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

June 20th, 2022

Research Center for Quantum Software

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

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

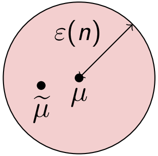
$$|X(\omega)\rangle = |X(\omega)_1\rangle \otimes |X(\omega)_2\rangle \otimes \cdots \otimes |X(\omega)_d\rangle$$

Calls to these routines are *samples*.

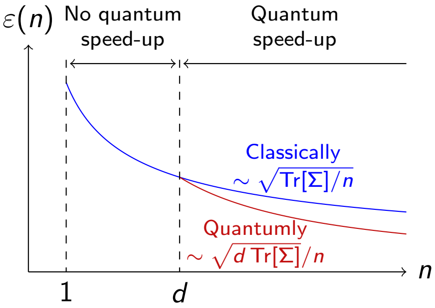
Results

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 $\mathrm{Var}[X]$ [Hei02;Mon15;HM19] Unknown $\mathrm{Var}[X]$ [Ham21]
	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right)$	*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 $\mathrm{Var}[X]$ [Hei02;Mon15;HM19] Unknown $\mathrm{Var}[X]$ [Ham21]
	$d \geq 1$	$\widetilde{\Theta}\left(\begin{cases} \frac{\sqrt{d\,\mathrm{Tr}[\Sigma]}}{n}, & \text{if } n \geq d \\ \sqrt{\frac{\mathrm{Tr}[\Sigma]}{n}}, & \text{if } n < d \end{cases}\right)$	*Our work*

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

Quantum algorithm outline

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.

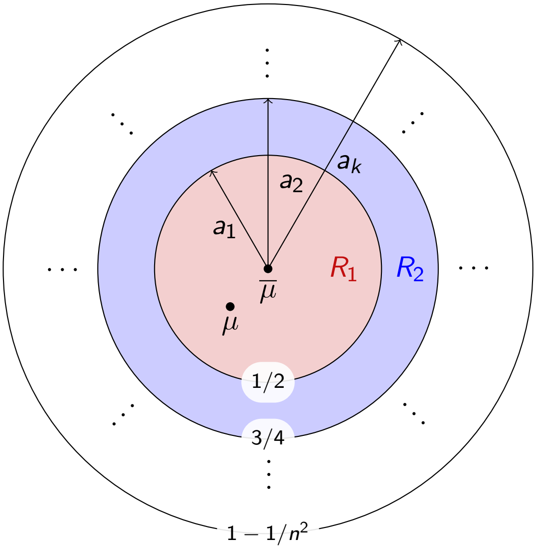
2. *Get an idea of the spread:*
   Estimate quantiles a_ℓ s.t.

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

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

3. *Estimate truncated mean on every ring:*

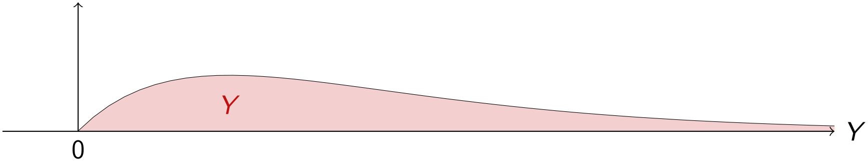
   $$\mathbb{E}[X] \approx \sum_{\ell=1}^{k} \mathbb{E}[X \cdot \mathbb{1}_{X \in R_\ell}].$$



