Near-optimal Quantum Algorithms for Multivariate Mean Estimation

arXiv:2111.09787

Arjan Cornelissen[1], Yassine Hamoudi[2], Sofiene Jerbi[3]

[1]QuSoft, University of Amsterdam
[2]Simons Institute for the Theory of Computing, University of California, Berkeley
[3]Institute for Theoretical Physics, University of Innsbruck

March 10th, 2022

QuSoft Research Center for Quantum Software

SIMONS INSTITUTE for the Theory of Computing

universität innsbruck

Problem statement

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Properties:

1. *Mean:*

$$\mu = \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) \in \mathbb{R}^d.$$

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Properties:

1. *Mean:*

$$\mu = \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) \in \mathbb{R}^d.$$

2. *Covariance matrix:*

$$\Sigma = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_d] \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_1, X_d] & \mathrm{Cov}[X_2, X_d] & \cdots & \mathrm{Var}[X_d] \end{bmatrix}$$

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Properties:

1. *Mean:*
$$\mu = \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) \in \mathbb{R}^d.$$
2. *Covariance matrix:*
$$\Sigma = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_d] \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_1, X_d] & \mathrm{Cov}[X_2, X_d] & \cdots & \mathrm{Var}[X_d] \end{bmatrix}$$

Multivariate mean estimation:

1. *Goal:* Approximate $\mu \in \mathbb{R}^d$.

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Properties:

1. *Mean:*

$$\mu = \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) \in \mathbb{R}^d.$$

2. *Covariance matrix:*

$$\Sigma = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_d] \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_1, X_d] & \mathrm{Cov}[X_2, X_d] & \cdots & \mathrm{Var}[X_d] \end{bmatrix}$$

Multivariate mean estimation:

1. *Goal:* Approximate $\mu \in \mathbb{R}^d$.
2. *Assumption:*

$$\mathrm{Tr}[\Sigma] = \sum_{j=1}^{d} \mathrm{Var}[X_j] < \infty.$$

Problem statement

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

Properties:

1. *Mean:*
$$\mu = \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}(\omega) X(\omega) \in \mathbb{R}^d.$$
2. *Covariance matrix:*
$$\Sigma = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_d] \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \cdots & \mathrm{Cov}[X_2, X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}[X_1, X_d] & \mathrm{Cov}[X_2, X_d] & \cdots & \mathrm{Var}[X_d] \end{bmatrix}$$

Multivariate mean estimation:

1. *Goal:* Approximate $\mu \in \mathbb{R}^d$.
2. *Assumption:*
$$\mathrm{Tr}[\Sigma] = \sum_{j=1}^{d} \mathrm{Var}[X_j] < \infty.$$

Applications:

1. Physics/chemistry simulations
2. Computer graphics
3. Finance
4. *Shadow tomography*

Access models

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^{d}$.

We want to approximate μ.

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

We want to approximate μ.

Classical access model:

1. Obtain outcome $\omega \sim \mathbb{P}$.
2. Function $\omega \mapsto X(\omega)$.

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

We want to approximate μ.

Classical access model:

1. Obtain outcome $\omega \sim \mathbb{P}$.
2. Function $\omega \mapsto X(\omega)$.

Quantum access model:

1. Distribution oracle:

$$|0\rangle \mapsto \sum_{\omega \in \Omega} \sqrt{\mathbb{P}(\omega)} \, |\omega\rangle.$$

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

We want to approximate μ.

Classical access model:

1. Obtain outcome $\omega \sim \mathbb{P}$.
2. Function $\omega \mapsto X(\omega)$.

Quantum access model:

1. Distribution oracle:
$$|0\rangle \mapsto \sum_{\omega \in \Omega} \sqrt{\mathbb{P}(\omega)} \, |\omega\rangle.$$
2. Random variable oracle:
$$|\omega\rangle \, |0\rangle \mapsto |\omega\rangle \, |X(\omega)\rangle.$$

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

We want to approximate μ.

Classical access model:

1. Obtain outcome $\omega \sim \mathbb{P}$.
2. Function $\omega \mapsto X(\omega)$.

Quantum access model:

1. Distribution oracle:
$$|0\rangle \mapsto \sum_{\omega \in \Omega} \sqrt{\mathbb{P}(\omega)} \, |\omega\rangle.$$
2. Random variable oracle:
$$|\omega\rangle \, |0\rangle \mapsto |\omega\rangle \, |X(\omega)\rangle.$$

Think of

$$|X(\omega)\rangle = |X(\omega)_1\rangle \otimes |X(\omega)_2\rangle \otimes \cdots \otimes |X(\omega)_d\rangle$$

Access models

We have

1. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$,
2. A random variable $X : \Omega \to \mathbb{R}^d$.

