# Fast quantum tomography with optimal error bounds

Mădălin Guță

School of Mathematical Sciences
University of Nottingham

M. G., J. Kahn, R. Kueng, J. A. Tropp, arXiv:1809.11162v1
A. Acharya, T. Kypraios, M.G., arXiv:1812.xxxxx

A. Acharya, M.G., J. Phys. A **50** 195301(2017)
A. Acharya, T. Kypraios, M.G., N. J. Phys. **18** 043018 (2016)
C. Butucea, M.G., T. Kypraios, N. J. Phys. **17** 113050 (2015)
M.G. T. Kypraios, I. Dryden, N.J. Phys, **14** 105002 (2012)
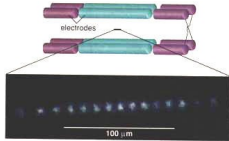
Quantum tomography
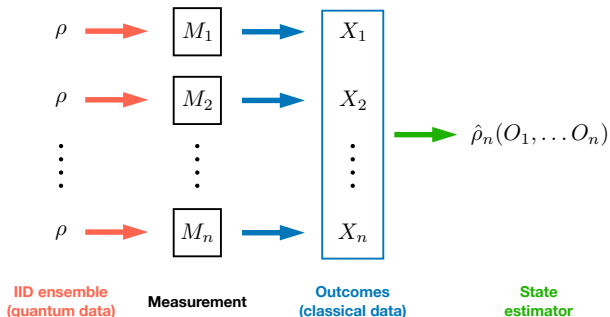
Key questions:

- measurement design: which measurements to perform?

- estimation method: which estimators are accurate and fast?

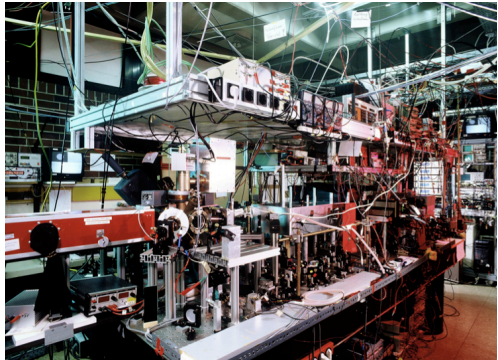- statistical model: how to deal with large dimensional systems?

- **Goal:** create an exotic state of multiple "qubits"

- **Validation:** statistical estimation from measurement outcomes

[Häffner *et al*, Nature 2005]

▶ $4^8 - 1 = 65\,535$ parameters

▶ $3^8 \times 100 = 656\,100$ measurements

▶ 10 hours measurement time

▶ days of computer time

Rainer Blatt's Lab, Innsbruck
"quantum computer" with 8 qubits (ions)

- States, measurements, data

- Least Squares estimator

- Projected Least Squares estimator

- Asymptotic rates

Quantum states

- Complex Hilbert space of 'wave functions' $\mathcal{H} = \mathbb{C}^d, L^2(\mathbb{R})$...

- State $=$ preparation: complex density matrix ρ on $\mathcal{H}$

  - $\rho = \rho^*$ (selfadjoint)
  - $\rho \geq 0$ (positive)
  - $\operatorname{Tr}(\rho) = 1$ (normalised)

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \rho_{d2} & \cdots & \rho_{dd} \end{pmatrix}$$
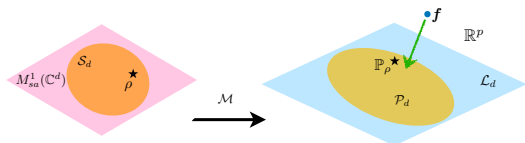
- Convex space of states $\mathcal{S}_d$

  - Pure states : one dimensional projection $P_\psi = |\psi\rangle\langle\psi| = \psi\psi^*$ with $\|\psi\| = 1$
  - Mixed states: convex combination of pure states

