delta^2$.

- $\mathrm{E}(X_B)/B \geq 1/(8\Delta^2)$ when $B \in [2\mathrm{E}(X)\Delta^2, 4\mathrm{E}(X)\Delta^2]$
- $\mathrm{E}(X_B)/B \leq 1/(16\Delta^2)$ when $B \geq 16\mathrm{E}(X)\Delta^2$

5 The algorithm

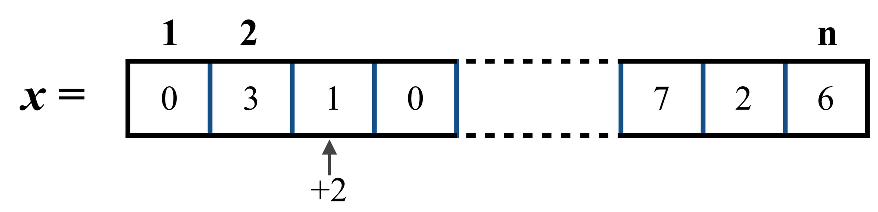
1. Set $B = 4H$ and $\tilde{\mu}_B = 0$.
2. While $\tilde{\mu}_B = 0$:
   - 2.1 Run the Amplitude-Estimation algorithm for Δ steps on X_B. Denote the result by $\tilde{\mu}_B$.
   - 2.2 Set $B = B/2$
3. Set $B = B/\epsilon$ and run the Amplitude-Estimation algorithm for $\Delta/\epsilon^{3/2}$ steps on X_B. Output the result as $\tilde{\mu}$.

Correctness: combine 4 with the fact (cf [Brassard et al.'02]) that the output of Amplitude-Estimation is 0 w.h.p. when the (normalized) estimated mean (*here:* $\mathrm{E}(X_B)/B$) is below the inverse-square of the number of samples (*here:* $1/\Delta^2$).

Step 3 can be refined to run in $\tilde{O}(\Delta/\epsilon)$.

6 New Applications

Frequency moments in the streaming model



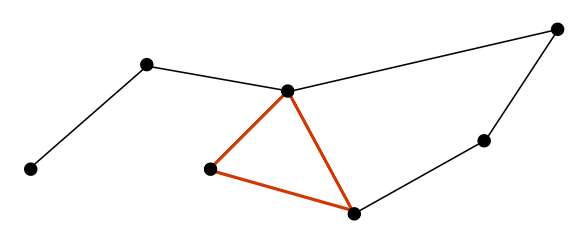
Model: stream of (classical) updates to x (turnstile model)

The frequency moment $F_k = \sum_{i=1}^n |x_i|^k$ can be estimated using P passes over the stream and a quantum memory of size $M = \frac{n^{1-2/k}}{P^2}$.

(vs classical memory of size $M = \frac{n^{1-2/k}}{P}$ [Monemizadeh, Woodruff'10])

- L_2-sampler $i \sim n$ with probability $|x_i|^2/F_2$
- Reversibility of linear sketch algorithms

Number of edges/triangles in the graph model with query access



Model: degrees, neighbors and edges can be queried in superposition.

Number m of edges: $\tilde{\Theta}\left(\frac{\sqrt{n}}{m^{1/4}}\right)$ queries

Number t of triangles: $\tilde{\Theta}\left(\frac{\sqrt{n}}{t^{1/6}} + \frac{m^{3/4}}{\sqrt{t}}\right)$ queries

(quadratic speed-ups over [Goldreich, Ron'08] [Eden, Levi, Ron'15])

- Variable-time amplitude-estimation
- Non-decr. f with $f(\mathrm{E}(X))^2 \geq \frac{\mathrm{E}(X^2)}{\mathrm{E}(X)^2}$

arXiv: 1807.06456