We want to approximate μ.

Classical access model:

1. Obtain outcome $\omega \sim \mathbb{P}$.
2. Function $\omega \mapsto X(\omega)$.

Quantum access model:

1. Distribution oracle:
$$|0\rangle \mapsto \sum_{\omega \in \Omega} \sqrt{\mathbb{P}(\omega)} |\omega\rangle.$$
2. Random variable oracle:
$$|\omega\rangle |0\rangle \mapsto |\omega\rangle |X(\omega)\rangle.$$

Think of
$$|X(\omega)\rangle = |X(\omega)_1\rangle \otimes |X(\omega)_2\rangle \otimes \cdots \otimes |X(\omega)_d\rangle$$

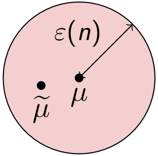
Calls to these routines are *samples*.

Results

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

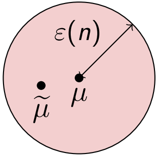
$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$



Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

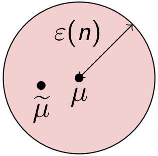


		$\varepsilon(n)$	Remarks
Classically	$d = 1$		
	$d \geq 1$		
Quantumly	$d = 1$		
	$d \geq 1$		

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

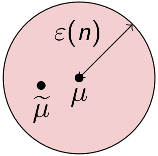


		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$		
	$d \geq 1$		

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

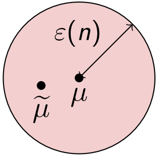


		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known $\mathrm{Var}[X]$ [Mon15;HM19]
	$d \geq 1$		

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

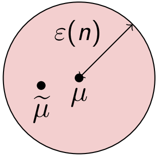


		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known Var[X] [Mon15;HM19] *Unknown* Var[X] *our work*
	$d \geq 1$		

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

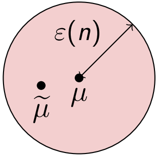


		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known Var[X] [Mon15;HM19] *Unknown* Var[X] *our work*
	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right)$	*Our work*

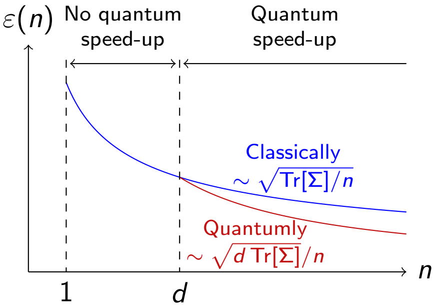
Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$



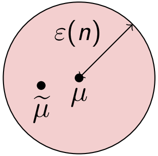
		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known Var[X] [Mon15;HM19] *Unknown Var[X] our work*
	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases}\frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d\end{cases}\right)$	*Our work*



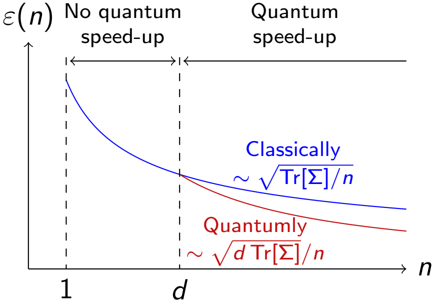
Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$



		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
Classically	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known Var[X] [Mon15;HM19] *Unknown Var[X] our work*
Quantumly	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right)$	*Our work*



Crucial observation: quantum speed-up only when $n \geq d$.

Results

Goal: Construct estimator $\widetilde{\mu}$, using n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \leq \varepsilon(n)\right] \geq \frac{2}{3}.$$

		$\varepsilon(n)$	Remarks
Classically	$d = 1$	$\Theta\left(\sqrt{\frac{\mathrm{Var}[X]}{n}}\right)$	Monte Carlo sampling
	$d \geq 1$	$\Theta\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}\right)$	Monte Carlo sampling
Quantumly	$d = 1$	$\widetilde{\Theta}\left(\frac{\sqrt{\mathrm{Var}[X]}}{n}\right)$	Known Var[X] [Mon15;HM19] Unknown Var[X] our work
	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right)$	Our work

Crucial observation: quantum speed-up only when $n \geq d$.

Quantum algorithm outline

Quantum algorithm outline

Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

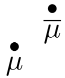
Quantum algorithm outline

Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.

$$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$

using $O(1)$ samples.