- States of low rank $r \ll d$: manifold of dimension $r(2 \cdot d - r) - 1$

$$\rho = \sum_{i=1}^{r} \lambda_i |\psi_i\rangle\langle\psi_i|$$

Measurements, empirical frequencies and estimators

■ Measurement with outcomes $\{1, \ldots, p\}$ as a direct map: $\mathcal{M} : M_{sa}(\mathbb{C}^d) \to \mathbb{R}^p$

$$\mathcal{M} : \rho \mapsto \mathbb{P}_\rho$$

$$\mathbb{P}_\rho(X = i) = \langle v_i | \rho | v_i \rangle, \qquad \sum_i |v_i\rangle\langle v_i| = \mathbf{1}$$
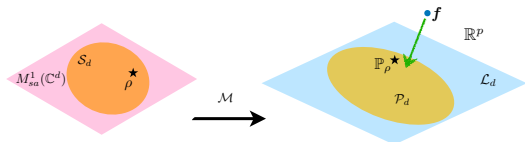
Measurements, empirical frequencies and estimators

- Measurement with outcomes $\{1, \ldots, p\}$ as a direct map: $\mathcal{M} : M_{sa}(\mathbb{C}^d) \to \mathbb{R}^p$

$$\mathcal{M} : \rho \quad \mapsto \quad \mathbb{P}_\rho$$
$$\mathbb{P}_\rho(X = i) \quad = \quad \langle v_i | \rho | v_i \rangle, \qquad \sum_i |v_i\rangle\langle v_i| = \mathbf{1}$$

- Measurement data: i.i.d. samples $X_1, \ldots, X_n \in \{1, \ldots, p\}$ from $\mathbb{P}_\rho$

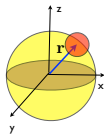- Empirical frequencies vector $\boldsymbol{f} = (f(1), \ldots, f(p))$

$$f(a) = \frac{\sharp\{i : X_i = a\}}{n}$$

- Estimator: recipe for 'projecting' $\boldsymbol{f}$ to a 'distribution' $\hat{\mathbb{P}} = \mathcal{M}(\hat{\rho})$

- Any state on $\mathbb{C}^2$ is parametrized by a 3-D Bloch vector $\mathbf{r} = (r_x, r_y, r_z)$ with $\|\mathbf{r}\| \leq 1$

$$\rho_{\mathbf{r}} = \tfrac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}$$

- Any state on $\mathbb{C}^2$ is parametrized by a 3-D Bloch vector $\mathbf{r} = (r_x, r_y, r_z)$ with $\|\mathbf{r}\| \leq 1$

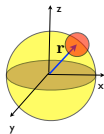$$\rho_{\mathbf{r}} = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}$$
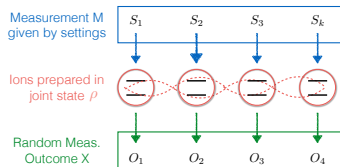

- **3 standard measurement bases** corresponding to $s = x, y, z$ spin observables

$$\underbrace{|e_x^{\pm}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}}_{s=x} \qquad \underbrace{|e_y^{\pm}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}}_{s=y} \qquad \underbrace{|e_z^{+}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ |e_z^{-}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{s=z}$$

- Probability distributions: $\mathbb{P}(o = \pm|s) = \frac{1 \pm r_s}{2} \longrightarrow r_s = \mathbb{P}(+|s) - \mathbb{P}(-|s)$

- n measurement repetitions $\longrightarrow$ count frequencies $\boldsymbol{f} = \{f(\pm|x), f(\pm|y), f(\pm|z)\}$

- Estimator: $\hat{r}_{x,y,z} := f(+|x,y,z) - f(-|x,y,z) \qquad \longrightarrow \qquad \hat{\rho} := \rho_{\hat{\mathbf{r}}},$

Measuring (correlated) states of multiple qubits

- k qubits system: $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 \cong \mathbb{C}^{2^k} = \mathbb{C}^d$

- Simultaneous, separate spin measurements on each qubit:
  - ▶ 3^k settings $\mathbf{s} = (s_1, \ldots, s_k) \in \{x, y, z\}^k$
  - ▶ 2^k outcomes $\mathbf{o} = (o_1, \ldots, o_k) \in \{+1, -1\}^k$
  - ▶ product basis vectors $|v_{\mathbf{s}}^{\mathbf{o}}\rangle := |v_{s_1}^{o_1} \otimes \cdots \otimes v_{s_k}^{o_k}\rangle$
  - ▶ probabilities $\mathbb{P}_\rho(X = \mathbf{o} \mid \mathbf{s}) = \langle v_{\mathbf{s}}^{\mathbf{o}}|\rho|v_{\mathbf{s}}^{\mathbf{o}}\rangle$