Quantile estimation

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

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

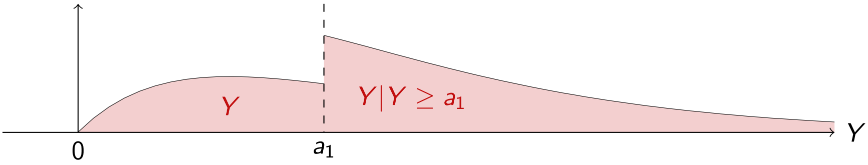


Quantile estimation

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

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

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



Quantile estimation

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

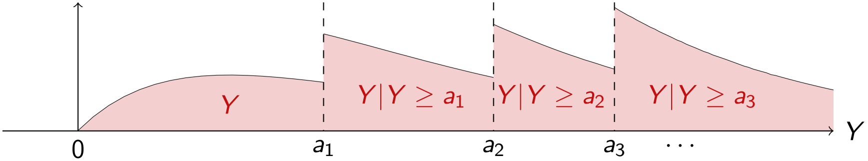
1. $\widetilde{O}(1)$ samples from Y $\hookrightarrow$ Let a_1 be the median $\mathbb{P}[Y \geq a_1] \approx \frac{1}{2}$

$\vdots$

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

$\vdots$

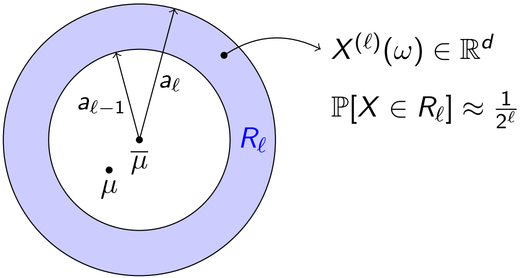
$2\log(n)$. $\widetilde{O}(1)$ samples from $Y | Y \geq a_{2\log(n)}$ $\hookrightarrow$ Let $a_{2\log(n)}$ be the median $\mathbb{P}[Y \geq a_{2\log(n)}] \approx \frac{1}{n^2}$
Requires $\widetilde{O}(n)$ samples from Y by *a.a.*



Mean estimation on the rings

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

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

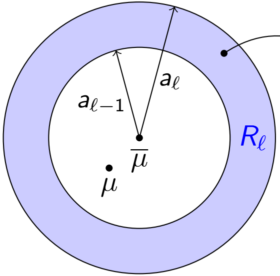


Mean estimation on the rings

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

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

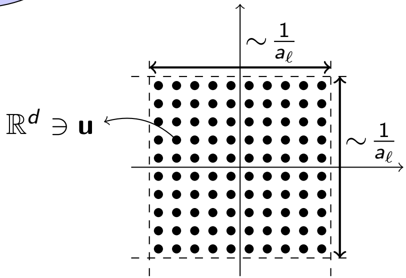
1. *Amplitude amplification on the ring:*
   Requires $\widetilde{O}(\sqrt{2^\ell})$ samples.
2. *Phase encoding techniques* [GSLW18]:
   $|\mathbf{u}\rangle \mapsto e^{i2^\ell \cdot \mathbf{u}^T \mathbb{E}[X^{(\ell)}]} |\mathbf{u}\rangle$
   Requires $\widetilde{O}(1)$ calls to 1.



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

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

For most $\mathbf{u}$:
$|\mathbf{u}^T X^{(\ell)}(\omega)| \leq 1$
(Naively: $\leq \sqrt{d}$)



Mean estimation on the rings

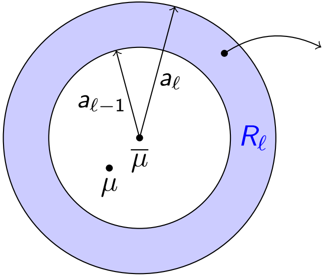
We consider $X^{(\ell)} = X \cdot \mathbb{1}_{X \in R_\ell} - \overline{\mu}$.

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

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

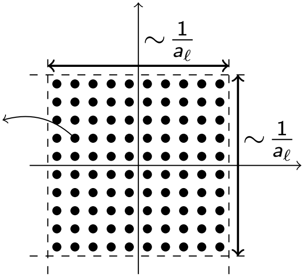
2. *Phase encoding techniques* [GSLW18]:
   $|\mathbf{u}\rangle \mapsto e^{i2^\ell \cdot \mathbf{u}^T \mathbb{E}[X^{(\ell)}]} |\mathbf{u}\rangle$
   Requires $\widetilde{O}(1)$ calls to 1.