Quantum algorithm outline

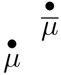
Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.

   $$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$

   using $O(1)$ samples.

2. *Get an idea of the spread:*



Quantum algorithm outline

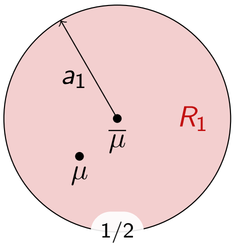
Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.

   $$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$

   using $O(1)$ samples.

2. *Get an idea of the spread:*



Quantum algorithm outline

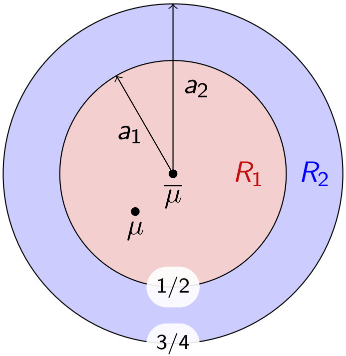
Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.

$$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$

   using $O(1)$ samples.

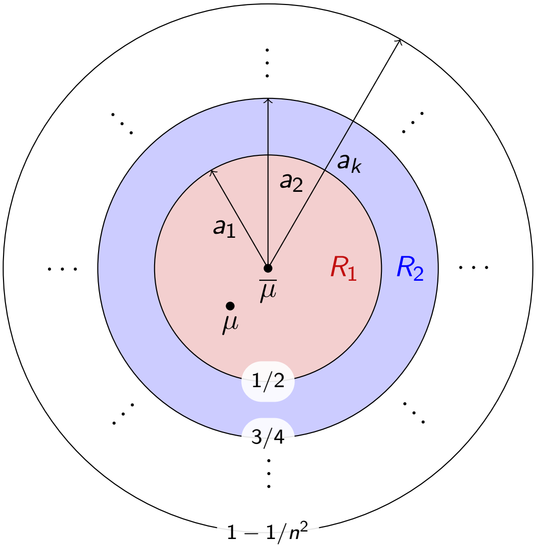
2. *Get an idea of the spread:*



Quantum algorithm outline

Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

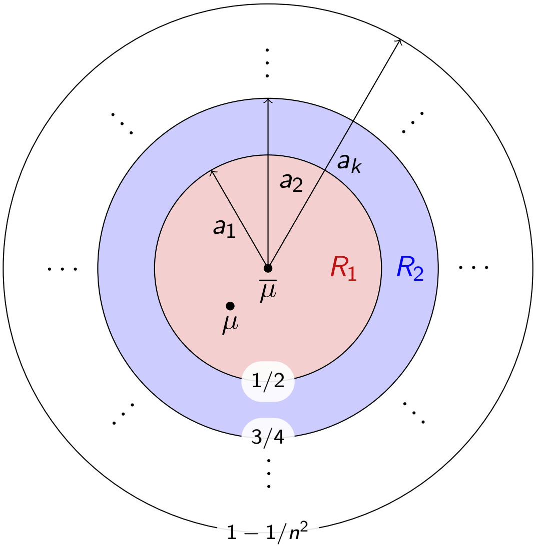
1. *Get a crude estimate:* $\overline{\mu}$ s.t.
$$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$
using $O(1)$ samples.
2. *Get an idea of the spread:*



Quantum algorithm outline

Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.
$$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$
using $O(1)$ samples.
2. *Get an idea of the spread:* Estimate quantiles a_ℓ s.t.
$$\mathbb{P}\left[\|X - \overline{\mu}\|_2 \geq a_\ell\right] \approx \frac{1}{2^\ell},$$
for $\ell \in \{1, \ldots, 2\log(n)\}$.



Quantum algorithm outline

Goal: Estimate $\mu = \mathbb{E}[X] \in \mathbb{R}^d$.

1. *Get a crude estimate:* $\overline{\mu}$ s.t.

$$\|\mu - \overline{\mu}\|_2 \leq \sqrt{\mathrm{Tr}[\Sigma]},$$

   using $O(1)$ samples.