- Measure n times in each setting: counts $\{N_{\mathbf{o}, \mathbf{s}}\} \longrightarrow$ frequencies $\boldsymbol{f} \in \mathbb{R}^{2^k \times 3^k}$

- 3^k columns of length 2^k

- one column for each measurement setting

- each column contains the counts totalling $n = 100$, of the $2^k = 16$ possible outcomes

- Total set of $3^k \times 2^k \gg 4^k$ projections is highly overcomplete in $M(\mathbb{C}^{2^k})$!

1	2	11	11	11	21	5	16	21	19	11	16	2	26	15	5
2	19	10	6	15	4	22	10	3	12	8	16	18	5	14	16
3	30	12	15	9	10	18	14	3	6	11	4	4	2	1	5
4	0	4	15	10	17	2	4	14	13	0	4	8	5	1	3
5	21	13	12	7	6	5	14	12	8	12	7	19	3	8	3
6	1	12	14	0	1	1	0	6	6	12	8	2	6	2	7
7	1	2	0	19	7	12	14	6	7	14	7	9	23	15	34
8	0	1	1	0	4	8	0	6	6	0	7	12	4	15	5
9	21	17	8	10	7	7	14	9	8	15	6	9	6	3	0
10	2	16	15	0	12	9	0	3	4	1	7	3	0	4	6
11	0	0	1	17	9	2	14	12	7	0	1	0	5	5	2
12	1	1	1	0	2	8	0	4	3	0	1	0	0	3	1
13	1	0	1	1	0	0	0	0	0	14	9	7	6	2	4
14	0	1	0	0	0	1	0	0	1	1	5	6	0	2	2
15	1	0	0	1	0	0	0	0	0	1	2	0	9	6	3
16	0	0	0	0	0	0	0	1	0	0	0	1	0	4	4

[Dataset 4 ions (from Blatt group, Innsbruck)]

2-design measurements

Definition

A measurement $\mathcal{M}$ given by $\{|v_i\rangle\langle v_i|\}_{i=1}^{p}$ is called a projective 2-design if

$$\sum_i \langle v_i|A|v_i\rangle^k = d\int \langle\psi|A|\psi\rangle^k \mu(d\psi), \qquad k = 1, 2$$

or equivalently,

$\sum_i |v_i\rangle\langle v_i| \otimes |v_i\rangle\langle v_i| = C P_{\text{Sym}}$, where P_{Sym} is the projection onto the symm. subsp. of $\mathbb{C}^d \otimes \mathbb{C}^d$

- Symmetrically informationally complete measurement:

$$|\langle v_i|v_j\rangle|^2 = \frac{1}{d^2(d+1)}, \qquad 1 \le i \ne j \le p = d^2$$

- Mutually unbiased bases: $d+1$ ONBs $\{|v_i^{(a)}\rangle\}$ with

$$\left|\left\langle v_i^{(a)} \middle| v_j^{(b)} \right\rangle\right|^2 = \frac{1}{d}, \qquad a \ne b$$

- Covariant measurement: continuous set of projections

$$M(dv) = d\,|v\rangle\langle v|\,\mu(dv)$$

- stabiliser states....

- States, measurements, data

- Least Squares estimator

- Projected Least Squares estimator

- Asymptotic rates
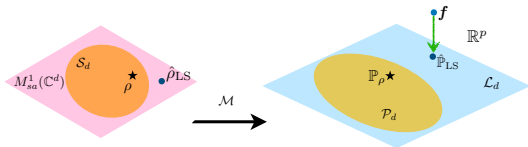
- Tomography as 'linear regression' problem

$$\boldsymbol{f} = \mathbb{P}_\rho + \boldsymbol{\epsilon} = \mathcal{M}\rho + \boldsymbol{\epsilon}$$

- Least squares estimator

$$\hat{\rho}_{\mathrm{LS}} \;=\; \operatorname*{argmin}_{\tau \in M(\mathbb{C}^d)} \|\mathcal{M}(\tau) - \boldsymbol{f}\| \;=\; \left(\mathcal{M}^\dagger \mathcal{M}\right)^{-1} \mathcal{M}^\dagger(\boldsymbol{f})$$