3. *Bernstein-Vazirani over the reals* [GAW18]:
   $$\|\widetilde{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}]\|_\infty = \widetilde{O}(a_\ell/(n\sqrt{2^\ell})).$$
   $$= \widetilde{O}(\sqrt{\mathrm{Tr}[\Sigma]}/n).$$
   Requires $\widetilde{O}(n/\sqrt{2^\ell})$ calls to 2.



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

For most $\mathbf{u}$:
$|\mathbf{u}^T X^{(\ell)}(\omega)| \leq 1$
(Naively: $\leq \sqrt{d}$)

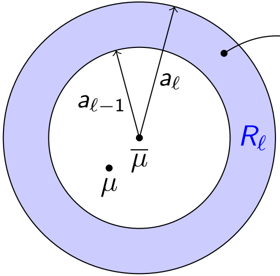


$f(\mathbf{u}) = \mathbf{u}^T \mathbb{E}[X^{(\ell)}]$
$\Rightarrow \nabla f = \mathbb{E}[X^{(\ell)}]$

Mean estimation on the rings

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

1. *Amplitude amplification on the ring:*
   Requires $\widetilde{O}(\sqrt{2^\ell})$ samples.
2. *Phase encoding techniques* [GSLW18]:
   $|\mathbf{u}\rangle \mapsto e^{i2^\ell \cdot \mathbf{u}^T \mathbb{E}[X^{(\ell)}]} |\mathbf{u}\rangle$
   Requires $\widetilde{O}(1)$ calls to 1.
3. *Bernstein-Vazirani over the reals* [GAW18]:
   $$\left\|\widetilde{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}]\right\|_\infty = \widetilde{O}\left(a_\ell/(n\sqrt{2^\ell})\right). = \widetilde{O}(\sqrt{\mathrm{Tr}[\Sigma]}/n).$$
   Requires $\widetilde{O}(n/\sqrt{2^\ell})$ calls to 2.
4. *Conversion to ℓ_2-norm* (Hölder's inequality):
   $$\left\|\widetilde{\mu}^{(\ell)} - \mathbb{E}[X^{(\ell)}]\right\|_2 = \widetilde{O}(\sqrt{d\,\mathrm{Tr}[\Sigma]}/n).$$



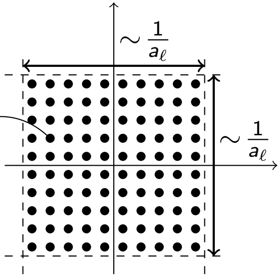
$X^{(\ell)}(\omega) \in \mathbb{R}^d$

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

For most $\mathbf{u}$:
$|\mathbf{u}^T X^{(\ell)}(\omega)| \leq 1$
(Naively: $\leq \sqrt{d}$)

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

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



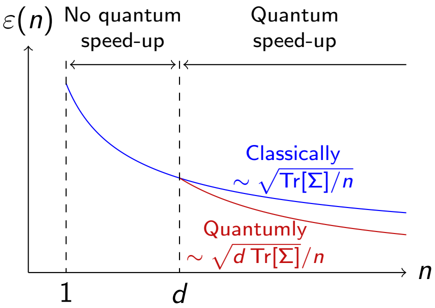
Concluding remarks

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

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!

hamoudi@berkeley.edu

References

[GAW18] A. Gilyén, S. Arunachalam, N. Wiebe. *Optimizing quantum optimization algorithms via faster quantum gradient computation*. `arXiv:1711.00465`.

[GLSW18] A., Y. Su, G. H. Low, N. Wiebe. *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*. `arXiv:1806.01838`.

[Ham22] Y. Hamoudi. *Quantum Sub-Gaussian Mean Estimator*. `arXiv:2108.12172`.

[Hei02] S. Heinrich. *Quantum Summation with an Application to Integration*. `arXiv:quant-ph/0105116`.

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

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

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

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

[LV22] J. C.H. Lee, P. Valiant. *Optimal Sub-Gaussian Mean Estimation in Very High Dimensions*. `arXiv:2011.08384`.

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