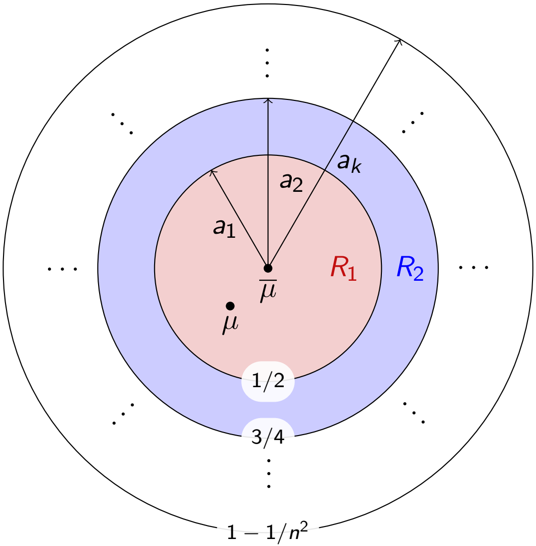
2. *Get an idea of the spread:* Estimate quantiles a_ℓ s.t.

$$\mathbb{P}\left[\|X - \overline{\mu}\|_2 \geq a_\ell\right] \approx \tfrac{1}{2^\ell},$$

   for $\ell \in \{1, \ldots, 2\log(n)\}$.

3. *Estimate truncated mean on every ring:*

$$\mathbb{E}[X] \approx \textstyle\sum_{\ell=1}^{k} \mathbb{E}[X \cdot \mathbb{1}_{X \in R_\ell}].$$



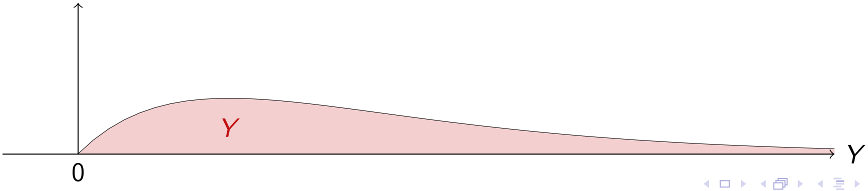
Quantile estimation

Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$. Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$. Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$



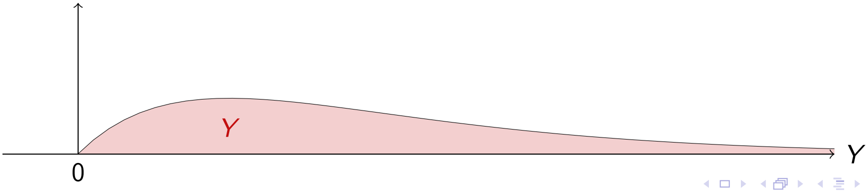
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. $\widetilde{O}(1)$ samples from Y
   Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



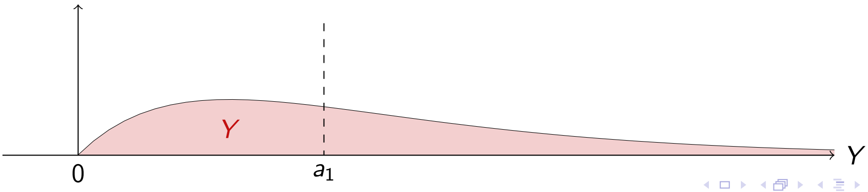
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. | $\widetilde{O}(1)$ samples from Y
   Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



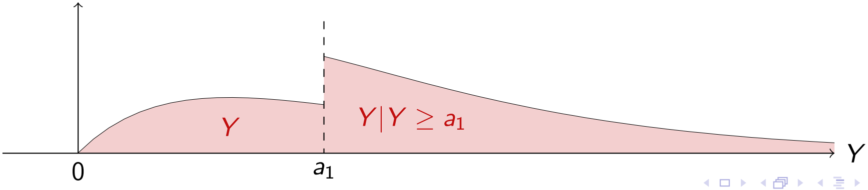
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. $\widetilde{O}(1)$ samples from Y
   Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



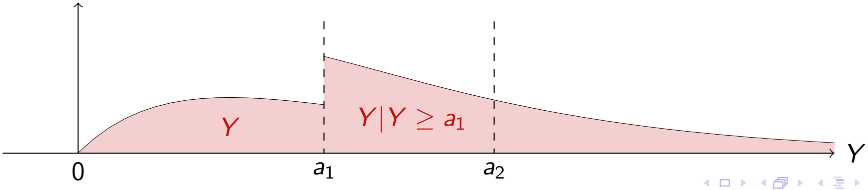
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. | $\widetilde{O}(1)$ samples from Y
   | Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



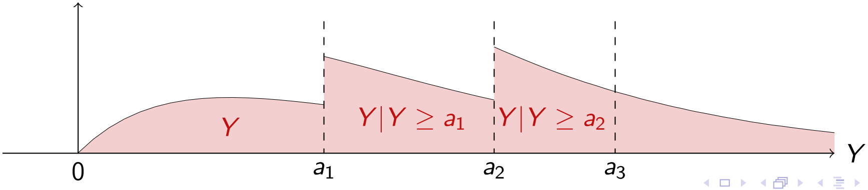
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. $\widetilde{O}(1)$ samples from Y
   Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