- Disadvantages
  - ▶ Least squares estimator is not a density matrix (not positive and trace one)
  - ▶ Least squares estimator is too "noisy" for low rank states
  - ▶ Least squares estimator minimises prediction rather than estimation error $\mathbb{E}\|\hat{\rho} - \rho\|_2^2$

Norm-error concentration bound for the least squares estimator[1] [2]

- operator-norm distance $\quad \|\rho - \tau\| = |\lambda_{max}(\Delta)|, \qquad \Delta := \rho - \tau$
- trace-norm distance $\quad \|\rho - \tau\|_1 = \sum_i |\lambda_i(\Delta)|^2 \leq d \cdot \|\rho - \tau\| \quad$ **(*)**

Theorem

The following concentration bound holds

$$\mathbb{P}\left[\|\hat{\rho}_{\mathrm{LS}} - \rho\| \geq \epsilon\right] \leq d e^{-\frac{3n\epsilon^2}{8g(d)}}$$

where

- $g(d) = 2d$ *for* 2-*design measurements*
- $g(d) \simeq d^{1.6}$ *for* *Pauli bases measurements*

Concentration inequality and $(*)$ *give upper bound for the trace-norm error*

- *2-design measurements:* $\qquad \mathbb{E}\|\rho - \hat{\rho}_{\mathrm{LS}}\|_1 \leq c_1 \log(d) \frac{d \cdot \sqrt{d}}{\sqrt{n}}$
- *Pauli bases measurements:* $\qquad \mathbb{E}\|\rho - \hat{\rho}_{\mathrm{LS}}\|_1 \leq C_1 \log(d) \frac{d \cdot d^{0.8}}{\sqrt{n}}$

---

[1] C. Butucea, M.G. and T. Kypraios, New Journal of Physics, **17**, 113050 (2015)

[2] M. G., J. Kahn, R. Kueng, J. A. Tropp, arXiv:1809.11162v1

Idea of the proof

- Write

$$\hat{\rho}_{\mathrm{LS}} - \rho = \frac{1}{n} \sum_{i=1}^{n} A_i$$

  where A_i are i.i.d. centred random matrices

- Use matrix Bernstein inequality[3] for i.i.d. $d \times d$ Hermitian matrices

$$\mathbb{P}\left(\left\| \frac{1}{n} \sum_{i=1}^{n} A_i \right\| \geq \epsilon \right) \leq 2d \exp\left(-\frac{n\epsilon^2/2}{W + \epsilon V/3} \right).$$

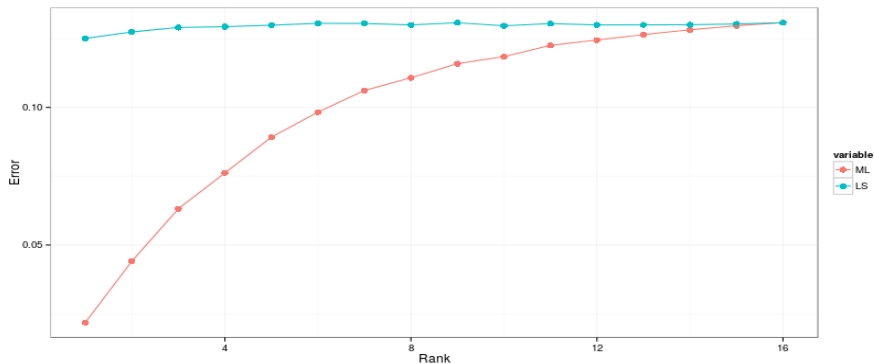 where $\|A_j\| \leq V$ and $\|\mathbb{E}(A_j^2)\| \leq W$