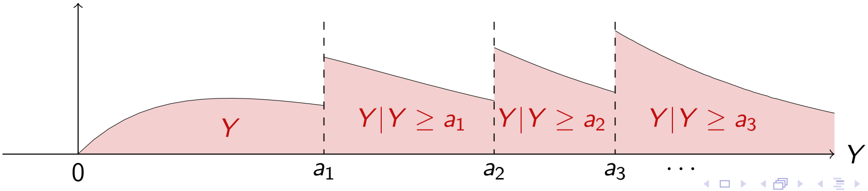
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$.

Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. $\widetilde{O}(1)$ samples from Y
   Requires $\widetilde{O}(1)$ calls

$\hookrightarrow$ Let a_1 be the median
$\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$



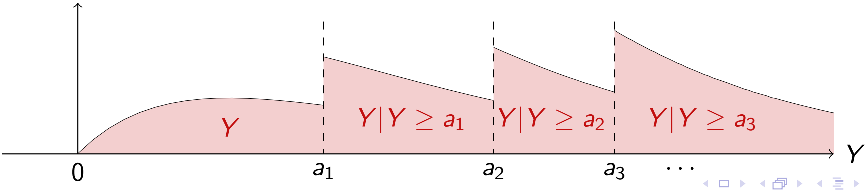
Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$. Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

1. $\widetilde{O}(1)$ samples from Y $\hookrightarrow$ Let a_1 be the median
   Requires $\widetilde{O}(1)$ calls $\quad \mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$

$\vdots$

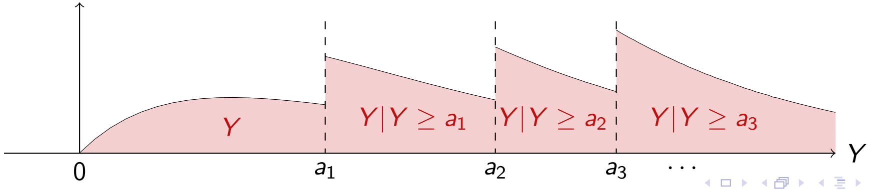
ℓ. $\widetilde{O}(1)$ samples from $Y | Y \geq a_{\ell-1}$ $\hookrightarrow$ Let a_ℓ be the median
   Requires $\widetilde{O}(\sqrt{2^{\ell-1}})$ calls $\quad \mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$



Quantile estimation

We let $Y = \|X - \overline{\mu}\|_2$. Want to find $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$

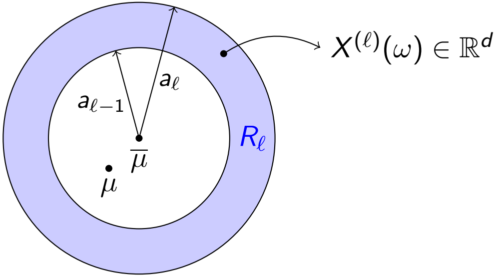
1.	$\widetilde{O}(1)$ samples from Y Requires $\widetilde{O}(1)$ calls	$\hookrightarrow$	Let a_1 be the median $\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ℓ.	$\widetilde{O}(1)$ samples from $Y \mid Y \geq a_{\ell-1}$ Requires $\widetilde{O}(\sqrt{2^{\ell-1}})$ calls	$\hookrightarrow$	Let a_ℓ be the median $\mathbb{P}[Y \geq a_\ell] \approx \frac{1}{2^\ell}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$2\log(n)$.	$\widetilde{O}(1)$ samples from $Y \mid Y \geq a_{2\log(n)}$ Requires $\widetilde{O}(n)$ calls	$\hookrightarrow$	Let $a_{2\log(n)}$ be the median $\mathbb{P}[Y \geq a_{2\log(n)}] \approx \frac{1}{n^2}$



Mean estimation on the rings

Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.
Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

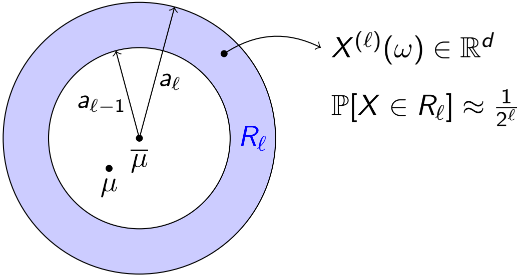


Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.

Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.

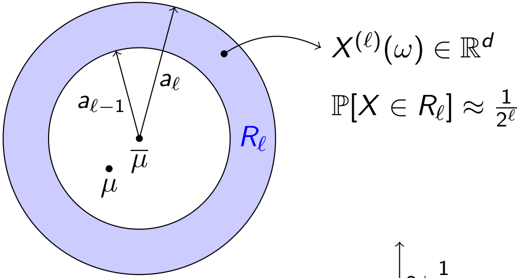


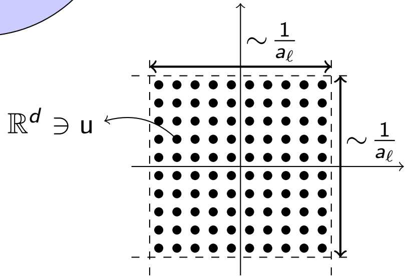
Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.

Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.

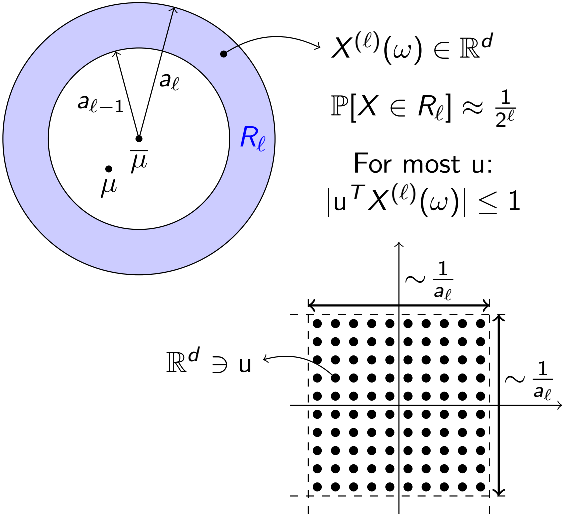




Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.
Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

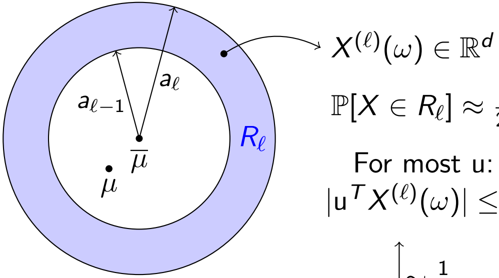
1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.



Mean estimation on the rings

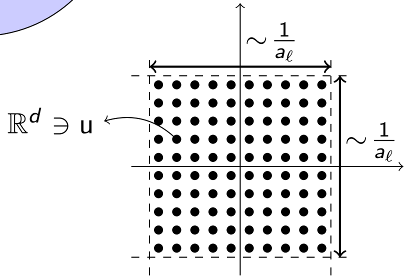
We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.
Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.
2. *Oracle conversion techniques* [GSLW18]:
   $|\mathrm{u}\rangle \mapsto e^{i2^\ell \cdot \mathrm{u}^T \mathbb{E}[X^{(\ell)}]} |\mathrm{u}\rangle$
   Requires $C_2 = \widetilde{O}(1)$ calls.



$\mathbb{P}[X \in R_\ell] \approx \frac{1}{2^\ell}$

For most u:
$|\mathrm{u}^T X^{(\ell)}(\omega)| \leq 1$

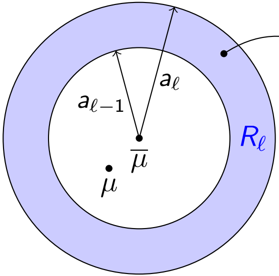


Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.

Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.
2. *Oracle conversion techniques* [GSLW18]:
   $|\mathrm{u}\rangle \mapsto e^{i2^\ell \cdot \mathrm{u}^T \mathbb{E}[X^{(\ell)}]} |\mathrm{u}\rangle$
   Requires $C_2 = \widetilde{O}(1)$ calls.