---

[3] J. A. Tropp, *Found Comput Math* **12** 389-434 (2012)
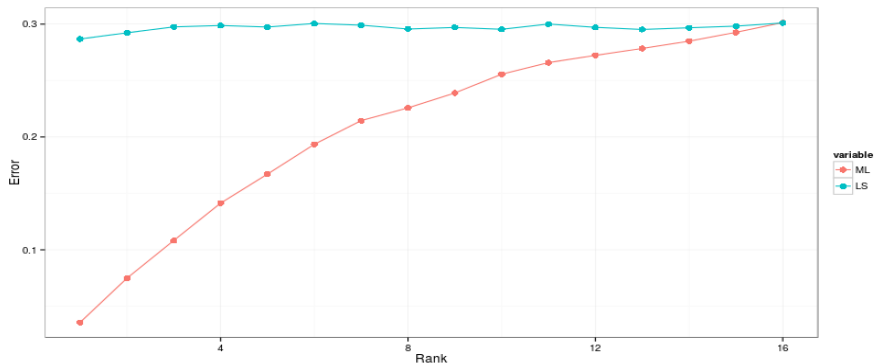
Trace-norm mean errors of LS (green) and ML (red) for rank-r states of 4 atoms, with Pauli bases measurement

Trace-norm mean errors of LS (green) and ML (red) for rank-r states of 4 atoms, with 200 random bases measurement

- States, measurements, data

- Least Squares estimator

- Projected Least Squares estimator

- Asymptotic rates
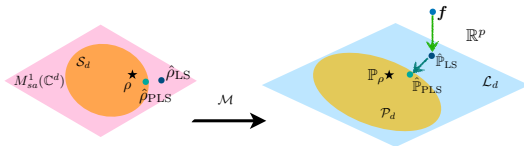
- Projected least squares estimator: closest state to the LS estimator

$$\hat{\rho}_{\mathrm{PLS}} = \operatorname*{argmin}_{\sigma \in \mathcal{S}_d} \|\sigma - \hat{\rho}_{\mathrm{LS}}\|_2^2 = \operatorname*{argmin}_{\sigma \in \mathcal{S}_d} \operatorname{Tr}\left[(\sigma - \hat{\rho}_{\mathrm{LS}})^2\right]$$

- Solution $\hat{\rho}_{\mathrm{PLS}}$ has the same eigenvectors as $\hat{\rho}_{\mathrm{LS}}$

$$\hat{\rho}_{\mathrm{LS}} = \sum_i \hat{\lambda}_i |\hat{\psi}_i\rangle\langle\hat{\psi}_i| \implies \hat{\rho}_{\mathrm{PLS}} = \sum_i \tilde{\lambda}_i |\hat{\psi}_i\rangle\langle\hat{\psi}_i|$$

- Simple algorithm for 'adjusting' the eigenvalues

Projection algorithm

- Repeat until all eigenvalues are positive:

  - truncate negative eigenvalues: $\qquad\qquad\qquad \hat{\lambda}_i < 0 \longrightarrow \hat{\lambda}_i = 0$

  - shift remaining eigenvalues so that $\sum_i \hat{\lambda}_i = 1$: $\qquad \hat{\lambda}_i \geq 0 \longrightarrow \hat{\lambda}_i - c$
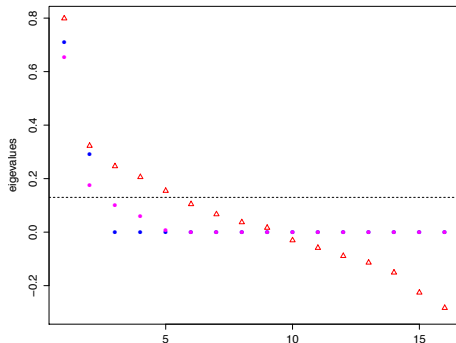
Eigenvalues of true state ρ (blue circles) versus LS (red triangles) and the PLS estimator (rose)

for a rank 2 state with $n = 20$ repetitions

Theorem

Let ρ be a state of *unknown rank r*.

The following concentration bound holds

$$\mathbb{P}\left[\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \geq \epsilon\right] \leq d\mathrm{e}^{-\frac{3n\epsilon^2}{128g(d)r^2}}$$

where

- $g(d) = 2d$ for *2-design measurements*
- $g(d) \simeq d^{1.6}$ for *Pauli bases measurements*

The concentration inequality implies the upper bound for the trace-norm error

- *2-design measurements:* $\mathbb{E}\|\rho - \hat{\rho}_{\mathrm{PLS}}\|_1 \leq c_1 \log(d)\frac{r\sqrt{d}}{\sqrt{n}}$

- *Pauli bases measurements:* $\mathbb{E}\|\rho - \hat{\rho}_{\mathrm{PLS}}\|_1 \leq C_1 \log(d)\frac{rd^{0.8}}{\sqrt{n}}$

[4] M. G., J. Kahn, R. Kueng, J. A. Tropp, arXiv:1809.11162v1

Proof idea

1) Using rank r assumption

$$\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \leq 2r\|\hat{\rho}_{\mathrm{PLS}} - \rho\|$$

2) Using definition of PLS as projection

$$\|\hat{\rho}_{\mathrm{PLS}} - \rho\| \leq \|\hat{\rho}_{\mathrm{PLS}} - \rho_{\mathrm{LS}}\| + \|\hat{\rho}_{\mathrm{LS}} - \rho\| \leq 2\|\hat{\rho}_{\mathrm{LS}} - \rho\|$$

3) Concentration bound for LS

$$\Pr\left[\|\hat{\rho}_{\mathrm{LS}} - \rho\| \geq \tau\right] \leq de^{-\frac{3n\tau^2}{8g(d)}} \quad \tau \in [0, 1].$$

1) + 2) + 3) $\implies \|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \leq 4r\|\hat{\rho}_{\mathrm{LS}} - \rho\| \implies$ concentration bound for PLS

$$\mathbb{P}\left[\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \geq \epsilon\right] \leq de^{-\frac{3n\epsilon^2}{128g(d)r^2}}$$

PLS + covariant measurement = optimal

Theorem

Let ρ be a state of *unknown rank r*.

The following concentration bound holds for *covariant measurements*

$$\mathbb{P}\left[\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \geq \epsilon\right] \leq \mathrm{e}^{2.2d - \frac{n\epsilon^2}{480r^2}}$$

Concentration inequality implies the upper bound for the trace-norm error

$$\mathbb{E}\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_1 \leq c_2 \frac{r\sqrt{d}}{\sqrt{n}}$$

- Universal lower bound [5] $\mathbb{E}\|\rho - \hat{\rho}\|_1 \geq C\frac{r\sqrt{d}}{\sqrt{n}}$ $\implies$ PLS is optimal (up to constant)

---

[5] J. Haah, *et al*, IEEE Trans. Inform. Theory (2017)

Proof technique

- Improve the concentration bound for the LS estimator

$$\mathbb{P}(\|\hat{\rho}_{\mathrm{LS}} - \rho\| \geq \epsilon) \quad \leq \quad ??$$

- Operator norm as max

$$\|A\| = \max_{y \in \mathbb{S}_d} |\langle y|A|y\rangle|$$

- Discretization by a covering net: replace the maximization over unit sphere $\mathbb{S}_d$ by a maximization over a finite point set $\mathcal{N} = \{|y_1\rangle, \dots |y_m\rangle\}$, $|\mathcal{N}| = \approx 3^{2d}$

$$\|A\| \leq 2 \max_i |\langle y_i|A|y_i\rangle|$$

- Classical concentration bound for $\langle y|\hat{\rho}_{\mathrm{LS}} - \rho|y\rangle = \frac{1}{n} \sum_{i=1}^{n} \langle y|A_i|y\rangle$

$$\mathbb{P}(|\langle y|\hat{\rho}_{\mathrm{LS}} - \rho|y\rangle| \geq t) \leq 2e^{-\frac{nt^2}{120}}$$

- Union bound

$$\mathbb{P}(\|\hat{\rho}_{\mathrm{LS}} - \rho\| \geq \epsilon) \leq \mathbb{P}\left(\max_i |\langle y_i|\hat{\rho}_{\mathrm{LS}} - \rho|y_i\rangle| \geq \frac{\epsilon}{2}\right) \leq 2 \cdot 3^{2d} \cdot e^{-\frac{n\epsilon^2}{480}}$$