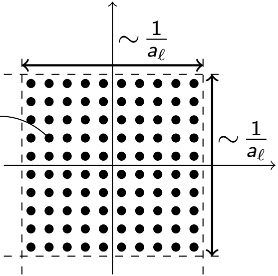
$X^{(\ell)}(\omega) \in \mathbb{R}^d$

$\mathbb{P}[X \in R_\ell] \approx \frac{1}{2^\ell}$

For most u:
$|\mathrm{u}^T X^{(\ell)}(\omega)| \leq 1$

$\mathbb{R}^d \ni \mathrm{u}$

$f(\mathrm{u}) = \mathrm{u}^T \mathbb{E}[X^{(\ell)}]$
$\Rightarrow \nabla f = \mathbb{E}[X^{(\ell)}]$

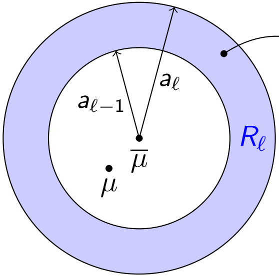


Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.

Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.
2. *Oracle conversion techniques* [GSLW18]:
   $|\mathrm{u}\rangle \mapsto e^{i2^\ell \cdot \mathrm{u}^T \mathbb{E}[X^{(\ell)}]} |\mathrm{u}\rangle$
   Requires $C_2 = \widetilde{O}(1)$ calls.
3. *Gradient estimation* [GAW18]:
   $\left\| \overline{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}] \right\|_\infty = \widetilde{O}(\sqrt{\mathrm{Tr}[\Sigma]}/n)$.
   Requires $n/(C_1 C_2)$ calls.



$X^{(\ell)}(\omega) \in \mathbb{R}^d$

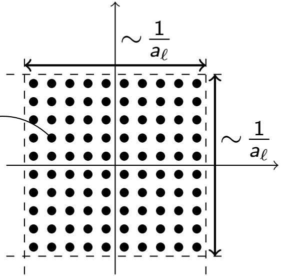
$\mathbb{P}[X \in R_\ell] \approx \frac{1}{2^\ell}$

For most u:
$|\mathrm{u}^T X^{(\ell)}(\omega)| \leq 1$

$\mathbb{R}^d \ni \mathrm{u}$

$f(\mathrm{u}) = \mathrm{u}^T \mathbb{E}[X^{(\ell)}]$

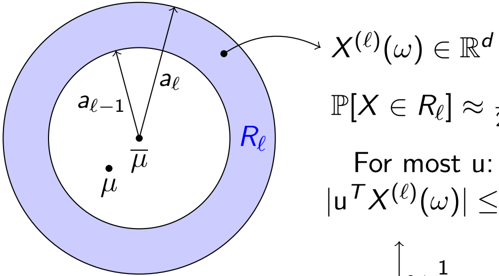
$\Rightarrow \nabla f = \mathbb{E}[X^{(\ell)}]$



Mean estimation on the rings

We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell}$.
Goal: Estimate $\mathbb{E}[X^{(\ell)}]$.

1. *Amplitude amplification on the ring:*
   Requires $C_1 = \widetilde{O}(\sqrt{2^\ell})$ calls.
2. *Oracle conversion techniques* [GSLW18]:
   $|\mathrm{u}\rangle \mapsto e^{i2^\ell \cdot \mathrm{u}^T \mathbb{E}[X^{(\ell)}]} |\mathrm{u}\rangle$
   Requires $C_2 = \widetilde{O}(1)$ calls.
3. *Gradient estimation* [GAW18]:
   $\left\|\overline{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}]\right\|_\infty = \widetilde{O}(\sqrt{\mathrm{Tr}[\Sigma]}/n)$.
   Requires $n/(C_1 C_2)$ calls.
4. *Conversion to ℓ_2* (Hölder's inequality):
   $\left\|\overline{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}]\right\|_2 = \widetilde{O}(\sqrt{d\,\mathrm{Tr}[\Sigma]}/n)$.

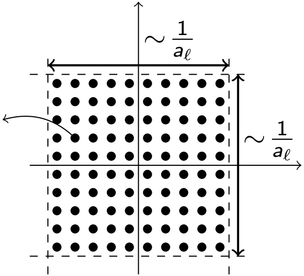


$\mathbb{P}[X \in R_\ell] \approx \frac{1}{2^\ell}$

For most u:
$|\mathrm{u}^T X^{(\ell)}(\omega)| \leq 1$

$\mathbb{R}^d \ni \mathrm{u}$

$f(\mathrm{u}) = \mathrm{u}^T \mathbb{E}[X^{(\ell)}]$
$\Rightarrow \nabla f = \mathbb{E}[X^{(\ell)}]$



Concluding remarks

Concluding remarks

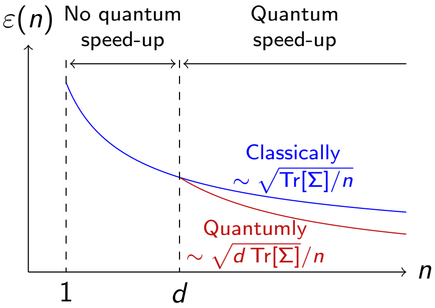
Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$



Concluding remarks

Open questions:

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Concluding remarks

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Open questions:

- Dependence on the failure probability δ?
  - Classically: [LM19;Hop20]
    $$\varepsilon(n) = O\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}} + \sqrt{\frac{\|\Sigma\| \log \frac{1}{\delta}}{n}}\right)$$
  - Constant prefactors: [LV20;LV22].