Trace-norm error of LS vs PLS vs ML for Pauli bases measurements
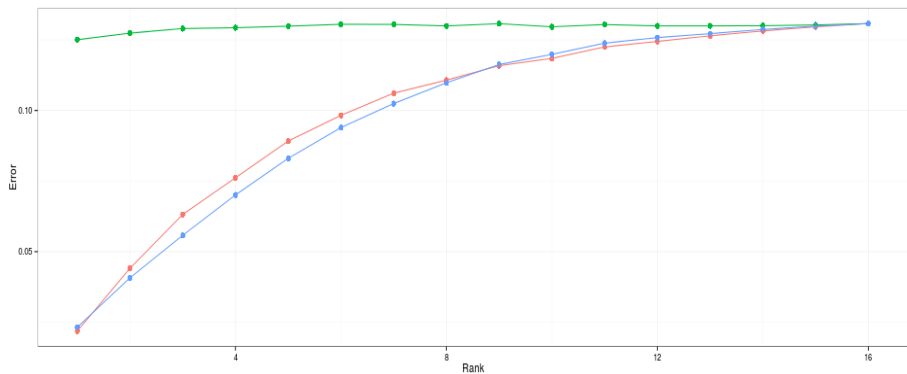
Trace-norm mean errors of LS (green), PLS(blue), ML (red) for rank-r states of 4 atoms, with Pauli bases measurement

Trace-norm mean errors of LS (green), PLS(blue), ML (red) for rank-r states of 4 atoms, with 200 random bases measurement

- States, measurements, data

- Least Squares estimator

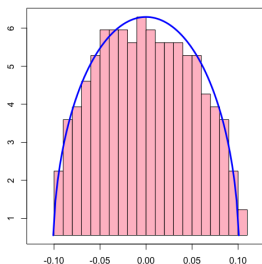- Projected Least Squares estimator

- Asymptotic rates

LS asymptotics for covariant measurements

- State ρ of low rank $r \ll d$ with eigenvalues $(\frac{1}{r}, \ldots, \frac{1}{r}, 0 \ldots, 0)$

- LS estimator
$$\hat{\rho}_{\mathrm{LS}} = \rho + \frac{1}{\sqrt{n}}\Delta = \frac{1}{r}\left(\begin{array}{cc} I & 0 \\ 0 & 0 \end{array}\right) + \frac{1}{\sqrt{n}}\left(\begin{array}{cc} A & B \\ B^* & C \end{array}\right)$$

- For large n the 'error matrix' Δ becomes Gaussian (approximately GUE)

- For large d (and low r) the eigenvalues of Δ (and C) follow a Wigner semicircle law

- Asymptotic rates: $\qquad \mathbb{E}\|\hat{\rho}_{\mathrm{LS}} - \rho\|_2^2 \approx \frac{d^2}{n}, \qquad \mathbb{E}\|\hat{\rho}_{\mathrm{LS}} - \rho\|_1 \approx 2\frac{\sqrt{d}}{\sqrt{n}}$
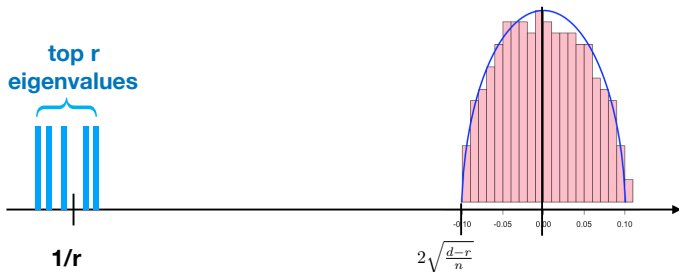
- For $n \gg d \gg 1 \implies$ LS has 2 well separated groups of eigenvalues
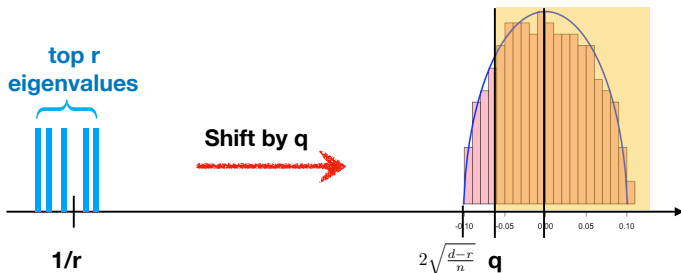
- For $n \gg d \gg 1$ $\implies$ LS has 2 well separated groups of eigenvalues

- Thresholding cut-off point q can be computed deterministically

$$rq = \int_{q}^{2\sqrt{\frac{d-r}{n}}} (x - q)\frac{n}{2\pi}\sqrt{\frac{4(d-r)}{n} - x^2}\, dx$$

- For $n \gg d \gg 1 \implies$ LS has 2 well separated groups of eigenvalues

- Thresholding cut-off point q can be computed deterministically

$$rq = \int_q^{2\sqrt{\frac{d-r}{n}}} (x - q)\frac{n}{2\pi}\sqrt{\frac{4(d-r)}{n} - x^2}\, dx$$

- For $n \gg d \gg 1$ $\implies$ LS has 2 well separated groups of eigenvalues

- Thresholding cut-off point q can be computed deterministically

$$rq = \int_q^{2\sqrt{\frac{d-r}{n}}} (x - q)\frac{n}{2\pi}\sqrt{\frac{4(d-r)}{n} - x^2} \, dx$$

- For $n \gg d \gg 1 \implies$ LS has 2 well separated groups of eigenvalues

- Thresholding cut-off point q can be computed deterministically

$$r q = \int_q^{2\sqrt{\frac{d-r}{n}}} (x - q)\frac{n}{2\pi} \sqrt{\frac{4(d-r)}{n} - x^2} \, dx$$

- Asymptotic rates: $\mathbb{E}\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_2^2 \approx \frac{6 \cdot r \cdot d}{n}, \qquad \mathbb{E}d_B(\hat{\rho}_{\mathrm{PLS}}, \rho) \approx \frac{r \cdot q(r,d)}{\sqrt{n}}$
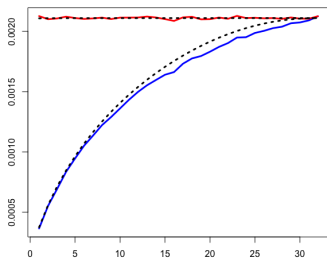
Frobenius error for LS (red) PLS (blue) 5 atoms
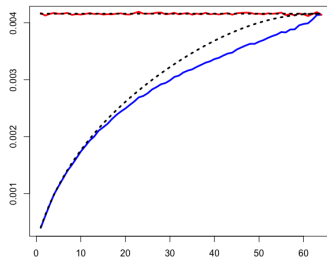
Frobenius error for LS (red) PLS (blue) 6 atoms
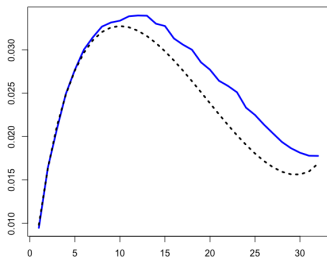
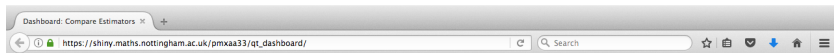Bures error for LS (red) PLS (blue) 5 atoms
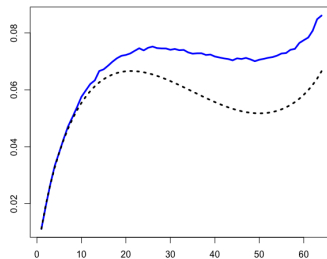
Bures error for LS (red) PLS (blue) 6 atoms

Online quantum tomography tools

- select estimators / states (or uploaded data) / error functions

  https://shiny.maths.nottingham.ac.uk/shiny/qt_dashboard/

Outlook
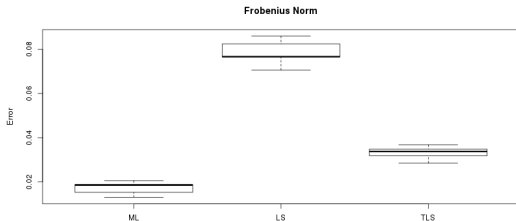
- PLS is faster than ML and has (almost) optimal rates for 2-design measuremens

- PLS achieves fundamental lower bound for covariant measurements

- Wigner semicircle used for computing asymptotic risk of PLS

Future work

- Confidence regions for PLS

- Can PLS be extended to other statistical models (eg MPS)