Concluding remarks

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Open questions:

- Dependence on the failure probability δ?
  - Classically: [LM19;Hop20]
    $$\varepsilon(n) = O\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}} + \sqrt{\frac{\|\Sigma\| \log\frac{1}{\delta}}{n}}\right)$$
  - Constant prefactors: [LV20;LV22].
- Optimality in different norms?

Concluding remarks

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Open questions:

- Dependence on the failure probability δ?
  - Classically: [LM19;Hop20]
    $$\varepsilon(n) = O\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}} + \sqrt{\frac{\|\Sigma\| \log \frac{1}{\delta}}{n}}\right)$$
  - Constant prefactors: [LV20;LV22].
- Optimality in different norms?
- Different access models?
  $$|\omega\rangle |0\rangle \mapsto |\omega\rangle |X(\omega)\rangle .$$
  $$|\omega\rangle |j\rangle \mapsto e^{iX(\omega)_j} |\omega\rangle |j\rangle .$$

Concluding remarks

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Open questions:

- Dependence on the failure probability δ?
  - Classically: [LM19;Hop20]
  $$\varepsilon(n) = O\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}} + \sqrt{\frac{\|\Sigma\| \log \frac{1}{\delta}}{n}}\right)$$
  - Constant prefactors: [LV20;LV22].
- Optimality in different norms?
- Different access models?
  $$|\omega\rangle |0\rangle \mapsto |\omega\rangle |X(\omega)\rangle .$$
  $$|\omega\rangle |j\rangle \mapsto e^{iX(\omega)_j} |\omega\rangle |j\rangle .$$
- Can prior knowledge on Σ help?

Concluding remarks

Main result:

Optimal estimator $\widetilde{\mu}$ with n samples, s.t.

$$\mathbb{P}\left[\|\mu - \widetilde{\mu}\|_2 \geq \varepsilon(n)\right] \leq \frac{1}{3},$$

has precision

$$\varepsilon(n) = \widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d, \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right).$$

Open questions:

- Dependence on the failure probability δ?
  - Classically: [LM19;Hop20]
    $$\varepsilon(n) = O\left(\sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}} + \sqrt{\frac{\|\Sigma\| \log \frac{1}{\delta}}{n}}\right)$$
  - Constant prefactors: [LV20;LV22].
- Optimality in different norms?
- Different access models?
  $$|\omega\rangle |0\rangle \mapsto |\omega\rangle |X(\omega)\rangle .$$
  $$|\omega\rangle |j\rangle \mapsto e^{iX(\omega)_j} |\omega\rangle |j\rangle .$$
- Can prior knowledge on Σ help?

Thanks for your attention!

`arjan@cwi.nl`

References

[GAW18] András Gilyén, Srinivasan Arunachalam, Nathan Wiebe. *Optimizing quantum optimization algorithms via faster quantum gradient computation*. `arXiv:1711.00465`.

[GLSW18] András Gilyén, Yuan Su, Guang Hao Low, Nathan Wiebe. *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*. `arXiv:1806.01838`.

[HM19] Yassine Hamoudi, Frédéric Magniez. *Quantum Chebyshev's Inequality and Applications*. `arXiv:1807.06456`.

[Hop20] Samuel B. Hopkins. *Mean Estimation with Sub-Gaussian Rates in Polynomial Time*. `arXiv:1809.07425`.

[LM19] Gabor Lugosi, Shahar Mendelson. *Mean estimation and regression under heavy-tailed distributions – a survey*. `arXiv:1906.04280`.

[LV20] Jasper C.H. Lee, Paul Valiant. *Optimal Sub-Gaussian Mean Estimation in* $\mathbb{R}$. `arXiv:2011.08384`.

[LV22] Jasper C.H. Lee, Paul Valiant. *Optimal Sub-Gaussian Mean Estimation in Very High Dimensions*.

[Mon15] Ashley Montanaro. *Quantum speedup of Monte Carlo methods*. `arXiv:1504.06987`